SECOND EDITION

CRC HANDBOOK OF THERMAL ENGINEERING

EDITED BY
RAJ P. CHHABRA

CRC Press
Taylor & Francis Group

CRC Handbook of Thermal Engineering

Second Edition

MECHANICAL and AEROSPACE ENGINEERING

Frank Kreith
Series Editor

Efstathios E. Michaelides
Co-Series Editor

RECENTLY PUBLISHED TITLES

Air Distribution in Buildings, *Essam E. Khalil*

Alternative Fuels for Transportation, *Edited by Arumugam S. Ramadhas*

Computer Techniques in Vibration, *Edited by Clarence W. de Silva*

CRC Handbook of Thermal Engineering, Second Edition, *Edited by Frank Kreith and Raj P. Chhabra*

Design and Control of Automotive Propulsion Systems, *Zongxuan Sun and Guoming (George) Zhu*

Distributed Generation: The Power Paradigm for the New Millennium,
Edited by Anne-Marie Borbely and Jan F. Kreider

Elastic Waves in Composite Media and Structures: With Applications to Ultrasonic Nondestructive Evaluation, *Subhendu K. Datta and Arvind H. Shah*

Elastoplasticity Theory, *Vlado A. Lubarda*

Energy Audit of Building Systems: An Engineering Approach, *Moncef Krarti*

Energy Conversion, Second Edition, *Edited by D. Yogi Goswami and Frank Kreith*

Energy Efficient Electrical Systems for Buildings, *Moncef Krarti*

Energy Efficiency and Renewable Energy Handbook, Second Edition,
Edited by D. Yogi Goswami and Frank Kreith

Energy Efficiency in the Urban Environment, *Heba Allah Essam E. Khalil and Essam E. Khalil*

Energy Management and Conservation Handbook, Second Edition, *Edited by Frank Kreith and D. Yogi Goswami*

Essentials of Mechanical Stress Analysis, *Amir Javidinejad*

The Finite Element Method Using MATLAB®, Second Edition, *Young W. Kwon and Hyochoong Bang*

Fluid Power Circuits and Controls: Fundamentals and Applications, *John S. Cundiff*

Fuel Cells: Principles, Design, and Analysis, *Shripad Revankar and Pradip Majumdar*

Fundamentals of Environmental Discharge Modeling, *Lorin R. Davis*

Handbook of Hydrogen Energy, *Edited by S.A. Sherif, D. Yogi Goswami, Elias K. Stefanakos, and Aldo Steinfeld*

Heat Transfer in Single and Multiphase Systems, *Greg F. Naterer*

Heating and Cooling of Buildings: Principles and Practice of Energy Efficient Design Third Edition,
T. Agami Reddy,Jan F. Kreider, Peter S. Curtiss, and Ari Rabl

Intelligent Transportation Systems: Smart and Green Infrastructure Design, Second Edition,
Sumit Ghosh and Tony S. Lee

Introduction to Biofuels, *David M. Mousdale*

Introduction to Precision Machine Design and Error Assessment, *Edited by Samir Mekid*

Introductory Finite Element Method, *Chandrakant S. Desai and Tribikram Kundu*

Large Energy Storage Systems Handbook, *Edited by Frank S. Barnes and Jonah G. Levine*

Machine Elements: Life and Design, *Boris M. Klebanov, David M. Barlam, and Frederic E. Nystrom*

Mathematical and Physical Modeling of Materials Processing Operations,
Olusegun Johnson Ilegbusi, Manabu Iguchi, and Walter E. Wahnsiedler

Mechanics of Composite Materials, *Autar K. Kaw*

Mechanics of Fatigue, *Vladimir V. Bolotin*

Mechanism Design: Enumeration of Kinematic Structures According to Function, *Lung-Wen Tsai*

Mechatronic Systems: Devices, Design, Control, Operation and Monitoring, *Edited by Clarence W. de Silva*

The MEMS Handbook, Second Edition (3 volumes), *Edited by Mohamed Gad-el-Hak*

- MEMS: Introduction and Fundamentals
- MEMS: Applications
- MEMS: Design and Fabrication

Multiphase Flow Handbook, Second Edition, *Edited by Efstathios E. Michaelides, Clayton T. Crowe, and John D. Schwarzkopf*

Nanotechnology: Understanding Small Systems, Third Edition, *Ben Rogers, Jesse Adams, and Sumita Pennathur*

Nuclear Engineering Handbook, Second Edition, *Edited by Kenneth D. Kok*

Optomechatronics: Fusion of Optical and Mechatronic Engineering, *Hyungsuck Cho*

Practical Inverse Analysis in Engineering, *David M. Trujillo and Henry R. Busby*

Pressure Vessels: Design and Practice, *Somnath Chattopadhyay*

Principles of Solid Mechanics, *Rowland Richards, Jr.*

Principles of Sustainable Energy Systems, Second Edition, *Edited by Frank Kreith with Susan Krumdieck, Co-Editor*

Thermodynamics for Engineers, *Kau-Fui Vincent Wong*

Vibration and Shock Handbook, *Edited by Clarence W. de Silva*

Vibration Damping, Control, and Design, *Edited by Clarence W. de Silva*

Viscoelastic Solids, *Roderic S. Lakes*

Weatherization and Energy Efficiency Improvement for Existing Homes: An Engineering Approach, *Moncef Krarti*

CRC Handbook of Thermal Engineering

Second Edition

Edited by

Raj P. Chhabra

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper
International Standard Book Number-13: 978-1-4987-1527-0 (Hardback)

Library of Congress Cataloging-in-Publication Data

Names: Chhabra, R. P., editor.
Title: CRC handbook of thermal engineering / edited by Raj P. Chhabra.
Other titles: Handbook of thermal engineering
Description: Second edition. | Boca Raton : Taylor & Francis, CRC Press, 2017. | Includes bibliographical references and index.
Identifiers: LCCN 2017015929 | ISBN 9781498715270 (hardback) | ISBN 9781315119717 (ebook)
Subjects: LCSH: Heat engineering--Handbooks, manuals, etc. | Thermodynamics--Handbooks, manuals, etc.
Classification: LCC TJ260 .C69 2017 | DDC 621.402--dc23
LC record available at https://lccn.loc.gov/2017015929

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Acknowledgments xi
Preface to the Second Edition xiii
Nomenclature xv
Editor xxi
Contributors xxiii

Chapter 1 Engineering Thermodynamics 1

1.1 Fundamentals 3
Michael J. Moran
1.2 Control Volume Applications 15
Michael J. Moran
1.3 Property Relations and Data 23
Michael J. Moran
1.4 Combustion 56
Michael J. Moran
1.5 Exergy Analysis 67
Michael J. Moran
1.6 Vapor and Gas Power Cycles 76
Michael J. Moran
1.7 Guidelines for Improving Thermodynamic Effectiveness 84
Michael J. Moran
1.8 Exergoeconomics 86
George Tsatsaronis
1.9 Design Optimization 94
George Tsatsaronis
1.10 Economic Analysis of Thermal Systems 102
George Tsatsaronis
1.11 Exergoenvironmental Analysis 106
George Tsatsaronis
1.12 Advanced Exergy-Based Methods 107
George Tsatsaronis

Chapter 2 Fluid Mechanics 113

2.1 Fluid Statics 115
Stanley A. Berger
2.2 Equations of Motion and Potential Flow 123
Stanley A. Berger
2.3 Similitude: Dimensional Analysis and Data Correlation 139
Stuart W. Churchill
2.4 Hydraulics of Pipe Systems 154
J. Paul Tullis and Blake Paul Tullis
2.5 Open Channel Flow 170
Frank M. White
2.6 External Incompressible Flows 178
John C. Leylegian

2.7 Compressible Flow 190
John C. Leylegian
2.8 Multiphase Flow 209
John C. Chen
2.9 Non-Newtonian Flows 223
Anoop K. Gupta, Raj P. Chhabra, Thomas F. Irvine, Jr., and Massimo Capobianchi

Chapter 3 Heat and Mass Transfer 249

3.1 Conduction Heat Transfer 250
Robert F. Boehm
3.2 Convection Heat Transfer 263
3.2.1 Natural Convection 263
Swati A. Patel, Raj P. Chhabra, George D. Raithby, and K.G. Terry Hollands
3.2.2 Forced Convection: External Flows 283
Anoop K. Gupta, Raj P. Chhabra, and N.V. Suryanarayana
3.2.3 Forced Convection: Internal Flows 305
Anoop K. Gupta, Raj P. Chhabra, and N.V. Suryanarayana
3.2.4 Convection Heat Transfer in Non-Newtonian Fluids 315
Swati A. Patel, Raj P. Chhabra, Thomas F. Irvine, Jr., and Massimo Capobianchi
3.3 Radiative Heat Transfer 341
Michael F. Modest
3.4 Phase-Change 368
3.4.1 Boiling and Condensation 368
Van P. Carey
3.4.2 Particle Gas Convection 388
John C. Chen
3.4.3 Melting and Freezing 393
Vasilios Alexiades and Jan Kośny
3.5 Mass Transfer 417
Anthony F. Mills

Chapter 4 Applications 451

4.1 Heat Exchangers for the Process and Energy Industries 459
Joshua D. Ramsey, Ken Bell, and Ramesh K. Shah
4.2 Application of Nanofluids in Heat Exchangers: Performance and Challenges 545
Bengt Sundén and Zan Wu
4.3 Convection Heat Transfer in Conduits with Nanofluids 575
Clement Kleinstreuer and Zelin Xu
4.4 Fouling in Crude Oil and Food Related Heat-Transfer Equipment 603
D. Ian Wilson and Graham T. Polley
4.5 Bioheat Transfer 629
John A. Pearce, Kenneth R. Diller, and Jonathan W. Valvano
4.6 Thermal Insulation 716
David W. Yarbrough
4.7 Energy Audit for Buildings 728
Moncef Krarti

4.8 Advanced Energy-Efficient Building Envelope Systems 749
Moncef Krarti and John Zhai
4.9 Use of Phase Change Materials in Buildings 781
Jan Kośny and David W. Yarbrough
4.10 Thermal Bridges in Building Structures 800
Jan Kośny and David W. Yarbrough
4.11 Compressors 818
Christian K. Bach, Ian H. Bell, Craig R. Bradshaw, Eckhard A. Groll, Abhinav Krishna, Orkan Kurtulus, Margaret M. Mathison, Bryce Shaffer, Bin Yang, Xinye Zhang, and Davide Ziviani
4.12 Pumps and Fans 853
Robert F. Boehm
4.13 Cooling Towers 864
Anthony F. Mills
4.14 Pinch Technology 875
Santanu Bandyopadhyay and Shankar Narasimhan
4.15 Air-Conditioning Systems 897
Donald L. Fenton
4.16 Heat Transfer Enhancement 933
Raj M. Manglik
4.17 Heat Pipes 953
Sameer Khandekar
4.18 Liquid Atomization and Spraying 986
Mario F. Trujillo and *Rolf D. Reitz*
4.19 Heat Transfer in Plasma Sprays 1003
Milind A. Jog
4.20 Thermal Processing and Preservation of Foods 1017
Prabhat Kumar and K.P. Sandeep
4.21 Thermal Conduction in Electronic Microstructures and Nanostructures ... 1036
Sanjiv Sinha, Krishna Valavala and Jun Ma
4.22 Role of Cooling in Electronics Reliability 1078
Pradeep Lall
4.23 Direct Contact Heat Transfer 1131
Harold R. Jacobs
4.24 Heat Transfer in Presence of Synthetic Jets 1153
Mangesh Chaudhari and Amit Agrawal
4.25 Temperature and Heat Transfer Measurements 1162
Robert J. Moffat and Tadhg O'Donovan
4.26 Flow Measurement 1188
Jungho Kim, S.A. Sherif, and Alan T. McDonald
4.27 Applications of Artificial Neural Networks and Genetic Methods in Thermal Engineering 1217
Arturo Pacheco-Vega, Gerardo Diaz, Mihir Sen, and K.T. Yang
4.28 Thermal Aspects of Paper Making 1270
Martine Rueff and Evelyne Mauret
4.29 Drying of Materials 1306
Pawel Wawrzyniak, Ireneusz Zbicinski, and Mariia Sobulska
4.30 Heat Transfer in Rotary Kilns 1338
P.S. Ghoshdastidar
4.31 Heat Transfer in Glass Manufacturing Processes 1352
Naveen Tiwari and Rajappa Tadepalli

4.32 Solar Hydrogen as a "Renewable Reductant": Points and Counterpoints ... 1369
Raj Ganesh S. Pala
4.33 Passive and Active Solar Distillation ... 1386
Desh Bandhu Singh and G.N. Tiwari

Chapter 5 Numerical Methods and Computational Tools ... 1441

5.1 Computer Aided Engineering ... 1442
Atul Sharma
5.2 Finite Difference Methods ... 1445
Atul Sharma
5.3 Finite Volume Method ... 1454
Atul Sharma
5.4 Finite Element Method ... 1468
Salil S. Kulkarni
5.5 Lattice Boltzmann Method ... 1496
K. Hrisheekesh and Amit Agrawal
5.6 Immersed Boundary Method for Fluid-Structure Interaction Simulations ... 1507
Rajneesh Bhardwaj and Atul Sharma
5.7 Numerical Methods for Multiphase Flows ... 1524
Shyamprasad Karagadde and Atul Sharma
5.8 Large Eddy Simulation for Wall-Bounded Flows ... 1545
Amitabh Bhattacharya
5.9 Software and Computer Codes ... 1558
Atul Sharma

Appendix A: Properties of Gases and Vapors ... 1559
Paul Norton

Appendix B: Properties of Liquids ... 1597

Appendix C: Properties of Solids ... 1601

Appendix D: SI Units and Conversion Factors ... 1605

Index ... 1613

Acknowledgments

I would like to place on record my deep appreciation and gratitude to the many individuals who have facilitated the publication of this revised *Handbook*. First of all, I am grateful to the authors who have contributed to this Handbook. I have really enjoyed working with them all. Clearly, the high quality of the work included here demonstrates their passion, dedication, and high level of professionalism. Without them, this updated *Handbook* would not have been possible! Next, I wish to thank numerous reviewers for their timely and constructive reports on various sections, especially those in Chapter 4, which have led to significant improvements to the material on engineering applications included here. During the past two years, it has been a wonderful experience to work with many individuals at Taylor and Francis, the publishers of this *Handbook*. Each one of them has been extremely helpful, not only in keeping this project under a tight leash, but also coming to my rescue whenever I was in a difficult situation. Some of these individuals deserve to be singled out here. First and foremost, I would like to express my deep gratitude to Jonathan Plant, the commissioning editor for this project, who patiently answered my endless queries via e-mail and over the telephone and dealt with several unusual requests regarding the artwork, copyright issues, etc. Thank you, Jonathan! Similarly, Kyra Lindholm, Amber Donley, Hector Mojena III, and Laurie Oknowsky went out of their way, and much beyond the call of duty, to help with the issues related to obtaining the necessary permissions to reproduce the copyrighted material in this *Handbook*. Finally, Richard Tressider and Mathi Ganesan (of Codemantra) converted the material received from the individual contributors in a range of formats and styles into the finished product—the *Handbook*! I thank them all for their help during the course of this project.

Preface to the Second Edition

In order to put the field of thermal engineering—thermodynamics and heat transfer—in perspective, one only needs to walk down the aisles of a modern department or hardware store, and look at the range of manufactured goods, appliances, foodstuffs, integrated circuit boards, textiles, papers, televisions, air conditioners, cell phones, etc. now used in everyday life. Their manufacture and satisfactory performance are strongly influenced by various aspects of thermal engineering. Similarly, in the technological world, processes such as the fabrication of metallic and plastic components, fractionation of crude oil into valuable and usable products, desalination of water, and energy generation and management (to name only a few) are all heavily dependent on the application of thermal engineering principles.

Indeed, it is no exaggeration to say that thermal engineering is an inseparable part of the natural and man-made world we inhabit. This *Handbook* brings together current state-of-the-art-knowledge in many key application areas. The objectives and goals of the new edition of this *Handbook* remain the same as that of the first edition published in 2000 under the leadership of Frank Kreith: *The purpose here is not to compete with the available traditional Handbooks of heat transfer that emphasize fundamentals, analytical approaches to thermal problems and elegant solutions of traditional problems in thermal sciences! Instead, this Handbook endeavors to provide most up to date and reliable information on specific topics of interest in a convenient form that is accessible to the working professionals in industry.*

The first three chapters provide a quick refresher on the basic principles of thermodynamics (Chapter 1), fluid mechanics (Chapter 2), and heat transfer (Chapter 3) to understand the wide-ranging specific industrial applications presented in Chapter 4. In view of the growing importance of CFD (computational fluid dynamics) tools used in solving thermal engineering problems, Chapter 5 provides an overview of the range of computational techniques in use now, along with their relative merits and demerits and domains of application. This chapter has been put together by a completely new group of contributors to best address the latest developments across the range of computational methods.

So what is new in this edition? The first three chapters dealing with the fundamentals of thermodynamics, fluid mechanics, and heat transfer have been thoroughly reviewed and revised as deemed appropriate. In Chapter 4 on applications, many new application areas have been added that entail the application of thermal engineering principles, including the drying of wet solids and the role of drying in paper making, glass processing, rotary kilns, solar distillation, solar hydrogen as a renewable reductant, plasma sprays, thermal insulation and thermal bridging in buildings and structures, energy audit and management in buildings, bioheat transfer, to mention a few. Notwithstanding the dynamic and ever-changing nature of the field of thermal engineering, certain choices had to be made in the early planning stages to limit the scope of the *Handbook*. Likewise, a few chapters have been dropped from the first edition. Much of the new material in this *Handbook* has gone through a peer review process and been carefully proofread. However, in a project of this magnitude, with authors from diverse backgrounds, it is inevitable that some errors have gone unspotted. As the editor, I would like to invite the working professionals and other users of this *Handbook* to send me feedback regarding the applications included here, or the topics that should be added to the future editions of this volume, or simply to report the errors spotted so that these can be corrected at the first opportunity.

Raj P. Chhabra
Kanpur, India

Nomenclature

Symbol	Quantity	Unit		Dimensions (MLtT)
		SI	English	
a	Velocity of sound	m/s	ft/s	$L\,t^{-1}$
a	Acceleration	m/s^2	ft/s^2	$L\,t^{-2}$
A	Area: A_c, cross-sectional area; A_p, projected area of a body normal to the direction of flow; A_q, area through which rate of heat flow is q; A_g, surface area; A_o, outside surface area; A_i inside surface area; A_f, fin surface area	m^2	ft^2	L^2
b	Breadth or width	m	ft	L
c	Specific heat; c_p, specific heat at constant pressure; c_v, specific heat at constant volume	J/kg K	$Btu/lb_m\,°R$	$L^2\,t^{-2}\,T^{-1}$
C	Constant or Coefficient; C_D, total drag coefficient; C_f, skin friction coefficient; C_{fx}, local value of C_f at distance x, from leading edge; $\overline{C}_f$, average value of C_f	none	none	—
C	Thermal capacity	J/K	Btu/°F	$M\,L^2\,t^{-2}\,T^{-1}$
$\dot{C}$	Hourly heat capacity rate; $\dot{C}_c$, hourly heat capacity rate of colder fluid in a heat exchanger; $\dot{C}_h$, hourly heat capacity of hotter fluid; C*, ratio of heat capacity rates in heat exchangers	W/K	Btu/h°F	$M\,L^2\,t^{-1}\,T^{-1}$
D	Diameter, D_H, hydraulic diameter; D_o, outside diameter; D_i, inside diameter	m	ft	L
e	Base of natural or Napierian logarithm	none	none	—
e	Total energy per unit mass	J/kg	Btu/lb_m	$L^2\,t^{-2}$
E	Total energy	J	Btu	$M\,L^2\,t^{-2}$
E	Emissive power of a radiating body; E_b, emissive power of a blackbody	W/m^2	$Btu/h{\cdot}ft^2$	$M\,t^{-2}$
E_λ	Monochromatic emissive power per micron at wavelength λ	W/m μm	$Btu/h{\cdot}ft^2$ micron	$M\,t^{-2}\,L^{-1}$
f	Darcy or fanning friction factor for flow through a pipe or duct	none	none	—
f′	Friction coefficient for flow over banks of tubes	none	none	—
F	Force; F_B, buoyant force	N	lb	$M\,L\,t^{-2}$
F_T	Temperature factor	none	none	—
$F_{1\text{-}2}$	Geometric shape factor for radiation from one blackbody to another	none	none	—
g	Acceleration due to gravity	m/s^2	ft/s^2	$L\,t^{-2}$
g_c	Dimensional conversion factor	$1.0\ kg{\cdot}m/N{\cdot}s^2$	$32.2\ ft{\cdot}lb_m/lb{\cdot}s^2$	
G	Mass velocity or flow rate per unit area	$kg/s{\cdot}m^2$	$lb_m/h{\cdot}ft^2$	$M\,L^{-2}\,t^{-1}$
G	Irradiation incident on unit surface in unit time	W/m^2	$Btu/h{\cdot}ft^2$	$M\,L^{-2}\,t^{-1}$
h	Enthalpy per unit mass	J/kg	Btu/lb_m	$L^2\,t^{-2}$
h	Local heat transfer coefficient; $\bar{h}$, average heat transfer coefficient $\bar{h}=\bar{h}_c+\bar{h}_r$; $\bar{h}_b$, heat transfer coefficient of a boiling liquid; $\bar{h}_c$, local convection heat transfer coefficient; $\bar{h}_c$, average heat transfer coefficient; $\bar{h}_r$, average heat transfer coefficient for radiation	$W/m^2{\cdot}K$	$Btu/h{\cdot}ft^2{\cdot}°F$	$M\,t^{-3}\,T^{-1}$

h_{fg}	Latent heat of condensation or evaporation	J/kg	Btu/lb_m	$L^2 t^{-2}$
H	Head, elevation of hydraulic grade line	m	ft	L
i	Angle between sun direction and surface normal	rad	deg	—
I	Moment of inertia	m^4	ft^4	L^4
I	Intensity of radiation	W/sr	Btu/h unit solid angle	$M L^2 t^{-3}$
I_λ	Intensity per unit wavelength	W/sr·μm	Btu/h·sr micron	$M L t^{-3}$
J	Radiosity	W/m^2	$Btu/h·ft^2$	$M L^{-2} t^{-1}$
k	Thermal conductivity; k_s, thermal conductivity of a solid; k_f, thermal conductivity of a fluid; k_g, thermal conductivity of a gas	W/m·K	Btu/h·ft°F	$M L^{-2} t^{-1} T^{-1}$
K	Thermal conductance; k_k, thermal conductance for conduction heat transfer; k_c, thermal convection conductance; K_r, thermal conduction for radiation heat transfer	W/K	Btu/h·ft°F	$M t^{-1} T^{-1}$
K	Bulk modulus of elasticity	Pa	lb/ft^2	$M L^{-1} t^{-2}$
log	Logarithm to the base 10	none	none	—
ln	Logarithm to the base e	none	none	—
l	Length, general or characteristic length of a body	m	ft	L
L	Lift force	N	lb	$M L t^{-2}$
L_f	Latent heat of solidification	J/kg	Btu/lb_m	$L^2 t^{-2}$
$\dot{m}$	Mass flow rate	kg/s	lb_m/s	$M t^{-1}$
m	Mass	kg	lb_m	M
M	Molecular weight	gm/gm mole	lb_m/lb mole	—
$\dot{M}$	Momentum per unit time	N	lb	MLt^{-2}
n	Manning roughness factor	none	none	—
N	Number of moles	none	none	—
NPSH	Net positive suction head	m	ft	L
N	Number in general; number of tubes, etc.	none	none	—
P	Static pressure; p_c, critical pressure; p_A, partial pressure of component A	N/m^2	psi or lb/ft^2 or atm	$M L^{-1} t^{-2}$
P	Wetted perimeter or height of weir	m	ft	L
q	Discharge per unit width	m^2/s	ft^2/s	$L^2 t^{-1}$
q	Rate of heat flow; q_k, rate of heat flow by conduction; q_r, rate of heat flow by radiation; q_c, rate of heat flow by convection; q_b, rate of heat flow by nucleate boiling	W	Btu/h	$M L^2 t^{-3}$
q'''	Rate of heat generation per unit volume	W/m^3	$Btu/h·ft^3$	$M L^{-1} t^{-3}$
q''	Rate of heat generation per unit area (heat flux)	W/m^2	$Btu/h·ft^2$	$M t^{-3}$
Q	Quantity of heat	J	Btu	$M L^2 t^{-3}$
r	Radius; r_H, hydraulic radius; r_i, inner radius; r_o, outer radius	m	ft	L
R	Thermal resistance; R_c, thermal resistance to convection heat transfer; R_k, thermal resistance to conduction heat transfer; R_r, to radiation heat transfer	K/W	h°F/Btu	$L T M^{-1}$
R_e	Electrical resistance	ohm	ohm	—
R	Universal gas constant	8.314 J/K·kg mole	1545 $ft·lb_f$/lb·mole°F	$L^2 t^{-2} T^{-1}$
s	Entropy per unit mass	J/kg·K	$ft·lb/lb_m·°R$	$L^2t^{-2}T^{-1}$
S	Entropy	J/K	ft·lb/°R	$ML^2t^{-2}T^{-1}$

S_L	Distance between centerlines of tubes in adjacent longitudinal rows	m	ft	L
S_T	Distance between centerlines of tubes in adjacent transverse rows	m	ft	L
t	Time	s	h or s	t
T	Temperature; T_b, temperature of bulk of fluid; T_f, mean film temperature; T_s, surface temperature, T_o, temperature of fluid far removed from heat source or sink; T_m, mean bulk temperature of fluid flowing in a duct; T_M, temperature of saturated vapor; T_{sl}, temperature of a saturated liquid; T_{fr}, freezing temperature; T_t, liquid temperature; T_{as}, adiabatic wall temperature	K or °C	°F or R	T
u	Internal energy per unit mass	J/kg	Btu/lb_m	$L^2\ t^{-2}$
u	Velocity in x direction; u′, instantaneous fluctuating x component of velocity; ū, average velocity	m/s	ft/s or ft/h	$L\ t^{-1}$
u*	Shear stress velocity	m/s	ft/s	Lt^{-1}
U	Internal energy	J	Btu	ML^2t^{-2}
U	Overall heat transfer coefficient	W/m^2K	$Btu/h{\cdot}ft^{2\circ}F$	$M\ t^{-3}\ T^{-1}$
U_∞	Free-stream velocity	m/s	ft/s	$L\ t^{-1}$
v	Specific volume	m^3/kg	ft^3/lb_m	$L^3\ M^{-1}$
v	Velocity in y direction; v′, instantaneous fluctuating y component of velocity	m/s	ft/s or ft/h	$L\ t^{-1}$
V	Volume	m^3	ft^3	L^3
$\dot{V}$	Volumetric flow rate	m^3/s	ft^3/s	$L^3\ t^{-1}$
W_s	Shaft work	m·N	ft·lb	ML^2t^{-2}
$\dot{W}$	Rate of work output or power	W	Btu/h	$M\ L^2\ t^{-3}$
x	Coordinate or distance from the leading edge; x_c, critical distance from the leading edge where flow becomes turbulent	m	ft	L
x	Quality	percent	percent	none
y	Coordinate or distance from a solid boundary measured in direction normal to surface	m	ft	L
z	Coordinate	m	ft	L
Z	Ratio of hourly heat capacity rates in heat exchangers	none	none	—

Greek Symbols

α	Absorptivity for radiation, α_λ, monochromatic absorptivity at wavelength λ	none	none	—
α	Thermal diffusivity = k/ρc	m^2/s	ft^2/s	$L^2\ t^{-1}$
β	Temperature coefficient of volume expansion	l/K	1/R	T^{-1}
β_k	Temperature coefficient of thermal conductivity	l/K	1/R	T^{-1}
γ	Specific heat ratio, c_p/c_v	none	none	—
Γ	Circulation	m^2	ft^2	L^2t^{-1}
Γ	Body force per unit mass	N/kg	lb/lb_m	$L\ t^{-2}$
Γ_c	Mass rate of flow of condensate per unit breadth = $\dot{m}/\pi D$ for a vertical tube	kg/s·m	$lb_m/h{\cdot}ft$	$M\ L^{-2}\ t^1$
δ	Boundary-layer thickness; δ_h, hydrodynamic boundary-layer thickness; δ_{th}, thermal boundary-layer thickness	m	ft	L
Δ	Difference between values	none	none	—

ε	Heat exchanger effectiveness	none	none	—
ε	Roughness height	m	ft	L
ε	Emissivity for radiation; ε_λ, monochromatic emissivity at wavelength λ; ε_φ, emissivity in direction φ	none	none	—
ε_H	Thermal eddy diffusivity	m^2/s	ft^2/s	$L^2\,t^{-1}$
ε_M	Momentum eddy diffusivity	m^2/s	ft^2/s	$L^2\,t^{-1}$
ζ	Ratio of thermal to hydrodynamic boundary-layer thickness, δ_h/δ_{th}	—	—	—
η	Efficiency; η_f, fin efficiency	none	none	—
λ	Wavelength; λ_{max}, wavelength at which monochromatic emissive power $E_{b\lambda}$ is a maximum	μm	micron	L
μ	Absolute viscosity	$N{\cdot}s/m^2$	$lb/ft{\cdot}s$	$M\,L^{-1}\,t^{-1}$
ν	Kinematic viscosity, μ/ρ	m^2/s	ft^2/s	$L^2\,t^{-1}$
ν_f	Frequency of radiation	1/s	1/s	t^{-1}
Φ	Velocity potential	m^2/s	ft^2/s	$L^2\,t^{-1}$
ρ	Mass density, 1/v; ρ_l, density of liquid; ρ_v, density of vapor	kg/m^3	lb_m/ft^3	$M\,L^{-3}$
τ	Shearing stress, τ_s, shearing stress at surface; τ_w, shear at wall of a tube or a duct	N/m^2	lb/ft^2	$M\,L^{-1}\,t^{-2}$
τ	Transmissivity for radiation	none	none	—
σ	Stefan-Boltzmann constant	W/m^2K^4	$Btu/h\,ft^2R^4$	$M\,t^{-3}\,T^{-4}$
σ	Surface tension	N/m	lb/ft	$M\,t^{-2}$
φ	Angle	rad	rad	—
ψ	Stokes' stream function	m^3/s	ft^3/s	$L^3\,t^{-1}$
ω	Angular velocity	1/s	1/s	t^{-1}
ω	Solid angle	sr	steradian	—

Dimensionless Numbers

Bi	Biot number
Ec	Eckert number
Eu	Euler number
Fo	Fourier modulus
Fr	Froude number
Gz	Graetz number
Gr	Grahsof number
Ja	Jakob number
Kn	Knudsen number
M	Mach number
Nu	Average Nusselt number; Nu_D, average diameter Nusselt number; Nu_x, local Nusselt number
Pe	Peclet number
Pr	Prandtl number
Ra	Rayleigh number
Re	Reynolds number; Re_x, local value of Re at a distance x from leading edge; Re_D, diameter Reynolds number; Re_b, bubble Reynolds number
Θ	Boundary Fourier modulus or dimensionless time
St	Stanton number

We	Weber number

Miscellaneous

$a > b$	a greater than b
$a < b$	a smaller than b
$\propto$	Proportional sign
$\simeq$	Approximately equal sign
∞	Infinity sign
Σ	Summation sign

Subscripts

c	critical condition
i	inlet
f	fin
u	unit quantities
w	wall or properties at wall temperature
c.s.	control surface
c.v.	control volume
o	stagnation or standard state condition; outlet or outside
1,2	inlet and outlet, respectively, of control volume

Note: Those symbols and subscripts that are not included in the above list are defined in the text.

Editor

Dr. Raj P. Chhabra is a professor of chemical engineering at the Indian Institute of Technology Kanpur. He received his MS degree from the Indian Institute of Science, Bangalore, and his PhD from Monash University, both in chemical engineering. Dr. Chhabra is a fellow of the Indian Academy of Sciences, the National Academy of Science (India), and the Indian Academy of Engineering. His current research interests include hydrodynamics of non-Newtonian fluid particle systems, multiphase flow in pipes and packed beds, interaction between nonlinear properties of fluids and shapes of particles, and the modeling of Newtonian and non-Newtonian flow in fibrous media.

Contributors

Amit Agrawal
Department of Mechanical Engineering
Indian Institute of Technology Bombay
Mumbai, India

Vasilios Alexiades
Department of Mathematics
University of Tennessee
Knoxville, Tennessee

Christian K. Bach
Department of Mechanical and Aerospace Engineering
Oklahoma State University
Stillwater, Oklahoma

Amitabh Bhattacharya
Department of Mechanical Engineering
Indian Institute of Technology Bombay
Mumbai, India

Santanu Bandyopadhyay
Department of Energy Science and Engineering
Indian Institute of Technology Bombay
Mumbai, India

Ian H. Bell
Ray W. Herrick Laboratories
Purdue University
West Lafayette, Indiana

Ken Bell
Department of Chemical Engineering
Oklahoma State University
Stillwater, Oklahoma

Stanley A. Berger
Department of Mechanical Engineering
University of California
Berkeley, California

Rajneesh Bhardwaj
Department of Mechanical Engineering
Indian Institute of Technology Bombay
Mumbai, India

Robert F. Boehm
Department of Mechanical Engineering
University of Nevada
Las Vegas, Nevada

Craig R. Bradshaw
Torad Engineeering, LLC
Cumming, Georgia

Massimo Capobianchi
Department of Mechanical Engineering
Gonzaga University
Spokane, Washington

Van P. Carey
Department of Mechanical Engineering
University of California
Berkeley, California

Mangesh Chaudhari
Department of Mechanical Engineering
Vishwakarma Institute of Technology
Pune, India

Raj P. Chhabra
Department of Chemical Engineering
Indian Institute of Technology
Kanpur, India

John C. Chen
Department of Chemical Engineering
Lehigh University
Bethlehem, Pennsylvania

Stuart W. Churchill
Department of Chemical and Bioengineering
University of Pennsylvania
Philadelphia, Pennsylvania

Gerardo Diaz
School of Engineering
University of California
Merced, California

Kenneth R. Diller
Department of Biomedical Engineering
University of Texas
Austin, Texas

Donald L. Fenton
Mechanical and Nuclear Engineering
Kansas State University
Manhattan, Kansas

P.S. Ghoshdastidar
Department of Mechanical Engineering
Indian Institute of Technology
Kanpur, India

Eckhard A. Groll
School of Mechanical Engineering
Purdue University
West Lafayette, Indiana

Anoop K. Gupta
Department of Chemical Engineering
National University of Singapore
Singapore

K.G. Terry Hollands
Department of Mechanical Engineering
University of Waterloo
Waterloo, Ontario, Canada

K. Hriseehkesh
Department of Mechanical Engineering
Indian Institute of Technology Bombay
Mumbai, India

Thomas F. Irvine, Jr.
Department of Mechanical Engineering
State University of New York
Stony Brook, New York

Harold R. Jacobs
Department of Mechanical Engineering
Pennsylvania State University
University Park, Pennsylvania

Milind A. Jog
Department of Mechanical Engineering
University of Cincinnati
Cincinnati, Ohio

Shyamprasad Karagadde
Department of Mechanical Engineering
Indian Institute of Technology Bombay
Mumbai, India

Sameer Khandekar
Department of Mechanical Engineering
Indian Institute of Technology
Kanpur, India

Jungho Kim
Department of Mechanical Engineering
University of Maryland
College Park, Maryland

Clement Kleinstreuer
Department of Mechanical & Aerospace Engineering
North Carolina State University
Raleigh, North Carolina

Jan Kośny
Fraunhofer Center for Sustainable Energy Systems
Boston, Massachusetts

Moncef Krarti
Department of Civil, Environmental, and Architectural Engineering
University of Colorado
Boulder, Colorado

Abhinav Krishna
Eaton Corporation
Southfield, Michigan

Salil S. Kulkarni
Department of Mechanical Engineering
Indian Institute of Technology
Mumbai, India

Prabhat Kumar
Pepsico Global R&D
Plano, Texas

Orkan Kurtulus
Ray W. Herrick Laboratories
Purdue University
West Lafayette, Indiana

Pradeep Lall
Motorola
Libertyville, Illinois
and
Department of Mechanical Engineering
Auburn University
Auburn, Alabama

John C. Leylegian
Department of Mechanical Engineering
Manhattan College
Riverdale, New York

Jun Ma
Department of Mechanical Science & Engineering
University of Illinois at Urbana–Champaign
Urbana, Illinois

Raj M. Manglik
Department of Mechanical, Industrial and Nuclear Engineering
University of Cincinnati
Cincinnati, Ohio

Margaret M. Mathison
Department of Mechanical Engineering
Iowa State University
Ames, Iowa

Evelyne Mauret
Centre National de la Recherche Scientifique
Université Grenoble Alpes
Grenoble, France

Alan T. McDonald
Department of Mechanical Engineering
Purdue University
West Lafayette, Indiana

Anthony F. Mills
Mechanical & Aerospace Engineering Department
University of California
Los Angeles, California

Michael F. Modest
Department of Mechanical Engineering
University of California
Merced, California

Robert J. Moffat
Department of Mechanical Engineering
Stanford University
Stanford, California

Michael J. Moran
Department of Mechanical Engineering
Ohio State University
Columbus, Ohio

Shankar Narasimhan
Department of Chemical Engineering
Indian Institute of Technology Madras
Chennai, India

Paul Norton
National Renewable Energy Laboratory
Golden, Colorado

Tadhg O'Donovan
Mechanical, Process & Energy Engineering
Heriot-Watt University
Edinburgh, United Kingdom

Arturo Pacheco-Vega
Department of Mechanical Engineering
California State University
Los Angeles, California

Raj Ganesh S. Pala
Department of Chemical Engineering
Indian Institute of Technology Kanpur
Kanpur, India

Swati A. Patel
Department of Chemical Engineering
Indian Institute of Technology
Ropar, India

John A. Pearce
Department of Electrical & Computer Engineering
University of Texas
Austin, Texas

Graham T. Polley
Cal Gavin Ltd
Minerva Mill
Alcester, United Kingdom

George D. Raithby
Department of Mechanical Engineering
University of Waterloo
Waterloo, Ontario, Canada

Joshua D. Ramsey
Department of Chemical Engineering
Oklahoma State University
Stillwater, Oklahoma

Rolf D. Reitz
Department of Mechanical Engineering
University of Wisconsin
Madison, Wisconsin

Martine Rueff
Centre National de la Recherche Scientifique
Université Grenoble Alpes
Grenoble, France

K.P. Sandeep
Department of Food, Bioprocessing, & Nutrition Sciences
North Carolina State University
Raleigh, North Carolina

Mihir Sen
Department of Aerospace and Mechanical Engineering
University of Notre Dame
South Bend, Indiana

Bryce Shaffer
Air Squared, Inc
Broomfield, Colorado

Ramesh K. Shah
Delphi Harrison Thermal Systems
Lockport, New York

Atul Sharma
Department of Mechanical Engineering
Indian Institute of Technology Bombay
Mumbai, India

S.A. Sherif
Department of Mechanical and Aerospace Engineering
University of Florida
Gainesville, Florida

Desh Bandhu Singh
Centre for Energy Studies
Indian Institute of Technology
New Delhi, India

Sanjiv Sinha
Department of Mechanical Science & Engineering
University of Illinois at Urbana–Champaign
Urbana, Illinois

Mariia Sobulska
Department of Heat & Mass Transfer
Lodz University of Technology
Lodz, Poland

Bengt Sundén
Division of Heat Transfer Energy Sciences
Lund University
Lund, Sweden

N.V. Suryanarayana
Department of Mechanical Engineering
Michigan Technological University
Houghton, Michigan

Rajappa Tadepalli
Saint-Gobain Research India
Chennai, India

G. N. Tiwari
Centre for Energy Studies
Indian Institute of Technology
New Delhi, India

Naveen Tiwari
Department of Chemical Engineering
Indian Institute of Technology
Kanpur, India

Mario F. Trujillo
Department of Mechanical Engineering
University of Wisconsin
Madison, Wisconsin

George Tsatsaronis
Institute for Energy Engineering
Technical University of Berlin
Berlin, Germany

Blake Paul Tullis
Department of Civil Engineering
Utah State University
Logan, Utah

J. Paul Tullis
Department of Civil Engineering
Utah State University
Logan, Utah

Krishna Valavala
Department of Mechanical Science & Engineering
University of Illinois at Urbana–Champaign
Urbana, Illinois

Jonathan W. Valvano
Department of Electrical and Computer Engineering
University of Texas
Austin, Texas

Pawel Wawrzyniak
Department of Heat & Mass Transfer
Lodz University of Technology
Lodz, Poland

Frank M. White
Department of Mechanical Engineering
University of Rhode Island
Kingston, Rhode Island

D. Ian Wilson
Department of Chemical Engineering & Biotechnology
University of Cambridge
Cambridge, England

Zan Wu
Division of Heat Transfer Energy Sciences
Lund University
Lund, Sweden

Zelin Xu
Department of Mechanical and Aerospace Engineering North Carolina State University
Raleigh, North Carolina

Bin Yang
Ray W. Herrick Laboratories
Purdue University
West Lafayette, Indiana

K.T. Yang
Department of Mechanical and Aerospace Engineering
University of Notre Dame
South Bend, Indiana

David W. Yarbrough
R&D Services, Inc.
Cookeville, Tennessee

Ireneusz Zbicinski
Faculty of Process and Environmental Engineering
Lodz University of Technology
Lodz, Poland

John Zhai
Department of Civil, Environmental and Architectural Engineering
University of Colorado
Boulder, Colorado

Xinye Zhang
Ray W. Herrick Laboratories
Purdue University
West Lafayette, Indiana

Davide Ziviani
Ray W. Herrick Laboratories
Purdue University
West Lafayette, Indiana

1 Engineering Thermodynamics

Michael J. Moran and George Tsatsaronis

CONTENTS

1.1 Fundamentals 3
Basic Concepts and Definitions 3
System 3
State, Property 3
Process, Cycle 4
Phase and Pure Substance 4
Equilibrium 4
Temperature 4
The First Law of Thermodynamics, Energy 5
Work 6
Energy 6
Heat 7
Power Cycles 8
The Second Law of Thermodynamics, Entropy 8
Kelvin–Planck Statement 8
Irreversibilities 9
Carnot Corollaries 10
Kelvin Temperature Scale 10
Carnot Efficiency 11
The Clausius Inequality 11
Entropy and Entropy Generation 12
1.2 Control Volume Applications 15
Conservation of Mass 15
Control Volume Energy Balance 15
Control Volume Entropy Balance 16
Control Volumes at Steady State 17
Internally Reversible Heat Transfer and Work 21
1.3 Property Relations and Data 23
Basic Relations for Pure Substances 23
Maxwell Relations 24
Specific Heats and Other Properties 25
P–v–T Relations 27
P–v–T Surface 29
Graphical Representations 31
Compressibility Charts 32
Equations of State 34
Gas Mixtures 35
Evaluating Δh, Δu, and Δs 37
Fundamental Thermodynamic Functions 39
Thermodynamic Data Retrieval 40
Ideal Gas Model 43

Polytropic Processes 45
Ideal Gas Mixtures 46
Moist Air 49
Generalized Charts for Enthalpy, Entropy, and Fugacity 50
Multicomponent Systems 52
Chemical Potential 54
1.4 Combustion 56
Reaction Equations 56
Property Data for Reactive Systems 58
Enthalpy of Formation 58
Absolute Entropy 60
Gibbs Function of Formation 63
Reaction Equilibrium 64
1.5 Exergy Analysis 67
Defining Exergy 67
Environment 67
Dead States 68
Exergy Balances 68
Control Volume Exergy Rate Balance 68
Chemical Exergy 69
Exergetic Efficiency 71
Introduction to Exergy Costing 74
1.6 Vapor and Gas Power Cycles 76
Rankine and Brayton Cycles 76
Otto, Diesel, and Dual Cycles 78
Carnot, Ericsson, and Stirling Cycles 82
1.7 Guidelines for Improving Thermodynamic Effectiveness 84
1.8 Exergoeconomics 86
Exergy Costing 87
Cost Balance 88
Auxiliary Costing Equations 89
General Example 90
Exergoeconomic Variables and Evaluation 91
1.9 Design Optimization 94
An Iterative Exergoeconomic Procedure for Optimizing the Design of a Thermal System 95
Case Study 96
First Design Case 98
Second Design Case 99
Third Design Case 100
Additional Iterations 101
1.10 Economic Analysis of Thermal Systems 102
Estimation of Total Capital Investment 102
Principles of Economic Evaluation 102
Annuities 104
Cost Escalation 104
Levelization 104
Parameters and Assumptions 105
Calculation of the Product Costs 105
1.11 Exergoenvironmental Analysis 106
1.12 Advanced Exergy-Based Methods 107
References 107

Although various aspects of what is now known as thermodynamics have been of interest since antiquity, formal study began only in the early 19th century through consideration of the motive power of *heat*: the capacity of hot bodies to produce *work*. Today, the scope is larger, dealing generally with *energy and entropy*, and with relationships among the *properties* of matter. Moreover, in recent decades, engineering thermodynamics has undergone a revolution, both in terms of the presentation of fundamentals and in the manner that it is applied. In particular, the second law of thermodynamics has emerged as an effective tool for engineering analysis and design.

1.1 FUNDAMENTALS

Michael J. Moran

Classical thermodynamics is concerned primarily with the macrostructure of matter. It addresses the gross characteristics of large aggregations of molecules and not the behavior of individual molecules. The microstructure of matter is studied in kinetic theory and statistical mechanics (including quantum thermodynamics). In this chapter, the classical approach to thermodynamics is featured.

Basic Concepts and Definitions

Thermodynamics is both a branch of physics and an engineering science. The scientist is normally interested in gaining a fundamental understanding of the physical and chemical behavior of fixed, quiescent quantities of matter and uses the principles of thermodynamics to relate the *properties* of matter. Engineers are generally interested in studying *systems* and how they interact with their *surroundings*. To facilitate this, engineers have extended the subject of thermodynamics to the study of systems through which matter flows.

System

In a thermodynamic analysis, the *system* is the subject of the investigation. Normally the system is a specified quantity of matter and/or a region that can be separated from everything else by a well-defined surface. The defining surface is known as the *control surface* or *system boundary.* The control surface may be movable or fixed. Everything external to the system is the *surroundings*. A system of fixed mass is referred to as a *control mass* or as a *closed system*. When there is flow of mass through the control surface, the system is called a *control volume*, or an *open system*. An *isolated* system is a closed system that does not interact in any way with its surroundings.

State, Property

The condition of a system at any instant of time is called its *state*. The state at a given instant of time is described by the properties of the system. A *property* is any quantity whose numerical value depends on the state but not the history of the system. The value of a property is determined in principle by some type of physical operation or test.

Extensive properties depend on the size or extent of the system. Volume, mass, energy, and entropy are examples of extensive properties. An extensive property is additive in the sense that its value for the whole system equals the sum of the values for its parts. *Intensive* properties are independent of the size or extent of the system. Pressure and temperature are examples of intensive properties.

A *mole* is a quantity of substance having a mass numerically equal to its molecular weight. Designating the molecular weight by M and the number of moles by n, the mass m of the substance is $m=nM$. One kilogram mole, designated kmol, of oxygen is 32.0 kg and one pound mole (lb mol) is 32.0 lb. When an extensive property is reported on a unit mass or a unit mole basis, it is called a *specific* property. An overbar is used to distinguish an extensive property written on a per mole basis

from its value expressed per unit mass. For example, the volume per mole is $\overline{v}$, whereas the volume per unit mass is v, and the two specific volumes are related by $\overline{v} = Mv$.

Process, Cycle

Two states are identical if, and only if, the properties of the two states are identical. When any property of a system changes in value, there is a change in state, and the system is said to undergo a *process*. When a system in a given initial state goes through a sequence of processes and finally returns to its initial state, it is said to have undergone a *cycle*.

Phase and Pure Substance

The term *phase* refers to a quantity of matter that is homogeneous throughout in both chemical composition and physical structure. Homogeneity in physical structure means that the matter is all *solid*, or all *liquid*, or all *vapor* (or, equivalently, all *gas*). *A* system can contain one or more phases. For example, a system of liquid water and water vapor (steam) contains *two* phases. A *pure substance* is one that is uniform and invariable in chemical composition. *A* pure substance can exist in more than one phase, but its chemical composition must be the same in each phase. For example, if liquid water and water vapor form a system with two phases, the system can be regarded as a pure substance because each phase has the same composition. The nature of phases that coexist in equilibrium is addressed by the *phase rule* (Section 1.3, Multicomponent Systems).

Equilibrium

Equilibrium means a condition of balance. In thermodynamics, the concept includes not only a balance of forces, but also a balance of other influences. Each kind of influence refers to a particular aspect of thermodynamic (complete) equilibrium. *Thermal* equilibrium refers to an equality of temperature, *mechanical* equilibrium to an equality of pressure, and *phase* equilibrium to an equality of chemical potentials (Section 1.3, Multicomponent Systems). *Chemical* equilibrium is also established in terms of chemical potentials (Section 1.4, Reaction Equilibrium). For complete equilibrium, the several types of equilibrium must exist individually.

To determine if a system is in thermodynamic equilibrium, one may think of testing it as follows: isolate the system from its surroundings and watch for changes in its observable properties. If there are no changes, it may be concluded that the system was in equilibrium at the moment it was isolated. The system can be said to be at an *equilibrium state*. When a system is *isolated*, it cannot interact with its surroundings; however, its state can change as a consequence of spontaneous events occurring internally as its intensive properties, such as temperature and pressure, tend toward uniform values. When all such changes cease, the system is in equilibrium. At equilibrium, temperature and pressure are uniform throughout. If gravity is significant, a pressure variation with height can exist, as in a vertical column of liquid.

Temperature

A scale of temperature independent of the *thermometric substance* is called a *thermodynamic* temperature scale. The Kelvin scale, a thermodynamic scale, can be elicited from the second law of thermodynamics (Section 1.1, The Second Law of Thermodynamics, Entropy). The definition of temperature following from the second law is valid over all temperature ranges and provides an essential connection between the several *empirical* measures of temperature. In particular, temperatures evaluated using a *constant volume gas thermometer* are identical to those of the Kelvin scale over the range of temperatures where gas thermometry can be used.

The empirical *gas scale* is based on the experimental observations that (1) at a given temperature level, all gases exhibit the same value of the product $p\bar{v}$ (p is pressure and $\bar{v}$ the specific volume on a molar basis) if the pressure is low enough and (2) the value of the product $p\bar{v}$ increases with the temperature level. On this basis, the gas temperature scale is defined by

$$T = \frac{1}{\bar{R}} \lim_{p \to 0} (p\bar{v}),$$

where T is temperature and $\bar{R}$ is the *universal gas constant*. The absolute temperature at the *triple point of water* (Section 1.3, *P–v–T* Relations) is fixed by international agreement to be 273.16 K on the *Kelvin* temperature scale. $\bar{R}$ is then evaluated experimentally as $\bar{R} = 8.314\ \text{kJ/kmol} \cdot \text{K}$ (1545 ft lbf/lbmol °R).

The *Celsius temperature scale* (also called the centigrade scale) uses the degree Celsius (°C), which has the same magnitude as the Kelvin. Thus, temperature *differences* are identical on both scales. However, the zero point on the Celsius scale is shifted to 273.15 K, as shown by the following relationship between the Celsius temperature and the Kelvin temperature:

$$T(^\circ\text{C}) = T(\text{K}) - 273.15. \tag{1.1.1}$$

On the Celsius scale, the triple point of water is 0.01°C and 0 K corresponds to –273.15°C.

Two other temperature scales are commonly used in engineering in the United States. By definition, the *Rankine scale*, the unit of which is the degree Rankine (°R), is proportional to the Kelvin temperature according to

$$T(^\circ\text{R}) = 1.8T(\text{K}). \tag{1.1.2}$$

The Rankine scale is also an absolute thermodynamic scale with an absolute zero that coincides with the absolute zero of the Kelvin scale. In thermodynamic relationships, temperature is always in terms of the Kelvin or Rankine scale unless specifically stated otherwise.

A degree of the same size as that on the Rankine scale is used in the *Fahrenheit scale*, but the zero point is shifted according to the relation

$$T(^\circ\text{F}) = T(^\circ\text{R}) - 459.67. \tag{1.1.3}$$

Substituting Equations 1.1.1 and 1.1.2 into Equation 1.1.3 gives

$$T(^\circ F) = 1.8T(^\circ C) + 32. \tag{1.1.4}$$

This equation shows that the Fahrenheit temperature of the *ice point* (0°C) is 32°F and of the *steam point* (100°C) is 212°F. The 100 Celsius or Kelvin degrees between the ice point and steam point corresponds to 180 Fahrenheit or Rankine degrees.

To provide a standard for temperature measurement taking into account both theoretical and practical considerations, the International Temperature Scale of 1990 (ITS-90) is defined in such a way that the temperature measured on it conforms with the thermodynamic temperature, the unit of which is the Kelvin, to within the limits of accuracy of measurement obtainable in 1990. Further discussion of ITS-90 is provided by Preston-Thomas (1990).

The First Law of Thermodynamics, Energy

Energy is a fundamental concept of thermodynamics and one of the most significant aspects of engineering analysis. Energy can be *stored* within systems in various macroscopic forms: kinetic energy, gravitational potential energy, and internal energy. Energy can also be *transformed* from

one form to another and *transferred* between systems. For closed systems, energy can be transferred by *work* and *heat transfer*. The total amount of energy is *conserved* in all transformations and transfers.

Work

In thermodynamics, the term *work* denotes a means for transferring energy. Work is an effect of one system on another that is identified and measured as follows: Work is done by a system on its surroundings if the *sole effect* on everything external to the system *could have been* the raising of a weight. The test of whether a work interaction has taken place is not that the elevation of a weight is actually changed, nor that a force actually acted through a distance, but that the sole effect *could be* the change in elevation of a weight. The magnitude of the work is measured by the number of standard weights that could have been raised. Since the raising of a weight is in effect a force acting through a distance, the work concept of mechanics is preserved. This definition includes work effects such as is associated with rotating shafts, displacement of the boundary, and the flow of electricity.

Work done *by* a system is considered positive: $W > 0$. Work done *on* a system is considered negative: $W < 0$. The time rate of doing work, or *power*, is symbolized by $\dot{W}$ and adheres to the same sign convention.

Energy

A closed system undergoing a process that involves only work interactions with its surroundings experiences an *adiabatic* process. On the basis of experimental evidence, it can be postulated that *when a closed system is altered adiabatically, the amount of work is fixed by the end states of the system and is independent of the details of the process*. This postulate, which is one way the *first law of thermodynamics* can be stated, can be made regardless of the type of work interaction involved, the type of process, or the nature of the system.

As the work in an adiabatic process of a closed system is fixed by the end states, an extensive property called *energy* can be defined for the system such that its change between two states is the work in an adiabatic process that has these as the end states. In engineering thermodynamics, the change in the energy of a system is considered to be made up of three macroscopic contributions: the change in *kinetic energy*, *KE*, associated with the motion of the system *as a whole* relative to an external coordinate frame, the change in gravitational *potential energy*, *PE*, associated with the position of the system *as a whole* in the Earth's gravitational field, and the change in *internal energy*, *U*, which accounts for all other energy associated with the system. Like kinetic energy and gravitational potential energy, internal energy is an extensive property.

In summary, the change in energy between two states of a closed system in terms of the work W_{ad} of an adiabatic process between these states is

$$\left(\mathrm{KE}_2 - \mathrm{KE}_1\right) + \left(\mathrm{PE}_2 - \mathrm{PE}_1\right) + \left(U_2 - U_1\right) = -W_{ad}, \tag{1.1.5}$$

where 1 and 2 denote the initial and final states, respectively, and the minus sign before the work term is in accordance with the previously stated sign convention for work. Since any arbitrary value can be assigned to the energy of a system at a given state 1, no particular significance can be attached to the value of the energy at state 1 or at *any* other state. Only *changes* in the energy of a system have significance.

The specific energy (energy per unit mass) is the sum of the specific internal energy, u, the specific kinetic energy, $v^2/2$, and the specific gravitational potential energy, gz, such that

$$\text{Specific energy} = u + \frac{v^2}{2} + gz, \tag{1.1.6}$$

where the velocity v and the elevation z are each relative to specified datums (often the Earth's surface) and g is the acceleration of gravity.

A property related to internal energy u, pressure p, and specific volume v is *enthalpy*, defined by

$$h = u + pv \tag{1.1.7a}$$

or on an extensive basis

$$H = U + pV \tag{1.1.7b}$$

Heat

Closed systems can also interact with their surroundings in a way that cannot be categorized as work, as, for example, a gas (or liquid) contained in a closed vessel undergoing a process while in contact with a flame. This type of interaction is called a *heat interaction*, and the process is referred to as *nonadiabatic*.

A fundamental aspect of the energy concept is that energy is conserved. Thus, since a closed system experiences precisely the same energy change during a nonadiabatic process as during an adiabatic process between the same end states, it can be concluded that the *net* energy transfer to the system in each of these processes must be the same. It follows that heat interactions also involve energy transfer. Denoting the amount of energy transferred *to* a closed system in heat interactions by Q, these considerations can be summarized by the *closed system energy balance*:

$$\left(U_2 - U_1\right) + \left(\mathrm{KE}_2 - \mathrm{KE}_1\right) + \left(\mathrm{PE}_2 - \mathrm{PE}_1\right) = Q - W. \tag{1.1.8}$$

The closed system energy balance expresses the conservation of energy principle for closed systems of all kinds.

The quantity denoted by Q in Equation 1.1.8 accounts for the amount of energy transferred to a closed system during a process by means other than work. On the basis of an experiment, it is known that such an energy transfer is induced only as a result of a temperature difference between the system and its surroundings and occurs only in the direction of decreasing temperature. This means of energy transfer is called an *energy transfer by heat*. The following sign convention applies:

$$Q > 0 : \text{heat transfer to the system}$$

$$Q < 0 : \text{heat transfer from the system}$$

The time rate of heat transfer, denoted by $\dot{Q}$, adheres to the same sign convention.

Methods based on experiment are available for evaluating energy transfer by heat. These methods recognize two basic transfer mechanisms: *conduction* and *thermal radiation*. In addition, theoretical and empirical relationships are available for evaluating energy transfer involving *combined* modes such as *convection*. Further discussion of heat-transfer fundamentals is provided in Chapter 3.

The quantities symbolized by W and Q account for *transfers* of energy. The terms *work* and *heat* denote different *means* whereby energy is transferred and not *what* is transferred. Work and heat are not properties, and it is improper to speak of work or heat "contained" in a system. However, to achieve economy of expression in subsequent discussions, W and Q are often referred to simply as work and heat transfer, respectively. This less formal approach is commonly used in engineering practice.

Power Cycles

Since energy is a property, over each cycle there is no net change in energy. Thus, Equation 1.1.8 reads for *any* cycle

$$Q_{\text{cycle}} = W_{\text{cycle}}.$$

That is, for *any* cycle the net amount of energy received through heat interactions is equal to the net energy transferred out in work interactions. A *power cycle*, or *heat* engine, is one for which a net amount of energy is transferred out by work: $W_{\text{cycle}} > 0$. This equals the net amount of energy transferred in by heat.

Power cycles are characterized both by addition of energy by heat transfer, Q_{A}, and inevitable rejections of energy by heat transfer, Q_{R}:

$$Q_{\text{cycle}} = Q_{\text{A}} - Q_{\text{R}}.$$

Combining the last two equations,

$$W_{\text{cycle}} = Q_{\text{A}} - Q_{\text{R}}.$$

The *thermal efficiency* of a heat engine is defined as the ratio of the net work developed to the total energy added by heat transfer:

$$\eta = \frac{W_{\text{cycle}}}{Q_{\text{A}}} = 1 - \frac{Q_{\text{R}}}{Q_{\text{A}}}. \tag{1.9}$$

The thermal efficiency is strictly less than 100%. That is, some portion of the energy Q_{A} supplied is invariably rejected $Q_{\text{R}} \neq 0$.

The Second Law of Thermodynamics, Entropy

Many statements of the second law of thermodynamics have been proposed. Each of these can be called a statement of the second law *or* a corollary of the second law since, if one is invalid, all are invalid. In every instance where a consequence of the second law has been tested directly or indirectly by experiment it has been verified. Accordingly, the basis of the second law, like every other physical law, is experimental evidence.

Kelvin–Planck Statement

The Kelvin–Planck statement of the second law of thermodynamics refers to a *thermal reservoir.* A thermal reservoir is a system that remains at a constant temperature even though energy is added or removed by heat transfer. A reservoir is an idealization, of course, but such a system can be approximated in a number of ways—by the Earth's atmosphere, large bodies of water (lakes, oceans), and so on. Extensive properties of thermal reservoirs, such as internal energy, can change in interactions with other systems even though the reservoir temperature remains constant, however.

The Kelvin–Planck statement of the second law can be given as follows: *It is impossible for any system to operate in a thermodynamic cycle and deliver a net amount of energy by work to its surroundings while receiving energy by heat transfer from a single thermal reservoir.* In other words, *a perpetual motion machine of the second kind* is impossible. Expressed analytically, the Kelvin–Planck statement is

$$W_{\text{cycle}} \leq 0 \left(\text{single reservoir}\right),$$

where the words *single reservoir* emphasize that the system communicates thermally only with a single reservoir as it executes the cycle. The "less than" sign applies when *internal irreversibilities* are present as the system of interest undergoes a cycle and the "equal to" sign applies only when no irreversibilities are present.

Irreversibilities

A process is said to be *reversible* if it is possible for its effects to be eradicated in the sense that there is some way by which *both* the system and its surroundings can be *exactly restored* to their respective initial states. A process is *irreversible* if there is no way to undo it. That is, there is no means by which the system and its surroundings can be exactly restored to their respective initial states. A system that has undergone an irreversible process is not necessarily precluded from being restored to its initial state. However, were the system restored to its initial state, it would not also be possible to return the surroundings to their initial state.

There are many effects whose presence during a process renders it irreversible. These include, but are not limited to, the following: heat transfer through a finite temperature difference; unrestrained expansion of a gas or liquid to a lower pressure; spontaneous chemical reaction; mixing of matter at different compositions or states; friction (sliding friction as well as friction in the flow of fluids); electric current flow through a resistance; magnetization or polarization with hysteresis; and inelastic deformation. The term *irreversibility* is used to identify effects such as these.

Irreversibilities can be divided into two classes, *internal* and *external*. Internal irreversibilities are those that occur within the system, while external irreversibilities are those that occur within the surroundings, normally the immediate surroundings. As this division depends on the location of the boundary there is some arbitrariness in the classification (by locating the boundary to take in the immediate surroundings, all irreversibilities are internal). Nonetheless, valuable insights can result when this distinction between irreversibilities is made. When internal irreversibilities are absent during a process, the process is said to be *internally reversible*. At every intermediate state of an internally reversible process of a closed system, all intensive properties are uniform throughout each phase present: the temperature, pressure, specific volume, and other intensive properties do not vary with position. The discussions to follow compare the actual and internally reversible process concepts for two cases of special interest.

For a gas as the system, the work of expansion arises from the force exerted by the system to move the boundary against the resistance offered by the surroundings:

$$W = \int_1^2 F\,\mathrm{d}x = \int_1^2 pA\,\mathrm{d}x,$$

where the force is the product of the moving area and the pressure exerted by the system there. Noting that $A\,\mathrm{d}x$ is the change in total volume of the system,

$$W = \int_1^2 p\,\mathrm{d}V.$$

This expression for work applies to both actual and internally reversible expansion processes. However, for an internally reversible process p is not only the pressure at the moving boundary but also the pressure of the entire system. Furthermore, for an internally reversible process the volume equals mv, where the specific volume v has a single value throughout the system at a given instant. Accordingly, the work of an internally reversible expansion (or compression) process is

$$W = m\int_1^2 p\,\mathrm{d}v. \tag{1.1.10}$$

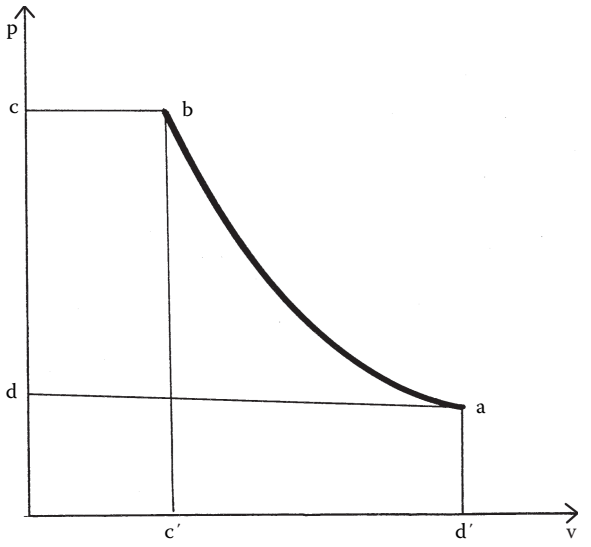


FIGURE 1.1.1 Internally reversible process on p–v coordinates.

When such a process of a closed system is represented by a continuous curve on a plot of pressure vs. specific volume, the area *under* the curve is the magnitude of the work per unit of system mass (area a–b–c'–d' of Figure 1.1.1, for example).

Although improved thermodynamic performance can accompany the reduction of irreversibilities, steps in this direction are normally constrained by a number of practical factors often related to costs. For example, consider two bodies able to communicate thermally. With a *finite* temperature difference between them, a spontaneous heat transfer would take place and, as noted previously, this would be a source of irreversibility. The importance of the heat-transfer irreversibility diminishes as the temperature difference narrows; and as the temperature difference between the bodies vanishes, the heat transfer approaches *ideality*. From the study of heat transfer it is known, however, that the transfer of a finite amount of energy by heat between bodies whose temperatures differ only slightly requires a considerable amount of time, a large heat-transfer surface area, or both. To approach *ideality*, therefore, a heat transfer would require an exceptionally long time and/or an exceptionally large area, each of which has cost implications constraining what can be achieved practically.

Carnot Corollaries

The two corollaries of the second law known as *Carnot* corollaries state: (1) the thermal efficiency of an irreversible power cycle is always less than the thermal efficiency of a reversible power cycle when each operates between the same two thermal reservoirs; (2) all reversible power cycles operating between the same two thermal reservoirs have the same thermal efficiency. A cycle is considered *reversible* when there are no irreversibilities within the system as it undergoes the cycle, and heat transfers between the system and reservoirs occur ideally (i.e., with a vanishingly small temperature difference).

Kelvin Temperature Scale

Carnot corollary 2 suggests that the thermal efficiency of a reversible power cycle operating between two thermal reservoirs depends only on the temperatures of the reservoirs and not on the nature of the substance making up the system executing the cycle or the series of processes. With Equation

1.1.9 it can be concluded that the ratio of the heat transfers is also related only to the temperatures, and is independent of the substance and processes:

$$\left(\frac{Q_C}{Q_H}\right)_{\text{rev cycle}} = \psi(T_C, T_H),$$

where Q_H *is* the energy transferred to the system by heat transfer from a *hot* reservoir at temperature T_H, and Q_C is the energy rejected from the system to a *cold* reservoir at temperature T_C. The words *rev cycle* emphasize that this expression applies only to systems undergoing reversible cycles while operating between the two reservoirs. Alternative temperature scales correspond to alternative specifications for the function ψ in this relation.

The Kelvin temperature scale is based on $\psi(T_C, T_H) = T_C / T_H$. Then

$$\left(\frac{Q_C}{Q_H}\right)_{\text{rev cycle}} = \frac{T_C}{T_H}. \tag{1.1.11}$$

This equation defines only a ratio of temperatures. The specification of the Kelvin scale is completed by assigning a numerical value to one standard reference state. The state selected is the same used to define the *gas scale*: at the triple point of water the temperature is specified to be 273.16 K. If a reversible cycle is operated between a reservoir at the reference state temperature and another reservoir at an unknown temperature T, then the latter temperature is related to the value at the reference state by

$$T = 273.16\left(\frac{Q}{Q'}\right)_{\text{rev cycle}},$$

where Q is the energy received by heat transfer from the reservoir at temperature T, and Q' is the energy rejected to the reservoir at the reference temperature. Accordingly, a temperature scale is defined that is valid over all ranges of temperature and that is independent of the thermometric substance.

Carnot Efficiency

For the special case of a reversible power cycle operating between thermal reservoirs at temperatures T_H and T_C on the Kelvin scale, combination of Equations 1.1.9 and 1.1.11 results in

$$\eta_{\max} = 1 - \frac{T_C}{T_H}, \tag{1.1.12}$$

called the *Carnot efficiency.* This is the efficiency of *all* reversible power cycles operating between thermal reservoirs at T_H and T_C. Moreover, it is the *maximum theoretical* efficiency that any power cycle, real or ideal, could have while operating between the same two reservoirs. As temperatures on the Rankine scale differ from Kelvin temperatures only by the factor 1.8, the above equation may be applied with either scale of temperature.

The Clausius Inequality

The Clausius inequality provides the basis for introducing two ideas instrumental for quantitative evaluations of processes of systems from a second law perspective: *entropy* and *entropy generation*. The Clausius inequality states that

$$\oint\left(\frac{\delta Q}{T}\right)_b \le 0, \tag{1.1.13a}$$

where δQ represents the heat transfer at a part of the system boundary during a portion of the cycle, and T is the absolute temperature at that part of the boundary. The symbol δ is used to distinguish

the differentials of *nonproperties*, such as heat and work, from the differentials of properties, written with the symbol *d*. The subscript b indicates that the integrand is evaluated at the boundary of the system executing the cycle. The symbol $\oint$ indicates that the integral is to be performed over all parts of the boundary and over the entire cycle. The Clausius inequality can be demonstrated using the Kelvin–Planck statement of the second law, and the significance of the inequality is the same: The equality applies when there are no internal irreversibilities as the system executes the cycle, and the inequality applies when internal irreversibilities are present.

The Clausius inequality can be expressed alternatively as

$$\oint \left(\frac{\delta Q}{T} \right)_b = -S_{\text{gen}}, \tag{1.1.13b}$$

where S_{gen} can be viewed as representing the *strength* of the inequality. The value of S_{gen} is positive when internal irreversibilities are present, zero when no internal irreversibilities are present, and can never be negative. Accordingly, S_{gen} is a measure of the irreversibilities present within the system executing the cycle. In the next section, S_{gen} is identified as the *entropy* generated (or *produced*) by internal irreversibilities during the cycle.

Entropy and Entropy Generation

Entropy

Consider two cycles executed by a closed system. One cycle consists of an internally reversible process A from state 1 to state 2, followed by an internally reversible process *C* from state 2 to state 1. The other cycle consists of an internally reversible process *B* from state 1 to state 2, followed by the same process *C* from state 2 to state 1 as in the first cycle. For these cycles, Equation 1.1.13b takes the form

$$\left(\int_1^2 \frac{\delta Q}{T} \right)_A + \left(\int_2^1 \frac{\delta Q}{T} \right)_C = -S_{\text{gen}} = 0$$

$$\left(\int_1^2 \frac{\delta Q}{T} \right)_B + \left(\int_2^1 \frac{\delta Q}{T} \right)_C = -S_{\text{gen}} = 0,$$

where S_{gen} has been set to zero since the cycles are composed of internally reversible processes. Subtracting these equations leaves

$$\left(\int_1^2 \frac{\delta Q}{T} \right)_A = \left(\int_1^2 \frac{\delta Q}{T} \right)_B.$$

Since A and B are arbitrary, it follows that the integral of $\delta Q/T$ has the same value for *any* internally reversible process between the two states: The value of the integral depends on the end states only. It can be concluded, therefore, that the integral defines the change in some property of the system. Selecting the symbol *S* to denote this property, its change is given by

$$S_2 - S_1 = \left(\int_1^2 \frac{\delta Q}{T} \right)_{\text{int rev}}, \tag{1.1.14a}$$

where the subscript *int rev* indicates that the integration is carried out for any internally reversible process linking the two states. This extensive property is called *entropy.*

Since entropy is a property, the change in entropy of a system in going from one state to another is the same for *all* processes, both internally reversible and irreversible, between these two states. In

other words, once the change in entropy between two states has been evaluated, this is the magnitude of the entropy change for *any* process of the system between these end states.

The definition of entropy change expressed on a differential basis is

$$dS = \left(\frac{\delta Q}{T} \right)_{\text{int rev}} . \tag{1.1.14b}$$

Equation 1.1.14b indicates that when a closed system undergoing an internally reversible process *receives* energy by heat transfer, the system experiences an *increase* in entropy. Conversely, when energy is *removed* from the system by heat transfer, the entropy of the system *decreases*. This can be interpreted to mean that an entropy transfer is *associated* with (or accompanies) heat transfer. The direction of the entropy transfer is the same as that of the heat transfer. In an *adiabatic* internally reversible process of a closed system, the entropy would remain constant. A constant entropy process is called an *isentropic* process.

On rearrangement, Equation 1.1.14b becomes

$$(\delta Q)_{\text{int rev}} = T \, \mathrm{d}S.$$

Then, for an internally reversible process of a closed system between states 1 and 2,

$$Q_{\text{int rev}} = m \int_1^2 T \, \mathrm{d}s. \tag{1.1.15}$$

When such a process is represented by a continuous curve on a plot of temperature vs. specific entropy, the area *under* the curve is the magnitude of the heat transfer per unit of system mass.

Entropy Balance

For a cycle consisting of an actual process from state 1 to state 2, during which internal irreversibilities are present, followed by an internally reversible process from state 2 to state 1, Equation 1.1.13b takes the form

$$\int_1^2 \left(\frac{\delta Q}{T} \right)_b + \int_2^1 \left(\frac{\delta Q}{T} \right)_{\text{int rev}} = -S_{\text{gen}},$$

where the first integral is for the actual process and the second integral is for the internally reversible process. Since no irreversibilities are associated with the internally reversible process, the term S_{gen} accounting for the effect of irreversibilities during the cycle can be identified with the actual process only.

Applying the definition of entropy change, the second integral of the foregoing equation can be expressed as

$$S_1 - S_2 = \int_2^1 \left(\frac{\delta Q}{T} \right)_{\text{int rev}}$$

Introducing this and rearranging the equation, the *closed system entropy balance* results

$$\underbrace{S_2 - S_1}_{\substack{\text{Entropy} \\ \text{change}}} = \underbrace{\int_1^2 \left(\frac{\delta Q}{T} \right)_b}_{\substack{\text{Entropy} \\ \text{transfer}}} + \underbrace{S_{\text{gen}}}_{\substack{\text{Entropy} \\ \text{generation.}}} \tag{1.1.16}$$

When the end states are fixed, the entropy change on the left side of Equation 1.1.16 can be evaluated independent of the details of the process from state 1 to state 2. However, the two terms on the right side depend explicitly on the nature of the process and cannot be determined solely from knowledge of the end states. The first term on the right side is associated with heat transfer to or from the system during the process. This term can be interpreted as the *entropy transfer associated with (or accompanying) heat transfer.* The direction of entropy transfer is the same as the direction of the heat transfer, and the same sign convention applies as for heat transfer: A positive value means that entropy is transferred into the system, and a negative value means that entropy is transferred out.

The entropy change of a system is not accounted for solely by entropy transfer, but is also due to the second term on the right side of Equation 1.1.16 denoted by S_{gen}. The term S_{gen} is positive when internal irreversibilities are present during the process and vanishes when internal irreversibilities are absent. This can be described by saying that entropy is *generated* (or produced) within the system by the action of irreversibilities. The second law of thermodynamics can be interpreted as specifying that entropy is generated by irreversibilities and conserved only in the limit as irreversibilities are reduced to zero. Since S_{gen} measures the effect of irreversibilities present within a system during a process, its value depends on the nature of the process and not solely on the end states. Entropy generation is *not* a property.

When applying the entropy balance, the objective is often to evaluate the entropy generation term. However, the value of the entropy generation for a given process of a system usually does not have much significance by itself. The significance is normally determined through comparison. For example, the entropy generation within a given component might be compared to the entropy generation values of the other components included in an overall system formed by these components. By comparing entropy generation values, the components where appreciable irreversibilities occur can be identified and rank ordered. This allows attention to be focused on the components that contribute most heavily to inefficient operation of the overall system.

To evaluate the entropy transfer term of the entropy balance requires information regarding both the heat transfer and the temperature on the boundary where the heat transfer occurs. The entropy transfer term is not always subject to direct evaluation, however, because the required information is either unknown or undefined, such as when the system passes through states sufficiently far from equilibrium. In practical applications, it is often convenient, therefore, to enlarge the system to include enough of the immediate surroundings that the temperature on the boundary of the *enlarged system* corresponds to the ambient temperature, T_{amb}. The entropy transfer term is then simply Q/T_{amb}. However, as the irreversibilities present would not be just those for the system of interest but those for the enlarged system, the entropy generation term would account for the effects of internal irreversibilities within the system *and* external irreversibilities present within that portion of the surroundings included within the enlarged system.

A form of the entropy balance convenient for particular analyses is the *rate form*:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \sum_j \frac{\dot{Q}_j}{T_j} + \dot{S}_{\text{gen}}, \tag{1.1.17}$$

where $\mathrm{d}S/\mathrm{d}t$ is the time rate of change of entropy of the system. The term $\dot{Q}_j/T_j$ represents the time rate of entropy transfer through the portion of the boundary whose instantaneous temperature is T_j. The term $\dot{S}_{\text{gen}}$ accounts for the time rate of entropy generation due to irreversibilities within the system.

For a system *isolated* from its surroundings, the entropy balance is

$$\left(S_2 - S_1\right)_{\text{isol}} = S_{\text{gen}}, \tag{1.1.18}$$

where S_{gen} is the total amount of entropy generated within the isolated system. Since entropy is generated in all actual processes, the only processes of an isolated system that actually can occur are those for which the entropy of the isolated system increases. This is known as the *increase of entropy principle*.

1.2 CONTROL VOLUME APPLICATIONS

Michael J. Moran

Since most applications of engineering thermodynamics are conducted on a control volume basis, the control volume formulations of the mass, energy, and entropy balances presented in this section are especially important. These are given here in the form of *overall* balances. Equations of change for mass, energy, and entropy in the form of differential equations are also available in the literature (see, e.g., Bird et al., 2002).

Conservation of Mass

When applied to a control volume, the principle of mass conservation states the time rate of accumulation of mass within the control volume equals the difference between the total rates of mass flow in and out across the boundary. An important case for engineering practice is one for which inward and outward flows occur, each through one or more ports. For this case, the conservation of mass principle takes the form

$$\frac{\mathrm{d}m_{\mathrm{cv}}}{\mathrm{d}t} = \sum_i \dot{m}_i - \sum_e \dot{m}_e . \tag{1.2.1}$$

The left side of this equation represents the time rate of change of mass contained within the control volume, $\dot{m}_i$ denotes the mass flow rate at an inlet, and $\dot{m}_e$ is the mass flow rate at an outlet.

The *volumetric flow rate* through a portion of the control surface with area dA is the product of the velocity component normal to the area, v_{n}, times the area: $v_{\mathrm{n}}\,\mathrm{d}A$. The *mass* flow rate through dA is $\rho(v_{\mathrm{n}}\,\mathrm{d}A)$. The mass rate of flow through a port of area A is then found by integration over the area

$$\dot{m} = \int_A \rho v_{\mathrm{n}}\,\mathrm{d}A.$$

For *one-dimensional* flow, the intensive properties are uniform with position over area A, and the last equation becomes

$$\dot{m} = \rho v A = \frac{vA}{v}, \tag{1.2.2}$$

where v denotes the specific volume and the subscript n has been dropped from velocity for simplicity.

Control Volume Energy Balance

When applied to a control volume, the principle of energy conservation states the time rate of accumulation of energy within the control volume equals the difference between the total incoming rate of energy transfer and the total outgoing rate of energy transfer. Energy can enter and exit a control volume by work and heat transfer. Energy also enters and exits with flowing streams of matter. Accordingly, for a control volume with one-dimensional flow at a single inlet and a single outlet,

$$\frac{\mathrm{d}(U + \mathrm{KE} + \mathrm{PE})_{\mathrm{cv}}}{\mathrm{d}t} = \dot{Q}_{\mathrm{cv}} - \dot{W} + \dot{m}\underline{\left(u_i + \frac{v_i^2}{2} + gz_i\right)} - \dot{m}\underline{\left(u_e + \frac{v_e^2}{2} + gz_e\right)}, \tag{1.2.3}$$

where the underlined terms account for the specific energy of the incoming and outgoing streams. The terms $\dot{Q}_{\mathrm{cv}}$ and $\dot{W}$ account, respectively, for the net rates of energy transfer by heat and work over the boundary (control surface) of the control volume.

Because work is always done on or by a control volume where matter flows across the boundary, the quantity $\dot{W}$ of Equation 1.2.3 can be expressed in terms of two contributions: one is the work associated with the force of the fluid pressure as mass is introduced at the inlet and removed at the exit. The other, denoted as $\dot{W}_{cv}$, includes *all other* work effects, such as those associated with rotating shafts, displacement of the boundary, and electrical effects. The work rate concept of mechanics allows the first of these contributions to be evaluated in terms of the product of the pressure force, pA, and velocity at the point of application of the force. To summarize, the work term $\dot{W}$ of Equation 1.2.3 can be expressed (with Equation 1.2.1) as

$$\begin{aligned}\dot{W} &= \dot{W}_{cv} + (p_e A_e)v_e - (p_i A_i)v_i \\ &= \dot{W}_{cv} + \dot{m}_e(p_e v_e) - \dot{m}_i(p_i v_i)\end{aligned} \tag{1.2.4}$$

The terms $\dot{m}_i\left(pv_i\right)$ and $\dot{m}_e\left(p_e v_e\right)$ account for the work associated with the pressure at the inlet and outlet, respectively, and are commonly referred to as *flow work*.

Substituting Equation 1.2.4 into Equation 1.2.3 and introducing the specific enthalpy h, the following form of the control volume energy rate balance results:

$$\frac{\mathrm{d}(U + \mathrm{KE} + \mathrm{PE})_{\mathrm{cv}}}{\mathrm{d}t} = \dot{Q}_{\mathrm{cv}} - \dot{W}_{\mathrm{cv}} + \dot{m}_i\left(h_i + \frac{v_i^2}{2} + gz_i\right) - \dot{m}_e\left(h_e + \frac{v_e^2}{2} + gz_e\right). \tag{1.2.5}$$

To allow for applications where there may be several locations on the boundary through which mass enters or exits, the following expression is appropriate:

$$\frac{\mathrm{d}(U + KE + PE)_{\mathrm{cv}}}{\mathrm{d}t} = \dot{Q}_{\mathrm{cv}} - \dot{W}_{\mathrm{cv}} + \sum_i \dot{m}_i\left(h_i + \frac{\mathrm{v}_i^2}{2} + \mathrm{gz}_i\right) - \sum_e \dot{m}_e\left(h_e + \frac{\mathrm{v}_e^2}{2} + \mathrm{gz}_e\right). \tag{1.2.6}$$

Equation 1.2.6 is an *accounting* rate balance for the energy of the control volume. It states that the time rate of accumulation of energy within the control volume equals the difference between the total rates of energy transfer in and out across the boundary. The mechanisms of energy transfer are heat and work, as for closed systems, and the energy accompanying the entering and exiting mass.

Control Volume Entropy Balance

Like mass and energy, entropy is an extensive property. And like mass and energy, entropy can be transferred into or out of a control volume by streams of matter. As this is the principal difference between the closed system and control volume forms, the control volume entropy rate balance is obtained by modifying Equation 1.1.17 to account for these entropy transfers. The result is

$$\underbrace{\frac{\mathrm{d}S_{\mathrm{cv}}}{\mathrm{d}t}}_{\substack{\text{Rate of}\\ \text{entropy}\\ \text{change}}} = \underbrace{\sum_j \frac{\dot{Q}_j}{T_j} + \sum_i \dot{m}_i s_i x}_{\substack{\text{Rate of}\\ \text{entropy}\\ \text{transfer}}} - \underbrace{\sum_e \dot{m}_e s_e + \dot{S}_{\mathrm{gen}}}_{\substack{\text{Rate of}\\ \text{entropy}\\ \text{generation}}} \tag{1.2.7}$$

where $\mathrm{d}S_{cv}/\mathrm{d}t$ represents the time rate of change of entropy within the control volume. The terms $\dot{m}_i s_i$ and $\dot{m}_e s_e$ account, respectively, for rates of entropy *transfer* into and out of the control volume associated with mass flow. One-dimensional flow is assumed at locations where mass enters and exits. $\dot{Q}_j$ represents the time rate of heat transfer at the location on the boundary where the

instantaneous temperature is T_j; and $\dot{Q}_j / T_j$ accounts for the associated rate of entropy *transfer*. $\dot{S}_{\text{gen}}$ denotes the time rate of entropy *generation* due to irreversibilities *within* the control volume. When a control volume comprises a number of components, $\dot{S}_{\text{gen}}$ is the sum of the rates of entropy generation of the components.

Control Volumes at Steady State

Engineering systems are often idealized as being at *steady state*, meaning that all properties are unchanging in time. For a control volume at steady state, the identity of the matter within the control volume changes continuously, but the total amount of mass remains constant. At steady state, Equation 1.2.1 reduces to

$$\sum_i \dot{m}_i = \sum_e \dot{m}_e . \tag{1.2.8a}$$

The energy rate balance of Equation 1.2.6 becomes, at steady state,

$$0 = \dot{Q}_{\text{cv}} - \dot{W}_{\text{cv}} + \sum_i \dot{m}_i \left(h_i + \frac{v_i^2}{2} + g z_i \right) - \sum_e \dot{m}_e \left(h_e + \frac{v_e^2}{2} + g z_e \right). \tag{1.2.8b}$$

At steady state, the entropy rate balance of Equation 1.2.7 reads

$$0 = \sum_j \frac{\dot{Q}_j}{T_j} + \sum_i \dot{m}_i s_i - \sum_e \dot{m}_e s_e + \dot{S}_{\text{gen}} . \tag{1.2.8c}$$

Mass and energy are conserved quantities, but entropy is not generally conserved. Equation 1.2.8a indicates that the total rate of mass flow into the control volume equals the total rate of mass flow out of the control volume. Similarly, Equation 1.2.8b states that the total rate of energy transfer into the control volume equals the total rate of energy transfer out of the control volume. However, Equation 1.2.8c shows that the rate at which entropy is transferred out *exceeds* the rate at which entropy enters, the difference being the rate of entropy generation within the control volume owing to irreversibilities.

Applications frequently involve control volumes having a single inlet and a single outlet, as, for example, the control volume of Figure 1.2.1 where heat transfer (if any) occurs at T_{b}: the temperature, or a suitable average temperature, on the boundary where heat transfer occurs. For this case

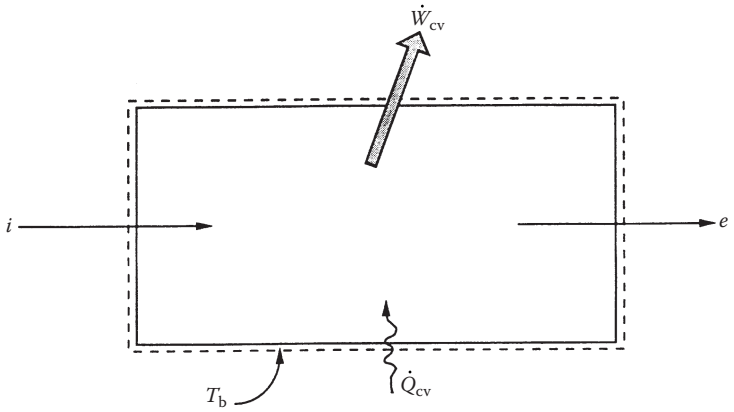


FIGURE 1.2.1 One-inlet, one-outlet control volume at steady state.

the mass rate balance, Equation 1.2.8a, reduces to $\dot{m}_i = \dot{m}_e$. Denoting the common mass flow rate by $\dot{m}$, Equations 1.2.8b and 1.2.8c read, respectively,

$$0 = \dot{Q}_{cv} - \dot{W}_{cv} + \dot{m}\left[(h_i - h_e) + \left(\frac{v_i^2 - v_e^2}{2}\right) + g(z_i - z_e)\right] \tag{1.2.9a}$$

$$0 = \frac{\dot{Q}_{cv}}{T_b} + \dot{m}(s_i - s_e) + \dot{S}_{gen}. \tag{1.2.10a}$$

When Equations 1.2.9a and 1.2.10a are applied to particular cases of interest, additional simplifications are usually made. The heat-transfer term Q_{cv} is dropped when it is insignificant relative to other energy transfers across the boundary. This may be the result of one or more of the following: (1) the outer surface of the control volume is insulated, (2) the outer surface area is too small for there to be effective heat transfer, (3) the temperature difference between the control volume and its surroundings is small enough that the heat transfer can be ignored, (4) the gas or liquid passes through the control volume so quickly that there is not enough time for significant heat transfer to occur. The work term $\dot{W}_{cv}$ drops out of the energy rate balance when there are no rotating shafts, displacements of the boundary, electrical effects, or other work mechanisms associated with the control volume being considered. The changes in kinetic and potential energy of Equation 1.2.9a are frequently negligible relative to other terms in the equation.

The special forms of Equations 1.2.9a and 1.2.10a listed in Table 1.2.1 are obtained as follows: When there is no heat transfer, Equation 1.2.10a gives

$$s_e - s_i = \frac{\dot{S}_{gen}}{\dot{m}} \geq 0 \quad \text{(no heat transfer)} \tag{1.2.10b}$$

Accordingly, when irreversibilities are present within the control volume, the specific entropy increases as mass flows from inlet to outlet. In the ideal case in which no internal irreversibilities are present, mass passes through the control volume with no change in its entropy—that is, *isentropically.*

For no heat transfer, Equation 1.2.9a gives

$$\dot{W}_{cv} = \dot{m}\left[(h_i - h_e) + \left(\frac{v_i^2 - v_e^2}{2}\right) + g(z_i - z_e)\right]. \tag{1.2.9b}$$

A special form that is applicable, at least approximately, to *compressors*, *pumps*, and *turbines* results from dropping the kinetic and potential energy terms of Equation 1.2.9b, leaving

$$\dot{W}_{cv} = \dot{m}(h_i - h_e) \quad \text{(compressors, pumps, and turbines)} \tag{1.2.9c}$$

In *throttling devices*, a significant reduction in pressure is achieved simply by introducing a restriction into a line through which a gas or liquid flows. For such devices $\dot{W}_{cv} = 0$ and Equation 1.2.9c reduces further to read

$$h_e \equiv h_i \quad \text{(throttling process)} \tag{1.2.9d}$$

TABLE 1.2.1
Energy and Entropy Balances for One-Inlet, One-Outlet Control Volumes at Steady State and No Heat Transfer

Energy balance

$$\dot{W}_{cv} = \dot{m}\left[(h_i - h_e) + \left(\frac{v_i^2 - v_e^2}{2}\right) + g(z_i - z_e)\right] \quad (1.2.9b)$$

Compressors, pumps, and turbines[a]

$$\dot{W}_{cv} = \dot{m}(h_i - h_e) \quad (1.2.9c)$$

Throttling

$$h_e \cong h_i \quad (1.2.9d)$$

Nozzles, diffusers[b]

$$v_e = \sqrt{v_i^2 + 2(h_i - h_e)} \quad (1.2.9f)$$

Entropy balance

$$s_e - s_i = \frac{\dot{S}_{gen}}{\dot{m}} \geq 0 \quad (1.2.10b)$$

[a] For an ideal gas with constant c_p, Equation 1′ of Table 1.3.6 allows Equation 1.2.9c to be written as

$$\dot{W}_{cv} = \dot{m}c_p(T_i - T_e) \quad (1.2.9c')$$

The power developed in an *isentropic process* is obtained with Equation 5′ of Table 1.3.6 as

$$\dot{W}_{cv} = \dot{m}c_pT_i\left[1 - (p_e/p_i)^{(k-1)/k}\right](s = c), \quad (1.2.9c'')$$

where $c_p = kR/(k-1)$.

[b] For an ideal gas with constant c_p, Equation 1′ of Table 1.3.6 allows Equation 1.2.9f to be written as

$$v_e = \sqrt{v_i^2 + 2c_p(T_i - T_e)} \quad (1.2.9f')$$

The exit velocity for an *isentropic process* is obtained with Equation 5′ of Table 1.3.6 as

$$v_e = \sqrt{v_i^2 + 2c_pT_i\left[1 - (p_e/p_i)^{(k-1)/k}\right]}\ (s = c) \quad (1.2.9f'')$$

where $c_p = kR/(k-1)$.

That is, in upstream and downstream of the throttling device, the specific enthalpies are equal.

A *nozzle* is a flow passage of varying cross-sectional area in which the velocity of a gas or liquid increases in the direction of flow. In a *diffuser*, the gas or liquid decelerates in the direction of flow. For such devices, $\dot{W}_{cv} = 0$. The heat transfer and potential energy change are also generally negligible. Then Equation 1.2.9b reduces to

$$0 = h_i - h_e + \frac{v_i^2 - v_e^2}{2}. \quad (1.2.9e)$$

Solving for the outlet velocity,

$$v_e = \sqrt{v_i^2 + 2(h_i - h_e)} \quad \text{(nozzle, diffuser)}. \quad (1.2.9f)$$

Further discussion of the flow-through nozzles and diffusers is provided in Chapter 2.

The mass, energy, and entropy rate balances, Equations 1.2.8, can be applied to control volumes with multiple inlets and/or outlets, as, for example, cases involving heat-recovery steam generators, feedwater heaters, and counterflow and crossflow heat exchangers. Transient (or unsteady) analyses can be conducted with Equations 1.2.1, 1.2.6, and 1.2.7. Illustrations of all such applications are provided by Moran et al. (2014).

Example 1.2.1

A turbine receives steam at 7 MPa, 440°C and exhausts at 0.2 MPa for subsequent process heating duty. If heat transfer and kinetic/potential energy effects are negligible, determine the steam mass flow rate, in kg/h, for a turbine power output of 30 MW when (1) the steam quality at the turbine outlet is 95%, (2) the turbine expansion is internally reversible.

***Solution*:** With the indicated idealizations, Equation 1.2.9c is appropriate. Solving, $\dot{m} = \dot{W}_{cv}/(h_i - h_e)$. Steam table data (Table A.5) at the inlet condition are $h_i = 3261.7$ kJ/kg, $s_i = 6.6022$ kJ/kg·K.

1. At 0.2 MPa and $x = 0.95$, $h_e = 2596.5$ kJ/kg. Then

$$\dot{m} = \frac{30\,\text{MW}}{(3261.7 - 2596.5)\text{kJ/kg}}\left(\frac{10^3\,\text{kJ/s}}{1\text{MW}}\right)\left(\frac{3600\,\text{s}}{1\text{h}}\right).$$
$$= 162{,}357\,\text{kg/h}$$

2. For an internally reversible expansion, Equation 1.2.10b reduces to give $s_e = s_i$. For this case, $h_e = 2499.6$ kJ/kg ($x = 0.906$), and $\dot{m} = 141{,}714$ kg/h.

Example 1.2.2

Air at 500°F, 150 lbf/in.^{2}, and 10 ft/s expands adiabatically through a nozzle and exits at 60°F, 15 lbf/in.^{2}. For a mass flow rate of 5 lb/s determine the exit area, in in.^{2}. Repeat for an isentropic expansion to 15 lbf/in.^{2}. Model the air as an ideal gas (Section 1.3, Ideal Gas Model) with specific heat $c_p = 0.24$ Btu/lb · °R ($k = 1.4$).

Solution: The nozzle exit area can be evaluated using Equation 1.2.2 together with the ideal gas equation, $v = RT/p$:

$$A_e = \frac{\dot{m}v_e}{v_e} = \frac{\dot{m}(RT_e/p_e)}{v_e}.$$

The exit velocity required by this expression is obtained using Equation 1.2.9f′ of Table 1.2.1:

$$v_e = \sqrt{v_i^2 + 2c_p(T_i - T_e)}$$
$$= \sqrt{\left(\frac{10\text{ ft}}{s}\right)^2 + 2\left(0.24\frac{\text{Btu}}{\text{lb}\cdot\text{R}}\right)\left(\frac{778.17\text{ ft}\cdot\text{lbf}}{1\,\text{Btu}}\right)(440°\text{R})\left(\frac{32.174\,\text{lb}\cdot\text{ft/s}^2}{1\text{lbf}}\right)}.$$
$$= 2299.5\,\text{ft/s}$$

Finally, with $R = \bar{R}/M = 53.33$ ft · lbf/lb · °R,

$$A_e = \frac{\left(5\text{lb/s}\right)\left(53.3\frac{\text{ft}\cdot\text{lbf}}{\text{lb}\cdot°\text{R}}\right)(520°\text{R})}{\left(2299.5\,\text{ft/s}\right)\left(15\text{ lbf/in.}^2\right)} = 4.02\text{ in.}^2$$

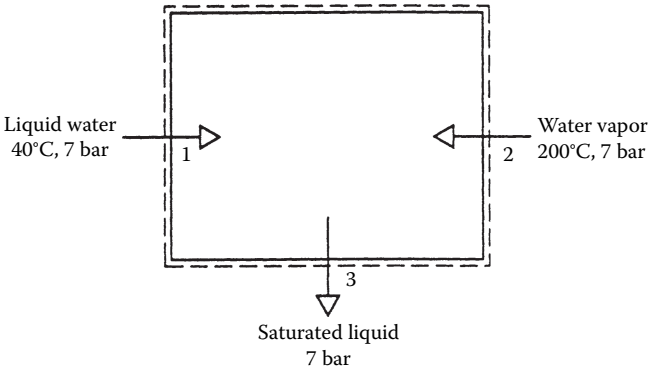


FIGURE 1.2.2 Open feed water heater.

Using Equation 1.2.9f″ in Table 1.2.1 for the isentropic expansion,

$$v_e = \sqrt{(10)^2 + 2(0.24)(778.17)(960)(32.174)\left[1-\left(\frac{15}{150}\right)^{0.4/1.4}\right]}.$$

$$= 2358.3\,\text{ft/s}$$

Then $A_e = 3.92$ in.^{2}.

Example 1.2.3

Figure 1.2.2 provides steady-state operating data for an open feedwater heater. Ignoring heat transfer and kinetic/potential energy effects, determine the ratio of mass flow rates, $\dot{m}_1/\dot{m}_2$.

Solution: For this case, Equations 1.2.8a and b reduce to read, respectively,

$$\dot{m}_1 + \dot{m}_2 = \dot{m}_3$$

$$0 = \dot{m}_1 h_1 + \dot{m}_2 h_2 - \dot{m}_3 h_3.$$

Combining and solving for the ratio $\dot{m}_1/\dot{m}_2$,

$$\frac{\dot{m}_1}{\dot{m}_2} = \frac{h_2 - h_3}{h_3 - h_1}.$$

Inserting steam table data, in kJ/kg, from Table A.5,

$$\frac{\dot{m}_1}{\dot{m}_2} = \frac{2844.8 - 697.2}{697.2 - 167.6} = 4.06.$$

Internally Reversible Heat Transfer and Work

For one-inlet, one-outlet control volumes at steady state, the following expressions give the heat-transfer rate and power in the absence of internal irreversibilities:

$$\left(\frac{\dot{Q}_{cv}}{\dot{m}}\right)_{\text{int rev}} = \int_1^2 T\,ds \tag{1.2.11}$$

$$\left(\frac{\dot{W}_{cv}}{\dot{m}}\right)_{\text{int rev}} = -\int_1^2 v\,\mathrm{d}p + \frac{v_1^2 - v_2^2}{2} + g(z_1 - z_2). \tag{1.2.12a}$$

(see, e.g., Moran et al. 2014).

If there is no significant change in kinetic or potential energy from inlet to outlet, Equation 1.2.12a reads

$$\left(\frac{\dot{W}_{cv}}{\dot{m}}\right)_{\text{int rev}} = -\int_1^2 v\,\mathrm{d}p \quad (\Delta ke = \Delta pe = 0). \tag{1.2.12b}$$

The specific volume remains approximately constant in many applications with liquids. Then Equation 1.2.12b becomes

$$\left(\frac{\dot{W}_{cv}}{\dot{m}}\right)_{\text{int rev}} = -v(p_2 - p_1) \quad (v = \text{constant}). \tag{1.2.12c}$$

When the states visited by a unit of mass flowing without irreversibilities from inlet to outlet are described by a continuous curve on a plot of temperature vs. specific entropy, Equation 1.2.11 implies that the area under the curve is the magnitude of the heat transfer per unit of mass flowing. When such an ideal process is described by a curve on a plot of pressure vs. specific volume, as shown in Figure 1.1.1, the magnitude of the integral $\int v\,\mathrm{d}p$ of Equation 1.2.12a and b is represented by the area *a–b–c–d behind* the curve. The area *a–b–c′–d′ under* the curve is identified with the magnitude of the integral $\int p\,\mathrm{d}v$ of Equation 1.1.10.

1.3 PROPERTY RELATIONS AND DATA

Michael J. Moran

Pressure, temperature, volume, and mass can be found experimentally. The relationships between the specific heats c_v and c_p and temperature at relatively low pressure are also accessible experimentally, as are certain other property data. Specific internal energy, enthalpy, and entropy are among those properties that are not so readily obtained in the laboratory. Values for such properties are calculated using experimental data of properties that are more amenable to measurement, together with appropriate property relations derived using the principles of thermodynamics. In this section, property relations and data sources are considered for *simple compressible systems*, which include a wide range of industrially important substances.

Property data are provided in the publications of the *National Institute of Standards and Technology* (formerly the U.S. Bureau of Standards), of professional groups such as the *American Society of Mechanical Engineering (ASME)*, the *American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE)*, and the *American Chemical Society*, and of corporate entities such as *Dupont* and *Dow Chemical*. Handbooks and property reference volumes such as those included in the list of references for this chapter are readily accessed sources of data. Property data are also retrievable from various commercial online data bases.

Basic Relations for Pure Substances

An energy balance in differential form for a closed system undergoing an internally reversible process in the absence of overall system motion and the effect of gravity reads

$$\mathrm{d}U = (\delta Q)_{\text{int rev}} - (\delta W)_{\text{int rev}}$$

From Equation 1.1.14b, $(\delta Q)_{\text{int rev}} = T\,\mathrm{d}s$. When consideration is limited to *simple compressible systems*: systems for which the only significant work in an internally reversible process is associated with volume change, $(\delta W)_{\text{int rev}} = p\,\mathrm{d}V$, the following equation is obtained:

$$\mathrm{d}U = T\,\mathrm{d}S - p\,\mathrm{d}V \tag{1.3.1a}$$

Introducing enthalpy, $H = U + pV$, the Helmholtz function, $\psi = U - TS$, and the Gibbs function, $G = H - TS$, three additional expressions are obtained:

$$\mathrm{d}H = T\,\mathrm{d}S + V\,\mathrm{d}p \tag{1.3.1b}$$

$$\mathrm{d}\psi = -p\,\mathrm{d}V - S\,\mathrm{d}T \tag{1.3.1c}$$

$$\mathrm{d}G = V\,\mathrm{d}p - S\,\mathrm{d}T \tag{1.3.1d}$$

Equations 1.3.1 can be expressed on a per-unit-mass basis as

$$\mathrm{d}u = T\,\mathrm{d}s - p\,\mathrm{d}v \tag{1.3.2a}$$

$$\mathrm{d}h = T\,\mathrm{d}s + v\,\mathrm{d}p \tag{1.3.2b}$$

$$\mathrm{d}\psi = -p\,\mathrm{d}v - s\,\mathrm{d}T \tag{1.3.2c}$$

$$\mathrm{d}g = v\,\mathrm{d}p - s\,\mathrm{d}T \tag{1.3.2d}$$

Similar expressions can be written on a per-mole basis.

TABLE 1.3.1
Relations from Exact Differentials

Function	Differential	Coefficients	Maxwell
General:			
$z = z(x, y)$	$dz = M(x, y)dx + N(x, y)dy$	$\left(\frac{\partial z}{\partial x}\right)_y = M$	$\left(\frac{\partial M}{\partial y}\right)_x = \left(\frac{\partial N}{\partial x}\right)_y$
		$\left(\frac{\partial z}{\partial y}\right)_x = N$	
Internal energy:			
$u(s, v)$	$du = Tds - pdv$	$\left(\frac{\partial u}{\partial s}\right)_v = T$	$\left(\frac{\partial T}{\partial v}\right)_s = -\left(\frac{\partial p}{\partial s}\right)_v$
		$\left(\frac{\partial u}{\partial v}\right)_s = -p$	
Enthalpy:			
$h(s, p)$	$dh = Tds + vdp$	$\left(\frac{\partial h}{\partial s}\right)_p = T$	$\left(\frac{\partial T}{\partial p}\right)_s = \left(\frac{\partial v}{\partial s}\right)_p$
		$\left(\frac{\partial h}{\partial p}\right)_s = v$	
Helmholtz function:			
$\psi(v, T)$	$d\psi = -pdv - sdT$	$\left(\frac{\partial \psi}{\partial v}\right)_T = -p$	$\left(\frac{\partial p}{\partial T}\right)_v = \left(\frac{\partial s}{\partial v}\right)_T$
		$\left(\frac{\partial \psi}{\partial T}\right)_v = -s$	
Gibbs function:			
$g(T, p)$	$dg = vdp - sdT$	$\left(\frac{\partial g}{\partial p}\right)_T = v$	$\left(\frac{\partial v}{\partial T}\right)_p = -\left(\frac{\partial s}{\partial p}\right)_T$
		$\left(\frac{\partial g}{\partial T}\right)_p = -s$	

Maxwell Relations

Since only properties are involved, each of the four differential expressions given by Equations 1.3.2 is an *exact* differential exhibiting the general form $dz = M(x, y)dx + N(x, y)dy$, where the second mixed partial derivatives are equal: $(\partial M/\partial y) = (\partial N/\partial x)$. Underlying these exact differentials are, respectively, functions of the form $u(s, v)$, $h(s, p)$, $\psi(v, T)$, and $g(T, p)$. From such considerations the *Maxwell relations* given in Table 1.3.1 can be established.

Example 1.3.1

Derive the Maxwell relation following from Equation 1.3.2a.

Solution: The differential of the function $u = u(s, v)$ is

$$du = \left(\frac{\partial u}{\partial s}\right)_v ds + \left(\frac{\partial u}{\partial v}\right)_s dv$$

By comparison with Equation 1.3.2a,

$$T = \left(\frac{\partial u}{\partial s}\right)_v, \quad -p = \left(\frac{\partial u}{\partial v}\right)_s$$

In Equation 1.3.2a, T plays the role of M, and $-p$ plays the role of N, so the equality of second mixed partial derivatives gives the Maxwell relation,

$$\left(\frac{\partial T}{\partial v}\right)_s = -\left(\frac{\partial p}{\partial s}\right)_v$$

Since each of the properties T, p, v, and s appears on the right side of two of the eight coefficients of Table 1.3.1, four additional property relations can be obtained by equating such expressions:

$$\left(\frac{\partial u}{\partial s}\right)_v = \left(\frac{\partial h}{\partial s}\right)_p, \quad \left(\frac{\partial u}{\partial v}\right)_s = \left(\frac{\partial \psi}{\partial v}\right)_T$$
$$\left(\frac{\partial h}{\partial p}\right)_s = \left(\frac{\partial g}{\partial p}\right)_T, \quad \left(\frac{\partial \psi}{\partial T}\right)_v = \left(\frac{\partial g}{\partial T}\right)_p$$

These four relations are identified in Table 1.3.1 by brackets. As any three of Equations 1.3.2 can be obtained from the fourth simply by manipulation, the 16 property relations in Table 1.3.1 can also be regarded as following from this single differential expression. Several additional first-derivative property relations can be derived (see, e.g., Zemansky, 1972).

Specific Heats and Other Properties

Engineering thermodynamics uses a wide assortment of thermodynamic properties and relations among these properties. Table 1.3.2 lists several commonly encountered properties.

Among the entries in Table 1.3.2 are the specific heats c_v and c_p. These intensive properties are often required for thermodynamic analysis and are defined as partial derivations of the functions $u(T, v)$ and $h(T, p)$, respectively,

TABLE 1.3.2
Symbols and Definitions for Selected Properties

Property	Symbol	Definition	Property	Symbol	Definition
Pressure	p		Specific heat, constant volume	c_v	$\left(\partial u/\partial T\right)_v$
Temperature	T		Specific heat, constant pressure	c_p	$\left(\partial h/\partial T\right)_p$
Specific volume	v		Volume expansivity	β	$\frac{1}{v}\left(\partial v/\partial T\right)_p$
Specific internal energy	u		Isothermal compressivity	κ	$-\frac{1}{v}\left(\partial v/\partial p\right)_T$
Specific entropy	s		Isentropic compressibility	α	$-\frac{1}{v}\left(\partial v/\partial p\right)_s$
Specific enthalpy	h	$u + pv$	Isothermal bulk modulus	B	$-v\left(\partial p/\partial v\right)_T$
Specific Helmholtz function	ψ	$u - Ts$	Isentropic bulk modulus	B_s	$-v\left(\partial p/\partial v\right)_s$
Specific Gibbs function	g	$h - Ts$	Joule–Thomson coefficient	μ_J	$\left(\partial T/\partial p\right)_h$
Compressibility factor	Z	pv/RT	Joule coefficient	η	$\left(\partial T/\partial v\right)_u$
Specific heat ratio	k	c_p/c_v	Velocity of sound	c	$\sqrt{-v^2\left(\partial p/\partial v\right)_s}$

$$c_v = \left(\frac{\partial u}{\partial T}\right)_v \tag{1.3.3}$$

$$c_p = \left(\frac{\partial h}{\partial T}\right)_p \tag{1.3.4}$$

Since u and h can be expressed either on a unit mass basis or a per-mole basis, values of the specific heats can be similarly expressed. Table 1.3.3 summarizes relations involving c_v and c_p. The property k, the specific heat ratio, is

$$k = \frac{c_p}{c_v} \tag{1.3.5}$$

TABLE 1.3.3
Specific Heat Relations

$$c_v = \left(\frac{\partial u}{\partial T}\right)_v = T\left(\frac{\partial s}{\partial T}\right)_v \tag{1}$$

$$= -T\left(\frac{\partial p}{\partial T}\right)_v \left(\frac{\partial v}{\partial T}\right)_s \tag{2}$$

$$c_p = \left(\frac{\partial h}{\partial T}\right)_p = T\left(\frac{\partial s}{\partial T}\right)_p \tag{3}$$

$$= T\left(\frac{\partial v}{\partial T}\right)_p \left(\frac{\partial p}{\partial T}\right)_s \tag{4}$$

$$c_p - c_v = T\left(\frac{\partial p}{\partial T}\right)_v \left(\frac{\partial v}{\partial T}\right)_p \tag{5}$$

$$= -T\left(\frac{\partial v}{\partial T}\right)_p^2 \left(\frac{\partial p}{\partial v}\right)_T \tag{6}$$

$$= \frac{Tv\beta^2}{\kappa} \tag{7}$$

$$c_p = \frac{1}{\mu_J}\left[T\left(\frac{\partial v}{\partial T}\right)_p - v\right] \tag{8}$$

$$c_v = \frac{1}{\eta}\left[T\left(\frac{\partial p}{\partial T}\right)_v - p\right] \tag{9}$$

$$k = \frac{c_p}{c_v} = \left(\frac{\partial v}{\partial p}\right)_T \left(\frac{\partial p}{\partial v}\right)_s \tag{10}$$

$$\left(\frac{\partial c_v}{\partial v}\right)_T = T\left(\frac{\partial^2 p}{\partial T^2}\right)_v \tag{11}$$

$$\left(\frac{\partial c_p}{\partial p}\right)_T = -T\left(\frac{\partial^2 v}{\partial T^2}\right)_p \tag{12}$$

Note: See, for example, Moran, M.J. and Shapiro, H.N. *Fundamentals of Engineering Thermodynamics*, 4th ed., Wiley, New York, 2000.

Values for c_v and c_p can be obtained via statistical mechanics using *spectroscopic* measurements. They can also be determined macroscopically through exacting property measurements. Specific heat data for common gases, liquids, and solids are provided by the handbooks and property reference volumes listed among the Chapter 1 references. Specific heats are also considered in Section 1.3 as a part of the discussions of the *incompressible model* and the *ideal gas model*. Figure 1.3.1 shows how c_p for water vapor varies as a function of temperature and pressure. Other gases exhibit similar behavior. The figure also gives the variation of c_p with temperature in the limit as pressure tends to zero (the ideal gas limit). In this limit c_p increases with increasing temperature, which is a characteristic exhibited by other gases as well.

The following two equations are often convenient for establishing relations among properties:

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z = 1 \tag{1.3.6a}$$

$$\left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z = -1 \tag{1.3.6b}$$

Their use is illustrated in Example 1.3.2.

Example 1.3.2

Obtain Equations 2 and 11 in Table 1.3.3 from Equation 1.

***Solution*:** Identifying *x*, *y*, *z* with *s*, *T*, and *v*, respectively, Equation 1.3.6b reads

$$\left(\frac{\partial T}{\partial v}\right)_s \left(\frac{\partial v}{\partial s}\right)_T \left(\frac{\partial s}{\partial T}\right)_v = -1$$

Applying Equation 1.3.6a to each of $(\partial T/\partial v)_s$ and $(\partial v/\partial s)_T$,

$$\left(\frac{\partial s}{\partial T}\right)_v = -\frac{1}{\left(\partial T/\partial v\right)_s \left(\partial v/\partial s\right)_T} = -\left(\frac{\partial v}{\partial T}\right)_s \left(\frac{\partial s}{\partial v}\right)_T$$

Introducing the Maxwell relation from Table 1.3.1 corresponding to $\psi(T, v)$,

$$\left(\frac{\partial s}{\partial T}\right)_v = -\left(\frac{\partial v}{\partial T}\right)_s \left(\frac{\partial p}{\partial T}\right)_v$$

With this, Equation 2 of Table 1.3.3 is obtained from Equation 1, which in turn is obtained in Example 1.3.3. Equation 11 of Table 1.3.3 can be obtained by differentiating Equation 1 with respect to specific volume at fixed temperature, and again using the Maxwell relation corresponding to ψ.

P–v–T Relations

Considerable pressure, specific volume, and temperature data have been accumulated for industrially important gases and liquids. These data can be represented in the form $p=f(v, T)$, called an *equation of state*. Equations of state can be expressed in tabular, graphical, and analytical forms.

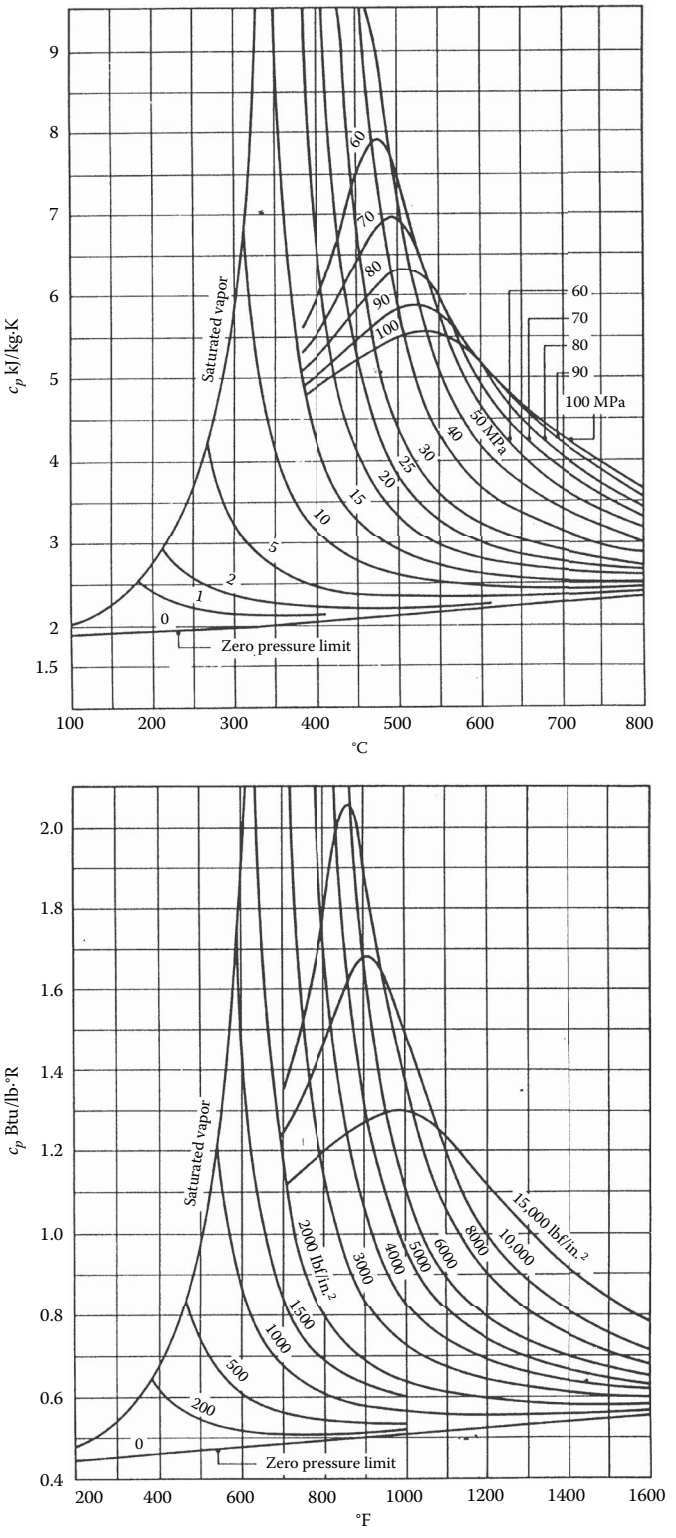


FIGURE 1.3.1 c_p of water vapor as a function of temperature and pressure. (Adapted from Keenan, J.H. et al., *Steam Tables—S.I. Units (English Units)*, John Wiley & Sons, New York, 1978.)

P–v–T Surface

The graph of a function $p=f(v, T)$ is a surface in three-dimensional space. Figure 1.3.2 shows the *p–v–T* relationship for water. Figure 1.3.2b shows the projection of the surface onto the pressure–temperature plane, called the *phase diagram*. The projection onto the *p–v* plane is shown in Figure 1.3.2c.

Figure 1.3.2 has three regions labeled solid, liquid, and vapor where the substance exists only in a single phase. Between the single-phase regions lie *two-phase* regions, where two phases coexist in equilibrium. The lines separating the single-phase regions from the two-phase regions are *saturation lines*. Any state represented by a point on a saturation line is a *saturation state*. The line separating the liquid phase and the two-phase liquid–vapor region is the saturated liquid line. The state denoted by *f* (Figure 1.3.2c) is a saturated liquid state. The saturated vapor line separates the vapor region and the two-phase liquid–vapor region. The state denoted by *g* is a saturated vapor state. The saturated liquid line and the saturated vapor line meet at the *critical point*. At the critical point, the pressure is the *critical pressure* p_c, and the temperature is the *critical temperature* T_c. Three phases can coexist in equilibrium along the line labeled *triple line*. The triple line projects onto a point on the phase diagram. The triple point of water is used in defining the Kelvin temperature scale (Section 1.1, Basic Concepts and Definitions; The Second Law of Thermodynamics, Entropy).

When a phase change occurs during constant pressure heating or cooling, the temperature remains constant as long as both phases are present. Accordingly, in the two-phase liquid–vapor region, a line of constant pressure is also a line of constant temperature. For a specified pressure, the corresponding temperature is called the *saturation temperature*. For a specified temperature,

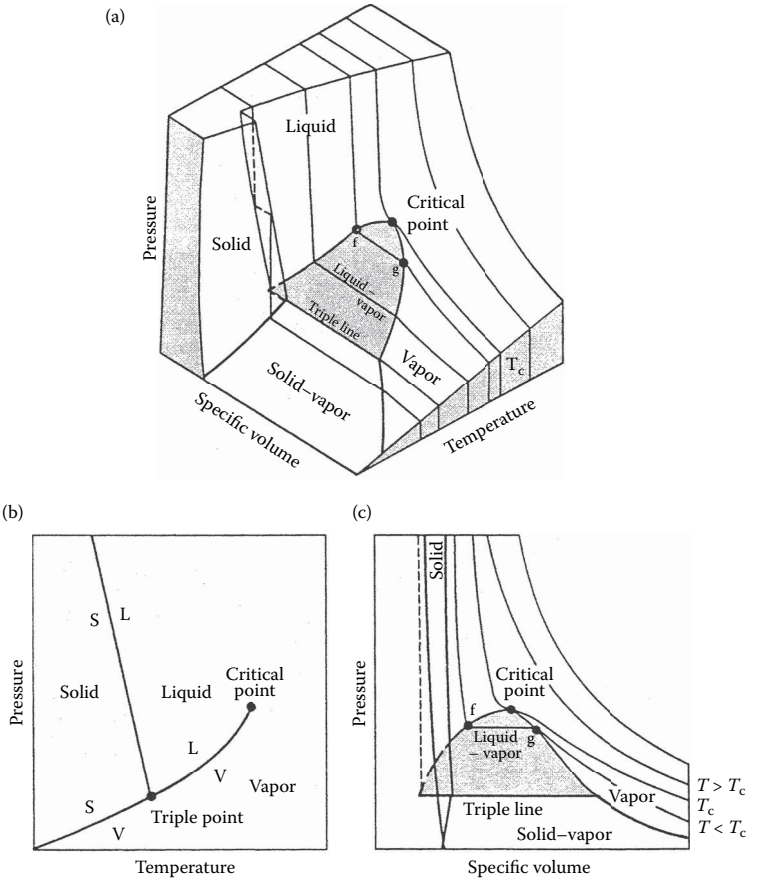


FIGURE 1.3.2 Pressure-specific volume–temperature surface and projections for water (not to scale).

the corresponding pressure is called the *saturation pressure*. The region to the right of the saturated vapor line is known as the *superheated vapor region* because the vapor exists at a temperature greater than the saturation temperature for its pressure. The region to the left of the saturated liquid line is known as the *compressed liquid region* because the liquid is at a pressure higher than the saturation pressure for its temperature.

When a mixture of liquid and vapor coexists in equilibrium, the liquid phase is a saturated liquid and the vapor phase is a saturated vapor. The total volume of any such mixture is $V = V_f + V_g$; or, alternatively, $mv = m_f v_f + m_g v_g$, where m and v denote mass and specific volume, respectively. Dividing by the total mass of the mixture m and letting the *mass fraction* of the vapor in the mixture, m_g/m, be symbolized by x, called the *quality*, the apparent specific volume v of the mixture is

$$\begin{aligned} v &= (1-x)v_f + xv_g \\ &= v_f + xv_{fg} \end{aligned} \tag{1.3.7a}$$

where $v_{fg} = v_g - v_f$. Expressions similar in form can be written for internal energy, enthalpy, and entropy:

$$\begin{aligned} u &= (1-x)u_f + xu_g \\ &= u_f + xu_{fg} \end{aligned} \tag{1.3.7b}$$

$$\begin{aligned} h &= (1-x)h_f + xh_g \\ &= h_f + xh_{fg} \end{aligned} \tag{1.3.7c}$$

$$\begin{aligned} s &= (1-x)s_f + xs_g \\ &= s_f + xs_{fg} \end{aligned} \tag{1.3.7d}$$

For the case of water, Figure 1.3.3 illustrates the phase change from solid to liquid (melting): *a*–*b*–*c*; from solid to vapor (sublimation): *a*′–*b*′–*c*′; and from liquid to vapor (vaporization): *a*″–*b*″–*c*″. During

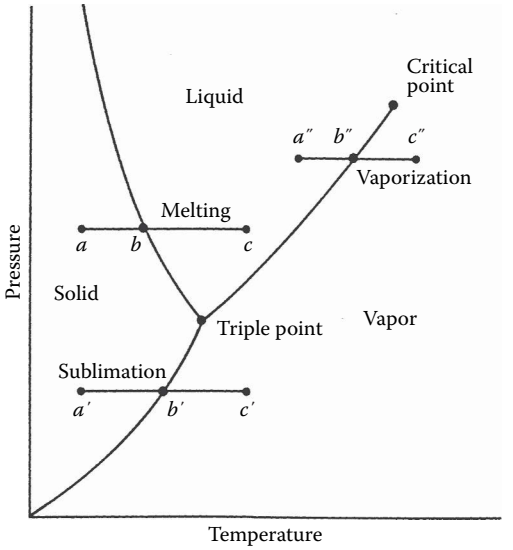


FIGURE 1.3.3 Phase diagram for water (not to scale).

any such phase change the temperature and pressure remain constant and thus are not independent properties. The *Clapeyron equation* allows the change in enthalpy during a phase change at fixed temperature to be evaluated from *p–v–T* data pertaining to the phase change. For vaporization, the Clapeyron equation reads

$$\left(\frac{\mathrm{d}p}{\mathrm{d}T}\right)_{\mathrm{sat}} = \frac{h_{\mathrm{g}} - h_{\mathrm{f}}}{T\left(v_{\mathrm{g}} - v_{\mathrm{f}}\right)} \tag{1.3.8}$$

where $(\mathrm{d}p/\mathrm{d}T)_{\mathrm{sat}}$ is the slope of the saturation pressure–temperature curve at the point determined by the temperature held constant during the phase change. Expressions similar in form to Equation 1.3.8 can be written for sublimation and melting.

The Clapeyron equation shows that the slope of a saturation line on a phase diagram depends on the signs of the specific volume and enthalpy changes accompanying the phase change. In most cases, when a phase change takes place with an increase in specific enthalpy, the specific volume also increases, and $(\mathrm{d}p/\mathrm{d}T)_{\mathrm{sat}}$ is positive. However, in the case of the melting of ice and a few other substances, the specific volume decreases on melting. The slope of the saturated solid–liquid curve for these few substances is negative, as illustrated for water in Figure 1.3.3.

Graphical Representations

The intensive states of a pure, simple compressible system can be represented graphically with any two independent intensive properties as the coordinates, excluding properties associated with motion and gravity. While any such pair may be used, there are several selections that are conventionally employed. These include the *p–T* and *p–v* diagrams in Figure 1.3.2, the *T–s* diagram in Figure 1.3.4, the *h–s* (Mollier) diagram in Figure 1.3.5, and the *p–h* diagram shown

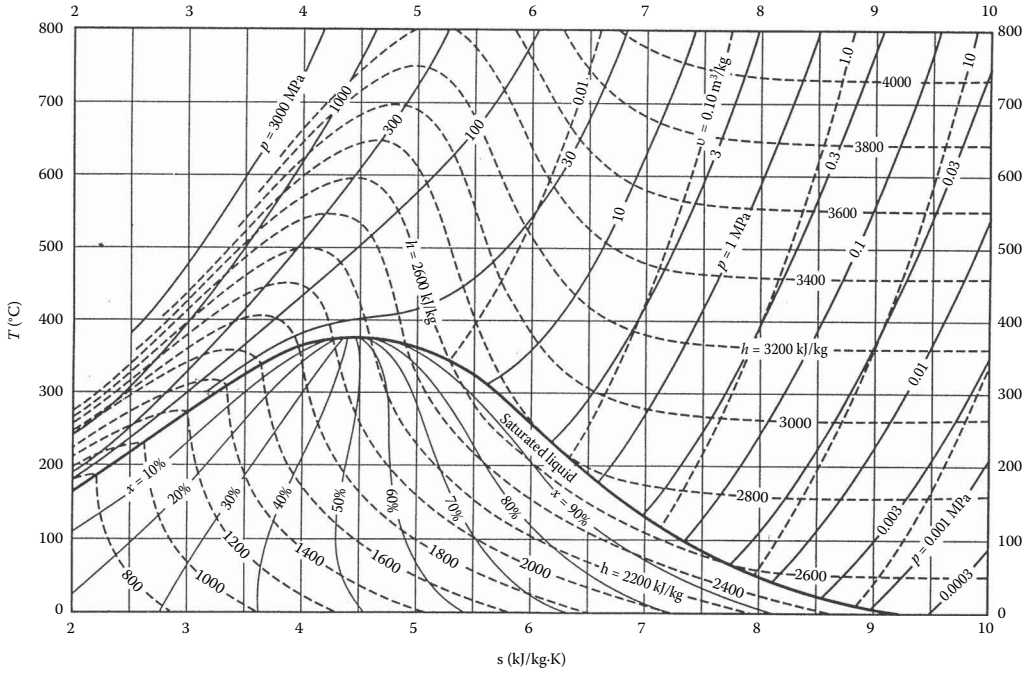


FIGURE 1.3.4 Temperature–entropy diagram for water. (From Jones, J.B. and Dugan, R.E., *Engineering Thermodynamics*, Prentice-Hall, Englewood Cliffs, NJ, 1996 based on data and formulations from Haar, L. et al., *NBS/NRC Steam Tables*. Hemisphere, Washington, DC, 1984.)

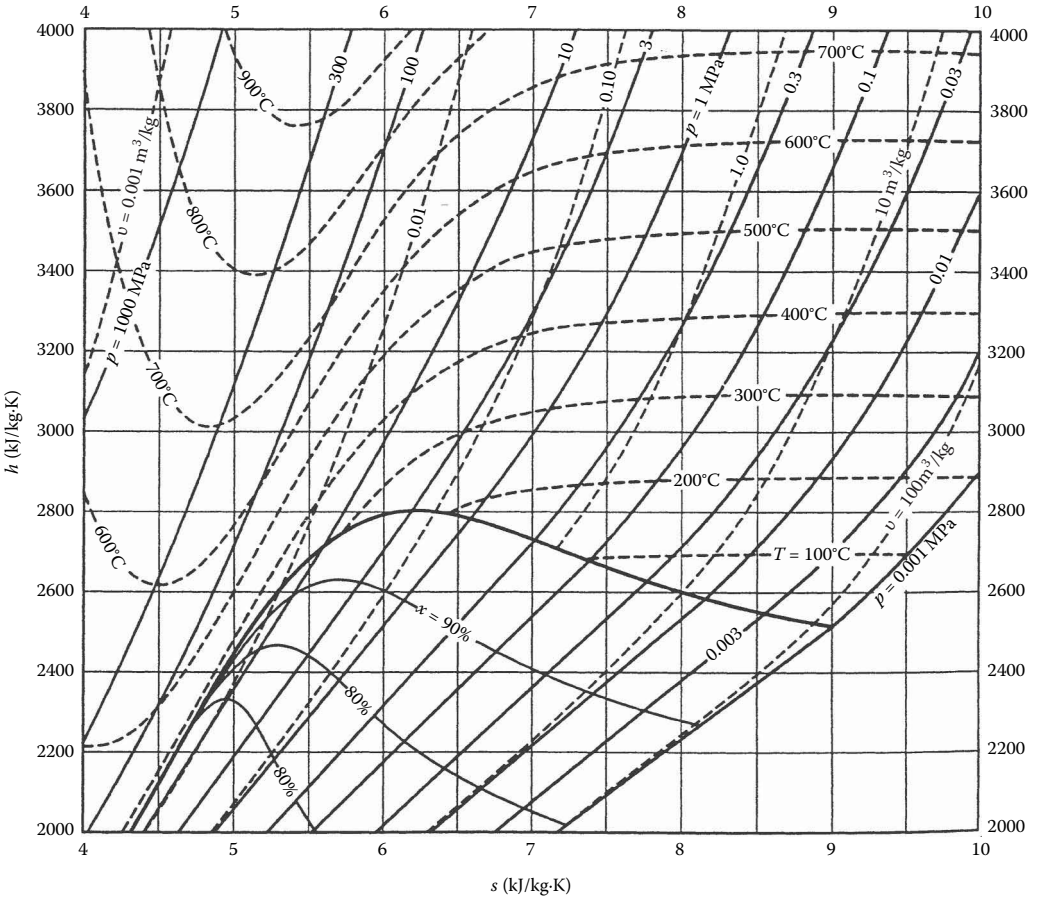


FIGURE 1.3.5 Enthalpy–entropy (Mollier) diagram for water. (From Jones, J.B. and Dugan, R.E., *Engineering Thermodynamics*, Prentice-Hall, Englewood Cliffs, NJ, 1996 based on data and formulations from Haar, L. et al., *NBS/NRC Steam Tables*, Hemisphere, Washington, DC, 1984.)

in Figure 1.3.6. The compressibility charts considered next use the compressibility factor as one of the coordinates.

Compressibility Charts

The p–v–T relation for a wide range of common gases is illustrated by the generalized compressibility chart shown in Figure 1.3.7. In this chart, the compressibility factor, Z, is plotted versus the *reduced* pressure, p_R, *reduced* temperature, T_R, and *pseudoreduced* specific volume, v'_R, where

$$Z = \frac{p\bar{v}}{\bar{R}T} \tag{1.3.9}$$

and

$$p_R = \frac{p}{p_c}, \quad T_R = \frac{T}{T_c}, \quad v'_R = \frac{\bar{v}}{\left(\bar{R}T_c/p_c\right)} \tag{1.3.10}$$

In these expressions, $\bar{R}$ is the universal gas constant, and p_c and T_c denote the critical pressure and temperature, respectively. Values of p_c and T_c are given for several substances in Table A.9. The

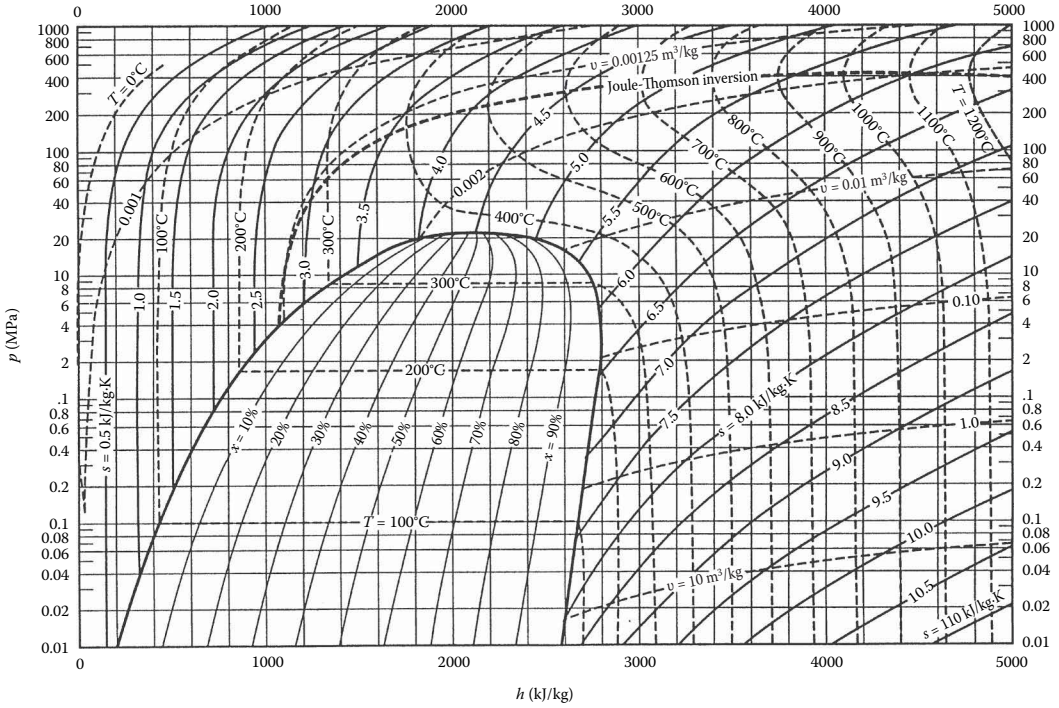


FIGURE 1.3.6 Pressure–enthalpy diagram for water. (From Jones, J.B. and Dugan, R.E., *Engineering Thermodynamics*, Prentice-Hall, Englewood Cliffs, NJ, 1996 based on data and formulations from Haar, L. et al., *NBS/NRC Steam Tables*, Hemisphere, Washington, DC, 1984.)

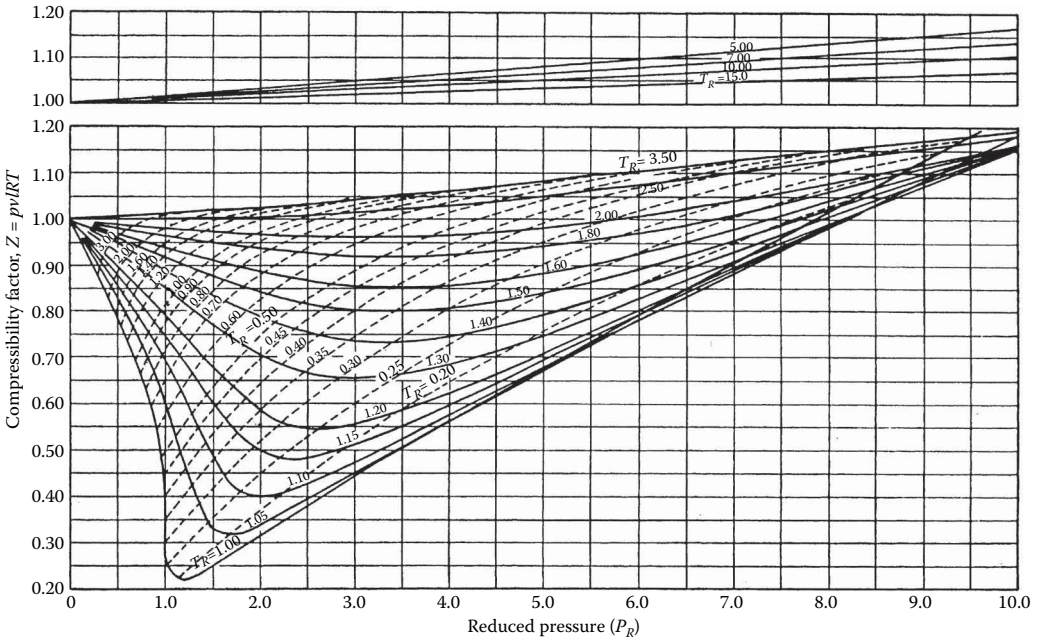


FIGURE 1.3.7 Generalized compressibility chart $\left(T_R = T/T_C,\ p_R = p/p_C,\ v'_R = \overline{v}p_C/RT_C\right)$ for $p_R \leq 10$. (From Obert, E.F. *Concepts of Thermodynamics*. McGraw-Hill, New York, 1960.)

reduced isotherms shown in Figure 1.3.7 represent the best curves fitted to the data of several gases. For the 30 gases used in developing the chart, the deviation of observed values from those of the chart is at most on the order of 5% and for most ranges is much less.*

Figure 1.3.7 gives a common value of about 0.27 for the compressibility factor at the critical point. As the critical compressibility factor for different substances actually varies from 0.23 to 0.33, the chart is inaccurate in the vicinity of the critical point. This source of inaccuracy can be removed by restricting the correlation to substances having essentially the same Z_c values, which is equivalent to including the critical compressibility factor as an independent variable: $Z=f(T_R, p_R, Z_c)$. To achieve greater accuracy, variables other than Z_c have been proposed as a third parameter—for example, the *acentric factor* (see, e.g., Reid and Sherwood, 1966).

Generalized compressibility data are also available in tabular form (see, e.g., Reid and Sherwood, 1966) and in equation form (see, e.g., Reynolds, 1979). The use of generalized data in any form (graphical, tabular, or equation) allows p, v, and T for gases to be evaluated simply and with reasonable accuracy. When accuracy is an essential consideration, generalized compressibility data should not be used as a substitute for p–v–T data for a given substance as provided by computer software, a table, or an equation of state.

Equations of State

Considering the isotherms shown in Figure 1.3.7, it is plausible that the variation of the compressibility factor might be expressed as an equation, at least for certain intervals of p and T. Two expressions can be written that enjoy a theoretical basis. One gives the compressibility factor as an infinite series expansion in pressure,

$$Z = 1 + \hat{B}(T)\,p + \hat{C}(T)\,p^2 + \hat{D}(T)\,p^3 + \cdots$$

and the other is a series in $1/\overline{v}$,

$$Z = 1 + \frac{B(T)}{\overline{v}} + \frac{C(T)}{\overline{v}^2} + \frac{D(T)}{\overline{v}^3} + \cdots$$

These expressions are known as *virial expansions*, and the coefficients $\hat{B}, \hat{C}, \hat{D}, \ldots$ and B, C, D, … are called *virial coefficients*. In principle, the virial coefficients can be calculated using expressions from statistical mechanics derived from consideration of the force fields around the molecules. Thus far only the first few coefficients have been calculated and only for gases consisting of relatively simple molecules. The coefficients also can be found, in principle, by fitting p–v–T data in particular realms of interest. Only the first few coefficients can be found accurately this way, however, and the result is a *truncated* equation valid only at certain states.

Over 100 equations of state have been developed in an attempt to portray accurately the p–v–T behavior of substances and yet avoid the complexities inherent in a full virial series. In general, these equations exhibit little in the way of fundamental physical significance and are mainly empirical in character. Most are developed for gases, but some describe the p–v–T behavior of the liquid phase, at least qualitatively. Every equation of state is restricted to particular states. The realm of applicability is often indicated by giving an interval of pressure, or density, where the equation can be expected to represent the p–v–T behavior faithfully. When it is not stated, the realm of applicability often may be approximated by expressing the equation in terms of the compressibility factor Z

* To determine Z for hydrogen, helium, and neon above a T_R of 5, the reduced temperature and pressure should be calculated using $T_R = T/(T_c + 8)$ and $P_R = p/(p_c + 8)$, where temperatures are in K and pressures are in atm.

and the reduced properties, and superimposing the result on a generalized compressibility chart or comparing with compressibility data from the literature.

Equations of state can be classified by the number of adjustable constants they involve. The Redlich–Kwong equation is considered by many to be the best of the two-constant equations of state. It gives pressure as a function of temperature and specific volume and thus is *explicit* in pressure:

$$p = \frac{\bar{R}T}{\bar{v} - b} - \frac{a}{\bar{v}\left(\bar{v} + b\right)T^{1/2}} \tag{1.3.11}$$

This equation is primarily empirical in nature, with no rigorous justification in terms of molecular arguments. Values for the Redlich–Kwong constants for several substances are provided in Table A.9. Modified forms of the equation have been proposed with the aim of achieving better accuracy.

Although the two-constant Redlich–Kwong equation performs better than some equations of state having several adjustable constants, two-constant equations tend to be limited in accuracy as pressure (or density) increases. Increased accuracy normally requires a greater number of adjustable constants. For example, the Benedict–Webb–Rubin equation, which involves eight adjustable constants, has been successful in predicting the *p–v–T* behavior of *light hydrocarbons*. The Benedict–Webb–Rubin equation is also explicit in pressure,

$$p = \frac{\bar{R}T}{\bar{v}} + \left(B\bar{R}T - A - \frac{C}{T^2}\right)\frac{1}{\bar{v}^2} + \frac{\left(b\bar{R}T - a\right)}{\bar{v}^3} + \frac{a\alpha}{\bar{v}^6} + \frac{c}{\bar{v}^3T^2}\left(1 + \frac{\gamma}{\bar{v}^2}\right)\exp\left(-\frac{\gamma}{\bar{v}^2}\right) \tag{1.3.12}$$

Values of the Benedict–Webb–Rubin constants for various gases are provided in the literature (see, e.g., Cooper and Goldfrank, 1967). A modification of the Benedict–Webb–Rubin equation involving 12 constants is discussed by Lee and Kessler (1975). Many multiconstant equations can be found in the engineering literature, and equations having 50 or more constants have been developed for representing the *p–v–T* behavior of different substances.

Gas Mixtures

Since an unlimited variety of mixtures can be formed from a given set of pure components by varying the relative amounts present, the properties of mixtures are reported only in special cases such as air. Means are available for predicting the properties of mixtures, however. Most techniques for predicting mixture properties are empirical in character and are not derived from fundamental physical principles. The realm of validity of any particular technique can be established by comparing predicted property values with empirical data. In this section, methods for evaluating the *p–v–T* relations for pure components are adapted to obtain plausible estimates for gas mixtures. The case of ideal gas mixtures is discussed in Section 1.3, Ideal Gas Model. In Section 1.3, Multicomponent Systems, some general aspects of property evaluation for multicomponent systems are presented.

The total number of moles of mixture, n, *is* the sum of the number of moles of the components, n_i:

$$n = n_1 + n_2 + \cdots + n_j = \sum_{i=1}^{j} n_i \tag{1.3.13}$$

The *relative* amounts of the components present can be described in terms of *mole fractions*. The mole fraction y_i of component i is $y_i = n_i/n$. The sum of the mole fractions of all components present is equal to unity. The apparent molecular weight M is the mole fraction average of the component molecular weights, such that

$$M = \sum_{i=1}^{j} y_i M_i \tag{1.3.14}$$

The *relative* amounts of the components present also can be described in terms of *mass fractions:* m_i/m, where m_i is the mass of component i and m is the total mass of mixture.

The p–v–T relation for a gas mixture can be estimated by applying an equation of state to the overall mixture. The constants appearing in the equation of state are *mixture values* determined with empirical combining rules developed for the equation. For example, mixture values of the constants a and b for use in the Redlich–Kwong equation are obtained using relations of the form

$$a = \left(\sum_{i=1}^{j} y_i a_i^{1/2} \right)^2, \quad b = \sum_{i=1}^{j} y_i b_i \tag{1.3.15}$$

where a_i and b_i are the values of the constants for component i. Combination rules for obtaining mixture values for the constants in other equations of state are also found in the literature.

Another approach is to regard the mixture as if it were a single pure component having critical properties calculated by one of several mixture rules. *Kay's rule* is perhaps the simplest of these, requiring only the determination of a mole fraction averaged critical temperature T_c and critical pressure p_c:

$$T_c = \sum_{i=1}^{j} y_i T_{c,i}, \quad p_c = \sum_{i=1}^{j} y_i p_{c,i} \tag{1.3.16}$$

where $T_{c,i}$ and $p_{c,i}$ are the critical temperature and critical pressure of component i, respectively. Using T_c and p_c, the mixture compressibility factor Z is obtained as for a single pure component. The unknown quantity from among the pressure p, volume V, temperature T, and total number of moles n of the gas mixture can then be obtained by solving $Z = pV/n\bar{R}T$.

Additional means for predicting the p–v–T relation of a mixture are provided by empirical mixture rules. Several are found in the engineering literature. According to the *additive pressure rule*, the pressure of a gas mixture is expressible as a sum of pressures exerted by the individual components:

$$p = p_1 + p_2 + p_3 \ldots]_{T,V} \tag{1.3.17a}$$

where the pressures p_1, p_2, etc., are evaluated by considering the respective components to be at the volume V and temperature T of the mixture. The additive pressure rule can be expressed alternatively as

$$Z = \sum_{i=1}^{j} y_i Z_i \Bigg]_{T,V} \tag{1.3.17b}$$

where Z is the compressibility factor of the mixture and the compressibility factors Z_i are determined assuming that component i occupies the entire volume of the mixture at the temperature T.

The *additive volume rule* postulates that the volume V of a gas mixture is expressible as the sum of volumes occupied by the individual components:

$$V = V_1 + V_2 + V_3 \ldots]_{p,T} \tag{1.3.18a}$$

where the volumes V_1, V_2, etc., are evaluated by considering the respective components to be at the pressure p and temperature T of the mixture. The additive volume rule can be expressed alternatively as

$$Z = \sum_{i=1}^{j} y_i Z_i \Bigg]_{p,T} \tag{1.3.18b}$$

where the compressibility factors Z_i are determined assuming that component i exists at the pressure p and temperature T of the mixture.

Evaluating Δh, Δu, and Δs

Using appropriate specific heat and p–v–T data, the changes in specific enthalpy, internal energy, and entropy can be determined between states of single-phase regions. Table 1.3.4 provides expressions for such property changes in terms of particular choices of the independent variables: temperature and pressure, and temperature and specific volume.

Taking Equation 1 of Table 1.3.4 as a representative case, the change in specific enthalpy between states 1 and 2 can be determined using the three steps shown in the accompanying property diagram. This requires knowledge of the variation of c_p with temperature at a fixed pressure p', and the variation of $[v - T(\partial v/\partial T)_p]$ with pressure at temperatures T_1 and T_2:

1-a: Since temperature is constant at T_1, the first integral of Equation 1 in Table 1.3.4 vanishes, and

$$h_a - h_1 = \int_{p_1}^{p'} \left[v - T\left(\partial v/\partial T\right)_p \right] \mathrm{d}p$$

a-b: Since pressure is constant at p', the second integral of Equation 1 vanishes, and

$$h_b - h_a = \int_{T_1}^{T_2} c_p\left(T, p'\right) \mathrm{d}T$$

b-2: Since temperature is constant at T_2, the first integral of Equation 1 vanishes, and

$$h_2 - h_b = \int_{p'}^{p_2} \left[v - T\left(\partial v/\partial T\right)_p \right] \mathrm{d}p$$

Adding these expressions, the result is $h_2 - h_1$. The required integrals may be performed numerically or analytically. The analytical approach is expedited when an equation of state explicit in specific volume is known.

Similar considerations apply to Equations 2 to 4 of Table 1.3.4. To evaluate $u_2 - u_1$ with Equation 3, for example, requires the variation of c_v with temperature at a fixed specific volume v', and the variation of $[T(\partial p/\partial T)_v - p]$ with specific volume at temperatures T_1 and T_2. An analytical approach to performing the integrals is expedited when an equation of state explicit in pressure is known.

As changes in specific enthalpy and internal energy are related through $h = u + pv$ by

$$h_2 - h_1 = \left(u_2 - u_1\right) + \left(p_2 v_2 - p_1 v_1\right) \tag{1.3.19}$$

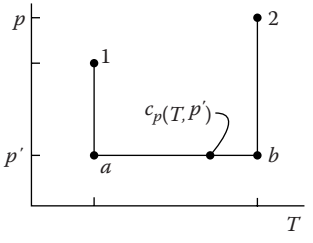
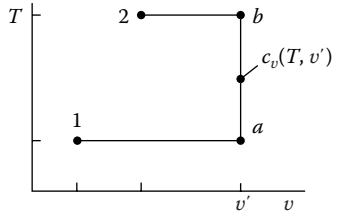
only one of $h_2 - h_1$ and $u_2 - u_1$ need be found by integration. The other can be evaluated from Equation 1.3.19. The one found by integration depends on the information available: $h_2 - h_1$ would be found when an equation of state explicit in v and c_p as a function of temperature at some fixed pressure is known, $u_2 - u_1$ would be found when an equation of state explicit in p and c_v as a function of temperature at some specific volume is known.

Example 1.3.3

Obtain Equation 1 of Table 1.3.3 and Equations 3 and 4 of Table 1.3.4.

Solution. With Equation 1.3.3 and the Maxwell relation corresponding to $\psi(T, v)$ from Table 1.3.1, Equations 3′ and 4′ of Table 1.3.4 become, respectively,

TABLE 1.3.4
Δh, Δu, Δs Expressions

Independent Properties:	**Temperature and Pressure**	**Temperature and Specific Volume**
Preferred data:	$v(T, p)$, $c_p(T, p)$	$p(T, v)$, $c_v(T, v)$
Property diagram:		
Property expressions:	$h(T, p)$:	$u(T, v)$:
	$$dh = \left(\frac{\partial h}{\partial T}\right)_p dT + \left(\frac{\partial h}{\partial p}\right)_T dp \quad (1')$$	$$du = \left(\frac{\partial u}{\partial T}\right)_v dT + \left(\frac{\partial u}{\partial v}\right)_T dv \quad (3')$$
	$c_p \qquad \left[v - T\left(\frac{\partial v}{\partial T}\right)_p\right]$	$c_v \qquad \left[T\left(\frac{\partial p}{\partial T}\right)_v - p\right]$
	$$\Delta h = \int c_p\, dT + \int \left[v - T\left(\frac{\partial v}{\partial T}\right)_p\right] dp \quad (1)$$	$$\Delta u = \int c_v\, dT + \int \left[T\left(\frac{\partial p}{\partial T}\right)_v - p\right] dv \quad (3)$$
	$s(T,p)$:	$s(T,v)$:
	$$ds = \left(\frac{\partial s}{\partial T}\right)_p dT + \left(\frac{\partial s}{\partial p}\right)_T dp \quad (2')$$	$$ds = \left(\frac{\partial s}{\partial T}\right)_v dT + \left(\frac{\partial s}{\partial v}\right)_T dv \quad (4')$$
	$\frac{c_p}{T} \qquad -\left(\frac{\partial v}{\partial T}\right)_p$	$\frac{c_v}{T} \qquad \left(\frac{\partial p}{\partial T}\right)_v$
	$$\Delta s = \int \frac{c_p}{T} dT - \int \left(\frac{\partial v}{\partial T}\right)_p dp \quad (2)$$	$$\Delta s = \int \frac{c_v}{T} dT + \int \left(\frac{\partial p}{\partial T}\right)_v dv \quad (4)$$

$$du = c_v\,dT + \left(\frac{\partial u}{\partial v}\right)_T dv$$

$$ds = \left(\frac{\partial s}{\partial T}\right)_v dT + \left(\frac{\partial p}{\partial T}\right)_v dv$$

Introducing these expressions for *du* and *ds* in Equation 1.3.2a, and collecting terms,

$$\left[T\left(\frac{\partial s}{\partial T}\right)_v - c_v\right]dT = \left[\left(\frac{\partial u}{\partial v}\right)_T + p - T\left(\frac{\partial p}{\partial T}\right)_v\right]dv$$

Since *T* and *v* are independent, the coefficients of d*T* and d*v* must vanish, giving, respectively,

$$\left(\frac{\partial s}{\partial T}\right)_v = \frac{c_v}{T}$$

$$\left(\frac{\partial u}{\partial v}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_v - p$$

The first of these corresponds to Equation 1 of Table 1.3.3 and Equation 4 of Table 1.3.4. The second of the earlier expressions establishes Equation 3 of Table 1.3.4. With similar considerations, Equation 3 of Table 1.3.3 and Equations 1 and 2 of Table 1.3.4 may be obtained.

Fundamental Thermodynamic Functions

A fundamental thermodynamic function is one that provides a complete description of the thermodynamic state. The functions $u(s, v)$, $h(s, p)$, $\psi(T, v)$, and $g(T, p)$ listed in Table 1.3.1 are fundamental thermodynamic functions.

In principle, all properties of interest can be determined from a fundamental thermodynamic function by differentiation and combination. Taking the function $\psi(T, v)$ as a representative case, the properties *v* and *T*, being the independent variables, are specified to fix the state. The pressure *p* and specific entropy *s* at this state can be determined by differentiation of $\psi(T, v)$, as shown in Table 1.3.1. By definition, $\psi = u - Ts$, so specific internal energy is obtained as

$$u = \psi + Ts$$

with *u*, *p*, and *v* known, the specific enthalpy can be found from the definition $h = u + pv$. Similarly, the specific Gibbs function is found from the definition $g = h - Ts$. The specific heat c_v can be determined by further differentiation $c_v = (\partial u/\partial T)_v$.

The development of a fundamental function requires the selection of a functional form in terms of the appropriate pair of independent properties and a set of adjustable coefficients that may number 50 or more. The functional form is specified on the basis of both theoretical and practical considerations. The coefficients of the fundamental function are determined by requiring that a set of selected property values and/or observed conditions be satisfied in a least-squares sense. This generally involves property data requiring the assumed functional form to be differentiated one or more times, for example, *p–v–T* and specific heat data. When all coefficients have been evaluated, the function is tested for accuracy by using it to evaluate properties for which accepted values are known such as *velocity of sound* and *Joule–Thomson* data. Once a suitable fundamental function is established, extreme accuracy in and consistency among the thermodynamic properties are possible. The properties of water tabulated by Keenan et al. (1969) and Haar et al. (1984) have been calculated from representations of the Helmholtz function.

Thermodynamic Data Retrieval

Tabular presentations of pressure, specific volume, and temperature are available for practically important gases and liquids. The tables normally include other properties useful for thermodynamic analyses, such as internal energy, enthalpy, and entropy. The various *steam* table*s* included in the references of this chapter provide examples. Computer software for retrieving the properties of a wide range of substances is also available, as, for example, the ASME Steam Tables (1993) and Bornakke and Sonntag (1996).

The sample *steam table* data presented in Table 1.3.5 are representative of data available for substances commonly encountered in mechanical engineering practice. Table A.5 and Figures 1.3.4 to 1.3.6 provide *steam table* data for a greater range of states. The form of the tables and figures, and how they are used are assumed to be familiar. In particular, the use of *linear interpolation* with such tables is assumed known.

Specific internal energy, enthalpy, and entropy data are determined relative to arbitrary datums, and such datums vary from substance to substance. Referring to Table 1.3.5a, the datum state for the specific internal energy and specific entropy of water is seen to correspond to saturated liquid water at 0.01°C (32.02°F), the triple point temperature. The value of each of these properties is set to zero at this state. If calculations are performed involving only differences in a particular specific property, the datum cancels. When there are changes in chemical composition during the process, special care should be exercised. The approach followed when composition changes due to chemical reaction is considered in Section 1.4.

Liquid water data (see Table 1.3.5d) suggests that at fixed temperature the variation of specific volume, internal energy, and entropy with pressure is slight. The variation of specific enthalpy with pressure at fixed temperature is somewhat greater because pressure is explicit in the definition of enthalpy. This behavior for v, u, s, and h is exhibited generally by liquid data and provides the basis for the following set of equations for estimating property data at liquid states from saturated liquid data:

$$v(T,p) \approx v_f(T) \tag{1.3.20a}$$

$$u(T,p) \approx u_f(T) \tag{1.3.20b}$$

$$h(T,p) \approx h_f(T) + v_f\left[p - p_{\text{sat}}(T)\right] \tag{1.3.20c}$$

$$s(T,p) \approx s_f(T) \tag{1.3.20d}$$

As before, the subscript f denotes the saturated liquid state at the temperature T, and p_{sat} *is* the corresponding saturation pressure. The second term on the right hand side of Equation 1.3.20c is often negligible, giving $h(T, p) \approx h_f(T)$, which is used in Example 1.2.3 to evaluate h_1.

In the absence of saturated liquid data, or as an alternative to such data, the *incompressible model* can be employed:

$$\text{Incompressible model}: \begin{cases} v = \text{constant} \\ u = u(T) \end{cases} \tag{1.3.21}$$

This model is also applicable to solids. Since internal energy varies only with temperature, the specific heat c_v is also a function of only temperature: $c_v(T) = du/dT$. Although specific volume is constant, enthalpy varies with both temperature and pressure, such that

$$h(T,p) = u(T) + pv \tag{1.3.22}$$

Differentiation of Equation 1.3.22 with respect to temperature at fixed pressure gives $c_p = c_v$. The common specific heat is often shown simply as c. Specific heat data for several liquids and solids are

TABLE 1.3.5
Sample Steam Table Data

(a) Properties of Saturated Water (Liquid–Vapor): Temperature Table

		Specific Volume (m^3/kg)		Internal Energy (kJ/kg)		Enthalpy (kJ/kg)			Entropy (kJ/kg·K)	
Temp. (°C)	Pressure (bar)	Saturated Liquid ($v_f \times 10^3$)	Saturated Vapor (v_g)	Saturated Liquid (u_f)	Saturated Vapor (u_g)	Saturated Liquid (h_f)	Evap. (h_{fg})	Saturated Vapor (h_g)	Saturated Liquid (s_f)	Saturated Vapor (s_g)
0.01	0.00611	1.0002	206.136	0.00	2375.3	0.01	2501.3	2501.4	0.0000	9.1562
4	0.00813	1.0001	157.232	16.77	2380.9	16.78	2491.9	2508.7	0.0610	9.0514
5	0.00872	1.0001	147.120	20.97	2382.3	20.98	2489.6	2510.6	0.0761	9.0257
6	0.00935	1.0001	137.734	25.19	2383.6	25.20	2487.2	2512.4	0.0912	9.0003
8	0.01072	1.0002	120.917	33.59	2386.4	33.60	2482.5	2516.1	0.1212	8.9501

(b) Properties of Saturated Water (Liquid–Vapor): Pressure Table

		Specific Volume (m^3/kg)		Internal Energy (kJ/kg)		Enthalpy (kJ/kg)			Entropy (kJ/kg·K)	
Pressure (bar)	Temp. (°C)	Saturated Liquid ($v_f \times 10^3$)	Saturated Vapor (v_g)	Saturated Liquid (u_f)	Saturated Vapor (u_g)	Saturated Liquid (h_f)	Evap. (h_{fg})	Saturated Vapor (h_g)	Saturated Liquid (s_f)	Saturated Vapor (s_g)
0.04	28.96	1.0040	34.800	121.45	2415.2	121.46	2432.9	2554.4	0.4226	8.4746
0.06	36.16	1.0064	23.739	151.53	2425.0	151.53	2415.9	2567.4	0.5210	8.3304
0.08	41.51	1.0084	18.103	173.87	2432.2	173.88	2403.1	2577.0	0.5926	8.2287
0.10	45.81	1.0102	14.674	191.82	2437.9	191.83	2392.8	2584.7	0.6493	8.1502
0.20	60.06	1.0172	7.649	251.38	2456.7	251.40	2358.3	2609.7	0.8320	7.9085

(Continued)

TABLE 1.3.5 (*Continued*)
Sample Steam Table Data

(c) Properties of Superheated Water Vapor

T (°C)	v (m³/kg)	u (kJ/kg)	h (kJ/kg)	s (kJ/kg·K)	v (m³/kg)	u (kJ/kg)	h (kJ/kg)	s (kJ/kg·K)
	$p = 0.06$ bar $= 0.006$ MPa ($T_{sat} = 36.16$°C)				$p = 0.35$ bar $= 0.035$ MPa ($T_{sat} = 72.69$°C)			
Sat.	23.739	2425.0	2567.4	8.3304	4.526	2473.0	2631.4	7.7158
80	27.132	2487.3	2650.1	8.5804	4.625	2483.7	2645.6	7.7564
120	30.219	2544.7	2726.0	8.7840	5.163	2542.4	2723.1	7.9644
160	33.302	2602.7	2802.5	8.9693	5.696	2601.2	2800.6	8.1519
200	36.383	2661.4	2879.7	9.1398	6.228	2660.4	2878.4	8.3237

(d) Properties of Compressed Liquid Water

T (°C)	$v \times 10^3$ (m³/kg)	u (kJ/kg)	h (kJ/kg)	s (kJ/kg·K)	$v \times 10^3$ (m³/kg)	u (kJ/kg)	h (kJ/kg)	s (kJ/kg·K)
	$p = 25$ bar $= 2.5$ MPa ($T_{sat} = 223.99$°C)				$p = 50$ bar $= 5.0$ MPa ($T_{sat} = 263.99$°C)			
20	1.0006	83.80	86.30	0.2961	0.9995	83.65	88.65	0.2956
80	1.0280	334.29	336.86	1.0737	1.0268	333.72	338.85	1.0720
140	1.0784	587.82	590.52	1.7369	1.0768	586.76	592.15	1.7343
200	1.1555	849.9	852.8	2.3294	1.1530	848.1	853.9	2.3255
Sat.	1.1973	959.1	962.1	2.5546	1.2859	1147.8	1154.2	2.9202

Source: Moran, M.J. and Shapiro, H.N., *Fundamentals of Engineering Thermodynamics*, 4th ed., Wiley, New York, 2000 as extracted from Keenan, J.H. et al., *Steam Tables*. Wiley, New York, 1969.

provided in Appendix tables B and C. As the variation of c with temperature is slight, c is frequently taken as constant.

When the incompressible model is applied, Equation 1.3.19 takes the form

$$\begin{aligned} h_2 - h_1 &= \int_{T_1}^{T_2} c(T)\,\mathrm{d}T + v(p_2 - p_1) \\ &= c_{\text{ave}}(T_2 - T_1) + v(p_2 - p_1) \end{aligned} \tag{1.3.23}$$

Also, as Equation 1.3.2a reduces to $\mathrm{d}u = T\,\mathrm{d}s$, and $\mathrm{d}u = c(T)\mathrm{d}T$, the change in specific entropy is

$$\begin{aligned} \Delta s &= \int_{T_1}^{T_2} \frac{c(T)}{T}\,\mathrm{d}T \\ &= c_{\text{ave}} \ln \frac{T_2}{T_1} \end{aligned} \tag{1.3.24}$$

Ideal Gas Model

Inspection of the generalized compressibility chart, Figure 1.3.7, shows that when p_R is small, and for many states when T_R is large, the value of the compressibility factor Z is close to 1. In other words, for pressures that are low relative to p_c, and for many states with temperatures high relative to T_c, the compressibility factor approaches a value of 1. Within the indicated limits, it may be assumed with reasonable accuracy that $Z = 1$, that is,

$$p\bar{v} = \bar{R}T \text{ or } pv = RT \tag{1.3.25a}$$

where $R = \bar{R}/M$ is the *specific* gas constant. Other forms of this expression in common use are

$$pV = n\bar{R}T, \quad pV = mRT \tag{1.3.25b}$$

Referring to Equation 3′ of Table 1.3.4, it can be concluded that $(\partial u/\partial v)_T$ vanishes identically for a gas whose equation of state is *exactly* given by Equation 1.3.25, and thus the specific internal energy depends only on temperature. This conclusion is supported by experimental observations beginning with the work of Joule, who showed that the internal energy of air at low density depends primarily on temperature.

These considerations allow for an *ideal gas model* of each real gas: (1) the equation of state is given by Equation 1.3.25 and (2) the internal energy and enthalpy are functions of temperature alone. The real gas approaches the model in the limit of low reduced pressure. At other states, the actual behavior may depart substantially from the predictions of the model. Accordingly, caution should be exercised when invoking the ideal gas model lest significant error is introduced.

Specific heat data for gases can be obtained by direct measurement. When extrapolated to zero pressure, ideal gas-specific heats result. Ideal gas-specific heats also can be calculated using molecular models of matter together with data from spectroscopic measurements. Table A.9 provides ideal gas-specific heat data for a number of substances. The following ideal gas-specific heat relations are frequently useful:

$$c_p(T) = c_v(T) + R \tag{1.3.26a}$$

$$c_p = \frac{kR}{k-1}, \quad c_v = \frac{R}{k-1} \tag{1.3.26b}$$

where $k = c_p/c_v$.

With the ideal gas model, Equations 1 through 4 of Table 1.3.4 give Equations 1 through 4 of Table 1.3.6, respectively. Equation 2 of Table 1.3.6 can be expressed alternatively using $s°(T)$ defined by

$$s^\circ(T) = \int_0^T \frac{c_p(T)}{T}\,dT \tag{1.3.27}$$

as

$$s(T_2, p_2) - s(T_1, p_1) = s^\circ(T_2) - s^\circ(T_1) - R\ln\frac{p_2}{p_1} \tag{1.3.28}$$

Expressions similar in form to Equations 1.3.26 through 1.3.38 can be written on a molar basis.

For processes of an ideal gas between states having the same specific entropy, $s_2 = s_1$, Equation 1.3.28 gives

$$\frac{p_2}{p_1} = \frac{\exp\left[s^\circ(T_2)/R\right]}{\exp\left[s^\circ(T_1)/R\right]}$$

or with $p_r = \exp[s°(T)/R]$

$$\frac{p_2}{p_1} = \frac{p_r(T_2)}{p_r(T_1)} \quad (s_2 = s_1) \tag{1.3.29a}$$

A relation between the specific volume and temperatures for two states of an ideal gas having the same specific entropy can also be developed:

$$\frac{v_2}{v_1} = \frac{v_r(T_2)}{v_r(T_1)} \quad (s_2 = s_1) \tag{1.3.29b}$$

TABLE 1.3.6
Ideal Gas Expressions for Δh, Δu, and Δs

Variable Specific Heats		Constant Specific Heats	
$h(T_2) - h(T_1) = \int_{T_1}^{T_2} c_p(T)\,dT$	(1)	$h(T_2) - h(T_1) = c_p(T_2 - T_1)$	(1′)
$s(T_2, p_2) - s(T_1, p_1) = \int_{T_1}^{T_2} \frac{c_p(T)}{T}\,dT - R\ln\frac{p_2}{p_1}$	(2)	$s(T_2, p_2) - s(T_1, p_1) = c_p\ln\frac{T_2}{T_1} - R\ln\frac{p_2}{p_1}$	(2′)
$u(T_2) - u(T_1) = \int_{T_1}^{T_2} c_v(T)\,dT$	(3)	$u(T_2) - u(T_1) = c_v(T_2 - T_1)$	(3′)
$s(T_2, v_2) - s(T_1, v_1) = \int_{T_1}^{T_2} \frac{c_v(T)}{T}\,dT + R\ln\frac{v_2}{v_1}$	(4)	$s(T_2, v_2) - s(T_1, v_1) = c_v\ln\frac{T_2}{T_1} + R\ln\frac{v_2}{v_1}$	(4′)
$s_2 = s_1$		$s_2 = s_1$	
$\frac{p_r(T_2)}{p_r(T_1)} = \frac{p_2}{p_1}$	(5)	$\frac{T_2}{T_1} = \left(\frac{p_2}{p_1}\right)^{(k-1)/k}$	(5′)
$\frac{v_r(T_2)}{v_r(T_1)} = \frac{v_2}{v_1}$	(6)	$\frac{T_2}{T_1} = \left(\frac{v_1}{v_2}\right)^{k-1}$	(6′)

Equations 1.3.29 are listed in Table 1.3.6 as Equations 5 and 6, respectively.

Table A.8 provides a tabular display of h, u, $s°$, p_r, and v_r versus temperature for air as an ideal gas. Tabulations of $\overline{h}$, $\overline{u}$, and $\overline{s}^\circ$ for several other common gases are provided in Table A.2. Property retrieval software also provides such data; see, e.g., Moran and Shapiro (2000). The use of data from Table A.8 for the nozzle of Example 1.2.2 is illustrated in Example 1.3.4.

When the ideal gas-specific heats are assumed constant, Equations 1 to 6 of Table 1.3.6 become Equations 1′ to 6′, respectively. The specific heat c_p is taken as constant in Example 1.2.2.

Example 1.3.4

Using data from Table A.8, evaluate the exit velocity for the nozzle of Example 1.2.2 and compare with the exit velocity for an isentropic expansion to 15 lbf/in.^{2}.

Solution: The exit velocity is given by Equation 1.2.9f

$$v_e = \sqrt{v_i^2 + 2\left(h_i - h_e\right)}$$

At 960 and 520°R, Table A.8 gives, respectively, h_i=231.06 Btu/lb and h_e=124.27 Btu/lb. Then

$$\mathrm{v}_e = \sqrt{\left(\frac{10\,\mathrm{ft}}{s}\right)^2 + 2\left(231.06 - 124.27\right)\left(\mathrm{Btu/lb}\right)\left(\frac{778.17\,\mathrm{ft\cdot lbf}}{1\mathrm{Btu}}\right)\left(\frac{32.174\,\mathrm{lb\cdot ft/s^2}}{1\mathrm{lbf}}\right)}$$

$$= 2312.5\,\mathrm{ft/s}$$

Using Equation 1.3.29a and p_r data from Table A.8, the specific enthalpy at the exit for an isentropic expansion is found as follows:

$$p_r\left(T_e\right) = p_r\left(T_i\right)\frac{p_e}{p_i} = 10.61\left(\frac{15}{150}\right) = 1.061$$

Interpolating with p_r data, h_e=119.54 Btu/lb. With this, the exit velocity is 2363.1 ft/s. The actual exit velocity is about 2% less than the velocity for an isentropic expansion, the maximum theoretical value. In this particular application, there is good agreement in each case between velocities calculated using Table A.8 data and, as in Example 1.2.2, assuming c_p constant. Such agreement cannot be expected generally, however. See, for example, the Brayton cycle data of Table 1.3.7.

Polytropic Processes

An internally reversible process described by the expression pv^n=constant is called a *polytropic process* and *n is* the *polytropic exponent*. Although this expression can be applied with real gas data, it most generally appears in practice together with the use of the ideal gas model. Table 1.3.7 provides several expressions applicable to polytropic processes and the special forms they take when the ideal gas model is assumed. The expressions for $\int p\,\mathrm{d}v$ and $\int v\,\mathrm{d}p$ have application to work evaluations with Equations 1.1.10 and 1.2.12, respectively. In some applications, it may be appropriate to determine n by fitting pressure-specific volume data.

Example 1.3.5 illustrates both the polytropic process and the reduction in the compressor work achievable by cooling a gas as it is compressed.

Example 1.3.5

A compressor operates at steady state with air entering at 1 bar, 20°C and exiting at 5 bar. (1) If the air undergoes a polytropic process with $n=1.3$, determine the work and heat transfer, each in kJ/kg of air flowing. Repeat for (2) an isothermal compression and (3) an isentropic compression.

Solution:

1. Using Equation 5′ of Table 1.3.7 together with Equation 1.2.12b,

$$\frac{\dot{W}_{cv}}{\dot{m}} = \frac{nRT_1}{n-1}\left[1-\left(\frac{p_2}{p_1}\right)^{(n-1)/n}\right]$$

$$= \left(\frac{1.3}{0.3}\right)\left(\frac{8.314}{28.97}\text{kJ/kg}\cdot\text{K}\right)(293\text{ K})\left[1-(5)^{0.3/1.3}\right]$$

$$= -163.9\text{kJ/kg}$$

(The area behind process 1–2 given in Figure 1.3.8, area 1–2–a–b, represents the magnitude of the work required, per unit mass of air flowing.) Also, Equation 1′ of Table 1.3.7 gives $T_2=425$ K. An energy rate balance at steady state and enthalpy data from Table A.8 gives

$$\frac{\dot{Q}_{cv}}{\dot{m}} = \frac{\dot{W}_{cv}}{\dot{m}} + h_2 - h_1$$

$$= -163.9 + (426.3 - 293.2) = -30.8\text{kJ/kg}$$

2. Using Equation 3′ of Table 1.3.7 together with Equation 1.2.12b,

$$\frac{\dot{W}_{cv}}{\dot{m}} = -RT\ln\frac{p_2}{p_1}$$

$$= -\left(\frac{8.314}{28.97}\right)(293)\ln 5$$

$$= -135.3\text{kJ/kg}$$

Area 1–2′–*a*–*b* in Figure 1.3.8 represents the magnitude of the work required, per unit of mass of air flowing. An energy balance reduces to give $\dot{Q}_{cv}/\dot{m} = \dot{W}_{cv}/\dot{m} = -135.3$kJ/kg.

3. For an isentropic compression, $\dot{Q}_{cv}=0$ and an energy rate balance reduces to give $\dot{W}_{cv}/\dot{m} = -(h_{2s} - h_1)$, where 2*s* denotes the exit state. With Equation 1.3.29a and p_r data, $h_{2s}=464.8$ kJ/kg ($T_{2s}=463$ K). Then $\dot{W}_{cv}/\dot{m} = -(464.8 - 293.2) = -171.6$ kJ/kg. Area 1–2*s*–*a*–*b* in Figure 1.3.8 represents the magnitude of the work required, per unit of mass of air flowing.

Ideal Gas Mixtures

When applied to an ideal gas mixture, the additive pressure rule (Section 1.3, *p*–*v*–*T* Relations) is known as the *Dalton model*. According to this model, each gas in the mixture acts as if it exists separately at the volume and temperature of the mixture. Applying the ideal gas equation of state to the mixture as a whole and to each component *i*, $pV = n\bar{R}T$, $p_iV = n_i\bar{R}T$, where p_i, the *partial pressure* of component *i*, is the pressure that component *i* would exert if n_i moles occupied the full volume *V* at the temperature *T*. Forming a ratio, the partial pressure of component *i* is

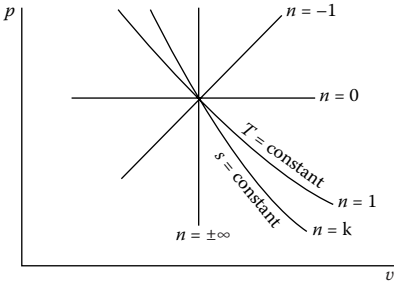
$$p_i = \frac{n_i}{n}p = y_i p \tag{1.3.30}$$

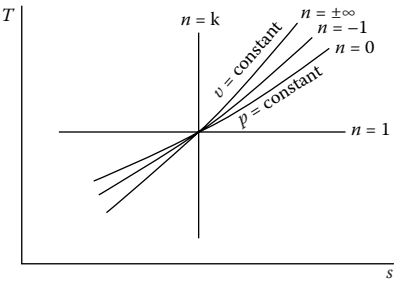
TABLE 1.3.7
Polytropic Processes: pv^n = Constant

General		Ideal Gas[a]	
$\frac{p_2}{p_1}=\left(\frac{v_1}{v_2}\right)^n$	(1)	$\frac{p_2}{p_1}=\left(\frac{v_1}{v_2}\right)^n=\left(\frac{T_2}{T_1}\right)^{n/(n-1)}$	(1′)
$n=0$: constant pressure $n=\pm\infty$: constant specific volume		$n=0$: constant pressure $n=\pm\infty$: constant specific volume $n=1$: constant temperature $n=k$: constant specific entropy when k is constant	
$n=1$		$n=1$	
$\int_1^2 p\,dv=p_1v_1\ln\frac{v_2}{v_1}$	(2)	$\int_1^2 p\,dv=RT\ln\frac{v_2}{v_1}$	(2′)
$-\int_1^2 v\,dp=-p_1v_1\ln\frac{p_2}{p_1}$	(3)	$-\int_1^2 v\,dp=-RT\ln\frac{p_2}{p_1}$	(3′)
$n\neq 1$		$n\neq 1$	
$\int_1^2 p\,dv=\frac{p_2v_2-p_1v_1}{1-n}=\frac{p_1v_1}{n-1}\left[1-\left(\frac{p_2}{p_1}\right)^{(n-1)/n}\right]$	(4)	$\int_1^2 p\,dv=\frac{R(T_2-T_1)}{1-n}=\frac{RT_1}{n-1}\left[1-\left(\frac{p_2}{p_1}\right)^{(n-1)/n}\right]$	(4′)
$-\int_1^2 v\,dp=\frac{n}{1-n}(p_2v_2-p_1v_1)=\frac{np_1v_1}{n-1}\left[1-\left(\frac{p_2}{p_1}\right)^{(n-1)/n}\right]$	(5)	$-\int_1^2 v\,dp=\frac{nR}{1-n}(T_2-T_1)=\frac{nRT_1}{n-1}\left[1-\left(\frac{p_2}{p_1}\right)^{(n-1)/n}\right]$	(5′)

Note: For polytropic processes of closed systems where volume change is the only work mode, Equations 2, 4, and 2′, 4′ are applicable with Equation 1.1.10 to evaluate the work. When each unit of mass passing through a one-inlet, one-exit control volume at steady state undergoes a polytropic process, Equations 3, 5, and 3′, 5′ are applicable with Equations 1.2.12a and b to evaluate the power. Also note that generally $-\int_1^2 v\,dp=n\int_1^2 p\,dv$.

[a]





where y_i is the mole fraction of component i. The sum of the partial pressures equals the mixture pressure.

The internal energy, enthalpy, and entropy of the mixture can be determined as the sum of the respective properties of the component gases, provided that the contribution from each gas is evaluated at the condition at which the gas exists in the mixture. On a *molar* basis,

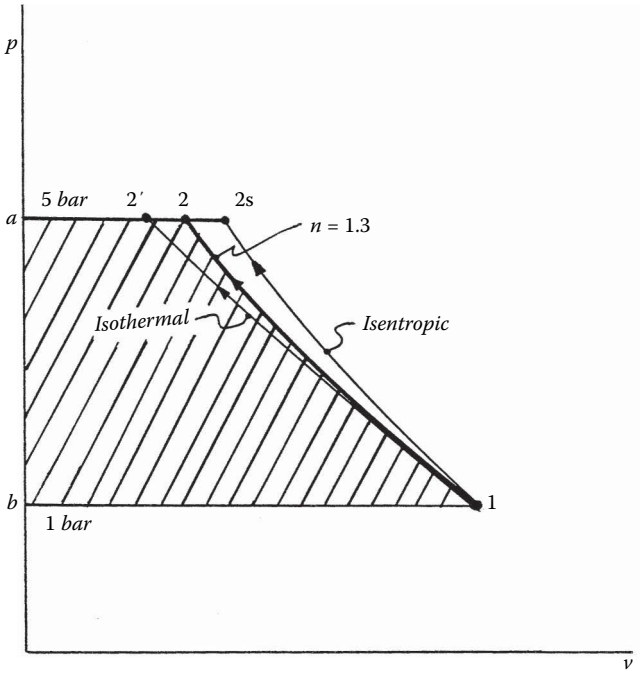


FIGURE 1.3.8 Internally reversible compression processes.

$$U = \sum_{i=1}^{j} n_i \overline{u}_i \text{ or } \overline{u} = \sum_{i=1}^{j} y_i \overline{u}_i \tag{1.3.31a}$$

$$H = \sum_{i=1}^{j} n_i \overline{h}_i \text{ or } \overline{h} = \sum_{i=1}^{j} y_i \overline{h}_i \tag{1.3.31b}$$

$$S = \sum_{i=1}^{j} n_i \overline{s}_i \text{ or } \overline{s} = \sum_{i=1}^{j} y_i \overline{s}_i \tag{1.3.31c}$$

The specific heats $\overline{c}_v$ and $\overline{c}_p$ for an ideal gas mixture in terms of the corresponding specific heats of the components are expressed similarly:

$$\overline{c}_v = \sum_{i=1}^{j} y_i \overline{c}_{vi} \tag{1.3.31d}$$

$$\overline{c}_p = \sum_{i=1}^{j} y_i \overline{c}_{pi} \tag{1.3.31e}$$

When working on a *mass* basis, expressions similar in form to Equations 1.3.31 can be written using *mass* and *mass fractions* in place of *moles* and *mole fractions*, respectively, and using u, h, s, c_p, and c_v in place of $\overline{u}$, $\overline{h}$, $\overline{s}$, $\overline{c}_p$, and $\overline{c}_v$, respectively.

The internal energy and enthalpy of an ideal gas depend only on temperature, and thus the $\overline{u}_i$ and $\overline{h}_i$ terms appearing in Equations 1.3.31 are evaluated at the temperature of the mixture. Since entropy depends on *two* independent properties, the $\overline{s}_i$ terms are evaluated either at the temperature

and the partial pressure p_i of component i, or at the temperature and volume of the mixture. In the former case,

$$S = \sum_{i=1}^{j} n_i \overline{s}_i \left(T, p_i\right)$$

$$= \sum_{i=1}^{j} n_i \overline{s}_i \left(T, x_i p\right) \tag{1.3.32}$$

Inserting the expressions for H and S given by Equations 1.3.31b and c into the Gibbs function, $G=H-TS$,

$$G = \sum_{i=1}^{j} n_i \overline{h}_i \left(T\right) - T \sum_{i=1}^{j} n_i \overline{s}_i \left(T, p_i\right)$$

$$= \sum_{i=1}^{j} n_i \overline{g}_i \left(T, p_i\right) \tag{1.3.33}$$

where the molar-specific Gibbs function of component i is $g_i(T, p_i)=h_i(T) - Ts_i(T, p_i)$. The Gibbs function of i can be expressed alternatively as

$$\overline{g}_i \left(T, p_i\right) = \overline{g}_i \left(T, p'\right) + \overline{R}T \ln\left(p_i / p'\right)$$

$$= \overline{g}_i \left(T, p'\right) + \overline{R}T \ln\left(x_i p / p'\right) \tag{1.3.34}$$

where p' is some specified pressure. Equation 1.3.34 is obtained by integrating Equation 1.3.2d at fixed temperature T from pressure p' to p_i.

Moist Air

An ideal gas mixture of particular interest for many practical applications is *moist air*. Moist air refers to a mixture of dry air and water vapor in which the dry air is treated as if it were a pure component. Ideal gas mixture principles usually apply to moist air. In particular, the *Dalton model* is applicable, and so the mixture pressure p is the sum of the partial pressures p_a and p_v of the dry air and water vapor, respectively.

Saturated air is a mixture of dry air and saturated water vapor. For saturated air, the partial pressure of the water vapor equals $p_{\text{sat}}(T)$, which is the saturation pressure of water corresponding to the dry-bulb (mixture) temperature T. The makeup of moist air can be described in terms of the *humidity ratio* (*specific humidity*) and the *relative humidity*. The bulb of a *wet-bulb thermometer* is covered with a wick saturated with liquid water, and the *wet-bulb* temperature of an air–water vapor mixture is the temperature indicated by such a thermometer exposed to the mixture.

When a sample of moist air is cooled at constant pressure, the temperature at which the sample becomes saturated is called the *dew point temperature*. Cooling below the dew point temperature results in the condensation of some of the water vapor initially present. When cooled to a final equilibrium state at a temperature below the dew point temperature, the original sample would consist of a gas phase of dry air and saturated water vapor in equilibrium with a liquid water phase.

Psychrometric charts are plotted with various moist air parameters, including the dry-bulb and wet-bulb temperatures, the humidity ratio, and the relative humidity, usually for a specified value of the mixture pressure such as 1 atm.

Generalized Charts for Enthalpy, Entropy, and Fugacity

The changes in enthalpy and entropy between two states can be determined in principle by correcting the respective property change determined using the ideal gas model. The corrections can be obtained, at least approximately, by inspection of the generalized enthalpy correction and entropy correction charts, Figures 1.3.9 and 1.3.10, respectively. Such data are also available in tabular form (see, e.g., Reid and Sherwood, 1966) and calculable using a generalized equation for the compressibility factor (Reynolds, 1979). Using the superscript * to identify ideal gas property values, the changes in specific enthalpy and specific entropy between states 1 and 2 are

$$\bar{h}_2 - \bar{h}_1 = \underline{\bar{h}_2^* - \bar{h}_1^*} - \underline{\bar{R}T_c\left[\left(\frac{\bar{h}^* - \bar{h}}{\bar{R}T_c}\right)_2 - \left(\frac{\bar{h}^* - \bar{h}}{\bar{R}T_c}\right)_1\right]} \tag{1.3.35a}$$

$$\bar{s}_2 - \bar{s}_1 = \underline{\bar{s}_2^* - \bar{s}_1^*} - \underline{\bar{R}\left[\left(\frac{\bar{s}^* - \bar{s}}{\bar{R}}\right)_2 - \left(\frac{\bar{s}^* - \bar{s}}{\bar{R}}\right)_1\right]} \tag{1.3.35b}$$

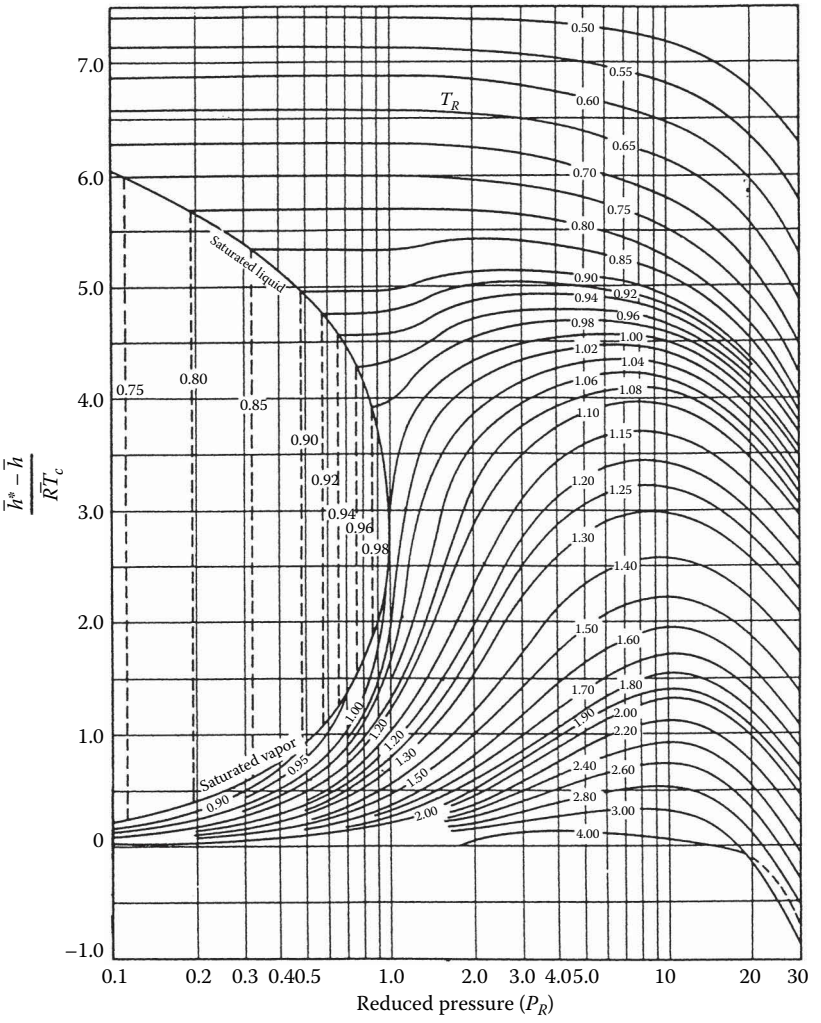


FIGURE 1.3.9 Generalized enthalpy correction chart. (Adapted from Van Wylen, G.J. and Sonntag, R.E., *Fundamentals of Classical Thermodynamics*, 3rd ed., English/SI., Wiley, New York, 1986.)

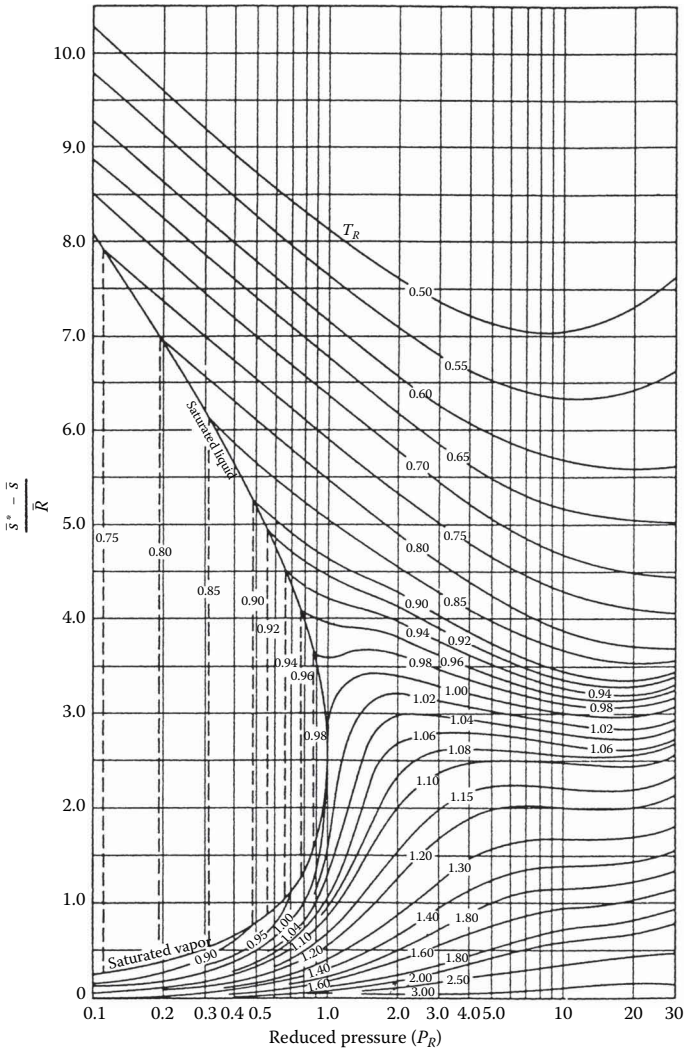


FIGURE 1.3.10 Generalized entropy correction chart. (Adapted from Van Wylen, G.J. and Sonntag, R.E., *Fundamentals of Classical Thermodynamics*, 3rd. ed., English/SI., Wiley, New York, 1986.)

The first underlined term on the right side of each expression represents the respective property change assuming ideal gas behavior. The second underlined term is the correction that must be applied to the ideal gas value to obtain the actual value. The quantities $\left(\overline{h}^* - \overline{h}\right)\Big/\overline{R}T_c$ and $\left(\overline{s}^* - \overline{s}\right)\Big/\overline{R}$ at state 1 would be read from the respective correction chart or table or calculated, using the reduced temperature T_{R1} and reduced pressure p_{R1} corresponding to the temperature T_1 and pressure p_1 at state 1, respectively. Similarly, $\left(\overline{h}^* - \overline{h}\right)\Big/\overline{R}T_c$ *and* $(\overline{s}^* - \overline{s})/\overline{R}$ at state 2 would be obtained using T_{R2} and p_{R2}. Mixture values for T_c and p_c determined by applying Kay's rule or some other mixture rule also can be used to enter the generalized enthalpy correction and entropy correction charts.

Figure 1.3.11 gives the *fugacity* coefficient, *f/p*, as a function of reduced pressure and reduced temperature. The fugacity *f* plays a similar role in determining the specific Gibbs function for a real gas as pressure plays for the ideal gas. To develop this, consider the variation of the specific Gibbs function with pressure at fixed temperature (from Table 1.3.1)

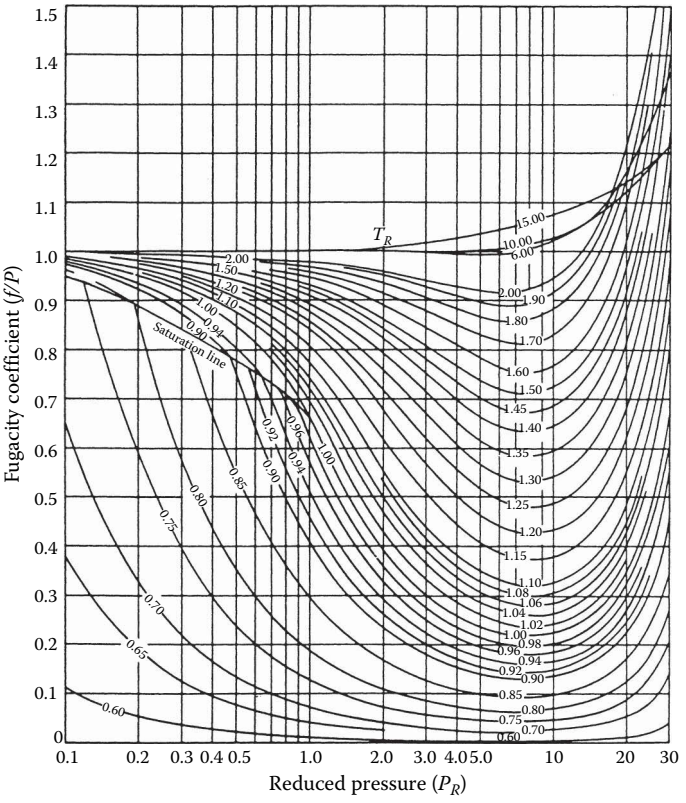


FIGURE 1.3.11 Generalized fugacity coefficient chart. (From Van Wylen, G.J. and Sonntag, R.E., *Fundamentals of Classical Thermodynamics*, 3rd ed., English/SI., Wiley, New York, 1986.)

$$\left(\frac{\partial g}{\partial p}\right)_T = v$$

For an ideal gas, integration at fixed temperature gives

$$g^* = RT \ln p + C(T)$$

where $C(T)$ is a function of integration. To evaluate g for a real gas, fugacity replaces pressure,

$$g = RT \ln f + C(T)$$

In terms of the fugacity coefficient, the departure of the real gas value from the ideal gas value at fixed temperature is then

$$g - g^* = RT \ln \frac{f}{p} \tag{1.3.36}$$

As pressure is reduced at fixed temperature, f/p tends to unity, and the specific Gibbs function is given by the ideal gas value.

Multicomponent Systems

In this section are presented some general aspects of the properties of multicomponent systems consisting of nonreacting mixtures. For a single-phase *multicomponent* system consisting of j

components, an extensive property X may be regarded as a function of temperature, pressure, and the number of moles of each component present in the mixture: $X=X(T, p, n_1, n_2,\ldots, n_j)$. Since X is mathematically *homogeneous of degree one* in the n's, the function is expressible as

$$X = \sum_{i=1}^{j} n_i \overline{X}_i \tag{1.3.37}$$

where the *partial molar property* $\overline{X}_i$ is by definition

$$\overline{X}_i = \left(\frac{\partial X}{\partial n_i} \right)_{T,p,n_l} \tag{1.3.38}$$

and the subscript n_l denotes that all n's except n_i are held fixed during differentiation. As $\overline{X}_i$ depends in general on temperature, pressure, and mixture composition: $\overline{X}_i\left(T, p, n_1, n_2, \ldots, n_j\right)$, the partial molal property $\overline{X}_i$ is an intensive property of the mixture and not simply a property of the ith component.

Selecting the extensive property X to be volume, internal energy, enthalpy, entropy, and the Gibbs function, respectively, gives

$$\begin{aligned} V &= \sum_{i=1}^{j} n_i \overline{V}_i, \quad U = \sum_{i=1}^{j} n_i \overline{U}_i \\ H &= \sum_{i=1}^{j} n_i \overline{H}_i, \quad S = \sum_{i=1}^{j} n_i \overline{S}_i \\ G &= \sum_{i=1}^{j} n_i \overline{G}_i \end{aligned} \tag{1.3.39}$$

where $\overline{V}_i$, $\overline{U}_i$, $\overline{H}_i$, $\overline{S}_i$, and $\overline{G}_i$ denote the respective partial molal properties.

When pure components, each initially at the same temperature and pressure, are mixed, the changes in volume, internal energy, enthalpy, and entropy on mixing are given by

$$\Delta V_{\text{mixing}} = \sum_{i=1}^{j} n_i \left(\overline{V}_i - \overline{v}_i \right) \tag{1.3.40a}$$

$$\Delta U_{\text{mixing}} = \sum_{i=1}^{j} n_i \left(\overline{U}_i - \overline{u}_i \right) \tag{1.3.40b}$$

$$\Delta H_{\text{mixing}} = \sum_{i=1}^{j} n_i \left(\overline{H}_i - \overline{h}_i \right) \tag{1.3.40c}$$

$$\Delta S_{\text{mixing}} = \sum_{i=1}^{j} n_i \left(\overline{S}_i - \overline{s}_i \right) \tag{1.3.40d}$$

where $\overline{v}_i$, $\overline{u}_i$, $\overline{h}_i$, and $\overline{s}_i$ denote the molar-specific volume, internal energy, enthalpy, and entropy of pure component i.

Chemical Potential

The partial molal Gibbs function of the *i*th component of a multicomponent system is the *chemical potential*, μ_i,

$$\mu_i = \overline{G}_i = \left(\frac{\partial G}{\partial n_i}\right)_{T,p,n_l} \tag{1.3.41}$$

Like temperature and pressure, the chemical potential, μ_i is an *intensive* property.

When written in terms of chemical potentials, Equation 1.3.37 for the Gibbs function reads

$$G = \sum_{i=1}^{j} n_i \mu_i \tag{1.3.42}$$

For a *single component system*, Equation 1.3.42 reduces to $G = n\mu$; that is, the chemical potential equals the molar Gibbs function. For an ideal gas mixture, comparison of Equations 1.3.33 and 1.3.42 suggests $\mu_i = \overline{g}_i(T, p_i)$; that is, the chemical potential of component *i* in an ideal gas mixture equals its Gibbs function per mole of gas *i* evaluated at the mixture temperature and the partial pressure of the *i*th gas of the mixture.

The chemical potential is a measure of the *escaping tendency* of a substance in a multiphase system: a substance tends to move from the phase having the higher chemical potential for that substance to the phase having a lower chemical potential. A necessary condition for *phase equilibrium is* that the chemical potential of each component has the same value in every phase.

The *Gibbs phase rule* gives the number *F* of independent intensive properties that may be arbitrarily specified to fix the intensive state of a system at equilibrium consisting of *N* nonreacting components present in *P* phases: $F = 2 + N - P$. *F* is called the *degrees of freedom* (or the *variance*). For water as a single component, for example, $N = 1$ and $F = 3 - P$.

- For a single phase, $P = 1$ and $F = 2$: Two intensive properties can be varied independently, say temperature *and* pressure, while maintaining a single phase.
- For two phases, $P = 2$ and $F = 1$: Only one intensive property can be varied independently if two phases are maintained—for example, temperature *or* pressure.
- For three phases, $P = 3$ and $F = 0$: There are no degrees of freedom; each intensive property of each phase is fixed. For a system consisting of ice, liquid water, and water vapor at equilibrium, there is a unique temperature: 0.01°C (32.02°F) and a unique pressure: 0.6113 kPa (0.006 atm).

The phase rule does not address the relative amounts that may be present in the various phases.

With $G = H - TS$ and $H = U + pV$, Equation 1.3.42 can be expressed as

$$U = TS - pV + \sum_{i=1}^{j} n_i \mu_i \tag{1.3.43}$$

from which can be derived

$$dU = T\,dS - p\,dV + \sum_{i=1}^{j} \mu_i\,dn_i \tag{1.3.44}$$

When the mixture composition is constant, Equation 1.3.44 reduces to Equation 1.3.1a.

Ideal Solution: The *Lewis–Randall rule* states that the fugacity $\overline{f}_i$ of each component *i* in an *ideal solution* is the product of its mole fraction and the fugacity of the pure component, f_i, at the same temperature, pressure, and state of aggregation (gas, liquid, or solid) as the mixture:

$$\overline{f}_i = y_i f_i \quad \text{(Lewis–Randall rule)} \tag{1.3.45}$$

The following characteristics are exhibited by an ideal solution: $\overline{V}_i = \overline{v}_i$, $\overline{U}_i = \overline{u}_i$, $\overline{H}_i = \overline{h}_i$. With these, Equations 1.3.40a, b, and c show that there is no change in volume, internal energy, or enthalpy on mixing pure components to form an ideal solution. The *adiabatic* mixing of different pure components would result in an increase in entropy, however, because such a process is irreversible.

The volume of an ideal solution is

$$V = \sum_{i=1}^{j} n_i \overline{v}_i = \sum_{i=1}^{j} V_i \quad \left(\text{ideal solution}\right) \tag{1.3.46}$$

where V_i is the volume that pure component i would occupy when at the temperature and pressure of the mixture. Comparing Equations 1.3.18a and 1.3.46, the *additive volume rule* is seen to be exact for ideal solutions. The internal energy and enthalpy of an ideal solution are

$$U = \sum_{i=1}^{j} n_i \overline{u}_i, \qquad H = \sum_{i=1}^{j} n_i \overline{h}_i \qquad \left(\text{ideal solution}\right) \tag{1.3.47}$$

where $\overline{u}_i$ and $\overline{h}_i$ denote, respectively, the molar internal energy and enthalpy of pure component i at the temperature and pressure of the mixture. Many gaseous mixtures at low to moderate pressures are adequately modeled by the Lewis-Randall rule. The ideal gas mixtures considered in Section 1.3, Ideal Gas Model, is an important special case. Some liquid solutions also can be modeled with the Lewis–Randall rule.

1.4 COMBUSTION

Michael J. Moran

The thermodynamic analysis of reactive systems is primarily an extension of principles presented in Sections 1.1 through 1.3. It is necessary, though, to modify the methods used to evaluate specific enthalpy and entropy.

Reaction Equations

In combustion reactions, rapid oxidation of combustible elements of the fuel results in energy release as combustion products are formed. The three major combustible chemical elements in most common fuels are carbon, hydrogen, and sulfur. Although sulfur is usually a relatively unimportant contributor to the energy released, it can be a significant cause of pollution and corrosion.

The emphasis in this section is on hydrocarbon fuels, which contain hydrogen, carbon, sulfur, and possibly other chemical substances. Hydrocarbon fuels may be liquids, gases, or solids such as coal. Liquid hydrocarbon fuels are commonly derived from crude oil through distillation and cracking processes. Examples are gasoline, diesel fuel, kerosene, and other types of fuel oils. The compositions of liquid fuels are commonly given in terms of mass fractions. For simplicity in combustion calculations, gasoline is often considered to be octane, C_8H_{18}, and diesel fuel is considered to be dodecane, $C_{12}H_{26}$. Gaseous hydrocarbon fuels are obtained from natural gas wells or are produced in certain chemical processes. Natural gas normally consists of several different hydrocarbons, with the major constituent being methane, CH_4. The compositions of gaseous fuels are commonly given in terms of mole fractions. Both gaseous and liquid hydrocarbon fuels can be synthesized from coal, oil shale, and tar sands. The composition of coal varies considerably with the location from which it is mined. For combustion calculations, the makeup of coal is usually expressed as an *ultimate analysis* giving the composition on a mass basis in terms of the relative amounts of chemical elements (carbon, sulfur, hydrogen, nitrogen, oxygen) and ash.

A fuel is said to have burned *completely* if all of the carbon present in the fuel is burned to carbon dioxide, all of the hydrogen is burned to water, and all of the sulfur is burned to sulfur dioxide. In practice, these conditions are usually not fulfilled and combustion is *incomplete*. The presence of carbon monoxide (CO) in the products indicates incomplete combustion. The products of combustion of *actual* combustion reactions and the relative amounts of the products can be determined with certainty only by experimental means. Among several devices for the experimental determination of the composition of products of combustion are the *Orsat analyzer, gas chromatograph, infrared analyzer*, and *flame ionization detector*. Data from these devices can be used to determine the makeup of the gaseous products of combustion. Analyses are frequently reported on a "dry" basis: Mole fractions are determined for all gaseous products as if no water vapor were present. Some experimental procedures give an analysis including the water vapor, however.

Since water is formed when hydrocarbon fuels are burned, the mole fraction of water vapor in the gaseous products of combustion can be significant. If the gaseous products of combustion are cooled at constant mixture pressure, the *dew point temperature* (Section 1.3, Ideal Gas Model) is reached when water vapor begins to condense. Corrosion of duct work, mufflers, and other metal parts can occur when water vapor in the combustion products condenses.

Oxygen is required in every combustion reaction. Pure oxygen is used only in special applications such as cutting and welding. In most combustion applications, air provides the needed oxygen. Idealizations are often used in combustion calculations involving air: (1) all components of air other than oxygen (O_2) *are* lumped with nitrogen (N_2). On a molar basis, air is then considered to be 21% oxygen and 79% nitrogen. With this idealization, the molar ratio of the nitrogen to the oxygen in combustion air is 3.76; (2) the water vapor present in air may be considered in writing the combustion equation or ignored. In the latter case, the combustion air is regarded as *dry*; (3) additional

simplicity results by regarding the nitrogen present in the combustion air as inert. However, if high-enough temperatures are attained, nitrogen can form compounds, often termed NO_X, such as nitric oxide and nitrogen dioxide. Even trace amounts of oxides of nitrogen appearing in the exhaust of internal combustion engines can be a source of air pollution.

The minimum amount of air that supplies sufficient oxygen for the complete combustion of all the combustible chemical elements is the *theoretical*, or *stoichiometric*, amount of air. In practice, the amount of air actually supplied may be greater than or less than the theoretical amount, depending on the application. The amount of air is commonly expressed as the *percent of theoretical air* or the *percent excess* (or *percent deficiency*) of air. The *air–fuel ratio* and its reciprocal *the fuel–air ratio*, each of which can be expressed on a mass or molar basis, are other ways that fuel–air mixtures are described. Another is the *equivalence ratio*: the ratio of the actual fuel–air ratio to the fuel–air ratio for complete combustion with the theoretical amount of air. The reactants form a *lean* mixture when the equivalence ratio is less than unity and a *rich* mixture when the ratio is greater than unity.

Example 1.4.1

Methane, CH_4, is burned with dry air. The molar analysis of the products on a dry basis is CO_2, 9.7%; CO, 0.5%; O_2, 2.95%; and N_2, 86.85%. Determine (1) the air–fuel ratio on both a molar and a mass basis, (2) the percent of theoretical air, (3) the equivalence ratio, and (4) the dew point temperature of the products, in °F, if the pressure is 1 atm.

***Solution*:**

1. The solution is conveniently conducted on the basis of 100 lbmol of dry products. The chemical equation then reads

$$a\mathrm{CH_4} + b(\mathrm{O_2} + 3.76\mathrm{N_2}) \rightarrow 9.7\mathrm{CO_2} + 0.5\mathrm{CO} + 2.95\mathrm{O_2} + 86.85\mathrm{N_2} + c\mathrm{H_2O},$$

   where N_2 is regarded as inert. Water is included in the products together with the assumed 100 lbmol of dry products. Balancing the carbon, hydrogen, and oxygen, the reaction equation is

$$10.2\mathrm{CH_4} + 23.1\left(\mathrm{O_2} + 3.76\mathrm{N_2}\right) \rightarrow 9.7\mathrm{CO_2} + 0.5\mathrm{CO} + 2.95\mathrm{O_2} + 86.85\mathrm{N_2} + 20.4\mathrm{H_2O}.$$

   The nitrogen also balances, as can be verified. This checks the accuracy of both the given product analysis and the calculations conducted to determine the unknown coefficients. Exact closure cannot be expected with measured data, however. On a molar basis, the air–fuel ratio is

$$\overline{\mathrm{AF}} = \frac{23.1(4.76)}{10.2} = 10.78 \frac{\text{lbmol (air)}}{\text{lbmol (fuel)}}.$$

   On a mass basis,

$$\mathrm{AF} = (10.78)\left(\frac{28.97}{16.04}\right) = 19.47 \frac{\text{lb (air)}}{\text{lb (fuel)}}.$$

2. The balanced chemical equation for the complete combustion of methane with the *theoretical* amount of air is

$$\mathrm{CH_4} + 2\left(\mathrm{O_2} + 3.76\mathrm{N_2}\right) \rightarrow \mathrm{CO_2} + 2\mathrm{H_2O} + 7.52\mathrm{N_2}.$$

The theoretical air–fuel ratio on a molar basis is

$$\left(\overline{AF}\right)_{theo} = \frac{2(4.76)}{1} = 9.52 \frac{\text{lbmol (air)}}{\text{lbmol (fuel)}}$$

The percent theoretical air is then

$$\% \text{ theoretical air} = \frac{\left(\overline{AF}\right)}{\left(\overline{AF}\right)_{theo}}$$

$$= \frac{10.78}{9.52} = 1.13(113\%).$$

3. Equivalence ratio $= (\overline{FA})/(\overline{FA})_{theo} = 9.52 / 10.78 = 0.88$. The reactants form a lean mixture.
4. To determine the dew point temperature requires the partial pressure p_v of the water vapor. The mole fraction of the water vapor is

$$y_v = \frac{20.4}{100 + 20.4} = 0.169.$$

Since $p = 1$ atm, $p_v = 0.169$ atm $= 2.48$ lbf/in.^{2}. With $p_{sat} = 2.48$ lbf/in.^{2}, the corresponding saturation temperature from the steam tables is 134°F. This is the dew point temperature.

Property Data for Reactive Systems

Tables of thermodynamic properties such as the steam tables provide values for the specific enthalpy and entropy relative to some arbitrary datum state where the enthalpy (or alternatively, the internal energy) and entropy are set to zero. When a chemical reaction occurs, however, reactants disappear and products are formed, and it is generally no longer possible to evaluate $\Delta\overline{h}$ and $\Delta\overline{s}$ so that these arbitrary datums cancel. Accordingly, special means are required to assign specific enthalpy and entropy for application to reacting systems.

Property data suited for the analysis of reactive systems are available from several sources. The encyclopedic *JANAF Thermochemical Tables* is commonly used. Data for a wide range of substances are retrievable from Knacke et al. (1991), which provides both tabular data and analytical expressions readily programmable for use with personal computers of the specific heat, enthalpy, entropy, and Gibbs function. Textbooks on engineering thermodynamics also provide selected data, as, for example, Moran et al. (2014).

Enthalpy of Formation

An enthalpy datum for reacting systems can be established by assigning arbitrarily a value of zero to the enthalpy of the *stable elements* at a *standard reference state* where the temperature is $T_{ref} = 298.15$ K (25°C) and the pressure is p_{ref}, which may be 1 bar or 1 atm depending on the data source. The term stable simply means that the particular element is chemically stable. For example, at the standard state, the stable forms of hydrogen, oxygen, and nitrogen are H_2, O_2, and N_2 and not the monatomic H, O, and N.

The molar enthalpy of a *compound* at the standard state equals its *enthalpy of formation*, symbolized here by $\overline{h}_f^\circ$. The enthalpy of formation is the energy released or absorbed when the compound is formed from its elements, the compound and elements all being at T_{ref} and p_{ref}. The enthalpy of formation may be determined by application of procedures from statistical thermodynamics using observed spectroscopic data. The enthalpy of formation also can be found in principle by measuring the heat transfer in a reaction in which the compound is formed from the elements. In this chapter, the superscript ° is used to denote p_{ref}. For the case of the enthalpy of formation, the reference temperature T_{ref} is also intended by this symbol. Table 1.4.1 gives the values of the enthalpy of formation of various substances at 298 K and 1 atm.

TABLE 1.4.1
Enthalpy of Formation, Gibbs Function of Formation, and Absolute Entropy of Various Substances at 298 K and 1 atm

Substance	Formula	$\bar{h}_f^\circ$ (kJ/kmol)	$\bar{g}_f^\circ$ (kJ/kmol)	$\bar{s}^\circ$ (kJ/kmol · K)
Carbon	C(s)	0	0	5.74
Hydrogen	H_2(g)	0	0	130.57
Nitrogen	N_2(g)	0	0	191.50
Oxygen	O_2(g)	0	0	205.03
Carbon monoxide	CO(g)	–110,530	–137,150	197.54
Carbon dioxide	CO_2(g)	–393,520	–394,380	213.69
Water	H_2O(g)	–241,820	–228,590	188.72
	H_2O(l)	–285,830	–237,180	69.95
Hydrogen peroxide	H_2O_2(g)	–136,310	–105,600	232.63
Ammonia	NH_3(g)	–46,190	–16,590	192.33
Oxygen	O(g)	249,170	231,770	160.95
Hydrogen	H(g)	218,000	203,290	114.61
Nitrogen	N(g)	472,680	455,510	153.19
Hydroxyl	OH(g)	39,460	34,280	183.75
Methane	CH_4(g)	–74,850	–50,790	186.16
Acetylene	C_2H_2(g)	226,730	209,170	200.85
Ethylene	C_2H_4(g)	52,280	68,120	219.83
Ethane	C_2H_6(g)	–84,680	–32,890	229.49
Propylene	C_3H_6(g)	20,410	62,720	266.94
Propane	C_3H_8(g)	–103,850	–23,490	269.91
Butane	C_4H_{10}(g)	–126,150	–15,710	310.03
Pentane	C_5H_{12}(g)	–146,440	–8,200	348.40
Octane	C_8H_{18}(g)	–208,450	17,320	463.67
	C_8H_{18}(l)	–249,910	6,610	360.79
Benzene	C_6H_6(g)	82,930	129,660	269.20
Methyl alcohol	CH_3OH(g)	–200,890	–162,140	239.70
	CH_3OH(l)	–238,810	–166,290	126.80
Ethyl alcohol	C_2H_5OH(g)	–235,310	–168,570	282.59
	C_2H_5OH(l)	–277,690	174,890	160.70

Source: Wark, K., *Thermodynamics*, 4th ed., McGraw-Hill, New York, 1983 as based on JANAF Thermochemical Tables, NSRDS-NBS-37, 1971; *Selected Values of Chemical Thermodynamic Properties*, NBS Tech. Note 270-3, 1968; and *API Research Project 44*, Carnegie Press, 1953.

The molar enthalpy of a substance at a state other than the standard state is found by adding the molar enthalpy change $\Delta\bar{h}$ between the standard state and the state of interest to the molar enthalpy of formation:

$$\bar{h}(T,p) = \bar{h}_f^\circ + \left[\bar{h}(T,p) - \bar{h}\left(T_{ref}, p_{ref}\right)\right] = \bar{h}_f^\circ + \Delta\bar{h}. \tag{1.4.1}$$

That is, the enthalpy of a substance is composed of $\bar{h}_f^\circ$, associated with the formation of the substance from its elements, and $\Delta\bar{h}$, associated with a change of state at constant composition. An arbitrarily chosen datum can be used to determine $\Delta\bar{h}$, since it is a *difference* at constant composition. Accordingly, $\Delta\bar{h}$ can be evaluated from sources such as the steam tables and the ideal gas tables.

The *enthalpy of combustion*, $\bar{h}_{RP}$, is the difference between the enthalpy of the products and the enthalpy of the reactants, each on a per-mole-of-fuel basis, when complete combustion occurs and both reactants and products are at the same temperature and pressure. For hydrocarbon fuels, the

enthalpy of combustion is negative in value since chemical internal energy is liberated in the reaction. The *heating value* of a fuel is a positive number equal to the magnitude of the enthalpy of combustion. Two heating values are recognized: the *higher* heating value and the *lower* heating value. The higher heating value is obtained when all the water formed by combustion is a liquid; the lower heating value is obtained when all the water formed by combustion is a vapor. The higher heating value exceeds the lower heating value by the energy that would be required to vaporize the liquid water formed at the specified temperature. Heating values are typically reported at a temperature of 25°C (77°F) and a pressure of 1 bar (or 1 atm). These values also depend on whether the fuel is a liquid or a gas. A sampling is provided on a unit mass-of-fuel basis in Table 1.4.2.

In the absence of work $\dot{W}_{cv}$ and appreciable kinetic and potential energy effects, the energy liberated on combustion is transferred from a reactor at steady state in two ways: the energy accompanying the exiting combustion products and by heat transfer. The temperature that would be achieved by the products in the limit of adiabatic operation is the *adiabatic flame* or *adiabatic combustion* temperature.

For a specified fuel and specified temperature and pressure of the reactants, the *maximum* adiabatic flame temperature is realized for complete combustion with the theoretical amount of air. Example 1.4.2 provides an illustration. The measured value of the temperature of the combustion products may be several hundred degrees below the calculated maximum adiabatic flame temperature, however, for several reasons including the following: (1) heat loss can be reduced but not eliminated; (2) once adequate oxygen has been provided to permit complete combustion, bringing in more air dilutes the combustion products, lowering the temperature; (3) incomplete combustion tends to reduce the temperature of the products, and combustion is seldom complete; and (4) as result of the high temperatures achieved, some of the combustion products may dissociate. Endothermic dissociation reactions also lower the product temperature.

Absolute Entropy

A common datum for assigning entropy values to substances involved in chemical reactions is realized through the *third law* of thermodynamics, which is based on experimental observations obtained primarily from studies of chemical reactions at low temperatures and specific heat measurements at temperatures approaching absolute zero. The third law states that the entropy of a pure crystalline substance is zero at the absolute zero of temperature, 0 K or 0°R. Substances not having a pure crystalline structure have a nonzero value of entropy at absolute zero.

TABLE 1.4.2
Heating Values in kJ/kg of Selected Hydrocarbons at 25°C

		Higher Value[a]		Lower Value[b]	
Hydrocarbon	**Formula**	**Liquid Fuel**	**Gas. Fuel**	**Liquid Fuel**	**Gas. Fuel**
Methane	CH_4	—	55,496	—	50,010
Ethane	C_2H_6	—	51,875	—	47,484
Propane	C_3H_8	49,973	50,343	45,982	46,352
n-Butane	C_4H_{10}	49,130	49,500	45,344	45,714
n-Octane	C_8H_{18}	47,893	48,256	44,425	44,788
n-Dodecane	$C_{12}H_{26}$	47,470	47,828	44,109	44,467
Methanol	CH_3OH	22,657	23,840	19,910	21,093
Ethanol	C_2H_5OH	29,676	30,596	26,811	27,731

[a] H_2O liquid in the products.
[b] H_2O vapor in the products.

The third law provides a datum relative to which the entropy of each substance participating in a reaction can be evaluated. The entropy relative to this datum is called the *absolute* entropy. The change in entropy of a substance between absolute zero and any given state can be determined from measurements of energy transfers and specific heat data or from procedures based on statistical thermodynamics and observed molecular data. Tables 1.4.1, A.2, and A.8 provide absolute entropy data for various substances. In these tables, $p_{\text{ref}} = 1$ atm.

When the absolute entropy is known at pressure p_{ref} and temperature T, the absolute entropy at the same temperature and *any* pressure p can be found from

$$\overline{s}(T, p) = \overline{s}(T, p_{\text{ref}}) + \left[\overline{s}(T, p) - \overline{s}(T, p_{\text{ref}})\right]. \tag{1.4.2}$$

For an ideal gas, the second term on the right side of Equation 1.4.2 can be evaluated by using Equation 1.3.28, giving

$$\overline{s}(T, p) = \overline{s}^{\circ}(T) - \overline{R} \ln \frac{p}{p_{\text{ref}}} \quad \text{(ideal gas)}. \tag{1.4.3}$$

In this expression, $\overline{s}^{\circ}(T)$ denotes the absolute entropy at temperature T and pressure p_{ref}.

The entropy of the ith component of an *ideal gas mixture* is evaluated at the mixture temperature T and the *partial* pressure p_i : $\overline{s}_i(T, p_i)$. For the ith component, Equation 1.4.3 takes the form

$$\begin{aligned} \overline{s}_i(T, p_i) &= \overline{s}_i^{\circ}(T) - \overline{R} \ln \frac{p_i}{p_{\text{ref}}} \\ &= \overline{s}_i^{\circ}(T) - \overline{R} \ln \frac{y_i p}{p_{\text{ref}}} \quad \text{(ideal gas)}, \end{aligned} \tag{1.4.4}$$

where $\overline{s}_i^{\circ}(T)$ is the absolute entropy of component i at temperature T and p_{ref}.

Example 1.4.2

Liquid octane at 25°C, 1 atm enters a well-insulated reactor and reacts with dry air entering at the same temperature and pressure. For steady-state operation and negligible effects of kinetic and potential energy, (1) determine the temperature of the combustion products for complete combustion with the theoretical amount of air, and (2) the rates of entropy generation and exergy destruction, each per kmol of fuel.

***Solution*:** For combustion of liquid octane with the theoretical amount of air, the chemical equation is

$$C_8H_{18}(\text{l}) + 12.5O_2 + 47N_2 \rightarrow 8CO_2 + 9H_2O(\text{g}) + 47N_2.$$

1. At steady state, the control volume energy rate balance reduces to read

$$0 = \frac{\dot{Q}_{cv}}{\dot{n}_F}^{0} - \frac{\dot{W}_{cv}}{\dot{n}_F}^{0} + \sum_{R} n_i \left(\overline{h}_f^{\circ} + \Delta\overline{h}\right)_i - \sum_{P} n_e \left(\overline{h}_f^{\circ} + \Delta\overline{h}\right)_e,$$

where R denotes reactants, P denotes products, and the symbols for enthalpy have the same significance as in Equation 1.4.1. Since the reactants enter at 25°C, the corresponding $(\Delta\overline{h})_i$ terms vanish, and the energy rate equation becomes

$$\sum_{P} n_e \left(\Delta\overline{h}\right)_e = \sum_{R} n_i \overline{h}_{fi}^{\circ} - \sum_{P} n_e \overline{h}_{fe}^{\circ}.$$

Introducing coefficients from the reaction equation, this takes the form

$$8\left(\Delta\overline{h}\right)_{CO_2} + 9\left(\Delta\overline{h}\right)_{H_2O(g)} + 47\left(\Delta\overline{h}\right)_{N_2} = \left[\left(\overline{h}_f^\circ\right)_{C_8H_{18}(l)} + 12.5\left(\overline{h}_f^\circ\right)_{O_2} + 47\left(\overline{h}_f^\circ\right)_{N_2}\right] - \left[8\left(\overline{h}_f^\circ\right)_{CO_2} + 9\left(\overline{h}_f^\circ\right)_{H_2O(g)} + 47\left(\overline{h}_f^\circ\right)_{N_2}\right].$$

Using data from Table 1.4.1 to evaluate the right side,

$$8\left(\Delta\overline{h}\right)_{CO_2} + 9\left(\Delta\overline{h}\right)_{H_2O(g)} + 47\left(\Delta\overline{h}\right)_{N_2} = 5{,}074{,}630 \text{ kJ/kmol (fuel)}.$$

Each $\Delta\overline{h}$ term on the left side of this equation depends on the temperature of the products, T_p, which can be solved for iteratively as $T_p = 2395$ K.

2. The entropy rate balance on a per-mole-of-fuel basis takes the form

$$0 = \sum_j \frac{\dot{Q}_j}{\dot{n}_F}\cancel{\frac{T_j}{T_j}}^{0} + \overline{s}_F + \left(12.5\overline{s}_{O_2} + 47\overline{s}_{N_2}\right) - \left(8\overline{s}_{CO_2} + 9\overline{s}_{H_2O(g)} + 47\overline{s}_{N_2}\right) + \frac{\dot{S}_{gen}}{\dot{n}_F}$$

or on rearrangement,

$$\frac{\dot{S}_{gen}}{\dot{n}_F} = \left(8\overline{s}_{CO_2} + 9\overline{s}_{H_2O(g)} + 47\overline{s}_{N_2}\right) - \overline{s}_F - \left(12.5\overline{s}_{O_2} + 47\overline{s}_{N_2}\right).$$

The absolute entropy of liquid octane from Table 1.4.1 is 360.79 kJ/kmol·K. The oxygen and nitrogen in the combustion air enter the reactor as components of an ideal gas mixture at T_{ref}, p_{ref}, With Equation 1.4.4, where $p = p_{ref}$, and absolute entropy data from Table 1.4.1,

$$\begin{aligned}\overline{s}_{O_2} &= \overline{s}_{O_2}^\circ\left(T_{ref}\right) - \overline{R}\ln y_{O_2}\\ &= 205.03 - 8.314\ln 0.21 = 218.01 \text{kJ/kmol}\cdot\text{K},\\ \overline{s}_{N_2} &= \overline{s}_{N_2}^\circ\left(T_{ref}\right) - \overline{R}\ln y_{N_2}\\ &= 191.5 - 8.314\ln 0.79 = 193.46 \text{kJ/kmol}\cdot\text{K}.\end{aligned}$$

The product gas exits as a gas mixture at 1 atm, 2395 K with the following composition: $y_{CO_2} = 8/64 = 0.125$, $y_{H_2O(g)} = 9/64 = 0.1406$, and $y_{N_2} = 47/64 = 0.7344$. With Equation 1.4.4, where $p = p_{ref}$, *and* absolute entropy data at 2395 K from Table A-23 of Moran et al. (2014),

$$\overline{s}_{CO_2} = 320.173 - 8.314\ln(0.125) = 337.46 \text{kJ/kmol}\cdot\text{K},$$

$$\overline{s}_{H_2O} = 273.986 - 8.314\ln(0.1406) = 290.30 \text{kJ/kmol}\cdot\text{K},$$

$$\overline{s}_{N_2} = 258.503 - 8.314\ln(0.7344) = 261.07 \text{kJ/kmol}\cdot\text{K}.$$

Inserting values, the rate of entropy generation is

$$\begin{aligned}\frac{\dot{S}_{gen}}{\dot{n}_F} &= 8(337.46) + 9(290.30) + 47(261.07) - 360.79 - 12.5(218.01) - 47(193.46)\\ &= 5404 \text{kJ/kmol}\cdot\text{K}.\end{aligned}$$

Using Equation 1.5.3 and assuming $T_0 = 298$ K, the rate of energy destruction is $\dot{E}_D/\dot{n}_F = 1.61\times10^6$ kJ/kmol.

Gibbs Function of Formation

Paralleling the approach used for enthalpy, a value of zero is assigned to the Gibbs function of each stable element at the standard state. The *Gibbs function of formation* of a compound equals the change in the Gibbs function for the reaction in which the compound is formed from its elements. Table 1.4.1 provides Gibbs function of formation data of various substances at 298 K and 1 atm.

The Gibbs function at a state other than the standard state is found by adding to the Gibbs function of formation the change in the specific Gibbs function $\Delta\overline{g}$ between the standard state and the state of interest:

$$\overline{g}(T,p) = \overline{g}_f^\circ + \left[\overline{g}(T,p) - \overline{g}\left(T_{\text{ref}}, p_{\text{ref}}\right)\right] = \overline{g}_f^\circ + \Delta\overline{g}, \tag{1.4.5a}$$

where

$$\Delta\overline{g} = \left[\overline{h}(T,p) - \overline{h}\left(T_{\text{ref}}, p_{\text{ref}}\right)\right] - \left[T\overline{s}(T,p) - T_{\text{ref}}\overline{s}\left(T_{\text{ref}}, p_{\text{ref}}\right)\right]. \tag{1.4.5b}$$

The Gibbs function of component i in an ideal gas mixture is evaluated at the partial pressure of component i and the mixture temperature.

As an application, the maximum theoretical work that can be developed, per mole of fuel consumed, is evaluated for the control volume of Figure 1.4.1, where the fuel and oxygen each enter in separate streams and carbon dioxide and water each exit separately. All entering and exiting streams are at the same temperature T and pressure p. The reaction is complete:

$$C_aH_b + \left(a + \frac{b}{4}\right)O_2 \rightarrow aCO_2 + \frac{b}{2}H_2O.$$

This control volume is similar to idealized devices such as a reversible fuel *cell* or a *van't Hoff equilibrium box*.

For steady-state operation, the energy rate balance reduces to give

$$\frac{\dot{W}_{cv}}{\dot{n}_F} = \frac{\dot{Q}_{cv}}{\dot{n}_F} + \overline{h}_F + \left(a + \frac{b}{4}\right)\overline{h}_{O_2} - a\overline{h}_{CO_2} - \frac{b}{2}\overline{h}_{H_2O},$$

where $\dot{n}_F$ denotes the molar flow rate of the fuel. Kinetic and potential energy effects are regarded as negligible. If heat transfer occurs only at the temperature T, an entropy balance for the control volume takes the form

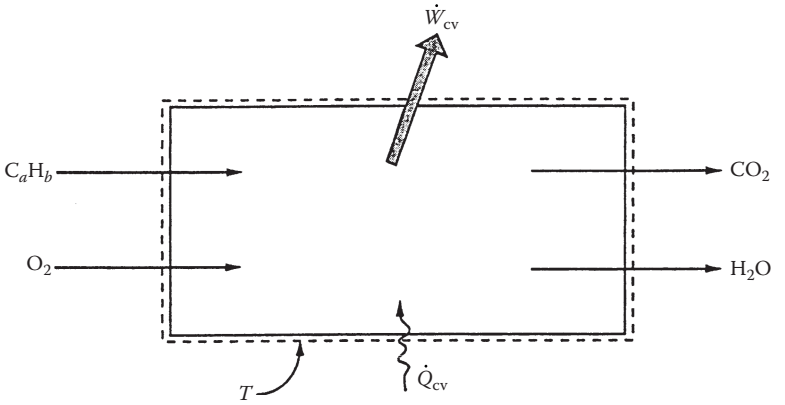


FIGURE 1.4.1 Device for evaluating maximum work.

$$0 = \frac{\dot{Q}_{cv}/\dot{n}_F}{T} + \bar{s}_F + \left(a + \frac{b}{4}\right)\bar{s}_{O_2} - a\bar{s}_{CO_2} - \frac{b}{2}\bar{s}_{H_2O} + \frac{\dot{S}_{gen}}{\dot{n}_F}.$$

Eliminating the heat transfer term from these expressions, an expression for the maximum theoretical value of the work developed per mole of fuel is obtained when the entropy generation term is set to zero:

$$\left(\frac{\dot{W}_{cv}}{\dot{n}_F}\right)_{\text{int rev}} = \left[\bar{h}_F + \left(a + \frac{b}{4}\right)\bar{h}_{O_2} - a\bar{h}_{CO_2} - \frac{b}{2}\bar{h}_{H_2O}\right](T,p) - T\left[\bar{s}_F + \left(a + \frac{b}{4}\right)\bar{s}_{O_2} - a\bar{s}_{CO_2} - \frac{b}{2}\bar{s}_{H_2O}\right](T,p).$$

This can be written alternatively in terms of the enthalpy of combustion as

$$\left(\frac{\dot{W}_{cv}}{\dot{n}_F}\right)_{\text{int rev}} = -\bar{h}_{RP}(T,p) - T\left[\bar{s}_F + \left(a + \frac{b}{4}\right)\bar{s}_{O_2} - a\bar{s}_{CO_2} - \frac{b}{2}\bar{s}_{H_2O}\right](T,p) \tag{1.4.6a}$$

or in terms of Gibbs functions as

$$\left(\frac{\dot{W}_{cv}}{\dot{n}_F}\right)_{\text{int rev}} = \left[\bar{g}_F + \left(a + \frac{b}{4}\right)\bar{g}_{O_2} - a\bar{g}_{CO_2} - \frac{b}{2}\bar{g}_{H_2O}\right](T,p). \tag{1.4.6b}$$

Equation 1.4.6b is used in the solution to Example 1.4.3.

Example 1.4.3

Hydrogen (H_2) and oxygen (O_2), each at 25°C, 1 atm, enter a fuel cell operating at steady state, and liquid water exits at the same temperature and pressure. The hydrogen flow rate is 2×10^{-4} kmol/s and the fuel cell operates isothermally at 25°C. Determine the maximum theoretical power the cell can develop, in kW.

***Solution*:** The overall cell reaction is $H_2 + 1/2\ O_2 \rightarrow H_2O(\ell)$, and Equations 1.4.6 are applicable. Selecting Equation 1.4.6b, and using Gibbs function data from Table 1.4.1,

$$\left(\frac{\dot{W}_{cv}}{\dot{n}_F}\right)_{\text{int rev}} = \left(\bar{g}_{H_2} + \frac{1}{2}\bar{g}_{O_2} - \bar{g}_{H_2O(l)}\right)(25°C, 1\text{atm})$$

$$= 0 + \frac{1}{2}(0) - (-237{,}180) = 237{,}180\ \text{kJ/kmol}.$$

Then,

$$\left(\dot{W}_{cv}\right)_{\text{int rev}} = \left(237{,}180\ \text{kJ/kmol}\right)\left(2\times10^{-4}\ \text{kmol/s}\right)\left(\frac{\text{kW}}{1\ \text{kJ/s}}\right) = 47.4\ \text{kW}.$$

Reaction Equilibrium

Let the objective be to determine the equilibrium composition of a system consisting of five gases A, B, C, D, and E, at a temperature T and pressure p, subject to a chemical reaction of the form

$$\nu_A A + \nu_B B \leftrightarrow \nu_C C + \nu_D D,$$

where the v's are stoichiometric coefficients. Component E is assumed to be inert and thus does not appear in the reaction equation. The equation suggests that at equilibrium the tendency of A and B to form C and D is just balanced by the tendency of C and D to form A and B.

At equilibrium, the temperature and pressure would be uniform throughout the system. Additionally, the *equation of reaction equilibrium* must be satisfied:

$$v_A\mu_A + v_B\mu_B = v_C\mu_C + v_D\mu_D, \tag{1.4.7a}$$

where the μ's are the chemical potentials (Section 1.3, Multicomponent Systems) of A, B, C, and D in the equilibrium mixture. In principle, the composition that would be present at equilibrium for a given temperature and pressure can be determined by solving this equation.

For ideal gas mixtures, the solution procedure is simplified by using the *equilibrium constant* $K(T)$ and the following equation:

$$\begin{aligned} K(T) &= \frac{y_C^{v_C} y_D^{v_D}}{y_A^{v_A} y_B^{v_B}} \left(\frac{p}{p_{ref}}\right)^{v_C+v_D-v_A-v_B} \\ &= \frac{n_C^{v_C} n_D^{v_D}}{n_A^{v_A} n_B^{v_B}} \left(\frac{p/p_{ref}}{n}\right)^{v_C+v_D-v_A-v_B}, \end{aligned} \tag{1.4.7b}$$

where y_A, y_B, y_C, and y_D denote the mole fractions of A, B, C, and D in the equilibrium mixture and $n=n_A+n_B+n_C+n_D+n_E$, where the n's denote the molar amounts of the gases in the mixture. Tabulations of $K(T)$ for each of several reactions of the form Equation 1.4.7a are provided in Table 1.4.3. An application of Equation 1.4.7b is provided in Example 1.4.4.

Example 1.4.4

One kmol of CO reacts with the theoretical amount of dry air to form an equilibrium mixture of CO_2, CO, O_2, and N_2 at 2500 K, 1 atm. Determine the amount of CO in the equilibrium mixture, in kmol.

***Solution*:** The reaction of CO with the theoretical amount of dry air to form CO_2, CO, O_2, and N_2 is

$$CO + \frac{1}{2}O_2 + 1.88N_2 \rightarrow zCO + \frac{z}{2}O_2 + (1-z)CO_2 + 1.88N_2,$$

where z is the amount of CO, in kmol, present in the equilibrium mixture. The total number of moles n is

$$n = z + \frac{z}{2} + (1-z) + 1.88 = \frac{5.76+z}{2}.$$

At equilibrium $CO_2 \leftrightarrow CO + 1/2O_2$; and Equation 1.4.7b takes the form

$$K = \frac{z\left(z/2\right)^{1/2}}{1-z}\left[\frac{p/p_{ref}}{(5.76+z)/2}\right]^{1/2},$$

where $p/p_{ref}=1$. At 2500 K, Table 1.4.3 gives $K=0.0363$. Solving iteratively, $z=0.175$.

TABLE 1.4.3
Logarithms to the Base 10 of the Equilibrium Constant *K*

	log10 K								
Temp (K)	**$H_2 \Leftrightarrow 2H$**	**$O_2 \Leftrightarrow 2O$**	**$N_2 \Leftrightarrow 2N$**	**$\frac{1}{2}O_2+\frac{1}{2}N_2 \Leftrightarrow NO$**	**$H_2O \Leftrightarrow H_2+\frac{1}{2}O_2$**	**$H_2O \Leftrightarrow OH+\frac{1}{2}H_2$**	**$CO_2 \Leftrightarrow CO+\frac{1}{2}O_2$**	**$CO+H_2 \Leftrightarrow CO+H_2O$**	**Temp (°R)**
298	–71.224	–81.208	–159.600	–15.171	–40.048	–46.054	–45.066	–5.018	537
500	–40.316	–45.880	–92.672	–8.783	–22.886	–26.130	–25.025	–2.139	900
1000	–17.292	–19.614	–43.056	–4.062	–10.062	–11.280	–10.221	–0.159	1800
1200	–13.414	–15.208	–34.754	–3.275	–7.899	–8.811	–7.764	+0.135	2160
1400	–10.630	–12.054	–28.812	–2.712	–6.347	–7.021	–6.014	+0.333	2520
1600	–8.532	–9.684	–24.350	–2.290	–5.180	–5.677	–4.706	+0.474	2880
1700	–7.666	–8.706	–22.512	–2.116	–4.699	–5.124	–4.169	+0.530	3060
1800	–6.896	–7.836	–20.874	–1.962	–4.270	–4.613	–3.693	+0.577	3240
1900	–6.204	–7.058	–19.410	–1.823	–3.886	–4.190	–3.267	+0.619	3420
2000	–5.580	–6.356	–18.092	–1.699	–3.540	–3.776	–2.884	+0.656	3600
2100	–5.016	–5.720	–16.898	–1.586	–3.227	–3.434	–2.539	+0.688	3780
2200	–4.502	–5.142	–15.810	–1.484	–2.942	–3.091	–2.226	+0.716	3960
2300	–4.032	–4.614	–14.818	–1.391	–2.682	–2.809	–1.940	+0.742	4140
2400	–3.600	–4.130	–13.908	–1.305	–2.443	–2.520	–1.679	+0.764	4320
2500	–3.202	–3.684	–13.070	–1.227	–2.224	–2.270	–1.440	+0.784	4500
2600	–2.836	–3.272	–12.298	–1.154	–2.021	–2.038	–1.219	+0.802	4680
2700	–2.494	–2.892	–11.580	–1.087	–1.833	–1.823	–1.015	+0.818	4860
2800	–2.178	–2.536	–10.914	–1.025	–1.658	–1.624	–0.825	+0.833	5040
2900	–1.882	–2.206	–10.294	–0.967	–1.495	–1.438	–0.649	+0.846	5220
3000	–1.606	–1.898	–9.716	–0.913	–1.343	–1.265	–0.485	+0.858	5400
3100	–1.348	–1.610	–9.174	–0.863	–1.201	–1.103	–0.332	+0.869	5580
3200	–1.106	–1.340	–8.664	–0.815	–1.067	–0.951	–0.189	+0.878	5760
3300	–0.878	–1.086	–8.186	–0.771	–0.942	–0.809	–0.054	+0.888	5940
3400	–0.664	–0.846	–7.736	–0.729	–0.824	–0.674	+0.071	+0.895	6120
3500	–0.462	–0.620	–7.312	–0.690	–0.712	–0.547	+0.190	+0.902	6300

Source: Based on data from the JANAF Thermochemical Tables, NSRDS-NBS-37, 1971.

1.5 EXERGY ANALYSIS

Michael J. Moran

The method of *exergy analysis* presented in this section enables the location, cause, and true magnitude of energy resource waste and loss to be determined. Such information can be used in the design of new energy-efficient systems and for improving the performance of existing systems. Exergy analysis also provides insights that elude a purely first-law approach. For example, on the basis of first-law reasoning alone, the condenser of a power plant may be mistakenly identified as the component primarily responsible for the plant's seemingly low overall performance. An exergy analysis correctly reveals not only that the condenser loss is relatively unimportant (see the last two rows of the Rankine cycle values of Table 1.6.2), but also that the steam generator is the principal site of thermodynamic inefficiency owing to combustion and heat-transfer irreversibilities within it.

When exergy concepts are combined with principles of engineering economy, the result is known as *thermoeconomics* or *exergoeconomics*. Thermoeconomics allows the real cost sources at the component level to be identified: capital investment costs, operating and maintenance costs, and the costs associated with the destruction and loss of exergy. Optimization of thermal systems can be achieved by a careful consideration of such cost sources. From this perspective, thermoeconomics is *exergy-aided cost minimization*.

Discussions of exergy analysis and thermoeconomics are provided by Bejan et al. (1996), Moran (1989), and Moran et al. (2014). The present section provides an introduction.

Defining Exergy

An opportunity for doing work exists whenever two systems at different states are placed in communication because, in principle, work can be developed as the two are allowed to come into equilibrium. When one of the two systems is a suitably idealized system called an *environment* and the other is some system of interest, *exergy* is the maximum theoretical useful work (shaft work or electrical work) obtainable as the systems interact to equilibrium, heat transfer occurring with the environment only. (Alternatively, exergy is the minimum theoretical useful work required to form a quantity of matter from substances present in the environment and to bring the matter to a specified state.) Exergy is a measure of the *departure* of the state of the system from that of the environment, and is therefore an attribute of the system and environment together. Once the environment is specified, however, a value can be assigned to exergy in terms of property values for the system only, so exergy can be regarded as an extensive property of the system.

Exergy can be destroyed and generally is not conserved. A limiting case is when exergy would be completely destroyed, as would occur if a system were to come into equilibrium with the environment *spontaneously* with no provision to obtain work. The capability to develop work that existed initially would be completely wasted in the spontaneous process. Moreover, since no work needs to be done to effect such a spontaneous change, the value of exergy can never be negative.

Environment

Models with various levels of specificity are employed for describing the environment used to evaluate exergy. Models of the environment typically refer to some portion of a system's surroundings, the intensive properties of each phase of which are uniform and do not change significantly as a result of any process under consideration. The environment is regarded as composed of common substances existing in abundance within the Earth's atmosphere, oceans, and crust. The substances are in their stable forms as they exist naturally, and there is no possibility of developing work from interactions—physical or chemical—between parts of the environment. Although the intensive properties of the environment are assumed to be unchanging, the extensive properties can change as a result of interactions with other systems. Kinetic and potential energies are

evaluated relative to coordinates in the environment, all parts of which are considered to be at rest with respect to one another.

For computational ease, the temperature T_0 and pressure p_0 of the environment are often taken as standard-state values, such as 1 atm and 25°C (77°F). However, these properties may be specified differently depending on the application. T_0 and p_0 might be taken as the average ambient temperature and pressure, respectively, for the location at which the system under consideration operates. Or, if the system uses atmospheric air, T_0 might be specified as the average air temperature. If both air and water from the natural surroundings are used, T_0 would be specified as the lower of the average temperatures for air and water.

Dead States

When a system is in equilibrium with the environment, the state of the system is called the *dead state*. At the dead state, the conditions of mechanical, thermal, and chemical equilibrium between the system and the environment are satisfied: the pressure, temperature, and chemical potentials of the system equal those of the environment, respectively. In addition, the system has no motion or elevation relative to coordinates in the environment. Under these conditions, there is no possibility of a spontaneous change within the system or the environment, nor can there be an interaction between them. The value of exergy is zero.

Another type of equilibrium between the system and environment can be identified. This is a restricted form of equilibrium where only the conditions of mechanical and thermal equilibrium must be satisfied. This state of the system is called the *restricted dead state*. At the restricted dead state, the fixed quantity of matter under consideration is imagined to be sealed in an envelope impervious to mass flow, at zero velocity and elevation relative to coordinates in the environment, and at the temperature T_0 and pressure p_0.

Exergy Balances

Exergy can be transferred by three means: exergy transfer associated with work, exergy transfer associated with heat transfer, and exergy transfer associated with the matter entering and exiting a control volume. All such exergy transfers are evaluated relative to the environment used to define exergy. Exergy is also destroyed by irreversibilities within the system or control volume.

Exergy balances can be written in various forms, depending on whether a closed system or control volume is under consideration and whether steady-state or transient operation is of interest. Owing to its importance for a wide range of applications, an exergy rate balance for control volumes at steady state is presented next.

Control Volume Exergy Rate Balance

At steady state, the control volume exergy rate balance takes the form

$$0 = \underbrace{\sum_j \dot{E}_{q\cdot j} - \dot{W}_{\mathrm{cv}} + \sum_i \dot{E}_i - \sum_e \dot{E}_e}_{\text{rates of exergy transfer}} - \underbrace{\dot{E}_{\mathrm{D}}}_{\text{rates of exergy destruction}} \tag{1.5.1a}$$

or

$$0 = \sum_j \left(1 - \frac{T_0}{T_j}\right)\dot{Q}_j - \dot{W}_{\mathrm{cv}} + \sum_i \dot{m}_i e_i - \sum_e \dot{m}_e e_e - \dot{E}_{\mathrm{D}}\,. \tag{1.5.1b}$$

$\dot{W}_{cv}$ has the same significance as in Equation 1.2.4: the work rate excluding the flow work. $\dot{Q}_j$ is the time rate of heat transfer at the location on the boundary of the control volume where the instantaneous temperature is T_j. The associated rate of exergy transfer is

$$\dot{E}_{q,j} = \left(1 - \frac{T_0}{T_j}\right)\dot{Q}_j. \tag{1.5.2}$$

As for other control volume rate balances, the subscripts i and e denote inlets and outlets, respectively. The exergy transfer rates at control volume inlets and outlets are denoted, respectively, as $\dot{E}_i = \dot{m}_i e_i$ and $\dot{E}_e = \dot{m}_e e_e$. Finally, $\dot{E}_D$ accounts for the time rate of exergy destruction due to irreversibilities within the control volume. The exergy destruction rate is related to the entropy generation rate by

$$\dot{E}_D = T_0 S_{gen}. \tag{1.5.3}$$

The specific exergy transfer terms e_i and e_e are expressible in terms of four components: physical exergy e^{PH}, kinetic exergy e^{KN}, potential exergy e^{PT}, and chemical exergy e^{CH}:

$$e = e^{PH} + e^{KN} + e^{PT} + e^{CH}. \tag{1.5.4}$$

The first three components are evaluated as follows:

$$e^{PH} = (h - h_0) - T_0(s - s_0), \tag{1.5.5a}$$

$$e^{KN} = \frac{1}{2}v^2, \tag{1.5.5b}$$

$$e^{PT} = gz. \tag{1.5.5c}$$

In Equation 1.5.5a, h_0 and s_0 denote, respectively, the specific enthalpy and specific entropy at the restricted dead state. In Equations 1.5.5b and 1.5.5c, v and z denote velocity and elevation relative to coordinates in the environment, respectively. The chemical exergy e^{CH} is considered next.

Chemical Exergy

To evaluate the chemical exergy, the exergy component associated with the departure of the chemical composition of a system from that of the environment, the substances comprising the system are referred to the properties of a suitably selected set of environmental substances. For this purpose, alternative models of the environment have been developed. For discussion, see, for example, Moran (1989) and Kotas (1995).

Exergy analysis is facilitated, however, by employing a *standard environment* and a corresponding table of *standard chemical exergies*. Standard chemical exergies are based on standard values of the environmental temperature T_0 and pressure p_0—for example, 298.15 K (25°C) and 1 atm, respectively. A standard environment is also regarded as consisting of a set of reference substances with standard concentrations reflecting as closely as possible the chemical makeup of the natural environment. The reference substances generally fall into three groups: gaseous components of the atmosphere, solid substances from the lithosphere, and ionic and nonionic substances from the oceans. The chemical exergy data of Table 1.5.1 correspond to two proposed standard exergy reference environments, called here model I and model II.

TABLE 1.5.1
Standard Molar Chemical Exergy, e^{CH} (kJ/kmol), of Various Substances at 298 K and p_0

Substance	Formula	Model I[a]	Model II[b]
Nitrogen	$N_2(g)$	640	720
Oxygen	$O_2(g)$	3,950	3,970
Carbon dioxide	$CO_2(g)$	14,175	19,870
Water	$H_2O(g)$	8,635	9,500
	$H_2O(l)$	45	900
Carbon (graphite)	C(s)	404,590	410,260
Hydrogen	$H_2(g)$	235,250	236,100
Sulfur	S(s)	598,160	609,600
Carbon monoxide	CO(g)	269,410	275,100
Sulfur dioxide	$SO_2(g)$	301,940	313,400
Nitrogen monoxide	NO(g)	88,850	88,900
Nitrogen dioxide	$NO_2(g)$	55,565	55,600
Hydrogen sulfide	$H_2S(g)$	799,890	812,000
Ammonia	$NH_3(g)$	336,685	337,900
Methane	$CH_4(g)$	824,350	831,650
Ethane	$C_2H_6(g)$	1,482,035	1,495,840
Methanol	$CH_3OH(g)$	715,070	722,300
	$CH_3OH(l)$	710,745	718,000
Ethyl alcohol	$C_2H_5OH(g)$	1,348,330	1,363,900
	$C_2H_5OH(l)$	1,342,085	1,357,700

[a] *Ahrendts, J. Die Exergie Chemisch Reaktionsfahiger Systeme, VDI-Forschungsheft. VDI-Verlag, Dusseldorf, 579, 1977.* Also see *Reference States, Energy—The International Journal*, *5:* 667–677, 1980. In Model I, $p_0 = 1.019$ atm. This model attempts to impose a criterion that the reference environment be in equilibrium. The reference substances are determined assuming restricted chemical equilibrium for nitric acid and nitrates and unrestricted thermodynamic equilibrium for all other chemical components of the atmosphere, the oceans, and a portion of the Earth's crust. The chemical composition of the gas phase of this model approximates the composition of the natural atmosphere.

[b] Szargut, J. et al., *Energy Analysis of Thermal, Chemical, and Metallurgical Processes*, Hemisphere, New York, 1988. In Model II, $p_0 = 1.0$ atm. In developing this model, a reference substance is selected for each chemical element from among substances that contain the element being considered and that are abundantly present in the natural environment, even though the substances are not in completely mutual stable equilibrium. An underlying rationale for this approach is that substances found abundantly in nature have little economic value. On an overall basis, the chemical composition of the exergy reference environment of Model II is closer than Model I to the composition of the natural environment, but the equilibrium criterion is not always satisfied.

Although the use of standard chemical exergies greatly facilitates the application of exergy principles, the term *standard* is somewhat misleading since there is no one specification of the environment that suffices for all applications. Still, chemical exergies calculated relative to alternative specifications of the environment are generally in good agreement. For a broad range of engineering applications, the simplicity and ease of use of standard chemical exergies generally outweigh any slight lack of accuracy that might result. In particular, the effect of slight variations in the values of T_0 and p_0 about the values used to determine the standard chemical exergies reported in Table 1.5.1 can be neglected.

The literature of exergy analysis provides several expressions allowing the chemical exergy to be evaluated in particular cases of interest. The molar chemical exergy of a gas mixture, for example, can be evaluated from

$$\overline{e}^{\mathrm{CH}} = \sum_{i=1}^{f} y_i \overline{e}_i^{\mathrm{CH}} + \overline{R}T_0 \sum_{i=1}^{f} y_i \ln y_i, \tag{1.5.6}$$

where $\overline{e}_i^{\mathrm{CH}}$ is the molar chemical exergy of the ith component.

Example 1.5.1

Ignoring the kinetic and potential exergies, determine the exergy rate, in kJ/kg, associated with each of the following streams of matter:

1. Saturated water vapor at 20 bar.
2. Methane at 5 bar, 25°C.

Let T_0=298 K, p_0=1.013 bar (1 atm).

Solution: Equation 1.5.4 reduces to read

$$e = \left(h - h_0\right) - T_0\left(s - s_0\right) + e^{\mathrm{CH}}$$

1. From Table A.5, h=2799.5 kJ/kg, s=6.3409 kJ/kg·K. At T_0=298 K (25°C), water would be a liquid; thus with Equations 1.3.20c and d, $h_0 \approx 104.9$ kJ/kg, $s_0 \approx 0.3674$ kJ/kg·K. Table 1.5.1 (model I) gives e^{CH}=45/18.02=2.5 kJ/kg. Then

$$\begin{aligned} e &= (2799.5 - 104.9) - 298(6.3409 - 0.3674) + 2.5 \\ &= 914.5 + 2.5 = 917.0 \text{ kJ/kg.} \end{aligned}$$

   Here the specific exergy is determined predominately by the physical component.
2. Assuming the ideal gas model for methane, $h - h_0 = 0$. Also, Equation 1.3.28 reduces to give $s - s_0 = -R \ln p/p_0$. Then, Equation 1.5.4 reads

$$e = RT_0 \ln p / p_0 + e^{\mathrm{CH}}.$$

   With e^{CH}=824,350/16.04=51,393.4 kJ/kg from Table 1.5.1 (model I),

$$\begin{aligned} e &= \left(\frac{8.314}{16.04}\text{kJ/kg}\cdot\text{K}\right)(298\text{K})\ln\frac{5}{1.013} + 51{,}393.4\text{kJ/kg} \\ &= 246.6 + 51{,}393.4 \\ &= 51{,}640\text{kJ/kg.} \end{aligned}$$

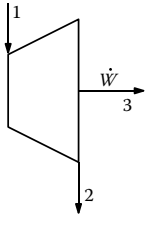
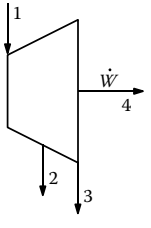
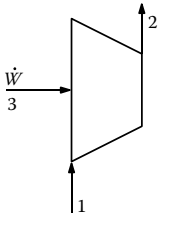
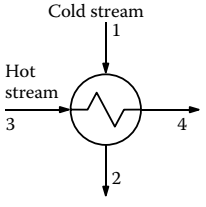
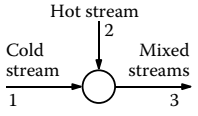
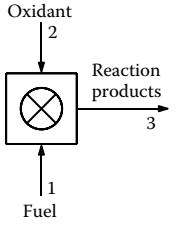
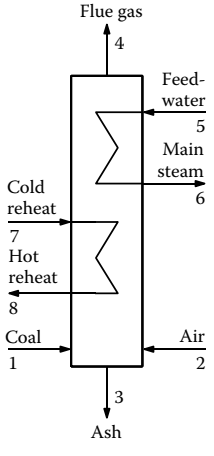
Here, the specific exergy is determined predominately by the chemical component.

The small difference between p_0=1.013 bar and the value of p_0 for model I has been ignored.

Exergetic Efficiency

The exergetic efficiency (second law efficiency, effectiveness, or rational efficiency) provides a true measure of the performance of a system from the thermodynamic viewpoint. To define the exergetic efficiency, both a *product* and a *fuel* for the system being analyzed are identified. The product represents the desired result of the system (power, steam, some combination of power and steam, etc.). Accordingly, the definition of the product must be consistent with the purpose of purchasing and using the system. The fuel represents the resources expended to generate the product and is not

TABLE 1.5.2
The Exergetic Efficiency for Selected Components at Steady State

Component	Turbine or Expander	Extraction Turbine	Compressor, Pump, or Fan	Heat Exchanger[a]	Mixing Unit[b]	Gasifier or Combustion Chamber	Boiler
							
$\dot{E}_P$	$\dot{W}$	$\dot{W}$	$\dot{E}_2 - \dot{E}_1$	$\dot{E}_2 - \dot{E}_1$	$\dot{m}_1(e_3 - e_1)$	$\dot{E}_3 - \dot{E}_2$	$(\dot{E}_6 - \dot{E}_5) + (\dot{E}_8 - \dot{E}_7)$
$\dot{E}_F$	$\dot{E}_1 - \dot{E}_2$	$\dot{E}_1 - \dot{E}_2 - \dot{E}_3$	$\dot{W}$	$\dot{E}_3 - \dot{E}_4$	$\dot{m}_2(e_2 - e_3)$	$\dot{E}_1$	$(\dot{E}_1 + \dot{E}_2) + (\dot{E}_3 + \dot{E}_4)$
ε	$\dfrac{\dot{W}}{\dot{E}_1 - \dot{E}_2}$	$\dfrac{\dot{W}}{\dot{E}_1 - \dot{E}_2 - \dot{E}_3}$	$\dfrac{\dot{E}_2 - \dot{E}_1}{\dot{W}}$	$\dfrac{\dot{E}_2 - \dot{E}_1}{\dot{E}_3 - \dot{E}_4}$	$\dfrac{\dot{m}_1(e_3 - e_1)}{\dot{m}_2(e_2 - e_3)}$	$\dfrac{\dot{E}_3 - \dot{E}_2}{\dot{E}_1}$	$\dfrac{(\dot{E}_6 - \dot{E}_5) + (\dot{E}_8 - \dot{E}_7)}{(\dot{E}_1 - \dot{E}_2) - (\dot{E}_3 - \dot{E}_4)}$

Note: For discussion, see Bejan et al. (1996).

[a] This definition assumes that the purpose of the heat exchanger is to heat the cold stream ($T_1 \geq T_0$). If the purpose of the heat exchanger is to provide cooling ($T_3 \geq T_0$), then the following relations should be used: $\dot{E}_P = \dot{E}_4 - \dot{E}_3$ and $\dot{E}_F = \dot{E}_1 - \dot{E}_2$.

[b] The purpose of the mixing unit is to heat the cold stream; otherwise a mixing unit may be considered a dissipative component, i.e. a component for which no product can be defined.

necessarily restricted to being an actual fuel such as a natural gas, oil, or coal. Both the product and the fuel are expressed in terms of exergy.

For a control volume at steady state whose exergy rate balance reads

$$\dot{E}_F = \dot{E}_P + \dot{E}_D + \dot{E}_L .$$

The exergetic efficiency is

$$\varepsilon = \frac{\dot{E}_P}{\dot{E}_F} = 1 - \frac{\dot{E}_D + \dot{E}_L}{\dot{E}_F}, \tag{1.5.7}$$

where the rates at which the fuel is supplied and the product is generated are $\dot{E}_F$ and $\dot{E}_P$, respectively. $\dot{E}_D$ and $\dot{E}_L$ denote the rates of exergy destruction and exergy loss, respectively. Exergy is destroyed by irreversibilities within the control volume, and exergy is lost from the control volume via stray heat transfer, material streams vented to the surroundings, and so on. The exergetic efficiency shows the percentage of the fuel exergy provided to a control volume that is found in the product exergy. Moreover, the difference between 100% and the value of the exergetic efficiency, expressed as a percent, is the percentage of the fuel exergy wasted in this control volume as exergy destruction and exergy loss.

To apply Equation 1.5.7, decisions are required concerning what are considered as the fuel and the product. Table 1.5.2 provides illustrations for several common components. Similar considerations are used to write exergetic efficiencies for systems consisting of several such components, as, for example, a power plant.

Exergetic efficiencies can be used to assess the thermodynamic performance of a component, plant, or industry relative to the performance of *similar* components, plants, or industries. By this means the performance of a gas turbine, for instance, can be gauged relative to the typical present-day performance level of gas turbines. A comparison of exergetic efficiencies for *dissimilar* devices—gas turbines and heat exchangers, for example—is generally not significant, however.

The exergetic efficiency is generally more meaningful, objective, and useful than other efficiencies based on the first or second law of thermodynamics, including the thermal efficiency of a power plant, the isentropic efficiency of a compressor or turbine, and the effectiveness of a heat exchanger. The thermal efficiency of a cogeneration system, for instance, is misleading because it treats both work and heat transfer as having equal thermodynamic value. The isentropic turbine efficiency (Equation 1.6.2a) does not consider that the working fluid at the outlet of the turbine has a higher temperature (and consequently a higher exergy that may be used in the next component) in the actual process than in the isentropic process. The heat exchanger effectiveness fails, for example, to identify the exergy destruction associated with the pressure drops of the heat exchanger working fluids.

Example 1.5.2

Evaluate the exergetic efficiency of the turbine in part (1) of Example 1.2.1 for $T_0 = 298$ K.

Solution: The exergetic efficiency from Table 1.5.2 is

$$\varepsilon = \frac{\dot{W}}{\dot{E}_1 - \dot{E}_2} = \frac{\dot{W}}{\dot{m}(e_1 - e_2)}.$$

Using Equations 1.5.4 and 1.5.5a, and noting that the chemical exergy at 1 and 2 cancels,

$$\varepsilon = \frac{\dot{W}}{\dot{m}\left[\left(h_1 - h_2\right) - T_0\left(s_1 - s_2\right)\right]}.$$

Since $\dot{W} = \dot{m}(h_1 - h_2)$,

$$\varepsilon = \frac{\dot{W}}{\dot{W} + \dot{m}T_0(s_2 - s_1)}.$$

Finally, using data from Example 1.2.1 and $s_2 = 6.8473$ kJ/kg·K,

$$\varepsilon = \frac{30\,\text{MW}}{30\,\text{MW} + \left(\frac{162{,}357}{3600}\text{kg/s}\right)(298\text{K})(6.8473 - 6.6022)\left(\text{kJ/kg}\cdot\text{K}\right)\left(\frac{1\,\text{MW}}{10^3\,\text{kJ/s}}\right)}$$

$$= \frac{30\,\text{MW}}{(30 + 3.29)\text{MW}} = 0.9(90\%).$$

Introduction to Exergy Costing

Since exergy measures the true thermodynamic values of the work, heat, and other interactions between the system and its surroundings as well as the effect of irreversibilities within the system, exergy is a rational basis for assigning costs. This aspect of thermoeconomics is called *exergy costing*. An introduction to exergy costing is given in the present discussion.

Referring to Figure 1.5.1 showing a steam turbine-electric generator at steady state, the total cost to produce the electricity and exiting steam equals the cost of the entering steam plus the cost of owning and operating the device. This is expressed by the *cost rate balance* for the turbine generator:

$$\dot{C}_e + \dot{C}_2 = \dot{C}_1 + \dot{Z}, \tag{1.5.8a}$$

where $\dot{C}_e$ is the cost rate associated with the electricity, $\dot{C}_1$ and $\dot{C}_2$ are the cost rates associated with the entering steam and exiting steam, respectively, and $\dot{Z}$ accounts for the cost rate associated with owning and operating the system, each *annualized* in $ per year.

With exergy costing, the cost rates $\dot{C}_1$, $\dot{C}_2$, and $\dot{C}_e$ are evaluated in terms of the associated rate of exergy transfer and a *unit cost*. Equation 1.5.8a then appears as

$$c_e\dot{W}_e + c_2\dot{E}_2 = c_1\dot{E}_1 + \dot{Z}. \tag{1.5.8b}$$

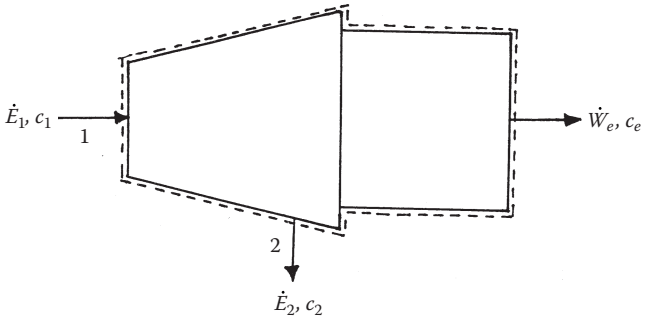


FIGURE 1.5.1 Steam turbine/electric generator used to discuss exergy costing.

The coefficients c_1, c_2, and c_e in Equation 1.5.8b denote the *average* cost per unit of exergy for the associated exergy rate. The unit cost c_1 of the entering steam would be obtained from exergy costing applied to the components upstream of the turbine. Assigning the same unit cost to the exiting steam: $c_2 = c_1$ on the basis that the purpose of the turbine generator is to generate electricity and thus all costs associated with owning and operating the system should be charged to the power, Equation 1.5.8b becomes

$$c_e \dot{W}_e = c_1(\dot{E}_1 - \dot{E}_2) + \dot{Z}. \tag{1.5.8c}$$

The first term on the right side accounts for the cost of the net exergy used, and the second term accounts for cost of the system itself. Introducing the exergetic efficiency from Table 1.5.2, the unit cost of the electricity is

$$c_e = \frac{c_1}{\varepsilon} + \frac{\dot{Z}}{\dot{W}_e}. \tag{1.5.9}$$

This equation shows, for example, that the unit cost of electricity would increase if the exergetic efficiency were to decrease owing to a deterioration of the turbine with use.

Example 1.5.3

A turbine generator with an exergetic efficiency of 90% develops 7×10^7 kW·h of electricity annually. The annual cost of owning and operating the system is $\$2.5 \times 10^5$. If the average unit cost of the steam entering the system is \$0.0165 $kW^{-1} \cdot h^{-1}$ of exergy, evaluate the unit cost of the electricity.

Solution: Substituting values into Equation 1.5.9,

$$c_e = \frac{\$0.0165/\text{kW} \cdot \text{h}}{0.9} + \frac{\$2.5 \times 10^5/\text{year}}{7 \times 10^7\,\text{kW} \cdot \text{h/year}}$$

$$= 0.0183 + 0.0036 = \$0.0219/\text{kW} \cdot \text{h}.$$

1.6 VAPOR AND GAS POWER CYCLES

Michael J. Moran

Vapor and gas power systems develop electrical or mechanical power from energy sources of chemical, solar, or nuclear origin. In *vapor* power systems, the *working fluid*, normally water, undergoes a phase change from liquid to vapor, and conversely. In *gas* power systems, the working fluid remains a gas throughout, although the composition normally varies owing to the introduction of a fuel and subsequent combustion. The present section introduces vapor and gas power systems.

The processes taking place in power systems are sufficiently complicated that idealizations are typically employed to develop tractable thermodynamic models. The *air-standard analysis* of gas power systems considered later in the present section is a noteworthy example. Depending on the degree of idealization, such models may provide only qualitative information about the performance of the corresponding real-world systems. Yet such information is frequently useful in gauging how changes in major operating parameters might affect actual performance. Elementary thermodynamic models can also provide simple settings to assess, at least approximately, the advantages and disadvantages of features proposed to improve thermodynamic performance.

Rankine and Brayton Cycles

In their simplest embodiments, vapor power and gas turbine power plants are represented conventionally in terms of four components in series, forming, respectively, the Rankine cycle and the Brayton cycle shown schematically in Table 1.6.1. The thermodynamically ideal counterparts of these cycles are composed of four internally reversible processes in series: two isentropic processes alternated with two constant pressure processes. Table 1.6.1 provides property diagrams of the actual and corresponding ideal cycles. Each actual cycle is denoted 1–2–3–4–1; the ideal cycle is 1–2*s*–3–4*s*–1. For simplicity, pressure drops through the boiler, condenser, and heat exchangers are not shown. Invoking Equation 1.2.11 for the ideal cycles, the heat added per unit of mass flowing is represented by the area *under* the isobar from state 2*s* to state 3: area *a*–2*s*–3–*b*–*a*. The heat rejected is the area *under* the isobar from state 4*s* to state 1: area *a*–1–4*s*–*b*–*a*. Enclosed area 1–2*s*–3–4*s*–1 represents the net heat added per unit of mass flowing. For any power cycle, the net heat added equals the net work done.

Expressions for the principal energy transfers are provided by Equations 1 through 4 shown on the schematics of Table 1.6.1. They are obtained by reducing Equation 1.2.9a with the assumptions of negligible heat loss and negligible changes in kinetic and potential energy from the inlet to the outlet of each component. All quantities are positive in the directions of the arrows on the figure. Using these expressions, the thermal efficiency is

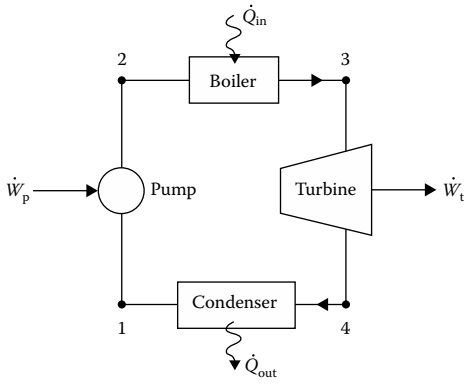
$$\eta = \frac{(h_3 - h_4) - (h_2 - h_1)}{h_3 - h_2}. \tag{1.6.1}$$

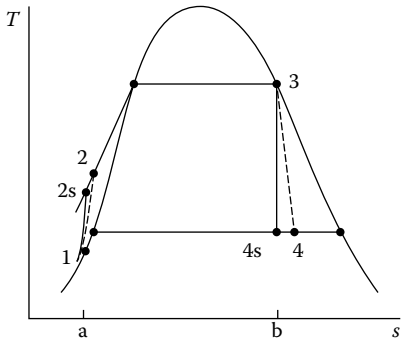
To obtain the thermal efficiency of the ideal cycle, h_{2_s} replaces h_2 and h_{4_s} replaces h_4 in Equation 1.6.1.

Decisions concerning cycle operating conditions normally recognize that the thermal efficiency tends to increase as the average temperature of heat addition increases and/or the temperature of heat rejection decreases. In the Rankine cycle, a high average temperature of heat addition can be achieved by superheating the vapor prior to entering the turbine, and/or by operating at an elevated steam generator pressure. In the Brayton cycle, an increase in the compressor pressure ratio p_2/p_1 tends to increase the average temperature of heat addition. Owing to materials limitations at elevated temperatures and pressures, the state of the working fluid at the turbine inlet must observe practical limits, however. The turbine inlet temperature of the Brayton cycle, for example, is controlled by providing air far in excess of what is required for combustion. In a Rankine cycle using water as the

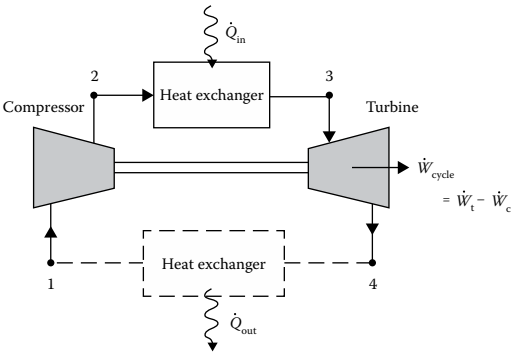
TABLE 1.6.1
Rankine and Brayton Cycles

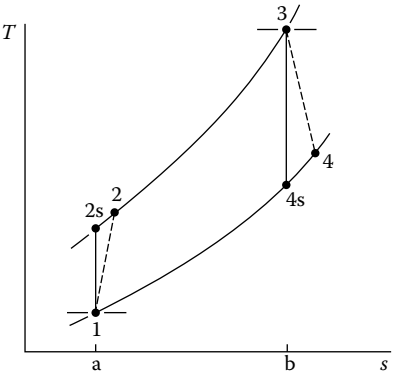
Rankine Cycle





Brayton Cycle





$$\left.\begin{matrix}\dot{W}_p \\ \dot{W}_c\end{matrix}\right\} = \dot{m}(h_2 - h_1) \quad (>0) \tag{1}$$

$$\dot{Q}_{in} = \dot{m}(h_3 - h_2) \quad (>0) \tag{2}$$

$$\dot{W}_t = \dot{m}(h_3 - h_4) \quad (>0) \tag{3}$$

$$\dot{Q}_{out} = \dot{m}(h_1 - h_4) \quad (>0) \tag{4}$$

working fluid, a low temperature of heat rejection is typically achieved by operating the condenser at a pressure below 1 atm. To reduce erosion and wear by liquid droplets on the blades of the Rankine cycle steam turbine, at least 90% quality should be maintained at the turbine exit: $x_4 > 0.9$.

The *back work ratio*, bwr, is the ratio of the work required by the pump or compressor to the work developed by the turbine:

$$\text{bwr} = \frac{h_2 - h_1}{h_3 - h_4}. \tag{1.6.2a}$$

As a relatively high specific volume vapor expands through the turbine of the Rankine cycle and a much lower specific volume liquid is pumped, the back work ratio is characteristically quite low in vapor power plants—in many cases on the order of 1% to 2%. In the Brayton cycle, however, both the turbine and compressor handle a relatively high specific volume gas, and the back ratio is much larger, typically 40% or more.

The effect of friction and other irreversibilities for flow through turbines, compressors, and pumps is commonly accounted for by an appropriate *isentropic efficiency*. The isentropic turbine efficiency is

$$\eta_t = \frac{h_3 - h_4}{h_3 - h_{4s}}. \tag{1.6.2b}$$

The isentropic compressor efficiency is

$$\eta_c = \frac{h_{2s} - h_1}{h_2 - h_1}. \tag{1.6.2c}$$

In the isentropic pump efficiency, η_p, which takes the same form as Equation 1.6.2c, the numerator is frequently approximated via Equation 1.2.12c *as* $h_{2s} - h_1 \approx v_1 \Delta p$, where Δp is the pressure rise across the pump.

Simple gas turbine power plants differ from the Brayton cycle model in significant respect. In actual operation, excess air is continuously drawn into the compressor, where it is compressed to a higher pressure; then fuel is introduced and combustion occurs; and finally the mixture of combustion products and air expands through the turbine and is subsequently discharged to the surroundings. Accordingly, the low-temperature heat exchanger shown by a dashed line in the Brayton cycle schematic of Table 1.6.1 is not an actual component, but included only to account formally for the cooling in the surroundings of the hot gas discharged from the turbine.

Another frequently employed idealization used with gas turbine power plants is that of an *air-standard analysis*. An air-standard analysis involves two major assumptions: (1) as shown by the Brayton cycle schematic of Table 1.6.1, the temperature rise that would be brought about by combustion is effected instead by a heat transfer from an external source and (2) the working fluid throughout the cycle is air, which behaves as an ideal gas. In a *cold* air-standard analysis, the specific heat ratio k for air is taken as constant. Equations 1 through 6 of Table 1.3.6 together with data from Table A.8 apply generally to air-standard analyses. Equations 1′ to 6′ of Table 1.3.6 apply to cold air-standard analyses, as does the following expression for the turbine power obtained from Table 1.2.1 (Equation 1.2.9c″):

$$\dot{W}_t = \dot{m} \frac{kRT_3}{k-1} \left[1 - \left(p_4 / p_3 \right)^{(k-1)/k} \right] \tag{1.6.3}$$

(Equation 1.6.3 also corresponds to Equation 5′ of Table 1.3.8 when $n = k$.) An expression similar in form can be written for the power required by the compressor.

For the simple Rankine and Brayton cycles of Table 1.6.1, the results of sample calculations are provided in Table 1.6.2. The Brayton cycle calculations are on an air-standard analysis basis.

Otto, Diesel, and Dual Cycles

Although most gas turbines are also internal combustion engines, the name is usually reserved to *reciprocating* internal combustion engines of the type commonly used in automobiles, trucks, and buses. Two principal types of reciprocating internal combustion engines are the *spark-ignition* engine and the *compression-ignition* engine. In a spark-ignition engine, a mixture of fuel and air is ignited by a spark plug. In a compression ignition engine, air is compressed to a high enough pressure and temperature that combustion occurs spontaneously when fuel is injected.

In a *four-stroke* internal combustion engine, a piston executes four distinct strokes within a cylinder for every two revolutions of the crankshaft. Figure 1.6.1 gives *a* pressure–displacement diagram as it might be displayed electronically. With the intake valve open, the piston makes an *intake stroke* to draw a fresh charge into the cylinder. Next, with both valves closed, the piston

TABLE 1.6.2
Sample Calculations for the Rankine and Brayton Cycles of Table 1.6.1

Rankine Cycle

Given data:

$p_1 = p_4 = 8\,\text{kPa}$ (saturated liquid at 1)

$T_3 = 480°\text{C}$ (superheated vapor at 3)

$p_2 = p_3 = 8\,\text{MPa}$

$\dot{W}_{\text{net}} = 100\,\text{MW}$

Ideal cycle: $\eta_t = \eta_p = 100\%$

Actual cycle: $\eta_t = 85\%, \eta_p = 70\%$

Parameter	Ideal Cycle	Actual Cycle
x_4	0.794	0.873
h_2 (kJ/kg)	181.9[a]	185.4
$\dot{m}$ (kg/h)	2.86×10^5	3.38×10^5
η(%)	39.7	33.6
$\dot{Q}_{\text{out}}$(MW)	151.9	197.6
$\dot{E}_{q,\text{out}}$(MW)[b]	8.2	10.7

[a] $h_{2s} \approx h_1 + v_1 \Delta p$

[b] Equation 1.5.2 with $T_0 = 298$ K, $T_j = T_{\text{sat}}\,(8\,\text{kPa}) = 315$ K

Brayton Cycle

Given data:

$p_1 = p_4 = 1\,\text{bar}$

$p_2 = p_3 = 10\,\text{bar}$

$T_3 = 1400\,\text{K}$

$\eta_t = \eta_c = 100\%$

Parameter	Air-Standard Analysis	Cold Air-Standard Analysis ($k = 1.4$)
T_2 (K)	574.1	579.2
T_4 (K)	787.7	725.1
$\dot{W}_{\text{net}}/\dot{m}$ (kJ/kg)	427.2	397.5
η (%)	45.7	48.2
Back work ratio	0.396	0.414

undergoes a *compression stroke* raising the temperature and pressure of the charge. A combustion process is then initiated, resulting in a high-pressure, high-temperature gas mixture. A *power* stroke follows the compression stroke, during which the gas mixture expands and work is done on the piston. The piston then executes an *exhaust stroke* in which the burned gases are purged from the cylinder through the open exhaust valve. Smaller engines operate on *two-stroke* cycles. In two-stroke engines, the intake, compression, expansion, and exhaust operations are accomplished in one revolution of the crankshaft. Although internal combustion engines undergo *mechanical* cycles, the cylinder contents do not execute a *thermodynamic* cycle, since matter is introduced with one composition and is later discharged at a different composition.

A parameter used to describe the performance of reciprocating piston engines is the *mean effective pressure* or mep. The mean effective pressure is the theoretical constant pressure that, if it acted on the piston during the power stroke, would produce the same *net* work as actually developed in one cycle. That is,

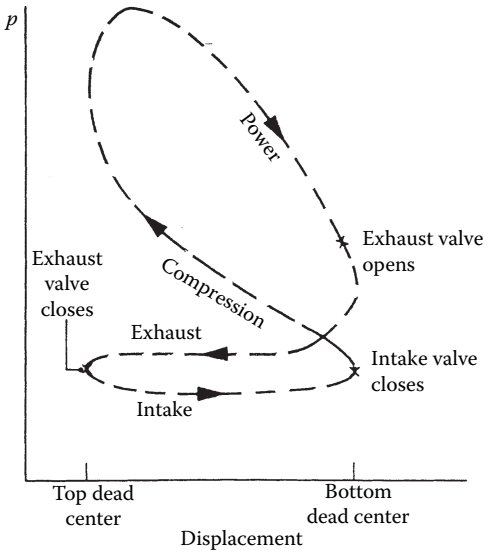


FIGURE 1.6.1 Pressure–displacement diagram for a reciprocating internal combustion engine.

$$\text{mep} = \frac{\text{Net work for one cycle}}{\text{Displacement volume}},$$

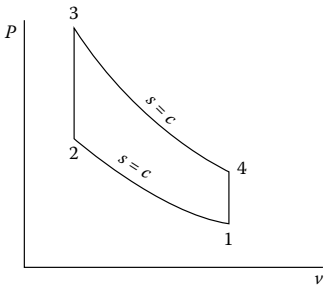
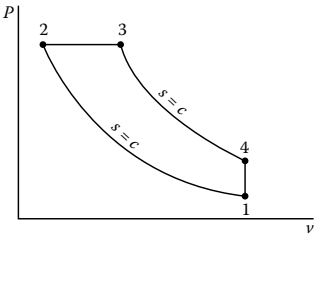
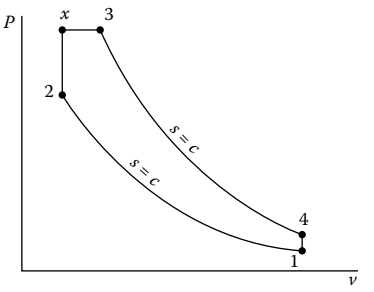
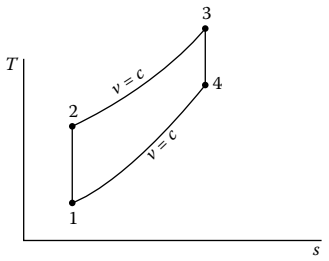
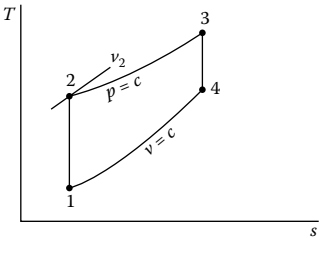
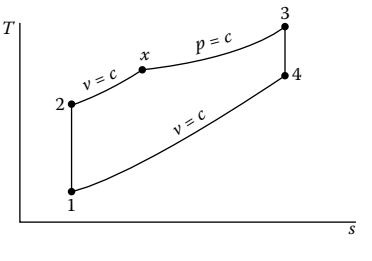
where the displacement volume is the volume swept out by the piston as it moves from the top dead center to the bottom dead center. For two engines of equal displacement volume, the one with a higher mean effective pressure would produce the greater net work and, if the engines run at the same speed, greater power.

Detailed studies of the performance of reciprocating internal combustion engines may take into account many features, including the combustion process occurring within the cylinder and the effects of irreversibilities associated with friction and with pressure and temperature gradients. Heat transfer between the gases in the cylinder and the cylinder walls, and the work required to charge the cylinder and exhaust the products of combustion also might be considered. Owing to these complexities, accurate modeling of reciprocating internal combustion engines normally involves computer simulation.

To conduct *elementary* thermodynamic analyses of internal combustion engines, considerable simplification is required. A procedure that allows engines to be studied *qualitatively* is to employ an *air-standard analysis* having the following elements: (1) a fixed amount of air modeled as an ideal gas is the system, (2) the combustion process is replaced by a heat transfer from an external source and generally represented in terms of elementary thermodynamic processes, (3) there are no exhaust and intake processes as in an actual engine; the cycle is completed by a constant volume heat rejection process, and (4) all processes are internally reversible.

The processes employed in air-standard analyses of internal combustion engines are selected to represent the events taking place within the engine simply and mimic the appearance of observed pressure–displacement diagrams. In addition to the constant volume heat rejection noted previously, the compression stroke and at least a portion of the power stroke are conventionally taken as isentropic. The heat addition is normally considered to occur at constant volume, constant pressure, or constant volume followed by a constant pressure process, yielding, respectively, the Otto, Diesel, and Dual cycles shown in Table 1.6.3. Referring to Table 1.6.3, the ratio v_1/v_2 is the *compression ratio, r.* For the Diesel cycle, the ratio v_3/v_2 is the *cutoff ratio,* r_c.

TABLE 1.6.3
Otto, Diesel, and Dual Cycles

(a) Otto Cycle	(b) Diesel Cycle	(c) Dual Cycle
		
		
$\frac{W_{23}}{m} = 0$	$\frac{W_{23}}{m} = p_2(v_3 - v_2)$	$\frac{W_{23}}{m} = 0, \frac{Q_{2x}}{m} = u_x - u_2$
$\frac{Q_{23}}{m} = u_3 - u_2$	$\frac{Q_{23}}{m} = h_3 - h_2$	$\frac{W_{x3}}{m} = p_3(v_3 - v_2), \frac{Q_{x3}}{m} = h_3 - h_x$
$\eta = 1 - \frac{u_4 - u_1}{u_3 - u_2}$	$\eta = 1 - \frac{u_4 - u_1}{h_3 - h_2}$	$\eta = 1 - \frac{u_4 - u_1}{(u_x - u_2) + (h_3 - h_x)}$

Reducing the closed system energy balance, Equation 1.1.8, gives the following expressions for heat and work applicable in each case shown in Table 1.6.3:

$$\frac{W_{12}}{m} = u_1 - u_2 \quad (< 0)$$

$$\frac{W_{34}}{m} = u_3 - u_4 \quad (> 0)$$

$$\frac{Q_{41}}{m} = u_1 - u_4 \quad (< 0).$$

Table 1.6.3 provides additional expressions for work, heat transfer, and thermal efficiency identified with each case individually. The thermal efficiency, evaluated from Equation 1.1.9, takes the form

$$\eta = 1 - \frac{|Q_{41}/m|}{Q_A/m}.$$

Equations 1 through 6 of Table 1.3.6 together with data from Table A.8 apply generally to air-standard analyses. In a *cold* air-standard analysis, the specific heat ratio k for air is taken as constant. Equations 1′ through 6′ of Table 1.3.6 apply to cold air-standard analyses, as does Equation 4′ of Table 1.3.8, with $n = k$ for the isentropic processes of these cycles.

As all processes are internally reversible, areas on the p–v and T–s diagrams of Table 1.6.3 can be interpreted, respectively, as work and heat transfer. Invoking Equation 1.1.10 and referring to the p–v diagrams, the areas under process 3–4 of the Otto cycle, process 2–3–4 of the Diesel cycle, and process x–3–4 of the Dual cycle represent the work done by the gas during the power stroke, per unit of mass. For each cycle, the area under the isentropic process 1–2 represents the work done on the gas during the compression stroke, per unit of mass. The enclosed area of each cycle represents the net work done per unit mass. With Equation 1.1.15 and referring to the T–s diagrams, the areas under process 2–3 of the Otto and Diesel cycles and under process 2–x–3 of the Dual cycle represent the heat added per unit of mass. For each cycle, the area under the process 4–1 represents the heat rejected per unit of mass. The enclosed area of each cycle represents the net heat added, which equals the net work done, each per unit of mass.

Carnot, Ericsson, and Stirling Cycles

Three thermodynamic cycles that exhibit the Carnot efficiency (Equation 1.1.12) are the Carnot, Ericsson, and Stirling cycles shown in Figure 1.6.2. Each case represents a reversible power cycle in which heat is added from an external source at a constant temperature T_H (process 2–3) and rejected to the surroundings at a constant temperature T_C (process 4–1). Carnot cycles can be configured both as vapor power cycles and as cycles executed by a gas in a piston-cylinder assembly (see Moran et al., 2014). Carnot cycles also can be executed in systems where a capacitor is charged and discharged, a paramagnetic substance is magnetized and demagnetized, and in other ways. Regardless of the type of device and the working substance used, the Carnot cycle always has the same four internally reversible processes in series: two isentropic processes alternated with two isothermal processes.

The Ericsson and Stirling cycles also consist of four internally reversible processes in series: heating from state 1 to state 2 (at constant pressure in the Ericsson cycle and at constant volume in the Stirling cycle), isothermal heating from state 2 to state 3 at temperature T_H, cooling from state 3 to state 4 (at constant pressure in the Ericsson cycle and at constant volume in the Stirling cycle), and isothermal cooling from state 4 to state 1 at temperature T_C. An ideal regenerator allows the heat

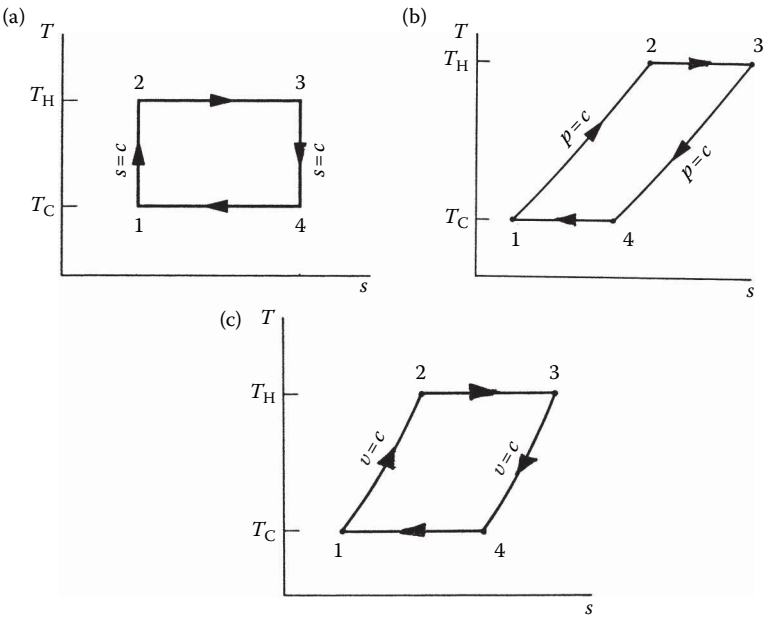


FIGURE 1.6.2 (a) Carnot, (b) Ericsson, and (c) Stirling cycles.

input required for process 1–2 to be obtained from the heat rejected in process 3–4. Accordingly, as in the Carnot cycle all the heat added externally occurs at T_H and all of the heat rejected to the surroundings occurs at T_C.

The Ericsson and Stirling cycles are principally of theoretical interest as examples of cycles that exhibit the same thermal efficiency as the Carnot cycle (Equation 1.1.12). However, a practical engine of the piston-cylinder type that operates on a closed regenerative cycle having features in common with the Stirling cycle has been under study in recent years. This engine, known as the *Stirling engine*, offers the opportunity for high efficiency together with reduced emissions from combustion products because the combustion takes place externally and not within the cylinder as in internal combustion engines. In the Stirling engine, energy is transferred to the working fluid from products of combustion, which are kept separate. It is an *external* combustion engine.

1.7 GUIDELINES FOR IMPROVING THERMODYNAMIC EFFECTIVENESS

Michael J. Moran

Thermal design frequently aims at the most effective system from the cost viewpoint. Still, in the cost optimization process, particularly of complex energy systems, it is often expedient to begin by identifying a design that is nearly optimal thermodynamically; such a design can then be used as a point of departure for cost optimization. Presented in this section are guidelines for improving the use of fuels (natural gas, oil, and coal) by reducing sources of thermodynamic inefficiency in thermal systems. Further discussion is provided by Bejan et al. (1996).

To improve thermodynamic effectiveness, it is necessary to deal directly with inefficiencies related to exergy destruction and exergy loss. The primary contributors to exergy destruction are chemical reaction, heat transfer, mixing, and friction, including unrestrained expansions of gases and liquids. To deal with them effectively, the principal sources of inefficiency not only should be understood qualitatively but also determined quantitatively, at least approximately. Design changes to improve effectiveness must be done judiciously, however, for the cost associated with different sources of inefficiency can be different. For example, the unit cost of the electrical or mechanical power required to provide for the exergy destroyed owing to a pressure drop is generally higher than the unit cost of the fuel required for the exergy destruction caused by combustion or heat transfer.

Since chemical reaction is a significant source of thermodynamic inefficiency, it is generally good practice to minimize the use of combustion. In many applications, the use of combustion equipment such as boilers is unavoidable, however. In these cases, a significant reduction in the combustion irreversibility by conventional means simply cannot be expected, for the major part of the exergy destruction introduced by combustion is an inevitable consequence of incorporating such equipment. Still, the exergy destruction in practical combustion systems can be reduced by minimizing the use of excess air and by preheating the reactants. In most cases, only a small part of the exergy destruction in a combustion chamber can be avoided by these means. Consequently, after considering such options for reducing the exergy destruction related to combustion, efforts to improve thermodynamic performance should focus on components of the overall system that are more amenable to betterment by cost-effective conventional measures. In other words, *some exergy destructions and energy losses can be avoided, others cannot. Efforts should be centered on those that can be avoided.*

Nonidealities associated with heat transfer also typically contribute heavily to inefficiency. Accordingly, unnecessary or cost-ineffective heat transfer must be avoided. Additional guidelines follow:

- The higher the temperature T at which a heat transfer occurs in cases where $T > T_0$, where T_0 denotes the temperature of the environment (Section 1.5), the more valuable the heat transfer and, consequently, the greater the need to avoid heat transfer to the ambient, to cooling water, or to a refrigerated stream. Heat transfer across T_0 should be avoided.
- The lower the temperature T at which a heat transfer occurs in cases where $T < T_0$, the more valuable the heat transfer and, consequently, the greater the need to avoid direct heat transfer to the ambient or a heated stream.
- Since exergy destruction associated with heat transfer between streams varies inversely with the temperature level, the lower the temperature level, the greater the need to minimize the stream-to-stream temperature difference.
- Avoid the use of intermediate heat-transfer fluids when exchanging energy by heat transfer between two streams.

Although irreversibilities related to friction, unrestrained expansion, and mixing are often secondary in importance to those of combustion and heat transfer, they should not be overlooked, and the following guidelines apply:

- Relatively more attention should be paid to the design of the lower temperature stages of turbines and compressors (the last stages of turbines and the first stages of compressors) than to the remaining stages of these devices.
- For turbines, compressors, and motors, consider the most thermodynamically efficient options.
- Minimize the use of throttling; check whether power recovery expanders are a cost-effective alternative for pressure reduction.
- Avoid processes using excessively large thermodynamic driving forces (differences in temperature, pressure, and chemical composition). In particular, minimize the mixing of streams differing significantly in temperature, pressure, or chemical composition.
- The greater the mass rate of flow, the greater the need to use the exergy of the stream effectively.
- The lower the temperature level, the greater the need to minimize friction.

Flowsheeting or *process simulation* softwares can assist efforts aimed at improving thermodynamic effectiveness by allowing engineers to readily model the behavior of an overall system, or system components, under specified conditions and do the required thermal analysis, sizing, costing, and optimization. Since process simulation is a rapidly evolving field, vendors should be contacted for up-to-date information concerning the features of flowsheeting software, including optimization capabilities (if any).

1.8 EXERGOECONOMICS

George Tsatsaronis

Exergoeconomics is an exergy-aided cost reduction method that combines exergy and cost analyses to provide the designer or operator of an energy conversion plant with information not available through conventional energy, exergy, or cost analyses. Exergoeconomics is also a powerful tool for understanding the interconnections between thermodynamics and economics and, thus, the behavior of an energy conversion plant from the cost viewpoint. For a more detailed presentation of exergoeconomics, readers may refer to the following references: Bejan et al. (1996), Lazzaretto and Tsatsaronis (2006), Tsatsaronis (1984, 1993, 1994, 1999a, 2007, 2011), and Tsatsaronis et al. (1985, 1986). Sections 1.2 through 1.5 outline the principles for conducting detailed thermodynamic evaluations of thermal systems. In particular, techniques have been developed for evaluating the thermodynamic inefficiencies of these systems: exergy destructions and exergy losses. However, we often need to know the costs (and the environmental impacts) associated with such inefficiencies. This information is very useful for improving the overall effectiveness of the system—that is, for reducing the costs (and the environmental impacts) associated with the final products produced by the system.

In addition, if a system has more than one product, as for example the net power and saturated vapor of the cogeneration system shown in Figure 1.8.1, we would want to know the production cost for each product. This is a common problem in chemical plants where electrical power, chilled water, compressed air, and steam at various pressure levels are generated in one department and sold to another. The plant operator wants to know the true cost at which each of the utilities is generated; these costs are then charged to the appropriate final products according to the type and amount of each utility used to generate a final product. In the design of a thermal system, such cost allocation assists in pinpointing cost-ineffective processes and operations and in identifying technical options that might improve the cost-effectiveness of the system.

Accordingly, the objective of an exergoeconomic analysis might be to (1) calculate separately the cost of each product generated by a system having more than one product, (2) estimate the cost associated with each inefficiency within the system, (3) understand the cost formation process and the flow of costs in the system, (4) optimize specific variables in a single component, or (5) minimize the costs associated with the overall system.

An exergoeconomic analysis must be preceded by an exergy analysis (see Section 1.5) and an economic analysis [see Bejan et al. (1996), Tsatsaronis and Winhold (1984), and Section 1.10]. Because of the variation of costs from year to year, we must use the cost levelization approach when we evaluate the design of a thermal system from the cost viewpoint. Therefore, the cost values used throughout the following discussion are levelized costs. For conciseness, the term levelized is omitted, however.

Cost accounting in exergoeconomics calls for the use of cost balances. Cost balances can be formulated for an overall system (subscript tot) operating at steady state

$$\dot{C}_{\mathrm{P,tot}} = \dot{C}_{\mathrm{F,tot}} + \dot{Z}^{\mathrm{Cl}}_{\mathrm{tot}} + \dot{Z}^{\mathrm{OM}}_{\mathrm{tot}} \tag{1.8.1}$$

and for each component (see Equations 1.8.6 and 1.8.7). Equation 1.8.1 indicates that the cost rate associated with the product of the system ($\dot{C}_{\mathrm{P}}$) equals the total rate of expenditures made to generate the product, namely, the fuel cost rate ($\dot{C}_{\mathrm{F}}$) and the cost rates associated with capital investment ($\dot{Z}^{\mathrm{CI}}$) and operating and maintenance (O&M) ($\dot{Z}^{\mathrm{OM}}$). Here, and throughout Sections 1.8, 1.9, and 1.10, the terms *fuel* and *product* are used in the sense introduced in Section 1.5. When referring to a single stream associated with a fuel or product, the expression *fuel stream* or *product stream* is used. The rates $\dot{Z}^{\mathrm{CI}}$ and $\dot{Z}^{\mathrm{OM}}$ are calculated by dividing the annual contribution of capital investment and the annual operating and maintenance costs, respectively, by the

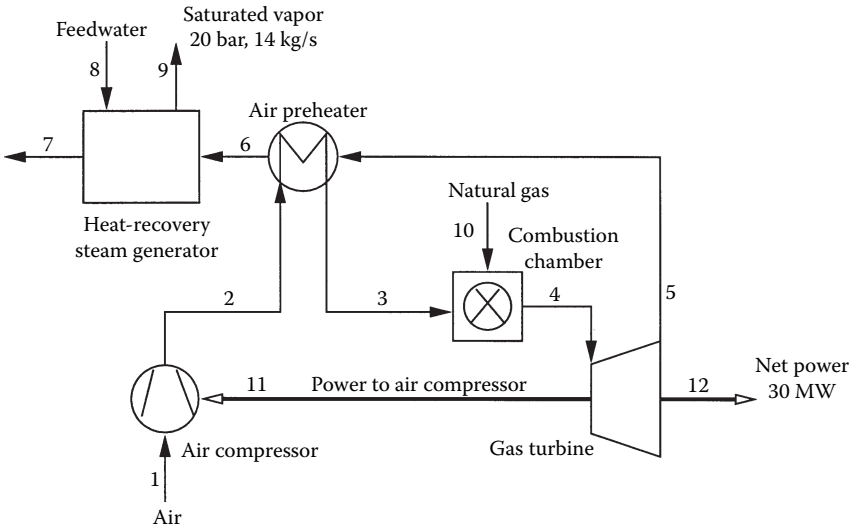


FIGURE 1.8.1 Cogeneration system.

number of time units (usually hours or seconds) of system operation per year. The sum of these two variables is denoted by $\dot{Z}$:

$$\dot{Z} = \dot{Z}^{\mathrm{CI}} + \dot{Z}^{\mathrm{OM}}. \tag{1.8.2}$$

In this section, we discuss the basic elements of exergoeconomics, which include exergy costing, cost balances, and means for costing various exergy transfers. The exergoeconomic variables defined later are used in the evaluation and optimization of the design and operation of thermal systems.

Exergy Costing

For a system operating at steady state, there may be a number of entering and exiting streams as well as both heat and work interactions with the surroundings. Associated with these transfers of matter and energy are exergy transfers into and out of the system and exergy destructions caused by the irreversibilities within the system. Since exergy measures the *true* thermodynamic value of such effects, and costs should only be assigned to commodities of value, it is meaningful to use exergy as a basis for assigning costs in thermal systems. Indeed, exergoeconomics rests on the notion that exergy is the *only* rational basis for assigning costs to the interactions a thermal system experiences with its surroundings and to the sources of inefficiencies within it. We refer to this approach as *exergy costing.*

In exergy costing, a cost rate is associated with each exergy transfer. Thus, for entering and exiting streams of matter with exergy transfers $\dot{E}_i$ and $\dot{E}_e$, respectively, power $\dot{W}$, and exergy transfer associated with heat transfer $\dot{E}_q$, we write, respectively,

$$\dot{C}_i = c_i \dot{E}_i,\ \dot{C}_e = c_e \dot{E}_e \tag{1.8.3}$$

$$\dot{C}_w = c_w \dot{W},\ \dot{C}_q = c_q \dot{E}_q, \tag{1.8.4}$$

where c_i, c_e, c_w, and c_q denote *average* costs per unit of exergy, for example, in dollars per gigajoule ($/GJ).

Exergy costing does not necessarily imply that costs associated with streams of matter are related *only* to the exergy rate of each respective stream. Nonexergy-related costs also can affect the total cost rate associated with material streams. Examples include the cost rates associated with a treated water stream at the outlet of a water treatment unit, an oxygen or nitrogen stream at the outlet of an air separation unit, a limestone stream supplied to a gasifier or fluidized-bed reactor, iron feedstock supplied to a metallurgical process, and an inorganic chemical fed to a chemical reactor. Accordingly, when significant nonexergy-related costs occur, the total cost rate associated with the material stream j, denoted by $\dot{C}_j^{\mathrm{TOT}}$, is given by

$$\dot{C}_j^{\mathrm{TOT}} = \dot{C}_j + \dot{C}_j^{\mathrm{NE}}, \tag{1.8.5}$$

where $\dot{C}_j$ is the cost rate directly related to the exergy of stream j (e.g., Equation 1.8.3), and $\dot{C}_j^{\mathrm{NE}}$ is the cost rate due to nonexergetic effects. The term $\dot{C}_j^{\mathrm{NE}}$ represents a convenient way for charging nonexergy-related costs from one component to other components that should bear such costs. More details about $\dot{C}_j^{\mathrm{NE}}$ are given in Bejan et al. (1996) and Tsatsaronis et al. (1986).

Cost Balance

Exergy costing usually involves cost balances formulated for each component separately. A cost balance applied to the kth system component indicates that the sum of cost rates associated with all exiting exergy transfers equals the sum of cost rates of all entering exergy transfers plus the appropriate charges due to capital investment ($\dot{Z}_k^{\mathrm{CL}}$) and operating and maintenance expenses ($\dot{Z}_k^{\mathrm{OM}}$). The sum of the last two terms is denoted by ($\dot{Z}_k$). For example, for a component receiving heat transfer (subscript q) and generating power (subscript w), we write

$$\sum_e \dot{C}_{e,k} + \dot{C}_{w,k} = \dot{C}_{q,k} + \sum_i \dot{C}_{i,k} + \dot{Z}_k. \tag{1.8.6}$$

This equation simply states that the total cost of the exiting exergy transfers equals the total expenditure to obtain them: the cost of the entering exergy streams plus the capital and other costs. When a component receives power (as in a compressor or a pump) the term $\dot{C}_{w,k}$ would move with its positive sign to the right side of this expression. The term $\dot{C}_{q,k}$ would appear with its positive sign on the left side if there is a heat transfer *from* the component. Cost balances are generally written so that all terms are positive.

Introducing the cost rate expressions of Equations 1.8.3 and 1.8.4, Equation 1.8.6 becomes

$$\sum_e (c_e \dot{E}_e)_k + c_{w,k}\dot{W}_k = c_{q,k}\dot{E}_{q,k} + \sum_i (c_i \dot{E}_i)_k + \dot{Z}_k. \tag{1.8.7}$$

The exergy rates exiting and entering the kth component denoted by $\dot{E}_{e,k}, \dot{W}_k, \dot{E}_{q,k}$, and $\dot{E}_{i,k}$ are calculated in an exergy analysis conducted at a previous stage. The term $\dot{Z}_k$ is obtained by first calculating the capital investment and O&M costs associated with the kth component and then computing the levelized values of these costs per unit of time (year, hour, or second) of system operation [see Bejan et al. (1996) and Tsatsaronis et al. (1984, 1986)].

The variables in Equation 1.8.7 are the levelized costs per unit of exergy for the exergy transfers associated with the kth component: $c_{e,k}$, $c_{w,k}$, $c_{q,k}$, and $c_{i,k}$. In analyzing a component, we may assume that the costs per exergy unit are known for all entering streams. These costs are known from the components they exit or, if a stream enters the overall system consisting of all components under consideration, from the purchase cost of this stream. Consequently, the unknown variables to be calculated with the aid of the cost balance for the kth component are the costs per exergy unit of the

exiting streams $c_{e,k}$ and, if power or useful heat are generated in that component, the cost per unit of exergy associated with the transfer of power $c_{w,k}$ or heat $c_{q,k}$. Some auxiliary equations are usually necessary to calculate these costs, as discussed next.

Auxiliary Costing Equations

Various approaches for formulating the auxiliary equations are suggested in the literature. However, the method recommended for obtaining an objective set of auxiliary equations consistent with the definitions of fuel and product is detailed by Lazzaretto and Tsatsaronis (2006) and summarized in the following three steps:

Step 1: Exergy Streams

All material and energy streams crossing the boundaries of the component being considered should be first identified, and the exergy values associated with these streams should be calculated.

Step 2: Definition of Fuel and Product

In evaluating the performance of a component, it is generally meaningful and appropriate to operate with exergy differences associated with each material stream between the inlet and outlet of the component. For example, in defining the product of a heat exchanger operating above ambient temperature, we consider only the exergy addition to the cold stream and not the sum of the exergies associated with the material streams at the outlet. Similarly, for defining the fuel of the heat exchanger, we consider only the exergy removal from the hot stream and not the sum of the exergies associated with the material streams at the inlet. Exergy differences (exergy additions to or removals from a stream) should be applied to all material streams undergoing a change of physical exergy (Equation 1.5.5a) and to some material streams undergoing a chemical exergy (Equation 1.5.6) conversion. This approach has been used in developing Table 1.5.1 for all cases except the gasifier/combustion chamber, which is considered next.

In many cases involving conversion of chemical exergy (e.g., conversion of chemical exergy of a solid fuel in chemical and thermal exergy through a gasification process), the purpose of owning and operating the component dictates that the chemical exergy at the outlet is considered on the product side and the chemical exergy of the fuel stream at the inlet on the fuel side. Thus, in the definition of the exergetic efficiency of a gasifier or combustion chamber in Table 1.5.1, the exergy of the "fuel" for the component equals the value of the exergy of the entering fuel stream.

Accordingly, when considering the fuel and product of a component, a decision must be made for each exergy stream with respect to whether an exergy difference or just the exergy values at the inlet or outlet should be included in the definitions of fuel and product. Then, the product consists of all the exergy values to be considered at the outlet plus all the exergy increases between inlet and outlet (i.e., the exergy additions to the respective material streams). Similarly, the fuel consists of the sum of all exergy rates to be considered at the inlet and all the exergy decreases between inlet and outlet (i.e. the exergy removals from the respective material streams).

Step 3: Auxiliary Equations

In general, if there are N_e exergy streams exiting the component being considered, we have N_e unknowns and only one equation, the costs balance (Equation 1.8.7). Therefore, we need to formulate $N_e - 1$ auxiliary equations. This is accomplished with the aid of the F and P rules presented next:

- The F rule refers to the *removal of exergy* from an exergy stream within the component being considered. The F rule states that the total cost associated with this removal of exergy must be equal to the average cost at which the removed exergy was supplied to the same stream in upstream components. The number of auxiliary equations provided by the

F rule is always equal to the number ($N_{e,\mathrm{F}}$) of exiting exergy streams that are considered in the definition of the fuel for the component.

- The P rule refers to the *supply of exergy* to an exergy stream within the component being considered and to the costing of streams associated with the product. The P rule states that each exergy unit is supplied to any stream associated with the product at the *same average cost,* c_{P}. This cost is calculated from the cost balance and the equations obtained by applying the F rule. The number of auxiliary equations provided by the P rule is always equal to $N_{e,\mathrm{P}} - 1$, where $N_{e,\mathrm{P}}$ is the number of exiting exergy streams that are included in the product definition.

Since the total number of exiting streams (N_c) is equal to the sum ($N_{e,\mathrm{F}}+N_{e,\mathrm{P}}$), the F and P rules together provide the required $N_e - 1$ auxiliary equations.

General Example

The general application of these steps may be demonstrated with the aid of Figure 1.8.2.

Step 1: Referring to Figure 1.8.2, there are seven exergy streams (1, 2, and 5 through 9) entering the component (subscript *i*) and seven exergy streams (3 through 8 and 10) exiting the component (subscript *e*). The streams shown in this figure are selected to cover all situations that might be encountered. In an actual component, however, not all of the streams shown in Figure 1.8.2 exist.

Step 2: The exergy streams 1 through 4 are associated with the chemical exergy or the total exergy of the corresponding material streams. The purpose of owning and operating the component dictates that the entering streams 1 and 2 should be part of the fuel, whereas the exiting streams 3 and 4 should be part of the product. For the exergy streams with the numbers 5, 6, 7, and 8, the purpose of the component dictates the consideration of the respective exergy differences between outlet and inlet. These are positive for streams 7 and 8 and negative for streams 5 and 6. Streams 9 at the inlet and 10 at the outlet represent exergy streams associated with the transport of mechanical, electrical, or thermal energy. We conclude that exergy streams 1, 2, 5, 6, and 9 are associated with the fuel, whereas streams 3, 4, 7, 8, and 10 are associated with the product of the component. Thus, the fuel and product of the component, respectively, are

$$\dot{E}_{\mathrm{F}} = \dot{E}_1 + \dot{E}_2 + \left(\dot{E}_{5i} - \dot{E}_{5e}\right) + \left(\dot{E}_{6i} - \dot{E}_{6e}\right) + \dot{E}_9 \tag{1.8.8}$$

$$\dot{E}_{\mathrm{P}} = \dot{E}_3 + \dot{E}_4 + \left(\dot{E}_{7e} - \dot{E}_{7i}\right) + \left(\dot{E}_{8e} - \dot{E}_{8i}\right) + \dot{E}_{10}. \tag{1.8.9}$$

Step 3: In accordance with Equation 1.8.7, the cost balance for the component is

$$\dot{C}_3 + \dot{C}_4 + \dot{C}_{5e} + \dot{C}_{6e} + \dot{C}_{7e} + \dot{C}_{8e} + \dot{C}_{10} = \dot{C}_1 + \dot{C}_2 + \dot{C}_{5i} + \dot{C}_{6i} + \dot{C}_{7i} + C_{8i} + \dot{C}_9 + \dot{Z}. \tag{1.8.10}$$

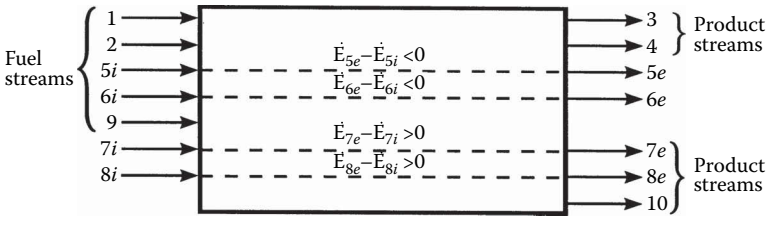


FIGURE 1.8.2 Schematic of a component in a thermal system to define fuel, product, and auxiliary equations.

By grouping the terms associated with fuel and product, we obtain

$$\dot{C}_3+\dot{C}_4+(\dot{C}_{7e}-\dot{C}_{7i})+(\dot{C}_{8e}-\dot{C}_{8i})+\dot{C}_{10}=\dot{C}_1+\dot{C}_2(\dot{C}_{5e}-\dot{C}_{5i})+(\dot{C}_{6e}-\dot{C}_{6i})+\dot{C}_9+\dot{Z}. \quad (1.8.11)$$

The F rule states that the total cost rate associated with the removal of exergy from stream 5 ($\dot{C}_{5i}-\dot{C}_{5e}$) must be equal to the *average* cost at which the removed exergy ($\dot{E}_{5i}-\dot{E}_{5e}$) was supplied to that stream in upstream components. Since each exergy unit was supplied to stream 5 at the average cost of c_{5i}, the F rule for stream 5 becomes

$$\dot{C}_{5i}-\dot{C}_{5e}=c_{5i}(\dot{E}_{5i}-\dot{E}_{5e}). \quad (1.8.12)$$

From this equation we obtain

$$c_{5e}=c_{5i}. \quad (1.8.13)$$

Similarly, for stream 6

$$c_{6e}=c_{6i}. \quad (1.8.14)$$

The P rule states that each exergy unit is *supplied* to all streams associated with the product at the *same* average cost, c_P. This rule leads to the following equation:

$$c_3=c_4=\frac{\dot{C}_{7e}-\dot{C}_{7i}}{\dot{E}_{7e}-\dot{E}_{7i}}=\frac{\dot{C}_{8e}-\dot{C}_{8i}}{\dot{E}_{8e}-\dot{E}_{8i}}=c_{10}=c_P. \quad (1.8.15)$$

Since we assume that the cost rate $\dot{Z}$ and all costs associated with all entering streams are known, we can calculate the unknowns, $\dot{C}_3$, $\dot{C}_4$, $\dot{C}_{5e}$, $\dot{C}_{6e}$, $\dot{C}_{7e}$, $\dot{C}_{8e}$, and $\dot{C}_{10}$, by solving the system of Equations 1.8.11, 1.8.13, 1.8.14, and 1.8.15.

Note that Equation 1.8.11 can be written as

$$c_P\dot{E}_P=\dot{C}_1+\dot{C}_2+(\dot{C}_{5i}-\dot{C}_{5e})+(\dot{C}_{6i}-\dot{C}_{6e})+\dot{C}_9+\dot{Z} \quad (1.8.16)$$

or

$$c_P\dot{E}_P=c_F\dot{E}_F+\dot{Z}, \quad (1.8.17)$$

where $\dot{E}_F$ $\dot{E}_P$, and c_P are given in Equations 1.8.8, 1.8.9, and 1.8.15, respectively. The variable c_F denotes the average cost at which each exergy unit of $\dot{E}_F$ is supplied to the component. Equation 1.8.16 demonstrates that the average cost (c_P) at which each exergy unit is supplied to all streams associated with the product can be calculated by combining the cost balance (Equation 1.8.16) with the equations obtained from the F rule (Equations 1.8.13 and 1.8.14).

After calculating the value of c_P, Equation 1.8.15 can be used to calculate the cost associated with each exiting stream included in the definition of the product. This suggests that the cost balance (Equation 1.8.16) should always be used to calculate the value of c_P with the aid of the auxiliary equations obtained from the F rule.

Exergoeconomic Variables and Evaluation

The cost balances together with the auxiliary equations formulated for each plant component form a system of linear equations, the solution of which provides the cost per exergy unit and the cost rates associated with all exergy streams in the system. The remaining exergoeconomic variables

are calculated from these cost rates and the known exergy rates using the equations for defining the respective exergoeconomic variables, as discussed next.

The exergoeconomic evaluation is conducted at the system component level using the following variables for the kth component. From the exergy analysis we know the

- Rate of exergy destruction $\dot{E}_{\mathrm{D},k}$:

$$\dot{E}_{\mathrm{D},k} = \dot{E}_{\mathrm{F},k} - \dot{E}_{\mathrm{P},k} - \dot{E}_{\mathrm{L},k} \tag{1.8.18}$$

- Exergetic efficiency ε_k:

$$\varepsilon_k = \dot{E}_{\mathrm{P},k} / \dot{E}_{\mathrm{F},k} = 1 - \left(\dot{E}_{\mathrm{D},k} + \dot{E}_{\mathrm{L},k}\right) / \dot{E}_{\mathrm{F},k} \tag{1.8.19}$$

- Exergy destruction ratio y_k:

$$y_{\mathrm{D},k} = \dot{E}_{\mathrm{D},k} / \dot{E}_{\mathrm{F,tot}} \tag{1.8.20}$$

In addition, we calculate the following variables from the exergoeconomic analysis:

- Cost per unit of fuel exergy $c_{\mathrm{F},k}$:

$$c_{\mathrm{F},k} = \dot{C}_{\mathrm{F},k} / \dot{E}_{\mathrm{F},k} \tag{1.8.21}$$

- Cost per unit of product exergy $c_{\mathrm{P},k}$:

$$c_{\mathrm{P},k} = \dot{C}_{\mathrm{P},k} / \dot{E}_{\mathrm{P},k} \tag{1.8.22}$$

- Cost rate associated with exergy destruction $\dot{C}_{\mathrm{D},k}$:

$$\dot{C}_{\mathrm{D},k} = c_{\mathrm{F},k} \dot{E}_{\mathrm{D},k} \tag{1.8.23}$$

- Cost rate associated with exergy losses $\dot{C}_{\mathrm{L},k}$

$$\dot{C}_{\mathrm{L},k} = c_{\mathrm{F},k}\, \dot{E}_{\mathrm{L},k} \tag{1.8.24}$$

- Cost rate associated with capital investment $\dot{Z}_k^{\mathrm{CI}}$
- Cost rate associated with operating and maintenance expenses $\dot{Z}_k^{\mathrm{OM}}$
- Sum $\dot{Z}_k$ of the cost rates associated with capital investment and O&M expenses:

$$\dot{Z}_k = \dot{Z}_k^{\mathrm{CI}} + \dot{Z}_k^{\mathrm{OM}} \tag{1.8.25}$$

- Relative cost difference r_k:

$$r_k = \frac{c_{\mathrm{P},k} - c_{\mathrm{F},k}}{c_{\mathrm{F},k}} = \frac{1-\varepsilon_k}{\varepsilon_k} + \frac{\dot{Z}_k}{c_{\mathrm{F},k}\dot{E}_{\mathrm{P},k}} \tag{1.8.26}$$

- Exergoeconomic factor f:

$$f_k = \frac{\dot{Z}_k}{\dot{Z}_k + c_{\mathrm{F},k}(\dot{E}_{\mathrm{D},k} + \dot{E}_{\mathrm{L},k})} \tag{1.8.27}$$

The value of $c_{F,k}$ depends on the relative position of the kth component in the system and on the interconnections among the kth component and the remaining components. As a general rule, the closer the kth component to the product (fuel) stream of the overall system, the larger (smaller) the value of $c_{F,k}$.

An exergoeconomic analysis

1. Identifies and compares the real cost sources in a system: Equations 1.8.23 through 1.8.25.
2. Illustrates the cost formation process within a system.
3. Calculates separately the cost at which each product stream is generated.
4. Most importantly, assists in the effective cost minimization in a thermal system, particularly its design.

Examples of design optimization (improvement) are given in Bejan et al. (1996), Lin and Tsatsaronis (1993), Tsatsaronis (1993, 2015), and Tsatsaronis et al. (1984, 1985, 1986, 1991, 1992, 1994). This point is discussed in Section 1.9.

1.9 DESIGN OPTIMIZATION

George Tsatsaronis

Design optimization of a thermal system means the modification of the structure and the design parameters of a system to minimize the total levelized cost of the system products under boundary conditions associated with available materials, financial resources, protection of the environment, and government regulation, together with the safety, reliability, operability, availability, and maintainability of the system. A truly optimized system is one for which the magnitude of every significant thermodynamic inefficiency (exergy destruction and exergy loss) is justified by considerations related to costs or is imposed by at least one of the above-mentioned boundary conditions. A *thermodynamic* optimization, which aims only at minimizing the thermodynamic inefficiencies, may be considered as a subcase of design optimization based on cost minimization.

An appropriate formulation of the optimization problem is usually the most important and sometimes the most difficult step of a successful optimization study. In optimization problems, we separate the independent variables into *decision variables* and *parameters*. The values of the decision variables are amenable to change. The values of the parameters are fixed by the particular application. In optimization studies, only the decision variables may be varied. The parameters are independent variables that are each given one specific and unchanging value in any particular model statement. The variables whose values are calculated from the independent variables using a mathematical model are the dependent variables.

However, the optimization of thermal systems involves complexities that render conventional mathematical optimization approaches ineffective. The reasons include the following:

- Some of the input data and functions required for the thermodynamic and, particularly, the economic model might not be available or might not be in the required form. For example, it is not always possible to express the purchased-equipment costs (PEC) as a function of the appropriate thermodynamic decision variables.
- A significant decrease in the product costs may be achievable only through changes in the structure of the system, but such changes are seldom elicited from conventional optimization techniques focusing on the optimization of a particular structure. Moreover, it is not always practical to develop a mathematical optimization model for every promising system structure.
- Even if all the required information is available, the complexity of the system might not allow a satisfactory mathematical model to be formulated and solved in a reasonable time. Very often developing an appropriate superstructure for the problem under consideration (to simultaneously optimize the structure of the system and the decision variables) results in a very complex optimization problem that cannot be handled by the available mathematical tools (e.g., mixed integer nonlinear programing).

In such cases, the application of exergoeconomic techniques may provide significant benefits for the optimization process. The more complex the thermal system, the larger the expected benefits, particularly when chemical reactions are involved. The interactions of exergoeconomics with several other areas during the optimization procedure are shown schematically in Figure 1.9.1. Exergoeconomics uses results from the synthesis, cost analysis, and simulation of thermal systems and provides useful information for the evaluation and optimization of these systems as well as for the application of expert systems to improve the design and operation of such systems.

The following section presents the main features of a general methodology that can be used to evaluate and iteratively optimize the design of a thermal system.

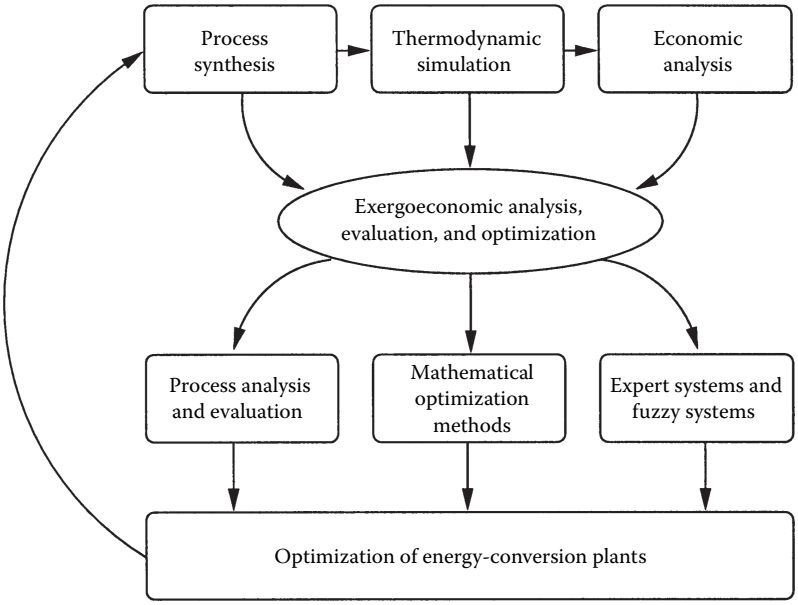


FIGURE 1.9.1 Interactions of exergoeconomics with other areas of engineering and optimization procedure.

An Iterative Exergoeconomic Procedure for Optimizing the Design of a Thermal System

The conventional approach to optimization is to iteratively optimize subsystems and/or ignore the influence of some structural changes and decision variables. An effective alternative approach for the optimization of complex systems is the following iterative exergoeconomic optimization technique that consists of seven steps:

1. In the first step, a workable design is developed. The guidelines presented in Section 1.7 and in Bejan et al. (1996), Lin and Tsatsaronis (1993), Linnhoff et al. (1982), Sama (1993), Tsatsaronis (1993), and Tsatsaronis and Pisa (1994) may assist in developing a workable design that is relatively close to the optimal design. Thus, the use of these guidelines can reduce the total number of required iterations.
2. A detailed exergoeconomic analysis and, if necessary, a pinch analysis are conducted for the design configuration developed in the previous step. The results are used to determine design changes that are expected to improve the design being considered. In this step, and in steps 3 through 7, we consider only changes in the decision variables that affect both the exergetic efficiency and the investment costs. The remaining decision variables are optimized in step 8.
3. A component optimization may be conducted for one or two components in isolation when in these components the sum of the cost rates $(\dot{Z}_k + \dot{C}_{D,k})$ is *significantly* higher than the same sum for the remaining components. For this, we assume that the costs per exergy unit remain constant for all inlet streams. Step 3 is meaningful only for components in which each of the terms $\dot{Z}_k$ and $\dot{C}_{D,k}$ has a significant contribution to the costs associated with the respective component. If not, step 3 should either be omitted or, preferably, replaced by (1) an efficiency maximization procedure when $\dot{C}_{D,k}$ is the dominating cost rate, or (2) an investment cost minimization procedure when $\dot{Z}_k$ is the dominating cost rate. A different approach is discussed by Tsatsaronis et al. (1991, 1994). The sum $(\dot{Z}_k + \dot{C}_{D,k})$ is a measure

of the economic importance of the kth component. Therefore, the components should be considered in order of descending value of this sum. The quality of this information is significantly improved when we consider only *avoidable* costs (see also Section 1.12) associated with both the capital investment and the cost of exergy destruction in the kth component (Tsatsaronis and Park, 2002).

4. The exergoeconomic factor f_k is used to identify the major cost source (capital investment or cost of exergy destruction).
   a. When the f_k value is high, investigate whether it is cost-effective to reduce the capital investment of the kth component, usually at the expense of the component efficiency.
   b. When the f_k value is low, try to improve the component efficiency by accepting an increase in the capital investment.
5. Eliminate any subprocesses that increase the exergy destruction or exergy loss without contributing to the reduction of capital investment or of fuel costs for other components.
6. Consider improving the exergetic efficiency of a component if it has a relatively low exergetic efficiency or relatively large values for the rate of exergy destruction, the exergy destruction ratio, or the exergy loss ratio.
7. Based on the results from steps 2 through 6, a new design is developed and the value of the objective function for this design is calculated. In comparison with the previous design, if the value of the objective function has been improved, we may decide to proceed with another iteration that involves steps 2 through 7. If, however, the value of the objective function is not better in the new design than in the previous one, we may either revise some design changes and repeat steps 2 through 7 or proceed with step 8.
8. In this step, we use an appropriate mathematical optimization technique to optimize the decision variables that affect the investment costs but not the exergetic efficiency. At the end of this step, a cost-optimal design is obtained.
9. Finally, a parametric study may be conducted to investigate the effect on the optimization results of some parameters used and/or assumptions made in the optimization procedure.

When applying this methodology, it is important to recognize that the values of all thermoeconomic variables depend on the component types: heat exchanger, compressor, turbine, pump, chemical reactor, and so forth. Accordingly, whether a particular value is judged to be high or low can be determined only with reference to a particular class of components. Application of fuzzy inference systems could also be useful in making such judgments (Cziesla and Tsatsaronis, 1999).

It is also important to consider the effects of contemplated design changes in one component on the performance of the remaining components. These effects may be determined either by inspection of the system flowsheets or by using a simulation program.

A cogeneration system studied by Frangopoulos (1994), Tsatsaronis and Pisa (1994), Valero et al. (1994), and von Spakovsky (1994) is used in the following case to demonstrate the application of steps 1, 2, 4, and 7. Examples of exergoeconomic evaluations and improvements of complex thermal systems are discussed by Tsatsaronis et al. (1991, 1992, 1994).

Case Study

Figure 1.8.1 presents the base-case design of a cogeneration system that develops a net power output of 30 MW and provides 14 kg/s of saturated water vapor at 20 bar. Not all data assumed for this system are realistic. The investment costs have been artificially increased to demonstrate the application of the exergoeconomic methodology to a variety of components. The optimization of a similar cogeneration system would in reality be significantly easier because the components of a gas turbine system would not be optimized individually, as done here.

The first five columns of Table 1.9.1 show relevant thermodynamic and economic data. The second column of Table 1.9.2 shows the assumed PEC for each component in the base-case design.

TABLE 1.9.1
Mass Flow Rate, Temperature, Pressure, Exergy Rate, and Cost Data for the Streams of the Cogeneration System

State	Stream	Mass Flow Rate $\dot{m}$ (kg/s)	Temperature T (K)	Pressure p (bar)	Exergy Flow Rate $\dot{E}$ (MW)	Cost Flow Rate $\dot{C}$ ($/h)	Cost per Exergy Unit c ($/GJ)
1	Air	91.28	298.1	1.01	0.000	0	0
2	Air	91.28	603.7	10.13	27.538	2756	27.80
3	Air	91.28	850.0	9.62	41.938	3835	25.40
4	Combustion products	92.92	1520.0	9.14	101.454	5301	14.51
5	Combustion products	92.92	1006.2	1.10	38.782	2026	14.51
6	Combustion products	92.92	779.8	1.07	21.752	1137	14.51
7	Combustion products	92.92	426.9	1.01	2.773	145	14.51
8	Water	14.00	298.1	20.00	0.062	0	0
9	Water	14.00	485.6	20.00	12.810	1256	27.23
10	Methane	1.64	298.1	12.00	84.994	1398	4.57
11	Power to air compressor	—	—	—	29.662	2003	18.76
12	Net power	—	—	—	30.000	2026	18.76

TABLE 1.9.2
Values of the Purchased-Equipment Costs (PEC) and the Thermoeconomic Variables for the Base Design Case ($T_3 = 850$ K; $T_4 = 1520$ K; $p_2/p_1 = 10$; $\eta_{sc} = \eta_{st} = 0.86$)[a]

Component	PEC (10^6 $)	ε (%)	$\dot{E}$ (MW)	y_D (%)	c_f ($/GJ)	c_P ($/GJ)	$\dot{C}_D$ ($/h)	$\dot{Z}$ ($/h)	$\dot{C}_D + \dot{Z}$ ($/h)	r (%)	f (%)
Combustion chamber	0.34	80.37	25.48	29.98	11.45	14.51	1050	68	1118	26.7	6.1
Gas turbine	3.74	95.20	3.01	3.54	14.51	18.76	157	753	910	29.2	82.7
Air compressor	3.73	92.84	2.12	2.50	18.76	27.80	143	753	896	48.2	84.0
HRSG	1.31	67.17	6.23	7.33	14.51	27.36	326	264	590	88.5	44.8
Air preheater	0.94	84.55	2.63	3.09	14.51	20.81	137	189	326	43.4	57.9

[a] For the overall plant, we have $\dot{C}_{P,tot} = \$3617/h$ and $\dot{C}_{L,tot} = \dot{C}_7 = \$145/h$.

These costs are obtained from the cost equations given in Appendix B of Bejan et al. (1996). The remaining direct costs, as well as the indirect costs, are estimated. The total capital investment of the cogeneration system in the base case is approximately $46 million in mid-1994 dollars. Table 7.9 in Bejan et al. (1996) summarizes the parameters and assumptions used in the economic analysis, which is based on the total revenue-requirement method (EPRI Technical Assessment Guide, 1991).

The year-by-year economic analysis results in the levelized annual costs for fuel ($\$10.4 \times 10^6$), operation and maintenance ($\$6.0 \times 10^6$), and carrying charges ($\$10.5 \times 10^6$) for a levelization time period of 10 years. These values are the corresponding levelized current-dollar costs obtained for the base case. The levelized costs are used as input data for the thermoeconomic analysis and optimization. The cost flow rates in the system are obtained by dividing the levelized annual costs by the number of hours of system operation per year.

The methodology introduced earlier will now be applied to the case study cogeneration system. The objectives are to identify the effects of the design variables on the costs and suggest values of

the design variables that would make the system more cost-effective. The key design variables—the decision variables—for the cogeneration system are the compressor pressure ratio p_2/p_1, the isentropic compressor efficiency η_{sc}, the isentropic turbine efficiency η_{st}, the temperature of the air entering the combustion chamber T_3, and the temperature of the combustion products entering the gas turbine T_4.

First Design Case

The following nominal values of the decision variables correspond to the first workable design (base-case design) developed for the cogeneration system of Figure 1.8.1 and Tables 1.9.1 and 1.9.2:

$$p_2/p_1 = 10,\ \eta_{sc} = \eta_{st} = 0.86,\ T_3 = 850\,\text{K},\ T_4 = 1520\,\text{K}.$$

The last two columns of Table 1.9.1 and the last 10 columns of Table 1.9.2 summarize the values of the thermoeconomic variables calculated for each component of the cogeneration system for the base-case design. In accordance with the methodology presented, the components are listed in order of descending value of the sum $(\dot{C}_D + \dot{Z})$.

The combustion chamber, the gas turbine, and the air compressor have the highest values of the sum $(\dot{Z} + \dot{C}_D)$ and are, therefore, the most important components from the thermoeconomic viewpoint. The low value of the exergoeconomic factor f for the combustion chamber shows that the costs associated with the combustion chamber are almost exclusively due to exergy destruction. A part of the exergy destruction in a combustion chamber can be avoided by preheating the reactants and by reducing the heat loss and the excess air, but this usually leads only to a small reduction in the cost of exergy destruction. For simplicity, we assume here that the heat loss cannot be further reduced. The excess air is determined by the desired temperature T_4 at the inlet to the gas turbine. The temperature T_4 is a key design variable for it affects both the thermodynamic performance and the investment costs of the overall system (exergy destruction and investment cost of the combustion chamber, gas turbine, air preheater, and heat-recovery steam generator (HRSG), as well as exergy loss associated with stream 7).

An increase in the heat-transfer rate within the air preheater, achieved through an increase in temperature T_3, also results in a decrease of the exergy destruction in the combustion chamber. Thus, the temperature T_3 is also a key design variable because, in addition to the combustion chamber, it affects the exergy loss associated with stream 7 as well as the performance and investment costs of the air preheater and the HRSG. Holding all other decision variables constant, the higher the temperature T_3, the smaller the average temperature difference in the air preheater and the HRSG. A decrease in the average temperature difference in these heat exchangers results in an increase in both the exergetic efficiency and the capital investment for each heat exchanger. In summary, by considering measures for reducing the high cost rate associated with the exergy destruction in the combustion chamber of the cogeneration system, two key design variables have been identified: temperatures T_3 and T_4. An increase in these temperatures reduces the $\dot{C}_D$ value for the combustion chamber and other components but increases their capital investment costs.

Turning next to the gas turbine, which has the second highest value of the sum $(\dot{Z} + \dot{C}_D)$, the relatively large value of the factor f suggests that the capital investment and O&M costs dominate. The capital investment cost of the gas turbine depends on temperature T_4, pressure ratio p_2/p_1, and isentropic efficiency η_{st}. To reduce the high $\dot{Z}$ value associated with the gas turbine, a reduction in the values of these variables should be considered.

The air compressor has the highest f value and the second highest relative cost difference r among all components. Thus, we would expect the cost-effectiveness of the entire system to improve if the $\dot{Z}$ value for the air compressor is reduced. This may be achieved by reducing the pressure ratio p_2/p_1 and/or the isentropic compressor efficiency η_{sc}.

The HRSG has the lowest exergetic efficiency and the highest r value among all the components. As the f value indicates, almost 45% of the relative cost difference is caused by the $\dot{Z}$ value in this component, with the remaining 55% caused by exergy destruction. Thus, we might conclude that a decrease of the exergy destruction in the HRSG could be cost-effective for the entire system, even if this would increase the investment cost associated with this component. The exergy destruction in the HRSG can be reduced by decreasing the values of T_6 and T_7. A decrease in the value of T_7 also results in a decrease in the exergy loss from the total system. In terms of the decision variables, increasing T_3 and/or decreasing T_4 at fixed values of the remaining decision variables would reduce temperatures T_6 and T_7.

The relatively high value of f in the air preheater suggests a reduction in the investment cost of this component. This can be achieved by decreasing T_3. It should be noted, however, that changes suggested by the evaluation of this component should only be considered if they do not contradict changes suggested by components with a larger value of $(\dot{C}_D + \dot{Z})$.

Summarizing the foregoing conclusions, the following changes in the design variables are expected to improve the cost-effectiveness of the system:

- Increase the value of T_3 as suggested by the evaluation of the combustion chamber and HRSG.
- Decrease the pressure ratio p_2/p_1 (and thus p_4/p_5) and the isentropic efficiencies η_{sc} and η_{st}, as suggested by the evaluation of the air compressor and gas turbine.
- Maintain T_4 fixed, since we get contradictory indications from the evaluation of the combustion chamber on one side and the gas turbine and HRSG on the other side.

Second Design Case

Contemplating the effects of changes made in accordance with the previous list in the values of the design variables used in the first iteration, the following new values are selected for the second design case:

$$T_3 = 870\,\text{K},\ T_4 = 1520\,\text{K}\,(\text{unchanged}),\ p_2/p_1 = 9,\ \eta_{sc} = 85\%,\ \eta_{st} = 85\%.$$

The new values of the thermoeconomic variables for each component are summarized in Table 1.9.3. In the new design, the combustion chamber, the gas turbine, and the air compressor also have the highest values of the sum $(\dot{C}_D + \dot{Z})$ and, therefore, are still the most important components from the thermoeconomic viewpoint. The high cost rate associated with the combustion chamber can be reduced by increasing the values of T_3 and T_4. In the evaluation of the cogeneration system, we should, however, consider that the value of the combustion chamber would always be the highest among the $(\dot{C}_D + \dot{Z})$ values for the components of the cogeneration system.

The gas turbine now has the highest f value. The reduction in this value from 82.7% in the base design to 81.0% in the new design is relatively small compared with a target value of below 75%. This observation suggests (1) a significant decrease in the values of η_{st} and/or p_2/p_1—that is, a decrease greater than the decrease in these variables in the previous step: from 86 to 85% and from 10% to 9%, respectively, and (2) a reduction in the value of T_4. Note that the decrease in the T_4 value contradicts the corresponding suggestion from the combustion chamber.

The high values of the exergoeconomic factor f and the relative cost difference r for the air compressor suggest a decrease in the values of the decision variables p_2/p_1 and η_{sc}.

The anticipated increase in the exergetic efficiency of the HRSG (see first iteration) was not realized because of the interdependence of the components. The reduction in the values of p_2/p_1, η_{sc}, and η_{st} for the compressor and the turbine leads to an increase in the temperature differences (and, therefore, a decrease in the exergetic efficiency) of the HRSG. Thus, the thermoeconomic evaluation of the HRSG suggests that the T_3 value increases and the T_4 value decreases.

TABLE 1.9.3
Component Data for the Second Iteration Case: (T_3=870 K; T_4=1520 K; p_2/p_1=9; η_{sc}=85%; η_{st}=85%)

Component	ε (%)	$\dot{E}_D$ (MW)	y_D (%)	c_F ($/GJ)	c_P ($/GJ)	$\dot{C}_D$ ($/h)	$\dot{Z}$ ($/h)	$\dot{C}_D+\dot{Z}$ ($/h)	r (%)	f (%)
Combustion	80.3	25.93	29.77	10.50	13.26	980	72	1052	26.3	6.8
chamber	(81.3)	(27.47)	(29.92)	(9.42)	(11.71)	(931)	(55)	(986)	(24.4)	(5.5)
Gas turbine	94.9	3.18	3.66	13.26	16.97	152	647	799	28.0	81.0
	(94.3)	(3.69)	(4.01)	(11.71)	(13.75)	(155)	(296)	(451)	(17.5)	(65.6)
Air compressor	92.1	2.34	2.69	16.97	23.96	143	546	689	41.2	79.2
	(90.5)	(2.99)	(3.25)	(13.75)	(18.38)	(148)	(324)	(472)	(33.6)	(68.7)
HRSG	66.6	6.40	7.35	13.26	25.60	305	261	566	93.1	46.1
	(67.6)	(6.10)	(6.65)	(11.71)	(23.51)	(257)	(284)	(541)	(100.7)	(52.5)
Air preheater	84.7	3.15	3.62	13.26	18.94	150	206	356	42.9	57.8
	(85.6)	(4.97)	(4.90)	(11.71)	(16.53)	(190)	(275)	(464)	(41.2)	(59.2)
Overall plant	49.1	41.01	47.09	4.57	21.80	675	1922	2597	377.0	74.0
	(46.6)	(44.79)	(48.79)	(4.57)	(19.06)	(736)	(1424)	(2160)	(317.2)	(65.9)

Note: For the overall plant in the new design case, we have $\dot{C}_{P,tot}$=$3355/h and $\dot{C}_{L,tot}=\dot{C}_7$=$157/h. The values given in parentheses refer to the third iteration case considered here (T_3=910 K; T_4=1480 K; p_2/p_1=7; η_{sc}=0.83). For the overall plant in this design case, we have $\dot{C}_{P,tot}$=$2934/h and $\dot{C}_{L,tot}=\dot{C}_7$=$167/h.

The relatively high value of f in the air preheater suggests a reduction in the T_3 value. As noted in the first iteration, however, changes suggested by the evaluation of this component should only be considered if they do not contradict changes suggested by components with a higher value of the sum $(\dot{C}_D+\dot{Z})$.

Summarizing the foregoing suggestions from the thermoeconomic evaluation of each component, the following changes in the decision variables are expected to improve the cost-effectiveness of the cogeneration system:

- Increase the value of T_3, as suggested by the evaluation of the combustion chamber and HRSG.
- Decrease the pressure ratio p_2/p_1 and the isentropic efficiencies η_{sc} and η_{st} as suggested by the evaluations of the air compressor and gas turbine.
- Decrease the temperature T_4, as suggested from the evaluations of the gas turbine and the HRSG.

Third Design Case

To illustrate the effect of the suggested changes in the decision variables on the overall costs, we use the following new set of values for the design variables:

$$T_3 = 910\,\text{K},\ T_4 = 1480\,\text{K},\ p_2/p_1 = 7,\ \eta_{sc} = 0.83,\ \text{and}\ \eta_{st} = 0.83.$$

The results from the thermoeconomic analysis for the last set of values are summarized in Table 1.9.3 by the values given in parentheses. A comparison of the corresponding values shown in Table 1.9.3 demonstrates the improvement in the cost-effectiveness of the last design case. As a result of these changes, the value of the objective function $\dot{C}_{P,tot}$ is reduced from $3355/h to $2934/h. The $\dot{C}_7$ value has increased from $157/h in the new case to $167/h in the last case. This increase, however, is outweighed by the decreases in the values of $(\dot{C}_{D,k}+\dot{Z}_k)$.

Additional Iterations

Additional iterations conducted in a similar way may be necessary to further decrease the value of the objective function and/or establish a nearly optimal design case. The cost-optimal values of the decision variables (Bejan et al., 1996) are

$$T_3 = 910\,\mathrm{K},\ T_4 = 1463.0\,\mathrm{K},\ p_2/p_1 = 5.77,\ \eta_{sc} = 0.811,\ \text{and}\ \eta_{st} = 0.845.$$

With these values, we obtain the objective function $\dot{C}_{P,tot} = \$2870/\mathrm{h}$. For the cost-optimal case, the exergetic efficiency of the overall system is 45.0%, the cost rate associated with the exergy loss is \$205/h, and the pinch temperature difference in the HRSG is 49.7 K.

1.10 ECONOMIC ANALYSIS OF THERMAL SYSTEMS

George Tsatsaronis

Generally, each company has a preferred approach for conducting an economic analysis and calculating the costs associated with the products generated in a thermal system. Here only a few general introductory comments are made. A detailed discussion of an economic analysis is given in Bejan et al. (1996).

A complete economic analysis consists of the following three steps:

Step 1: Estimate the required total capital investment.

Step 2: Determine the economic, financial, operating, and market-associated parameters for the detailed cost calculation.

Step 3: Calculate the cost of all products generated by the system.

These steps are now discussed in turn.

Estimation of Total Capital Investment

The capital needed to purchase the land, build all the necessary facilities, and purchase and install the required machinery and equipment for a system is called *fixed-capital investment*. The fixed-capital investment represents the total system cost, assuming a zero-time design, and construction period (overnight construction). The *total capital investment* is the sum of the fixed-capital investment and other outlays (e.g., startup costs, working capital, costs of licensing, research and development, as well as interest during construction).

The costs of all permanent equipment, materials, labor, and other resources involved in the fabrication, erection, and installation of the permanent facilities are the *direct costs*. The *indirect costs* (e.g., costs associated with engineering, supervision, and construction, including contractor's profit and contingencies) do not become a permanent part of the facilities but are required for the orderly completion of the project. The fixed-capital investment is the sum of all direct and indirect costs.

The *purchased equipment costs* are estimated with the aid of vendors' quotations, quotations from experienced professional cost estimators, calculations using extensive cost databases, or estimation charts. If necessary, the cost estimates of purchased equipment must be adjusted for size (with the aid of scaling exponents) and for time (with the aid of cost indices). The remaining direct costs are associated with equipment installation, piping, instrumentation, controls, electrical equipment and materials, land, civil structural and architectural work, and service facilities. These direct costs, indirect costs, and other outlays, if they cannot be estimated directly, are calculated as a percentage of the purchased equipment costs or alternatively as a percentage of the fixed-capital investment.

For more detailed information on the estimation of the total capital investment, readers should consult the references by Bejan et al. (1996), Chauvel et al. (1976), Douglas (1988), Garrett (1989), Guthrie (1969, 1970, 1974), Humphreys (1991), Perry and Chilton (1973), Peters and Timmerhaus (1991), Ulrich (1984), and Valle-Riestra (1983).

Principles of Economic Evaluation

Decisions about expenditures generally require consideration of the *time value of money*: If an amount of money P (*present value*) is deposited in an account earning i percent interest per time period, and the interest is compounded at the end of each of n time periods, the amount will grow to a *future value* (F):

$$F = P(1 + i)^n . \tag{1.10.1}$$

In engineering economy, the unit of time is usually taken as the year. If compounding occurs p times per year ($p \geq 1$) for a total number of n years ($n \geq 1$) and i is the *annual rate of return* (or *cost-of- money rate*), Equation 1.10.1 becomes

$$F = P\left(1 + \frac{i}{p}\right)^{np}. \tag{1.10.2}$$

Then the annual effective rate of return is

$$i_{\text{eff}} = \left(1 + \frac{i}{p}\right)^{p} - 1, \tag{1.10.3}$$

and the general equation for calculating the future value becomes

$$F = P(1 + i_{\text{eff}})^{n}. \tag{1.10.4}$$

When Equation 1.10.4 is used to calculate the present value (P) of a future amount (F), the term i_{eff} is called effective *discount rate*. Table 1.10.1 summarizes some basic formulas and factors used in economic analysis.

TABLE 1.10.1
Summary of Basic Formulas and Factors Used in Economic Analysis

Formula[a]	Factor
$\frac{F}{P} = \left(1 + i_{\text{eff}}\right)^{n}$	Single-payment compound-amount factor (SPCAF)
$\frac{P}{F} = \frac{1}{\left(1 + i_{\text{eff}}\right)^{n}}$	Single-payment present-worth factor or single-payment discount factor (SPDF)
$\frac{F}{A} = \frac{\left(1 + i_{\text{eff}}\right)^{n} - 1}{i_{\text{eff}}}$	Uniform-series compound-amount factor (USCAF)
$\frac{A}{F} = \frac{i_{\text{eff}}}{\left(1 + i_{\text{eff}}\right)^{n} - 1}$	Uniform-series sinking fund factor (USSFF)
$\frac{P}{A} = \frac{\left(1 + i_{\text{eff}}\right)^{n} - 1}{i_{\text{eff}}\left(1 + i_{\text{eff}}\right)^{n}}$	Uniform-series present-worth factor (USPWF)
$\frac{A}{P} = \frac{i_{\text{eff}}\left(1 + i_{\text{eff}}\right)^{n}}{\left(1 + i_{\text{eff}}\right)^{n} - 1}$	Capital-recovery factor (CRF)
$\frac{C_k}{P_0} = \frac{\left(1 + i_{\text{eff}}\right)^{n}}{\left(1 + i_{\text{eff}}\right)^{n} - 1}$	Capitalized-cost factor (CCF)
$\frac{A}{P_0} = \frac{k\left(1 - k^{n}\right)}{1 - k}\text{CRF}$	Constant-escalation levelization factor (CELF)

[a] In these expressions, cost-of-money compounding and ordinary annuities are assumed. The exponent n denotes the number of years and i_{eff} is the effective rate of return. The factor k is defined by Equation 1.10.9.

Annuities

An annuity is a series of *equal-amount* money transactions occurring at equal time intervals or periods (usually years). The most common type of annuity is the *ordinary annuity*, which involves money transactions occurring at the end of each period.

If an amount A is deposited at the end of each period in an account earning i_{eff} percent per period, the future sum F (*amount of the annuity*) accrued at the end of the nth period is

$$F = A\frac{(1+i_{\text{eff}})^n - 1}{i_{\text{eff}}}. \tag{1.10.5}$$

The term $[(1+i_{\text{eff}})^n - 1]/i_{\text{eff}}$ is called the *uniform-series compound amount factor* and the reciprocal $i_{\text{eff}} / [(1+i_{\text{eff}})^n - 1]$ is called the *uniform-series sinking fund factor.* These terms are also listed in Table 1.10.1.

The *present value* or *present worth of an annuity* (P) is defined as the amount of money that would have to be invested at the beginning of the first time interval at an effective compound rate of return per period i_{eff} to yield a total amount at the end of the last time interval equal to the amount of the annuity:

$$\frac{P}{A} = \frac{(1+i_{\text{eff}})^n - 1}{i_{\text{eff}}(1+i_{\text{eff}})^n}. \tag{1.10.6}$$

The expression on the right side of this equation is called *uniform-series present-worth factor.* The reciprocal of this factor is the *capital-recovery factor* (CRF, see Table 1.10.1), which may be used to convert the expenditures associated with capital investment into an annuity.

Cost Escalation

The *real escalation rate* of an expenditure is the annual rate of expenditure change caused by factors such as resource depletion, increased demand, and technological advances. The *nominal* (or *apparent*) *escalation rate* (r_{n}) is the total annual rate of change in an expenditure and includes the effects of both the real escalation rate (r_{r}) and the *inflation rate* (r_{i}):

$$(1+r_{\mathbf{n}}) = (1+r_{\mathbf{n}})(1+r_{\mathbf{r}}). \tag{1.10.7}$$

In the analysis of thermal systems, we often assume that all costs except fuel costs and the prices of byproducts change annually with the average inflation rate (i.e., $r_{\text{r}} = 0$). When fossil-fuel costs are expected over a long period of future years to increase on the average faster than the predicted inflation rate, a positive real escalation rate for fuel costs may be appropriate for the economic analysis of the thermal system being studied.

Levelization

Cost escalation applied to an expenditure over an n-year period results in a nonuniform cost schedule in which the expenditure at any year is equal to the previous year expenditure multiplied by $(1+r_{\text{n}})$. The *constant-escalation levelization factor* (CELF, see Table 1.10.1) is used to express the relationship between the value of the expenditure at the beginning of the first year (P_0) and an equivalent annuity (A), which is now called a *levelized value*. This levelization factor depends on

the effective annual cost-of-money rate i_{eff}, the number of time periods n, and the nominal escalation rate r_{n}:

$$\frac{A}{P_0} = \text{CELF} = \frac{k(1-k^n)}{1-k}\text{CRF}, \tag{1.10.8}$$

where

$$k = \frac{1+r_{\text{n}}}{1+i_{\text{eff}}}. \tag{1.10.9}$$

The concept of levelization is general and is defined as the use of time-value-of-money arithmetic to convert a series of varying quantities to a financially equivalent constant quantity (annuity) over a specified time interval. In exergoeconomics, the concept of levelization is applied to calculate representative values for all considered time periods: the levelized annual fuel costs, the levelized annual operating and maintenance expenses, and the levelized total cost associated with the capital investment. Among all available approaches, the *total revenue requirement method* (Bejan et al., 1996; Tsatsaronis et al., 1984, 1986) is the most detailed and most appropriate method for that purpose.

Parameters and Assumptions

To conduct a detailed economic analysis, several parameters must be specified and assumptions must be made for the entire life of the system being analyzed. These include

- The estimated total capital investment
- Current fuel cost and operating and maintenance expenditures
- Average capacity factors, average general inflation rate, and average nominal escalation rate of each expenditure
- Beginning and length of the design and construction period and of the operation period
- Allocation of investment expenditures to the individual years of design and construction
- Plant financing sources and associated required returns on capital
- Tax rates
- Depreciation method used for tax purposes
- Insurance cost

These parameters and assumptions are discussed in Bejan et al. (1996) and the EPRI Technical Assessment Guide (1991).

Calculation of the Product Costs

In a conventional economic analysis, we must identify among all product streams of a thermal system the *main product* stream; its cost is calculated from the overall cost balance. To proceed in this way, it is necessary to make assumptions with respect to the *selling prices* of all the remaining product streams, which are now *byproduct* streams.

However, when exergoeconomics is applied to a thermal system, the *total* annual levelized costs associated with owning and operating the system are apportioned among the product streams using the exergy costing approach and the auxiliary costing equations as discussed in Section 1.8. Thus, the cost associated with each product stream is calculated directly, and there is no need to use the prices of the byproducts.

1.11 EXERGOENVIRONMENTAL ANALYSIS

George Tsatsaronis

To improve the performance of an energy conversion system from the environmental point of view, we must understand the formation of environmental impact within the system. The term environment here refers to the physical environment, which is different from the thermodynamic environment used in exergy analysis. A so-called exergoenvironmental analysis rests on the notion that exergy, being the commodity of value, is the only rational basis for assigning not only monetary values but also environmental impact values to the transport of exergy and to the inefficiencies within a system (Meyer et al, 2009). We refer to this approach as *exergoenvironmental costing*.

The exergoenvironmental analysis consists of three steps: the first step is an exergetic analysis of the overall energy conversion system. In the second step, a life-cycle assessment (LCA) of (1) each relevant system component and (2) all relevant input streams to the overall system is carried out. In the third step, the environmental impact obtained from the LCA is assigned to the exergy streams in the system, *exergoenvironmental variables* are calculated, and an *exergoenvironmental evaluation* is conducted. With the aid of an exergoenvironmental analysis, the most important components with the highest environmental impact and the sources of this impact are identified. In the following the last two steps of an exergoenvironmental analysis are discussed.

An LCA is conducted for the input streams, supplied to the overall system, especially the fuel stream(s), and for the full life cycle of components. Inventories of elementary flows (i.e., consumption of natural resources and energy as well as emissions) are calculated following the guidelines of standard approaches. The environmental impact assessment can be performed with an indicator, e.g., the Eco-Indicator 99 (Goedkoop and Spriensma, 2000). This is based on the definition of three damage categories: human health, ecosystem quality, and natural resources. The numbers obtained from the three categories are aggregated and the result is expressed as Eco-Indicator points (pts), where a higher damage corresponds to a higher Eco-Indicator value.

In analogy to the assignment of costs to exergy streams in exergoeconomics, an environmental impact rate and an environmental impact per unit of exergy are assigned to each exergy stream in an exergoenvironmental evaluation. The *environmental impact rate* is the environmental impact expressed, for example, in Eco-Indicator points per unit of time of system operation (Pts/s or mPts/s). The specific (exergy-based) environmental impact (also called the specific *environmental cost*) is the average environmental impact associated with the production of the stream per unit of exergy of the same stream [Pts/(kJ exergy) or mPts/(MJ exergy)].

The further procedure in an exergoenvironmental analysis is identical with that in an exergoeconomic analysis: We only need to replace the term cost (in the exergoeconomic analysis) by the term environmental impact (in the exergoenvironmental analysis). Thus, we formulate an environmental impact balance for each system component, we apply the same auxiliary equations as in exergoeconomics, and we calculate the average environmental impact associated with fuel, product, and exergy destruction within each system component.

Hence, the exergoenvironmental analysis is the only approach that allows engineers to calculate the environmental impact associated with the real thermodynamic inefficiencies (exergy destruction and exergy loss) in the same way as the exergoeconomic analysis provides the cost associated with the same inefficiencies. For a detailed presentation of exergoenvironmental analysis, readers may refer to the following references: Meyer et al. (2009), Petrakopoulou et al. (2011a,b), Tsatsaronis (2011), and Tsatsaronis and Morosuk (2015).

1.12 ADVANCED EXERGY-BASED METHODS

George Tsatsaronis

The exergetic, exergoeconomic, and exergoenvironmental analyses discussed so far provide very useful information (not obtainable by other means) for the improvement of an energy conversion system from the viewpoints of thermodynamics, economics, and ecology. However, these approaches do not fully consider (1) the mutual interdependencies among the system components and (2) the real potential for improving each component of the energy conversion system (Tsatsaronis, 1999b). Advanced exergy-based analyses enable these evaluations.

The interactions among the components of the same system can be estimated and the quality of the conclusions obtained from an exergoeconomic and exergoenvironmental evaluation can be improved by splitting

- The exergy destruction within each (important) system component, the cost of exergy destruction, and the environmental impact associated with exergy destruction
- The component-related investment cost, and the environmental impact associated with manufacturing and installing each important system component

into endogenous/exogenous and avoidable parts (e.g., Tsatsaronis, 2008; Tsatsaronis and Morosiuk, 2008; Tsatsaronis, 2011). We call the analyses based on such splitting *advanced* (*exergetic*, *exergoeconomic*, or *exergoenvironmental*) *analyses*.

Endogenous (exergy destruction, capital investment cost, and construction-of component-related environmental impact) is that part of the considered variable (associated with a system component) that is obtained when all other components operate ideally and the component being considered operates with the same efficiency as in the real system. The *exogenous* part of the variable is the difference between the total value of the variable for the same component and the endogenous part.

The *unavoidable* exergy destruction, cost, or environmental impact cannot be further reduced due to technological limitations such as availability and cost of materials and manufacturing methods. The difference between total and unavoidable values for a component is the *avoidable value*. In the improvement efforts, only the avoidable values should be considered.

By combining the two splitting approaches, we obtain the following:

- The *avoidable endogenous* part of a variable that can be reduced by improving the considered component from the exergetic, economic, and environmental points of view.
- The *avoidable exogenous* part of the same variable that can be reduced by a structural improvement of the overall system, or by improving the remaining components, and always, of course, by improving the component being considered.

These variables are used in the iterative improvement of the system and in establishing research priorities.

The advanced exergy-based methods significantly improve the quality of the obtained information and the optimization procedures. For a detailed presentation of the advanced exergy-based methods and for applications, readers may refer to Boyano et al. (2012), Morosuk and Tsatsaronis (2009a,b, 2011, 2012), Morosuk et al. (2012), Petrakopoulou et al. (2011a,b, 2012, 2014), Tsatsaronis (2008, 2011), Tsatsaronis and Morosuk (2008, 2009, 2015), and Tsatsaronis and Park (2002).

REFERENCES

Ahrendts, J. Reference states, *Int. J. Energy* 5: 667–677, 1980.

ASHRAE Handbook 1993 Fundamentals. American Society of Heating, Refrigerating, and Air Conditioning Engineers, Atlanta, GA, 1993.

ASME Steam Tables, 6th ed., ASME Press, Fairfield, NJ, 1993.
Baasel, W.D., *Preliminary Chemical Engineering Plant Design*, 2nd ed., Van Nostrand Reinhold, New York, 1990.
Bejan, A., Tsatsaronis, G., and Moran, M., *Thermal Design and Optimization*, John Wiley & Sons, New York, 1996.
Biegler, L.T., Chemical process simulation, *Chem. Eng. Progr.*, 85: 50–61, 1989.
Bird, R.B., Stewart, W.E., and Lightfoot, E.N., *Transport Phenomena*, 2nd ed., John Wiley & Sons, New York, 2002.
Bolz, R.E. and Tuve, G.L., Eds., *Handbook of Tables for Applied Engineering Science*, 2nd ed., Chemical Rubber Company, Cleveland, 1973.
Bornakke, C. and Sonntag, R.E., *Tables of Thermodynamic and Transport Properties*, John Wiley & Sons, New York, 1996.
Boyano, A., Blanco-Marigota, A.M., Morosuk, T., and Tsatsaronis, G., Steam methane reforming system for hydrogen production: Advanced exergetic analysis, *Int. J. Thermodynamics*, 15(1): 1–9, 2012.
Chauvel, A. et al., *Manual of Economic Analysis of Chemical Processes*, McGraw-Hill, New York, 1976.
Cooper, H.W. and Goldfrank, J.C., B-W-R constants and new correlations, *Hydrocarbon Proc.* 46(12): 141–146, 1967.
Cziesla, F. and Tsatsaronis, G., Iterative exergoeconomic evaluation and improvement of thermal power plants using fuzzy inference systems, *Proceedings of International Conference on Efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy Systems,* Tokyo, June 8–10: 380–385, 1999.
Douglas, J.M., *Conceptual Design of Chemical Processes*, McGraw-Hill, New York, 1988.
Edgar, T.F. and Himmelblau, D.M., *Optimization of Chemical Processes*, McGraw-Hill, New York, 1988.
EPRI Technical Assessment Guide (TAG™), Electric Power Research Institute, TR-100281, 3, Rev. 6, 1991.
Frangopoulos, C.A., Application of the thermoeconomic functional approach to the CGAM problem, *Int. J. Energy*, 19: 323–342, 1994.
Garrett, D.E., *Chemical Engineering Economics*, Van Nostrand Reinhold, New York, 1989.
Goedkopp, M. and Spriensma, R., The Eco-Indicator 99: A damage oriented method for life cycle impact assessment, Methodology Report, Amersfoort, The Netherlands (http://www.pre.nl), 2000.
Gray, D.E., Ed., *American Institute of Physics Handbook,* McGraw-Hill, New York, 1972.
Guthrie, K.M., Data and techniques for preliminary capital cost estimating, *Chem. Eng.*, 24: 114–142, 1969.
Guthrie, K.M., Capital and operating costs for 54 chemical processes, *Chem. Eng.*, 15: 140–156, 1970.
Guthrie, K.M., *Process Plant Estimating, Evaluation and Control*, Craftsman, Solana Beach, CA, 1974.
Haar, L., Gallagher, J.S., and Kell, G.S., *NBS/NRC Steam Tables*, Hemisphere, New York, 1984.
Handbook of Chemistry and Physics, annual ed., CRC Press, Boca Raton, FL.
Humphreys, K.K., *Jelen's Cost and Optimization Engineering*, 3rd ed., McGraw-Hill, New York, 1991.
JANAF Thermochemical Tables, 3rd ed., American Chemical Society and the American Institute of Physics for the National Bureau of Standards, Washington, DC, 1986.
Jones, J.B. and Dugan, R.E., *Engineering Thermodynamics*, Prentice-Hall, Englewood Cliffs, NJ, 1996.
Keenan, J.H., Chao, J., and Kaye, J., *Gas Tables International Version*, 2nd ed., John Wiley & Sons, New York (1980, English Units; 1983, SI Units).
Keenan, J.H., Keyes, F.G., Hill, P.G., and Moore, J.G., *Steam Tables*, John Wiley & Sons, New York, 1978.
Knacke, O., Kubaschewski, O., and Hesselmann, K., *Thermochemical Properties of Inorganic Substances*, 2nd ed., Springer-Verlag, Berlin, 1991.
Kotas, T.J., *The Exergy Method of Thermal Plant Analysis*, Krieger, Melbourne, FL, 1995.
Lazzaretto, A. and Tsatsaronis, G., *SPECO*: A systematic and general methodology for calculating efficiencies and costs in thermal systems, *Int. J. Energy* 31:1257–1289, 2006.
Lee, B.I. and Kessler, M.G., A generalized thermodynamic correlation based on three-parameter corresponding states, *AIChE J.*, 21: 510–527, 1975.
Lee, Y.D., Ahn, K.Y., Morosuk, T., and Tsatsaronis, G., Environmental impact assessment of a solid-oxide fuel-cell-based combined-heat-and-power-generation system, *Int. J. Energy*, 79: 455–466, 2015.
Liley, P.E., Reid, R.C., and Buck, E., Physical and chemical data, in *Perrys' Chemical Engineers Handbook*, Perry, R.H. and Green, D.W., Eds, 6th ed., McGraw-Hill, New York, 1984.
Liley, P.E., Thermodynamic properties of substances, in *Marks' Standard Handbook for Mechanical Engineers*, Avallone, E.A. and Baumeister, T., Eds, 9th ed., McGraw-Hill, New York, 1987.

Lin, L. and Tsatsaronis, G., Cost optimization of an advanced IGCC power plant concept design, in *Thermodynamics and the Design, Analysis and Improvement of Energy Systems*, Richter, H.J., Ed., AES Vol. 10, ASME, New York: 156–166, 1993.

Linnhoff, B. et al., *A User Guide on Process Integration for the Efficient Use of Energy*, Institution of Chemical Engineers, Rugby, Warks, UK, 1982.

Meyer, L., Tsatsaronis, G., Buchgeister, J., and Schebek, L. Exergoenvironmental analysis for the evaluation of the environmental impact of energy conversion systems, *Int. J. Energy* 34: 75–89, 2009.

Moran, M.J., *Availability Analysis—A Guide to Efficient Energy Use*, ASME Press, New York, 1989.

Moran, M.J. and Shapiro, H.N., *Fundamentals of Engineering Thermodynamics*, 4th ed., John Wiley & Sons, New York, 2000a.

Moran, M.J. and Shapiro, H.N., IT: *Interactive Thermodynamics*, Computer software to accompany *Fundamentals of Engineering Thermodynamics*, 4th ed., Developed by Intellipro Inc., John Wiley & Sons, New York, 2000b.

Morosuk, T. and Tsatsaronis, G., Advanced exergy analysis for chemically reacting systems—Application to a simple open gas-turbine system, *Int. J. Thermodyn.*, 12(3): 105–111, 2009a.

Morosuk, T. and Tsatsaronis, G., Advanced exergetic evaluation of refrigeration machines using different working fluids, *Int. J. Energy*, 34: 2248–2258, 2009b.

Morosuk, T. and Tsatsaronis, G., Comparative evaluation of LNG-based cogeneration systems using advanced exergetic analysis, *Int. J. Energy*, 36(3): 3771–3778, 2011.

Morosuk, T. and Tsatsaronis, G., 3-D exergy-based methods for improving energy-conversion systems, *Int. J. Thermodyn.*, 15(4): 201–213, 2012.

Morosuk, T., Tsatsaronis, G., and Zhang, C., Conventional thermodynamic and advanced exergetic analysis of a refrigeration machine using a Voorhees' compression process, *Energy Convers. Manag.*, 60: 143–151, 2012.

Moran, M.J., Shapiro, H.N., Boettner, D.D., and Bailey, M.B., *Fundamentals of Engineering Thermodynamics*, 8th ed., John Wiley & Sons, New York, 2014.

Obert, E.F., *Concepts of Thermodynamics*, McGraw-Hill, New York, 1960.

Papalambros, P. Y. and Wilde, D. J., *Principles of Optimal Design-Modeling and Computation*, Cambridge University Press, Cambridge, UK, 1988.

Perry, J.H. and Chilton, C.H., *Chemical Engineer's Handbook*, 5th ed., McGraw-Hill, New York, 1973.

Peters, M.S. and Timmerhaus, K.D., *Plant Design and Economics for Chemical Engineers*, 4th ed., McGraw-Hill, New York, 1991.

Petrakopoulou, F., Boyano, A., Cabrera, M. and Tsatsaronis, G. Exergoeconomic and exergoenvironmental analyses of a combined cycle power plant with chemical looping technology, *Int. J. Greenh. Gas Control*, 5(3): 475–482, 2011a.

Petrakopoulou, F., Tsatsaronis, G., Boyano, A., and Morosuk, T., Exergoeconomic and exergoenvironmental evaluation of power plants including CO2 capture, *Chem. Eng. Res. Design*, 89(9): 1461–1469, 2011b.

Petrakopoulou, F., Tsatsaronis, G., Boyano, A., Morosuk, T., and Paitazoglou, C., Environmental evaluation of a power plant using conventional and advanced exergy-based methods, *Int. J. Energy*, 45(1): 23–30, 2012.

Petrakopoulou, F., Tsatsaronis, G., and Morosuk, T., Assessment of a power plant with CO2 capture using an advanced exergoenvironmental analysis, *J. Energy Resour. Technol.*, 136(2): 1–10, art. no. 022001, 2014.

Preston-Thomas, H., The international temperature scale of 1990 (ITS 90), *Metrologia*, 27: 3–10, 1990.

Reid, R.C. and Sherwood, T.K., *The Properties of Gases and Liquids*, 2nd ed., McGraw-Hill, New York, 1996.

Reid, R.C., Prausnitz, J.M., and Poling, B.E., *The Properties of Gases and Liquids*, 4th ed., McGrawHill, New York, 1987.

Reklaitis, G.V., Ravindran, A., and Ragsdell, K.M., *Engineering Optimization*, John Wiley & Sons, New York, 1983.

Reynolds, W.C., Thermodynamic properties in SI—Graphs, tables and computational equations for 40 substances, Department of Mechanical Engineering, Stanford University, Palo Alto, CA, 1979.

Sama, D.A., The use of the second law of thermodynamics, in *Proceedings of International Conference on ENSEC '93*J, Szargut, Z., Kolenda, G., Tsatsaronis, and A. Ziebik, Eds, ASME, New York: 53–76, 1993.

Stephan, K., Tables, *in Dubbel Handbook of Mechanical Engineering*, Beitz, W. and Kuttner, K.-H., Eds, Springer-Verlag, London, 1994.

Szargut, J., Morris, D.R., and Steward, F.R., *Exergy Analysis of Thermal, Chemical, and Metallurgical Processes*, Hemisphere, New York, 1988.

Tsatsaronis, G., Thermoeconomic analysis and optimization of energy systems, *Prog. Energy Combust. Sci*., 19: 227–257, 1993.

Tsatsaronis, G., Combination of exergetic and economic analysis in energy-conversion processes, in *Energy Economics and Management in Industry, Vol. 1* (Proceedings of the European Congress, Algarve, Portugal, April 2–5, 1984, Pergamon Press, Oxford, UK): 151–157, 1984.

Tsatsaronis, G. (Ed.), Invited papers on exergoeconomics, *Int. J. Energy* (Special Issue) 19(3): 279–257, 1994.

Tsatsaronis, G., Design optimization using exergoeconomics, in *Thermodynamic Optimization of Complex Energy Systems*, Bejan A. and Mamut E., Eds, Kluwer Academic Publishers, Lancaster, UK: 101–115, 1999a.

Tsatsaronis, G., Strengths and limitations of exergy analysis, in *Thermodynamic Optimization of Complex Energy Systems*, Bejan A. and Mamut E., Eds, Kluwer Academic Publishers, Lancaster, UK: 93–100, 1999b.

Tsatsaronis, G., Definitions and nomenclature in exergy analysis and exergoeconomics, *Int. J. Energy*, 32(4): 249–253, 2007.

Tsatsaronis, G., Recent developments in exergy analysis and exergoeconomics, *Int. J. Exergy,* 5(5–6): 489–499, 2008.

Tsatsaronis, G., Exergoeconomics and exergoenvironmental Analysis, in *Thermodynamics and the Destruction of Resources*, Bakshi, B.R., Gutowski, T.G., and Sekulic, D.P., Eds, Cambridge University Press, New York: 377–401, (Chapter 15), 2011.

Tsatsaronis, G., Krause, A., Lin, L., and Tawfik, T., Thermoeconomic evaluation of the design of a pressurized fluidized-bed hydroretorting plant, final report prepared for the Institute of Gas Technology and the Department of Energy, DE-AC21-87MC11089, 1992.

Tsatsaronis, G., Lin, L., Pisa, J., and Tawfik, T., Thermoeconomic design optimization of a KRW-based IGCC power plant, final report prepared for the U.S. Department of Energy, Morgantown Energy Technology Center, DE-FC21-98MC26019, 1991.

Tsatsaronis, G., Lin, L., Pisa, J., and Tawfik, T., Optimization of an IGCC power plant. Parts I and II, in *Thermodynamics and the Design, Analysis and Improvement of Energy Systems*, R.F. Boehm et al., Eds, AES Vol. 27, ASME, New York: 37–67, 1992.

Tsatsaronis, G., Lin, L., and Tawfik, T.J., Exergoeconomic evaluation of a KRW-based IGCC power plant, *Eng. Gas Turbines Power*, 116: 300–306, 1994.

Tsatsaronis, G. and Morosuk, T., A general exergy-based method for combining a cost analysis with an environmental impact analysis, Parts I and II, *Proceedings of the ASME International Mechanical Engineering Congress and Exposition, Boston, MA, November 2–6, 2008*: IMECE2008-67218; IMECE2008-67219, 2008.

Tsatsaronis, G. and Morosuk, T., Advances in exergy-based methods for improving energy conversion systems, in *Optimization Using Exergy-Based Methods and Computational Fluid Dynamics*, Tsatsaronis, G. and Boyano, A. Eds, Clausthal-Zellerfeld: Papierflieger Verlag, Clausthal-Zellerfeld, Germany: 1–10, 2009.

Tsatsaronis, G. and Morosuk, T., Understanding the formation of costs and environmental impacts using exergy-based methods, in *Energy Security and Development*, Chapter 18, B.S. Reddy and S. Ulgiati, Eds, Springer, New Delhi: 271–291, 2015.

Tsatsaronis, G. and Park, M.-H., On avoidable and unavoidable exergy destructions and investment costs in thermal systems, *Energy Conversion and Management*, 43: 1259–1270, 2002.

Tsatsaronis, G. and Pisa, J., Exergoeconomic evaluation and optimization of energy systems, *Int. J. Energy* 19: 287–321, 1994.

Tsatsaronis, G. and Winhold, M., Thermoeconomic analysis of power plants. EPRI AP-3651, RP 2029–8. Final Report. Electric Power Research Institute, Palo Alto, CA, 1984.

Tsatsaronis, G. and Winhold, M., Exergoeconomic analysis and evaluation of energy conversion plants. *Int. J. Energy*, 10: 69–94, 1985.

Tsatsaronis, G., Winhold, M., and Stojanoff, C.G., Thermoeconomic analysis of a gasification-combined-cycle power plant. EPRI AP-4734, RP 2029-8. Final report. Electric Power Research Institute, Palo Alto, CA, 1986.

Ulrich, G.D., *A Guide to Chemical Engineering Process Design and Economics*, John Wiley & Sons, New York, 1984.

Valero, A., Lozano, M.A., Serra, L., and Torres, C., Application of the exergetic cost theory to the CGAM problem, *Int. J. Energy*, 19: 365–381, 1994.

Valle-Riestra, F.J., *Project Evaluation in the Chemical Process Industries*, McGraw-Hill, New York, 1983.

Van Wylen, G.J., Sonntag, R.E., and Bornakke, C., *Fundamentals of Classical Thermodynamics*, 4th ed., John Wiley & Sons, New York, 1994.

von Spakovksy, M.R., Application of engineering functional analysis to the analysis and optimization of the CGAM problem, *Int. J. Energy*, 19: 343–364, 1994.

Wark, K., *Thermodynamics*, 4th ed., McGraw-Hill, New York, 1983.

Zemansky, M.W., Thermodynamic symbols, definitions, and equations, in *American Institute of Physics Handbook*, D.E. Gray, Ed., McGraw-Hill, New York, 1972.

2 Fluid Mechanics

Stanley A. Berger, Stuart W. Churchill, J. Paul Tullis, Blake Paul Tullis, Frank M. White, John C. Leylegian, John C. Chen, Anoop K. Gupta, Raj P. Chhabra, Thomas F. Irvine, Jr., and Massimo Capobianchi

CONTENTS

2.1 Fluid Statics 115
Equilibrium of a Fluid Element 115
Hydrostatic Pressure 115
Manometry 116
Hydrostatic Forces on Submerged Objects 116
Hydrostatic Forces in Layered Fluids 119
Buoyancy 119
Stability of Submerged and Floating Bodies 120
Pressure Variation in Rigid-Body Motion of a Fluid 121
Further Information 122
2.2 Equations of Motion and Potential Flow 123
Integral Relations for a Control Volume 123
Reynolds Transport Theorem 123
Conservation of Mass 123
Conservation of Momentum 124
Conservation of Energy 125
Differential Relations for Fluid Motion 127
Mass Conservation–Continuity Equation 127
Momentum Conservation 127
Analysis of Rate of Deformation 129
Relationship between Forces and Rate of Deformation 129
The Navier–Stokes Equations 131
Energy Conservation: The Mechanical and Thermal Energy Equations 131
Boundary Conditions 133
Vorticity in Incompressible Flow 133
Stream Function 135
Inviscid Irrotational Flow: Potential Flow 136
Further Information 138
2.3 Similitude: Dimensional Analysis and Data Correlation 139
Dimensional Analysis 139
Correlation of Experimental Data and Theoretical Values 146
References 153
2.4 Hydraulics of Pipe Systems 154
Basic Computations 154
Fluid Friction 155

Minor Losses 157
Pressure Class Guidelines 158
External Loads 158
Limiting Velocities 160
Valves 160
Control Valves 160
Check Valves 164
Air Valves 165
Pump Selection 165
Filling and Flushing 168
System Monitoring and Control 168
References 169
Further Information 169
2.5 Open Channel Flow 170
Definition 170
Uniform Flow 170
Critical Flow 170
Hydraulic Jump 172
Weirs 173
Gradually Varied Flow 174
Nomenclature 177
References 178
2.6 External Incompressible Flows 179
Introduction and Scope 179
Boundary Layers 179
Drag 184
Lift 187
Boundary Layer Control 188
Computation vs. Experiment 189
References 190
Further Information 190
2.7 Compressible Flow 191
Introduction 191
One-Dimensional Flow 192
Normal Shock Wave 193
One-Dimensional Flow with Heat Addition 196
One-Dimensional Flow with Area Change 198
Two-Dimensional Supersonic Flow 203
References 208
Further Information 208
2.8 Multiphase Flow 209
Introduction 209
Fundamentals 209
Gas–Liquid Two-Phase Flow 212
Flow Regimes 212
Void Fractions 214
Pressure Drop 217
Gas–Solid, Liquid–Solid Two-Phase Flows 217
Flow Regimes 218
Minimum Fluidization 218
Terminal Slip Velocity 220

Pneumatic Conveying 220
References 222
2.9 Non-Newtonian Flows 223
Introduction 223
Classification of Non-Newtonian Behavior 223
Types of Time-Independent Fluids and Constitutive Equations 228
Rheological Equation of State 229
Fully Developed Laminar Flow in a Tube 232
Laminar Flow in Noncircular Ducts 239
Turbulent Flow in Circular and Noncircular Ducts 240
Effect of Viscoelasticity 243
References 246
Further Information 248

2.1 FLUID STATICS

Stanley A. Berger

Equilibrium of a Fluid Element

If the sum of the external forces acting on a fluid element is zero, the fluid will be either at rest or moving as a solid body—in either case, we say the fluid element is in equilibrium. In this section we consider fluids in such an equilibrium state. For fluids in equilibrium, the only internal stresses acting will be normal forces, since the shear stresses depend on velocity gradients, and all such gradients, by the definition of equilibrium, are zero. If one then carries out a balance between the normal surface stresses and the body forces, assumed proportional to volume or mass, such as gravity, acting on an elementary prismatic fluid volume, the resulting equilibrium equations, after shrinking the volume to zero, show that the normal stresses at a point are the same in all directions, and since they are known to be negative, this common value is denoted by $-p$, p being the pressure.

Hydrostatic Pressure

If we carry out an equilibrium of forces on an elementary volume element $dxdydz$, the forces being pressures acting on the faces of the element and gravity acting in the $-z$ direction, we obtain

$$\frac{\partial p}{\partial x} = \frac{\partial p}{\partial y} = 0, \quad \text{and} \quad \frac{\partial p}{\partial z} = -\rho g = -\gamma \tag{2.1.1}$$

The first two of these imply that the pressure is the same in all directions at the same vertical height in a gravitational field. The third, where γ is the specific weight, shows that the pressure increases with depth in a gravitational field, the variation depending on $\rho(z)$. For homogeneous fluids, for which ρ = constant, this last equation can be integrated immediately, yielding

$$p_2 - p_1 = -\rho g\left(z_2 - z_1\right) = -\rho g\left(h_2 - h_1\right) \tag{2.1.2}$$

or

$$p_2 + \rho g h_2 = p_1 + \rho g h_1 = \text{constant} \tag{2.1.3}$$

where h denotes the elevation. These are the equations for the hydrostatic pressure distribution.

When applied to problems where a liquid, such as the ocean, lies below the atmosphere, with a constant pressure p_{atm}, h is usually measured (positive) downward from the ocean/atmosphere interface and p at any distance h below this interface differs from p_{atm} by an amount

$$p - p_{\text{atm}} = \rho g h \tag{2.1.4}$$

Pressures may be given either as *absolute pressure*, pressure measured relative to absolute vacuum, or *gauge pressure*, pressure measured relative to atmospheric pressure.

Manometry

The hydrostatic pressure variation may be employed to measure pressure differences in terms of heights of liquid columns—such devices are called manometers and are commonly used in wind tunnels and a host of other applications and devices. Consider, for example the U-tube manometer shown in Figure 2.1.1 filled with liquid of specific weight γ, the left leg open to the atmosphere and the right to the region whose pressure p is to be determined. In terms of the quantities shown in the figure, in the left leg

$$p_0 - \rho g h_2 = p_{\text{atm}} \tag{2.1.5a}$$

and in the right leg

$$p_0 - \rho g h_1 = p \tag{2.1.5b}$$

the difference being

$$p - p_{\text{atm}} = -\rho g \left(h_1 - h_2\right) = -\rho g d = -\gamma d \tag{2.1.6}$$

which determines p in terms of the height difference $d = h_1 - h_2$ between the levels of the fluid in the two legs of the manometer.

Hydrostatic Forces on Submerged Objects

The force acting on a submerged object due to the hydrostatic pressure is given by

$$\boldsymbol{F} = \iint p \; \boldsymbol{dA} = \iint p \cdot n \; \boldsymbol{dA} = \iint \rho g h \; \boldsymbol{dA} + p_0 \iint \boldsymbol{dA} \tag{2.1.7}$$

where h is the variable vertical depth of the element $\boldsymbol{dA}$ and p_0 is the pressure at the surface. In turn we consider plane and nonplanar surfaces.

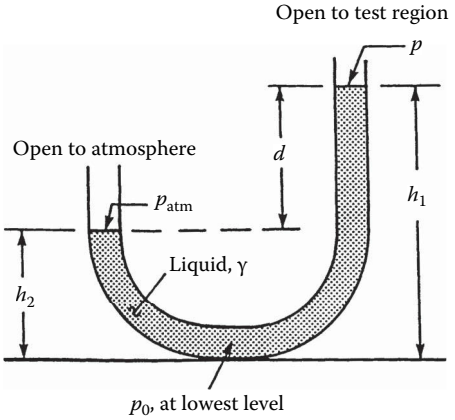


FIGURE 2.1.1 U-tube manometer.

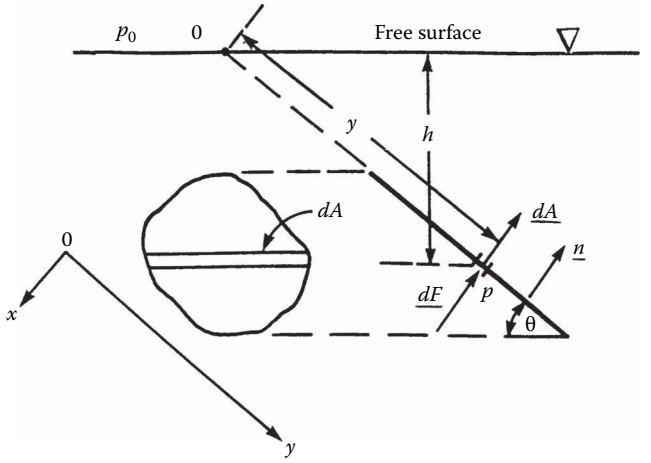


FIGURE 2.1.2 Hydrostatic force on a plane surface.

Forces on Plane Surfaces

Consider the planar surface A at an angle θ to a free surface shown in Figure 2.1.2. The force on one side of the planar surface, from Equation 2.1.7, is

$$\boldsymbol{F} = \rho g \boldsymbol{n} \iint_A h \; dA + p_0 A \boldsymbol{n} \tag{2.1.8}$$

but $h = y \sin\theta$, so

$$\iint_A h \; dA = \sin\theta \iint_A y \; dA = y_c A \sin\theta = h_c A \tag{2.1.9}$$

where the subscript c indicates the distance measured to the centroid of the area A. The total force (on one side) is then

$$\boldsymbol{F} = \gamma h_c \boldsymbol{A} + p_0 \boldsymbol{A} \tag{2.1.10}$$

Hence, the magnitude of the force is independent of the angle θ and is equal to the pressure at the centroid, $\gamma h_c + p_0$, times the area. If we use gauge pressure, the term $p_0 A$ in Equation 2.1.10 is dropped.

Since p is not evenly distributed over A, but varies with depth, F does not act through the centroid. The point of action of $\boldsymbol{F}$, called the *center of pressure*, can be determined by considering moments in Figure 2.1.2. The moment of the hydrostatic force acting on the elementary area dA about the axis perpendicular to the page passing through the point 0 on the free surface is

$$y \; dF = y(\gamma \, y \sin\theta \, dA) \; = \gamma \; y^2 \sin\theta \, dA \tag{2.1.11}$$

so if y_{cp} denotes the distance to the center of pressure,

$$y_{cp} F = \gamma \sin\theta \iint y^2 dA = \gamma \sin\theta I_x \tag{2.1.12}$$

where I_x is the moment of inertia of the plane area with respect to the axis formed by the intersection of the plane containing the planar surface and the free surface (say $0x$). Dividing by $F = \gamma h_c A = \gamma \; y_c \sin\theta A$ gives

$$y_{\text{cp}} = \frac{I_x}{y_c A} \tag{2.1.13}$$

By using the parallel axis theorem, $I_x = I_{xc} + Ay_c^2$, where I_{xc} is the moment of inertia with respect to an axis parallel to $0x$ passing through the centroid, Equation 2.1.13 becomes

$$y_{\text{cp}} = y_c + \frac{I_{xc}}{y_c A} \tag{2.1.14}$$

which shows that, in general, the center of pressure lies below the centroid.

Similarly, we find x_{cp} by taking moments about the y axis, specifically

$$x_{\text{cp}} F = \gamma \sin\theta \iint xy \; dA = \gamma \sin\theta \, I_{xy} \tag{2.1.15}$$

or

$$x_{\text{cp}} = \frac{I_{xy}}{y_c A} \tag{2.1.16}$$

where I_{xy} is the product of inertia with respect to the x and y axes. Again, by the parallel axis theorem, $I_{xy}=I_{xyc}+Ax_cy_c$, where the subscript c denotes the value at the centroid, so Equation 2.1.16 can be written

$$x_{\text{cp}} = x_c + \frac{I_{xyc}}{y_c A} \tag{2.1.17}$$

This completes the determination of the center of pressure (x_{cp}, y_{cp}). Note that if the submerged area is symmetrical with respect to an axis passing through the centroid and parallel to either the x or y axes that $I_{xyc}=0$ and $x_{\text{cp}}=x_c$; also that as y_c increases, $y_{\text{cp}} \to y_c$.

Centroidal moments of inertia and centroidal coordinates for some common areas are shown in Figure 2.1.3.

Forces on Curved Surfaces

On a curved surface the forces on individual elements of area differ in direction so a simple summation of them is not generally possible, and the most convenient approach to calculating the pressure force on the surface is by separating it into its horizontal and vertical components.

A free-body diagram of the forces acting on the volume of fluid lying above a curved surface together with the conditions of static equilibrium of such a fluid column leads to the results that:

1. The horizontal components of force on a curved submerged surface are equal to the forces exerted on the planar areas formed by the projections of the curved surface onto vertical planes normal to the directions of these components, the lines of action of these forces being calculated as described earlier for planar surfaces; and
2. The vertical component of force on a curved submerged surface is equal in magnitude to the weight of the entire column of fluid lying above the curved surface, and acts through the center of mass of this volume of fluid.

Since the three components of force, two horizontal and one vertical, calculated as above, need not meet at a single point, there is, in general, no single resultant force. They can, however, be combined and considered to be equivalent to a single force at any arbitrary point of application together with a moment about that point.

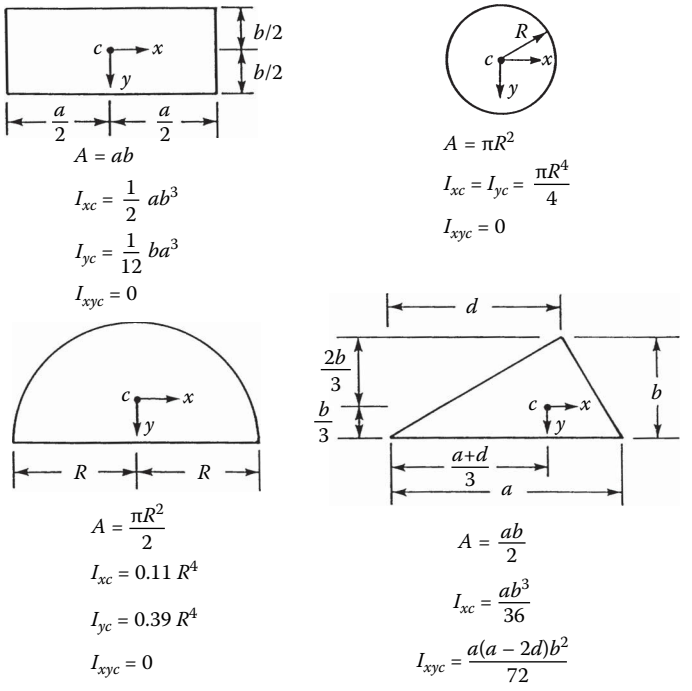


FIGURE 2.1.3 Centroidal moments of inertia and coordinates for some common areas.

Hydrostatic Forces in Layered Fluids

All of the above results which employ the linear hydrostatic variation of pressure are valid only for homogeneous fluids. If the fluid is heterogeneous, consisting of individual layers each of constant density, then the pressure varies linearly with a different slope in each layer and the preceding analyses must be remedied by computing and summing their separate contributions to the forces and moments.

Buoyancy

The same principles used above to compute hydrostatic forces can be used to calculate the net pressure force acting on completely submerged or floating bodies. These laws of buoyancy, the principles of Archimedes, are that:

1. A completely submerged body experiences a vertical upward force equal to the weight of the displaced fluid; and
2. A floating or partially submerged body displaces its own weight in the fluid in which it floats (i.e., the vertical upward force is equal to the body weight).

The line of action of the buoyancy force in both (1) and (2) passes through the centroid of the displaced volume of fluid; this point is called the *center of buoyancy*. (This point need not correspond to the center of mass of the body, which could have nonuniform density. In the above it has been assumed that the displaced fluid has a constant γ. If this is not the case, such as in a layered fluid, the magnitude of the buoyant force is still equal to the weight of the displaced fluid, but the line of action of this force passes through the center of gravity of the displaced volume, not the centroid.)

If a body has a weight exactly equal to that of the volume of fluid it displaces, it is said to be *neutrally buoyant* and will remain at rest at any point where it is immersed in a (homogeneous) fluid.

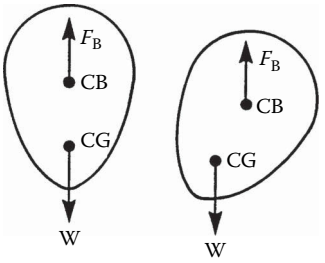


FIGURE 2.1.4 Stability for a submerged body: (left) equilibrium position and (right) slightly rotated.

Stability of Submerged and Floating Bodies

Submerged Body

A body is said to be in stable equilibrium if when given a slight displacement from the equilibrium position the forces thereby created tend to restore it back to its original position. The forces acting on a submerged body are the buoyancy force, F_B, acting through the center of buoyancy, denoted by CB, and the weight of the body, W, acting through the center of gravity denoted by CG (see Figure 2.1.4). We see from Figure 2.1.4 that if the CB lies above the CG, a rotation from the equilibrium position creates a restoring couple which will rotate the body back to its original position—thus, this is a *stable* equilibrium situation. The reader will readily verify that when the CB lies below the CG, the couple that results from a rotation from the vertical increases the displacement from the equilibrium position—thus, this is an *unstable* equilibrium situation.

Partially Submerged Body

The stability problem is more complicated for floating bodies because as the body rotates the location of the center of buoyancy may change. To determine stability in these problems requires that we determine the location of the *metacenter*. This is done for a symmetric body by tilting the body through a small angle $\Delta\theta$ from its equilibrium position and calculating the new location of the center of buoyancy CB′; the point of intersection of a vertical line drawn upward from CB′ with the line of symmetry of the floating body is the metacenter, denoted by M in Figure 2.1.5, and it is independent of $\Delta\theta$ for small angles. If M lies above the CG of the body, we see from Figure 2.1.5 that rotation of the body leads to a restoring couple, whereas M lying below the CG leads to a couple

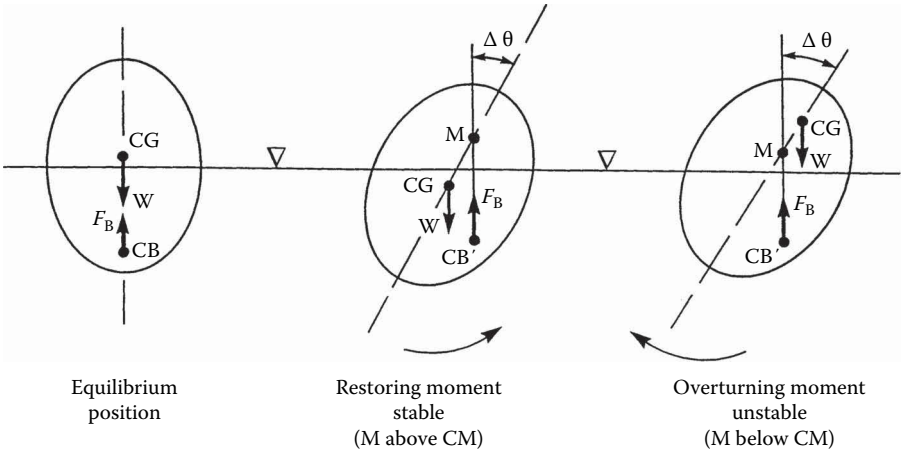


FIGURE 2.1.5 Stability for a partially submerged body.

which will increase the displacement. Thus, the stability of the equilibrium depends on whether M lies above or below the CG. The directed distance from CG to M is called the *metacentric height*, so equivalently the equilibrium is stable if this vector is positive and unstable if it is negative; stability increases as the metacentric height increases. For geometrically complex bodies, such as ships, the computation of the metacenter can be quite complicated.

Pressure Variation in Rigid-Body Motion of a Fluid

In rigid-body motion of a fluid all the particles translate and rotate as a whole, there is no relative motion between particles, and hence no viscous stresses since these are proportional to velocity gradients. The equation of motion is then a balance among pressure, gravity, and the fluid acceleration, specifically,

$$\nabla p = \rho(g - a) \tag{2.1.18}$$

where a is the uniform acceleration of the body. Equation 2.1.18 shows that the lines of constant pressure, including a free surface if any, are perpendicular to the direction $g-a$. Two important applications of this are to a fluid in uniform linear translation and rigid-body rotation. While such problems are not, strictly speaking, fluid statics problems, their analysis and the resulting pressure variation results are similar to those for static fluids.

Uniform Linear Acceleration

For a fluid partially filling a large container moving to the right with constant acceleration $\boldsymbol{a}=(a_x, a_y)$ the geometry of Figure 2.1.6 shows that the magnitude of the pressure gradient in the direction $\boldsymbol{n}$ normal to the accelerating free surface, in the direction $g-a$, is

$$\frac{dp}{dn} = \rho\left[a_x^2 + (g + a_y)^2\right]^{1/2} \tag{2.1.19}$$

and the free surface is oriented at an angle to the horizontal

$$\theta = \tan^{-1}\left(\frac{a_x}{g + a_y}\right) \tag{2.1.20}$$

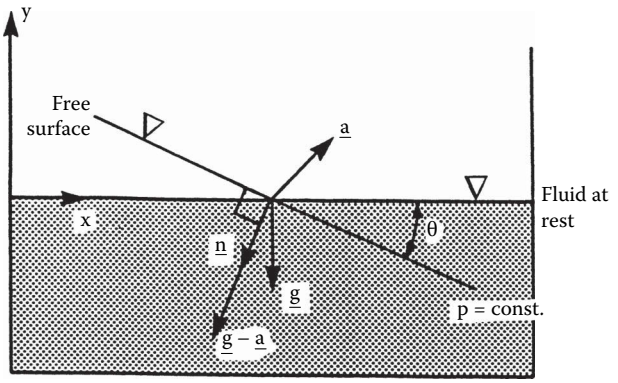


FIGURE 2.1.6 A fluid with a free surface in uniform linear acceleration.

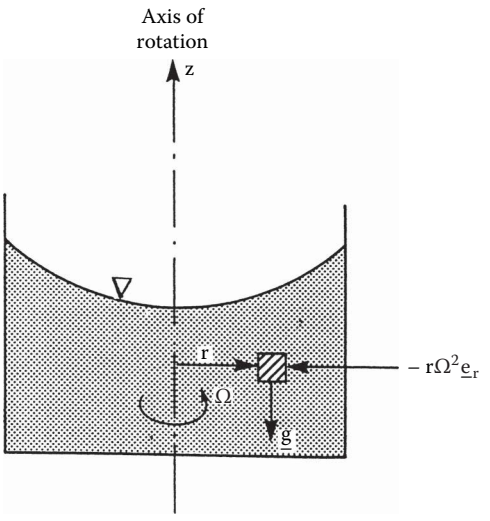


FIGURE 2.1.7 A fluid with a free surface in rigid-body rotation.

Rigid-Body Rotation

Consider the fluid-filled circilar cylinder rotating uniformly with angular velocity $\Omega=\Omega\boldsymbol{e}_z$ (Figure 2.1.7).

The only acceleration is the centripetal acceleration $\Omega\times(\Omega\times\boldsymbol{r})=-r\Omega^2\boldsymbol{e}_r$ so Equation 2.1.18 becomes

$$\nabla p=\frac{\partial p}{\partial r}\boldsymbol{e}_r+\frac{\partial p}{\partial z}\boldsymbol{e}_z=\rho(\boldsymbol{g}-\boldsymbol{a})=\rho(r\Omega^2\boldsymbol{e}_r-g\boldsymbol{e}_z) \tag{2.1.21}$$

or

$$\frac{\partial p}{\partial r}=\rho r\Omega^2,\qquad \frac{\partial p}{\partial z}=-\rho g=-\gamma \tag{2.1.22}$$

Integration of these equations leads to

$$p=p_0-\gamma z+\frac{1}{2}\rho r^2\Omega^2 \tag{2.1.23}$$

where p_0 is the reference pressure (at $r=z=0$). This result shows that at any fixed r the pressure varies hydrostatically in the vertical direction, while the constant pressure surfaces, including the free surface, are paraboloids of revolution.

FURTHER INFORMATION

The reader may find more detail and additional information on the topics in this section in any one of the many excellent introductory texts on fluid mechanics, such as

Munson, B.R., Young, D.F., and Okiishi, T.H. 1998. *Fundamentals of Fluid Mechanics*, 3rd ed., John Wiley & Sons, New York.

White, F.M. 1999. *Fluid Mechanics*, 4th ed., McGraw-Hill, New York.

2.2 EQUATIONS OF MOTION AND POTENTIAL FLOW

Stanley A. Berger

Integral Relations for a Control Volume

Like most physical conservation laws those governing the motion of a fluid apply to material particles or systems of such particles. This so-called Lagrangian viewpoint is generally not as useful in practical fluid flows as an analysis based on the flow through fixed (or deformable) control volumes (CV)—the Eulerian viewpoint. The relationship between these two viewpoints can be deduced from the Reynolds transport theorem, from which we also most readily derive the governing integral and differential equations of motion.

Reynolds Transport Theorem

The *extensive* quantity B, a scalar, vector, or tensor, is defined as any property of a given quantity of fluid (e.g., momentum, energy) and b as the corresponding value per unit mass (the *intensive* value). The Reynolds transport theorem for a moving and arbitrarily deforming control volume CV, with boundary CS (see Figure 2.2.1), states that

$$\frac{d}{dt}\left(B_{\text{system}}\right) = \frac{d}{dt}\left(\iiint_{\text{CV}} \rho b \; d\upsilon\right) + \iint_{\text{CS}} \rho b(\boldsymbol{V}_r \cdot \boldsymbol{n})\, dA \tag{2.2.1}$$

where B_{system} is the total quantity of B in the system (any mass of fixed identity), $\boldsymbol{n}$ is the outward normal to the CS, $\boldsymbol{V}_r = \boldsymbol{V}(\boldsymbol{r}, t) - \boldsymbol{V}_{\text{CS}}(\boldsymbol{r}, t)$, the velocity of the fluid particle, $\boldsymbol{V}(\boldsymbol{r}, t)$, relative to that of the CS, $\boldsymbol{V}_{\text{CS}}(\boldsymbol{r}, t)$, and d/dt on the left-hand side is the derivative following the fluid particles, i.e., the fluid mass comprising the system. The theorem states that the time rate of change of the total B in the system is equal to the rate of change within the CV plus the net flux of B through the CS. To distinguish between the d/dt which appears on the two sides of Equation 2.2.1 but which have different interpretations, the derivative on the left-hand side, following the system, is denoted by D/Dt and is called the material derivative. This notation is used in what follows. For any function $f(x, y, z, t)$,

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + (\boldsymbol{V} \cdot \nabla) f$$

For a CV fixed with respect to the reference frame, Equation 2.2.1 reduces to

$$\frac{D}{Dt}\left(B_{\text{system}}\right) = \frac{d}{dt}\iiint_{\substack{\text{CV} \\ \text{(fixed)}}} (\rho b)\, d\upsilon + \iint_{\text{CS}} \rho b(\boldsymbol{V} \cdot \boldsymbol{n})\, dA \tag{2.2.2}$$

(The time derivative operator in the first term on the right-hand side may be moved inside the integral, in which case it is then to be interpreted as the partial derivative $\partial/\partial t$.)

Conservation of Mass

If we apply Equation 2.2.2 for a fixed control volume, with B_{system} the total mass in the system, then since conservation of mass requires that $DB_{\text{system}}/Dt = 0$ there follows, since $b = B_{\text{system}}/m = 1$,

$$\iiint_{\substack{\text{CV} \\ \text{(fixed)}}} \frac{\partial \rho}{\partial t}\, d\upsilon + \iint_{\text{CS}} \rho(\boldsymbol{V} \cdot \boldsymbol{n})\, dA = 0 \tag{2.2.3}$$

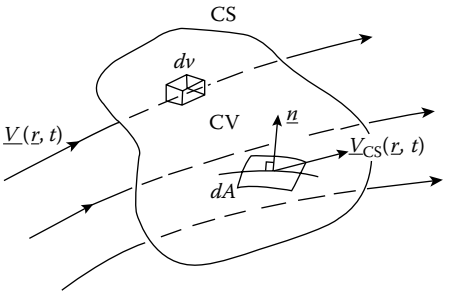


FIGURE 2.2.1 Control volume.

This is the integral form of the conservation of mass law for a fixed control volume. For a steady flow, Equation 2.2.3 reduces to

$$\iint_{CS} \rho(\boldsymbol{V} \cdot \boldsymbol{n}) dA = 0 \tag{2.2.4}$$

whether compressible or incompressible. For an incompressible flow, ρ=constant, so

$$\iint_{CS} (\boldsymbol{V} \cdot \boldsymbol{n}) dA = 0 \tag{2.2.5}$$

whether the flow is steady or unsteady.

Conservation of Momentum

The conservation of (linear) momentum states that

$$\boldsymbol{F}_{\text{total}} \equiv \sum (\text{external forces acting on the fluid system}) = \frac{D\boldsymbol{M}}{Dt} \equiv \frac{D}{Dt}\left(\iiint_{\text{system}} \rho \boldsymbol{V} \, d\upsilon \right) \tag{2.2.6}$$

where $\boldsymbol{M}$ is the total system momentum. For an arbitrarily moving, deformable control volume it then follows from Equation 2.2.1 with b set to V,

$$\boldsymbol{F}_{\text{total}} = \frac{d}{dt}\left(\iiint_{cv} \rho \boldsymbol{V} \, d\upsilon \right) + \iint_{CS} \rho \boldsymbol{V} (\boldsymbol{V}_r \cdot \boldsymbol{n}) dA \tag{2.2.7}$$

This expression is only valid in an inertial coordinate frame. To write the equivalent expression for a noninertial frame, we must use the relationship between the acceleration a_1 in an inertial frame and the acceleration a_R in a noninertial frame,

$$\boldsymbol{a}_1 = \boldsymbol{a}_R + \frac{d^2 \boldsymbol{R}}{dt^2} + 2\Omega \times \boldsymbol{V} + \Omega \times (\Omega \times r) + \frac{d\Omega}{dt} \times r \tag{2.2.8}$$

where $\boldsymbol{R}$ is the position vector of the origin of the noninertial frame with respect to that of the inertial frame, Ω is the angular velocity of the noninertial frame, and $\boldsymbol{r}$ and $\boldsymbol{V}$ the position and velocity vectors in the noninertial frame. The third term on the right-hand side of Equation 2.2.8 is the Coriolis acceleration, and the fourth term is the centripetal acceleration. For a noninertial frame Equation 2.2.7 is then

$$\boldsymbol{F}_{\text{total}} - \iiint_{\text{system}} \left[\frac{d^2\boldsymbol{R}}{dt^2} + 2\Omega \times V + \Omega \times (\Omega \times r) + \frac{d\Omega}{dt} \times r \right] \rho \, d\upsilon = \frac{D}{Dt} \left(\iiint_{\text{system}} \rho \boldsymbol{V} \, d\upsilon \right)$$

$$= \frac{d}{dt} \left(\iiint_{\text{CV}} \rho \boldsymbol{V} \, d\upsilon \right) + \iint_{\text{CS}} \rho \boldsymbol{V} \cdot (\boldsymbol{V}_r \cdot \boldsymbol{n}) dA \tag{2.2.9}$$

where the frame acceleration terms of Equation 2.2.8 have been brought to the left-hand side because to an observer in the noninertial frame they act as "apparent" body forces.

For a fixed control volume in an inertial frame for steady flow, it follows from the above that

$$\boldsymbol{F}_{\text{total}} = \iint_{\text{CS}} \rho \boldsymbol{V} (\boldsymbol{V} \cdot \boldsymbol{n}) dA \tag{2.2.10}$$

This expression is the basis of many control volume analyses for fluid flow problems.

The cross product of $\boldsymbol{r}$, the position vector with respect to a convenient origin, with the momentum Equation 2.2.6 written for an elementary particle of mass dm, noting that $(d\boldsymbol{r}/dt) \times \boldsymbol{V} = 0$, leads to the integral moment of momentum equation

$$\sum \boldsymbol{M} - \boldsymbol{M}_I = \frac{D}{Dt} \iiint_{\text{system}} \rho (\boldsymbol{r} \times \boldsymbol{V}) d\upsilon \tag{2.2.11}$$

where ΣM is the sum of the moments of all the external forces acting on the system about the origin of $\boldsymbol{r}$, and $\boldsymbol{M}_1$ is the moment of the apparent body forces (see Equation 2.2.9). The right-hand side can be written for a control volume using the appropriate form of the Reynolds transport theorem.

Conservation of Energy

The conservation of energy law follows from the first law of thermodynamics for a moving system

$$\dot{Q} - \dot{W} = \frac{D}{Dt} \left(\iiint_{\text{system}} \rho e \, d\upsilon \right) \tag{2.2.12}$$

where $\dot{Q}$ is the rate at which heat is added to the system, $\dot{W}$ the rate at which the system works on its surroundings, and e is the total energy per unit mass. For a particle of mass dm the contributions to the specific energy e are the internal energy u, the kinetic energy $V^2/2$, and the potential energy, which in the case of gravity, the only body force we shall consider, is gz, where z is the vertical displacement opposite to the direction of gravity. (We assume no energy transfer owing to chemical reaction as well as no magnetic or electric fields.) For a fixed control volume it then follows from Equation 2.2.2 [with $b = e = u + (V^2/2) + gz$] that

$$\dot{Q} - \dot{W} = \frac{d}{dt} \left(\iiint_{\text{CV}} \rho \left(u + \frac{1}{2} V^2 + gz \right) d\upsilon \right) + \iint_{\text{CS}} \rho \left(u + \frac{1}{2} V^2 + gz \right) (\boldsymbol{V} \cdot \boldsymbol{n}) \, dA \tag{2.2.13}$$

Example 2.2.1

An incompressible fluid flows through a pump at a volumetric flow rate $\hat{Q}$. The (head) loss between Sections 1 and 2 (see Figure 2.2.2) is equal to $\beta \rho V_1^2/2$ (V is the average velocity at the section). Calculate the power that must be delivered by the pump to the fluid to produce a given increase in pressure, $\Delta p = p_2 - p_1$.

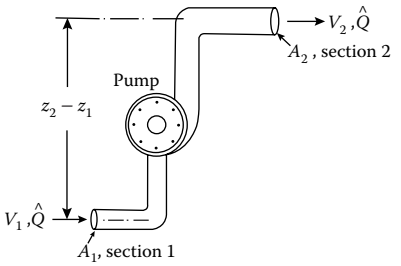


FIGURE 2.2.2 Pump producing pressure increase.

Solution: The principal equation needed is the energy Equation 2.2.13. The term $\dot{W}$, the rate at which the system does work on its surroundings, for such problems has the form

$$\dot{W} = -\dot{W}_{\text{shaft}} + \iint_{\text{CS}} p\boldsymbol{V}\cdot\boldsymbol{n}\ dA \tag{P.2.2.1}$$

where $\dot{W}_{\text{shaft}}$ represents the work done on the fluid by a moving shaft, such as by turbines, propellers, fans, etc., and the second term on the right side represents the rate of working by the normal stress, the pressure, at the boundary. For a steady flow in a control volume coincident with the physical system boundaries and bounded at its ends by Sections 1 and 2, Equation 2.2.13 reduces to ($u \equiv 0$),

$$\dot{Q} + \dot{W}_{\text{shaft}} - \iint_{\text{CS}} p\boldsymbol{V}\cdot\boldsymbol{n}\,dA = \iint_{\text{CS}} \left(\frac{1}{2}\rho V^2 + \gamma z\right)(\boldsymbol{V}\cdot\boldsymbol{n})\ dA \tag{P.2.2.2}$$

Using average quantities at Sections 1 and 2, and the continuity Equation 2.2.5, which reduces in this case to

$$V_1A_1 = V_2A_2 = \hat{Q} \tag{P.2.2.3}$$

we can write Equation P.2.2.2 as

$$\dot{Q} + \dot{W}_{\text{shaft}} - (p_2 - p_1)\hat{Q} = \left[\frac{1}{2}\rho\left(V_2^2 - V_1^2\right) + \gamma\left(z_2 - z_1\right)\right]\hat{Q} \tag{P.2.2.4}$$

$\dot{Q}$, the rate at which heat is added to the system, is here equal to $-\beta\rho V_1^2/2$, the head loss between Sections 1 and 2. Equation P.2.2.4 then can be rewritten

$$\dot{W}_{\text{shaft}} = \beta\rho\frac{V_1^2}{2} + (\Delta p)\hat{Q} + \frac{1}{2}\rho\left(V_2^2 - V_1^2\right)\hat{Q} + \gamma\left(z_2 - z_1\right)\hat{Q}$$

or, in terms of the given quantities,

$$\dot{W}_{\text{shaft}} = \frac{\beta\rho\hat{Q}^2}{A_1^2} + \left(\Delta p\right)\hat{Q} + \frac{1}{2}\rho\frac{\hat{Q}^3}{A_2^2}\left(1 - \frac{A_2^2}{A_1^2}\right) + \gamma\left(z_2 - z_1\right)\hat{Q} \tag{P.2.2.5}$$

Thus, for example, if the fluid is water ($\rho \approx 1000\,\text{kg/m}^3$, $\gamma = 9.8\,\text{kN/m}^3$), $\hat{Q} = 0.5\,\text{m}^3/\text{s}$, the heat loss is $0.2\rho V_1^2/2$, and $\Delta p = p_2 - p_1 = 2\times10^5\,\text{N/m}^2 = 200\,\text{kPa}$, $A_1 = 0.1\,\text{m}^2 = A_2/2$, $(z_2 - z_1) = 2\,\text{m}$, we find, using Equation P.2.2.5

$$\dot{W}_{\text{shaft}} = \frac{0.2(1000)(0.5)^2}{(0.1)^2} + \left(2\times10^5\right)(0.5) + \frac{1}{2}(1000)\frac{(0.5)^3}{(0.2)^2}(1-4) + \left(9.8\times10^3\right)(2)(0.5)$$
$$= 5{,}000 + 10{,}000 - 4{,}688 + 9{,}800 = 20{,}112 \text{ W}$$
$$= \frac{20{,}112}{745.7}\text{hp} = 27 \text{ hp}$$

Differential Relations for Fluid Motion

In the previous section the conservation laws were derived in integral form. These forms are useful in calculating, generally using a control volume analysis, gross features of a flow. Such analyses usually require some *a priori* knowledge or assumptions about the flow. In any case, an approach based on integral conservation laws cannot be used to determine the point-by-point variation of the dependent variables, such as velocity, pressure, temperature, etc. To do this requires the use of the differential forms of the conservation laws, which are presented below.

Mass Conservation–Continuity Equation

Applying Gauss's theorem (the divergence theorem) to Equation 2.2.3 we obtain

$$\iiint_{\substack{\text{CV}\\(\text{fixed})}}\left[\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho V)\right]d\upsilon = 0 \tag{2.2.14}$$

which, because the control volume is arbitrary, immediately yields

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{V}) = 0 \tag{2.2.15}$$

This can also be written as

$$\frac{D\rho}{Dt} + \rho\nabla\cdot\boldsymbol{V} = 0 \tag{2.2.16}$$

using the fact that

$$\frac{D\rho}{Dt} = \frac{\partial\rho}{\partial t} + \boldsymbol{V}\cdot\nabla\rho \tag{2.2.17}$$

Special cases:

1. Steady flow $[(\partial/\partial t)\,() \equiv 0]$

$$\nabla.(\rho\boldsymbol{V}) = 0 \tag{2.2.18}$$

2. Incompressible flow ($D\rho/D\text{t}=0$)

$$\nabla\cdot V = 0 \tag{2.2.19}$$

Momentum Conservation

We note first, as a consequence of mass conservation for a system, that the right-hand side of Equation 2.2.6 can be written as

$$\frac{D}{Dt}\left(\iiint_{\text{system}} \rho V \; d\upsilon\right) \equiv \iiint_{\text{system}} \rho \frac{DV}{Dt} d\upsilon \tag{2.2.20}$$

The total force acting on the system which appears on the left-hand side of Equation 2.2.6 is the sum of body forces $\boldsymbol{F}_b$ and surface forces $\boldsymbol{F}_s$. The body forces are often given as forces per unit mass (e.g., gravity), and so can be written as

$$F_b = \iiint_{\text{system}} \rho f \; d\upsilon \tag{2.2.21}$$

The surface forces are represented in terms of the second-order stress tensor* $\underline{\sigma} = \{\sigma_{ij}\}$, where σ_{ij} is defined as the force per unit area in the i direction on a planar element whose normal lies in the j direction.† From elementary angular momentum considerations for an infinitesimal volume it can be shown that σ_{ij} is a symmetric tensor, and therefore has only six independent components. The total surface force exerted on the system by its surroundings is given by

$$F_s = \iint_{\substack{\text{system} \\ \text{surface}}} \underline{\sigma} \cdot \boldsymbol{n} \; dA, \text{ with } i - \text{component } F_{s_i} = \iint \sigma_{ij} n_j \; dA \tag{2.2.22}$$

The integral momentum conservation law Equation 2.2.6 can then be written

$$\iiint_{\text{system}} \rho \frac{DV}{Dt} d\upsilon = \iiint_{\text{system}} \rho f \; d\upsilon + \iint_{\substack{\text{system} \\ \text{surface}}} \underline{\sigma} \cdot \boldsymbol{n} \; dA \tag{2.2.23}$$

The application of the divergence theorem to the last term on the right-side of Equation 2.2.23 leads to

$$\iiint_{\text{system}} \rho \frac{DV}{Dt} d\upsilon = \iiint_{\text{system}} \rho f \; d\upsilon + \iiint_{\text{system}} \nabla \cdot \underline{\sigma} \; d\upsilon \tag{2.2.24}$$

where $\nabla \cdot \sigma \equiv \{\partial \sigma_{ij} / x_j\}$. Since Equation 2.2.24 holds for any material volume, it follows that

$$\rho \frac{DV}{Dt} = \rho \boldsymbol{f} + \nabla \cdot \underline{\sigma} \tag{2.2.25}$$

With the decomposition of $\boldsymbol{F}_{\text{total}}$ above, Equation 2.2.10 can be written as

$$\iiint_{\text{CV}} \rho f \; d\upsilon + \iint_{\text{CS}} \underline{\sigma} \cdot \boldsymbol{n} \; dA = \iint_{\text{CS}} \rho \boldsymbol{V} (\boldsymbol{V} \cdot \boldsymbol{n}) dA \tag{2.2.26}$$

If ρ is uniform and $\boldsymbol{f}$ is a conservative body force, i.e., $\boldsymbol{f} = -\nabla \Psi$, where Ψ is the force potential, then Equation 2.2.26, after application of the divergence theorem to the body force term, can be written

---

* We shall assume the reader is familiar with elementary Cartesian tensor analysis and the associated subscript notation and conventions. The reader for whom this is not true should skip the details and concentrate on the final principal results and equations given at the ends of the next few subsections.

† This assignment of roles to the first and second subscripts of the stress tensor is a convention and is far from universal. Frequently, their roles are reversed, the first subscript denoting the direction of the normal to the planar element, the second the direction of the force.

$$\iint_{CS}\left(-\rho\Psi n+\underline{\sigma}\cdot\boldsymbol{n}\right)dA=\iint_{CS}\rho V(\boldsymbol{V}\cdot\boldsymbol{n})\,dA \tag{2.2.27}$$

It is in this form, involving only integrals over the surface of the control volume, that the integral form of the momentum equation is used in control volume analyses, particularly in the case when the body force term is absent.

Analysis of Rate of Deformation

The principal aim of the following two subsections is to derive a relationship between the stress and the rate of strain to be used in the momentum Equation 2.2.25. The reader less familiar with tensor notation may skip these sections, apart from noting some of the terms and quantities defined therein, and proceed directly to Equations 2.2.38 or 2.2.39.

The relative motion of two neighboring points P and Q, separated by a distance η, can be written (using $\boldsymbol{u}$ for the local velocity)

$$\boldsymbol{u}(Q)=\boldsymbol{u}(P)+(\nabla\boldsymbol{u})\eta$$

or, equivalently, writing $\nabla\boldsymbol{u}$ as the sum of antisymmetric and symmetric tensors,

$$\boldsymbol{u}(Q)=\boldsymbol{u}(P)+\frac{1}{2}\left((\nabla\boldsymbol{u})-(\nabla\boldsymbol{u})^{*}\right)\eta+\frac{1}{2}\left((\nabla\boldsymbol{u})+(\nabla\boldsymbol{u})^{*}\right)\eta \tag{2.2.28}$$

where $\nabla\boldsymbol{u}=\{\partial u_i/\partial x_j\}$, and the superscript * denotes transpose, so $(\nabla\boldsymbol{u})^{*}=\{\partial u_j/\partial x_i\}$. The second term on the right-hand side of Equation 2.2.28 can be rewritten in terms of the *vorticity*, $\nabla\times\boldsymbol{u}$, so Equation 2.2.28 becomes

$$\boldsymbol{u}(Q)=\boldsymbol{u}(P)+\frac{1}{2}(\nabla\times\boldsymbol{u})\times\eta+\frac{1}{2}\left((\nabla\boldsymbol{u})+(\nabla\boldsymbol{u})^{*}\right)\eta \tag{2.2.29}$$

which shows that the local velocity field consists of a rigid-body translation, a rigid-body rotation with angular velocity 1/2 ($\nabla\times\boldsymbol{u}$), and a velocity or rate of deformation. The coefficient of $\boldsymbol{\eta}$ in the last term in Equation 2.2.29 is defined as the rate-of-strain tensor and is denoted by $\underline{e}$, in subscript form

$$e_{ij}=\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right) \tag{2.2.30}$$

From $\underline{e}$ we can define a rate-of-strain central quadric, along the principal axes of which the deforming motion consists of a pure straining extension or contraction.

Relationship between Forces and Rate of Deformation

We are now in a position to determine the required relationship between the stress tensor $\underline{\sigma}$ and the rate of deformation. Assuming that in a static fluid the stress reduces to a (negative) hydrostatic or thermodynamic pressure, equal in all directions, we can write

$$\underline{\sigma}=-p\underline{I}+\underline{\tau}\qquad\text{or}\qquad\sigma_{ij}=-p\delta_{ij}+\tau_{ij} \tag{2.2.31}$$

where $\underline{\tau}$ is the viscous part of the total stress and is called the deviatoric stress tensor, $\underline{I}$ is the identity tensor, and δ_{ij} is the corresponding Kronecker delta ($\delta_{ij}=0$ if $i\neq j$; $\delta_{ij}=1$ if $i=j$). We make further assumptions that (1) the fluid exhibits no preferred directions; (2) the stress is independent of any previous history of distortion; and (3) that the stress depends only on the

local thermodynamic state and the kinematic state of the immediate neighborhood. Precisely, we assume that $\underline{\tau}$ is linearly proportional to the first spatial derivatives of $\boldsymbol{u}$, the coefficient of proportionality depending only on the local thermodynamic state. These assumptions and the relations below which follow from them are appropriate for a Newtonian fluid. Most common fluids, such as air and water under most conditions, are Newtonian, but there are many other fluids, including many which arise in industrial applications, which exhibit so-called non-Newtonian properties. The study of such non-Newtonian fluids, such as viscoelastic fluids, is the subject of the field of rheology (See Section 2.9 in this chapter).

With the Newtonian fluid assumptions above, and the symmetry of $\underline{\tau}$ which follows from the symmetry of $\underline{\sigma}$, one can show that the viscous part $\underline{\tau}$ of the total stress can be written as

$$\underline{\tau} = \lambda(\nabla \cdot \boldsymbol{u})\underline{I} + 2\mu\underline{e} \tag{2.2.32}$$

so the total stress for a Newtonian fluid is

$$\underline{\sigma} = -p\underline{I} + \lambda(\nabla \cdot \boldsymbol{u})\underline{I} + 2\mu\underline{e} \tag{2.2.33}$$

or, in subscript notation

$$\sigma_{ij} = -p\delta_{ij} + \lambda\left(\frac{\partial u_k}{\partial x_k}\right)\delta_{ij} + \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{2.2.34}$$

(the Einstein summation convention is assumed here, namely, that a repeated subscript, such as in the second term on the right-hand side above, is summed over; note also that $\nabla \cdot \boldsymbol{u} = \partial u_k/\partial x_k = e_{kk}$). The coefficient λ is called the "second viscosity" and μ the "absolute viscosity," or more commonly the "dynamic viscosity," or simply the "viscosity." For a Newtonian fluid λ and μ depend only on local thermodynamic state, primarily on the temperature.

We note, from Equation 2.2.34, that whereas in a fluid at rest the stress is an isotropic normal stress (equal to p in all directions), this is not the case for a moving fluid, since in general $\sigma_{11} \neq \sigma_{22} \neq \sigma_{33}$. To have an analogous quantity to p for a moving fluid, the values of we define the pressure in a moving fluid as the negative mean normal stress, denoted, say, by $\overline{p}$

$$\overline{p} = -\frac{1}{3}\sigma_{ij} \tag{2.2.35}$$

(σ_{ii} is the trace of $\underline{\sigma}$ and an invariant of $\underline{\sigma}$, independent of the orientation of the axes). From Equation 2.2.34

$$\overline{p} = -\frac{1}{3}\sigma_{ii} = p - \left(\lambda + \frac{2}{3}\mu\right)\nabla \cdot \boldsymbol{u} \tag{2.2.36}$$

For an incompressible fluid $\nabla \cdot \boldsymbol{u} = 0$ and hence $\overline{p} \equiv p$. The quantity $(\lambda + 2/3\mu)$ is called the bulk viscosity. If one assumes that the deviatoric stress tensor τ_{ij} makes no contribution to the mean normal stress, it follows that $\lambda + (2/3)\mu = 0$, so again $\overline{p} = p$. This condition, $\lambda = -(2/3)\mu$, is called the Stokes assumption or hypothesis. If neither the incompressibility nor the Stokes assumptions are made, the difference between $\overline{p}$ and p is usually still negligibly small because $(\lambda + 2/3\mu)\,\nabla u \ll p$ in most fluid flow problems. If the Stokes hypothesis is made, as is often the case in fluid mechanics, Equation 2.2.34 becomes

$$\sigma_{ij} = -p\delta_{ij} + 2\mu\left(e_{ij} - \frac{1}{3}e_{kk}\delta_{ij}\right) \tag{2.2.37}$$

The Navier–Stokes Equations

Substitution of Equation 2.2.33 into 2.2.25, since $\nabla\cdot(\phi\underline{I})=\nabla\phi$, for any scalar function ϕ, yields (replacing $\boldsymbol{u}$ in Equation 2.2.33 by $\boldsymbol{V}$)

$$\rho\frac{D\boldsymbol{V}}{Dt}=\rho\boldsymbol{f}-\nabla p+\nabla(\lambda\nabla\cdot\boldsymbol{V})+\nabla\cdot(2\mu\underline{e}) \tag{2.2.38}$$

These equations are the Navier–Stokes equations (although the name is as often given to the full set of governing conservation equations). With the Stokes assumption ($\lambda=-2/3\mu$), Equation 2.2.38 becomes

$$\rho\frac{D\boldsymbol{V}}{Dt}=\rho\boldsymbol{f}-\nabla p+\nabla\cdot\left[2\mu\left(\underline{\underline{e}}-\frac{1}{3}e_{kk}\underline{\underline{I}}\right)\right] \tag{2.2.39}$$

If the Eulerian frame is not an inertial frame, then one must use the transformation to an inertial frame either using Equation 2.2.8 or the "apparent" body force formulation, Equation 2.2.9.

Energy Conservation: The Mechanical and Thermal Energy Equations

In deriving the differential form of the energy equation we begin by assuming that heat enters or leaves the material or control volume by heat conduction across the boundaries, the heat flux per unit area being $\boldsymbol{q}$. It then follows that

$$\dot{Q}=-\iint\boldsymbol{q}\cdot\boldsymbol{n}\;dA=-\iiint\nabla\cdot\boldsymbol{q}\;d\upsilon \tag{2.2.40}$$

The work-rate term $\dot{W}$ can be decomposed into the rate of work done against body forces, given by

$$-\iiint\rho\boldsymbol{f}\cdot\boldsymbol{V}\,d\upsilon \tag{2.2.41}$$

and the rate of work done against surface stresses, given by

$$-\iint_{\substack{\text{system}\\\text{surface}}}\boldsymbol{V}\cdot(\underline{\sigma}\boldsymbol{n})dA \tag{2.2.42}$$

Substitution of these expressions for $\dot{Q}$ and $\dot{W}$ into Equation 2.2.12, use of the divergence theorem, and conservation of mass lead to

$$\rho\frac{D}{Dt}\left(u+\frac{1}{2}V^2\right)=-\nabla\cdot\boldsymbol{q}+\rho\boldsymbol{f}\cdot\boldsymbol{V}+\nabla\cdot(\boldsymbol{V}\underline{\sigma}) \tag{2.2.43}$$

(note that a potential energy term is no longer included in e, the total specific energy, as it is accounted for by the body force rate-of-working term $\rho f\cdot V$).

Equation 2.2.43 is the total energy equation showing how the energy changes as a result of working by the body and surface forces and heat transfer. It is often useful to have a purely thermal energy equation. This is obtained by subtracting from Equation 2.2.43 the dot product of $\boldsymbol{V}$ with the momentum Equation 2.2.25, after expanding the last term in Equation 2.2.43, resulting in

$$\rho\frac{Du}{Dt}=\frac{\partial V_i}{\partial x_j}\sigma_{ij}-\nabla\cdot\boldsymbol{q} \tag{2.2.44}$$

With $\sigma_{ij}=-p\delta_{ij}+\tau_{ij}$, and the use of the continuity equation in the form of Equation 2.2.16, the first term on the right-hand side of Equation 2.2.44 may be written

$$\frac{\partial V_i}{\partial x_j}\sigma_{ij} = -\rho\frac{D\left(\frac{p}{\rho}\right)}{Dt}+\frac{Dp}{Dt}+\Phi \tag{2.2.45}$$

where Φ is the rate of dissipation of mechanical energy per unit mass due to viscosity, and is given by

$$\Phi \equiv \frac{\partial V_i}{\partial x_j}\tau_{ij} = 2\mu\left(e_{ij}e_{ij}-\frac{1}{3}e_{kk}^2\right)=2\mu\left(e_{ij}-\frac{1}{3}e_{kk}\delta_{ij}\right)^2 \tag{2.2.46}$$

With the introduction of Equation 2.2.45, Equation 2.2.44 becomes

$$\rho\frac{De}{Dt} = -p\nabla\cdot\boldsymbol{V}+\Phi-\nabla\cdot\boldsymbol{q} \tag{2.2.47}$$

or

$$\rho\frac{Dh}{Dt} = \frac{Dp}{Dt}+\Phi-\nabla\cdot\boldsymbol{q} \tag{2.2.48}$$

where $h=e+(p/\rho p)$ is the specific enthalpy. Unlike the other terms on the right-hand side of Equation 2.2.47, which can be negative or positive, Φ is always nonnegative and represents the increase in internal energy (or enthalpy) owing to irreversible degradation of mechanical energy. Finally, from elementary thermodynamic considerations

$$\frac{Dh}{Dt} = T\frac{DS}{Dt}+\frac{1}{\rho}\frac{Dp}{Dt}$$

where S is the entropy, so Equation 2.2.48 can be written as

$$\rho T\frac{DS}{Dt} = \Phi-\nabla\cdot\boldsymbol{q} \tag{2.2.49}$$

If the heat conduction is assumed to obey the Fourier heat conduction law, so $\boldsymbol{q}=-k\nabla T$, where k is the thermal conductivity, then in all of the above equations

$$-\nabla\cdot\boldsymbol{q} = \nabla\cdot(k\nabla T) = k\nabla^2 T \tag{2.2.50}$$

the last of these equalities holding only if $k=$constant.

In the event the thermodynamic quantities vary little, the coefficients of the constitutive relations for $\underline{\sigma}$ and $\boldsymbol{q}$ may be taken to be constant and the above equations simplified accordingly.

We note also that if the flow is incompressible, then the mass conservation, or continuity, equation simplifies to

$$\nabla\cdot\boldsymbol{V} = 0 \tag{2.2.51}$$

and the momentum Equation 2.2.38 to

$$\rho\frac{D\boldsymbol{V}}{Dt} = \rho\boldsymbol{f}-\nabla p+\mu\nabla^2\boldsymbol{V} \tag{2.2.52}$$

where ∇^2 is the Laplacian operator. The small temperature changes, compatible with the incompressibility assumption, are then determined, for a perfect gas with constant k and specific heats, by the energy equation rewritten for the temperature, in the form

$$\rho c_v \frac{DT}{Dt} = k\nabla^2 T + \Phi \tag{2.2.53}$$

Boundary Conditions

The appropriate boundary conditions to be applied at the boundary of a fluid in contact with another medium depends on the nature of this other medium—solid, liquid, or gas. We discuss a few of the more important cases here in turn:

1 *At a solid surface:* $\boldsymbol{V}$ and T are continuous. Contained in this boundary condition is the "no-slip" condition, namely, that the tangential velocity of the fluid in contact with the boundary of the solid is equal to that of the boundary. For an inviscid fluid, the no-slip condition does not apply, and only the normal component of velocity is continuous. If the wall is permeable, the tangential velocity is continuous and the normal velocity is arbitrary; the temperature boundary condition for this case depends on the nature of the injection or suction at the wall.

2. *At a liquid/gas interface*: For such cases the appropriate boundary conditions depend on what can be assumed about the gas the liquid is in contact with. In the classical liquid free-surface problem, the gas, generally atmospheric air, can be ignored and the necessary boundary conditions are that (1) the normal velocity in the liquid at the interface is equal to the normal velocity of the interface and (2) the pressure in the liquid at the interface exceeds the atmospheric pressure by an amount equal to

$$\Delta p = p_{\text{liquid}} - p_{\text{atm}} = \sigma\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{2.2.54}$$

where R_1 and R_2 are the radii of curvature of the intercepts of the interface by two orthogonal planes containing the vertical axis and σ is the interfacial tension. If the gas is a vapor which undergoes nonnegligible interaction and exchanges with the liquid in contact with it, the boundary conditions are more complex. Then, in addition to the above conditions on normal velocity and pressure, the shear stress (momentum flux) and heat flux must be continuous as well.

For interfaces in general the boundary conditions are derived from continuity conditions for each "transportable" quantity, namely continuity of the appropriate intensity across the interface and continuity of the normal component of the flux vector. Fluid momentum and heat are two such transportable quantities, the associated intensities are velocity and temperature, and the associated flux vectors are stress and heat flux. (The reader should be aware of circumstances where these simple criteria do not apply, for example, the velocity slip and temperature jump for a rarefied gas in contact with a solid surface.)

Vorticity in Incompressible Flow

With μ=constant, ρ=constant, and $\boldsymbol{f}=-\boldsymbol{g}=-g\boldsymbol{k}$ the momentum equation reduces to the form (see Equation 2.2.52)

$$\rho\frac{D\boldsymbol{V}}{Dt} = -\nabla p - \rho g\boldsymbol{k} + \mu\nabla^2 V \tag{2.2.55}$$

With the vector identities

$$(\boldsymbol{V}\cdot\nabla)\boldsymbol{V} = \nabla\left(\frac{V^2}{2}\right) - \boldsymbol{V}\times(\nabla\times\boldsymbol{V}) \tag{2.2.56}$$

and

$$\nabla^2 \boldsymbol{V} = \nabla(\nabla\cdot\boldsymbol{V}) \; - \nabla\times \; (\nabla\times\boldsymbol{V}) \tag{2.2.57}$$

and defining the *vorticity*

$$\zeta \equiv \nabla\times\boldsymbol{V} \tag{2.2.58}$$

Equation 2.2.55 can be written, noting that for incompressible flow $\nabla\cdot\boldsymbol{V}=0$,

$$\rho\frac{\partial \boldsymbol{V}}{\partial t} + \nabla\left(p + \frac{1}{2}\rho V^2 + \rho g z \right) = \rho \boldsymbol{V}\times\xi - \mu\nabla\times\xi \tag{2.2.59}$$

The flow is said to be *irrotational* if

$$\zeta \equiv \nabla\times\boldsymbol{V} = 0 \tag{2.2.60}$$

from which it follows that a *velocity potential* Φ can be defined

$$\boldsymbol{V} = \nabla\Phi \tag{2.2.61}$$

Setting $\zeta=0$ in Equation 2.2.59, using Equation 2.2.61, and then integrating with respect to all the spatial variables, leads to

$$\rho\frac{\partial \Phi}{\partial t} + \left(p + \frac{1}{2}\rho V^2 + \rho g z \right) = F(t) \tag{2.2.62}$$

(the arbitrary function $F(t)$ introduced by the integration can either be absorbed in Φ, or is determined by the boundary conditions). Equation 2.2.62 is the unsteady *Bernoulli equation* for irrotational, incompressible flow. (Irrotational flows are always potential flows, even if the flow is compressible. Because the viscous term in Equation 2.2.59 vanishes identically for $\zeta=0$, it would appear that the above Bernoulli equation is valid even for viscous flow. Potential solutions of hydrodynamics are in fact exact solutions of the full Navier–Stokes equations. Such solutions, however, are not valid near solid boundaries or bodies because the no-slip condition generates vorticity and causes nonzero ζ; the potential flow solution is invalid in all those parts of the flow field that have been "contaminated" by the spread of the vorticity by convection and diffusion. See below.)

The curl of Equation 2.2.59, noting that the curl of any gradient is zero, leads to

$$\rho\frac{\partial \zeta}{\partial t} = \rho\nabla\times(\boldsymbol{V}\times\zeta) - \mu\nabla\times\nabla\times\zeta \tag{2.2.63}$$

but

$$\begin{aligned} \nabla^2\zeta &= \nabla(\nabla\cdot\zeta) - \nabla\times\nabla\times\zeta \\ &= -\nabla\times\nabla\times\zeta \end{aligned} \tag{2.2.64}$$

since div curl () ≡ 0, and therefore also

$$\nabla\times\left(\boldsymbol{V}\times\zeta\right) \equiv \zeta(\nabla\boldsymbol{V}) + \boldsymbol{V}\nabla\cdot\zeta - \boldsymbol{V}\nabla\zeta - \zeta\nabla\cdot\boldsymbol{V} \tag{2.2.65}$$

$$= \zeta(\nabla\boldsymbol{V}) - \boldsymbol{V}\nabla\zeta \tag{2.2.66}$$

Equation 2.2.63 can then be written as

$$\frac{D\boldsymbol{\zeta}}{Dt} = (\boldsymbol{\zeta}\cdot\nabla)\boldsymbol{V} + \nu\nabla^2\boldsymbol{\zeta} \tag{2.2.67}$$

where $\nu=\mu/\rho$ is the kinematic viscosity. Equation 2.2.67 is the vorticity equation for incompressible flow. The first term on the right, an inviscid term, increases the vorticity by vortex stretching. In inviscid, two-dimensional flow both terms on the right-hand side of Equation 2.2.67 vanish, and the equation reduces to $D\zeta/Dt=0$, from which it follows that the vorticity of a fluid particle remains constant as it moves. This is Helmholtz's theorem. As a consequence it also follows that if $\zeta=0$ initially, $\zeta\equiv 0$ always; i.e., *initially irrotational flows remain irrotational* (for inviscid flow). Similarly, it can be proved that $D\Gamma/Dt=0$; i.e., the circulation around a material closed circuit remains constant, which is Kelvin's theorem.

If $\nu\neq 0$, Equation 2.2.67 shows that the vorticity generated, say, at solid boundaries, diffuses and stretches as it is convected.

We also note that for steady flow the Bernoulli equation reduces to

$$p+\frac{1}{2}\rho V^2+\rho g z = \text{constant} \tag{2.2.68}$$

valid for steady, irrotational, incompressible flow.

Stream Function

For two-dimensional planar, incompressible flows ($\boldsymbol{V}=(u,\nu)$), the continuity equation

$$\frac{\partial u}{\partial x}+\frac{\partial \nu}{\partial y}=0 \tag{2.2.69}$$

can be identically satisfied by introducing a stream function Ψ, defined by

$$u=\frac{\partial\psi}{\partial y}, \qquad \upsilon=-\frac{\partial\psi}{\partial x} \tag{2.2.70}$$

Physically Ψ is a measure of the flow between streamlines. (Stream functions can be similarly defined to satisfy identically the continuity equations for incompressible cylindrical and spherical axisymmetric flows; and for these flows, as well as the above planar flow, also when they are compressible, but only then if they are steady.) Continuing with the planar case, we note that in such flows there is only a single nonzero component of vorticity, given by

$$\boldsymbol{\zeta}=(0,0,\zeta_z)=\left(0,0,\frac{\partial\nu}{\partial x}-\frac{\partial u}{\partial y}\right) \tag{2.2.71}$$

With Equation 2.2.70

$$\zeta_z=-\frac{\partial^2\psi}{\partial x^2}-\frac{\partial^2\psi}{\partial y^2}=-\nabla^2\psi \tag{2.2.72}$$

For this two-dimensional flow, Equation 2.2.67 reduces to

$$\frac{\partial\zeta_z}{\partial t}+\nu\frac{\partial\zeta_z}{\partial x}+\nu\frac{\partial\zeta_z}{\partial y}=\nu\left(\frac{\partial^2\zeta_z}{\partial x^2}+\frac{\partial^2\zeta_z}{\partial y^2}\right)=\nu\nabla^2\zeta_z \tag{2.2.73}$$

Substitution of Equation 2.2.72 into Equation 2.2.73 yields an equation for the stream function

$$\frac{\partial\left(\nabla^2\psi\right)}{\partial t}+\frac{\partial\psi}{\partial y}\frac{\partial\left(\nabla^2\psi\right)}{\partial x}-\frac{\partial\psi}{\partial x}\frac{\partial}{\partial y}\left(\nabla^2\psi\right)=\nu\nabla^2\psi \tag{2.2.74}$$

where $\nabla^4=\nabla^2\,(\nabla^2)$. For uniform flow past a solid body, for example, this equation for Ψ would be solved subject to the boundary conditions:

$$\begin{aligned}&\frac{\partial\psi}{\partial x}=0,\quad \frac{\partial\psi}{\partial y}=V_\infty \qquad \text{at infinity}\\ &\frac{\partial\psi}{\partial x}=0,\quad \frac{\partial\psi}{\partial y}=0 \qquad \text{at the body (no-slip)}\end{aligned} \tag{2.2.75}$$

For the special case of irrotational flow, it follows immediately from Equations 2.2.70 and 2.2.71 with $\zeta_z=0$, that Ψ satisfies the Laplace equation

$$\nabla^2\psi=\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}=0 \tag{2.2.76}$$

Inviscid Irrotational Flow: Potential Flow

For irrotational flows, we have already noted that a velocity potential Φ can be defined such that $\boldsymbol{V}=\nabla\Phi$. If the flow is also incompressible, so $\nabla\cdot\boldsymbol{V}=0$, it then follows that

$$\nabla\cdot(\nabla\Phi)=\nabla^2\Phi=0 \tag{2.2.77}$$

so Φ satisfies Laplace's equation. (Note that unlike the stream function Ψ, which can only be defined for two-dimensional flows, the above considerations for Φ apply to flow in two and three dimensions. On the other hand, the existence of Ψ does not require the flow to be irrotational, whereas the existence of Φ does.)

Since Equation 2.2.77 with appropriate conditions on $\boldsymbol{V}$ at boundaries of the flow completely determines the velocity field, and the momentum equation has played no role in this determination, we see that inviscid irrotational flow—*potential theory*—is a purely kinematic theory. The momentum equation enters only after Φ is known in order to calculate the pressure field consistent with the velocity field $\boldsymbol{V}=\nabla\Phi$.

For both two-and three-dimensional flows the determination of Φ makes use of the powerful techniques of potential theory, well developed in the mathematical literature. For two-dimensional planar flows, the techniques of complex variable theory are available, since Φ may be considered as either the real or imaginary part of an analytic function (the same being true for Ψ, since for such two-dimensional flows Φ and Ψ are conjugate variables.)

Because the Laplace equation, obeyed by both Φ and Ψ, is linear, complex flows may be built up from the superposition of simple flows; this property of inviscid irrotational flows underlies nearly all solution techniques in this area of fluid mechanics.

Example 2.2.2

A two-dimensional inviscid irrotational flow has the velocity potential

$$\Phi = x^2-y^2 \tag{P.2.2.6}$$

What two-dimensional potential flow does this represent?

Solution: It follows from Equations 2.2.61 and 2.2.70 that for two-dimensional flows, in general

$$u = \frac{\partial \Phi}{\partial x} = \frac{\partial \psi}{\partial y}, \qquad \upsilon = \frac{\partial \Phi}{\partial y} = -\frac{\partial \psi}{\partial x} \tag{P.2.2.7}$$

It follows from Equation P.2.2.6 that

$$u = \frac{\partial \psi}{\partial y} = 2x, \qquad \upsilon = -\frac{\partial \psi}{\partial x} = -2y \tag{P.2.2.8}$$

Integration of Equation P.2.2.8 yields

$$\Psi = 2xy \tag{P.2.2.9}$$

The streamlines, Ψ=constant, and equipotential lines, Φ=constant, both families of hyperbolas and each family the orthogonal trajectory of the other, are shown in Figure 2.2.3. Because the x and y axes are streamlines, Equations P.2.2.6 and P.2.2.9 represent the inviscid irrotational flow in a right-angle corner. By symmetry, they also represent the planar flow in the upper half-plane directed toward a stagnation point at $x=y=0$ (see Figure 2.2.4). In polar coordinates (r, θ), with corresponding velocity components (u_r, u_θ), this flow is represented by

$$\Phi = r^2 \cos 2\theta, \qquad \psi = r^2 \sin 2\theta \tag{P.2.2.10}$$

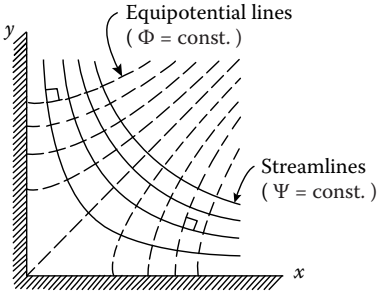


FIGURE 2.2.3 Potential flow in a 90° corner.

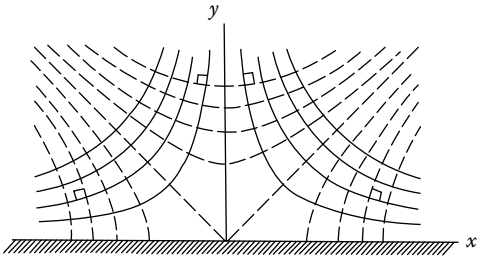


FIGURE 2.2.4 Potential flow impinging against a flat (180°) wall (plane stagnation-point flow).

with

$$u_r = \frac{\partial \Phi}{\partial r} = \frac{1}{r}\frac{\partial \psi}{\partial \theta} = 2r\cos 2\theta$$
$$u_\theta = \frac{1}{r}\frac{\partial \Phi}{\partial \theta} = -\frac{\partial \psi}{\partial r} = -2r\sin 2\theta \qquad \text{(P.2.2.11)}$$

For two-dimensional planar potential flows we may also use complex variables, writing the complex potential $f(z)=\Phi+i\Psi$ as a function of the complex variable $z=x+iy$, where the complex velocity is given by $f'(z)=w(z)=u-iv$. For the flow above,

$$f(z) = z^2 \qquad \text{(P.2.2.12)}$$

Expressions such as Equation P.2.2.12, where the right-hand side is an analytic function of z, may also be regarded as a conformal mapping, which makes available as an aid in solving two-dimensional potential problems all the tools of this branch of mathematics.

FURTHER INFORMATION

More detail and additional information on the topics in this section may be found in more advanced books on fluid dynamics, such as

Batchelor, G.K. 1967. *An Introduction to Fluid Dynamics*, Cambridge University Press, Cambridge, U.K.
Panton, R.L. 2005. *Incompressible Flow*, 3rd ed., John Wiley & Sons, New York.
Pozrikidis, C. 1997. *Introduction to Theoretical and Computational Fluid Dynamics*, Oxford University Press, New York.
Sherman, F.S. 1990. *Viscous Flow*, McGraw-Hill, New York.
Warsi, Z.U.A. 1999. *Fluid Dynamics*: *Theoretical and Computational Approaches*, 2nd ed., CRC Press, Boca Raton, FL.

2.3 SIMILITUDE: DIMENSIONAL ANALYSIS AND DATA CORRELATION

Stuart W. Churchill

Dimensional Analysis

Similitude refers to the formulation of a description for physical behavior that is general and independent of the individual dimensions, physical properties, forces, etc. In this subsection, the treatment of similitude is restricted to *dimensional analysis*; for a more general treatment see Zlokarnik (1991, 2006). The full power and utility of dimensional analysis is often underestimated and underutilized by engineers. This technique may be applied to a complete mathematical model or to a simple listing of the variables that define the behavior. Only the latter application is described here. For a description of the application of dimensional analysis to a mathematical model see Hellums and Churchill (1964).

General Principles

Dimensional analysis is based on the principle that all additive or equated terms of a complete relationship between the variables must have the same net dimensions. The analysis starts with the preparation of a list of the individual dimensional variables (dependent, independent, and parametric) that are presumed to define the behavior of interest. The performance of dimensional analysis in this context is reasonably simple and straightforward; the principal difficulty and uncertainty arise from the identification of the variables to be included or excluded. If one or more important variables are inadvertently omitted, the reduced description achieved by dimensional analysis will be incomplete and inadequate as a guide for the correlation of a full range of experimental data or theoretical values. The familiar band of plotted values in many graphical correlations is more often a consequence of the omission of one or more variables than of inaccurate measurements. If, on the other hand, one or more irrelevant or unimportant variables are included in the listing, the consequently reduced description achieved by dimensional analysis will result in one or more unessential dimensionless groupings. Such excessive dimensionless groupings are generally less troublesome than missing ones because the redundancy will ordinarily be revealed by the process of correlation. Excessive groups may, however, suggest unnecessary experimental work or computations, or result in misleading correlations. For example, real experimental scatter may inadvertently and incorrectly be correlated in all or in part with the variance of the excessive grouping.

In consideration of the inherent uncertainty in selecting the appropriate variables for dimensional analysis, it is recommended that this process be interpreted as *speculative* and subject to correction on the basis of experimental data or other information. Speculation may also be utilized as a formal technique to identify the effect of eliminating a variable or of combining two or more. The latter aspect of speculation, which may be applied either to the original listing of dimensional variables or to the resulting set of dimensionless groups, is often of great utility in identifying possible limiting behavior or dimensionless groups of marginal significance. The systematic speculative elimination of all but the most certain variables, one at a time, followed by regrouping, is recommended as a general practice. The additional effort as compared with the original dimensional analysis is minimal, but the possible return is very high. A general discussion of this process may be found in Churchill (1981).

The minimum number of independent dimensionless groups i required to describe the fundamental and parametric behavior is (Buckingham, 1914)

$$i = n - m \tag{2.3.1}$$

where n is the number of variables and m is the number of fundamental dimensions such as mass M, length L, time θ, and temperature T that are introduced by the variables. The inclusion of redundant dimensions such as force F and energy E that may be expressed in terms of mass, length, time, and temperature is at the expense of added complexity and is to be avoided. (Of course, mass could be replaced by force or temperature by energy as alternative fundamental dimensions.) In some rare cases i is actually greater than $n-m$. Then

$$i = n - k \tag{2.3.2}$$

where k is the maximum number of the chosen variables that cannot be combined to form a dimensionless group. Determination of the minimum number of dimensionless groups is helpful if the groups are to be chosen by inspection, but is unessential if the algebraic procedure described below is utilized to determine the groups themselves since the number is then obvious from the final result.

The *particular* minimal set of dimensionless groups is arbitrary in the sense that two or more of the groups may be multiplied together to any positive, negative, or fractional power as long as the number of independent groups is unchanged. For example, if the result of a dimensional analysis is

$$\phi\{X, Y, Z\} = 0 \tag{2.3.3}$$

where X, Y, and Z are independent dimensionless groups, an equally valid expression is

$$\phi\{XY^{1/2}, Z/Y^2, Z\} = 0 \tag{2.3.4}$$

Dimensional analysis itself does not provide any insight as to the best choice of equivalent dimensionless groupings, such as between those of Equations 2.3.3 and 2.3.4. However, isolation of each of the variables that are presumed to be the most important in a separate group may be convenient in terms of interpretation and correlation. Another possible criterion in choosing between alternative groupings may be the relative invariance of a particular one. The functional relationship provided by Equation 2.3.3 may equally well be expressed as

$$X = \phi\{Y, Z\} \tag{2.3.5}$$

where X is implied to be the dependent grouping and Y and Z to be independent or parametric groupings.

Three primary methods of determining a minimal set of dimensionless variables are (1) by inspection; (2) by combination of the residual variables, one at a time, with a set of chosen variables that cannot be combined to obtain a dimensionless group; and (3) by an algebraic procedure. These methods are illustrated in the examples that follow.

Example 2.3.1: Fully Developed Flow of Water through a Smooth Round Pipe

Choice of Variables. The shear stress τ_w on the wall of the pipe may be postulated to be a function of the density ρ and the dynamic viscosity μ of the water, the inside diameter D of the pipe, and the space-mean of the time-mean velocity u_m. The limitation to fully developed flow is equivalent to a postulate of independence from distance x in the direction of flow, and the specification of a smooth pipe is equivalent to the postulate of independence from the roughness e of the wall. The choice of τ_w rather than the pressure drop per unit length $-dP/dx$ avoids the need to include the acceleration due to gravity g and the elevation z as variables. The choice of u_m rather than the volumetric rate of flow V, the mass rate of flow w, or the mass rate of flow per unit area G is arbitrary but has some important consequences as noted below. The postulated dependence may be expressed functionally as $\phi\{\tau_w, \rho, \mu, D, u_m\}=0$ or $\tau_w=\phi\{\rho, \mu, D, u_m\}$.

Tabulation. Next prepare a tabular listing of the variables and their dimensions:

	τ_w	ρ	μ	D	u_m
M	1	1	1	0	0
L	–1	–3	–1	1	1
θ	–2	0	–1	0	–1
T	0	0	0	0	0

Minimal Number of Groups. The number of postulated variables is 5. Since the temperature does not occur as a dimension for any of the variables, the number of fundamental dimensions is 3. From Equation 2.3.1, the minimal number of dimensionless groups is $5-3=2$. From inspection of the above tabulation, a dimensionless group cannot be formed from as many as three variables such as D, μ, and ρ. Hence, Equation 2.3.2 also indicates that $i=5-3=2$.

Method of Inspection. By inspection of the tabulation or by trial and error it is evident that only two independent dimensionless groups may be formed. One such set is

$$\phi\left\{\frac{\tau_w}{\rho u_m^2}, \frac{Du_m\rho}{\mu}\right\}=0$$

Method of Combination. The residual variables τ_w and μ may be combined in turn with the non-combining variables ρ, D, and u_m to obtain two groups such as those above.

Algebraic Method. The algebraic method makes formal use of the postulate that the functional relationship between the variables may in general be represented by a power series. In this example such a power series may be expressed as

$$\tau_w = \sum_{i=1}^{N} A_i \rho^{a_i} \mu^{b_i} D^{c_i} u_m^{d_i}$$

where the coefficients A_i are dimensionless. Each additive term on the right-hand side of this expression must have the same net dimensions as τ_w. Hence, for the purposes of dimensional analysis, only the first term need be considered and the indexes may be dropped. The resulting highly restricted expression is $\tau_w = A\varrho^a \mu^b D^c u_m^d$. Substituting the dimensions for the variables gives

$$\frac{M}{L\theta^2} = A\left(\frac{M}{L^3}\right)^a \left(\frac{M}{L\theta}\right)^b L^c \left(\frac{L}{\theta}\right)^d$$

Equating the sum of the exponents of M, L, and θ on the right-hand side of the above expression with those of the left-hand side produces the following three simultaneous linear algebraic equations: $1=a+b$; $-1=-3a-b+c+d$; and $-2=-b-d$, which may be solved for a, c, and d in terms of b to obtain $a=1-b$, $c=-b$, and $d=2-b$. Substitution then gives $\tau_w = A\rho^{1-b}\mu^b D^{-b} u_m^{2-b}$ which may be regrouped as

$$\frac{\tau_w}{\rho u_m^2} = A\left(\frac{\mu}{Du_m\rho}\right)^b$$

Since this expression is only the first term of a power series, it should *not* be interpreted to imply that $\tau_w/\varrho u_m^2$ necessarily proportional to some power at $\mu/Du_m\rho$ but instead only the equivalent of the expression derived by the method of inspection. The inference of a power dependence between the dimensionless groups is the most common and serious error in the use of the algebraic method of dimensional analysis.

Speculative Reductions. Eliminating ρ from the original list of variables on speculative grounds leads to

$$\phi\left\{\frac{\tau_w D}{\mu u_m}\right\} = 0$$

or its exact equivalent:

$$\frac{\tau_w D}{\mu u_m} = A$$

The latter expression with $A=8$ is actually the exact solution for the laminar regime ($Du_m\rho/\mu < 1800$). A relationship that does not include ρ may alternatively be derived directly from the solution by the method of inspection as follows. First, ρ is eliminated from one group, say $\tau_w/\rho u_m^2$, by multiplying it with $Du_m\rho/\mu$ to obtain

$$\phi\left\{\frac{\tau_w D}{\mu u_m}, \frac{Du_m\rho}{\mu}\right\} = 0$$

The remaining group containing ρ is now simply dropped. Had the original expression been composed of three independent groups each containing ρ, that variable would have to be eliminated from two of them before dropping the third one.

The relationships that are obtained by the speculative elimination of μ, D, and u_m, one at a time, do not appear to have any range of physical validity. Furthermore, if w or G had been chosen as the independent variable rather than u_m, the limiting relationship for the laminar regime would not have been obtained by the elimination of ρ.

Alternative Forms. The solution may also be expressed in an infinity of other forms such as

$$\phi\left\{\frac{\tau_w D^2\rho}{\mu^2}, \frac{Du_m\rho}{\mu}\right\} = 0$$

If τ_w is considered to be the principal dependent variable and u_m the principal independent variable, this latter form is preferable in that these two quantities do not then appear in the same grouping. On the other hand, if D is considered to be the principal independent variable, the original formulation is preferable. The variance of $\tau_w/\rho u_m^2$ is less than that of $\tau_w D/\mu u_m$ and $\tau_w D^2\rho/\mu^2$ in the turbulent regime while that of $\tau_w D/\mu u_m$ is zero in the laminar regime. Such considerations may be important in devising convenient graphical correlations.

Alternative Notations. The several solutions above are more commonly expressed as

$$\phi\left\{\frac{f}{2}, \mathrm{Re}\right\} = 0$$

$$\phi\left\{\frac{f\,\mathrm{Re}}{2}, \mathrm{Re}\right\} = 0$$

or

$$\phi\left\{\frac{f\,\mathrm{Re}^2}{2}, \mathrm{Re}\right\} = 0$$

where $f = 2\tau_w/\rho u_m^2$ is the *Fanning friction factor* and $Re = Du_m\rho/\mu$ is the *Reynolds number*.

The more detailed forms, however, are to be preferred for purposes of interpretation or correlation because of the explicit appearance of the individual, physically measurable variables.

Addition of a Variable. The above results may readily be extended to incorporate the roughness e of the pipe as a variable. If two variables have the same dimensions, they will always appear as

a dimensionless group in the form of a ratio, in this case e appears most simply as e/D. Thus, the solution becomes

$$\phi\left\{\frac{\tau_w}{\rho u_m^2}, \frac{Du_m\rho}{\mu}, \frac{e}{D}\right\} = 0$$

Surprisingly, as contrasted with the solution for a smooth pipe, the speculative elimination of μ and hence of the group $Du_m\rho/\mu$ now results in a valid asymptote for $Du_m\rho/\mu \to \infty$ and all finite values of e/D, namely,

$$\phi\left\{\frac{\tau_w}{\rho u_m^2}, \frac{e}{D}\right\} = 0$$

Example 2.3.2: Fully Developed Forced Convection in Fully Developed Flow in a Round Tube

It may be postulated for this process that $h=\phi\{D, u_m, \rho, \mu, k, c_p\}$, where here h is the local heat-transfer coefficient, and c_p and k are the specific heat capacity and thermal conductivity, respectively, of the fluid. The corresponding tabulation is

	h	D	u_m	ρ	μ	K	c_p
M	1	0	0	1	1	1	0
L	0	1	1	–3	–1	1	2
θ	–3	0	–1	0	–1	–3	–2
T	–1	0	0	0	0	–1	–1

The number of variables is 7 and the number of independent dimensions is 4, as is the number of variables such as D, u_m, ρ, and k that cannot be combined to obtain a dimensionless group. Hence, the minimal number of dimensionless groups is $7-4=3$. The following acceptable set of dimensionless groups may be derived by any of the procedures illustrated in Example 2.3.1:

$$\frac{hD}{k} = \phi\left\{\frac{Du_m\rho}{\mu}, \frac{c_p\mu}{k}\right\}$$

Speculative elimination of μ results in

$$\frac{hD}{k} = \phi\left\{\frac{Du_m\rho c_p}{k}\right\}$$

which has often erroneously been inferred to be a valid asymptote for $c_p\mu/k \to 0$. Speculative elimination of D, u_m, ρ, k, and c_p individually also does not appear to result in expressions with any physical validity. However, eliminating c_p and ρ or u_m gives a valid result for the laminar regime, namely,

$$\frac{hD}{k} = A$$

The general solutions for flow and convection in a smooth pipe may be combined to obtain

$$\frac{hD}{k} = \phi\left\{\frac{\tau_w D^2 \rho}{\mu^2}, \frac{c_p \mu}{k}\right\}$$

which would have been obtained directly had u_m been replaced by τ_w in the original tabulation. This latter expression proves to be superior in terms of speculative reductions. Eliminating D results in

$$\frac{h\mu}{k(\tau_w \rho)^{1/2}} = \phi\left\{\frac{c_p \mu}{k}\right\}$$

which may be expressed in the more conventional form of

$$\mathrm{Nu} = \mathrm{Re}\left(\frac{f}{2}\right)^{1/2} \phi\{\mathrm{Pr}\}$$

where $\mathrm{Nu} = hD/k$ is the *Nusselt number* and $\mathrm{Pr} = c_p\mu/k$ is the *Prandtl number*. This result appears to be a valid asymptote for $\mathrm{Re} \to \infty$ and a good approximation for even moderate values (Re >5000) for large values of Pr. Elimination of μ as well as D results in

$$\frac{h}{c_p(\tau_w \rho)^{1/2}} = A$$

or

$$\mathrm{Nu} = A\,\mathrm{Re}\,\mathrm{Pr}\left(\frac{f}{2}\right)^{1/2}$$

which appears to be an approximate asymptote for $\mathrm{Re} \to \infty$ and $\mathrm{Pr} \to 0$. Elimination of both c_p and ρ again yields the appropriate result for laminar flow, indicating that ρ rather than u_m is the meaningful variable to eliminate in this respect.

The numerical value of the coefficient A in the several expressions above depends on the mode of heating, a true variable, but one from which the purely functional expressions are independent. If j_w, the heat flux density at the wall, and $T_w - T_m$, the temperature difference between the wall and the bulk of the fluid, were introduced as variables in place of $h \equiv j_w/(T_w - T_m)$, another group such as $c_p(T_w - T_m)(D\rho/\mu)^2$ or $\rho c_p(T_w - T_m)/\tau_w$ or $c_p(T_w - T_m)/u_m^2$ which represents the effect of viscous dissipation, would be obtained. This effect is usually but not always negligible.

Example 2.3.3: Free Convection from a Vertical Isothermal Plate

The behavior for this process may be postulated to be represented by

$$h = \phi\{g, \beta, (T_w - T_\infty), x, \mu, \rho, c_p, k\}$$

where g is the acceleration due to gravity, β is the volumetric coefficient of expansion with temperature, T_∞ is the unperturbed temperature of the fluid, and x is the vertical distance along the plate. The corresponding tabulation is

	h	g	β	$T_w - T_\infty$	x	μ	ρ	c_p	k
M	1	0	0	0	0	1	1	0	1
L	0	1	0	0	1	–1	–3	2	1
θ	–3	–2	0	0	0	–1	0	–2	–3
T	–1	0	–1	1	0	0	0	–1	1

The minimal number of dimensionless groups indicated by both methods is $9-4=5$. A satisfactory set of dimensionless groups, as found by any of the methods illustrated in Example 2.3.1, is

$$\frac{hx}{k}=\phi\left\{\frac{\rho^2 g x^3}{\mu^2},\frac{c_p\mu}{k},\beta\left(T_w-T_\infty\right),c_p\left(T_w-T_\infty\right)\left(\frac{\rho x}{\mu}\right)^2\right\}$$

It may be reasoned that the buoyant force which generates the convective motion must be proportional to $\rho g\beta(T_w-T_\infty)$, thus, g in the first term on the right-hand side must be multiplied by $\beta(T_w-T_\infty)$, resulting in

$$\frac{hx}{k}=\phi\left\{\frac{\rho^2 g\beta\left(T_w-T_\infty\right)x^3}{\mu^2},\frac{c_p\mu}{k},\beta\left(T_w-T_\infty\right),c_p\left(T_w-T_\infty\right)\left(\frac{\rho x}{\mu}\right)^2\right\}$$

The effect of expansion other than on the buoyancy is now represented by $\beta(T_w-T_\infty)$, and the effect of viscous dissipation by $c_p(T_w-T_\infty)(\rho x/\mu)^2$. Both effects are negligible for all practical circumstances. Hence, this expression may be reduced to

$$\frac{hx}{k}=\phi\left\{\frac{\rho^2 g\beta\left(T_w-T_\infty\right)x^3}{\mu^2},\frac{c_p\mu}{k}\right\}$$

or

$$\mathrm{Nu}_x=\phi\left\{\mathrm{Gr}_x,\mathrm{Pr}\right\}$$

where $\mathrm{Nu}_x=hx/k$ and $\mathrm{Gr}_x=\rho^2 g\beta(T_w-T_\infty)x^3/\mu^2$ is the *Grashof number*.

Elimination of x speculatively now results in

$$\frac{hx}{k}=\left(\frac{\rho^2 g\beta\left(T_w-T_\infty\right)x^3}{\mu^2}\right)^{1/3}\phi\left\{\mathrm{Pr}\right\}$$

or

$$\mathrm{Nu}_x=\mathrm{Gr}_x^{1/3}\phi\{\mathrm{Pr}\}$$

This expression appears to be a valid asymptote for $\mathrm{Gr}_x\to\infty$ and a good approximation for the entire turbulent regime. Eliminating μ speculatively rather than x results in

$$\frac{hx}{k}=\phi\left\{\frac{\rho^2 c_p^2 g\beta\left(T_w-T_\infty\right)x^3}{k^2}\right\}$$

or

$$\mathrm{Nu}_x=\phi\left\{\mathrm{Gr}_x\mathrm{Pr}^2\right\}$$

The latter expression appears to be a valid asymptote for $\mathrm{Pr}\to 0$ for all Gr_x, that is, for both the laminar and the turbulent regimes. The development of a valid asymptote for large values of Pr requires more subtle reasoning. First $c_p\mu/k$ is rewritten as $\mu/\rho\alpha$ where $\alpha=k/\rho c_p$. Then ρ is eliminated speculatively except as it occurs in $\rho g\beta(T_w-T_\infty)$ and $k/\rho c_p$. The result is

$$\frac{hx}{k}=\phi\left\{\frac{c_p\rho^2 g\beta(T_w-T_\infty)x^3}{\mu k}\right\}$$

or

$$\mathrm{Nu}_x=\phi\left\{\mathrm{Ra}_x\right\}$$

where

$$\mathrm{Ra}_x = \frac{c_p \rho^2 g \beta (T_w - T_\infty) x^3}{\mu k} = \mathrm{Gr}_x \mathrm{Pr}$$

is the *Rayleigh number*. The above expression for Nu_x appears to be a valid asymptote for $\mathrm{Pr} \to \infty$ and a reasonable approximation for even moderate values of Pr for all Gr_x, that is, for both the laminar and the turbulent regimes.

Eliminating x speculatively from the above expressions for small and large values of Pr results in

$$\mathrm{Nu}_x = A\left(\mathrm{Gr}_x \mathrm{Pr}^2\right)^{1/3} = A\left(\mathrm{Ra}_x \mathrm{Pr}\right)^{1/3}$$

and

$$\mathrm{Nu}_x = B\left(\mathrm{Gr}_x \mathrm{Pr}\right)^{1/3} = B\left(\mathrm{Ra}_x\right)^{1/3}$$

The former appears to be a valid asymptote for $\mathrm{Pr} \to 0$ and $\mathrm{Gr}_x \to \infty$ and a reasonable approximation for very small values of Pr in the turbulent regime, while the latter is well confirmed as a valid asymptote for $\mathrm{Pr} \to \infty$ and $\mathrm{Gr}_x \to \infty$ and as a good approximation for moderate and large values of Pr over the entire turbulent regime. The expressions in terms of Gr_x are somewhat more complicated than those in terms of Ra_x, but are to be preferred since Gr_x is known to characterize the transition from laminar to turbulent motion in natural convection just as Re_D does in forced flow in a channel. The power of speculation combined with dimensional analysis is well demonstrated by this example in which valid asymptotes are thereby attained for several regimes.

Correlation of Experimental Data and Theoretical Values

Correlations of experimental data are generally developed in terms of dimensionless groups rather than in terms of the separate dimensional variables in the interests of compactness and in the hope of greater generality. For example, a complete set of graphical correlations for the heat-transfer coefficient h of Example 2.3.2 in terms of each of the six individual independent variables and physical properties might approach book length, whereas the dimensionless groupings both imply that a single plot with one parameter should be sufficient. Furthermore, the reduced expression for the turbulent regime implies that a plot of $\mathrm{Nu}/\mathrm{Re}\, f^{1/2}$ vs. Pr should demonstrate only a slight parametric dependence on Re or $\mathrm{Re}\, f^{1/2}$. Of course, the availability of a separate correlation for f as a function of Re is implied.

Theoretical values, that is, ones obtained by numerical solution of a mathematical model in terms of either dimensional variables or dimensionless groups, are presumably free from imprecision. Even so, because of their discrete form, the construction of a correlation or correlations for such values may be essential for the same reasons as for experimental data.

Graphical correlations have the merit of revealing general trends, of providing a basis for evaluation of the choice of coordinates, and most of all of displaying visually the scatter of the individual experimental values about a curve representing a correlation or their behavior on the mean. As mentioned in the previous subsection, the omission of a variable may, however give the false impression of experimental error in such a plot. On the other hand, correlating equations are far more convenient as an input to a computer than is a graphical correlation. These two formats thus have distinct and complementary roles; both should generally be utilized. The merits and demerits of various graphical forms of correlations are discussed in detail by Churchill (1979), while the use of logarithmic and arithmetic coordinates, the effects of the appearance of a variable in both coordinates, and the effects of the distribution of error between the dependent and independent variable are further illustrated by Wilkie (1985).

Churchill and Usagi (1972, 1974) proposed general usage of the following expression for the formulation of correlating equations:

$$y^n\{x\} = y_0^n\{x\} + y_\infty^n\{x\} \tag{2.3.6}$$

where $y_0\{x\}$ and $y_\infty\{x\}$ denote asymptotes for small and large values of x, respectively, and n is an arbitrary exponent. For convenience and simplicity, Equation 2.3.6 may be rearranged in either of the following two forms:

$$\left(Y(x)\right)^n = 1 + Z^n\{x\} \tag{2.3.7}$$

or

$$\left(\frac{Y\{x\}}{Z\{x\}}\right)^n = 1 + \frac{1}{Z^n\{x\}} \tag{2.3.8}$$

where $Y\{x\} \equiv y\{x\}/y_0\{x\}$ and $Z\{x\} \equiv y_\infty\{x\}/y_0\{x\}$. Equations 2.3.6, 2.3.7, and 2.3.8 are hereafter denoted collectively as the CUE (Churchill–Usagi equation). The principal merits of the CUE as a canonical expression for correlation are its simple form, generality, and minimal degree of explicit empiricism, namely, only that of the exponent n, since the asymptotes $y_0\{x\}$ and $y_\infty\{x\}$ are ordinarily known in advance from theoretical considerations or well-established correlations. Furthermore, as will be shown, the CUE is quite insensitive to the numerical value of n. Although the CUE is itself very simple in form, it is remarkably successful in representing closely very complex behavior, even including the dependence on secondary variables and parameters, by virtue of the introduction of such dependencies through $y_0\{x\}$ and $y_\infty\{x\}$. In the rare instances in which such dependencies are not represented in the asymptotes, n may be correlated as a function of the secondary variables and/or parameters. Although the CUE usually produces very close representations, it is empirical and not exact. In a few instances, numerical values of n have been derived or rationalized on theoretical grounds, but even then some degree of approximation is involved. Furthermore, the construction of a correlating expression in terms of the CUE is subject to the following severe limitations:

1. The asymptotes $y_0\{x\}$ and $y_\infty\{x\}$ must intersect once and only once;
2. The asymptotes $y_0\{x\}$ and $y_\infty\{x\}$ must be free of singularities. Even though a singularity occurs beyond the asserted range of the asymptote, it will persist and disrupt the prediction of the CUE, which is intended to encompass all values of the independent variable *x;* and
3. The asymptotes must both be upper or lower bounds.

In order to avoid or counter these limitations, it may be necessary to modify or replace the asymptotes with others. Examples of this process are provided below. A different choice for the dependent variable may be an option in this respect. The suitable asymptotes for use in Equation 2.3.6 may not exist in the literature and therefore may need to be devised or constructed. See, for example, Churchill (1988b) for guidance in this respect. Integrals and derivatives of the CUE are generally awkward and inaccurate, and may include singularities not present or troublesome in the CUE itself. It is almost always preferable to develop a separate correlating equation for such quantities using derivatives or integrals of $y_0\{x\}$ and $y_\infty\{x\}$, simplified or modified as appropriate.

The Evaluation of *n*

Equation 2.3.6 may be rearranged as

$$n = \frac{\ln\left\{1+\left(\dfrac{y_\infty\{x\}}{y_0\{x\}}\right)^n\right\}}{\ln\left\{\dfrac{y\{x\}}{y_0\{x\}}\right\}} \tag{2.3.9}$$

and solved for n by iteration for any known value of $y\{x\}$, presuming that $y_0\{x\}$ and $y_\infty\{x\}$ are known. If $y\{x^*\}$ is known, where x^* represents the value of x at the point of intersection of the asymptotes, that is, for $y_0\{x\} = y_\infty\{x\}$, Equation 2.3.9 reduces to

$$n = \frac{\ln\{2\}}{\ln\left\{\dfrac{y\{x^*\}}{y_0\{x^*\}}\right\}} \tag{2.3.10}$$

and iterative determination of n is unnecessary.

A graphical and visual method of evaluation of n is illustrated in Figure 2.3.1 in which $Y\{Z\}$ is plotted vs. Z for $0 \le Z \le 1$ and $Y\{Z\}/Z$ vs. $1/Z$ for $0 \le 1/Z \le 1$ in arithmetic coordinates with n as a parameter. Values of $y\{x\}$ may be plotted in this form and the best overall value of n selected visually (as illustrated in Figure 2.3.2). A logarithmic plot of $Y\{Z\}$ vs. Z would have less sensitivity relative to the dependence on n. (See, for example, Figure 1 of Churchill and Usagi, 1972.). Figure 2.3.1 explains in part the success of the CUE. Although y and x may both vary from 0 to ∞, the composite variables plotted in Figure 2.3.1 are highly constrained in that the compound independent variables Z and $1/Z$ vary only between 0 and 1, while for $n \ge 1$, the compound dependent variables $Y\{Z\}$ and $Y\{Z\}/Z$ vary only from 1 to 2.

Because of the relative insensitivity of the CUE to the numerical value of n, an integer or a ratio of two small integers may be chosen in the interest of simplicity and without significant loss of accuracy. For example, the maximum variance in Y (for $0 \le Z \le 1$) occurs at $Z = 1$ and increases only $100(2^{1/20} - 1) = 3.5\%$ if n is decreased from 5 to 4. If $y_0\{x\}$ and $y_\infty\{x\}$ are both lower bounds, n will be positive, and if they are both upper bounds, n will be negative. To avoid extending Figure 2.3.1 for negative values of n, $1/y\{x\}$ may simply be interpreted as the dependent variable.

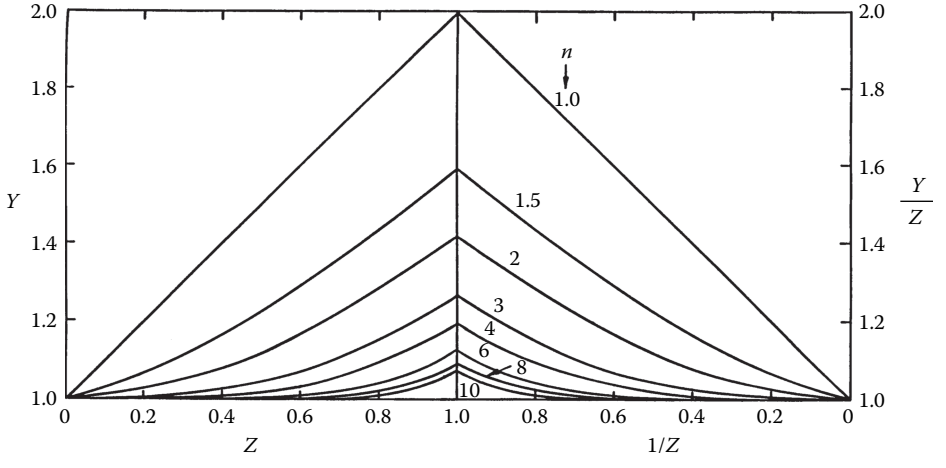


FIGURE 2.3.1 Arithmetic, split-coordinate plot of Equation 2.3.10. (From Churchill, S.W. and Usagi, R., *AIChE J.*, 18, 1123, 1972. With permission from the American Institute of Chemical Engineers.)

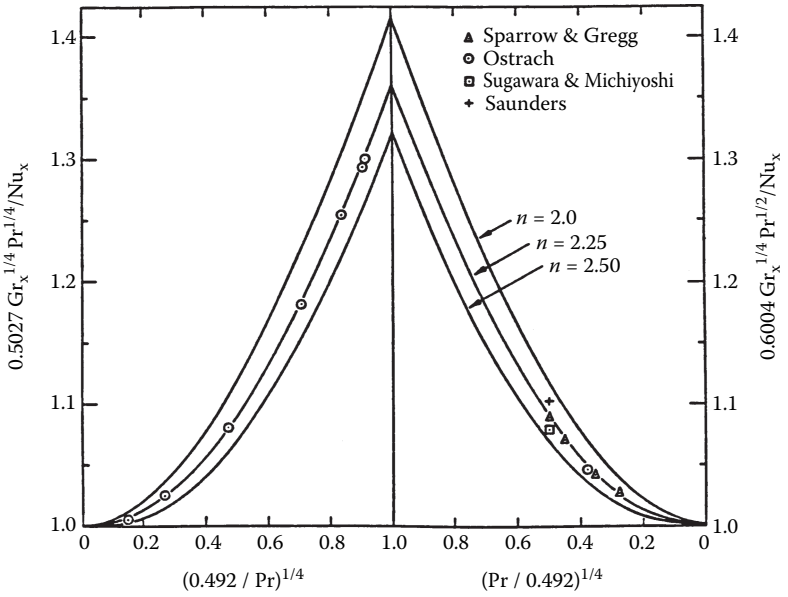


FIGURE 2.3.2 Arithmetic, split-coordinate plot of computed values and experimental data for laminar free convection from an isothermal vertical plate. (From Churchill, S.W. and Usagi, R., *AIChE J.*, 18, 1124, 1972. With permission from the American Institute of Chemical Engineers.)

Intermediate Regimes

Equations 2.3.6, 2.3.7, and 2.3.8 imply a slow, smooth transition between $y_0\{x\}$ and $y_\infty\{x\}$ and, moreover, one that is symmetrical with respect to $x^*(Z=1)$. Many physical systems demonstrate instead a relatively abrupt transition, as for example from laminar to turbulent flow in a channel or along a flat plate. The CUE may be applied serially as follows to represent such behavior if an expression $y_i\{x\}$ is postulated for the intermediate regime. First, the transition from the initial to the intermediate regime is represented by

$$y_1^n = y_0^n + y_i^n \tag{2.3.11}$$

Then the transition from this combined regime to the final regime by

$$y^m = y_1^m + y_\infty^m = \left(y_0^n + y_i^n\right)^{m/n} + y_\infty^m \tag{2.3.12}$$

Here, and throughout the balance of this subsection, in the interests of simplicity and clarity, the functional dependence of all the terms on x is implied rather than written out explicitly. If y_0 is a lower bound and y_i is implied to be one, y_1 and y_∞ must be upper bounds. Hence, n will then be positive and m negative. If y_0 and y_i are upper bounds, y_1 and y_∞ must be lower bounds; then n will be negative and m positive. The reverse formulation starting with y_∞ and y_1 leads by the same procedure to

$$y^n = y_0^n + \left(y_i^m + y_\infty^m\right)^{n/m} \tag{2.3.13}$$

If the intersections of y_i with y_0 and y_∞ are widely separated with respect to x, essentially the same pair of values for n and m will be determined for Equations 2.3.12 and 2.3.13, and the two representations for y will not differ significantly. On the other hand, if these intersections are close in terms of x, the pair of values of m and n may differ significantly and one representation may be quite superior to the other. In some instances a singularity in y_0 or y_∞ may be tolerable in either Equation

2.3.12 or 2.3.13 because it is overwhelmed by the other terms. Equations 2.3.12 and 2.3.13 have one hidden flaw. For $x \to 0$, Equation 2.3.12 reduces to

$$y \to y_0 \left[1 + \left(\frac{y_\infty}{y_0} \right)^m \right]^{1/m} \tag{2.3.14}$$

If y_0 is a lower bound, m is necessarily negative, and values of y less than y_0 are predicted. If y_0/y_∞ is sufficiently small or if m is sufficiently large in magnitude, this discrepancy may be tolerable. If not, the following alternative expression may be formulated, again starting from Equation 2.3.11:

$$\left(y^n - y_0^n \right)^m = y_i^{nm} + \left(y_\infty^n - y_0^n \right)^m \tag{2.3.15}$$

Equation 2.3.15 is free from the flaw identified by means of Equation 2.3.14 and invokes no additional empiricism, but a singularity may occur at $y_\infty = y_0$, depending on the juxtapositions of y_0, y_i, and y_∞. Similar anomalies occur for Equation 2.3.13 and the corresponding analog of Equation 2.3.14, as well as for behavior for which $n < 0$ and $m > 0$. The preferable form among these four is best chosen by trying each of them.

One other problem with the application of the CUE for a separate transitional regime is the formulation of an expression for $y_i\{x\}$, which is ordinarily not known from theoretical considerations. Illustrations of the empirical determination of such expressions for particular cases may be found in Churchill and Usagi (1974), Churchill and Churchill (1975), and Churchill (1976, 1977), as well as in Example 2.3.5 below.

Example 2.3.4: Pressure Gradient in Flow through a Packed Bed of Spheres

The pressure gradient at asymptotically low rates of flow (the creeping regime) can be represented by the Kozeny–Carman equation, $\Phi = 150\ \mathrm{Re}_p$, and at asymptotically high rates of flow (the inertial regime) by the Burke–Plummer equation, $\Phi = 1.75\ (\mathrm{Re}_p)^2$, where $\Phi = \rho\varepsilon^2 d_p(-\mathrm{d}P_f/\mathrm{d}x)\mu^2(1-\varepsilon)$, $\mathrm{Re}_p = d_p u_0 \rho/\mu(1-\varepsilon)$, d_p = diameter of spherical particles, m, ε = void fraction of bed of spheres, $\mathrm{d}P_f/\mathrm{d}x$ = dynamic pressure gradient (due to friction), Pa/m, and u_0 = superficial velocity (in absence of the spheres), m/s. For the origin of these two asymptotic expressions see Churchill (1988a). They both have a theoretical structure, but the numerical coefficients of 150 and 1.75 are basically empirical. These equations are both lower bounds and have one intersection. Experimental data are plotted in Figure 2.3.3, which has the form of Figure 2.3.1 with $\Upsilon = \Phi/150\ \mathrm{Re}_p$, $Y/Z = \Phi/(1.75\ \mathrm{Re}_p)^2$ and $Z = 1.75\mathrm{Re}_p^2/150\mathrm{Re}_p = \mathrm{Re}_p/85.7$. A value of $n = 1$ is seen to represent these data reasonably well on the mean, resulting in

$$\Phi = 150\mathrm{Re}_p + 1.75\left(\mathrm{Re}_p\right)^2$$

which was originally proposed as a correlating equation by Ergun (1952) on the conjecture that the volumetric fraction of the bed in "turbulent" flow is proportional to Re_p. The success of this expression in conventional coordinates is shown in Figure 2.3.4. The scatter, which is quite evident in the arithmetic split coordinates of Figure 2.3.3, is strongly suppressed in a visual sense in the logarithmic coordinates of Figure 2.3.4.

Example 2.3.5: Friction Factor for Commercial Pipes for All Conditions

The serial application of the CUE is illustrated here by the construction of a correlating equation for both smooth and rough pipes in the turbulent regime followed by combination of that expression with ones for the laminar and transitional regimes.

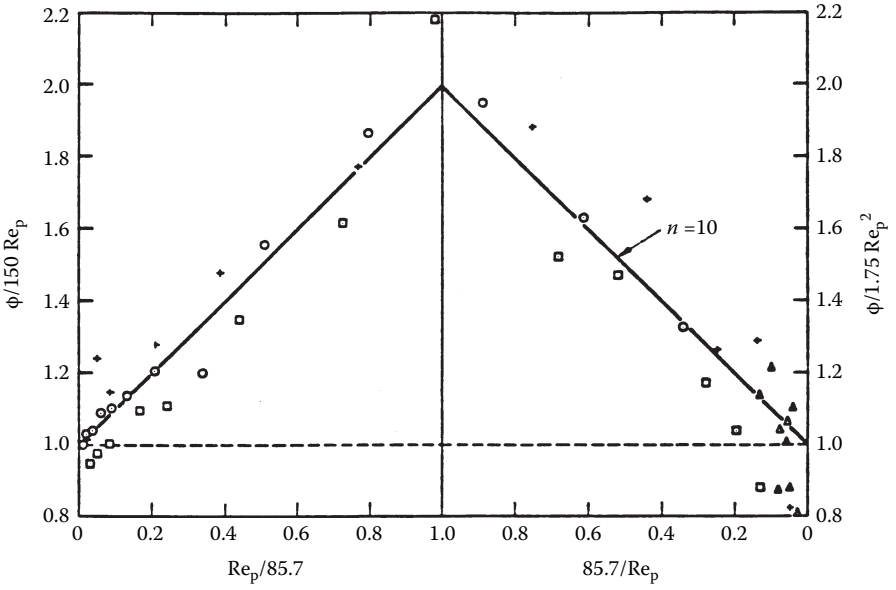


FIGURE 2.3.3 Arithmetic, split-coordinate plot of experimental data for the pressure drop in flow through a packed bed of spheres. (From Churchill, S.W. and Usagi, R., *AIChE J.*, 18, 1123, 1972. With permission from the American Institute of Chemical Engineers.)

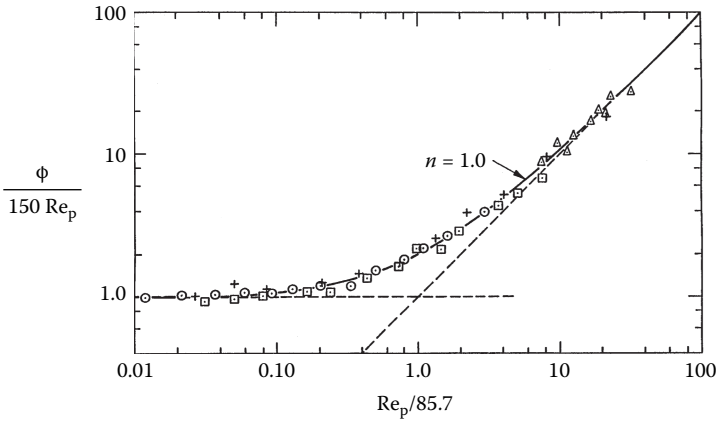


FIGURE 2.3.4 Logarithmic correlation of experimental data for the pressure drop in flow through a packed bed of spheres. (From Churchill, S.W. and Usagi, R., *AIChE J.*, 18, 1123, 1972. With permission from the American Institute of Chemical Engineers.)

The Turbulent Regime. The Fanning friction factor, f_F, for turbulent flow in a smooth round pipe for asymptotically large rates of flow (say $Re_D > 5000$) may be represented closely by the empirical expression:

$$\left(\frac{2}{f_F}\right)^{1/2} = 1.70 - \frac{310}{Re_D\left(\frac{f_F}{2}\right)^{1/2}} + \left(\frac{100}{Re_D\left(\frac{f_F}{2}\right)^{1/2}}\right)^2 + \frac{1}{0.436}\ln\left\{Re_D\left(\frac{f}{2}\right)^{1/2}\right\}$$

A corresponding empirical representation for naturally rough pipe is

$$\left(\frac{2}{f_F}\right)^{1/2} = 4.46 + \frac{1}{0.436}\ln\left\{\frac{D}{e}\right\}$$

Direct combination of these two expressions in the form of the CUE does not produce a satisfactory correlating equation, but their combination in the following rearranged forms:

$$e^{0.436(2/f_F)^{1/2}} = 2.10\,\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2} - e^{-135/\mathrm{Re}_D(f_F/2)^{1/2}} + e^{\left(43.6\Big/\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2}\right)^2}$$

and

$$\dot{e}^{0.430}\left(2/f_F\right)^{1/2} = 6.69\left(\frac{D}{e}\right)$$

with $n=-1$ results in, after the reverse rearrangement,

$$\left(\frac{2}{f_F}\right)^{1/2} = 1.70 + \frac{310}{\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2}} + \left(\frac{100}{\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2}}\right)^2 + \frac{1}{0.436}\ln\left\{\frac{\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2}}{1+0.301\left(\frac{e}{D}\right)\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2}}\right\}$$

The exact equivalent of this expression in structure but with the slightly modified numerical coefficients of 0.300, 2.46, and 0.304 was postulated by Colebrook (1938–1939) to represent his own experimental data. The coefficients of the expression given here are presumed to be more accurate, but the difference in the predictions of f_F with the two sets of coefficients is within the band of uncertainty of the experimental data. The turbulent regime of the "friction-factor" plot in most textbooks and handbooks is simply a graphical representation of the Colebrook equation. Experimental values are not included in such plots since *e*, the effective roughness of commercial pipes, is simply a correlating factor that forces agreement with the Colebrook equation. Values of *e* for various types of pipe in various services are usually provided in an accompanying table, that thereby constitutes an integral part of the correlation.

The Laminar Region. The Fanning friction factor in the laminar regime of a round pipe ($\mathrm{Re}_d < 1800$) is represented exactly by the following theoretical expression known as Poiseuille's law: $f_F = 16/\mathrm{Re}_D$. This equation may be rearranged as follows for convenience in combination with that for turbulent flow:

$$\left(\frac{2}{f_F}\right)^{1/2} = \frac{\mathrm{Re}_D\left(f_F/2\right)^{1/2}}{8}$$

The Transitional Regime. Experimental data as well as semitheoretical computed values for the limiting behavior in the transition may be represented closely by $(f_F/2) = (\mathrm{Re}_D/37500)^2$. This expression may be rewritten, in terms of $(2/f_F)^{1/2}$, and $\mathrm{Re}_D(f_F/2)^{1/2}$, as follows:

$$\left(\frac{f_F}{2}\right)^{1/2} = \left(\frac{37500}{\mathrm{Re}_D\left(f_F/2\right)^{1/2}}\right)^{1/2}$$

Overall Correlation. The following correlating equation for all $\mathrm{Re}_D(f_F/2)^{1/2}$ and e/D may now be constructed by the combination of the expressions for the turbulent and transition regimes in the form of the CUE with $n=8$, and then that expression and that for the laminar regime with $n=-12$,

both components being chosen on the basis of experimental data and predicted values for the full regime of transition:

$$\left(\frac{2}{f_F}\right)^{1/2} = \left[\left(\frac{8}{\mathrm{Re}_D(f_F/2)^{1/2}}\right)^{12} + \left(\left[\frac{37500}{\mathrm{Re}_D(f_F/2)^{1/2}}\right]\right.\right.$$

$$\left.\left. + \left|1.706 - \frac{322.4}{\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2}} + \left(\frac{95.2}{\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2}}\right)^2 + \frac{1}{0.436}\ln\left\{\frac{\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2}}{1+0.301\left(\frac{e}{D}\right)\mathrm{Re}_D\left(\frac{f_F}{2}\right)^{1/2}}\right\}\right|^8\right)^{-3/2}\right]^{-1/12}$$

The absolute value signs are only included for aesthetic reasons; the negative values of the logarithmic term for very small values of $\mathrm{Re}_D(f_F/2)^{1/2}$ do not affect the numerical value of $(2/f_F)^{1/2}$ in the regime in which they occur. This overall expression appears to have a complicated structure, but it may readily be recognized to reduce to its component parts when the corresponding term is large with respect to the other two. It is insensitive to the numerical values of the two arbitrary exponents. For example, doubling their values would have almost no effect on the predictions of $(f_F/2)^{1/2}$. The principal uncertainty is associated with the expression for the transition regime, but the overall effect of the corresponding term is very small. The uncertainties associated with this correlating equation are common to most graphical correlations and algebraic expressions for the friction factor, and are presumed to be fairly limited in magnitude and to be associated primarily with the postulated value of e. Although the overall expression is explicit in $\mathrm{Re}_D(f_F/2)^{1/2}$ rather than Re_D, the latter quantity may readily be obtained simply by multiplying the postulated value of $\mathrm{Re}_D(f_F/2)^{1/2}$ by the computed values of $(2/f_F)^{1/2}$.

REFERENCES

Buckingham, E. 1914. On physically similar systems: illustrations of the use of dimensional equations. *Phys. Rev., Ser. 2*, 4(4):345–375.

Churchill, S.W. 1976. A comprehensive correlating equation for forced convection from plates. *AIChE J.* 22(2):264–268.

Churchill, S.W. 1977. Comprehensive correlating equations for heat, mass and momentum transfer in fully developed flow in smooth tubes. *Ind. Eng. Chem. Fundam.* 16(1):109–116.

Churchill, S.W. 1979. *The Interpretation and Use of Rate Data. The Rate Process Concept*. Hemisphere Publishing Corp., Washington, DC.

Churchill, S.W. 1981. The use of speculation and analysis in the development of correlations. *Chem. Eng. Commun.* 9:19–38.

Churchill, S.W. 1988a. Flow through porous media, Chapter 19 in *Laminar Flows. The Practical Use of Theory*, pp. 501–538, Butterworths, Boston.

Churchill, S.W. 1988b. Derivation, selection, evaluation and use of asymptotes. *Chem. Eng. Technol.* 11:63–72.

Churchill, S.W. and Churchill, R.U. 1975. A general model for the effective viscosity of pseudoplastic and dilatant fluids. *Rheol. Acta*, 14:404–409.

Churchill, S.W. and Usagi, R. 1972. A general expression for the correlation of rates of transfer and other phenomena. *AIChE J.* 18(6):1121–1128.

Churchill, S.W. and Usagi, R. 1974. A standardized procedure for the production of correlations in the form of a common empirical equation. *Ind. Eng. Chem. Fundam.* 13(l):39–44.

Colebrook, C.R. 1938–1939. Turbulent flow in pipes with particular reference to the transition region between the smooth and rough pipe laws. *J. Inst. Civ. Eng.* 11(5024):133–156.

Ergun, S. 1952. Fluid flow through packed beds. *Chem. Eng. Prog.* 48(2):81–96.

Hellums, J.D. and Churchill, S.W. 1964. Simplifications of the mathematical description of boundary and initial value problems. *AIChE J.* 10(1):110–114.

Wilkie, D. 1985. The correlation of engineering data reconsidered. *Int. J. Heat Fluid Flow.* 8(2):99–103.

Zlokarnik, M. 1991. *Dimensional Analysis and Scale-Up in Chemical Engineering*. Springer-Verlag, Berlin.

Zlokarnik, M. 2006. Scale-up in chemical engineering, 2nd., Wiley-VCH, Weinheim, Germany.

2.4 HYDRAULICS OF PIPE SYSTEMS

J. Paul Tullis and Blake Paul Tullis

Basic Computations

Equations

Solving fluid flow problems involves the application of one or more of the three basic equations: continuity, momentum, and energy. These three basic tools are developed from the law of conservation of mass, Newton's second law of motion, and the first law of thermodynamics.

The simplest form of the continuity equation is for one-dimensional incompressible steady flow in a conduit. Applying continuity between any two sections gives

$$A_1V_1 = A_2V_2 = Q \tag{2.4.1}$$

For a variable density fluid, the equation can be written as

$$\rho_1A_1V_1 = \rho_2A_2V_2 = \dot{m} \tag{2.4.2}$$

where A is the cross-sectional area of the pipe, V is the mean flow velocity, Q is the flow rate, ρ is the fluid density, $\dot{m}$ is the mass flow rate, and the subscripts represent different system locations. The equations are valid for any rigid conduit as long as there is no addition or loss of liquid within the section.

For steady-state pipe flow, the momentum equation relates the net force in a given direction (F_x) acting on a control volume (a section of the fluid inside the pipe), to the net momentum flux through the control volume:

$$F_x = \rho_2A_2V_2V_{2x} - \rho_1A_1V_1V_{1x} \tag{2.4.3}$$

For incompressible flow, this equation can be reduced to

$$F_x = \rho Q\left(V_{2x} - V_{1x}\right) \tag{2.4.4}$$

These equations can easily be applied to a three-dimensional flow problem by adding equations for the y- and z-directions.

The following is a general form of the energy equation (see Chapter 1) applicable to incompressible pipe or duct flow:

$$\frac{P_1}{\gamma} + Z_1 + \frac{V_1^2}{2g} = \frac{p_2}{\gamma} + z_2 + \frac{V_2^2}{2g} - H_p + H_t + H_f \tag{2.4.5}$$

The units are energy per unit weight of liquid: ft·lb/lb or N·m/N, which reduce to ft or m. The first three terms are pressure head (P/γ), elevation head (Z) (relative to some datum), and velocity head ($V^2/2g$). The last three terms on the right side of the equation are the total dynamic head added by a pump (H_p) or removed by a turbine (H_t) and the friction plus minor head losses (H_f). The sum of the first three terms in Equation 2.4.5 is defined as the total head, and the sum of the pressure and elevation heads is referred to as the piezometric head.

The purpose of this section is to determine the pressure changes resulting from incompressible flow in pipe systems. Since pipes of circular cross-sections are most common in engineering application, the analysis in this section will be performed for circular geometry. However, the results can

be generalized for a pipe of noncircular geometry by substituting for the diameter D in any of the equations, the hydraulic diameter, D_h, defined as follows:

$$D_h = 4 \times \frac{\text{The cross-sectional area}}{\text{The wetted perimeter}}$$

The analysis in this section can also be applied to gases and vapors, provided the Mach number in the duct does not exceed 0.3. For greater values of the Mach number, the compressibility effect becomes significant and the reader is referred to Section 2.7 on compressible flow.

Fluid Friction

The method for calculating friction loss in pipes and ducts varies depending on whether the flow is laminar or turbulent. Most pipe flow is in the turbulent range.

The Reynolds number is the ratio of inertia forces to viscous forces and is a convenient parameter for predicting if a flow condition will be laminar or turbulent. It is defined as follows:

$$\text{Re} = \frac{\rho VD}{\mu} = \frac{VD}{\upsilon} \tag{2.4.6}$$

where V is the mean flow velocity, D is the diameter, ρ is the fluid density, μ is the dynamic viscosity, and ν is the kinematic viscosity.

Friction loss (H_f) depends on pipe diameter (D), length (L), roughness (e), fluid density (ρ) or specific weight (γ), viscosity (ν), and flow velocity (V). Dimensional analysis can be used to provide a functional relationship between the friction loss H_f, pipe dimensions, fluid properties, and flow parameters. The resulting equation is called the Darcy–Weisbach equation:

$$H_f = \frac{fLV^2}{2gd} = \frac{flQ^2}{1.23gD^5} \tag{2.4.7}$$

The friction factor f is a measure of pipe roughness. It has been evaluated experimentally for numerous pipes. The friction factor data were used to create the Moody diagram shown as in Figure 2.4.1. For $Re < 2000$, the flow in a pipe will be laminar and f, which is only a function of Re, is calculated as follows:

$$f = \frac{64}{\text{Re}} \tag{2.4.8}$$

For Reynolds numbers between 2000 and 4000, the flow is unstable as a result of the onset of turbulence (critical zone in Figure 2.4.1). In this range, friction loss calculations are difficult because it is impossible to determine a unique value of f. For $Re > 4000$, the flow becomes turbulent and f is a function of both Re and relative pipe roughness (e/D, e is the equivalent roughness height and D is the pipe diameter). At high Re, f eventually becomes independent of Re and depends only on e/d. This flow regime is referred to as fully turbulent flow and is represented in Figure 2.4.1 where family of friction factor curves become horizontal. The Re at which this occurs varies with e/D. Laminar flow in pipes is unusual. For example, for water flowing in a 0.3-m-diameter pipe, the velocity would have to be less than 0.02 m/s for laminar flow to exist. Therefore, most practical pipe flow problems are in the turbulent region.

Using the Moody diagram in Figure 2.4.1 to get f requires that Re and e/d be known. Calculating Re is direct if the water temperature, velocity, and pipe diameter are known. The problem is obtaining a good value for e. Typical values of e are listed in Figure 2.4.1. These values should be considered as guides only and not used if more exact values can be obtained from the pipe supplier.

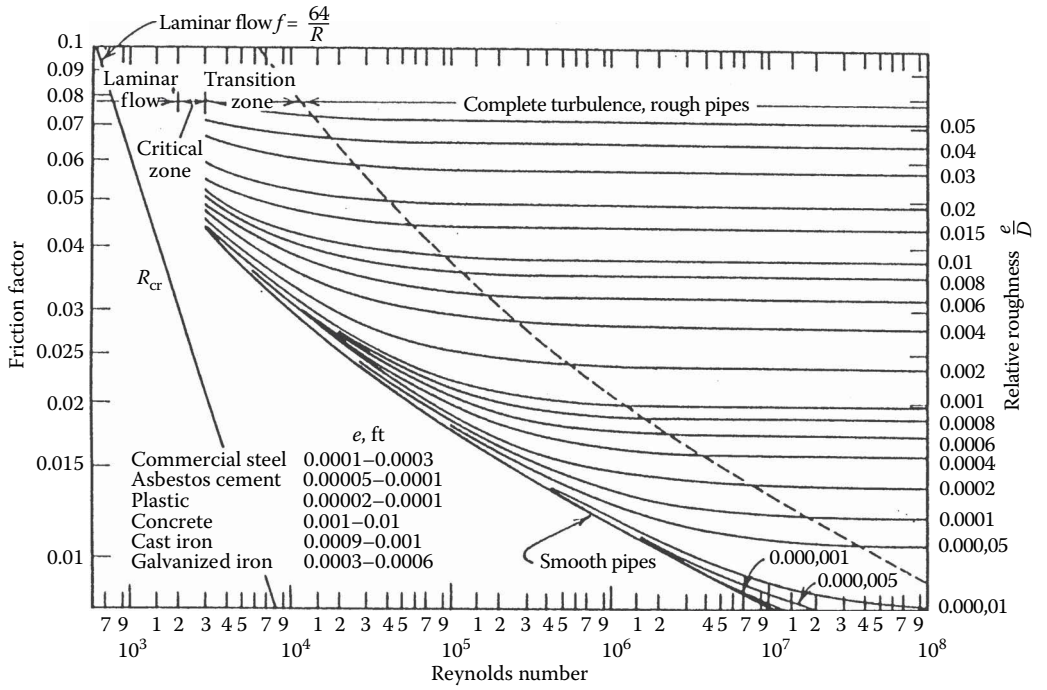


FIGURE 2.4.1 The Moody diagram.

Since roughness and inside pipe diameter may vary with time due to buildup of solid deposits or organic growths, f is also time dependent. Manufacturing tolerances also cause variations in the pipe diameter and surface roughness. Because of these factors, the friction factor for any pipe can only be approximated. A designer is required to use good engineering judgment in selecting a design value for f so that proper allowance is made for these uncertainties.

For noncircular pipes, the only change in the friction loss equation is the use of an equivalent diameter—based on the hydraulic radius (R_h), i.e., $D=4R_h$—in place of the circular pipe diameter D. R_h is the ratio of the flow area to the wetted perimeter.

Swamee and Jaine (1976), among others, have developed relationships that approximate the Moody diagram data. The Swamee–Jain equation is as follows:

$$f = 0.25\left[\mathrm{Ln}\left(\frac{e/D}{3.7}\right) + \frac{5.74}{\mathrm{Re}^{0.9}}\right]^{-2} \tag{2.4.9}$$

The practical problem of how to obtain a reliable value for e still remains. It cannot be directly measured but must be determined from friction loss pipe tests.

An exact solution using the Darcy–Weisbach equation for flow or pipe diameter requires a trial-and-error or iterative solution technique because of the dependency of f on Re [$\mathrm{Re}=f(V, D)$]. A typical approach to solving this problem is to estimate a reasonable fluid velocity to calculate Re and obtain f from the Moody diagram or Equation 2.4.9. Next, calculate a new velocity and Re value and repeat until the solution converges. The availability of solvers and programmable calculators has greatly simplified the solution to pipe flow and pipe sizing problems. The following is a visual basic program (Qswamee) that can be incorporated into a spreadsheet program to solve for the flow rate between two reservoirs via a single pipe when f is unknown.

```
Function Qswamee(dz,L,K,D,e,nu,g)
Dim f,Re,Q,A,Pi As Double
Pi=3.14159
A=Pi * D ^ 2 * 0.25
Qswamee=(dz * 2 * g * A ^ 2 / (0.018 * L / D+K)) ^ 0.5
Do
Q=Qswamee
Re=Q * D / (A * nu)
f=0.25 / (Log(e / (3.7 * D)+5.74 / (Re ^ 0.9))/Log(10)) ^ 2
Qswamee=(dz * 2 * g * A ^ 2 / (f * L / D+K)) ^ 0.5
Loop Until (Qswamee - Q) < 1e-06
End Function
```

The variables that follow the program name on the first line are values passed in through the function call at the spreadsheet level (dz is the change in elevation between the reservoirs, L is the pipe length, K is the sum of the minor loss coefficients, D is the pipe diameter, e is the pipe roughness, nu is the kinematic viscosity, and g is the gravitational constant). The second line defines additional variables used in the program. The initial flow rate (Qswamee) is approximated using an f value of 0.018. This value is stored as Q, Re, and f are calculated and an updated Qswamee is determined. This process continues until the difference between Q and Qswamee <1E-06 (user-defined accuracy limit). The Qswamee user-defined function returns the flow rate solution to the spreadsheet.

For long gravity flow pipelines, the general goal is to select the smallest pipe diameter that can pass the required flow without friction and minor losses exceeding the available head. For pumped systems, the pipe diameter selection must be based on an economic analysis that compares the pipe cost with the cost of building and operating the pumping facility.

Minor Losses

Flow through valves, orifices, elbows, transitions, etc., causes flow separation, which results in the generation and dissipation of turbulent eddies. For short systems containing many bends, valves, tees, etc., local or minor losses can exceed friction losses. The head loss h_l associated with the dissipation caused by a minor loss is proportional to the velocity head and can be accounted for as a minor or local loss using the following equation:

$$h_l = K_l \frac{Q^2}{(2gA_m^2)} \tag{2.4.10}$$

where K_l is the minor loss coefficient and A is the pipe area. The summation of the friction and minor losses in a pipe system can be expressed as follows:

$$H_f = h_f + h_l \tag{2.4.11}$$

$$H_f = \left[\Sigma\left(\frac{fL}{2gdA_p^2}\right) + \Sigma\left(\frac{K_l}{2gA_m^2}\right)\right]Q^2 = CQ^2 \tag{2.4.12}$$

where

$$C = \Sigma\left(\frac{fL}{2gdA_p^2}\right) + \Sigma\left(\frac{K_l}{2gA_m^2}\right) \tag{2.4.13}$$

A_p is the flow area of the pipe, and A_m is the flow area of the minor loss.

The challenging part of making minor loss calculations is obtaining reliable values of K_l. If the pipe is long, the friction losses may be large compared with the minor losses and approximate values of K_l will be sufficient. However, for short systems containing many pipefittings, the local losses can represent a significant portion of the total system losses, and they should be accurately determined. Numerous factors influence K_l. For example, for elbows, K_l is influenced by the shape of the conduit (rectangular vs. circular), by the radius of the bend, the bend angle, the Reynolds number, and the length of the outlet pipe. For dividing or combining tees or *Y*-branches, the percent division of flow and the change in pipe diameter must also be included when estimating K_l. Another important minor loss consideration is when local losses are installed close together. Depending on the type, orientation, and spacing of the components, the total loss coefficient may be greater or less than the simple sum of the individual K_l values.

Comparing the magnitude of $\sum\left(fL/2gdA_p^2\right)$ to $\sum(k_l/2gA_m^2)$ will determine how much care should be given to the selection of the K_l values. Typical values of K_l are listed in Table 2.4.1 (Tullis, 1989). When more comprehensive information on loss coefficients is needed, the reader is referred to Miller (1990).

Pressure Class Guidelines

Procedures for selecting the pressure class of pipe vary with the type of pipe material. Guidelines for different types of materials are available from professional societies and from the pipe manufacturers. These specifications should be obtained and studied for the pipe materials being considered.

The primary factors governing the selection of a pipe pressure class are (1) the maximum steady-state operating pressure, (2) surge/transient pressures, (3) external earth loads and live loads, (4) variation of pipe properties with temperature or long-time loading effects, and (5) damage that could result from handling, shipping, and installing or reduction in strength due to chemical attack or other aging factors. The influence of the first three items can be quantified, but the last two are subjective and are generally accounted for with a safety factor, which is the ratio of the burst pressure to the rated pressure. There is no single standard that specifies how large the safety factor should be or how many of the five preceding factors should be covered by the safety factor. Safety factor typically varies between 2 and 3.

When selecting a pipe pressure class, it is recommended that a transient computer analysis be included as part of that process. Such an analysis can determine the safe opening and closing times for valves, evaluate the influence of starting and stopping pumps on the maximum anticipated system pressure, and determine if surge protection devices are needed. Based on the results of computer simulations, select a pipe pressure class such that the rated pressure for the pipe is higher than the highest expected pressure caused by normal operation and by anticipated transient pressures. Be sure to select a safety factor large enough to account for the impact of earth and live loads and the effects of items 4 and 5 in the preceding list. Unusual or unexpected transient pressures are normally covered by the safety factor.

Selection of pressure class (or wall thickness) for larger pipes is sometimes more dependent on collapse pressure (due to an internal vacuum) or on handling loads than it is on operating pressure. A thin-wall, large-diameter pipe may be adequate for resisting relatively high internal pressures but may collapse under negative internal pressure.

External Loads

In some situations the external load is the controlling factor in determining if the pipe will collapse. The magnitude of the external load depends on the diameter and wall thickness of the pipe, the pipe material, the ovality (out of roundness) of the pipe cross section, the trench width, depth of cover, the specific weight of the soil, the degree of soil saturation, the type of backfill material, the method used to backfill, the degree of compaction, and live loads. The earth load increases with width

TABLE 2.4.1
Minor Loss Coefficients

Item	K_l Typical Value	Typical Range
Pipe inlets		
Inward projecting pipe	0.78	0.5–0.9
Sharp cornerflush	0.50	—
Slightly rounded	0.20	0.04–0.5
Bell mouth	0.04	0.03–0.1
Expansions[a]	$(1-A_1/A_2)^2$(based on V_1)	
Contractions[b]	$(1/C_c-1)^2$ (based on V_2)	

A_2/A_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
C_c	0.624	0.632	0.643	0.659	0.681	0.712	0.755	0.813	0.892

Bends[c]		
Short radius, $r/d=1$		
90	—	0.24
45	—	0.1
30	—	0.06
Long radius, $r/d=1.5$		
90	—	0.19
45	—	0.09
30	—	0.06
Mitered (1 m)		
90	1.1	—
60	0.50	0.40–0.59
45	0.3	0.35–0.44
30	0.15	0.11–0.19
Tees	c	—
Diffusers	[c]	—
Valves		
Check valves		
Swing check[d]	1.0	0.29–2.2
Tilt disk[d]	1.2	0.27–2.62
Lift check[d]	4.6	0.85–9.1
Double door[d]	1.32	1.0–1.8
Full-open gate valve	0.15	0.1–0.3
Full-open butterfly valve	0.2	0.2–0.6
Full-open globe valve	4.0	3–10

[a] See Streeter and Wylie (1975, p. 304).
[b] See Streeter and Wylie (1975, p. 305).
[c] See Miller (1990).
[d] See Kalsi Engineering and Tullis Engineering Consultants (1993).

and depth of the trench, whereas the live load reduces as depth of cover increases. The cumulative effect of all these sources of external loading requires considerable study and analysis. As the pipe diameter increases, the need for a more thorough analysis of earth and live loads becomes more important.

There are no simple guidelines for evaluating external pipe loads. Because of the complexity of this analysis, one option is to simply use a higher pressure class of pipe in areas where there will be large external loads. One should consider the consequence of a pipe failure caused by external loads

compared with the cost of using a thicker pipe. Those interested in the details of performing calculations of earth loading can review Watkins and Anderson (2000) and Spranger and Handy (1973).

Limiting Velocities

Upper and lower velocity both present concerns. If the velocity is too low, problems may develop due to settling of suspended solids and air collecting at various locations in the pipe. The safe lower velocity limit to avoid collecting air and sediment depends on the amount and type of sediment and on the pipe diameter and pipe profile. Velocities greater than about 1 m/s (3 ft/s) are typically required to move air pockets to air release valves. Higher velocities are usually needed to keep sediment in suspension.

Potential problems associated with high velocities are (1) erosion of the pipe wall or liner (especially if coarse suspended sediment is present), (2) increased chance of cavitation at control valves, (3) increased pumping costs, (4) difficulty with trapping air in the pipeline and its removal at air release valves, and (5) increased risk of hydraulic transients. Each of these should be considered before making the final pipe diameter selection. A typical upper velocity for many applications is about 6 m/s (~20 ft/s).

Valves

Valves serve a variety of functions. Some function as isolation valves that are either full open or closed. Control valves are used to regulate flow or pressure and must operate over a wide range of valve openings and flow conditions. Check valves prevent reverse flow, and air valves release air during initial filling, release air that is collected during operation, and admit air when the pipe is drained.

Control Valves

Hydraulic criteria considered when selecting the proper type and size of flow or pressure control valve should include the following:

1. The valve should not produce excessive pressure drop when full open.
2. The valve should have the ability to control flow or pressure drop over at least 50% of its movement.
3. The valve should not be subjected to excessive cavitation.
4. Valve opening/closing times should be selected so pressure transients caused by valve operation will not exceed the safe limits of the system.
5. The system should be analyzed to ensure that the operating torque or thrust does not exceed the capacity of the operator or the safe limit of the valve shaft and connections for any anticipated flow condition.

Controllability. For many flow control applications, it is desirable to select a control valve that has linear control characteristics. This means that if you close the valve, say 10%, the flow reduces about 10%. Unfortunately, this is sometimes not possible because the ability of a valve to control flow depends as much on the system as it does on the design of the valve. The same valve that operates linearly in one system may not in another. To demonstrate the relationship between valve performance and system requirements, consider a valve that will be used to control the flow between two reservoirs with an elevation difference of ΔZ. System A is a short pipe (0.3 m diameter, 100 m long, $\Delta Z = 10$ m) where pipe friction is small ($fL/2gdA_p^2 = 46.9$), and System B is a long pipe (0.3 m diameter, 10,000 m long, $\Delta Z = 200$ m) with high friction ($fL/2gdA_p^2 = 4690$). Initially, assume that the same valve will be used in both pipes and it will be

the same size as the pipe. The flow can be calculated using the energy equation (Equation 2.4.5) and the system loss equation (Equation 2.4.12):

$$Q = \sqrt{\frac{\Delta Z}{\left[\Sigma\left(\frac{fL}{2gdA_p^2}\right)+\Sigma\left(\frac{k_l}{2gA_m^2}\right)\right]}} \tag{2.4.14}$$

For the full-open valve, assume $K_l = 0.2$, and at 50% open $K_l = 9.0$. Correspondingly, $k_l/2gA_m^2 = 1.905$ and 85.7. For System A (short pipe), the flow with the valve full open will be 0.453 m^3/s and at 50% open the flow will be 0.275 m^3/s; a reduction of 39%. Repeating these calculations over the full range of valve openings would show that the flow for System A reduces almost linearly as the valve closes.

For System B (long pipe), the flow with the valve full open will be 0.206 m^3/s and at 50% open 0.205 m^3/s, a reduction of less than 1%. Because the valve loss at 50% closed, expressed by $k_l/2gA_m^2$, is still small compared to the friction term (4690), the valve will not start to control the flow until it has closed considerably more than 50%. A line-size valve is obviously not a good choice for a control valve in System B. One solution to this problem is to use a smaller valve. If the valve installed in System B was half the pipe diameter, closing it 50% would reduce the flow by about 22%. Even though the valve is half the diameter of the pipe, the reduction in flow with the reduced sized valve wide open would be insignificant.

The range of opening over which the valve controls the flow also has a significant effect on the safe closure time for control valves. Transient pressures are created each time there is an incremental reduction in the flow. If the valve does not begin to control until it has closed more than 50% closed, the effective valve closure time will be less than half the total closing time. This will increase the magnitude of the transients that will be generated unless the closing time is increased.

Torque/Thrust. If a valve will be subjected to unusually high velocities or large pressure drops, it is advisable to analyze the system to determine the maximum torque or thrust that the valve will be subjected to in order to insure that the valve shaft, connections, and operator are properly sized. If the maximum operating force exceeds operator capacity, the valve may not be able to open and close all flow conditions. If the shaft and connectors are underdesigned, the valve may fail and slam shut causing a severe transient. Valve manufacturers are responsible for determining the maximum allowable torque or thrust that their valves can tolerate but the end user needs to evaluate their system to determine if those maximum conditions might be exceeded in their system.

One situation where one may overlook the condition causing maximum torque is with parallel pumps that have a discharge control valve. The maximum system flow occurs with all pumps operating. However, the maximum flow velocity and maximum torque for each pump-discharge valve will occur when only one pump is operating. One specific example (Tullis, 1989) showed that the torque on a pump-discharge butterfly valve was three times higher when one pump was operating compared with three pumps operating in parallel.

Cavitation. Cavitation is frequently an important consideration in selection and operation of control valves. It is necessary to determine if cavitation will exist, evaluate its intensity, and estimate its effect on the system and environment. Cavitation can cause noise, vibration, erosion damage and can decrease valve performance. A thorough system analysis should consider the full range of operation. Some valves cavitate worst at small openings and others cavitate heavily near full open. It depends on both the system and the valve design. If cavitation is ignored in the design and selection of the valves, repairs and replacement of the valves may be costly.

One step in a cavitation analysis is selecting the acceptable level of cavitation. Experimental data are available for four limits: incipient (light, intermittent noise), critical (light, steady noise), incipient damage (pitting damage begins), and choking (very heavy damage and performance drops off). Cavitation data are available for each of these limits for a number of different valve types (Tullis, 1989, 1993).

Choosing the appropriate cavitation limit depends on several factors related to the operating requirements, expected life, location of the device, details of the design, and economics. For long-term operation of a control valve in a system where noise can be tolerated, the valve should never operate at or beyond incipient damage. In systems where noise is objectionable, critical cavitation would be a better operating limit. For short-term operation, choking cavitation may be appropriate.

Choking cavitation is sometimes misunderstood. Some have erroneously taught that onset of choking cavitation represents onset of cavitation. In reality, choking is the point where the intensity of cavitation is near its maximum and significant noise, vibrations, and erosion damage may occur both within the valve and in the downstream pipe. Choking cavitation is generally appropriate as a design limit only for valves that operate for short periods of time, such as a pressure relief valve. The intensity of cavitation and the corresponding noise vibration and erosion damage at the valve are at their maximum just before a valve chokes. If the valve operates beyond choking (sometimes referred to as supercavitation), the collapse of the large vapor cavity can occur at a significant distance downstream from the valve where serious vibration and material erosion problems can occur.

If the cavitation analysis indicates that the valve, orifice, or other device will be operating at a cavitation level greater than can be tolerated, various techniques can be used to limit the level of cavitation. One option is to select a different type of valve. Recent developments in valve design have produced a new generation of valves that are more resistant to cavitation. Most of them operate on the principle of dropping the pressure in stages. They usually have multiple paths with numerous sharp turns or flow restrictions in series. Two drawbacks to these valves are that they often have high-pressure drops when full open, and they can only be used with clean fluids. An alternate approach is to use a conventional valve in series with a number of orifice plates.

For some applications, another way to suppress cavitation is to use a free-discharge valve that is vented so cavitation cannot occur. There are valves specifically designed for this application. Some conventional valves can also be used for free discharge, if they can be adequately vented.

Injecting air to suppress cavitation downstream from conventional valves is a technique that has been used for many years with varying degrees of success. The most common mistake is injecting the air in the wrong location so the air does not get to the cavitation zone. If an adequate amount of air is injected into the proper region, the noise, vibrations, and erosion damage can be significantly reduced. If the system can tolerate some air being injected, aeration is often the cheapest and best remedy for cavitation.

For additional information on predicting when a valve will cavitate and how to suppress the cavitation, see Tullis (1989, 1993).

Transients. Transient pressures can occur during filling and flushing air from the line, while operating valves, and when starting or stopping pumps. The basic cause of high transient pressures is rapid changes in velocity. The larger the incremental velocity change and the faster that change takes place, the greater the transient pressures. If adequate design provisions and operational procedures are not established, the transient pressure can easily exceed the safe operating pressure of the pipe.

In order to perform a thorough transient analysis and ensure that the system will function as required, it is necessary to establish written operating procedures, preferably before or during the design process. A transient computer program can then be used to evaluate each operating condition and determine the highest transient pressure for each section of the system. As part of the analysis, the program can vary the operating times for the valves, simulate starting and stopping pumps and it can evaluate the influence of adding protective measures—like surge towers or air chambers.

Transients caused by slow velocity changes, such as the rise and fall of the water level in a surge tank, are called surges. Surge analysis or "rigid column theory" involves mathematical or numerical solution of simple ordinary differential equations. The compressibility of the fluid and the elasticity of the conduit are ignored, and the entire column of fluid is assumed to move as a rigid body.

When changes in velocity occur rapidly, both the compressibility of the liquid and the elasticity of the pipe must be included in the analysis. This procedure is often called "elastic" or "water

hammer" analysis and involves tracking acoustic pressure waves through the pipe. The solution requires solving partial differential equations.

An equation predicting the head rise ΔH caused by a sudden change of velocity $\Delta V = V_2 - V_1$ can be derived by applying the unsteady momentum equation to a control volume of a section of the pipe where the change of flow is occurring. Consider a partial valve closure that instantly reduces the velocity by an amount ΔV. Reduction of the velocity results in an increase in the pressure upstream from the valve creating a pressure wave of magnitude ΔH which travels up the pipe at the acoustic velocity a. The increased pressure compresses the liquid and expands the pipe. The transient head rise due to an incremental change in velocity (ΔV) is

$$\Delta H = -a\Delta V / g \tag{2.4.15}$$

The derivation of Equation 2.4.15 is based on an assumption of an instant velocity change. For a valve closing at the end of the pipe, instant closure actually refers to a finite time. It is the longest time that a valve can be closed and still cause a pressure rise equal to that of an instant closure. Normally, instant closure time refers to any valve closure time $\leq 2L/a$ s (which is the time required for the pressure wave to travel to and from the other end of a pipe of length L traveling at wave speed a). When closing a valve at a time of $2L/a$, the maximum head rise at the valve will be the same as if the valve were closed instantly. The $2L/a$ time is therefore often defined as the instant closure time.

For a valve at the end of a long pipeline, the effective instant closure time can be considerably greater than $2L/a$. This is because when the friction loss coefficient fL/d is much greater than the loss coefficient for the valve K_1, the valve can be closed a considerable amount before the flow actually begins to change. This dead time must be added to the $2L/a$ time to identify the effective instant closure time. To avoid the maximum potential transient pressure rise, the valve must be closed much slower than the effective instant closure time.

Computational techniques for estimating transient pressures caused by unsteady flow in pipelines are too complex to be done with simple hand calculations. The solution involves solving partial differential equations based on the equations of motion and continuity. These equations are often solved by the method of characteristics. This technique transforms the equations into total differential equations. After integration, the equations can be solved numerically by finite differences. This analysis provides equations that can be used to predict the flow and head at any interior pipe section at any time (Tullis, 1989; Wiley and Streeter, 1993).

To complete the analysis, equations describing the boundary conditions are required. Typical boundary conditions are the connection of a pipe to a reservoir, a valve, a change in pipe diameter or material, pipe junctions, etc. Friction loss is included in the development of the basic equations, and minor losses are handled as boundary conditions. The analysis properly models friction and the propagation and reflections of multiple pressure waves. It can also be used for surge calculations.

It is recommended that every pipe system should have at least a cursory transient analysis performed to identify the possibility of serious transients and decide whether or not a detailed analysis is necessary. If an analysis indicates that transients are a problem, the types of solutions available to the engineer include

1. Increasing the closing time of control valves
2. Using a smaller valve to provide better control
3. Designing special facilities for filling, flushing, and removing air from pipelines
4. Increasing the pressure class of the pipeline
5. Limiting the flow velocity
6. Using pressure relief valves, surge tanks, air chambers, etc.

Check Valves

Check valves are most commonly installed on the discharge side of pumps to prevent reverse flow when the pump is stopped. Selecting the wrong type or size of check valve can result in poor performance, severe transients, and frequent repairs (Kalsi and Tullis, 1993). Proper check valve selection requires understanding the characteristics of the various types of check valves and analyzing how they will function as a part of the system in which they will be installed. A check valve that operates satisfactorily in one system may be inadequate in another. Each valve type has unique characteristics that give it advantages or disadvantages compared with the others. The characteristics of check valves that describe their hydraulic performance and that should be considered in the selection process include

1. Opening characteristics, i.e., velocity vs. disk position data
2. Velocity required to fully open and firmly backseat the disk
3. Pressure drop at maximum flow
4. Stability of the disk at partial openings
5. Sensitivity of disk flutter to upstream disturbances
6. Speed of valve closure compared with the rate of flow reversal of the system

Disk stability varies with valve design, flow rate, disk position, and upstream disturbances and is an important factor in determining the useful life of a check valve. For most applications, it is preferable to size the check valve so that the disk is fully open and firmly backseated at normal flow rates. One of the worst design errors is to oversize a check valve that is located just downstream from a disturbance such as a pump, elbow, or control valve. If the disk does not fully open, it may be subjected to severe motion that will accelerate wear. To avoid this problem, it may be necessary to select a check valve that is smaller than the pipe size.

The transient pressure rise generated at check valve closure is another important consideration. The pressure rise is a function of how fast the valve disk closes compared with how fast the flow in the system reverses. The speed that the flow in a system reverses depends on the system. If a check valve that closes relatively slowly is installed in a system where rapid flow reversals occur, the disk can slam shut causing a pressure transient (Thorley, 1989).

The closing speed of a check valve varies with valve design and is influenced by the mass of the disk, the forces closing the disk, and the distance of travel from full open to closed. Fast closing valves have the following properties: the disk (including all moving parts) is lightweight, closure is assisted by springs, and the full stroke of the disk is short. Swing check valves are the slowest closing valves because they violate all three of these criteria; i.e., they have heavy disks, no springs, and long disk travel. The nozzle check valve is one of the fastest closing valves because the closing element is light, is spring loaded, and has a short stroke. The silent, duo, double door, and lift check valves with springs are similar to nozzle valves in their closing times, mainly because of the closing force of the spring. For descriptions of these check valves, see (Tullis, 1989; Kalsi and Tullis, 1993).

Systems where rapid flow reversals occur include parallel pumps, where one pump is stopped while the others are still operating, and systems that have air chambers or surge tanks close to the check valve. For these systems, there is a high-energy source downstream from the check valve to cause a quick flow reversal. As the disk nears its seat, it starts to restrict the reverse flow. This builds up the pressure, accelerates the disk, and slams it into the seat. Results of laboratory experiments, field tests, and computer simulations show that dramatic reductions in the transient pressures at disk closure can be achieved by replacing a slow-closing swing check valve with a fast-acting check valve. For example, in a system containing parallel pumps where the transient was generated by stopping one of the pumps, the peak transient pressure was reduced from 745 to 76 kPa when a swing check was replaced with a nozzle check valve.

Air Valves

There are three types of automatic air valves: (1) air/vacuum valves, (2) air release valves, and (3) combination valves. The air/vacuum valve contains a large orifice that releases large quantities of air at near atmospheric pressure while the pipe is being filled and admits large quantities of air into the pipe when the pipe is being drained. These valves typically contain a float, which rises and closes the outlet orifice as the valve body fills with water. Once the pipeline is pressurized, this type of valve cannot reopen to remove air that may subsequently accumulate unless the pressure inside the pipe becomes negative. If the pressure becomes negative during a transient or while draining, the float drops and admits air into the line. For thin-walled pipes that can collapse under internal vacuums, the air/vacuum valves should be sized for a full pipe break at the lowest pipe elevation. The air/vacuum valve must have the ability to supply an airflow equal to the maximum drainage rate of the water from the pipe break and at an internal vacuum pressure that is above the pipe collapse pressure of the pipe. Air valve sizing charts are provided by manufacturers.

Air release valves contain a small orifice that is designed to release small quantities of pressurized air. The mechanism opening and closing the orifice is typically controlled by a plunger activated by a float at the end of a lever arm. As air accumulates in the valve body, the float drops and opens the orifice. As the air is expelled, the float rises and closes the orifice.

The combination valve is actually two valves in one: a large valve that functions as an air/vacuum valve and a small one that functions as an air release valve. The installation can either consist of an air/vacuum valve and an air release valve plumbed in parallel, or the two can be integrated into a single valve body. Most air valve installations require combination valves.

One caution is that large-diameter manual air release valves should be avoided because improper operation of them can cause pressure transients. If the system is filled and pressurized with the manual air valves closed, the trapped air will be pressurized to full system pressure. When the manual air valve is opened, the air will be released at sonic velocity. The rapid depressurization of the air can cause rapid acceleration of the liquid upstream from the air valve and a high-pressure rise can occur when the water hits the air valve and instantly decelerates. If manual air valves are installed, they should be small so the air release rate is controlled to a safe rate. The suggested definition of a "small" manual air valve is that the air valve is about 1/20 of the pipe diameter.

Locating air valves in a piping system depends on the pipe profile, pipe length, and flow rates. Preferably, pipes should be laid to grade with valves placed at prominent high points and at regular intervals between high points (not more than a few thousand feet apart). At the high point of the pipeline, be sure to install air/vacuum valves of sufficient size to protect the pipe from collapse in the event of a major pipe break.

Pump Selection

Optimizing the life of a water supply system requires proper selection, operation, and maintenance of the pumps. During the selection process, the designer must be concerned about matching the pump performance to the system requirements and must anticipate problems that will be encountered when the pumps are started or stopped and when the pipe is filled and drained.

Selecting a pump for a particular service requires matching the system requirements to the capabilities of the pump. The process consists of developing a system equation by applying the energy equation to evaluate the pumping head required to overcome the elevation difference, friction, and minor losses. For a pump supplying water between two reservoirs, the pump head required to produce a given discharge can be expressed as follows:

$$H_p = \Delta Z + H_f \tag{2.4.16}$$

or

$$H_p = \Delta Z + CQ^2 \tag{2.4.17}$$

in which the constant C is defined by Equation 2.4.13.

Figure 2.4.2 shows a system curve for a pipe having an 82-m elevation lift and moderate friction losses. If the elevation of either reservoir is a variable, then there is not a single curve but a family of curves corresponding to differential reservoir elevations.

The three pump curves (A, B, and C) shown in Figure 2.4.2 represent different impeller diameters installed in the same pump casing with the pump operating at full speed. The intersections of the system curve with the pump curves identify the flow that each pump will supply if installed in that system. For the system curve shown, a single A pump would supply a flow of about 550 lpm at a head of 104 m and an efficiency of about 85%. A single C pump would supply a flow of about 300 lpm at a head of 88 m and an efficiency of about 60%. The net positive suction head (NPSH) and brake horsepower (bhp) can be obtained from their respective curves.

The selection process for a fixed speed pump is more complex when the system demand varies, either due to variations in the water surface elevation or to changing flow requirements. If the system must operate over a range of reservoir elevations, the pump should be selected so that the system curve, based on the mean (or the most frequently encountered) water level, intersects the pump curve near the midpoint of the best efficiency range. If the water level variation is not too great, the pump may be able to operate efficiently over the complete flow range.

The problem of pump selection also becomes more difficult when planning for future demands or if the pumps are required to supply a varying flow. If the flow range is large, multiple pumps or a variable-speed drive may be needed. Selection of multiple pumps and the decision about installing them in parallel or in series depend on the discharge requirement and the amount of friction in the system. Parallel installations are most effective for low-friction systems. Series pumps work best in high-friction systems.

For parallel pump operation, the combined two-pump curve is constructed by adding the flow of each pump at a common total dynamic head (H_p). Such a curve for two C pumps in parallel is shown in Figure 2.4.3 (labeled two pumps). The intersection of the two-pump curve with the system curve identifies the combined flow for the two pumps. The efficiency for each pump is determined by projecting horizontally to the left to intersect the single-pump curve. For this

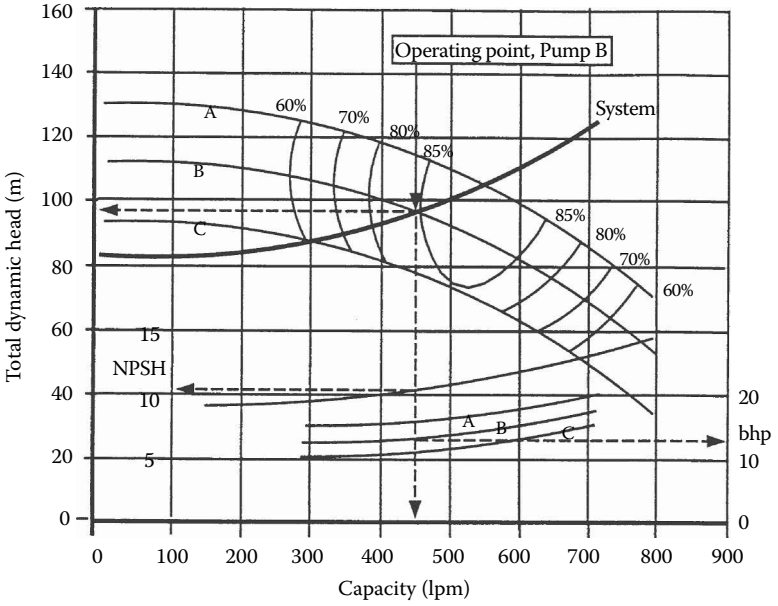


FIGURE 2.4.2 Pump selection for a single pump.

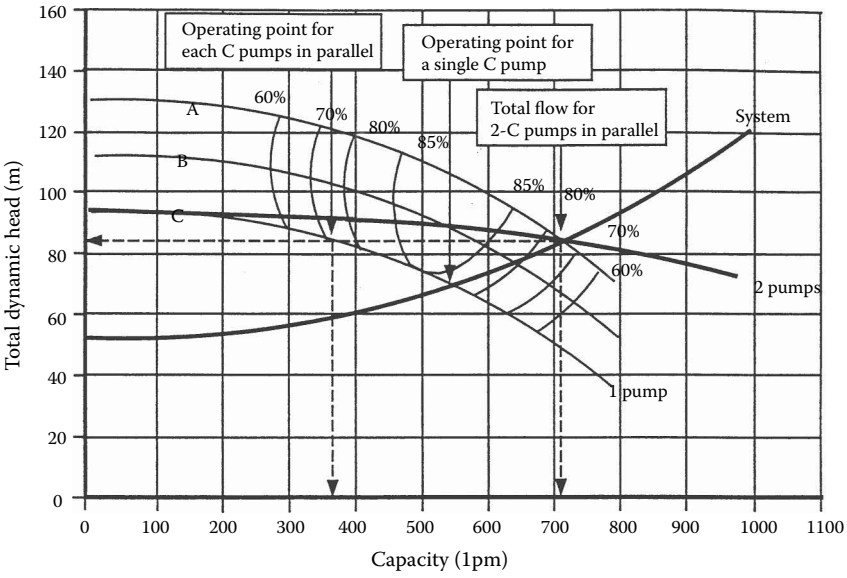


FIGURE 2.4.3 Selection of parallel pumps.

example with two C pumps operating in parallel, the flow will be 710 lpm at a head of 82 m and at an efficiency of each will be about 72%. For the two pumps to operate in the most efficient way, the selection should be made so the system curve intersects the single-pump curve to the right of its best efficiency point.

Starting a constant speed pump with the pipeline empty will result in filling at a very rapid rate because initially there is little friction to build backpressure. As a result, the pump will operate at a flow well above the design flow. This may cause the pump to cavitate, but the more serious problem is the possibility of high pressures generated by the rapid filling of the pipe. Provisions should be made to control the rate of filling to a safe rate. Start-up transients are often controlled by starting the pump against a partially open discharge valve located near the pump and using a bypass line around the pump. This allows the system to be filled slowly and safely. If the pipe remains full and no air is trapped, after the initial filling, subsequent start-up of the pumps generally does not create any serious problem. The preferred solution to avoiding rapid filling is to use variable speed pumps so the fill rate can be set at a safe rate.

For some systems, stopping the pump either intentionally or by a power failure can generate high pressures that can damage the pipe. If the design process does not consider these potential problems, the system may not function trouble free. Downtime and maintenance costs may be high. Not all systems will experience start-up and shutdown problems, but the design should at least consider the possibility. The problem is more severe for pipelines that have a large elevation change and multiple high points. The magnitude of the transient is related to the length and profile of the pipeline, the pump characteristics, the magnitude of the elevation change, and the type of check valve used. The downsurge caused by stopping the pump can cause column separation and high pressures due to flow reversals and closure of the check valves. Surge-protection equipment can be added to such systems to prevent damage and excessive maintenance.

Numerous mechanical devices and techniques have been used to suppress pump shutdown transients. These include increasing the rotational inertia of the pump, using surge tanks or air chambers near the pump, pressure relief valves, vacuum-breaking valves, and surge-anticipating valves. Selection of the proper transient control device will improve reliability, extend the economic life of the system, and reduce maintenance.

Filling and Flushing

If a piping system is short, the process of filling and flushing to remove trapped air poses little problem in either time or potential threat to the system. Simply fill slowly at a fill rate of about 0.25–0.5 mps (1–2 fps) and release the air either through automatic air valves or small manually opened bleed valves. Once the initial air is released, slowly increase the flow velocity to about 1–2 mps (3–6 fps) to flush any remaining air trapped to the air release valves.

When a pipeline is long, and especially if it has an undulating profile containing multiple high and low points, greater precaution is needed when filling and flushing. Few professional organizations publish guidelines for filling and flushing pipelines. The American Water Works Association (AWWA) is one organization that publishes limited guidelines for filling PVC pipe in their Standard: ANSI/AWWA C605-05. Their recommendations are conservative and incomplete. They simply suggest that pipelines be filled at a velocity of about 0.3 mps (1 fps) and then flushed at a velocity greater than 1 mps (3 fps). The 0.3 m/s fill limit is not the appropriate filling limit for all pipelines. It is simply a suggestion or a "rule of thumb," not a requirement and failing to strictly adhere to it does not automatically endanger a pipeline.

The appropriate rate to safely fill a pipeline depends on the length and profile of the pipeline and whether the pipeline is filled as a closed pipe (end valve closed) or an open pipe (end valve open). Consider a pipe that is lying on a flat surface connected to a pump with the end valve open. Or, alternately, consider the same pipe connected to a reservoir with a fill valve at the reservoir and no valve at the outlet end. Such a pipeline (short or long) can be safely filled by simply turning the pump on or opening the inlet valve. Because there is no closed or partially closed valve, and no high and low points in the line to restrict the movement of the air or water within the pipe, the pipe can be filled at a high velocity without endangering the pipe. The rapidly moving water simply forces the air to exit the end of the pipe as the water fills the pipe. The pipe gradually comes up to its normal operating pressure and no pressure surge is created. Using the suggested AWWA filling of 0.3 mps (1 fps) for such a pipe would be inappropriate and a waste of time because the pipe would never fill at such a low fill rate.

If it is necessary to fill a pipe as a closed system, the AWWA recommended 0.3 mps (1 fps) filling limit is a reasonable guideline because a closed pipe must be filled slowly to avoid the rapid pressurization of air that will be trapped in the pipe. This is not the preferred method of filling a long pipeline because it is inevitable that large air pockets will be trapped at high pressure. Once a closed pipe has been filled, pressurized and no more air can be vented, the trapped air must then be flushed to air release valves or out the end of the pipe by opening the end valve to establish flow. It is dangerous to flush high-pressure air pockets through the system and out of the end of the pipe. It is much safer to keep the pressure inside the pipe to a minimum while filling and flushing.

Another reason why it is best to avoid filling a pipe as a closed pipe is that it creates the possibility of subjecting the entire pipeline to the shutoff pressure of the upstream reservoir or pump, which may over-pressurize sections of the pipeline.

A safer way to fill a pipeline is to operate the system as an open pipe (end valve open). This keeps the pressure of the trapped air at a minimum while filling and flushing. Start filling at a rate of about 0.5–1 mps (1.5–3 fps). When liquid begins to flow out the end of the pipeline, increase the flow rate to about 2 mps (6 fps). Monitor flow and pressure in the pipeline and adjust the inflow rate and/or the opening of the valve at the end of the pipe to slowly bring the pipeline up to normal pressure.

System Monitoring and Control

Key elements such as flow rate, pressure, valve position, and pump operation should be continuously monitored and recorded by a System Control and Data Acquisition System (SCADA). Programmable Logic Controllers (PLCs) can be used to assist in operating the system and can provide warnings and take action when the flow or pressure is outside the safe or required operating limits.

Written operating instructions should be developed during the design phase and used after construction to safely operate the system. Such instructions should include normal operation and emergency shutdown procedures. Preparing the operating instructions during the design of the system and using them to perform computer simulations of the various operating scenarios will provide assurance that the system design will provide the required safety and flexibility so the system can meet all expectation. The operating instructions can be used to guide the steady-state and transient computer program simulations of system performance during the design phase.

REFERENCES

ASCE. 1992. *Pressure Pipeline Design for Water and Wastewater.* Prepared by the Committee on Pipeline Planning of the Pipeline Division of the American Society of Civil Engineers, New York.

ANSI/AWWA Standard C605-95. Sept. 2005. American Water Works Association, Denver, CO.

Kalsi Engineering and Tullis Engineering Consultants. 1993. *Application Guide for Check Valves in Nuclear Power Plants,* Revision 1, NP-5479. Prepared for Nuclear Maintenance Applications Center, Charlotte, NC.

Miller, D.S. 1990. *Internal Flow Systems—Design and Performance Prediction,* 2nd ed., Gulf Publishing Company, Houston, TX.

Morrison, E.B. 1969. Monograph for the design of thrust blocks. *Civil Eng.*, 39, 55–51.

Spanger, M.G. and Handy, R.L. 1973. *Soil Engineering,* 3rd ed., In text Eductational Publishers, New York, Chapters 25 and 26.

Streeter, V.L. and Wylie, E.B. 1975. *Fluid Mechanics,* 6th ed., McGraw-Hill, New York, 752 pp.

Swamee, P. and Jain, A. 1976. Explicit equations for pipe-flow problems. ASCE *J. Hydraul. Div.*, 102 (5), 657–664.

Thorley, A.R.D. 1989. Check valve behavior under transient flow conditions: A state-of-the-art review. *ASME J. Fluids Eng.,* 111(2), 173–183.

Tullis, J.P. 1989. *Hydraulics of Pipelines—Pumps, Valves, Cavitation, Transients,* John Wiley & Sons, New York.

Tullis, J.P. 1993. Cavitation guide for control valves, NUREG/CR-6031, U.S. Nuclear Regulatory Commission, Washington, DC.

Watkins, R.K. and Anderson, L.R. 2000. *Structural Mechanics of Buried Pipes*, CRC Press, New York.

Wood, D.J. 1966. An explicit friction factor relationship. *Civil Eng.,* 36, 60–61.

Wylie, E.B. and Streeter, V.L. 1993. *Fluid Transients in Systems,* Prentice-Hall, Englewood Cliffs, NJ.

Further Information

Bean, H.S., ed. 1971. *Fluid Meters, Their Theory and Application,* 6th ed., The American Society of Mechanical Engineers, New York.

Handbook of PVC: Design and Construction. 1979. Uni-Bell Plastic Pipe Association, Dallas, TX.

King, H.W. 1954. *Handbook of Hydraulics—For the Solution of Hydraulic Problems,* 4th ed., Revised by E.F. Brater. McGraw-Hill, New York.

Stephenson, D. 1981. *Pipeline Design for Water Engineers,* 2nd ed., Elsevier Scientific Publishing Company, Distributed in the U.S. by Gulf Publishing Company, Houston, TX.

Stutsman, R.D. 1993. Steel penstocks, ASCE Manuals and Reports on Engineering Practice No. 79, Energy Division, American Society of Civil Engineers, New York.

Tullis, J.P. 1996. Valves, in *The Engineering Handbook,* Dorf, R.C., Ed., CRC Press, Boca Raton, FL.

Watkins, R.K. and Spangler, M.G. 1958. Some characteristics of the modulus of passive resistance of soil: a study in similitude. *Highway Research Board Proceedings.* Vol. 37, 576–583.

2.5 OPEN CHANNEL FLOW

Frank M. White

Definition

The term *open channel flow* denotes the gravity-driven flow of a liquid with a free surface. Technically, we may study any flowing liquid and any gas interface. In practice, the vast majority of open channel flows concern water flowing beneath atmospheric air in artificial or natural channels. Excellent accounts of open channel flows are available in books by Brater (1976), Chow (1959), Henderson (1966) and Sellin (1970) for instance.

The geometry of an arbitrary channel is shown in Figure 2.5.1. The area A is for the water cross section only, and b is its top width. The wetted perimeter P covers only the bottom and sides, as shown, not the surface (whose air resistance is neglected). The water depth at any location is y, and the channel slope is θ, often denoted as $S_o = \sin\theta$. All of these parameters may vary with distance x along the channel. In unsteady flow (not discussed here), they may also vary with time.

Uniform Flow

A simple reference condition, called *uniform flow*, occurs in a long straight prismatic channel of constant slope S_o. There is no acceleration, and the water flows at constant depth with fluid weight exactly balancing the wetted wall shear force: $\rho g L A \sin\theta = \tau_w P L$, where L is the channel length. Thus, $\tau_w = \rho g R_h S_o$, where $R_h = A/P$ is called the *hydraulic radius* of the channel. If we relate wall shear stress to the Darcy friction factor f, $\tau_w = (f/8)\rho V^2$, we obtain the basic uniform flow open channel relation:

$$\text{Uniform flow}: V = \sqrt{\frac{8g}{f}}\sqrt{R_h S_o}, \text{ where} \sqrt{\frac{8g}{f}} = C = \text{Chézy coefficient} \tag{2.5.1}$$

Antoine Chézy first derived this formula in 1769. It is satisfactory to base f upon the pipe-flow Moody diagram (Figure 2.4.1) using the hydraulic diameter, $D_h = 4R_h$, as a length scale. That is, $f = fcn\,(VD_h/\nu, \varepsilon/D_h)$ from the Moody chart. In ordinary practice, however, engineers assume fully rough, high-Reynolds-number flow and use Robert Manning's century-old correlation:

$$C = \frac{\zeta}{n} R_h^{1/6}, \quad \text{or} \quad V_{\text{uniform}} \approx \frac{\zeta}{n} R_h^{2/3} S_o^{1/2} \quad \text{and} \quad Q = VA \tag{2.5.2}$$

where ζ is a conversion factor equal to 1.0 in SI units and 1.486 in English units. The quantity n is Manning's roughness parameter, with typical values, along with the associated roughness heights ε, listed in Table 2.5.1.

Critical Flow

Since the surface is always atmospheric, pressure head is not important in open channel flows. Total energy E relates only to velocity and elevation:

$$\text{Specific energy } E = y + \frac{V^2}{2g} = y + \frac{Q^2}{2gA^2}$$

At a given volume flow rate Q, the energy passes through a minimum at a condition called *critical flow*, where $dE/dy = 0$, or $dA/dy = b = gA^3/Q^2$:

$$A_{\text{crit}} = \left(\frac{bQ^2}{g}\right)^{1/3} \quad ; V_{\text{crit}} = \frac{Q}{A_{\text{crit}}} = \left(\frac{gA_{\text{crit}}}{b}\right)^{1/2} \tag{2.5.3}$$

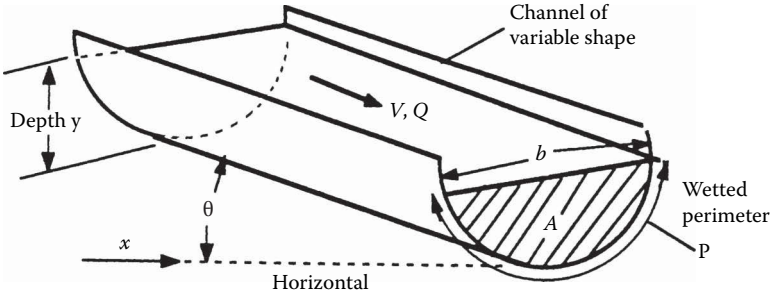


FIGURE 2.5.1 Definition sketch for an open channel.

TABLE 2.5.1
Average Roughness Parameters for Various Channel Surfaces

	n	Average Roughness Height ε	
Artificial Lined Channels		**ft**	**mm**
Glass	0.010±0.002	0.0011	0.3
Brass	0.011±0.002	0.0019	0.6
Steel; smooth	0.012±0.002	0.0032	1.0
Painted	0.014±0.003	0.0080	2.4
Riveted	0.015±0.002	0.012	3.7
Cast iron	0.013±0.003	0.0051	1.6
Cement; finished	0.012±0.002	0.0032	1.0
Unfinished	0.014±0.002	0.0080	2.4
Planed wood	0.012±0.002	0.0032	1.0
Clay tile	0.014±0.003	0.0080	2.4
Brickwork	0.015±0.002	0.012	3.7
Asphalt	0.016±0.003	0.018	5.4
Corrugated metal	0.022±0.005	0.12	37
Rubble masonry	0.025±0.005	0.26	80
Excavated Earth Channels			
Clean	0.022±0.004	0.12	37
Gravelly	0.025±0.005	0.26	80
Weedy	0.030±0.005	0.8	240
Stony; cobbles	0.035±0.010	1.5	500
Natural Channels			
Clean and straight	0.030±0.005	0.8	240
Sluggish, deep pools	0.040±0.010	3	900
Major rivers	0.035±0.010	1.5	500
Floodplains			
Pasture, farmland	0.035±0.010	1.5	500
Light brush	0.05±0.02	6	2000
Heavy brush	0.075±0.025	15	5000
Trees	0.15±0.05	—	—

where b is the top-surface width as shown in Figure 2.5.1. The velocity V_{crit} equals the speed of propagation of a surface wave along the channel. Thus, we may define the Froude number Fr of a channel flow, for any cross section, as $Fr = V/V_{crit}$. The three regimes of channel flow are

$$Fr < 1: \text{subcritical flow}; \quad Fr = 1: \text{critical flow}; \quad Fr > 1: \text{supercritical flow}$$

There are many similarities between Froude number in channel flow and Mach number in variable-area duct flow (see Section 2.6).

For a rectangular duct, $A = by$, we obtain the simplified formulas

$$V_{crit} = \sqrt{gy} \quad \text{and} \quad Fr = \frac{V}{\sqrt{gy}} \tag{2.5.4}$$

independent of the width of the channel.

Example 2.5.1

Water ($\rho = 998\,\text{kg/m}^3$, $\mu = 0.001$ kg/m·s) flows uniformly down a half-full brick 1-m-diameter circular channel sloping at 1°. Estimate (a) Q; and (b) the Froude number.

Solution 2.5.1 (a). First compute the geometric properties of a half-full circular channel:

$$A = \frac{\pi}{8}(1\,\text{m})^2 = 0.393\,\text{m}^2; \quad P = \frac{\pi}{2}(1\,\text{m}) = 1.57\,\text{m}; \quad R_h = \frac{A}{P} = \frac{0.393}{1.57} = 0.25$$

From Table 2.5.1, for brickwork, $n \approx 0.015$. Then, Manning's formula, Equation 2.5.2 predicts

$$V = \frac{\zeta}{n} R_h^{1/6} S R_o^{1/2} = \frac{1.0}{0.015}(0.25)^{1/6}(\sin 1°)^{1/2} \approx 3.49\,\frac{\text{m}}{\text{s}}; \quad Q = 3.49(0.393) \approx 1.37\,\frac{\text{m}^3}{\text{s}}$$

The uncertainty in this result is about ±10%. The flow rate is quite large (21,800 gal/min) because 1°, although seemingly small, is a substantial slope for a water channel.

One can also use the Moody chart. With $V \approx 3.49$ m/s, compute $\text{Re} = \rho V D_h/\mu \approx 3.49 \times 10^6$ and $\varepsilon/D_h \approx 0.0037$, then compute $f \approx 0.0278$ from the Moody chart. Equation 2.5.1 then predicts

$$V = \sqrt{\frac{8g}{f} R_h S_o} = \sqrt{\frac{8(9.81)}{0.0278}(0.25)(\sin 1°)} \approx 3.51\,\frac{\text{m}}{\text{s}}; \quad Q = VA \approx 1.38\,\frac{\text{m}^3}{\text{s}}$$

The two values of Q are remarkably close. With Q known from part (a), compute the critical conditions from Equation 2.5.3:

$$A_{crit} = \left(\frac{bQ^2}{g}\right)^{1/3} = \left[\frac{1.0(1.37)^2}{9.81}\right] = 0.576\ \text{m}^2, \quad V_{crit} = \frac{Q}{A_{crit}} = \frac{1.37}{0.576} = 2.38\,\frac{\text{m}}{\text{s}}$$

Hence

$$\text{Fr} = \frac{V}{V_{crit}} = \frac{3.49}{2.38} \approx 1.47 \ (\text{supercritical})$$

Again the uncertainty is approximately ±10%, primarily because of the need to estimate the brick roughness.

Hydraulic Jump

In gas dynamics (Section 2.6), a supersonic gas flow may pass through a thin normal shock and exit as a subsonic flow at higher pressure and temperature. By analogy, a supercritical open channel flow

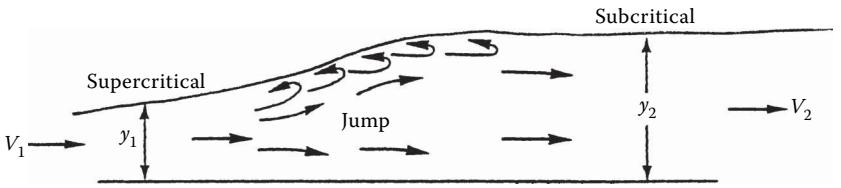


FIGURE 2.5.2 A two-dimensional hydraulic jump.

may pass through a *hydraulic jump* and exit as a subcritical flow at greater depth, as in Figure 2.5.2. Application of the continuity and momentum balances to a jump in a rectangular channel yields

$$V_2 = V_1 \frac{y_1}{y_2} \quad ; y_2 = \frac{y_1}{2}\left[-1+\sqrt{1+8\mathrm{Fr}_1^2}\right] \quad \text{where} \quad \mathrm{Fr}_1 = \frac{V}{\sqrt{g y_1}} > 1 \tag{2.5.5}$$

Both the normal shock and the hydraulic jump are dissipative processes: the entropy increases and the effective energy decreases. For a rectangular jump,

$$\Delta E = E_1 - E_2 = \frac{\left(y_2 - y_1\right)^3}{4 y_1 y_2} > 0 \tag{2.5.6}$$

For strong jumps, this loss in energy can be up to 85% of E_1. The second law of thermodynamics requires $\Delta E > 0$ and $y_2 > y_1$ or, equivalently, $\mathrm{Fr}_1 > 1$,

Note from Figure 2.5.2 that a hydraulic jump is not thin. Its total length is approximately four times the downstream depth. Jumps also occur in nonrectangular channels, and the theory is much more algebraically laborious.

Weirs

If an open channel flow encounters a significant obstruction, it will undergo rapidly varied changes which are difficult to model analytically but can be correlated with experiment. An example is the *weir* in Figure 2.5.3 (colloquially called a *dam*), which forces the flow to deflect over the top. If $L \ll Y$, the weir is termed *sharp-crested*; if $L = O(Y)$ it is *broad-crested*. Small details, such as the upper front corner radius or the crest roughness, may be significant. The crest is assumed level and of width b into the paper.

If there is a free overfall, as in Figure 2.5.3, the flow accelerates from subcritical upstream to critical over the crest to supercritical in the overfall. There is no flow when the excess upstream

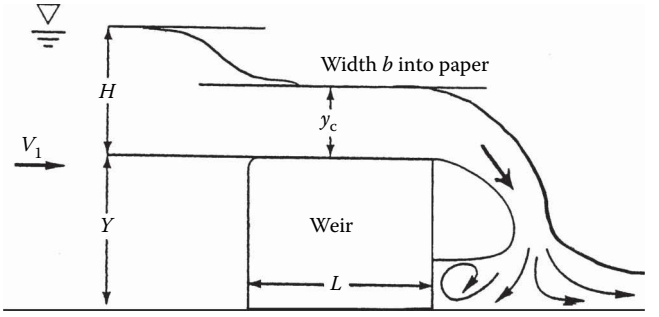


FIGURE 2.5.3 Geometry and notation for flow over a weir.

depth $H=0$. A simple Bernoulli-type analysis predicts that the flow rate Q over a wide weir is approximately proportional to $bg^{1/2}H^{3/2}$. An appropriate correlation is thus

$$Q_{\text{weir}} = C_{\text{d}} b g^{1/2} H^{3/2}, \quad \text{where } C_d = \text{dimensionless weir coefficient} \tag{2.5.7}$$

If the upstream flow is turbulent, the weir coefficient depends only upon geometry, and Reynolds number effects are negligible. If the weir has sidewalls and is narrow, replace width b by $(b-0.1H)$.

Two recommended empirical correlations for Equation 2.5.7 are as follows:

$$\begin{aligned} &\text{Sharp-crested:} \quad C_{\text{d}} \approx 0.564 + 0.0846\frac{H}{Y} \quad \text{for } \frac{L}{Y} < 0.07 \\ &\text{Broad-crested:} \quad C_{\text{d}} \approx 0.462 \quad \text{for } 0.08 < \frac{H}{L} < 0.33 \end{aligned} \tag{2.5.8}$$

These data are for wide weirs with a sharp upper corner in front. Many other weir geometries are discussed in the references for this section. Of particular interest is the sharp-edged vee-notch weir, which has no length scale b. If 2θ is the total included angle of the notch, the recommended correlation is

$$\text{Vee-notch, angle } 2\theta: \quad Q \approx 0.44 \tan\theta \; g^{1/2} H^{5/2} \quad \text{for} \quad 10^\circ < \theta \le 50^\circ \tag{2.5.9}$$

The vee-notch is more sensitive at low flow rates (large H for a small Q) and thus is popular in laboratory measurements of channel flow rates.

A weir in the field will tend to spring free and form a natural *nappe*, or air cavity, as in Figure 2.5.3. Narrow weirs, with sidewalls, may need to be aerated artificially to form a nappe and keep the flow from sliding down the face of the weir. The correlations above assume nappe formation. Further details can be found in Ackers et al. (1978), Boss and Coworkers (1984, 1985), Spitzer (1991), etc.

Gradually Varied Flow

Return to Figure 2.5.1 and suppose that (y, A, b, P, S_o) are all variable functions of horizontal position x. If these parameters are slowly changing, with no hydraulic jumps, the flow is termed *gradually varied* and satisfies a simple one-dimensional first-order differential equation if Q=constant:

$$\frac{dy}{dx} \approx \frac{S_o - S}{1 - \dfrac{V^2 b}{gA}}, \quad \text{where } V = \frac{Q}{A} \quad \text{and} \quad S = \frac{f}{D_h}\frac{V^2}{2g} = \frac{n^2 V^2}{\zeta^2 R_h^{4/3}} \tag{2.5.10}$$

The conversion factor $\zeta^2=1.0$ for SI units and 2.208 for English units. If flow rate, bottom slope $S_o(x)$, channel geometry, and surface roughness are known, we may solve for $y(x)$ for any given initial condition $y=y_0$ at $x=x_0$. The solution is computed by any common numerical method and is especially easy with a spreadsheet.

Recall from Equation 2.5.3 that the term $V^2b/(gA)=\text{Fr}^2$, so the sign of the denominator in Equation 2.5.10 depends upon whether the flow is sub- or supercritical. The mathematical behavior of Equation 2.5.10 differs also. If Fr is near unity, the change dy/dx will be very large, which probably violates the basic assumption of "gradual" variation.

For a given flow rate and local bottom slope, two reference depths are useful and may be computed in advance:

1. The *normal* depth y_n for which Equation 2.5.2 yields the flow rate.
2. The *critical* depth y_c for which Equation 2.5.3 yields the flow rate.

Comparison of these two, and their relation to the actual local depth y, specifies the type of solution curve being computed. The five bottom-slope regimes (mild M, critical C, steep S, horizontal H, and adverse A) create 12 different solution curves, as illustrated in Figure 2.5.4. All of these may be readily generated by a numerical solution of Equation 2.5.10. The following example illustrates a typical solution to a gradually varied flow problem.

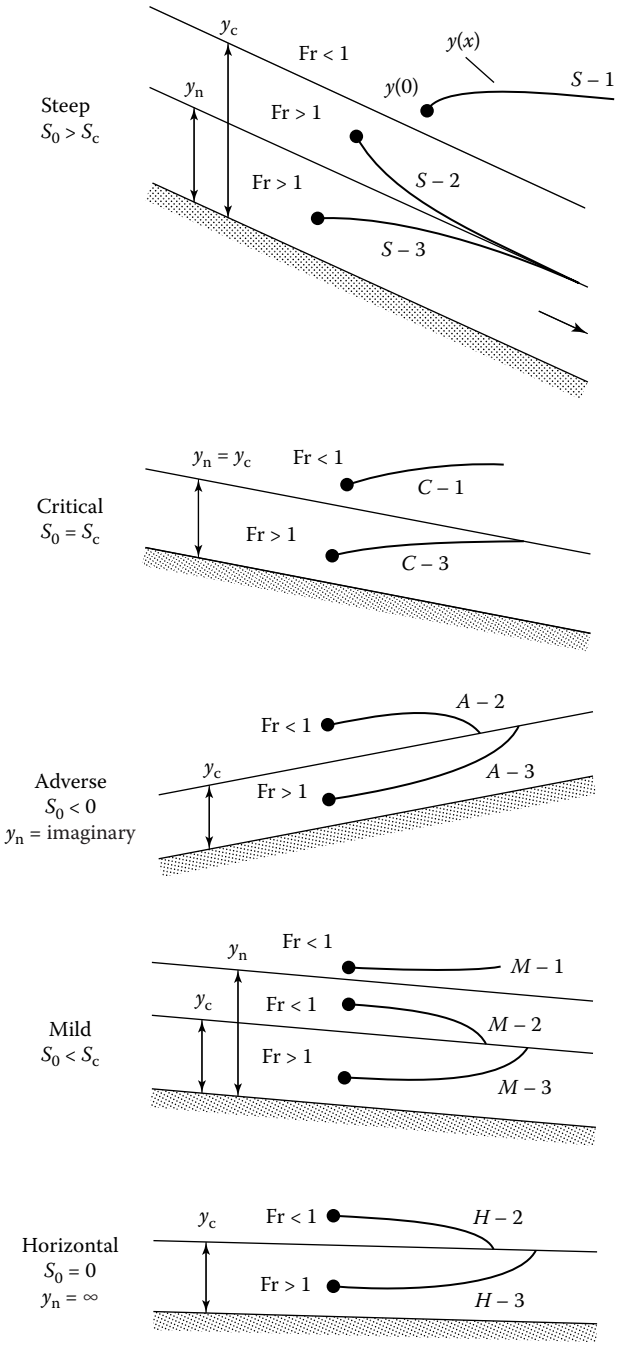


FIGURE 2.5.4 Classification of solution curves for gradually varied flow.

Example 2.5.2

Water, flowing at 2.5 m^3/s in a rectangular gravelly earth channel 2 m wide, encounters a broad-crested weir 1.5 m high. Using gradually varied theory, estimate the water depth profile back to 1 km upstream of the weir. The bottom slope is 0.1°.

Solution. We are given Q, $Y=1.5$ m, and $b=2$ m. We may calculate excess water level H at the weir (see Figure 2.5.3) from Equations 2.5.7 and 2.5.8:

$$Q = 2.5\frac{\text{m}^3}{\text{s}} = C_d b_{\text{eff}} g^{1/2} H^{3/2} = 0.462\left(2.0 - 0.1H\right)\left(9.81\right)^{1/2} H^{3/2}, \quad \text{solve for } H \approx 0.94 \text{ m}$$

Since the weir is not too wide, we have subtracted 0.1 H from b as recommended. The weir serves as a "control structure" which sets the water depth just upstream. This is our initial condition for gradually varied theory: $y(0)=Y+H=1.5+0.94\approx 2.44$ m at $x=0$. Before solving Equation 2.5.10, we find the normal and critical depths to get a feel for the problem:

$$\text{Normal depth:} \quad Q = 2.5\frac{\text{m}^3}{\text{s}} = \frac{1.0}{0.025}\left(2.0y_n\right)\left(\frac{2.0y_n}{2.0+2y_n}\right)^{2/3} \sqrt{\sin 0.1°}, \text{ solve } y_n \approx 1.14 \text{ m}$$

$$\text{Critical depth:} \quad A_c = 2.0y_c = \left(\frac{bQ^2}{g}\right)^{1/3} = \left[\frac{2.0\left(2.5\right)^2}{9.81}\right]^{1/3}, \quad \text{solve } y_c \approx 0.54 \text{ m}$$

We have taken $n\approx 0.025$ for gravelly earth, from Table 2.5.1. Since $y(0) > y_n > y_c$, we are on a mild slope $M-1$ "backwater" curve, as in Figure 2.5.4. For our data, Equation 2.5.10 becomes

$$\frac{dy}{dx} \approx \frac{S_o - n^2 Q^2 / \left(\zeta^2 A^2 R_h^{4/3}\right)}{1 - Q^2 b / \left(gA^3\right)}$$

where $Q=2.5$, $b=2$, $\zeta=1$, $A=2y$, $S_o=\sin 0.1°$, $R_h=2y/(2+2y)$, $g=9.81$, $y(0)=2.44$ at $x=0$.

Integrate numerically backward, that is, for $\Delta x < 0$, until $x=-1$ km$=-1000$ m. The complete solution curve is shown in Figure 2.5.5. The water depth decreases upstream and is approximately $y\approx 1.31$ m at $x=-1000$ m. If slope and channel width remain constant, the water depth asymptotically approaches the normal depth y_n far upstream.

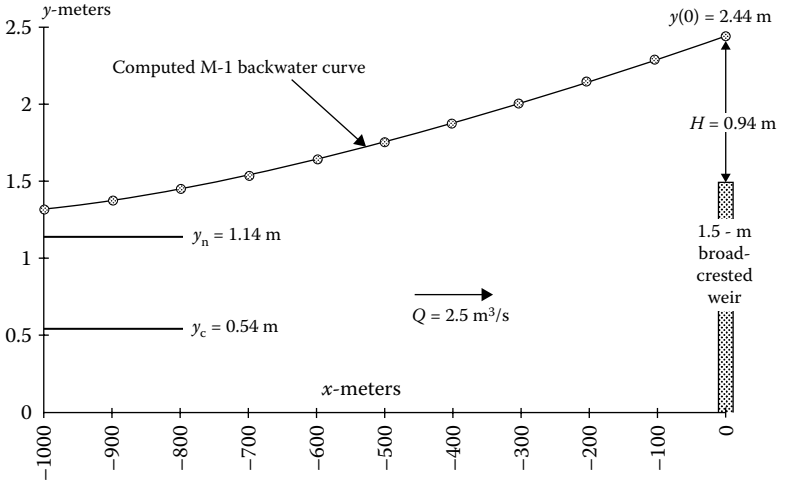


FIGURE 2.5.5 Backwater solution curve for Example 2.5.2.

REFERENCES

Ackers, P., et al. 1978. *Weirs and Flumes for Flow Measurement*, John Wiley & Sons, New York.
Bos, M.G. 1985. *Long-Throated Flumes and Broad-Crested Weirs*, Martinus Nijhoff (Kluwer), Dordrecht, The Netherlands.
Bos, M.G., Replogle, J.A., and Clemmens, A.J. 1984. *Flow-Measuring Flumes for Open Channel Systems*, John Wiley & Sons, New York.
Brater, E.F. 1976. *Handbook of Hydraulics*, 6th ed., McGraw-Hill, New York.
Chow, V.T. 1959. *Open Channel Hydraulics*, McGraw-Hill, New York.
French, R.H. 1985. *Open Channel Hydraulics*, McGraw-Hill, New York.
Henderson, F.M. 1966. *Open Channel Flow*, Macmillan, New York.
Sellin, R.H.J. 1970. *Flow in Channels*, Gordon & Breach, London.
Spitzer, D.W. (Ed.). 1991. *Flow Measurement: Practical Guides for Measurement and Control*, Instrument Society of America, Research Triangle Park, NC.

2.6 EXTERNAL INCOMPRESSIBLE FLOWS

John C. Leylegian

Introduction and Scope

Potential flow theory (Section 2.2) assumes a fluid to be incompressible and inviscid. There are no shear stresses; pressure is the only stress acting on a fluid particle. Because of the assumption of zero viscosity, potential flow theory predicts no drag force when an object moves through a fluid, which obviously is not correct, because all real fluids are viscous and lead to drag forces (note that although viscosity is required for any object to experience drag, the drag itself will often be due to the resulting pressure field). The objective of this section is to consider the behavior of a viscous, incompressible fluid flowing over an object.

A number of phenomena that occur in external flow at high Reynolds number (for most of which viscous effects are negligible) over an object are shown in Figure 2.6.1. As the flow reaches the object, the streamlines are diverted around the object; the freestream flow divides at the stagnation point. Fluid at the object surface moves at the velocity of the body as a result of the *no-slip condition*. Boundary layers form on the surfaces of the body; flow in the boundary layers is initially laminar, then if conditions are favorable, a *transition* to turbulent flow may occur (points "T").

As the boundary layers develop, they will cause a slight displacement of the streamlines of the external flow (their thickness is greatly exaggerated in Figure 2.6.1). *Separation* may occur if the fluid enters a region of increasing pressure on the rear of the body (points "S"); at separation, the flow adjacent to the body surface reverses direction. Fluid that was in the boundary layers thickens and forms a *wake* behind the object.

The Bernoulli equation is valid for steady, incompressible, inviscid flow. It may be used to predict pressure variations outside the boundary layer. The stagnation pressure is constant in the uniform inviscid flow far from an object, and the Bernoulli equation reduces to

$$p_\infty + \frac{1}{2}\rho U_\infty^2 = p + \frac{1}{2}\rho U^2, \tag{2.6.1}$$

where p is the freestream pressure, ρ is the fluid density, U is the freestream velocity, and the subscript ∞ denotes freestream conditions far upstream of the body. Therefore, the local pressure can be determined if the local freestream velocity is known.

Boundary Layers

The Boundary Layer Concept

The *boundary layer* is the thin region near the surface of a body in which viscous effects are important. By recognizing that viscous effects are concentrated near the surface of an object, Prandtl showed that only the Euler equations for inviscid flow need be solved in the region outside the boundary layer. Inside the boundary layer, the elliptic Navier–Stokes equations are simplified to the boundary layer equations, which are of parabolic form, and are easier to solve. The thin boundary layer has negligible pressure variation across it; pressure from the freestream is impressed upon the boundary layer.

The development of a boundary layer over a flat plate in a semi-infinite fluid is shown in Figure 2.6.2. Because the boundary layer is thin, there is negligible disturbance of the inviscid flow outside the boundary layer, and for a flat plate the *pressure gradient* along the surface is close to zero. Transition from laminar to turbulent boundary layer flow on a flat plate occurs when Reynolds number based on x (the distance downstream of the leading edge of the plate) exceeds $Re_x = 500{,}000$. Transition may occur earlier if the surface is rough, pressure increases in the flow direction, or separation occurs. Transition

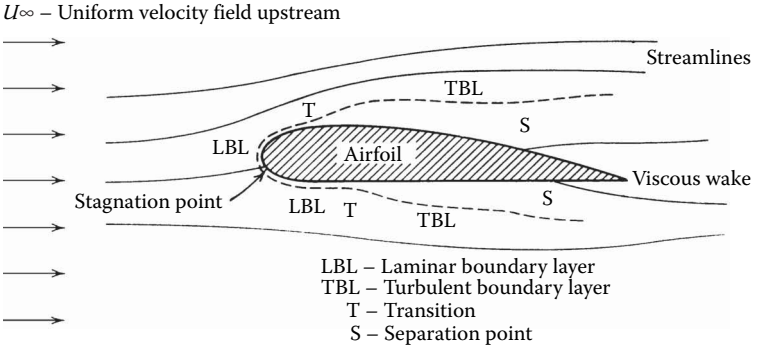


FIGURE 2.6.1 Viscous flow around an airfoil (boundary layer thickness exaggerated for clarity).

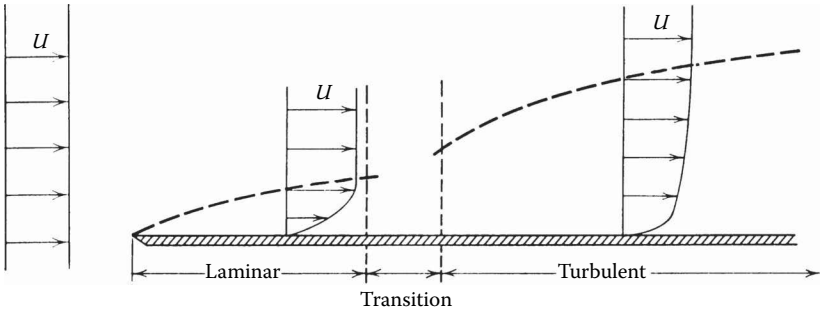


FIGURE 2.6.2 Boundary layer on a flat plate (vertical thickness exaggerated for clarity).

may also occur later if the surface is very smooth, if the pressure decreases in the flow direction, or if the flow has no disturbances. Following transition, the turbulent boundary layer thickens more rapidly than the laminar boundary layer as a result of increased shear stress at the body surface.

Boundary Layer Thickness Definitions

The boundary layer disturbance thickness, δ, is usually defined as the distance, y, from the surface to the point where the velocity within the boundary layer, u, is within 1% of the local freestream velocity, U. As shown in Figure 2.6.2, the boundary layer velocity profile merges smoothly and asymptotically into the freestream, making δ difficult to measure. For this reason and for their physical significance, we define two integral measures of boundary layer thickness. The displacement thickness, δ^*, is defined as

$$\frac{\delta^*}{\delta} = \int_0^\infty \left(1 - \frac{u}{U}\right) d\left(\frac{y}{\delta}\right) \tag{2.6.2}$$

The displacement thickness is the distance the solid boundary would have to be displaced into the freestream in a frictionless flow to produce the mass flow deficit caused by the presence of the boundary layer. The momentum thickness, θ, is defined as

$$\frac{\theta}{\delta} = \int_0^\infty \frac{u}{U}\left(1 - \frac{u}{U}\right) d\left(\frac{y}{\delta}\right) \tag{2.6.3}$$

The momentum thickness is the distance the solid boundary would have to be moved into the freestream to produce the same deficit in momentum flux as that caused by the boundary layer (momentum flux is momentum per unit time passing a cross section).

Because the displacement and momentum thicknesses are defined in terms of integrals for which the integrand vanishes in the freestream, they are easier to evaluate experimentally than disturbance thickness δ.

Exact Solution of the Laminar Flat-Plate Boundary Layer

Blasius obtained an exact solution for laminar boundary layer flow on a flat plate. He assumed a thin boundary layer to simplify the streamwise momentum equation. He also assumed *similar* velocity profiles in the boundary layer, so that when written as $u/U = f(y/\delta)$, velocity profiles do not vary with x. He used a similarity variable $\eta = y\,(U/\nu x)^{-1/2}$ to reduce the partial differential equations of motion and continuity to a single third-order ordinary differential equation.

Blasius used numerical methods to solve the ordinary differential equation; as a result, the velocity profile cannot be expressed as a function. The principal results of the Blasius solution may be expressed as

$$\frac{\delta}{x} = \frac{5}{\sqrt{Re_x}} \tag{2.6.4}$$

and

$$C_f = \frac{\tau_w}{\frac{1}{2}\rho U^2} = \frac{0.664}{\sqrt{Re_x}}, \tag{2.6.5}$$

where Re_x is the Reynolds number, C_f is the skin friction coefficient, and τ_w is the wall shear stress. These results show that the laminar boundary layer thickness on a flat plate varies as $x^{1/2}$ and wall shear stress varies as $1/x^{1/2}$.

Approximate Solutions

The Blasius solution cannot be expressed in closed form and is limited to laminar flow. Therefore, approximate methods that may give solutions for both laminar and turbulent flow in closed form are desirable. One such method is the *momentum integral equation* (MIE), which may be developed by integrating the boundary layer equation across the boundary layer or by applying the streamwise momentum equation to a differential control volume (Pritchard and Mitchell, 2015). The result is the ordinary differential equation:

$$\frac{d\theta}{dx} = \frac{\tau_w}{\rho U^2} - \left(\frac{\delta^*}{\theta} + 2\right)\frac{\theta}{U}\frac{dU}{dx}. \tag{2.6.6}$$

The first term on the right side of Equation 2.6.6 contains the influence of wall shear stress. Since τ_w is always positive in the direction of the flow, it always causes θ to increase with x. The second term on the right side contains the pressure gradient, which can have either sign. Therefore, the effect of the pressure gradient can be to either increase or decrease the rate of growth of boundary layer thickness.

Equation 2.6.6 is an ordinary differential equation that can be solved for θ as a function of x on a flat plate (zero pressure gradient), provided a reasonable shape is assumed for the boundary layer velocity profile and shear stress is expressed in terms of the other variables. Results for laminar and turbulent flat-plate boundary layer flows are discussed below.

TABLE 2.6.1
Exact and Approximate Solutions for Laminar Boundary Layer Flow over a Flat Plate at Zero Incidence

Velocity Distribution				
$\frac{u}{U}=f\left(\frac{y}{\delta}\right)=f(\eta)$	$\frac{\theta}{\delta}$	$\frac{\delta^*}{\delta}$	$a=\frac{\delta}{x}\sqrt{Re_x}$	$b=C_f\sqrt{Re_x}$
$f(\eta)=2\eta-\eta^2$	2/15	1/3	5.48	0.730
$f(\eta)=3/2\eta-1/2\eta^3$	39/280	3/8	4.64	0.647
$f(\eta)=\sin(\pi/2\eta)$	$(4-\pi)/2\pi$	$(\pi-2)/\pi$	4.80	0.654
Exact	0.133	0.344	5.00	0.664

Laminar Boundary Layers

A reasonable approximation to the laminar boundary layer velocity profile is to express u as a polynomial in y. The resulting solutions for δ and τ_w have the same dependence on x as the exact Blasius solution. Numerical results are presented in Table 2.6.1. Comparing the approximate and exact solutions shows remarkable agreement in view of the approximations used in the analysis. The trends are predicted correctly and the approximate values are within 10% of the exact values.

Turbulent Boundary Layers

The turbulent velocity profile may be expressed well using a power law, $u/U=(y/\delta)^{1/n}$, where n is an integer between 6 and 10 (frequently 7 is chosen). For turbulent flow, it is not possible to express shear stress directly in terms of a simple velocity profile; an empirical correlation is required. Using a pipe flow data correlation gives

$$\frac{\delta}{x}=\frac{0.382}{Re_x^{1/5}} \tag{2.6.7}$$

and

$$C_f=\frac{\tau_w}{\frac{1}{2}\rho U^2}=\frac{0.0594}{Re_x^{1/5}}. \tag{2.6.8}$$

These results characterize the turbulent boundary layer on a flat plate. They show that the turbulent boundary layer thickness varies as $x^{4/5}$ and wall shear stress varies as $1/x^{1/5}$.

Approximate results for laminar and turbulent boundary layers are compared in Table 2.6.2. At a Reynolds number of 1×10^6, wall shear stress for the turbulent boundary layer is nearly six times as large as for the laminar layer. For a turbulent boundary layer, thickness increases five times faster with distance along the surface than for a laminar layer. These approximate results give a physical feel for relative magnitudes in the two cases.

The MIE cannot be solved in closed form for flows with nonzero pressure gradients. However, the role of the pressure gradient can be understood qualitatively by studying the MIE.

Effect of Pressure Gradient

Boundary layer flow with favorable, zero, and adverse pressure gradients is depicted schematically in Figure 2.6.3. (Assume a thin boundary layer, so the flow on the lower surface behaves as external

TABLE 2.6.2
Thickness and Skin Friction Coefficient for Laminar and Turbulent Boundary Layers on a Flat Plate

	Boundary Layer Thickness/x		Skin Friction Coefficient		Turbulent/Laminar Ratio	
Reynolds Number	Laminar BL	Turbulent BL	Laminar BL	Turbulent BL	BL Thickness	Skin Friction
2×10^5	0.0112	0.0333	0.00148	0.00517	2.97	3.48
5×10^5	0.00707	0.0277	0.000939	0.00431	3.92	4.58
1×10^6	0.00500	0.0241	0.000664	0.00375	4.82	5.64
2×10^6	0.00354	0.0210	0.000470	0.00326	5.93	6.95
5×10^6	0.00224	0.0175	0.000297	0.00272	7.81	9.15
1×10^7	0.00158	0.0152	0.000210	0.00236	9.62	11.3
2×10^7	0.00112	0.0132	0.000148	0.00206	11.8	13.9
5×10^7	0.000707	0.0110	0.0000939	0.00171	15.6	18.3

Note: BL=boundary layer.

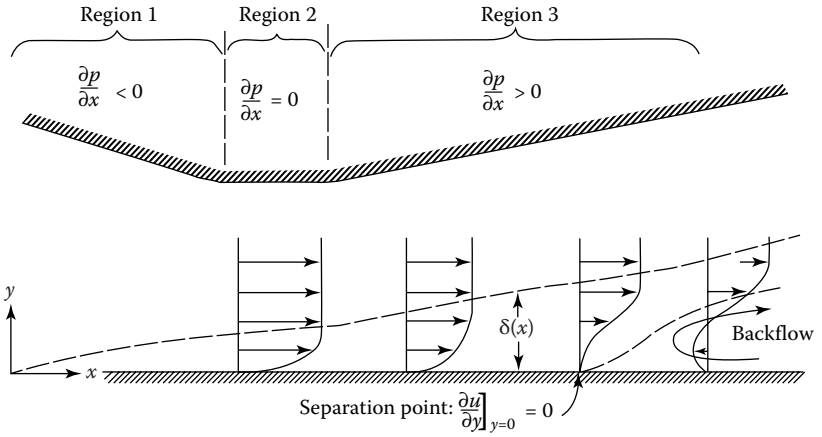


FIGURE 2.6.3 Boundary layer flow with pressure gradient (thickness exaggerated for clarity).

flow on a flat surface, with the pressure gradient generated by the upper surface geometry impressed on the boundary layer.) The pressure gradient is favorable when $\partial p/\partial x < 0$, zero when $\partial p/\partial x = 0$, and adverse when $\partial p/\partial x > 0$, as indicated for Regions 1, 2, and 3.

Viscous shear always causes a net retarding force on any fluid particle within the boundary layer. For zero pressure gradient, shear forces alone can never bring the particle to rest. (Recall that for laminar and turbulent boundary layers the shear stress varied as $1/x^{1/2}$ and $1/x^{1/5}$, respectively; shear stress never becomes zero for finite x.) Hence, since shear stress is given by $\tau_w = \mu\ \partial u/\partial y)_{y=0}$, the velocity gradient cannot be zero. Therefore, flow cannot separate in a zero pressure gradient; shear stresses alone can never cause flow separation.

In the favorable pressure gradient of Region 1, pressure forces tend to maintain the motion of the particle, so flow cannot separate. In the adverse pressure gradient of Region 3, pressure forces oppose the motion of a fluid particle. An adverse pressure gradient is a necessary condition for flow separation.

Velocity profiles for laminar and turbulent boundary layers are shown in Figure 2.6.2. It is easy to see that the turbulent velocity profile has much more momentum than the laminar profile. Therefore, the turbulent velocity profile can resist separation in an adverse pressure gradient better than the laminar profile.

The freestream velocity distribution must be known before the MIE can be applied. We obtain a first approximation by applying inviscid flow theory (e.g., potential flow) to calculate the flow field around the object. Much effort has been devoted to calculation of velocity distributions over objects of known shape (the "direct" problem) and to determine shapes to produce a desired pressure distribution (the "inverse" problem). Detailed discussion of such calculation schemes is beyond the scope of this section; the state-of-the art continues to progress.

Drag

Any object immersed in a viscous fluid flow experiences a net force from the shear stresses and pressure differences caused by the fluid motion. *Drag* is the force component parallel to the freestream velocity vector, while *lift* is the force component normal to the freestream velocity vector. *Streamlining* is the art of shaping a body to reduce fluid dynamic drag. Airfoils (and hydrofoils) are designed to produce lift in air (or water); they are streamlined to reduce drag and thus to attain high lift–drag ratios.

In general, lift and drag cannot be predicted analytically for flows with separation, but progress continues on computational fluid dynamics methods. For most engineering applications, aerodynamic bodies are designed based on a combination of past experience and theoretical/numerical work, and the resulting drag and lift forces are calculated from experimentally derived coefficients, discussed below.

The drag *coefficient* is defined as

$$C_D = \frac{F_D}{\frac{1}{2}\rho V^2 A}, \tag{2.6.9}$$

where $\frac{1}{2}\rho V^2$ is the dynamic pressure and A is a characteristic area upon which the coefficient is based. Common practice is to base the drag coefficient on the projected *frontal area* (Pritchard and Mitchell, 2015).

Similitude was treated in Section 2.3. In general, the drag coefficient may be expressed as a function of Reynolds number, Mach number, Froude number, relative roughness, depth of submergence divided by length, and so on. In this section, we consider neither high-speed flow nor free-surface effects, so we will consider only Reynolds number and roughness effects on drag coefficient.

Friction Drag

The total friction drag force acting on a plane surface aligned with the flow direction can be found by integrating the shear stress distribution along the surface. The drag coefficient for this case is defined as friction force divided by dynamic pressure and *wetted area* in contact with the fluid. Since shear stress is a function of Reynolds number, so is drag coefficient (see Figure 2.6.4). In Figure 2.6.4, transition occurs at $Re_x = 5 \times 10^5$; the dashed line represents the drag coefficient at larger Reynolds numbers. A number of empirical correlations may be used to model the variation in C_D shown in Figure 2.6.4 (Schlichting, 1979).

As seen in the figure, extending the laminar boundary layer line to higher Reynolds numbers shows that it is beneficial to delay transition to the highest possible Reynolds number. Some results are presented in Table 2.6.3; drag is reduced more than 50% by artificially extending laminar boundary layer flow to $Re_L = 10^6$.

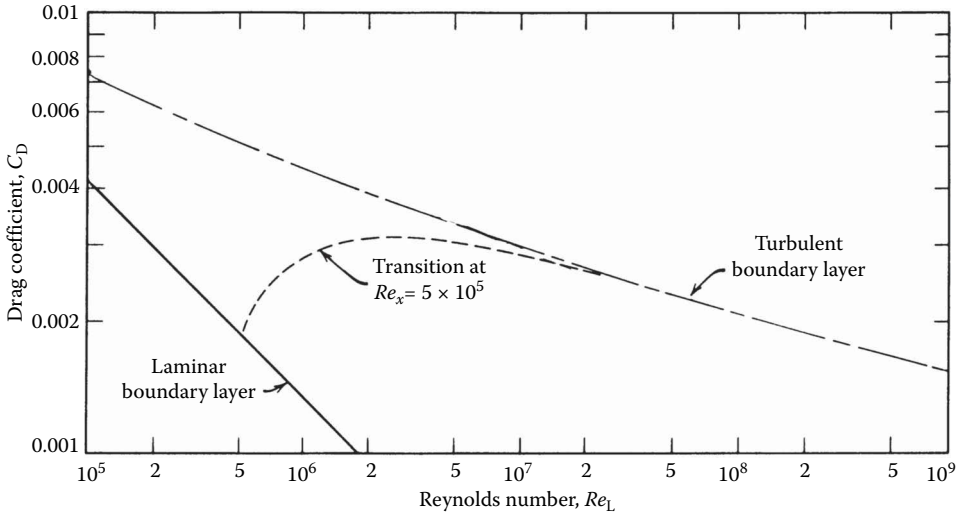


FIGURE 2.6.4 Drag coefficient vs. Reynolds number for a smooth flat plate parallel to the flow.

TABLE 2.6.3
Drag Coefficients for Laminar, Turbulent, and Transition Boundary Layers on a Flat Plate

	Drag Coefficient			Laminar/	% Drag
Reynolds Number	Laminar BL	Turbulent BL	Transition	Transition	Reduction
2×10^5	0.00297	0.00615	—	—	—
5×10^5	0.00188	0.00511	0.00189	—	—
1×10^6	0.00133	0.00447	0.00286	0.464	53.6
2×10^6	0.000939	0.00394	0.00314	0.300	70.0
5×10^6	0.000594	0.00336	0.00304	0.195	80.5
1×10^7	0.000420	0.00300	0.00284	0.148	85.2
2×10^7	0.000297	0.00269	0.00261	0.114	88.6
5×10^7	0.000188	0.00235	0.00232	0.081	9.19

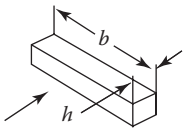
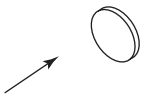
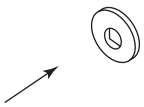
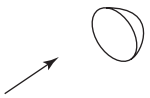
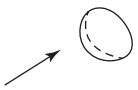
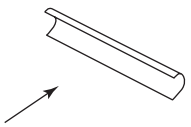
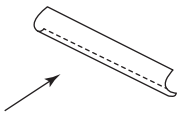
Note: BL = Boundary layer.

Pressure Drag

A thin flat surface normal to the flow has no area parallel to the flow direction. Therefore, there can be no friction force parallel to the flow; all drag is caused by pressure forces. Drag coefficients for objects with sharp edges tend to be independent of Reynolds number (for $Re > 1000$), because the separation points are fixed by the geometry of the object. Drag coefficients for selected objects are shown in Table 2.6.4.

Certain aspects of geometry have a significant effect on drag. For instance, rounding the edges that face the flow reduces drag markedly. This effect may be seen when comparing the drag coefficients for the hemisphere and C-section shapes facing into and away from the flow. It should also be noted that the drag coefficient for a two-dimensional object (long square cylinder) is about twice that for the corresponding three-dimensional object (square cylinder with $b/h = 1$).

TABLE 2.6.4
Drag Coefficient Data for Selected Objects ($Re > 1000$)[a]

Object	Diagram		$C_D(Re* \geq 10^3)$
Square prism		$b/h = \infty$	2.05
		$b/h = 1$	1.05
Disk			1.17
Ring			1.20[b]
Hemisphere (open end facing flow)			1.42
Hemisphere (open end facing downstream)			0.38
C-section (open side facing flow)			2.30
C-section (open side facing downstream)			1.20

[a] Data from Hoerner, S.F., *Fluid-Dynamic Drag*, 2nd ed., Published by the author, Midland Park, NJ, 1965.
[b] Based on ring area.

Friction and Pressure Drag: Bluff Bodies

Both friction and pressure forces contribute to the drag of *bluff bodies* (see Shapiro, 1960, for a good discussion of the mechanisms of drag). As an example, consider the drag coefficient for a smooth sphere shown in Figure 2.6.5. Transition from laminar to turbulent flow in the boundary layers on the forward portion of the sphere causes a dramatic dip in drag coefficient at the *critical Reynolds number* ($Re_D \approx 2 \times 10^5$). The turbulent boundary layer is better able to resist the adverse pressure gradient on the rear of the sphere, so separation is delayed and the wake is smaller, causing less pressure drag.

Surface roughness (or freestream disturbances) can reduce the critical Reynolds number. Many examples of this phenomenon may be seen. For example, dimples on a golf ball cause the boundary layer to become turbulent and, therefore, lower the drag coefficient in the range of speeds encountered in a drive. The stitches on a baseball or softball can have the same effect during a

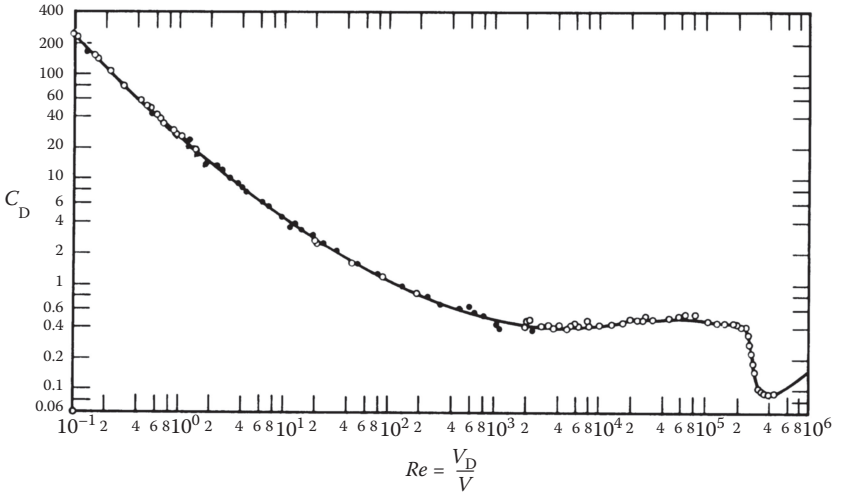


FIGURE 2.6.5 Drag coefficient vs. Reynolds number for a smooth sphere. (From Schlichting, H., *Boundary Layer Theory*, 7th ed., McGraw-Hill, New York, 1979. With permission.)

pitch. Roughness on race swimsuits and bicycle outfits also lower drag by causing transition to turbulent flow.

Streamlining

The most significant example of streamlining is adding a faired tail section to reduce the region of separated flow on the downstream portion of an object (at high Reynolds number, where pressure forces dominate drag). The pressure increase is taken over a longer distance, decreasing the magnitude of the adverse pressure gradient and delaying separation. However, adding a faired tail increases surface area, causing skin friction drag to increase. Thus, streamlining must be optimized for each shape.

Front contours are of principal importance in road vehicle design; the angle of the back glass also is important, but in most cases the entire rear end cannot be made long enough to control separation and reduce drag significantly.

Lift

Lift *coefficient* is defined as

$$C_L = \frac{F_L}{\frac{1}{2}\rho V^2 A}. \tag{2.6.10}$$

Note that lift coefficient is based on projected *planform area*.

Airfoils

Airfoils are shaped to produce lift primarily by accelerating flow over the upper surface to produce a low-pressure region. Because the flow must again decelerate, inevitably there must be a region of adverse pressure gradient near the rear of the upper surface (pressure distributions are shown clearly in Hazen, 1965).

Lift and drag coefficients for airfoil sections depend on Reynolds number and *angle of attack* between the chord line and the free stream flow direction. The *chord line* is the straight line joining the leading and trailing edges of the airfoil (Abbott and von Doenhoff, 1959).

As the angle of attack is increased, the minimum pressure point moves forward on the upper surface and the minimum pressure becomes lower. This change in the pressure profile increases the adverse pressure gradient. At some angle of attack, the adverse pressure gradient is strong enough to cause the boundary layer to separate completely from the upper surface, causing the airfoil to *stall*. The separated flow alters the pressure distribution, sharply reducing lift.

Increasing the angle of attack also causes the drag coefficient to increase. At some angle of attack below stall, the ratio of lift to drag, the *lift–drag* ratio, reaches a maximum, optimum, value.

Drag due to Lift

For wings (airfoils of finite span), lift and drag also are functions of aspect ratio. Lift is smaller and drag is larger for an airfoil of finite span compared with one of infinite span, because end effects cause the lift vector to rotate rearward. For a given geometric angle of attack, this rotation of the lift vector reduces the effective angle of attack, reducing lift. The additional component of lift acting in the flow direction increases drag; the increase in drag due to lift is called *induced drag.*

The effective aspect ratio includes the effect of planform shape. When written in terms of effective aspect ratio, the drag of a finite-span wing is

$$C_D = C_{D,\infty} + \frac{C_L^2}{\pi\,\mathrm{ar}} \tag{2.6.11}$$

where ar is the effective aspect ratio and the subscript ∞ refers to the infinite span section drag coefficient. The aspect ratio is defined as

$$\mathrm{ar} = \frac{b^2}{A}, \tag{2.6.12}$$

where b is the span of the wing and A is the planform area. In the case of a rectangular wing, the aspect ratio reduces to b/c, where c is the chord length. For further details, consult the references.

The lift coefficient must increase to support aircraft weight as speed is reduced. Therefore, induced drag can increase rapidly at low flight speeds. For this reason, minimum allowable flight speeds for commercial aircraft are closely controlled by the FAA.

Boundary Layer Control

The major part of the drag on an airfoil or wing is caused by skin friction. Therefore, it is important to maintain laminar flow in the boundary layers as far aft as possible; laminar flow sections are designed to do this. It also is important to prevent flow separation and to achieve high lift to reduce takeoff and landing speeds. These topics fall under the general heading of boundary layer control.

Profile Shaping

Boundary layer transition on a conventional airfoil section usually occurs almost immediately after the minimum pressure at about 25% chord aft of the leading edge. Transition can be delayed by shaping the profile to maintain a favorable pressure gradient over more of its length. The U.S. National Advisory Committee for Aeronautics (NACA) developed several series of profiles that delayed transition to 60% or 65% of chord, reducing drag coefficients (in the design range) 60% compared with conventional sections of the same thickness ratio (Abbott and von Doenhoff, 1959).

Flaps and Slats

Flaps are movable sections near the trailing edge of a wing. They extend and/or deflect to increase wing area and/or increase wing camber (curvature), to provide higher lift than the clean wing. Many aircraft also are fitted with leading edge slats which open to expose a slot from the pressure side

of the wing to the upper surface. The open slat increases the effective radius of the leading edge, improving maximum lift coefficient. The slot allows energized air from the pressure surface to flow into the low-pressure region atop the wing, energizing the boundary layers and delaying separation and stall.

Suction and Blowing

Suction removes low-energy fluid from the boundary layer, reducing the tendency for early separation. Blowing via high-speed jets directed along the surface reenergizes low-speed boundary layer fluid. The objective of both approaches is to delay separation, thus increasing the maximum lift coefficient the wing can achieve. Powered systems add weight and complexity; they also require bleed air from the engine compressor, reducing thrust or power output.

Moving Surfaces

Many schemes have been proposed to utilize moving surfaces for boundary layer control. Motion in the direction of flow reduces skin friction, and thus the tendency to separate; motion against the flow has the opposite effect. The aerodynamic behavior of sports balls—baseballs, golf balls, and tennis balls—depends significantly on aerodynamic side force (lift, down force, or side force) produced by spin. These effects are discussed at length in Pritchard and Mitchell (2015) and its references.

Computation vs. Experiment

Experiments cannot yet be replaced completely by analysis. Progress in modeling, numerical techniques, and computer power continues to be made, but the role of the experimentalist likely will remain important for the foreseeable future.

Computational Fluid Dynamics (CFD)

Computation of fluid flow requires accurate mathematical modeling of flow physics and accurate numerical procedures to solve the equations. The basic equations for laminar boundary layer flow are well known. For turbulent boundary layers, generally it is not possible to resolve the solution space into sufficiently small cells to allow direct numerical simulation. Instead, empirical models for the turbulent stresses must be used. Advances in computer memory storage capacity and speed (e.g., through use of massively parallel processing) continue to increase the resolution that can be achieved.

A second source of error in CFD work results from the numerical procedures required to solve the equations. Even if the equations are exact, approximations must be made to discretize and solve them using finite-difference or finite-volume methods. Whichever is chosen, the solver must guard against introducing numerical instability, round-off errors, and numerical diffusion (Zikanov, 2010). Because of the time and difficulty involved in developing CFD code, engineers resort to commercially available packages such as *ANSYS Fluent* or *STAR-CD.*

Role of the Wind Tunnel

Traditionally, wind tunnel experiments have been conducted to verify the design and performance of components and complete aircraft. Design verification of a modern aircraft may require expensive scale models, several thousand hours of wind tunnel time at many thousands of dollars an hour, and additional full-scale flight testing.

New wind tunnel facilities continue to be built and old ones refurbished. This trend of maintaining and improving experimental facilities indicates a need for continued experimental work in developing and optimizing aircraft configurations.

Many experiments are designed to produce baseline data to validate computer codes. Such systematic experimental data can help to identify the strengths and weaknesses of computational methods.

CFD tends to become only indicative of trends when massive zones of flow separation are present. Takeoff and landing configurations of conventional aircraft, with landing gear, high-lift devices, and flaps extended, tend to need final experimental confirmation and optimization. Many studies of vertical takeoff and vectored thrust aircraft require testing in wind tunnels.

REFERENCES

Abbott, I.H. and von Doenhoff, A.E. 1959. *Theory of Wing Sections, Including a Summary of Airfoil Data*, Dover, New York.

ANSYS Fluent. ANSYS, Inc., Southpointe, 2600 ANSYS Drive, Canonsburg, PA 15317 (http://www.ansys.com/Products/Fluids/ANSYS-Fluent).

Hazen, D.C. 1965. *Boundary Layer Control*, film developed by the National Committee for Fluid Mechanics Films (NCFMF) and available on videotape from Encyclopaedia Britannica Educational Corporation, Chicago, IL.

Hoerner, S.F. 1965. *Fluid-Dynamic Drag*, 2nd ed., Published by the author, Midland Park, NJ.

Pritchard, P.J. and Mitchell, J.W. 2015. *Fox and McDonald's Introduction to Fluid Mechanics*, 9th ed., John Wiley & Sons, New York.

Schlichting, H. 1979. *Boundary-Layer Theory*, 7th ed., McGraw-Hill, New York.

Shapiro, A.H. 1960. *The Fluid Dynamics of Drag*, film developed by the National Committee for Fluid Mechanics Film (NCFMF) and available on videotape from Encyclopaedia Britannica Educational Corporation, Chicago, IL.

STAR-CD. Adapco, 60 Broadhollow Road, Melville, NY 11747 (www.cd-adapco.com).

Zikanov, O. 2010. *Essential Computational Fluid Dynamics*, John Wiley & Sons, New York.

Further Information

A comprehensive source of basic information is the *Handbook of Fluid Dynamics*, edited by Victor L. Streeter (McGraw-Hill, New York, 1960).

Timely reviews of important topics are published in the *Annual Review of Fluid Mechanics* series (Annual Reviews, Inc., Palo Alto, CA.). Each volume contains a cumulative index.

ASME (American Society of Mechanical Engineers, New York, NY) publishes the *Journal of Fluids Engineering* quarterly. *JFE* contains fluid machinery and other engineering applications of fluid mechanics.

The monthly *AIAA Journal* and bimonthly *Journal of Aircraft* (American Institute for Aeronautics and Astronautics, New York) treat aerospace applications of fluid mechanics.

2.7 COMPRESSIBLE FLOW

John C. Leylegian

Introduction

This section deals with compressible flow. One- or two-dimensional steady, inviscid flows for which specific heats may be considered constant, i.e., calorically perfect gas behavior, are considered. Readers are referred to other sources of information for unsteady effects, viscous effects, and three-dimensional flows.

The term *compressible flow* is routinely used to define variable density flow. In many cases, these density variations are caused by the pressure changes from one point to another. *Compressibility* can be defined as the fractional change in volume of a fluid per unit change in pressure. It is a property of the fluid and is defined as follows:

$$\tau = \frac{1}{\rho}\frac{d\rho}{dp}$$

where τ is the compressibility of the fluid, ρ is the density, and p is the pressure being exerted on the fluid. A more precise definition of compressibility is obtained if we take into account the thermal and frictional losses. If during the compression, the temperature of the gas is held constant, it is called the isothermal compressibility and can be written as follows:

$$\tau_T = \frac{1}{\rho}\left(\frac{\partial \rho}{\partial p}\right)_T$$

However, if the compression process is reversible and adiabatic, it is called the isentropic compressibility and can be written as follows:

$$\tau_s = \frac{1}{\rho}\left(\frac{\partial \rho}{\partial p}\right)_s$$

Gases in general have high compressibility (τ_T for air is 10^{-5} m^2/N at 1 atm) as compared with liquids (τ_T for water is 5×10^{-10} m^2/N at 1 atm).

However, it can be seen that a large compressibility is not the only requirement for compressible flow; in fact, in many instances, air may be treated as incompressible. Based on the definition of compressibility, a large change in pressure arising with large changes in velocity would cause large changes in density. The change in velocity to cause a significant change in density may be found by comparing the flow velocity to the *speed of sound, a*, which is the speed of propagation of small pressure disturbances. The speed of sound for a gas is defined as follows:

$$a^2 = \left(\frac{\partial p}{\partial \rho}\right)_s \text{ or } a = \sqrt{\left(\frac{\partial p}{\partial \rho}\right)_s}$$

In an isentropic process for a perfect gas, the pressure and density are related through

$$\frac{p}{\rho^\gamma} = \text{constant}$$

Using this relation along with the ideal gas equation of state $p=\rho RT$, we can show that for a perfect gas

$$a=\sqrt{\gamma RT}=\sqrt{\frac{\gamma p}{\rho}}$$

where γ is the ratio of specific heats c_p/c_v, R is the gas constant, and T is the temperature. For air under normal conditions, γ is 1.4 and R is 287 $m^2/s^2{\cdot}K$, so that the speed of sound for air becomes $a=20.045\sqrt{T}$ m/s, where T is in Kelvin. The gas used has an effect on the speed of sound; e.g., helium has higher values of both γ and R so that the speed of sound for helium becomes $a=58.834\sqrt{T}$ m/s. This dependence of sound speed on the type of gas is employed in the design and use of shock tubes and shock tunnels.

Another important parameter in compressible flows is the *Mach number, M*, which is defined as the ratio of the gas velocity to the speed of sound or

$$M=\frac{V}{a}$$

where V is the velocity of gas. Depending upon the Mach number of the flow, we can define the following flow regimes:

$$M \ll 1 \text{ Incompressible flow}$$

$$M < 1 \text{ Subsonic flow}$$

$$M \approx 1 \text{ Transonic flow}$$

$$M > 1 \text{ Supersonic flow}$$

$$M \gg 1 \text{ Hypersonic flow}$$

Subsonic through hypersonic flows are compressible in nature. In these flows, the velocity is appreciable compared with the speed of sound, and the fractional changes in pressure, temperature, and density are all of significant magnitude. We will restrict ourselves in this section to subsonic through supersonic flows only.

Before we move on to study these flows, let us define one more term. Let us consider a gas with static pressure p and temperature T, traveling at some velocity V and corresponding Mach number M. If this gas is brought isentropically to rest, the resulting pressure and temperature of the gas are referred to as *stagnation pressure* p_0 and *stagnation temperature* T_0 (they are also called the total pressure and total temperature). The speed of sound at this stagnation state is called the *stagnation speed of sound* and is denoted as a_0.

One-Dimensional Flow

In one-dimensional flow, the flow properties vary only in one coordinate direction. In a steady, truly one-dimensional flow between two states 1 and 2, conservation of mass, momentum, and energy leads to the following equations:

$$\begin{aligned}
&\rho_1 u_1 = \rho_2 u_2 \\
&p_1 + \rho_1 u_1^2 = p_2 + \rho_2 u_2^2 \\
&h_1 + \frac{u_1^2}{2} + q = h_2 + \frac{u_2^2}{2}
\end{aligned} \tag{2.7.1}$$

where q is the heat added per unit mass of the gas. These equations neglect body forces, viscous stresses, and heat transfer due to thermal conduction and diffusion. The energy equation for a calorically perfect gas, for which $h=c_pT$, in the absence of external heat transfer, becomes

$$c_pT_1+\frac{u_1^2}{2}=c_pT_2+\frac{u_2^2}{2} \tag{2.7.2}$$

Using $c_p=\gamma R/(\gamma-1)$ and $a^2=\gamma RT$, the above equation can be written as follows:

$$\frac{a_1^2}{\gamma-1}+\frac{u_1^2}{2}=\frac{a_2^2}{\gamma-1}+\frac{u_2^2}{2} \tag{2.7.3}$$

Since Equation 2.7.3 is written for no heat addition, it holds for an adiabatic flow. If the energy equation is applied to the stagnation state, it can be written as follows:

$$\begin{aligned} c_pT+\frac{u^2}{2}&=c_pT_0\\ \frac{T_0}{T}&=1+\frac{\gamma-1}{2}M^2 \end{aligned} \tag{2.7.4}$$

It is worth mentioning that in arriving at Equation 2.7.4, only adiabatic flow is assumed, whereas the stagnation state is defined as that where the gas is brought to rest isentropically. Therefore, the definition of stagnation temperature is less restrictive than the general definition of the stagnation state. According to the general definition of isentropic flow, it is both adiabatic and reversible. For an isentropic flow

$$\frac{p_0}{p}=\left(\frac{\rho_0}{\rho}\right)^{\gamma}=\left(\frac{T_0}{T}\right)^{\gamma/(\gamma-1)} \tag{2.7.5}$$

From Equations 2.7.4 and 2.7.5, we can write

$$\frac{p_0}{p}=\left(1+\frac{\gamma-1}{2}M^2\right)^{\gamma/(\gamma-1)} \tag{2.7.6}$$

$$\frac{\rho_0}{\rho}=\left(1+\frac{\gamma-1}{2}M^2\right)^{1/(\gamma-1)} \tag{2.7.7}$$

Values of stagnation conditions are tabulated in Anderson (2003) as a function of M for $\gamma=1.4$.

Normal Shock Wave

A shock wave is a very thin region (of the order of a few molecular mean free paths) across which the static pressure, temperature, and density increase whereas the velocity decreases. If the shock wave is perpendicular to the flow, it is called a *normal shock wave*. The flow is supersonic ahead of the normal shock wave and subsonic behind it. Figure 2.7.1 shows the flow conditions across a normal shock wave which is treated as a discontinuity. Since there is no heat added or removed, the

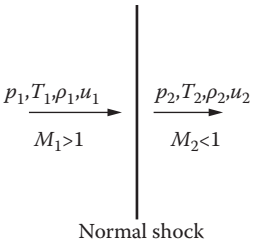


FIGURE 2.7.1 Flow conditions across a normal shock.

flow across the shock wave is adiabatic. By using Equation2.7.2, the normal shock equations can be written as follows:

$$\begin{aligned} \rho_1 u_1 &= \rho_2 u_2 \\ p_1 + \rho_1 u_1^2 &= p_2 + \rho_2 u_2^2 \\ h_1 + \frac{u_1^2}{2} &= h_2 + \frac{u_2^2}{2} \end{aligned} \tag{2.7.8}$$

Equation 2.7.8 is applicable to a general fluid; however, for a calorically perfect gas, we can use the relations $p = \rho RT$ and $h = c_p T$ to derive a number of equations relating flow conditions downstream of the normal shock to those upstream. These equations (also known as Rankine–Hugoniot relations) are

$$\begin{aligned} \frac{p_2}{p_1} &= 1 + \frac{2\gamma}{\gamma+1}\left(M_1^2 - 1\right) \\ \frac{\rho_2}{\rho_1} &= \frac{u_1}{u_2} = \frac{(\gamma+1)M_1^2}{2 + (\gamma-1)M_1^2} \\ \frac{T_2}{T_1} &= \frac{h_2}{h_1} = \left[1 + \frac{2\gamma}{\gamma+1}\left(M_1^2 - 1\right)\right]\left[\frac{2 + (\gamma-1)M_1^2}{(\gamma+1)M_1^2}\right] \\ M_2^2 &= \frac{1 + \frac{\gamma-1}{2}M_1^2}{\gamma M_1^2 - \frac{\gamma-1}{2}} \end{aligned} \tag{2.7.9}$$

Again, the values of p_2/p_1, ρ_2/ρ_1, T_2/T_1, etc. are tabulated in Anderson (2003) as a function of M_1 for $\gamma = 1.4$. Let us examine some limiting cases. As $M_1 \to 1$, Equation 2.7.9 yields $M_2 \to 1$, $p_2/p_1 \to 1$, $\rho_2/\rho_1 \to 1$, and $T_2/T_1 \to 1$. This is the case of an extremely weak normal shock across which no finite changes occur, i.e., a sound wave. On the other hand, as $M_1 \to \infty$, Equation 2.7.9 yields

$$M_2 \to \sqrt{\frac{\gamma-1}{2\gamma}} = 0.378; \quad \frac{\rho_2}{\rho_1} \to \frac{\gamma+1}{\gamma-1} = 6; \quad \frac{p_2}{p_1} \to \infty; \quad \frac{T_2}{T_1} \to \infty$$

However, the calorically perfect gas assumption no longer remains valid as $M_1 \to \infty$, since large changes in temperature result in significant changes to specific heat, as well as dissociation and ionization of gases. These phenomena are beyond the scope of this section.

Let us now examine why the flow ahead of a normal shock wave must be supersonic even though Equation 2.7.8 holds for $M_1 < 1$ as well as $M_1 > 1$. From the second law of thermodynamics, the entropy change across the normal shock can be written as follows:

$$s_2 - s_1 = c_p \ln \frac{T_2}{T_1} - R \ln \frac{p_2}{p_1}$$

By using Equation 2.7.9, it becomes

$$s_2 - s_1 = c_p \ln \left\{ \left[1 + \frac{2\gamma}{\gamma+1}\left(M_1^2 - 1\right) \right] \left[\frac{2 + (\gamma-1)M_1^2}{(\gamma+1)M_1^2} \right] \right\} - R \ln \left[1 + \frac{2\gamma}{\gamma+1}\left(M_1^2 - 1\right) \right] \tag{2.7.10}$$

Equation 2.7.10 shows that the entropy change across the normal shock is a function of M_1 only. Using Equation 2.7.10, it can be shown that

$$\begin{aligned} s_2 - s_1 &= 0 \quad \text{for } M_1 = 1 \\ &< 0 \quad \text{for } M_1 < 1 \\ &> 0 \quad \text{for } M_1 > 1 \end{aligned}$$

Since it is necessary that $s_2 - s_1 \geq 0$ from the second law, $M_1 \geq 1$. This, in turn, requires that $p_2/p_1 \geq 1$, $\rho_2/\rho_1 \geq 1$, $T_2/T_1 \geq 1$, and $M_2 \leq 1$.

We now examine how the stagnation conditions change across a normal shock wave. For a calorically perfect gas, the energy equation in Equation 2.7.9 gives

$$c_p T_{01} = c_p T_{02} \text{ or } T_{01} = T_{02}$$

In other words, the total temperature remains constant across a stationary normal shock wave.

Let us now apply the entropy change relation across the shock using the stagnation conditions.

$$s_2 - s_1 = c_p \ln \frac{T_{02}}{T_{01}} - R \ln \frac{p_{02}}{p_{01}}$$

Note that entropy at stagnation conditions is the same as at the static conditions since to arrive at stagnation conditions, the gas is brought to rest isentropically. Since $T_{02} = T_{01}$

$$\begin{aligned} s_2 - s_1 &= -R \ln \frac{p_{02}}{p_{01}} \\ \frac{p_{02}}{p_{01}} &= e^{-(s_2 - s_1)/R} \end{aligned} \tag{2.7.11}$$

Since $s_2 > s_1$ across the normal shockwave, Equation 2.7.11 gives $p_{02} < p_{01}$ or, in other words, the total pressure decreases across a shock wave.

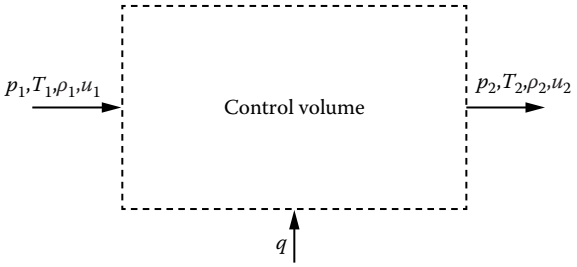


FIGURE 2.7.2 One-dimensional control volume with heat addition.

One-Dimensional Flow with Heat Addition

Consider one-dimensional flow through a control volume with inlet location 1 and exit at 2, as shown in Figure 2.7.2. A specified amount of heat per unit mass, q, is added to the control volume. The governing equations relating conditions 1 and 2 can be written as follows:

$$\begin{aligned} \rho_1 u_1 &= \rho_2 u_2 \\ p_1 + \rho_1 u_1^2 &= p_2 + \rho_2 u_2^2 \\ h_1 + \frac{u_1^2}{2} + q &= h_2 + \frac{u_2^2}{2} \end{aligned} \tag{2.7.12}$$

The following relations can be derived from Equation 2.7.12 for a calorically perfect gas:

$$q = c_p \left(T_{02} - T_{01} \right) \tag{2.7.13}$$

$$\frac{p_2}{p_1} = \frac{1 + \gamma M_1^2}{1 + \gamma M_2^2} \tag{2.7.14}$$

$$\frac{p_{02}}{p_{01}} = \frac{1 + \gamma M_1^2}{1 + \gamma M_2^2} \left(\frac{1 + \dfrac{\gamma - 1}{2} M_2^2}{1 + \dfrac{\gamma - 1}{2} M_1^2} \right)^{\gamma/(\gamma - 1)} \tag{2.7.15}$$

$$\frac{T_2}{T_1} = \left(\frac{1 + \gamma M_1^2}{1 + \gamma M_2^2} \right)^2 \left(\frac{M_2}{M_1} \right)^2 \tag{2.7.16}$$

$$\frac{T_{02}}{T_{01}} = \left(\frac{1 + \gamma M_1^2}{1 + \gamma M_2^2} \right)^2 \left(\frac{M_2}{M_1} \right)^2 \left(\frac{1 + \dfrac{\gamma - 1}{2} M_2^2}{1 + \dfrac{\gamma - 1}{2} M_1^2} \right) \tag{2.7.17}$$

$$\frac{\rho_2}{\rho_1} = \left(\frac{1 + \gamma M_2^2}{1 + \gamma M_1^2} \right)^2 \left(\frac{M_1}{M_2} \right)^2 \tag{2.7.18}$$

Equation 2.7.13 indicates that the effect of heat addition is to directly change the stagnation temperature T_0 of the flow. Table 2.7.1 shows some physical trends due to heat addition for subsonic and supersonic flows. With heat extraction, the trends in Table 2.7.1 are reversed.

Figure 2.7.3 shows a plot between enthalpy and entropy (known as the Mollier diagram) based on the earlier equations, for one-dimensional flow with heat addition. This curve is called the Rayleigh curve and is drawn for a set of given initial conditions. Each point on this curve corresponds to a different amount of heat added or removed. It is seen from this curve that heat addition always drives the Mach numbers toward 1. For a certain amount of heat addition, the flow will become sonic. For this condition, the flow is said to be *choked*. Any further increase in heat addition is not possible without adjustment in initial conditions. For example, if more heat is added in region 1, which is initially supersonic, than allowed for attaining Mach 1 in region 2, then a normal shock will form inside the control volume which will suddenly change the conditions in region 1 to subsonic. Similarly, in the case of an initially subsonic flow corresponding to region 1′, any heat addition beyond that is needed to attain Mach 1 in region 2, the conditions in region 1′ will adjust to a lower subsonic Mach number through a series of pressure waves.

Equations 2.7.13 through 2.7.18 may be used to determine the exit conditions for a flow with heat transfer. However, these equations are not easy to work with. Therefore, we establish a reference state for flow with heat transfer at Mach 1. We denote this sonic reference state using an asterisk (*):

TABLE 2.7.1
Effect of Heat Addition on Subsonic and Supersonic Flow

	$M_1<1$	$M_1>1$
M_2	Increases	Decreases
p_2	Decreases	Increases
T_2	Increases for $M_1 < \gamma^{-1/2}$ and decreases for $M_1 > \gamma^{-1/2}$	Increases
u_2	Increases	Decreases
T_{02}	Increases	Increases
p_{02}	Decreases	Decreases

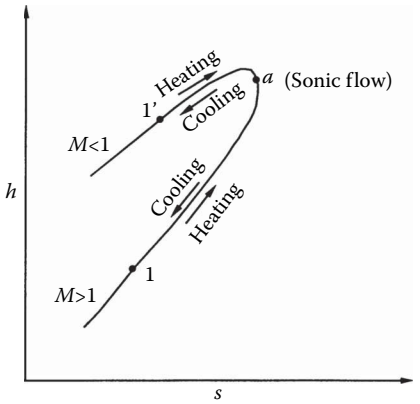


FIGURE 2.7.3 The Rayleigh curve.

$$\frac{p}{p^*} = \frac{1+\gamma}{1+\gamma M^2} \tag{2.7.19}$$

$$\frac{p_0}{p_0^*} = \frac{1+\gamma}{1+\gamma M^2}\left(\frac{2+(\gamma-1)M^2}{\gamma+1}\right)^{\gamma/(\gamma-1)} \tag{2.7.20}$$

$$\frac{T}{T^*} = M^2\left(\frac{1+\gamma}{1+\gamma M^2}\right)^2 \tag{2.7.21}$$

$$\frac{T_0}{T_0^*} = \frac{(\gamma+1)M^2}{\left(1+\gamma M^2\right)^2}\left[2+(\gamma-1)M^2\right] \tag{2.7.22}$$

$$\frac{\rho}{\rho^*} = \frac{1}{M^2}\left(\frac{1+\gamma M^2}{1+\gamma}\right) \tag{2.7.23}$$

Equations 2.7.13 and 2.7.19 through 2.7.23 are used to solve problems involving flow with heat transfer. The functions shown earlier are tabulated as a function of M for $\gamma = 1.4$ in Anderson (2003).

Similar to the preceding heat addition or extraction relationships, we can also develop relationships for one-dimensional steady, adiabatic flow but with frictional effects due to viscosity. In this case, the momentum equation is modified due to frictional shear stress. For details, readers are referred to Anderson (2003).

One-Dimensional Flow with Area Change

In one-dimensional flow with area change, in addition to flow conditions, the area of duct also changes with x. The governing equations for this flow can be written in a differential form as follows using an infinitesimal control volume shown in Figure 2.7.4:

$$\mathrm{d}(\rho u A) = 0 \tag{2.7.24}$$

$$\mathrm{d}p + \rho u\,\mathrm{d}u = 0 \tag{2.7.25}$$

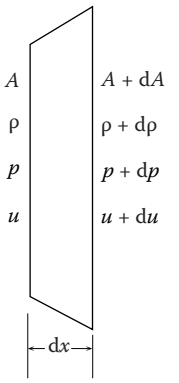


FIGURE 2.7.4 Control volume for one-dimensional flow with area change.

$$dh + u\,du = 0 \tag{2.7.26}$$

Equation 2.7.24 can be written as follows:

$$\frac{d\rho}{\rho} + \frac{du}{u} + \frac{dA}{A} = 0 \tag{2.7.27}$$

which can be further written as follows for an isentropic flow:

$$\frac{1}{M^2 - 1}\frac{du}{u} = \frac{dA}{A} \tag{2.7.28}$$

Some very useful physical insights can be obtained from this area–velocity relation.

- For subsonic flow ($0 \leq M < 1$), an increase in area results in decrease in velocity, and vice versa.
- For supersonic flow ($M > 1$), an increase in area results in increase in velocity, and vice versa.
- For sonic flow ($M = 1$), $dA/A = 0$, which corresponds to a minimum or maximum in the area distribution, but based on the area–velocity relationship for subsonic and supersonic flow, it can be seen that a minimum in area is the only physical solution.

Figure 2.7.5 shows the preceding results in a schematic form.

Based on this discussion, for a gas to accelerate isentropically from subsonic to supersonic conditions, or to decelerate from supersonic to subsonic conditions, it must flow through a convergent–divergent duct, also known as the de Laval nozzle. The minimum area of the nozzle at which the flow becomes sonic is called the throat. This physical observation forms the basis of designing supersonic wind tunnels shown schematically in Figure 2.7.6. In general, in a supersonic wind tunnel, a stagnant gas is first expanded to the desired supersonic Mach number. The supersonic flow enters the test section where it passes over a model being tested. The flow then is slowed down by compressing it through a second convergent–divergent duct, also known as a diffuser, before it is exhausted to the atmosphere.

Now, using the equations for one-dimensional flow with area change and the isentropic flow conditions, we can derive a relation for the area ratio that is needed to accelerate or decelerate the gas to sonic conditions. Denoting the sonic conditions by an asterisk, we can write $u^* = a^*$, but do not confuse this sonic reference state with the sonic reference state mentioned earlier for flow with heat transfer. The area is denoted as A^*, and would be the minimum area for the throat of the nozzle. From Equation 2.7.24, we have

$$\rho u A = \rho^* u^* A^*$$

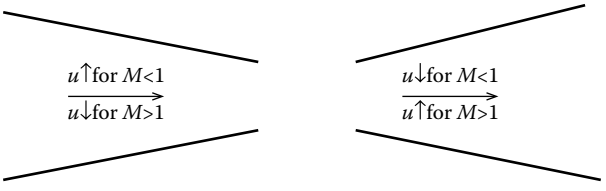


FIGURE 2.7.5 Compressible flow in converging and diverging ducts.

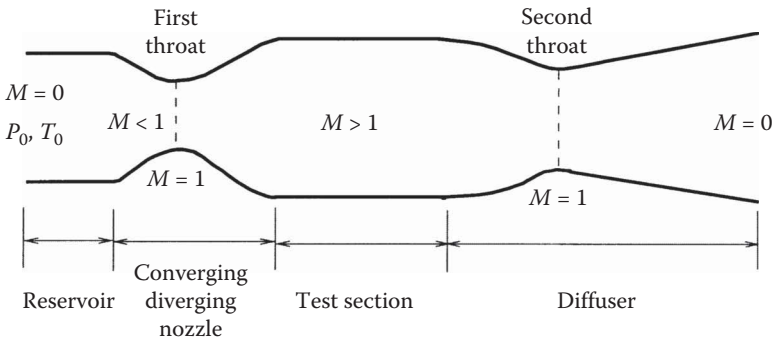


FIGURE 2.7.6 Schematic of a typical supersonic wind tunnel.

$$\frac{A}{A^*} = \frac{\rho^* u^*}{\rho u} = \frac{\rho^*}{\rho_0}\frac{\rho_0}{\rho}\frac{u^*}{u} \tag{2.7.29}$$

Under isentropic conditions

$$\frac{\rho_0}{\rho} = \left(1 + \frac{\gamma - 1}{2} M^2\right)^{1/(\gamma-1)} \tag{2.7.30}$$

$$\frac{\rho_0}{\rho^*} = \left(1 + \frac{\gamma - 1}{2}\right)^{1/(\gamma-1)} = \left(\frac{\gamma + 1}{2}\right)^{1/(\gamma-1)} \tag{2.7.31}$$

Also, $u^*/u = a^*/u$. Let us define a Mach number $M^* = u/a^*$. M^* is known as the *characteristic Mach number* and it is related to the local Mach number by the following relation:

$$M^{*2} = \frac{\frac{\gamma+1}{2} M^2}{1 + \frac{\gamma-1}{2} M^2} \tag{2.7.32}$$

Using Equations 2.7.30 through 2.7.32 in Equation 2.7.29, we can write

$$\left(\frac{A}{A^*}\right)^2 = \frac{1}{M^2}\left[\left(\frac{2}{\gamma+1}\right)\left(1 + \frac{\gamma-1}{2} M^2\right)\right]^{(\gamma+1)/(\gamma-1)} \tag{2.7.33}$$

Equation 2.7.26 is called the area Mach number relation. Figure 2.7.7 shows a plot of A/A^* against Mach number. A/A^* is always ≥1 for physically viable solutions.

The area Mach number relation indicates that for a given Mach number, there is only one area ratio A/A^*. This is a very useful relation and is frequently used to design convergent–divergent nozzles to produce a desired Mach number. Values of A/A^* are tabulated as a function of M in Anderson (2003).

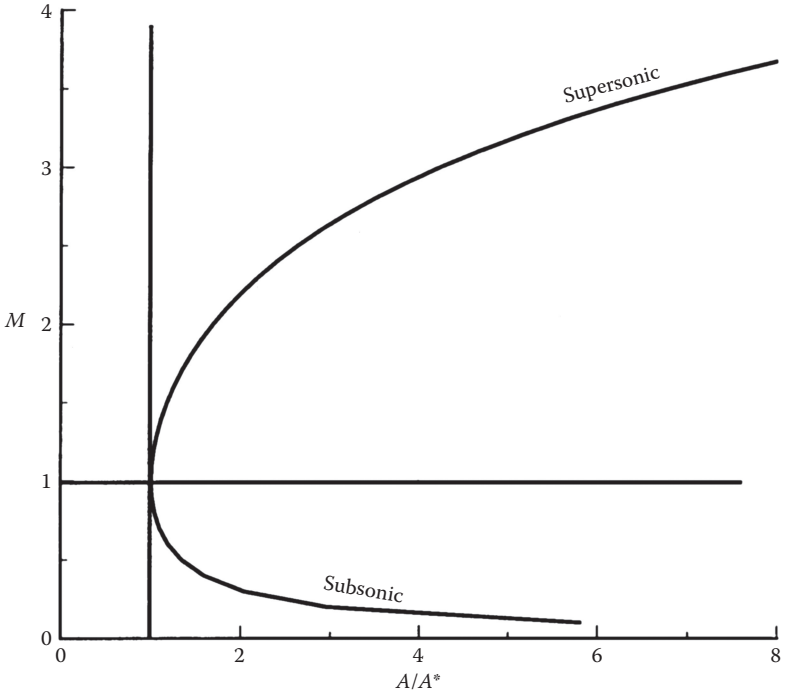


FIGURE 2.7.7 Variation of area ratio A/A^* as a function of Mach number for a one-dimensional flow with area change.

Equation 2.7.33 can also be written in terms of pressure as follows:

$$\frac{A}{A^*} - \frac{\left[1-\left(\frac{p}{p_0}\right)^{(\gamma-1)/\gamma}\right]^{1/2}\left(\frac{p}{p_0}\right)^{1/\gamma}}{\left(\frac{\gamma-1}{2}\right)^{1/2}\left(\frac{2}{\gamma+1}\right)^{(\gamma+1)/2(\gamma-1)}} \tag{2.7.34}$$

Nozzle Flow

Using the area relations, we can now plot the distributions of Mach number and pressure along a nozzle. Figure 2.7.8 shows pressure and Mach number distributions along a given nozzle and the wave configurations for several exit pressures. For curves *a* and *b*, the flow stays subsonic throughout and the exit pressure controls the flow in the entire nozzle. On curve *c*, the throat has just become sonic, and so the pressure at the throat, and upstream of it, can decrease no further. There is another exit pressure corresponding to curve *j* ($p_j < p_c$) for which a supersonic isentropic solution exists. But if the pressure lies between p_c and p_j, there is no isentropic solution possible. For example, for an exit pressure p_d, a shock will form in the nozzle at location *s* which will raise the pressure to p_d and turn the flow subsonic. The pressure will then rise to p_d as the subsonic flow goes through an increasing area duct. The location, *s*, depends on the exit pressure. Various possible situations are shown in Figure 2.7.8. It is clear that if the exit pressure is equal to or below p_f, the flow within the nozzle is fully supersonic. This is the principle used in designing supersonic wind tunnels by operating from a high-pressure reservoir or into a vacuum receiver, or both.

Diffuser

If a nozzle discharges directly into the receiver, the minimum pressure ratio for full supersonic flow in the test section is

$$\left(\frac{p_0}{p_E}\right)_{min} = \frac{p_0}{p_f}$$

where p_f is the value of p_E at which the normal shock stands right at the nozzle exit. However, by adding an additional diverging section, known as a diffuser, downstream of the test section as shown in Figure 2.7.9, it is possible to operate the tunnel at a lower pressure ratio than p_0/p_f. This change in operating conditions happens because the diffuser can now decelerate the subsonic flow downstream of the shock isentropically to a stagnation pressure p_0'. The pressure ratio required then is the ratio of stagnation pressures across a normal shock wave at the test section Mach number. In practice, the diffuser gives lower than expected recovery as a result of viscous losses caused by the interaction of shock wave and the boundary layer, which are neglected here.

The operation of supersonic wind tunnels can be made even more efficient; i.e., they can be operated at even lower pressure ratios than p_0/p_0', by using the approach shown in Figure 2.7.6 where the diffuser has a second throat. It can slow down the flow to subsonic Mach numbers isentropically and, ideally, can provide complete recovery, giving $p_0' = p_0$. However, due to other considerations, such as the starting process of the wind tunnel and viscous effects, it is not realized in real life. Rather, the tunnel must undergo a "starting" process. As the receiver pressure is decreased, the flow through the nozzle will accelerate, and as the pressure is dropped low enough to achieve sonic flow at the throat, a normal shock will sit in the diverging portion, as mentioned in the previous section. As the receiver pressure is decreased further, the shock will migrate through the diverging portion of the nozzle, until it sits at the nozzle exit. Further decrease in receiver pressure will make the shock jump to the diverging portion of the diffuser, and the wind tunnel test section will be at the proper conditions. The receiver pressure may be increased slightly, to move the shock closer to the diffuser

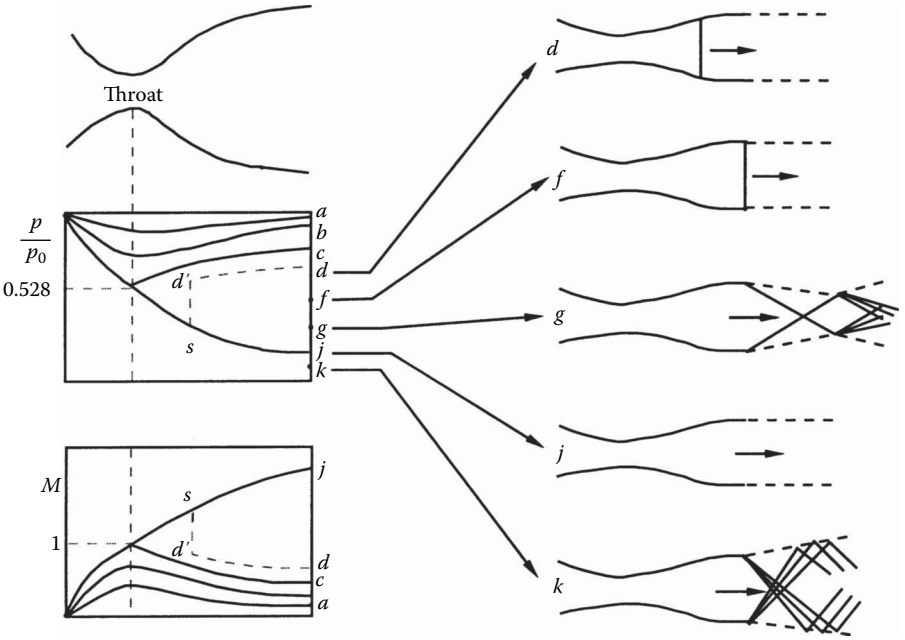


FIGURE 2.7.8 Effect of exit pressure on flow through a nozzle.

throat and decrease the stagnation pressure loss in the tunnel. At the end of the starting process, a weak normal shock wave sits just downstream of the diffuser throat. The area ratio (nozzle throat to diffuser throat) required to allow this configuration to operate is equal to the stagnation pressure ratio across a normal shock at the wind tunnel operation Mach number. If the diffuser throat is too small, the tunnel will not start properly. If the receiver pressure were to increase, the shock could jump back up to the nozzle, resulting in subsonic conditions in the test section. This phenomenon is known as unstart, and the starting process would have to be performed again.

Two-Dimensional Supersonic Flow

When supersonic flow goes over a wedge or an expansion corner, it goes through an oblique shock or expansion waves, respectively, to adjust to the change in surface geometry. Figure 2.7.10 shows the two flow situations. In Figure 2.7.10a, an oblique shock abruptly turns the flow parallel to the wedge surface. The Mach number after the shock is less than that ahead of it, and the pressure, temperature, and density all increase. In the case of an expansion corner, oblique expansion waves smoothly turn the flow to become parallel to the surface downstream of the expansion corner. In this case, the Mach number increases, and the pressure, temperature, and density decrease as the flow goes through the expansion corner. Oblique shocks and expansion waves occur in two- and

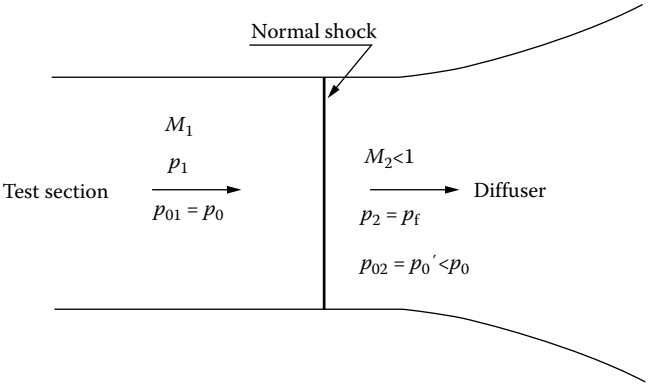


FIGURE 2.7.9 Normal shock diffuser.

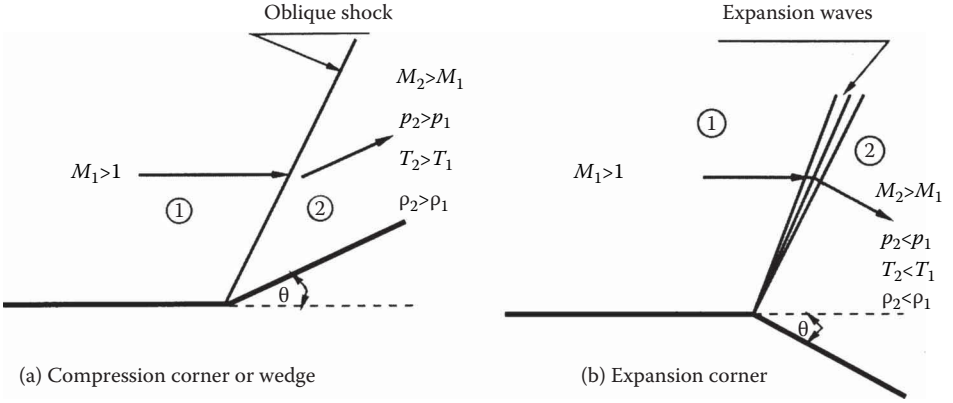


FIGURE 2.7.10 Supersonic flow over a corner.

three-dimensional supersonic flows. In this section, we will restrict ourselves to steady, two-dimensional supersonic flows.

Oblique Shock Waves

The oblique shock can be treated as the normal shock by superimposing a suitable uniform velocity v parallel to the wave. In this way, the resultant velocity before the shock can be adjusted to any desirable flow direction by adjusting the magnitude of v. As shown in Figure 2.7.11, the resultant velocity ahead of the shock is $w_1 = \sqrt{u_1^2 + v^2}$ and its direction relative to the shock is given by $\beta = \tan^{-1}(u_1/v)$. On the downstream side of the shock, since u_2 is always less than u_1, the flow always turns toward the shock. The magnitude of u_2 can be determined by the normal shock relations corresponding to velocity u_1 and the magnitude of v is such that the flow downstream of the shock turns parallel to the surface. Since imposition of a uniform velocity does not affect the pressure, temperature, etc., we can use normal shock relations with Mach number replaced in them to correspond to velocity u_1 or u_1/a_1, in other words $M_1 \sin\beta$. Thus, the oblique shock relations become

$$\frac{p_2}{p_1} = 1 + \frac{2\gamma}{\gamma+1}\left(M_1^2 \sin^2\beta - 1\right) \tag{2.7.35}$$

$$\frac{\rho_2}{\rho_1} = \frac{(\gamma+1)M_1^2 \sin^2\beta}{(\gamma-1)M_1^2 \sin^2\beta + 2} \tag{2.7.36}$$

$$\frac{T_2}{T_1} = \frac{a_2^2}{a_1^2} = \left[1 + \frac{2\gamma}{\gamma+1}\left(M_1^2 \sin^2\beta - 1\right)\right]\left[\frac{2 + (\gamma-1)M_1^2 \sin^2\beta}{(\gamma+1)M_1^2 \sin^2\beta}\right] \tag{2.7.37}$$

The Mach number M_2 $(= w_2/a_2)$ can be obtained by using a Mach number corresponding to velocity u_2 $\left(= w_2 \sin(\beta - \theta)\right)$ in the normal shock relation for the Mach number. In other words

$$M_2^2 \sin^2(\beta - \theta) = \frac{1 + \frac{\gamma-1}{2} M_1^2 \sin^2\beta}{\gamma M_1^2 \sin^2\beta - \frac{\gamma-1}{2}} \tag{2.7.38}$$

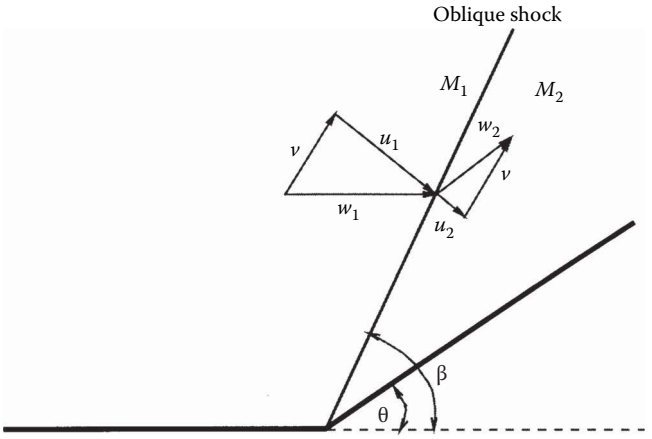


FIGURE 2.7.11 Oblique shock on a wedge.

To derive a relation between the wedge angle θ and the wave angle β, we have from Figure 2.7.11

$$\tan\beta = \frac{u_1}{v} \text{ and } \tan(\beta-\theta) = \frac{u_2}{v}$$

so that

$$\frac{\tan(\beta-\theta)}{\tan\beta} = \frac{u_2}{u_1} = \frac{\rho_1}{\rho_2} = \frac{(\gamma-1)M_1^2\sin^2\beta+2}{(\gamma+1)M_1^2\sin^2\beta}$$

This can be simplified to

$$\tan\theta = 2\cot\beta\frac{M_1^2\sin^2\beta-1}{M_1^2(\gamma+\cos 2\beta)+2} \tag{2.7.39}$$

Dennard and Spencer (1964) have tabulated oblique shock properties as a function of M_1. Let us now make some observations from the preceding relations.

From the normal shock relations, $M_1 \sin\beta \geq 1$. This condition defines a minimum wave angle for a given Mach number. The maximum wave angle, of course, corresponds to the normal shock or $\beta = \pi/2$. Therefore, the wave angle β has the following range:

$$\sin^{-1}\frac{1}{M} \leq \beta \leq \frac{\pi}{2} \tag{2.7.40}$$

Equation 2.7.39 becomes zero at the two limits of β. Figure 2.7.12 shows a plot of θ against β for various values of M_1. For each value of M_1, there is a maximum value of θ. For $\theta < \theta_{max}$, there are

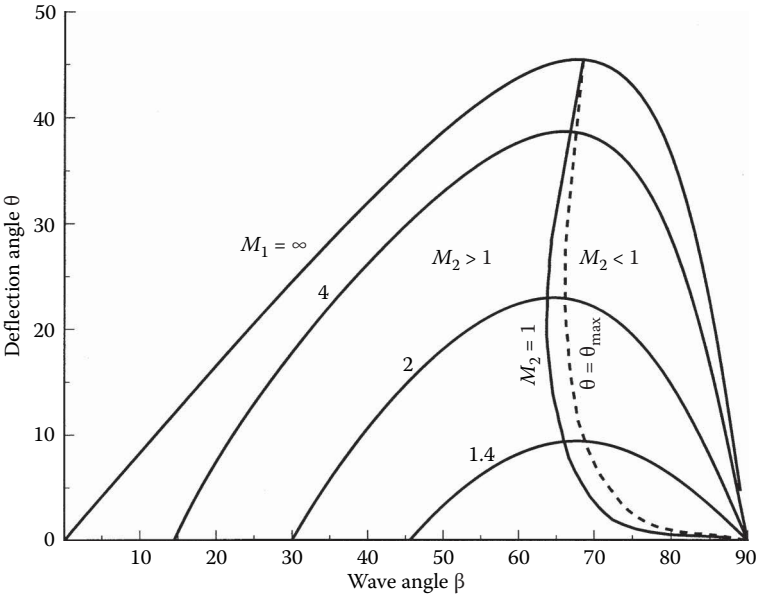


FIGURE 2.7.12 Oblique shock characteristics.

two possible solutions having different values of β. The larger value of β gives the stronger shock in which the flow becomes subsonic. A locus of solutions for which $M_2 = 1$ is also shown in the figure. It is seen from the figure that with weak shock solution, the flow remains supersonic except for a small range of θ slightly smaller than θ_{max}.

Let us now consider the limiting case of θ going to zero for the weak shock solution. As θ decreases to zero, β decreases to the limiting value μ, given by

$$M_1^2 \sin^2 \mu - 1 = 0$$
$$\mu = \sin^{-1} \frac{1}{M_1} \tag{2.7.41}$$

For this angle, the oblique shock relations show no jump in flow quantities across the wave or, in other words, there is no disturbance generated in the flow. This angle μ is called the *Mach angle* and the lines at inclination μ are called *Mach lines.*

Thin-Airfoil Theory

For a small deflection angle Δθ, it can be shown that the change in pressure in a supersonic flow at Mach M_1 between 1.2 and 5 is given approximately by

$$\frac{\Delta p}{p_1} \approx \frac{\gamma M_1^2}{\sqrt{M_1^2 - 1}} \Delta\theta \tag{2.7.42}$$

This expression holds for both compression and expansion. The Mach number range given here is based on the terms dropped in the linearization of the conservation equations. If Δp is measured with respect to the freestream pressure, p_1, and all deflections to the freestream direction, we can write Equation 2.7.35 as follows:

$$\frac{p - p_1}{p_1} = \frac{\gamma M_1^2}{\sqrt{M_1^2 - 1}} \theta \tag{2.7.43}$$

where θ is positive for a compression and negative for expansion. Let us define a pressure coefficient C_p, as follows:

$$C_p = \frac{p - p_1}{q_1}$$

where q_1 is the dynamic pressure and is equal to $\gamma p_1 M_1^2/2$. Equation 2.7.43 then gives

$$C_p = \frac{2\theta}{\sqrt{M_1^2 - 1}} \tag{2.7.44}$$

Equation 2.7.44 states that the pressure coefficient is proportional to the local flow deflection. This relation can be used to develop supersonic thin-airfoil theory. As an example, for a flat plate at angle of attack α_0 (shown in Figure 2.7.13), the pressure coefficients on the upper and lower surfaces are

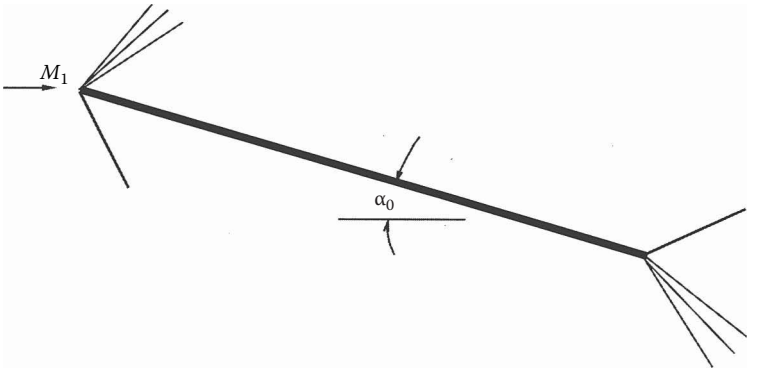


FIGURE 2.7.13 Lifting flat plate.

$$C_p = \mp \frac{2\alpha_0}{\sqrt{M_1^2 - 1}}$$

The lift and drag coefficients can be written as follows:

$$C_L = \frac{(p_L - p_U)c\cos\alpha_0}{q_1 c} = (C_{pL} - C_{pU})\cos\alpha_0$$

$$C_D = \frac{(p_L - p_U)c\sin\alpha_0}{q_1 c} = (C_{pL} - C_{pU})\sin\alpha_0$$

where c is the chord length of the plate. Since α_0 is small, we can write

$$C_L = \frac{4\alpha_0}{\sqrt{M_1^2 - 1}}, \quad C_D = \frac{4\alpha_0^2}{\sqrt{M_1^2 - 1}} \tag{2.7.45}$$

A similar type of expression can be obtained for an arbitrary thin airfoil that has thickness, camber, and angle of attack. Figure 2.7.14 shows such an airfoil. The pressure coefficients on the upper and lower surfaces can be written as follows:

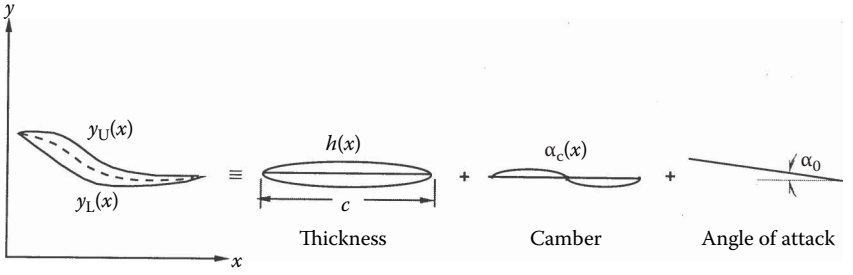


FIGURE 2.7.14 Arbitrary thin airfoil and its components.

$$C_{p_U} = \frac{2}{\sqrt{M_1^2 - 1}} \frac{dy_U}{dx}, \quad C_{p_L} = \frac{2}{\sqrt{M_1^2 - 1}} \left(\frac{dy_L}{dx} \right) \tag{2.7.46}$$

For the thin airfoil, the profile may be resolved into three separate components, as shown in Figure 2.7.14. The local slope of the airfoil can be obtained by superimposing the local slopes of the three components as follows:

$$\begin{aligned} \frac{dy_U}{dx} &= -\left(\alpha_0 + \alpha_c(x)\right) + \frac{dh}{dx} = -\alpha(x) + \frac{dh}{dx} \\ \frac{dy_L}{dx} &= -\left(\alpha_0 + \alpha_c(x)\right) - \frac{dh}{dx} = -\alpha(x) - \frac{dh}{dx} \end{aligned} \tag{2.7.47}$$

where $\alpha = \alpha_0 + \alpha_c(x)$ is the local total angle of attack of the camber line. The lift and drag for the thin airfoil are given by

$$L = q_1 \int_0^c \left(C_{p_L} - C_{p_U} \right) dx$$

$$D = q_1 \int_0^c \left[C_{p_L} \left(-\frac{dy_L}{dx} \right) + C_{p_U} \left(\frac{dy_U}{dx} \right) \right] dx$$

Let us define an average value of $\alpha(x)$ as follows:

$$\bar{\alpha} = \frac{1}{c} \int_0^c \alpha(x) dx$$

Using Equation 2.7.47 and the fact that $\bar{\alpha}_0 = \alpha$ and $\bar{\alpha}_c = 0$ by definition, the lift and drag coefficients for the thin airfoil can be written as follows:

$$\begin{aligned} C_L &= \frac{4\alpha_0}{\sqrt{M_1^2 - 1}} \\ C_D &= \frac{4}{\sqrt{M_1^2 - 1}} \left[\overline{\left(\frac{dh}{dx} \right)^2} + \overline{\alpha_C^2(x)} + \alpha_0^2 \right] \end{aligned} \tag{2.7.48}$$

Equation 2.7.48 shows that the lift coefficient depends only on the mean angle of attack whereas the drag coefficient is a linear combination of the drag due to thickness, drag due to camber, and drag due to lift (or mean angle of attack).

REFERENCES

Anderson, J.D. 2003. *Modern Compressible Flowwith Historical Perspective*, 3rd ed., McGraw-Hill, New York.

Dennard, J.S. and Spencer, P.B. 1964. Ideal-Gas Tables for Oblique-Shock Flow Parameters in Air at Mach Numbers from 1.05 to 12.0. NASA TN D-2221.

Liepmann, H.W. and Roshko, A. 1966. *Elements of Gas Dynamics*, John Wiley & Sons, New York.

Further Information

As mentioned in the beginning, this section discussed only one- or two-dimensional steady, inviscid compressible flows and assumes ideal gas behavior. Even this discussion was quite brief because of space limitations. For more details on the subject as well as for compressible unsteady viscous flows, readers are referred to Anderson (2003) and Liepmann and Roshko (1966).

2.8 MULTIPHASE FLOW

John C. Chen

Introduction

Classic study of fluid mechanics concentrates on the flow of a single homogeneous phase, e.g., water, air, steam. However, many industrially important processes involve simultaneous flow of multiple phases, e.g., gas bubbles in oil, wet steam, dispersed particles in gas or liquid, foams, emulsions. Examples include vapor–liquid flow in refrigeration systems, steam–water flows in boilers and condensers, vapor–liquid flows in distillation columns, pneumatic hydraulic transport of solid particulates. In spite of their importance, multiphase flows are often neglected in standard textbooks. Fundamental understanding and engineering design procedures for multiphase flows are not nearly so well developed as those for single-phase flows. An added complexity is the need to predict the relative concentrations of the different phases in the multiphase flows, a need that doesn't exist for single-phase flows.

Inadequate understanding not withstanding, a significant amount of data have been collected and combinations of theoretical models and empirical correlations are used in engineering calculations. This knowledge base is briefly summarized in this section and references are provided for additional information. While discussions are provided of solid–gas flows and solid–liquid flows, primary emphasis is placed on multiphase flow of gas–liquids since this is the most often encountered class of multiphase flows in industrial applications.

A multiphase flow occurs whenever two or more of the following phases occur simultaneously: gas/vapor, solids, single-liquid phase, multiple (immiscible) liquid phases. Every possible combination has been encountered in some industrial process, the most common being the simultaneous flow of vapor/gas and liquid (as encountered in boilers and condensers). All multiphase flow problems have features which are characteristically different from those found in single-phase problems. First, the relative concentration of different phases is usually a dependent parameter of great importance in multiphase flows, while it is a parameter of no consequence in single-phase flows. Second, the spatial distribution of the various phases in the flow channel strongly affects the flow behavior, again a parameter that is of no concern in single-phase flows. Finally, since the density of various phases can differ by orders of magnitude, the influence of gravitational body force on multiphase flows is of much greater importance than in the case of single-phase flows. In any given flow situation, the possibility exists for the various phases to assume different velocities, leading to the phenomena of slip between phases and consequent interfacial momentum transfer. Of course, the complexity of laminar/turbulent characteristics occurs in multiphase flows as in single-phase flows, with the added complexity of interactions between phases altering the laminar/turbulent flow structures. These complexities increase exponentially with the number of phases encountered in the multiphase problem. Fortunately, a large number of applications occur with just two phase flows, or can be treated as pseudo-two-phase flows.

Two types of analysis are used to deal with two-phase flows. The simpler approach utilizes homogeneous models which assume that the separate phases flow with the same identical local velocity at all points in the fluid. The second approach recognizes the possibility that the two phases can flow at different velocities throughout the fluid, thereby requiring separate conservation equations for mass and momentum for each phase. Brief descriptions of both classes of models are given below.

Fundamentals

Consider n phases in concurrent flow through a duct with cross-sectional area A_c. Fundamental quantities that characterize this flow are

$\dot{m}_i$ = mass flow rate of *i*th phase
u_i = velocity of *i*th phase
α_i = volume fraction of *i*th phase in channel

Basic relationships between these and related parameters are

$$G_i = \text{mass flux of } i\text{th phase} = \frac{\dot{m}_i}{A_c} \tag{2.8.1}$$

$$v_i = \text{superficial velocity of } i\text{th phase} = \frac{G_i}{\rho_i} \tag{2.8.2}$$

$$u_i = \text{actual velocity of } i\text{th phase} = \frac{v_i}{\alpha_i} \tag{2.8.3}$$

$$x_i = \text{flow quality of } i\text{th phase} = \frac{\dot{m}_i}{\sum_i^n m_i} = \frac{G_i}{\sum_{i=1}^n G_i} \tag{2.8.4}$$

$$\alpha_i = \text{volume fraction of } i\text{th phase} = \frac{\left(\frac{x_i}{\rho_i u_i}\right)}{\sum_{i=1}^n \left(\frac{x_i}{\rho_i u_i}\right)} \tag{2.8.5}$$

In most engineering calculations, the above parameters are defined as average quantities across the entire flow area, A_c. It should be noted, however, that details of the multiphase flow could involve local variations across the flow area. In the latter situation, G_i, v_i, and α_i are often defined on a local basis, varying with transverse position across the flow area.

Pressure drop along the flow channel is associated with gravitational body force, acceleration forces, and frictional shear at the channel wall. The total pressure gradient along the flow axis can be represented as

$$\frac{dP}{dz} = \left(\frac{dP}{dz}\right)_g + \left(\frac{dP}{dz}\right)_a + \left(\frac{dP}{dz}\right)_f \tag{2.8.6}$$

where

$$\left(\frac{dP}{dz}\right)_g = -g\cos\theta \cdot \sum_{i=1}^n \alpha_i \rho_i \tag{2.8.7}$$

θ = angle of channel from vertical and the flow is in the upward direction.

and

$$\left(\frac{dP}{dz}\right)_a = -\sum_{i=1}^{n} G_i \frac{du_i}{dz} \tag{2.8.8}$$

$$\left(\frac{dP}{dz}\right)_f = -\frac{\rho u^2}{2D} f \tag{2.8.9}$$

ρ = density of multiphase mixture

$$= \sum_{i=1}^{n} \rho_i \alpha_i \tag{2.8.10}$$

u = an average mixture velocity

$$= \frac{1}{\rho} \sum_{i=1}^{n} G_i \tag{2.8.11}$$

f = equivalent Darcy friction factor for the multiphase flow

In applications, the usual requirement is to determine pressure gradient (dP/dz) and the volume fractions (α_i). The latter quantities are of particular importance since the volume fraction of individual phases affects all three components of the pressure gradient, as indicated in Equations 2.8.7 through 2.8.11. Correlations of various types have been developed for prediction of the volume fractions, all but the simplest of which utilize empirical parameters and functions.

The simplest flow model is known as the homogeneous equilibrium model (HEM), wherein all phases are assumed to be in neutral equilibrium. One consequence of this assumption is that individual phase velocities are equal for all phases everywhere in the flow system, i.e., there is no-slip between the phases:

$$u_i = u \text{ for all } i \tag{2.8.12}$$

This assumption permits direct calculation of the volume fractions from known mass qualities:

$$\alpha_i = \frac{x_i}{\rho_i \sum_{i=1}^{n} \left(\frac{x_i}{\rho_i}\right)} \tag{2.8.13}$$

The uniform velocity for all phases is the same as mixture velocity:

$$u = \frac{1}{\rho} \sum_{i=1}^{n} G_i \tag{2.8.14}$$

where

$$\frac{1}{\rho} = \sum_{i=1}^{n} \left(\frac{x_i}{\rho_i}\right) \tag{2.8.15}$$

This homogeneous model permits direct evaluation of all three components of axial pressure gradient, if flow qualities (x_i) are known:

$$\left(\frac{dP}{dz}\right)_g = -\frac{g\cos\theta}{\sum_{i=1}^{n}\left(\frac{x_i}{\rho_i}\right)} \tag{2.8.16}$$

$$\left(\frac{dP}{dz}\right)_a = -\left(\sum_{i=1}^{n} G_i\right)\cdot\frac{du}{dz} \tag{2.8.17}$$

$$\left(\frac{dP}{dz}\right)_f = -\frac{\rho u^2}{2D}\cdot f \tag{2.8.18}$$

where u and ρ are given by Equations 2.8.14 and 2.8.15.

Predicting the coefficient of friction, f, remains a problem, even in the homogeneous model. For cases of fully turbulent flows, experience has shown that a value of 0.02 may be used as a first-order approximation for f, i.e., $f = 0.02$. More-accurate estimates require empirical correlations, specific to particular classes of multiphase flows and subcategories of flow regimes.

The following parts of this section consider the more common situations of two-phase flows and describe improved design methodologies specific to individual situations.

Gas–Liquid Two-Phase Flow

The most common case of multiphase flow is two-phase flow of gas and liquid, as encountered in steam generators and refrigeration systems. A great deal has been learned about such flows, including delineation of flow patterns in different flow regimes, methods for estimating volume fractions (gas void fractions), and two-phase pressure drops.

Flow Regimes

A special feature of multiphase flows is their ability to assume different spatial distributions of the two phases. These different flow patterns have been classified in flow regimes, which are themselves altered by the direction of flow relative to gravitational acceleration. Figures 2.8.1 and 2.8.2 (Delhaye, 1981; Michaelides et al, 2016) show the flow patterns commonly observed for co-current flow of gas and liquid in vertical and horizontal channels, respectively. For a constant liquid flow rate, the gas phase tends to be distributed as small bubbles at low gas flow rates. Increasing gas flow rate causes agglomeration of bubbles into larger slugs and plugs. Further increasing gas flow rate causes separation of the phases into annular patterns wherein liquid concentrates at the channel wall and gas flows in the central core for vertical ducts. For horizontal ducts, gravitational force tends to drain the liquid annulus toward the bottom of the channel, resulting in stratified and stratified wavy flows. This downward segregation of the liquid phase can be overcome by kinetic forces at high flow rates, causing stratified flows to revert to annular flows. At high gas flow rates, more of the liquid tends to be entrained as dispersed drops; in the limit one obtains completely dispersed mist flow.

Flow pattern maps are utilized to predict flow regimes for specific applications. The first generally successful flow map was that of Baker (1954) for horizontal flow, reproduced here in Figure 2.8.3. For vertical flows, the map of Hewitt and Roberts (1969), duplicated in Figure 2.8.4, provides a simple method for determining flow regimes. Parameters used for the axial coordinates of these flow maps are defined as follows:

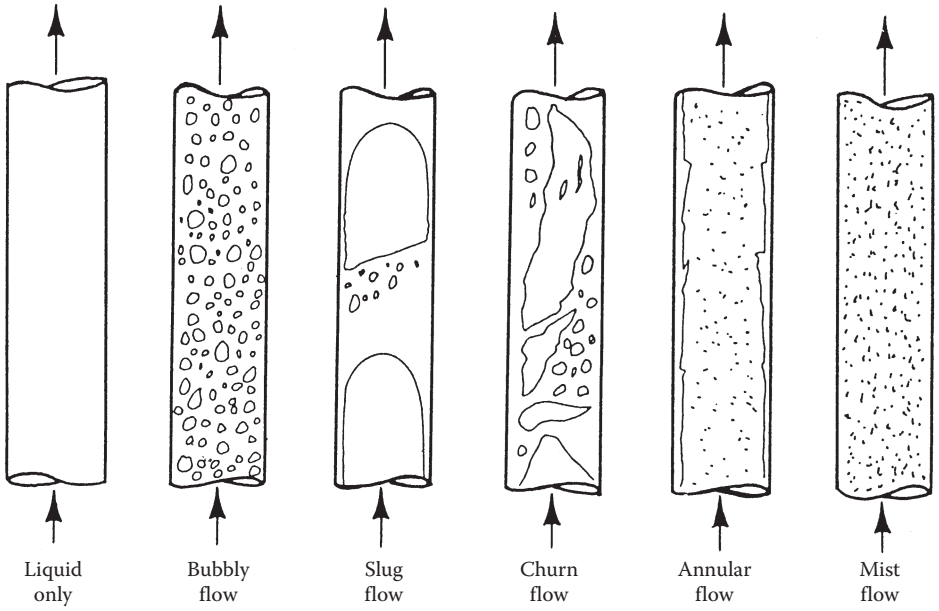


FIGURE 2.8.1 Flow patterns in gas–liquid vertical flow. (From Lahey, R.T., Jr. and Moody, F.I., 1977, *The Thermal Hydraulics of a Boiling Water Nuclear Reactor*, The American Nuclear Society, LaGrange, IL. With permission.)

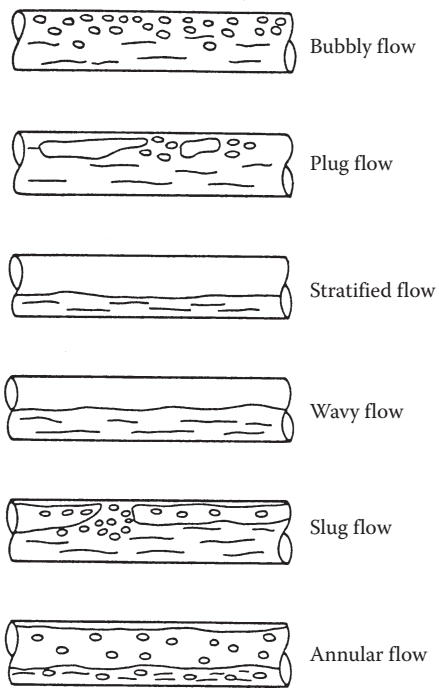


FIGURE 2.8.2 Flow patterns in gas–liquid horizontal flow.

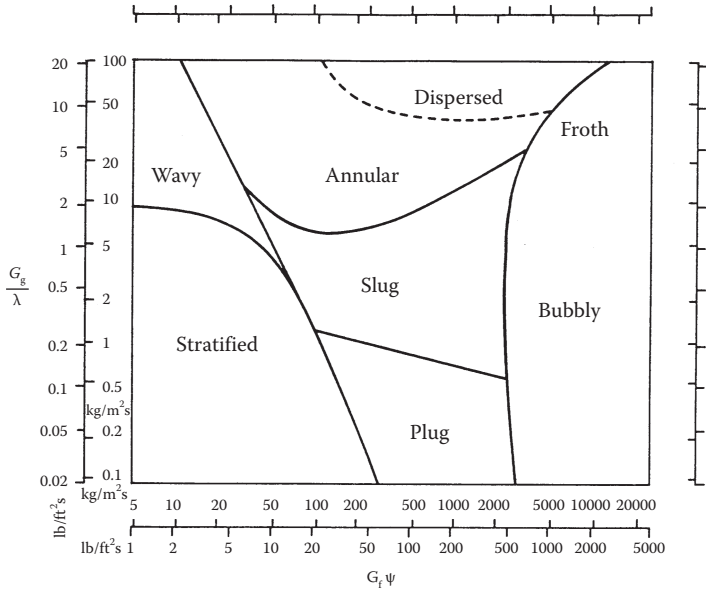


FIGURE 2.8.3 Flow pattern map for horizontal flow (Baker, 1954). (From Collier, J.G., 1972, *Convective Boiling and Condensation*, McGraw-Hill, London. With permission.)

$$\lambda = \left(\frac{\rho_g \rho_l}{\rho_a \rho_w}\right)^{1/2} \tag{2.8.19}$$

$$\psi = \left(\frac{\sigma_w}{\sigma}\right)\left[\left(\frac{\mu_l}{\mu_w}\right)\left(\frac{\rho_w}{\rho_l}\right)^2\right]^{1/3} \tag{2.8.20}$$

$$j = \text{volumetric flux}, \frac{G}{\rho} \tag{2.8.21}$$

Void Fractions

In applications of gas–liquid flows, the volume fraction of gas (α_g) is commonly called "void fraction" and is of particular interest. The simplest method to estimate void fraction is by the HEM. From Equation 2.8.13, the void fraction can be estimated as

$$\alpha_g = \frac{x_g}{x_g + (1 - x_g)\dfrac{\rho_g}{\rho_l}} \tag{2.8.22}$$

where α_g, x_g, ρ_g, ρ_ϵ are cross-sectional averaged quantities.

In most instances, the homogeneous model tends to overestimate the void fraction. Improved estimates are obtained by using separated-phase models which account for the possibility of slip between gas and liquid velocities. A classic separated-phase model is that of Lockhart and Martinelli (1949). The top portion of Figure 2.8.5 reproduces the Lockhart–Martinelli correlation for void fraction (shown as α) as a function of the parameter X which is defined as

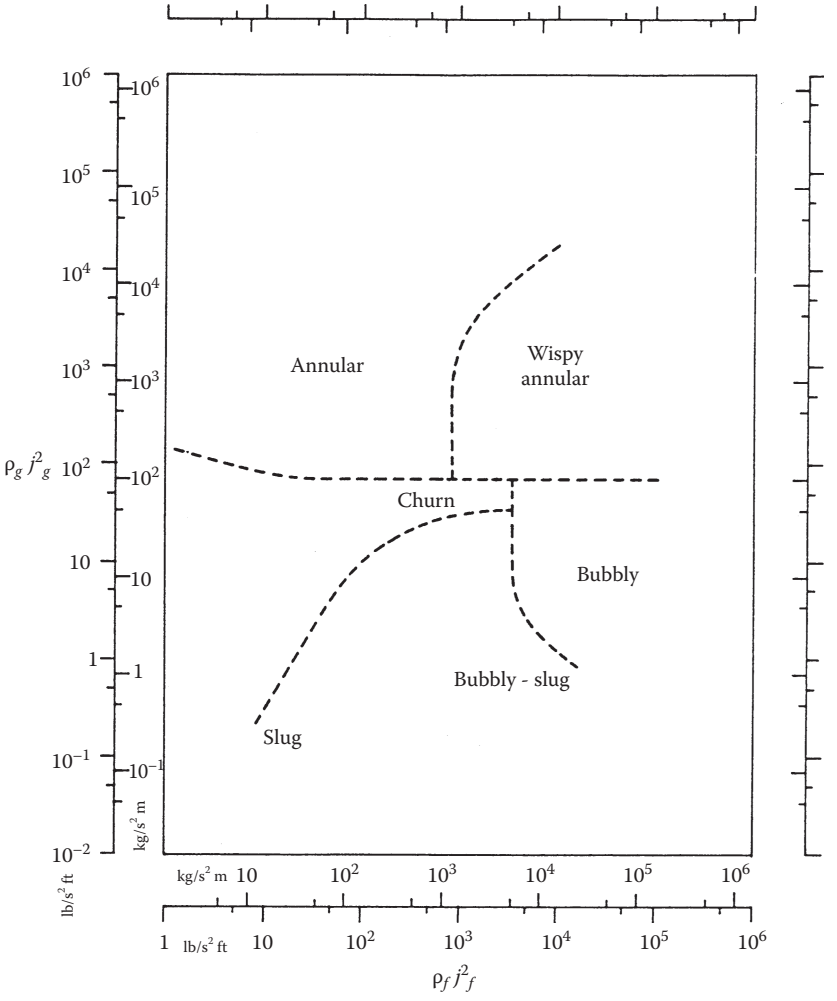


FIGURE 2.8.4 Flow pattern map for vertical flow (Hewitt and Roberts, 1969). (From Collier, J.G., 1972, *Convective Boiling and Condensation*, McGraw-Hill, London. With permission.)

$$X = \left[\left(\frac{dP}{dz} \right)_{fl} \Big/ \left(\frac{dP}{dz} \right)_{fg} \right]^{1/2} \tag{2.8.23}$$

where

$\left(\frac{dP}{dz} \right)_{fl}$ = frictional pressure gradient of liquid phase flowing alone in channel

$\left(\frac{dP}{dz} \right)_{fg}$ = frictional pressure gradient of gas phase flowing alone in channel

Often, flow rates are sufficiently high such that each phase if flowing alone in the channel would be turbulent. In this situation the parameter X can be shown to be

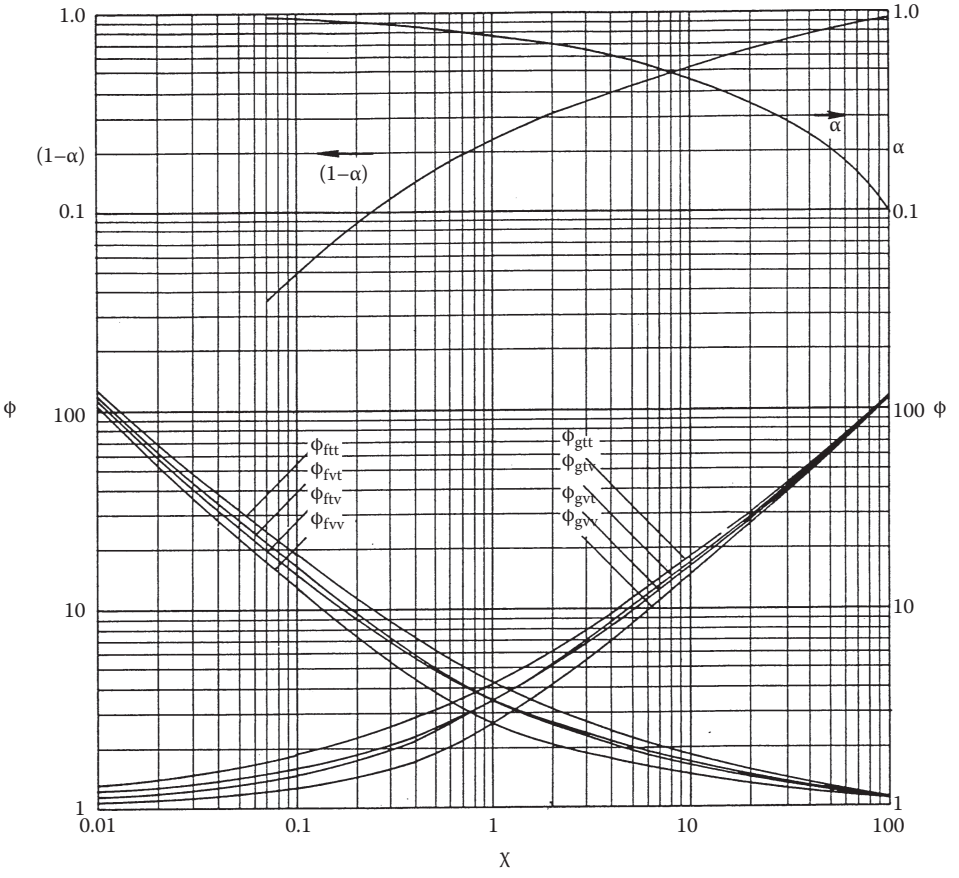


FIGURE 2.8.5 Correlations for void fraction and frictional pressure drop (Lockhart and Martinelli, 1949). (From Collier, J.G., 1972, *Convective Boiling and Condensation*, McGraw-Hill, London. With permission.)

$$X_{tt} = \left(\frac{1-x_g}{x_g}\right)^{0.9}\left(\frac{\rho_g}{\rho_l}\right)^{0.5}\left(\frac{\mu_l}{\mu_g}\right)^{0.1} \tag{2.8.24}$$

Another type of separated-phase model is the drift-flux formulation of Wallis (1969). This approach focuses attention on relative slip between phases and results in slightly different expressions depending on the flow regime. For co-current upflow in two of the more common regimes, the drift-flux model gives the following relationships between void fraction and flow quality:

Bubbly flow or churn-turbulent flow:

$$\alpha_g = \frac{x_g}{\left(\frac{u_o\rho_g}{G}\right) + C_o\left[x_g + (1-x_g)\frac{\rho_g}{\rho_l}\right]} \tag{2.8.25}$$

Dispersed drop (mist) flow:

$$x_g = \frac{1-(1-\alpha_g)\left(\frac{u_o\rho_l}{G}\alpha_g^2 + 1\right)}{1-(1-\alpha_g)\left(1-\frac{\rho_l}{\rho_g}\right)} \tag{2.8.26}$$

where u_o = terminal rise velocity of bubble, in bubbly flow, or terminal fall velocity of drop in churn-turbulent flow

C_o = an empirical distribution coefficient $\simeq 1.2$

Pressure Drop

Equations 2.8.16 through 2.8.18 permit calculation of the two-phase pressure drop by the homogeneous model, if the friction coefficient (f) is known. One useful method for estimating (f) is to treat the entire two-phase flow as if it were all liquid, except flowing at the two-phase mixture velocity. By this approach the frictional component of the two-phase pressure drop becomes

$$\left(\frac{dP}{dz}\right)_f = \left[1 + x_g\left(\frac{\rho_l}{\rho_g} - 1\right)\right] \cdot \left(\frac{dP}{dz}\right)_{flG} \tag{2.8.27}$$

where $(dP/dz)_{flG}$ = frictional pressure gradient if entire flow (of total mass flux G) flowed as liquid in the channel.

The equivalent frictional pressure drop for the entire flow as liquid, $(dP/dz)_{fl_g}$, can be calculated by standard procedures for single-phase flow. In using Equations 2.8.16 through 2.8.18, the void fraction would be calculated with the equivalent homogeneous expression Equation 2.8.13.

A more accurate method to calculate two-phase pressure drop is by the separated-phases model of Lockhart and Martinelli (1949). The bottom half of Figure 2.8.5 shows empirical curves for the Lockhart–Martinelli frictional multiplier, ϕ:

$$\phi_i = \left[\left(\frac{dP}{dz}\right)_f \Big/ \left(\frac{dP}{dz}\right)_{fi}\right]^{1/2} \tag{2.8.28}$$

where (i) denotes either the fluid liquid phase (f) or gas phase (g). The single-phase frictional gradient is based on the ith phase flowing alone in the channel, in either viscous laminar (v) or turbulent (t) modes. The most common case is where each phase flowing alone would be turbulent, whence one could use Figure 2.8.5 to obtain

$$\begin{aligned}\left(\frac{dP}{dz}\right)_f &= \text{frictional pressure gradient for two-phase flow} \\ &= \phi_{gtt}^2 \cdot \left(\frac{dP}{dz}\right)_{fg}\end{aligned} \tag{2.8.29}$$

where $(dP/dz)_{fg}$ is calculated for gas phase flowing alone and $X = X_{tt}$ as given by Equation 2.8.24.

The correlation of Lockhart–Martinelli has been found to be adequate for two-phase flows at low-to-moderate pressures, i.e., with reduced pressures less than 0.3. For applications at higher pressures, the revised models of Martinelli and Nelson (1984) and Thom (1964) are recommended.

Gas–Solid, Liquid–Solid Two-Phase Flows

Two-phase flows can occur with solid particles in gas or liquid. Such flows are found in handling of granular materials and heterogeneous reaction processing. Concurrent flow of solid particulates with a fluid phase can occur with various flow patterns, as summarized below.

Flow Regimes

Consider vertical upflow of a fluid (gas or liquid) with solid particles. Figure 2.8.6 illustrates the major flow regimes that have been identified for such two-phase flows. At low flow rates, the fluid phase percolates between stationary particles; this is termed flow through a fixed bed. At some higher velocity a point is reached when the particles are all suspended by the upward flowing fluid, the drag force between particles and fluid counterbalancing the gravitational force on the particles. This is the point of minimum fluidization, marking the transition from fixed to fluidized beds. Increase of fluid flow rate beyond minimum fluidization causes instabilities in the two-phase mixture, and macroscopic bubbles or channels of fluid are observed in the case of gaseous fluids. In the case of liquids, the two-phase mixture tends to expand, often without discrete bubbles or channels. Further increase of fluid velocity causes transition to turbulent fluidization wherein discrete regions of separated phases (fluid slugs or channels and disperse suspensions of particles) can coexist. Depending on specific operating conditions (e.g., superficial fluid velocity, particle size, particle density, etc.), net transport of solid particles with the flowing fluid can occur at any velocity equal to or greater than that associated with slug flow and turbulent flow. Further increases in fluid velocity increase the net transport of solid particles. This can occur with large-scale clusters of solid particles (as exemplified by the fast fluidization regime) or with dilute dispersions of solid particles (as often utilized in pneumatic conveying). For engineering application of fluid–solid two-phase flows, the important thresholds between flow regimes are marked by the fluid velocity for minimum fluidization, terminal slip, and saltation threshold.

Minimum Fluidization

The transition from flow through packed beds to the fluidization regime is marked by the minimum fluidization velocity of the fluid. On a plot pressure drop vs. superficial fluid velocity, the point of

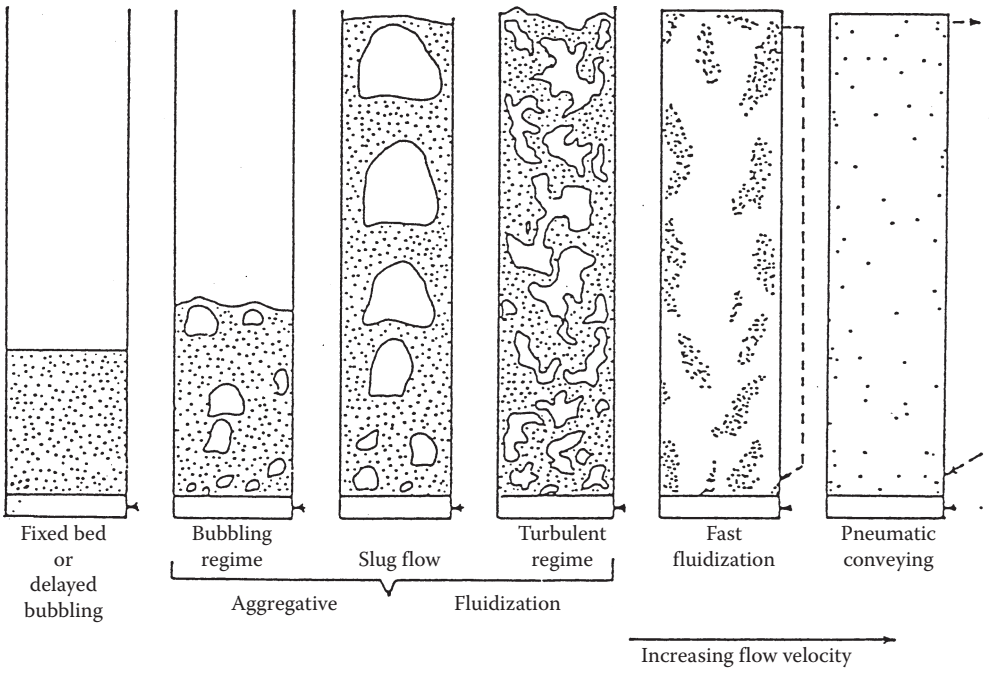


FIGURE 2.8.6 Flow patterns for vertical upflow of solid particles and gas or liquid. (From Chen, J.C., 1994, *Proc. Xth Int. Heat Transfer Conf.*, Brighton, UK, 1:369–386. With permission.)

minimum fluidization is marked by a transition from a linearly increasing pressure drop to a relatively constant pressure drop as shown in Figure 2.8.7 for typical data, for two-phase flow of gas with sand particles of 280 μm mean diameter (Chen, 1996). The threshold fluid velocity at minimum fluidization is traditionally derived from the Carman–Kozeny equation,

$$U_{mf} = \frac{(\rho_s - \rho_f)(\phi\, dp)^2\, g}{150\,\mu_f} \cdot \frac{\alpha_{mf}^2}{(1-\alpha_{mf})} \tag{2.8.30}$$

where ϕ = sphericity of particles (unity for spherical particles)

α_{mf} = volumetric fraction of fluid at minimum fluidization

Small, light particles have minimum fluidization voidage (α_{mf}) of the order 0.6, while larger particles such as sand have values closer to 0.4.

An alternative correlation for estimating the point of minimum fluidization is that of Wen and Yu (1966):

$$\frac{U_{mf} d_p \rho_f}{\mu_f} = (33.7 + 0.041 Ga)^{0.5} - 33.7 \tag{2.8.31}$$

where $Ga = \rho_f d_p^3 (\rho_s - \rho_f)\, g / \mu_f^2$.

When the fluid velocity exceeds U_{mf}, the two-phase mixture exists in the fluidized state in which the pressure gradient is essentially balanced by the gravitational force on the two-phase mixture:

$$\frac{dP}{dz} = g\left[\alpha_s \rho_s + \alpha_f \rho_f\right] \tag{2.8.32}$$

This fluidized state exists until the fluid velocity reaches a significant fraction of the terminal slip velocity, beyond which significant entrainment and transport of the solid particles occur.

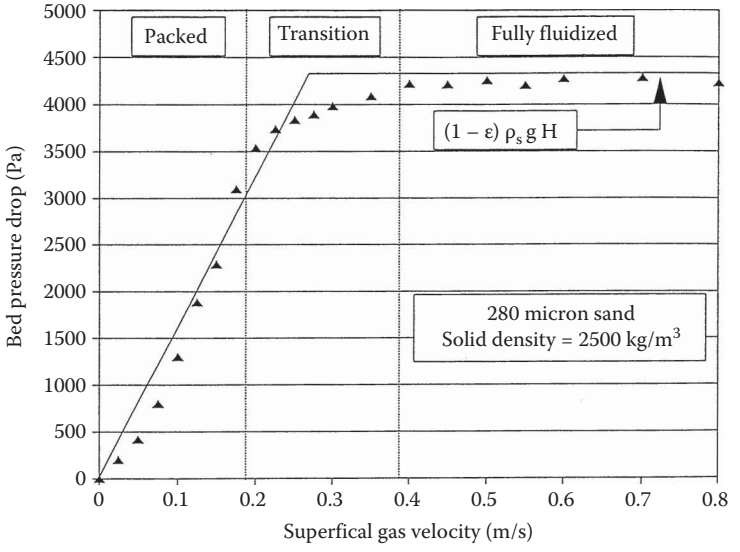


FIGURE 2.8.7 Transition at minimum fluidization. (From Chen, J.C., 1996, In *Annual Reviews of Heat Transfer*, Vol. VII, Begal House, Washington, DC With permission.)

Terminal Slip Velocity

For an isolated single particle, the maximum velocity relative to an upflowing fluid is the terminal slip velocity. At this condition, the interfacial drag of the fluid on the particle exactly balances the gravitational body force on the particle:

$$U_t = \left(U_f - U_s\right)_t = \left[\frac{4d_p\left(\rho_s - \rho_f\right)}{3\rho_f} \cdot \frac{1}{C_D}\right]^{1/2} \tag{2.8.33}$$

where C_D = coefficient of drag on the particle.

The coefficient of drag on the particle (C_D) depends on the particle Reynolds number:

$$\mathrm{Re}_p = \frac{\rho_f d_p\left(U_f - U_s\right)}{\mu_f} \tag{2.8.34}$$

The following expressions may be used to estimate C_D as appropriate:

$$\begin{aligned} C_D &= \frac{32}{\mathrm{Re}_p}, \qquad \mathrm{Re}_p \le 1 \\ C_D &= \frac{18.5}{\mathrm{Re}_p^{0.67}}, \qquad 1 \le \mathrm{Re}_p \le 10^3 \end{aligned} \tag{2.8.35}$$

Pneumatic Conveying

A desirable mode of pneumatic conveying is two-phase flow with solid particles dispersed in the concurrent flowing fluid. Such dispersed flows can be obtained if the fluid velocity is sufficiently high. For both horizontal and vertical flows, there are minimum fluid velocities below which saltation

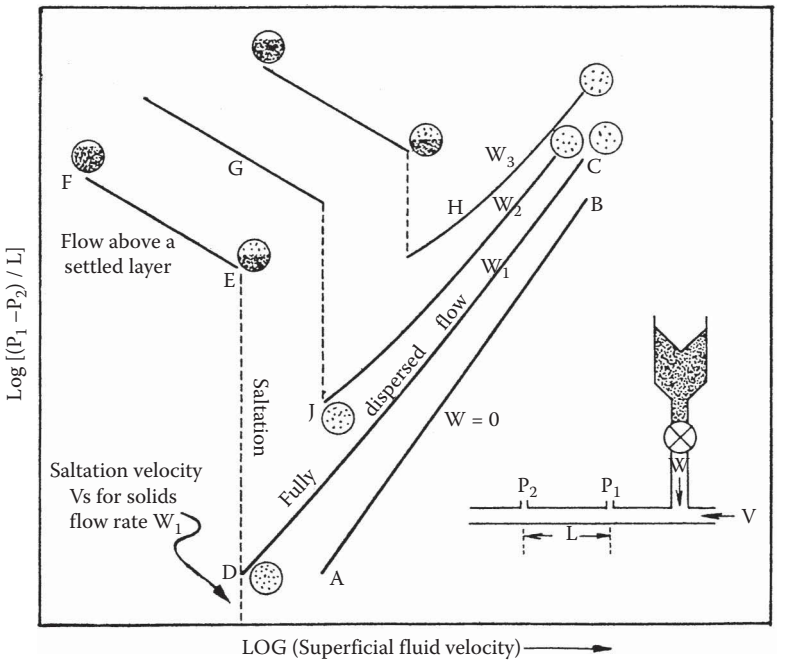


FIGURE 2.8.8 Flow characteristics in horizontal pneumatic conveying. (From Zeng, F.A. and Othmer, D.F., 1960, *Fluidization and Fluid-Particle Systems*, Reinhold, New York. With permission.)

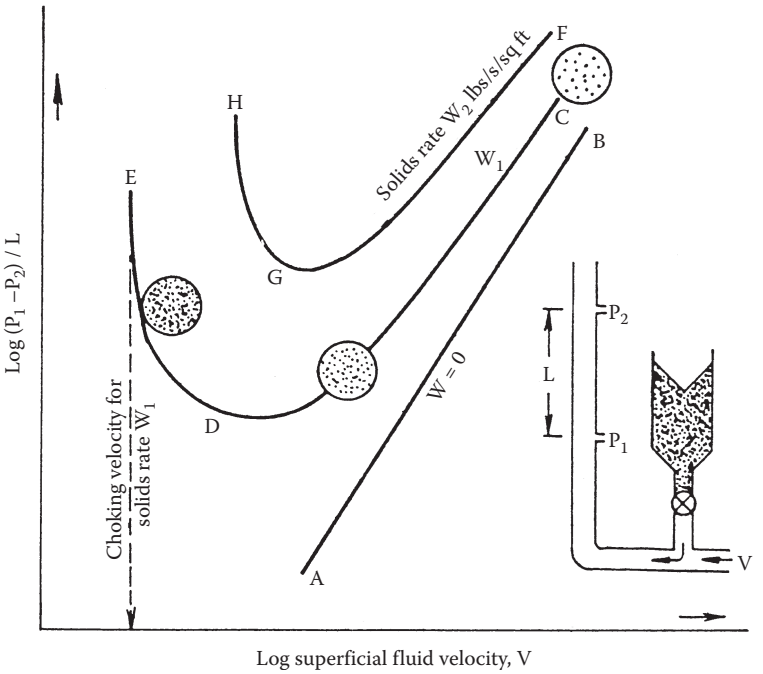


FIGURE 2.8.9 Flow characteristics in vertical pneumatic conveying. (From Zeng, F.A. and Othmer, D.F., 1960, *Fluidization and Fluid-Particle Systems*, Reinhold, New York. With permission.)

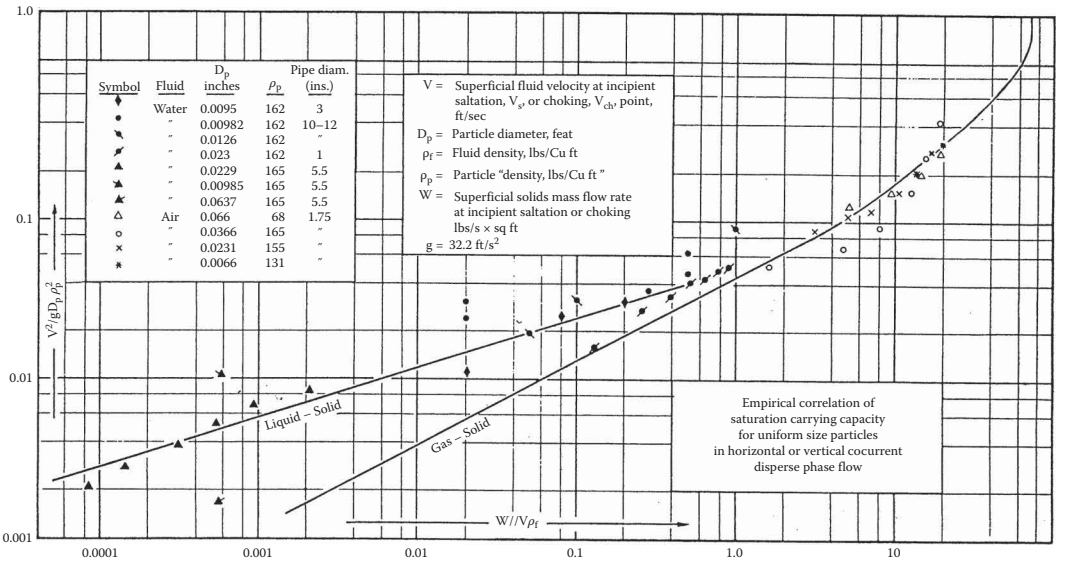


FIGURE 2.8.10 Correlations for limiting velocities in pneumatic conveying. (From Zeng, F.A. and Othmer, D.F., 1960, *Fluidization and Fluid-Particle Systems*, Reinhold, New York. With permission.)

of the solid particles due to gravitational force occurs, leading to settling of the solid particles in horizontal channels and choking of the particles in vertical channels. Figures 2.8.8 and 2.8.9, from Zenz and Othmer (1960), show these different regimes of pneumatic conveying for horizontal and vertical transport, respectively. Figure 2.8.8 shows that for a given rate of solids flow (W) there is a minimum superficial fluid velocity below which solid particles tend to settle into a dense layer at the

bottom of the horizontal channels. Above this saltation threshold, fully dispersed two-phase flow is obtained. In the case of vertical transport illustrated in Figure 2.8.9, there is a minimum fluid velocity below which solid particles tend to detrain from the two-phase suspension. This choking limit varies not only with particle properties but also with the actual rate of particle flow. Well-designed transport systems must operate with superficial fluid velocities greater than these limiting saltation and choking velocities.

Zenz and Othmer (1960) recommend the empirical correlations represented in Figure 2.8.10 estimating limiting superficial fluid velocities at incipient saltation or choking, for liquid or gas transport of uniformly sized particles. Note that these correlations are applicable for either horizontal or vertical concurrent flow. Figure 2.8.10 is duplicated from the original source and is based on parameters in engineering units, as noted in the figure. To operate successfully in dispersed pneumatic conveying of solid particles, the superficial fluid velocity must exceed that determined from the empirical correlations of Figure 2.8.10.

More detailed treatment of multiphase flows can be found in several books, e.g., see Wilson et al. (2006) for slurry flows, klinzing et al. (1997) for preumatic conveying and Michaelides et al. (2016) for gas-liquid and systems.

REFERENCES

Baker, O. 1954. Design of pipelines for simultaneous flow of oil and gas, *Oil Gas J.* 53:185–195.

Chen, J.C. 1994. Two-phase flow with and without phase changes: suspension flows. Keynote lecture, *Proc. Xth Int. Heat Transfer Conf.*, Brighton, UK 1:369–386.

Chen, J.C. 1996. Heat transfer to immersed surfaces in bubbling fluidized beds, in *Annual Review of Heat Transfer*, Vol. VII, Bengel House, Washington, D.C.

Collier, J.G. 1972. *Convective Boiling and Condensation*, McGraw-Hill, London.

Delhaye, J.M. 1981. Two-phase flow patterns, in *Two-Phase Flow and Heat Transfer*, A.E. Bergles, J.G. Collier, J.M. Delhaye, G.F. Hewitt, and F. Mayinger, Eds., Hemisphere Publishing, McGraw-Hill, New York.

Hewitt, G.F. and Roberts, D.N. 1969. Studies of two-phase flow patterns by simultaneous x-ray and flash photography, Report AERE-M 2159.

Klinzing, G.E., Marchus, R.D., Rizk, F and Leung, L.S. 1997. Pneumatic conveying of solids, 2nd edn., Springer, London, U.K.

Lahey, R.T., Jr. and Moody, F.I. 1977. *The Thermal Hydraulics of a Boiling Water Nuclear Reactor*, The American Nuclear Society, La Grange, IL.

Lockhart, R.W. and Martinelli, R.C. 1949. Proposed correlation of data for isothermal two-phase two-component flow in pipes, *Chem. Eng. Progr.* 45:39.

Martinelli, R.C. and Nelson, D.B. 1984. Prediction of pressure drop during forced-circulation boiling of water, *Trans. ASME* 70:695–702.

Michaelides, E.E., Schwartzkopf, J.D. and Crowe, C., eds. 2016. Multiphase Flow Handbook, 2nd edn, CRC press, Boca Raton, FL.

Thom, J.R.S. 1964. Prediction of pressure drop during forced circulation boiling of water, *Int. J. Heat Mass Transfer* 7:709–724.

Wallis, G.B. 1969. *One-Dimensional Two-Phase Flow*, McGraw-Hill, New York.

Wen, C.Y. and Yu, Y.H. 1966. A generalized method of predicting the minimum fluidization velocity, *AIChE J.* 12:610–612.

Wilson, K.C., Addie, G.R., Sellgren, A and Clift, R. 2006. Slurry transport using centrifugal pumps, 3rd edn., Springer, New York.

Zenz, F.A. and Othmer, D.F. 1960. *Fluidization and Fluid-Particle Systems*, Reinhold, New York.

2.9 NON-NEWTONIAN FLOWS

Anoop K. Gupta, Raj P. Chhabra, Thomas F. Irvine, Jr., and Massimo Capobianchi

Introduction

It is now readily acknowledged that most structured fluids (polymer solutions and melts, gas–liquid dispersions, foams, slurries and suspensions, worm-like micellar solutions containing surfactant molecules, for instance) exhibit what is known as complex non-Newtonian behavior under the application of an external stress. Conversely, such fluids do not conform to Newton's law of viscosity and indeed even a small concentration of high molecular weight polymers in water gives rise to numerous spectacular phenomena like drag reduction in turbulent pipe flow, die swell, bouncing balls; see Boger and Walters (1992) for a range of unusual effects observed with such systems. Indeed, none of these effects can be explained even qualitatively using the postulate of Newtonian fluid behavior. While detailed listings of systems displaying the so-called non-Newtonian or complex rheological behavior are available in several books (Bird et al., 1987; Barnes, 1989; Darby, 1976; Chhabra and Richardson, 2008; Chhabra, 2016), typical examples include coal/water, clay/water slurries, processed foodstuffs (like jams, jellies, and marmalades), polymers and composites, shaving foams, personal care products, biological fluids (like blood, saliva, and semen), etc.

In the following sections, the definition and classification of non-Newtonian fluids and the prediction of pressure drops in both laminar and turbulent flow in ducts of various cross-sectional shapes for different classes of non-Newtonian fluids are discussed. The analogous information for boundary layer flows is provided in Section 3.2.4.

Classification of Non-Newtonian Behavior

Newtonian Fluid Behavior

It is convenient to first define the familiar Newtonian fluid behavior. For a one-dimensional steady shearing flow (with one component of velocity, V_x, which varies in the lateral direction, y), the Newtonian fluid behavior is characterized by the linear relationship between the shear stress (τ_{yx}) and the velocity gradient (also called the rate of strain or simply the shear rate, $\dot{\gamma}_{yx}$ or dV_x / dy) as

$$\tau_{yx} = \mu\dot{\gamma}_{yx} = \mu \frac{dV_x}{dy}. \tag{2.9.1}$$

The constant of proportionality, μ, variously known as the viscosity or dynamic viscosity, is a true material constant which is influenced by pressure and temperature for each substance, but it is independent of shear rate or shear stress for a Newtonian fluid. For instance, water has a viscosity of 10^{-3} Pa·s at room temperature and atmospheric pressure, whereas the corresponding value for air is of the order ~ 10^{-5} Pa·s. Other related terms are fluidity which is simply the reciprocal of viscosity, and kinematic viscosity which is defined as the viscosity divided by density; the latter is also known as the momentum diffusivity and it plays the same role in diffusive momentum transfer processes as the thermal diffusivity in heat transfer. Returning to Equation 2.9.1, at a given temperature and pressure, the flow behavior of a Newtonian fluid is thus characterized by a single flow property, μ, which is given by the slope of the shear stress–shear rate data (see Figure 2.9.1 for three different fluids). Most low molecular weight (<~10, 000 Daltons) substances such as organic liquids, molten salts and metals, gases, all conform to Equation 2.9.1 and are thus accordingly called Newtonian fluids. For most

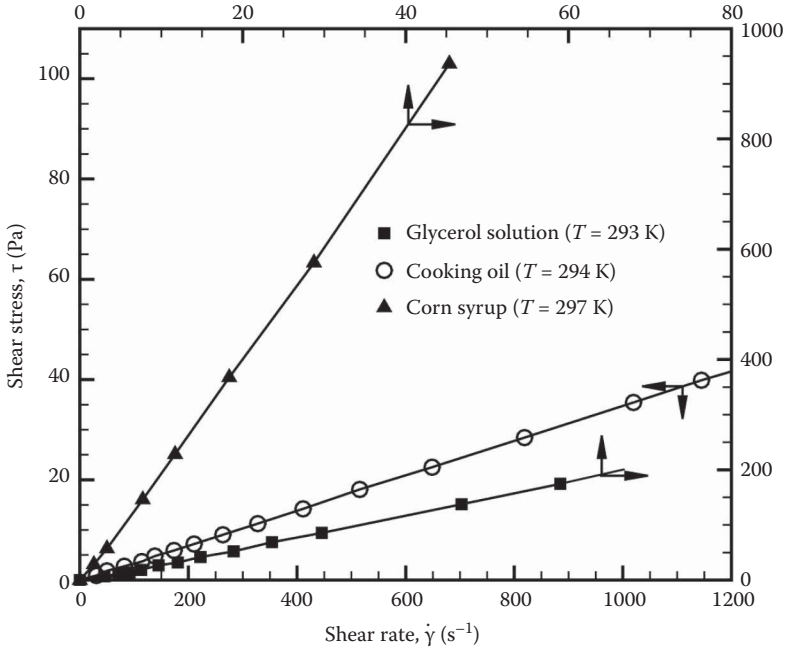


FIGURE 2.9.1 Typical shear stress–shear rate data for Newtonian fluids (glycerol solution, cooking oil, and corn syrup).

liquids, the value of viscosity decreases with temperature and increases with pressure. For gases, their viscosity increases with both temperature and pressure (Poling et al., 2000; Haynes, 2015).

It must be recognized here that Equation 2.9.1 relates to the simplest case of unidimensional and unidirectional flow situation. For the general case of 3D flows (Figure 2.9.2), it is easy to imagine that there will be two tangential stress components and one normal stress components corresponding to each coordinate direction thereby yielding a stress tensor ($\underline{\underline{S}}$) consisting of nine components altogether as follows:

$$\underline{\underline{S}} = \begin{pmatrix} -p+\tau_{xx} & \tau_{yx} & \tau_{zx} \\ \tau_{xy} & -p+\tau_{yy} & \tau_{zy} \\ \tau_{xz} & \tau_{yz} & -p+\tau_{zz} \end{pmatrix}. \tag{2.9.2}$$

In Equation 2.9.2, each row corresponds to the stress components acting in the same direction, whereas each column contains the stress components acting on the same surface.

The requirements of the tracelessness $\left(\sum(-p+\tau_{ii})=0\right)$ and of symmetry $\left(\tau_{ij}=\tau_{ji}\right)$ of the stress tensor imply that the state of stress at a point in a fluid can be described in terms of three independent shear components and two normal stress differences. For instance, in the Cartesian coordinate system, these are given as $\tau_{xy}(=\tau_{yx})$, $\tau_{yz}(=\tau_{zy})$, and $\tau_{zx}(=\tau_{xz})$, and the two normal stress differences are defined as

$$N_1 = \tau_{xx} - \tau_{yy};\ N_2 = \tau_{yy} - \tau_{zz}. \tag{2.9.3}$$

For a Newtonian substance, the tangential and normal stress components are related linearly to the components of the rate of deformation tensor via its viscosity; see Equations 2.2.32 through 2.2.37 in Chapter 2 of this handbook. For an incompressible Newtonian fluid subject to simple shearing

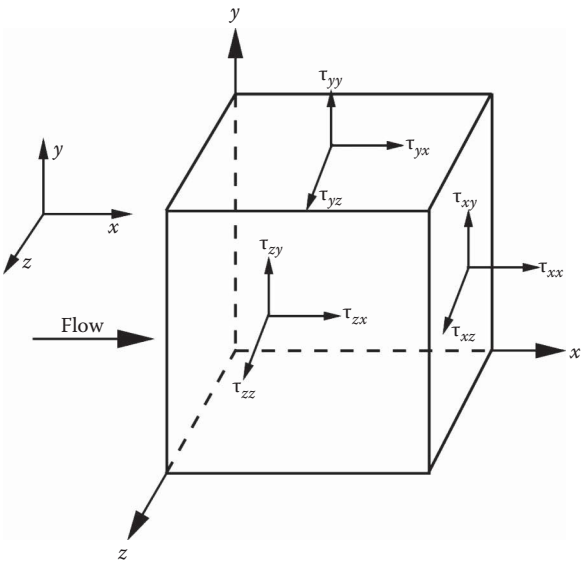


FIGURE 2.9.2 Stress components in a 3D flow.

motion, $\tau_{xx} = \tau_{yy} = \tau_{zz} = 0$; hence, the complete definition of a Newtonian fluid requires a fluid to satisfy the complete Navier–Stokes equations (see Equations 2.2.38 and 2.2.39) rather than simply exhibiting a constant viscosity in one-dimensional shear.

Non-Newtonian Fluid Behavior

In view of the preceding discussion, a non-Newtonian fluid is one which does not satisfy the Navier–Stokes equations, or in simple shear, it does not conform to Equation 2.9.1. Since the ensuing discussion is limited mainly to one-dimensional shearing, hereafter the relevant components of the stress and the rate-of-strain tensors are denoted simply as τ and $\dot{\gamma}$, respectively, without any subscripts. The simplest signature of non-Newtonian fluid behavior is when the shear stress (τ)–shear rate ($\dot{\gamma}$) data does not yield a linear relationship or does not pass through the origin or both, as shown in Figure 2.9.3. In spite of this complexity, it is still customary to define the so-called apparent viscosity, defined as, $|\tau|/|\dot{\gamma}|$, which is obviously not a constant. In fact, it may increase with shear rate (shear thickening or dilatant behavior, see Figure 2.9.4), or it may decrease with shear rate (shear thinning or pseudoplastic, see Figure 2.9.5). This is in stark contrast to the notion of constant viscosity in the context of Newtonian fluids for which the viscosity is an intrinsic property determined by the structure of the molecule. It is standard practice, though unscientific as well as arbitrary, to further categorize non-Newtonian flow behavior into three types:

1. Systems in which the value of the shear rate $\dot{\gamma}$ at a point within the fluid is determined solely by the current value of the stress at that point or vice versa. This category of fluid behavior is variously known as *purely viscous*, *generalized Newtonian fluids* (GNF), *time-independent* or *inelastic* fluids.
2. Systems in which the relationship between τ–$\dot{\gamma}$ also shows further dependence on the duration of shearing and/or the kinematic history of the sample. This type of fluid behavior is known as time-dependent. If the viscosity of a substance decreases with time at a fixed shear rate (see Figure 2.9.6 for instance), it is called thixotropy. The viscosity (or shear stress) may or may not approach an equilibrium value even after prolonged shearing. On the other hand, there are substances which exhibit the reverse trend, that is, their apparent viscosity increases with the duration of shearing (see Figure 2.9.7), though this behavior

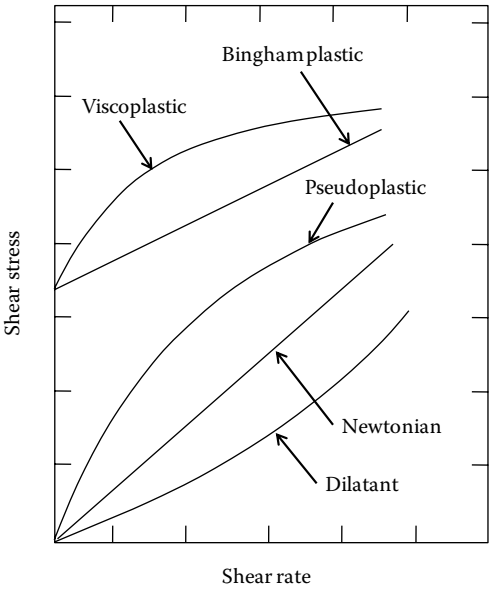


FIGURE 2.9.3 Qualitative flow curves for time-independent fluid behavior.

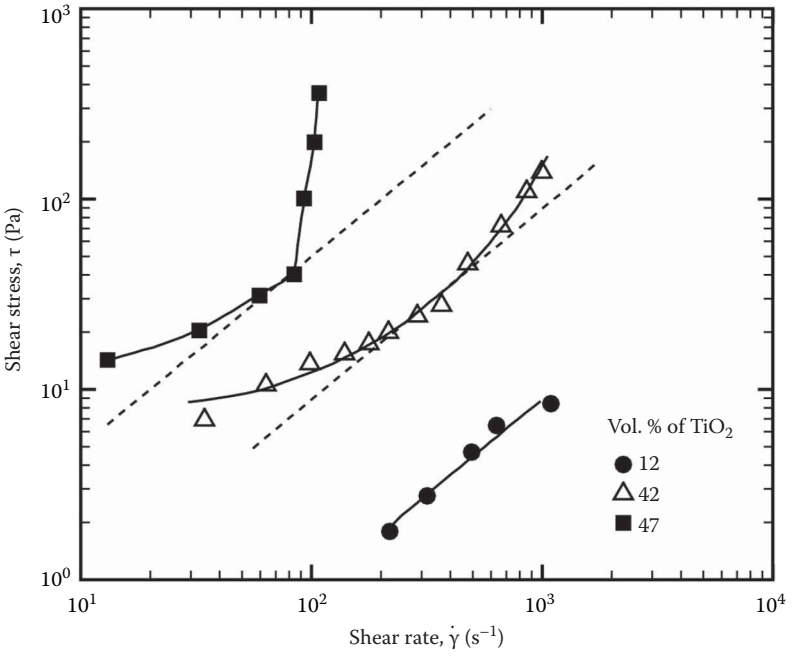


FIGURE 2.9.4 Typical shear stress–shear rate data for TiO_2 suspensions displaying shear-thickening behavior.

is less common. This phenomenon is known as anti-thixotropy, or negative thixotropy or rheopexy. More detailed discussion of this subclass of fluids and their constitutive modelling can be found in Barnes (1997), Mujumdar et al. (2002), Dullaert and Mewis (2005, 2006), and Mewis and Wagner (2012).

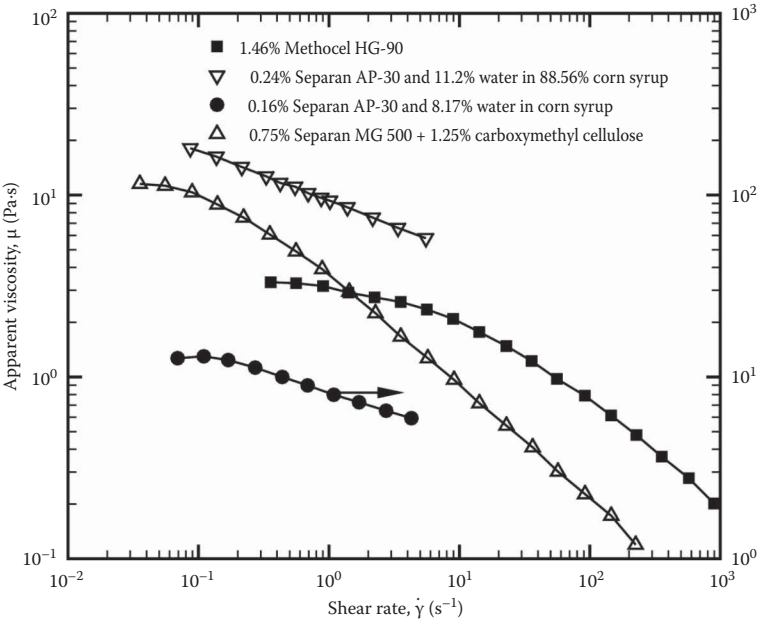


FIGURE 2.9.5 Representative plots for apparent viscosity vs. shear rate data for different pseudoplastic polymer solutions.

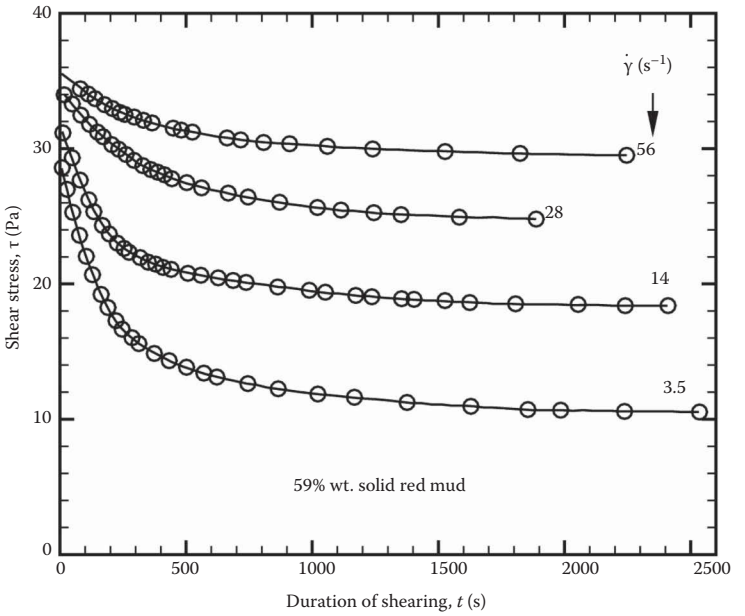


FIGURE 2.9.6 Typical experimental data showing thixotropic behavior in a red mud suspension. (From Ngugen, Q.D., Uhlherr, P.H.T., *Proceedings of 3rd National Conference Rheology*, Melbourne, Australia, 63–68, 1983. With permission.)

3. Finally, systems which display a blend of viscous fluid-like and elastic solid-like behavior. For example, these substances show partial elastic recovery recoil, creep, normal stress

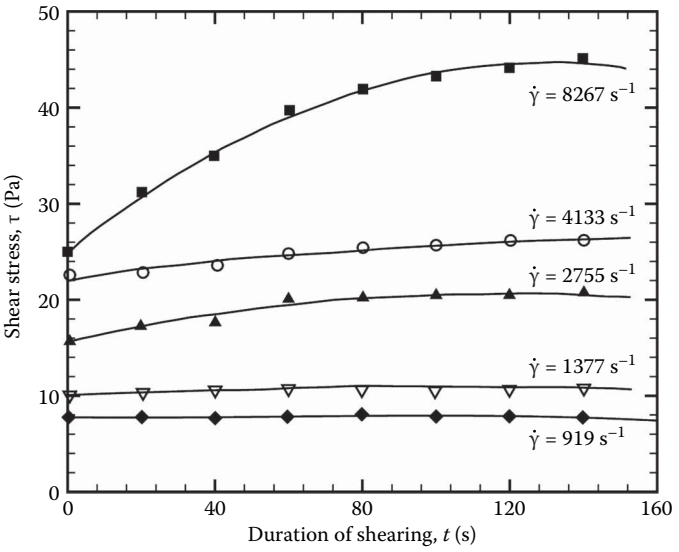


FIGURE 2.9.7 Rheopexy behavior in a saturated polyester. Replotted From Steg, I., Katz, D., J. *Appl. Polym. Sci.*, 9, 3177–3193, 1965.

effects, strain hardening, memory effects, etc. Accordingly, these are called viscoelastic or elastic-viscous. In a nutshell, the value of the shear rate ($\dot{\gamma}$) at a point in the fluid is determined by the current value of the stress as well as by the value of the stress in the past.

As noted earlier, while the foregoing classification is convenient, it is quite arbitrary because most real non-Newtonian fluids often display a combination of two or even all three types of characteristics under appropriate conditions, for example, most polymer solutions exhibit shear-dependent viscosity and viscoelastic behavior. Generally, it is, however, possible to identify the dominant non-Newtonian aspect and to use it as the basis for the process engineering design calculations. The ensuing discussion is mainly limited to the laminar and turbulent flow of time-independent fluids in ducts of circular and noncircular cross sections, followed by a short discussion about the viscoelastic effects in this flow situation. Detailed descriptions can be found in several books (Darby, 1976; Govier and Aziz, 1982; Bird et al., 1987; Chhabra and Richardson, 2008).

Types of Time-Independent Fluids and Constitutive Equations

For time-independent fluids, their steady shear behavior can evidently be described by a function, $g(\tau)$, of the form:

$$\dot{\gamma} = g(\tau). \tag{2.9.4}$$

Or, its inverse form,

$$\tau = g^{-1}(\dot{\gamma}). \tag{2.9.5}$$

Depending upon the form of these functions, three possibilities exist:

1. Shear-thinning or pseudoplastic behavior (Figure 2.9.5)
2. Shear-thickening or dilatant behavior (Figure 2.9.4)
3. Viscoplastic behavior with or without shear-thinning effects (Figure 2.9.8)

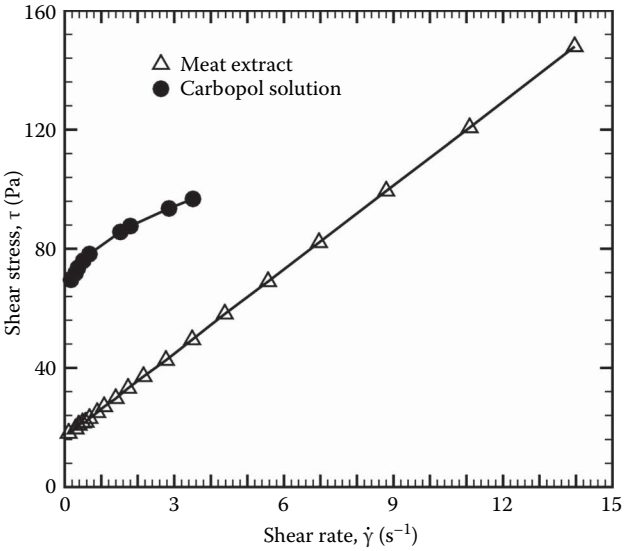


FIGURE 2.9.8 Representative shear stress–shear rate data showing viscoplastic behavior in a meat extract (Bingham plastic) and in an aqueous carbopol polymer solution (yield-pseudoplastic).

While these three possibilities are shown schematically in Figure 2.9.3, suffice it to add here that a viscoplastic substance deforms like an elastic solid for stress levels below its yield stress and it exhibits a constant (Bingham plastic) or shear-thinning (Herschel–Bulkley) viscosity for stress levels in excess of the fluid yield stress. Figure 2.9.8 shows viscoplastic behavior in a meat product and in an aqueous polymer solution. Though much confusion exists in the literature whether a true yield stress exists or not and/or about its measurement, the notion of an apparent yield stress has proved to be of value in approximating the flow behavior of many industrially important fluids (Barnes, 1999).

Rheological Equation of State

A constitutive equation (or the rheological equation of state) describes the response of a material to externally applied stress. Thus, Hooke's law of elasticity is a constitutive relation for an elastic solid, as is the ideal gas law for gases, or Equation 2.9.1 for a Newtonian fluid in unidirectional shearing flow. The literature abounds with many such relations for time-independent fluids entailing 2–6 (or even more) material constants.

From an engineering application standpoint, simpler (with fewer constants) constitutive relations are normally satisfactory as far as the prediction of pressure drop and pumping power for flow in pipes and conduits is concerned. In this context, four of the most popular models are considered here, namely, the power-law (or Ostwald de Waele) model and modifications thereof, and the Herschel–Bulkley model.

Power-Law Model

In simple shearing motion, the power-law fluid model is written as

$$\tau = K(\dot{\gamma})^n \tag{2.9.6}$$

or in terms of the apparent viscosity as

$$\mu = \frac{\tau}{\dot{\gamma}} = K(\dot{\gamma})^{n-1}, \tag{2.9.7}$$

where K and n are known as the power-law consistency coefficient (Pa·s^n) and the power-law behavior index, respectively. Naturally, $n = 1$ corresponds to Newtonian fluid behavior (Equation 2.9.1). For $n > 1$, Equation 2.9.7 predicts shear-thickening or dilatant behavior characterized by an increasing apparent viscosity with shear rate, as shown in Figure 2.9.4. Barnes (1989) has provided a thorough review of shear-thickening behavior in suspensions. For $n < 1$, Equation 2.9.6 predicts shear-thinning behavior, that is, decreasing apparent viscosity with shear rate. This is possibly the commonest type of time-independent fluid behavior exhibited by most polymeric fluids, suspensions and dispersions, as seen in Figure 2.9.5 (Chhabra and Richardson, 2008). Indeed, it is not uncommon to have the value of the power-law behavior index as small as $n = 0.15-0.2$ for some clay suspensions, for instance. At vanishingly small values of shear rate ($\dot{\gamma} \to 0.$), many polymer systems (solutions and melts) exhibit a constant viscosity (the so-called zero shear viscosity, μ_0) and clearly, the power-law model, given by Equation 2.9.6 or 2.9.7, leads to infinite viscosity in this limit for $n < 1$. For example, if one considers the flow of a pseudoplastic fluid ($n < 1$) through a circular duct, because of the symmetry, the shear rate (velocity gradient) becomes zero, and thus the apparent viscosity from Equation 2.9.6 becomes infinite, at the center of the pipe. This poses conceptual difficulties especially when performing numerical analyses on such systems. In this case, the power-law equation is not applicable and a more general constitutive equation is needed.

Modified Power-Law Constitutive Equation

Dunleavy and Middleman (1966) proposed a generalization of the power-law equation, which extends the shear rate range spanning the zero shear viscosity and onset of shear-thinning regimes, and is given by

$$\mu = \frac{\mu_0}{1 + \dfrac{\mu_0}{K}\dot{\gamma}^{1-n}}. \tag{2.9.8}$$

Examination of Equation 2.9.8 reveals that at low shear rates, the second term in the denominator becomes small compared with unity $\left(\text{i.e., } \mu_0 / K\dot{\gamma}^{n-1} << 1\right)$ and the apparent viscosity becomes constant equal to μ_0. This represents the Newtonian plateau at vanishingly low shear rates, as seen in Figure 2.9.5. On the other hand, as the second term in the denominator becomes large compared with unity, Equation 2.9.8 reduces to Equation 2.9.7 and represents the shear-thinning region or the power-law region.

An important advantage of the modified power-law equation is that it retains the rheological properties K and n of the power-law model plus the additional property μ_0. Thus, as will be shown later, in the flow and heat-transfer equations, the same dimensionless groups as in the power-law model will appear plus an additional dimensionless parameter which describes in which region (power-law or zero shear or both) a particular system is operating. Also, solutions using the modified power-law model will have Newtonian and power-law solutions as the limiting conditions.

Equation 2.9.8 describes the flow curve for a pseudoplastic fluid ($n < 1$). For a dilatant fluid, ($n > 1$), an appropriate modified power-law model is given by

$$\mu = \mu_0 \left[1 + \frac{K}{\mu_0}\dot{\gamma}^{n-1}\right] \tag{2.9.9}$$

Extended Modified Power-Law (EMPL) Constitutive Equation

For some non-Newtonian fluids, once the apparent viscosity transitions from Newtonian to power-law behavior, it follows the power-law model indefinitely as the shear rate increases. For such fluids, the modified power-law equation is appropriate. Other fluids at sufficiently high shear rates transition smoothly back to Newtonian behavior with a (constant) apparent viscosity, μ_∞, that is less than μ_0 for pseudoplastic fluids, and greater than μ_0 for dilatant fluids. Thus, constitutive equations for

such fluids necessarily require more parameters than μ_0, K, and n. A model that spans from the zero to the high shear rate Newtonian regions and that contains only one additional parameter, μ_∞, is the extended modified power-law constitutive equation. For pseudoplastic fluids, it is given by

$$\mu = \frac{\mu_0 - \mu_\infty}{1 + \frac{\mu_0}{K}\dot{\gamma}^{1-n}} + \mu_\infty . \tag{2.9.10}$$

This equation may be recognized as the Cross (1965) model except with the Cross parameters explicitly expressed in terms of K and n. An analogous equation for dilatant fluids is given by

$$\mu = \frac{\mu_\infty - \mu_0}{1 + \frac{\mu_\infty}{K}\dot{\gamma}^{1-n}} + \mu_0 . \tag{2.9.11}$$

It may be shown that from the zero shear rate Newtonian region through the power-law region, these equations follow the corresponding modified power-law models if μ_∞ differs sufficiently from μ_0 by four orders of magnitude or more. Otherwise, for more modest differences, the low and high shear rate transition regions merge to establish a smooth flow curve with a power-law region whose slope corresponds to an effective n that is closer to 1 than the physical value of n (Figure 2.9.9). The result is that flows of such fluids fail to reach power-law solutions even when all of the shear rates are in the power-law region.

Finally, it should be emphasized that when using analytical or numerical solutions, care needs to be exercised to ensure that the rheology model used in the solution matches the behavior of the fluid at hand. This is particularly true for dilatant fluids as their rheology may be more complicated

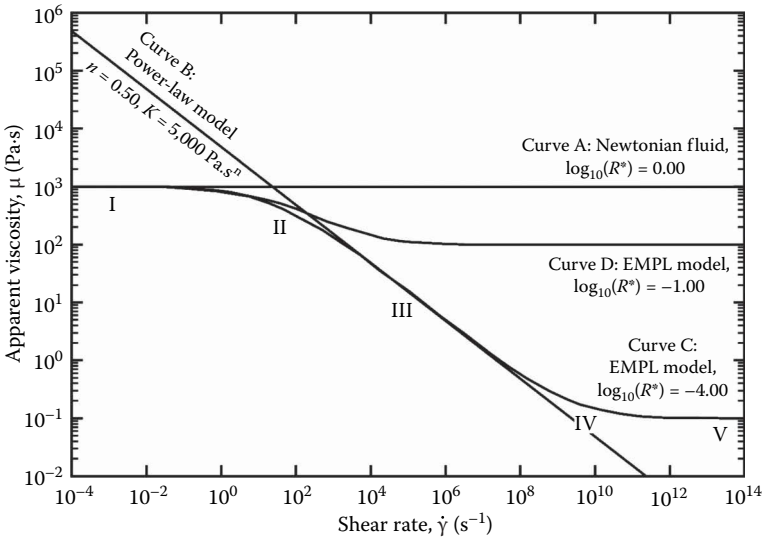


FIGURE 2.9.9 Flow curves of four related fluids. Curve A: a Newtonian fluid ($n=1$, $\log_{10} R^* = \log_{10}(\mu_\infty/\mu_0) = 0$) with a viscosity of 10^3 Pa·s. Curve B: a power-law model ($n=0.5$, $K = 5 \times 10^3$ Pa·s^n). Curve C: a strongly non-Newtonian pseudoplastic fluid ($n=0.5$, $K = 5 \times 10^3$ Pa·s^n, $\log_{10} R^* = -4.0$ and $\mu_0 = 10^3$ Pa·s) plotted using EMPL model, Equation 2.9.10. Also shown are the shear rate regimes: I-zero shear rate Newtonian region; II-low shear rate transition region; III-power-law region; IV-high shear rate transition region; V-high shear rate Newtonian region. Curve D: a weakly non-Newtonian pseudoplastic fluid ($n=0.5$, $K = 5 \times 10^3$ Pa·s^n, $\log_{10} R^* = -1.0$ and $\mu_0 = 10^3$ Pa·s). (From Capobianchi, M., Aziz, A., *ASME J. Heat Transfer*, 134, 122502, 2012. With permission.)

than what can be modeled by the above equations. For example, at low shear rates, dilatant fluids may first shear thin similar to a pseudoplastic fluid, then shear thicken as the shear rate increases and, after plateauing at a maximum apparent viscosity, begin to shear thin again as the shear rate increases further. Analytical and numerical solutions that use the above dilatant fluid equations would then be inapplicable or at best be only an approximation.

Viscoplastic Models

When a fluid exhibits both shear-thinning and yield stress characteristics, it is customary to use the Herschel–Bulkley viscosity model as

$$\tau = \tau_0 + K(\dot{\gamma})^n \quad |\tau| > |\tau_0|, \tag{2.9.12a}$$

$$\dot{\gamma} = 0 \quad |\tau| \le |\tau_0|, \tag{2.9.12b}$$

or, in terms of the apparent viscosity,

$$\mu = \frac{\tau}{\dot{\gamma}} = \frac{\tau_0}{\dot{\gamma}} + K(\dot{\gamma})^{n-1} \quad |\tau| > |\tau_0|, \tag{2.9.13a}$$

$$\mu \to \infty \quad |\tau| \le |\tau_0|. \tag{2.9.13b}$$

Though the Herschel–Bulkley fluid model, Equations 2.9.12 and 2.9.13, is a three parameter model similar to the modified power-law model, it contains several other viscosity equations as limiting cases. Firstly, for $n = 1$, it reduces to the familiar Bingham fluid model. For $\tau_0 = 0$, it reverts to the power-law model, Equation 2.9.7, and finally, for $n = 1$ and $\tau_0 = 0$, Newtonian fluid behavior is recovered.

These four models will now be used to study fully developed laminar and turbulent flow in circular and noncircular ducts in the following sections, especially with the objective of predicting the pressure drop in a given application.

Fully Developed Laminar Flow in a Tube

From a theoretical standpoint, one can combine Cauchy's momentum equation with a suitable constitutive equation for the extra stress tensor, τ, such as that given by Equation 2.9.6 or 2.9.12a, followed by an integration (and using the no-slip condition at the wall) to obtain the detailed velocity distribution across the pipe cross section. The resulting velocity distribution can be used to obtain expressions for the mean and maximum velocity, plug radius and velocity, etc., for a specific fluid model. A summary of the resulting expressions for a range of viscosity models is available in Skelland (1967), Govier and Aziz (1982), and Chhabra and Richardson (2008).

Herschel–Bulkley Fluids

It is useful to begin with the case of the Herschel–Bulkley fluid model, for it degenerates into various models under appropriate conditions. The volumetric flow rate, Q, for the Herschel–Bulkley model fluid is given by

$$Q = \pi R^3 n \left(\frac{\tau_w}{K}\right)^{1/n} (1-\phi)^{\frac{n+1}{n}} \left\{\frac{(1-\phi)^2}{(3n+1)} + \frac{2\phi(1-\phi)}{2n+1} + \frac{\phi^2}{n+1}\right\}, \tag{2.9.14}$$

where $\phi = (\tau_0 / \tau_w) < 1$ and $\tau_w = (\Delta p/L)(D/4)$ is the shear stress at the wall of the pipe. Evidently, for a given value of the pressure gradient $(\Delta p/L)$, one can readily calculate the corresponding flow

rate (Q). However, the reverse calculation (i.e., for a known value of Q) of ($\Delta p/L$) necessitates an iterative solution to Equation 2.9.14.

As noted above, Equation 2.9.14 can be reduced to the corresponding expression for power-law fluids by substituting $\phi = 0$ (i.e., $\tau_0 = 0$):

$$Q = \frac{\pi R^3 n}{(3n+1)}\left(\tau_w / K\right)^{1/n}. \tag{2.9.15}$$

It is customary to introduce the familiar Fanning friction factor, f as

$$f = \frac{\tau_w}{\left(1/2\right)\rho V^2} = \frac{\left(D\Delta p / 4L\right)}{\left(1/2\right)\rho V^2}. \tag{2.9.16}$$

Now, Equation 2.9.15 can be recast as

$$f = \frac{16}{Re_{PL}}, \tag{2.9.17}$$

where the corresponding Reynolds number, Re_{PL}, is defined as

$$Re_{PL} = \frac{\rho V^{2-n} D^n}{8^{n-1} K\left(\frac{3n+1}{4n}\right)^n}. \tag{2.9.18}$$

Equation 2.9.18 thus ensures that the friction factor values for power-law fluids coincide with the familiar expression of $f = (16/Re)$ for Newtonian fluids. In contrast, it is very cumbersome to organize Equation 2.9.14 in this form. However, for the case of the Bingham fluid ($n = 1$), Equation 2.9.14 reduces to

$$Q = \frac{\pi R^4}{8K}\left(\frac{-\Delta p}{L}\right)\left(1 - \frac{4}{3}\phi + \frac{1}{3}\phi^4\right). \tag{2.9.19}$$

This can be rearranged in terms of the friction factor f as

$$f = \frac{16}{Re_B}\left[1 + \frac{1}{6}\frac{He}{Re_B} - \frac{1}{3}\frac{He^4}{f^3 Re_B^7}\right], \tag{2.9.20}$$

where the Reynolds number, Re_B, and Hedstrom number, $He = Bi \cdot Re_B$, respectively, are defined as

$$Re_B = \frac{\rho VD}{K};\ He = \frac{\rho D^2 \tau_0}{K^2}. \tag{2.9.21}$$

A more convenient form of Equation 2.9.18 is given by

$$f \cdot Re_B = 16\left[1 + \frac{Bi}{6} - \frac{1}{3}\frac{Bi^4}{(fRe_B)^3}\right]. \tag{2.9.22}$$

Equation 2.9.22 suggests the product $f \cdot Re_B$ to be a function of the Bingham number, $Bi = \tau_0 D / KV$, only. Figure 2.9.10 shows a plot of Equation 2.9.22 where it is clearly seen that the value of $(f \cdot Re_B)$ veers away from the Newtonian value of 16 only for $Bi > 0.1$ or so. Evidently, the utility of Equation

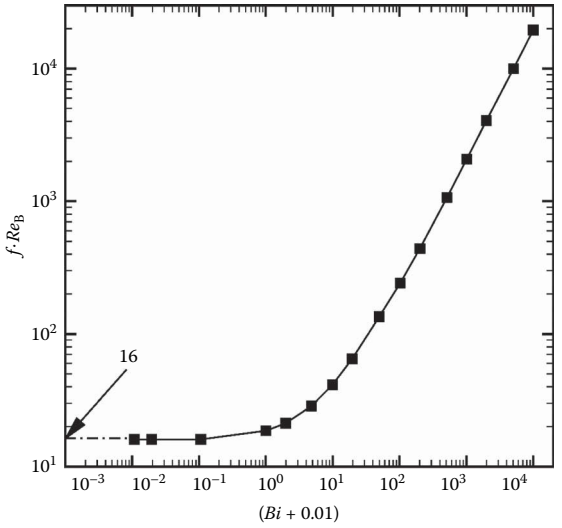


FIGURE 2.9.10 Effect of Bingham number on the value of $(f \cdot Re_B)$ as predicted by Equation 2.9.22.

(2.9.22) in practice is severely hampered due to the occurrence of the fourth order term. Ihle and Tamburrino (2012) have introduced the following two parameters to obtain appropriate explicit forms of Equation 2.9.22 as follows:

$$\alpha_1 = \frac{Bi}{Re_B};\ \alpha_2 = \frac{2}{Re_B}\left(\frac{Bi}{3} + 2\right). \tag{2.9.23}$$

The friction factor, f, is now given by

$$f = \alpha_2 + (\alpha_2^2 + \alpha_3)^{1/2} + \left[2\alpha_2^2\left\{1 + \left(1 + \frac{\alpha_3}{\alpha_2^2}\right)^{-1/2}\right\} - \alpha_3\right]^{1/2}, \tag{2.9.24}$$

where $\alpha_3 = \varepsilon_1 + \varepsilon_2$ when

$$\varepsilon_i = \frac{2}{3}\alpha_1^{4/3}\left\{\left(\frac{3}{2}\alpha_2\right)^2 + (-1)^i\left[\left(\frac{3}{2}\alpha_2\right)^4 - \alpha_1^4\right]^{1/2}\right\}^{1/3}. \tag{2.9.25}$$

Hence, ε_1 and ε_2 can be evaluated using Equation 2.9.25 by using $i=1$ and $i=2$, respectively. It needs to be emphasized here that for $\alpha_1=0$ (i.e., $Bi=0$), Equation 2.9.24 coincides with the expected limit of $f=16/\mathrm{Re}$ (Newtonian fluids). For a fixed value of the Reynolds number, Ihle and Tamburrino (2012) expanded Equation 2.9.24 around $Bi=0$ as

$$f \approx 4\alpha_2 - \frac{1}{12}\frac{\alpha_1^4}{\alpha_2^3} - \frac{1}{192}\frac{\alpha_1^8}{\alpha_2^7} - \cdots \tag{2.9.26}$$

Of course, Equation 2.9.26 is likely to work well for small values of Bingham number, but the error is only 10% at $Bi=10^4$. The main virtue of Equation 2.9.26 lies in its simplicity and the fact that one can calculate the value of the friction factor for given values of the Reynolds number and Bingham number without resorting to a trial-and-error procedure.

The corresponding development for the cases of the modified and extended modified power-law fluid models given by Equations 2.9.8 and 2.9.10 are presented next.

Modified Power-Law Model Fluids

For laminar fully developed flow of a modified power-law fluid in a circular duct, the product of the friction factor and a certain Reynolds number is a constant depending on the value of the flow index, n, and the shear rate parameter, β defined below, that is,

$$f \cdot Re_{\mathrm{m}} = \text{ constant } (n,\beta), \tag{2.9.27}$$

where f is the Fanning friction factor given by Equation 2.9.16, and Re_{m} is the modified power-law Reynolds number, that is,

$$Re_{\mathrm{m}} = \frac{\rho VD}{\mu^*}, \tag{2.9.28}$$

$$\mu^* = \frac{\mu_0}{1+\beta}, \tag{2.9.29}$$

$$\beta = \frac{\mu_0}{K}\left(\frac{V}{D}\right)^{1-n}. \tag{2.9.30}$$

β is the shear rate parameter mentioned previously which can be calculated by the designer for a given application (known values of V and D) and a certain pseudoplastic fluid (μ_0, K, n). The solution for a circular tube has been calculated by Brewster and Irvine (1987) and the results are shown in Figure 2.9.11 and in Table 2.9.1. Referring to Figure 2.9.11, we can see that when the values of $\log_{10}\beta$ is less than approximately –2, the duct is operating in the Newtonian ($\mu^* = \mu_0$) region. Note that in the Newtonian region, $\mathrm{Re_m}$ also reverts to the Newtonian Reynolds number.

When the value of $\log_{10}\beta$ is approximately in the range $-2 \le \log_{10}\beta \le 2$, the duct is operating in the transition region, that is, the Newtonian and shear-thinning region, and the values of $f \cdot Re_{\mathrm{m}}$ must be obtained from Figure 2.9.11 or from Table 2.9.1.

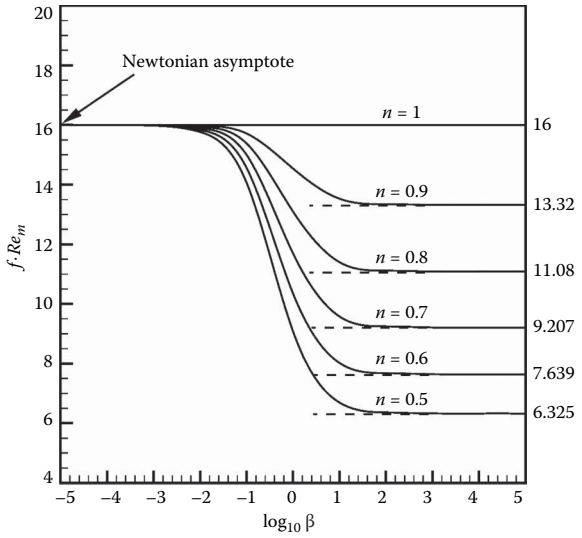


FIGURE 2.9.11 Product of friction factor and modified Reynolds number vs. $\log_{10}\beta$ for a circular duct. (Modified from Brewster, A.A. and Irvine, T.F., *Wärme und Stoffübertragung*, 21, 83–86, 1987.)

TABLE 2.9.1
Summary of Computed Values of $f \cdot Re_m$ for Various Values of n and β for a Circular Duct

	Value of $f \cdot Re_m$					
β	**n=1**	**n=0.9**	**n=0.8**	**n=0.7**	**n=0.6**	**n=0.5**
10^{-5}	16	16	16	16	16	16
10^{-4}	16	16	15.999	15.999	15.998	15.998
10^{-3}	16	15.997	15.993	15.988	15.983	15.976
10^{-2}	16	15.968	15.93	15.884	15.830	15.764
10^{-1}	16	15.713	15.380	14.997	14.559	14.061
10^{0}	16	14.538	13.094	11.690	10.346	9.0748
10^{1}	16	13.527	11.399	9.577	8.0205	6.6928
10^{2}	16	13.343	11.115	9.2463	7.6790	6.3628
10^{3}	16	13.323	11.084	9.2113	7.6433	6.3285
10^{4}	16	13.321	11.081	9.2078	7.6398	6.3250
10^{5}	16	13.321	11.081	9.2075	7.6393	6.3248
Exact solution	16	13.321	11.081	9.2073	7.6393	6.3245

Source: Brewster, A.A. and Irvine, T.F., *Wärme und Stoffübertragung*, 21, 83–86, 1987.

For $\log_{10} \beta > \sim 2$, the duct is operating completely in the power-law region and in this case the power-law friction factor–Reynolds number relations can be used, given by Equations 2.9.17 and 2.9.18. In this region, Re_m becomes the power-law Reynolds number given by

$$Re_g = \frac{\rho V^{2-n} D^n}{K}. \tag{2.9.31}$$

For convenience, Brewster and Irvine (1987) have presented the following correlation which agrees within 0.1% with the results tabulated in Table 2.9.1:

$$f \cdot Re_m = \frac{4(1+\beta)}{\frac{1}{64} + \frac{\beta}{2^{2n+3}\left(\frac{3n+1}{4n}\right)^n}}. \tag{2.9.32}$$

Thus, Equation 2.9.32 contains all of the information required to calculate the circular tube laminar fully developed pressure drop for a pseudoplastic fluid depending upon the shear rate region(s) under consideration. Note that in scaling such non-Newtonian systems, both Re_m and β must be held constant. Modified power-law solutions have been reported for two other duct shapes. Park et al. (1993) have presented the friction factor–Reynolds number relations for rectangular ducts and Capobianchi and Irvine (1992) for concentric annular ducts.

Extended Modified Power-Law Model Fluids

Solutions for the friction factor-modified Reynolds number product using the extended modified power-law models are generally functions of three parameters: the flow index, n, a shear rate parameter, β^*, and the ratio of the limiting Newtonian viscosities, R^*, defined by

$$\beta^* = \frac{1+\beta}{1+\beta_\infty}; \; R^* = \frac{\beta_\infty}{\beta} = \frac{\mu_\infty}{\mu_0}, \tag{2.9.33}$$

Here, β is defined by Equation 2.9.30 above, and β_∞ is defined similarly except with μ_0 replaced by μ_∞. For solutions that use the extended modified power-law equations, β^* serves the analogous role to that of β above. Namely, β^* determines the shear rate region where the system is operating. β^* may be shown to range from 1 to $1/R^*$ for pseudoplastic fluids, and from $1/R^*$ to 1 for dilatant fluids. At the left endpoint of these ranges, the system is operating in the low shear rate Newtonian region, whereas it is in the high shear rate Newtonian region at the right endpoint. Hence, solutions at the endpoints approach Newtonian values. At the midpoint of the $\log_{10}(\beta^*)$ range, the system may be shown to be operating in the power-law region. Values then approach power-law values, but only if, in addition, μ_∞ differs from μ_0 by approximately four orders of magnitude or more. Otherwise, even when operating in the power-law region, values for fluids with smaller differences in their limiting viscosities fall between the power-law and Newtonian values, approaching the latter as $R^* \to 1$. This behavior is a direct consequence of the effective n discussed above departing further from the physical value of n and approaching 1 as $\mu_\infty \to \mu_0$ (i.e., $R^* \to 1$). Finally, when β^* is to the left or right of the midpoint of the $\log_{10}(\beta^*)$ range (excluding the endpoints), the system is operating in the low or high shear rate transition region, respectively. Results are then again between the Newtonian and power-law asymptotes. This behavior may be clearly seen in the square duct solutions shown in Figure 2.9.12 from Capobianchi (2008).

Solutions for the Darcy friction factor-modified Reynolds number product for laminar, hydrodynamically fully developed flows within ducts using the extended modified power-law models have appeared in the literature. Capobianchi (2008) solved the rectangular channel problem, including the parallel-plate duct, and provided correlation equations for the results. Similarly, Brewster (2013) solved the circular tube problem and also provided correlation equations. That study used the nondimensionalization scheme described above but defined an additional parameter, a dimensionless

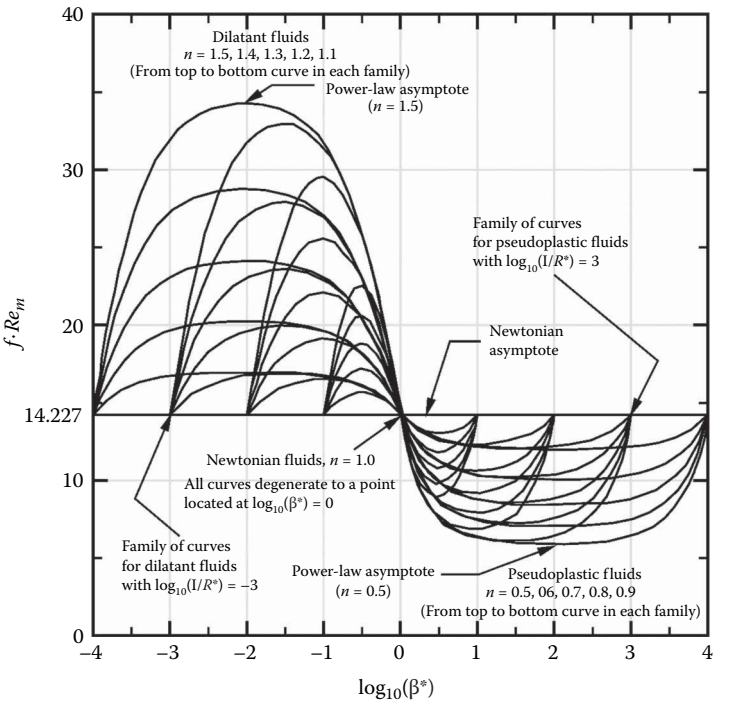


FIGURE 2.9.12 Fanning friction factor-modified Reynolds number product for laminar flow in the hydrodynamically fully developed region of a square duct with selected features indicated. (Modified from Capobianchi, M., *Int. J. Heat Mass Transfer*, 51, 1393–1401, 2008. With permission.)

reference viscosity, that is a function of only β^* and R^*, and whose range is 0–1. Results were displayed as a function of this parameter rather than β^* directly. Details may be found in the aforementioned references. Additional studies using the extended modified power-law models that include heat transfer and buoyancy-induced flows may be found in Section 3.2.4.

At this juncture, it is worthwhile to make two observations: the foregoing approaches are based on the choice of a specific viscosity model (power-law, or Bingham plastic, or Herschel–Bulkley, etc.). The material parameters appearing in these equations (n, K, τ_0, μ_0, μ_∞ or others) must be evaluated from the shear stress–shear rate data obtained using rheometers like a Couette cell, cone and plate, parallel plate or a capillary viscometer, for instance. The measurement of such data is detailed in several excellent books on rheometry (Whorlow, 1992; Macosko, 1994; Coussot, 2005). Bearing in mind that most non-Newtonian substances are characterized by the presence of transient structures (bubbles, droplets, flocs, micelles, aggregates and clusters of fine particles, entangled polymer molecules, for instance), such measurements are far from being straight forward. Some useful tips in this regard can be found in Coussot (2005) and Chhabra and Richardson (2008). Secondly, the preceding discussion necessitates *a priori* selection of a specific viscosity model. It is, however, possible to defer this choice to a later stage by developing a generalized approach for the laminar flow of time-independent fluids in a tube.

In the next section, we outline a generalized approach to estimate the pressure gradient in a pipe for the laminar flow of purely viscous fluids which will be subsequently extended to the case of noncircular ducts.

Generalized Approach for Laminar Flow of Purely Viscous Fluids

For the fully developed, steady and incompressible flow of a purely viscous fluid, the shear stress τ varies linearly with the radial position from being zero at the center of the pipe to its maximum value τ_w at the wall of the pipe. This combined with the general form of Equation 2.9.4 is possible to derive the following relationship (Chhabra and Richardson, 2008):

$$Q = \frac{\pi R^3}{\tau_w^3} \int_0^{\tau_w} \tau^2 g(\tau)\, d\tau. \tag{2.9.34}$$

Without assuming a specific form of the viscosity model, $g(\tau)$, one can deduce the following expression for the true shear rate at the tube wall:

$$\dot{\gamma}_w = \left(\frac{8V}{D}\right)\left(\frac{3n'+1}{4n'}\right), \tag{2.9.35}$$

where

$$n' = \frac{\log \tau_w}{\log\left(\dfrac{8V}{D}\right)}.$$

One can readily evaluate the value of n' by using the pressure gradient ($\tau_w = (D/4)(\Delta P / L)$)-volumetric flow rate ($8V / D$) information obtained using small diameter tubes. Naturally, $n' = 1$ denotes the case of Newtonian fluid behavior, $n' < 1$ corresponds to shear thinning, and finally, $n' > 1$ pertains to shear thickening or dilatant behavior. In the absence of end and slip effects, such a curve ($\log \tau_w$ vs. $\log 8V/D$) is unique irrespective of the pipe diameter as long as the laminar flow regime occurs. Over the range of shear rates over which n' is approximately constant, one can write a power-law type equation for this segment of the flow curve as

$$\tau_w = K' \left(\frac{8V}{D}\right)^{n'}. \tag{2.9.36}$$

Combining Equation 2.9.36 with the definition of the fanning friction factor, f, one obtains

$$f = \frac{16}{Re_{\text{gen}}}, \tag{2.9.37}$$

where the generalized Reynolds number, Re_{gen}, (also known as the Metzner–Reed Reynolds number) is defined as

$$Re_{\text{gen}} = \frac{\rho D^{n'} V^{2-n'}}{8^{n'-1} K'} = \frac{\rho VD}{\mu_{\text{eff}}}. \tag{2.9.38}$$

And, in turn, the effective viscosity, μ_{eff}, is given by

$$\mu_{\text{eff}} = K'\left(\frac{8V}{D}\right)^{n'-1}. \tag{2.9.39}$$

The similarity between Equations 2.9.39 and 2.9.7 is striking, except for the fact that n and K are true power-law constants whereas n' and K' appearing in Equation 2.9.36 are apparent constants because $\left(8V/D\right)$ is only the nominal (exact for Newtonian fluids) shear rate at the pipe wall. We conclude this section by noting the following relationships between n' and K' and the true rheological constants (Chhabra and Richardson, 2008).

For power-law model:

$$n = n';\ K' = K\left(\frac{3n+1}{4n}\right)^n. \tag{2.9.40}$$

For Bingham plastic model:

$$K' = \tau_w\left[\frac{K}{\tau_w\left(1-\frac{4}{3}\phi+\frac{\phi^4}{3}\right)}\right]^{n'}. \tag{2.9.41}$$

Similar correspondence can be established for any other viscosity model following the procedure outlined by Steffe (1996) and Chhabra and Richardson (2008).

Laminar Flow in Noncircular Ducts

Tubes of noncircular cross sections are finding increasing applications in the novel design of heat exchangers, heat sources and sinks used in electronic component cooling applications. Notwithstanding the significance of the detailed velocity profiles, it is readily conceded that the prediction of the pressure drop to sustain a given flow rate of a fluid with known rheology is the key parameter in such applications. While it is possible to solve the momentum equations to obtain the velocity profiles, secondary flows, etc., much effort has been expended in developing reliable schemes for the prediction of pressure drop in noncircular ducts. Perhaps the earliest attempt is that of Kozicki et al. (1967) who introduced two geometric parameters (a, b) characterizing the shape of the noncircular duct whose values can be evaluated using experimental throughput-pressure drop data or from analysis for a Newtonian fluid in the same duct. The values of a and b are assumed to be independent of the fluid rheology and extensive comparisons between the predictions and observations in wide-ranging shapes of duct (square, rectangular, triangular, elliptical, semicircular, for instance) indicate this to be a reasonable approximation, at least for purely viscous fluids (Chhabra

and Richardson, 2008). The values of a and b combined with the introduction of the familiar hydraulic diameter, D_h, allow the generalized Reynolds number for power-law fluids to be recast as

$$Re_g = \frac{\rho V^{2-n} D_h^n}{8^{n-1} K \left(b + \frac{a}{n} \right)^n}, \tag{2.9.42}$$

and in the laminar flow region, the Fanning friction factor, $f = 16/Re_{gen}$ applies. The values of the geometric parameters (a, b) for a range of shapes are listed in Table 2.9.2. The validity of this approach has been demonstrated extensively, for example, see Chhabra and Richardson (2008). Some further attempts have been made to improve upon the initial proposition of Kozicki et al. (1967). For instance, Delplace and Leuliet (1995) demonstrated that the two parameters a and b can be effectively reduced to one parameter, β_0, defined as follows:

$$\beta_0 = \frac{48}{f \cdot Re}, \tag{2.9.43}$$

$$a = \frac{1}{1+\beta_0}; \; b = \frac{\beta_0}{1+\beta_0}. \tag{2.9.44}$$

The value of β_0 can be estimated simply from the knowledge of the so-called Poiseuille number, that is, $f \cdot Re$ for Newtonian fluids. However, this approach presumes that the values of a and b always add up to 1 which results in some loss of accuracy. It is also worthwhile to add here that some researchers (Muzychka and Edge, 2008) have argued that the use of $\sqrt{A}$, where A is the flow area, as the characteristic linear dimension together with an aspect ratio as opposed to the use of the hydraulic diameter in conjunction with geometric parameters (a, b) leads to equally good predictions of the friction factor for power-law fluids in noncircular ducts.

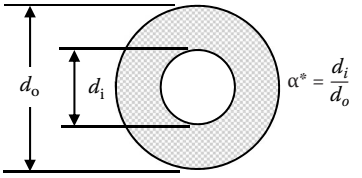
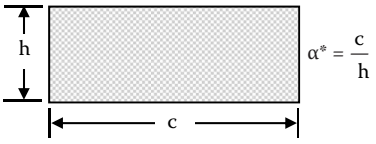
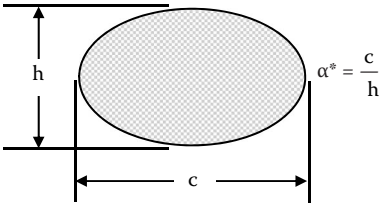
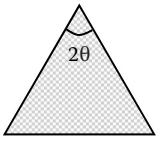
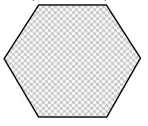
Turbulent Flow in Circular and Noncircular Ducts

The prediction of pressure gradient in the turbulent flow of even purely viscous fluids is far more difficult than that for Newtonian fluids. The main difficulty stems from the fact that while $\tau_w = (D/4)(\Delta p / L)$ still gives the true value of the wall shear stress, the corresponding shear rate cannot be evaluated. In a pioneering study, Dodge and Metzner (1959) combined dimensional considerations with experimental data for a series of non-Newtonian fluids (purely viscous) to put forward the following correlation for friction factor for smooth circular pipes:

$$\frac{1}{\sqrt{f}} = \frac{4}{(n')^{3/4}} \log \left[Re_{gen} f^{(2-n')/2} \right] - \frac{0.4}{(n')^{1.2}}. \tag{2.9.45}$$

Equation 2.9.45 is plotted in Figure 2.9.13. Three observations about Equation 2.9.45 are in order at this juncture. First, Equation 2.9.45 reduces to the familiar Nikuradse equation for Newtonian fluids, that is, $n' = 1$. Second, the limited numerical results on the turbulent flow of power-law fluids in circular tubes are in line with these predictions thereby lending support to the validity of Equation 2.9.42 (Rudman et al., 2004; Gnambode et al., 2015). Finally, attention is drawn to the fact that the values of n' and K' appearing in the definition of the Re_{gen} and in Equation 2.9.45 need to be evaluated from the laminar $\tau_w - (8V/D)$ data spanning up to the expected range of τ_w value in the turbulent flow conditions. This constrains the applicability of this method rather severely. Irvine (1988), on the other hand, put forward the following explicit expression for friction factor:

TABLE 2.9.2
Constants a and b for Various Duct Geometrics Used in the Method due to Kozicki et al. (1967)

Geometry	α^*	a	b
Concentric annuli 	0.1	0.4455	0.9510
	0.2	0.4693	0.9739
	0.3	0.4817	0.9847
	0.4	0.4890	0.9911
	0.5	0.4935	0.9946
	0.6	0.4965	0.9972
	0.7	0.4983	0.9987
	0.8	0.4992	0.9994
	0.9	0.4997	1.0000
	1.0[a]	0.5000	1.0000
Rectangular 	0.0	0.5000	1.0000
	0.25	0.3212	0.8482
	0.5	0.2440	0.7276
	0.75	0.2178	0.6866
	1.00	0.2121	0.8766
Elliptical 	0.00	0.3084	0.9253
	0.10	0.3018	0.9053
	0.20	0.2907	0.8720
	0.30	0.2796	0.8389
	0.40	0.2702	0.8107
	0.50	0.2629	0.7886
	0.60	0.2575	0.7725
	0.70	0.2538	0.7614
	0.80	0.2515	0.7546
	0.90	0.2504	0.7510
	1.00[b]	0.2500	0.7500
Isosceles triangular 	2θ (deg)		
	10	0.1547	0.6278
	20	0.1693	0.6332
	40	0.1840	0.6422
	60	0.1875	0.6462
	80	0.1849	0.6438
	90	0.1830	0.6395
Regular polygon (*N* sides) 	*N*		
	4	0.2121	0.6771
	5	0.2245	0.6966
	6	0.2316	0.7092
	8	0.2391	0.7241

[a] Parallel plates.
[b] Circular.

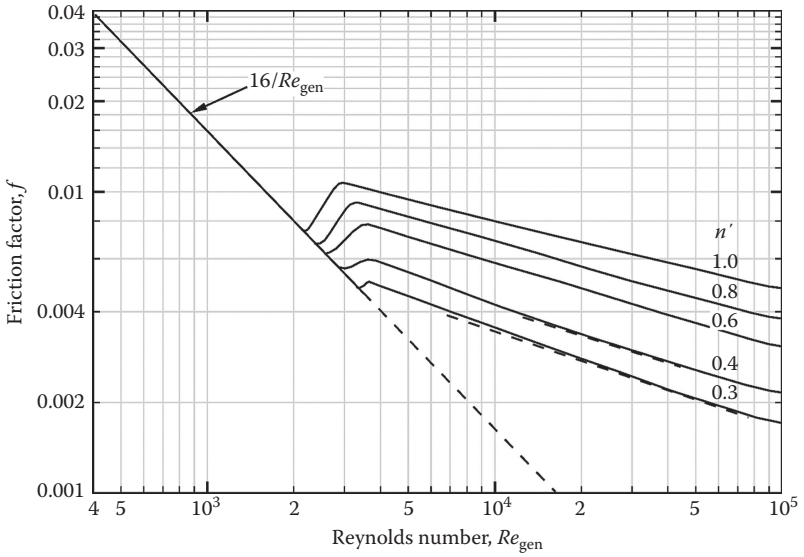


FIGURE 2.9.13 Friction factor–Reynolds number behavior for time-independent fluids as predicted by Equation 2.9.45.

$$f = \mathrm{F}(n) Re_{\mathrm{gen}}^{-\frac{1}{3n+1}}, \tag{2.9.46}$$

where $$\mathrm{F}(n) = \frac{2^{n+4}}{7^{7n}} \left(\frac{4n}{3n+1} \right)^{3n^2}$$

Both Equations 2.9.45 and 2.9.46 predict the value of friction factor with an accuracy of ~8–10% over the range of conditions as $\sim 0.35 \le n' \le 1$ and $Re_{\mathrm{gen}} \le 5 \times 10^4$. Further detailed discussion of the comparative performance of these as well as the numerous other friction factor correlations are available in the literature (Heywood and Cheng, 1984; Chhabra and Richardson, 2008; van den Heever et al., 2014).

For the case of Bingham plastic fluids in circular tubes, the following method due to Darby et al. (1992) has been shown to be reliable. In this approach, a weighted average friction factor is used as

$$f = \left(f_{\mathrm{L}}^{e} + f_{\mathrm{T}}^{e} \right)^{1/e}, \tag{2.9.47}$$

where f_{L} is given by the solution of the laminar expression, Equation 2.9.20 or 2.9.22, and the value of f_{T} is obtained from the following equation:

$$f_T = 10^{a_0} Re_{\mathrm{B}}^{-0.193}, \tag{2.9.48}$$

where

$$a_0 = -1.47 \left[1 + 0.146 \exp\left(-2.9 \times 10^{-5} He \right) \right], \tag{2.9.49}$$

and

$$e = 1.7 + \frac{4 \times 10^4}{Re_{\mathrm{B}}}. \tag{2.9.50}$$

This empirical formulation has been tested over wide range of conditions as $D \leq 335\,\text{mm}$, $10^3 \leq He \leq 6.6 \times 10^7$, and $Re_B \leq 3.4 \leq 10^5$.

Using the geometric parameters a and b, Equation 2.9.45 has also been extended to the turbulent flow of viscous fluids in noncircular ducts as

$$\frac{1}{\sqrt{f}} = \frac{4}{(n')^{3/4}} \log\left(Re_{gen} f^{(2-n')/2} \right) - \frac{0.4}{(n')^{1.2}} + 4(n')^{1/4} \log\left\{ \frac{4(a+bn')}{3n'+1} \right\}. \tag{2.9.51}$$

The limited experimental data available for flow in triangular, rectangular, and square ducts are in agreement with Equation 2.9.51 (Chhabra and Richardson, 2008). This section is concluded by noting that the value of $Re_{gen} <\sim 2100$ does seem to provide a conservative estimate of the cessation of the laminar flow regime both in circular and noncircular ducts for purely viscous fluids.

Effect of Viscoelasticity

In spite of the fact that much confusion exists in the literature concerning the role of fluid elasticity on pipe flow characteristics, some general inferences can, however, be drawn. In the laminar flow regime, the friction factor–Reynolds number relationships discussed in the preceding section are not influenced by the elasticity. However, for noncircular ducts, viscoelastic effects manifest itself in the form of recirculating regions (vortex formation) for certain shapes, though this does not appear to influence significantly the value of pressure gradient. In contrast, fluid elasticity gives rise to the well-known drag reduction effect under turbulent flow conditions in circular pipes. Indeed, it is possible to lower the pumping power requirements by up to 85% by simply adding trace amounts (parts per million) of high molecular weight polymers to a solvent (mostly water) under appropriate conditions in terms of the Reynolds number. Numerous hypotheses have been put forward to explain this phenomena (also known as the Toms effect), but none has proved to be completely satisfactory (Sellin et al., 1982; Gyr and Bewersdorff, 1995). Dosages of polymers (like polyethylene oxide, polyacrylamide, etc.) as small as 5 ppm can yield significant drag reduction, as shown in Figure 2.9.14 for polyethylene oxide polymer. Though initially the degree of drag reduction increases with

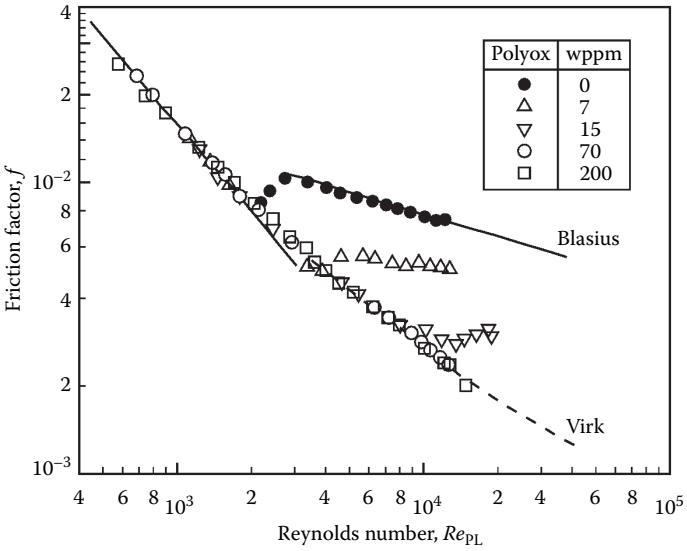


FIGURE 2.9.14 Drag reduction in friction factors for polyethylene oxide (Polyox) solutions in a small-diameter capillary tube. (From Cho, Y.I., Harnett, J.P., *Adv. Heat Transfer*, 15, 59–141, 1982. With permission.)

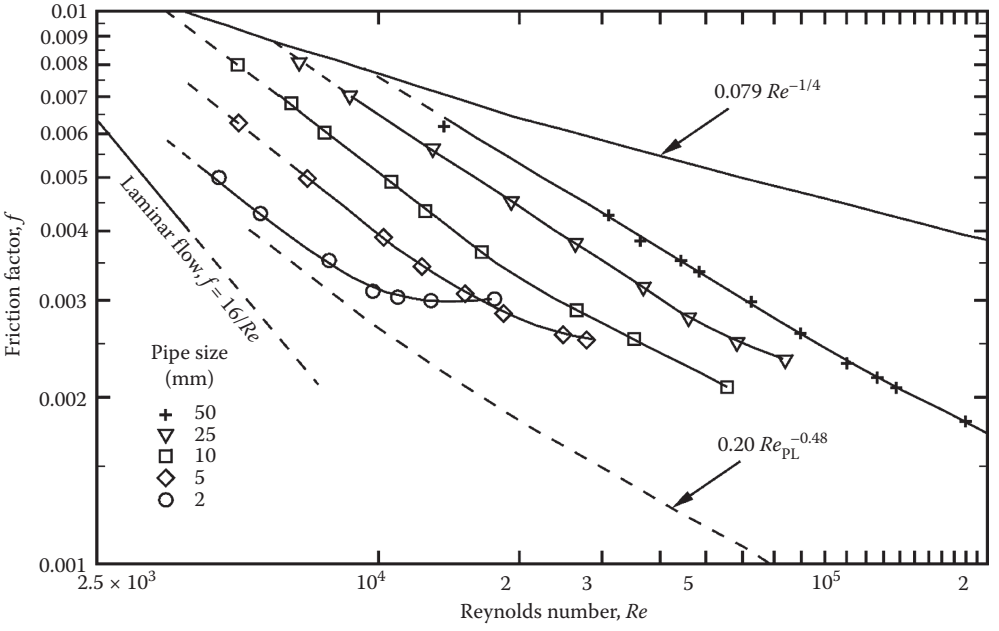


FIGURE 2.9.15 Effect of pipe diameter on friction factor (drag reduction) –10 wppm solutions of Alcomer 110L in water in pipes of various diameters. (From Hayes, J.W., Hutton, J.F., *Prog. Heat Mass Transfer*, 5, 195–209, 1972. With permission.)

polymer concentration (which varies from one polymer to another depending upon its molecular weight and architecture, and in some cases even with the pipe diameter, for instance), it reaches the so-called Virk's maximum drag reduction limit as seen in Figure 2.9.14. On the other hand, the effect of pipe diameter on drag reduction is depicted in Figure 2.9.15.

As of now, no reliable mechanism and/or predictive expression is available for predicting drag reduction quantitatively. Suffice it to add here, however, that the value of friction factor has an upper bound given by Blasius equation (pure water) $f = 0.079Re^{-1/4}$ and the lower bound is given by the Virk's maximum drag reduction asymptote approximated as $f = 0.20\,Re_{PL}^{-0.48}$.

However, no satisfactory information is currently available either to predict the onset or the extent of drag reduction in general (Gyr and Bewersdorff, 1995). In summary, our current understanding of the effect of viscoelasticity in turbulent flow regime in circular pipes is nowhere near as extensive or coherent as that for Newtonian fluids.

Nomenclature

A	Flow area, m^2
a, b	Geometric parameters, Equation 2.9.44, dimensionless
a_0	Constant, Equation 2.9.49, dimensionless
Bi	Bingham number (= $\tau_0 D / KV$), dimensionless
c	Duct width in Kozicki et al. (1967) approximation, Table 2.9.2, m
d_i	Concentric duct inner diameter in Kozicki et al. (1967) approximation, Table 2.9.2, m
d_0	Concentric duct outer diameter in Kozicki et al. (1967) approximation, Table 2.9.2, m
D	Pipe diameter, m

D_h	Hydraulic diameter (= 4 × A / wetted perimeter), m
e	Constant, Equation 2.9.47, dimensionless
f	Fanning friction factor, Equation 2.9.16, dimensionless
f_L	Fanning friction factor in laminar regime, dimensionless
f_T	Fanning friction factor in turbulent regime, dimensionless
$F(n)$	Function of n in Equation 2.9.46, dimensionless
$g(\tau)$	Shear rate function, s^{-1}
h	Duct height in Kozicki et al. (1967) approximation, Table 2.9.2, m
He	Hedstrom number, Equation 2.9.21, dimensionless
i	Index in Equation 2.9.25, dimensionless
K	Viscosity model parameter, Equation 2.9.6 and Equation 2.9.12a, Pa·s^n
K'	Apparent consistency coefficient, Equation 2.9.36, Pa·s^n
L	Pipe length, m
N	Number of sides in polygon duct in Kozicki et al. (1967) approximation, Table 2.9.2, dimensionless
N_1	First normal stress difference, Equation 2.9.3, Pa
N_2	Second normal stress difference, Equation 2.9.3, Pa
n	Viscosity model parameter, Equations 2.9.6 and 2.9.12a, dimensionless
n'	Apparent power-law index, Equation 2.9.36, dimensionless
p	Pressure, Pa
$\Delta p / L$	Pressure gradient, Pa/m
Q	Volumetric flow rate, m^3/s
R	Pipe radius, m
R*	Limiting viscosity ratio, μ_∞/μ_0, dimensionless
Re	Newtonian Reynolds number, dimensionless
Re_B	Bingham Reynolds number, Equation 2.9.21, dimensionless
Re_g	Power-law Reynolds number, Equation 2.9.31, dimensionless
Re_{gen}	Generalized Reynolds number, Equation 2.9.38, dimensionless
Re_m	Modified Reynolds number, Equation 2.9.28, dimensionless
Re_{PL}	Power-law Reynolds number, Equation 2.9.18, dimensionless
$\underline{\underline{S}}$	Stress tensor, Pa
t	Time, s
T	Temperature, K
V	Average velocity, m/s
V_x	x-Component of velocity, m/s
x, y, z	Cartesian coordinates

Greek Letters

α_1, α_2, α_3	Parameters, Equation 2.9.23, dimensionless
α^*	Geometric parameter in Kozicki et al. (1967) approximation, Table 2.9.2, dimensionless
β	Shear rate parameter based on μ_0, Equation 2.9.30, dimensionless
β_0	Constant, Equation 2.9.43, dimensionless
β_∞	Shear rate parameter, same as Equation 2.9.30 except based on μ_∞, dimensionless
β^*	Shear rate parameter, dimensionless
$\dot{\gamma}$	Rate of shear, s^{-1}
$\dot{\gamma}_{ij}$	Component of rate-of-strain tensor, s^{-1}

$\dot{\gamma}_w$	Rate of shear at the duct wall, s^{-1}
ε_i	Constants, Equation 2.9.25, dimensionless
θ	Vertex angle in isosceles duct in Kozicki et al. (1967) approximation, Table 2.9.2, degrees
μ	Viscosity, Pa·s
μ_{eff}	Effective viscosity, Pa·s
μ_0	Zero shear rate viscosity, Pa·s
μ_∞	Infinite shear rate viscosity, Pa·s
μ^*	Reference viscosity, Equation 2.9.29, Pa·s
ρ	Density, kg/m^3
τ	Extra stress tensor, Pa
τ_{ii}	Normal component of extra stress tensor, Pa
τ_{ij}	Shearing component of extra stress tensor, Pa
τ_0	Yield stress, Pa
τ_w	Wall shear stress, Pa
ϕ	Ratio $\left(= \tau_0 / \tau_w\right)$, dimensionless

REFERENCES

Barnes, H.A. 1989. Shear-thickening ("dilatancy") in suspensions of non-aggregating solid particles dispersed in Newtonian liquids, *Journal of Rheology*, 33: 329–366.

Barnes, H.A. 1997. Thixotropy—A review, *Journal of Non-Newtonian Fluid Mechanics*, 70: 1–33.

Barnes, H.A. 1999. The yield stress—A review, *Journal of Non-Newtonian Fluid Mechanics*, 81: 133–178.

Bird, R.B., Armstrong, R.C. and Hassager, O. 1987. *Dynamics of Polymer Liquids Vol. 1: Fluid Dynamics*, 2nd ed., Wiley, New York.

Boger, D.V. and Walters, K. 1992. *Rheological Phenomena in Focus*, Elsevier, Amsterdam.

Brewster, A.A. 2013. Pressure drop predictions for laminar fully-developed flows of purely-viscous non-Newtonian fluids in circular ducts, *ASME Journal of Fluids Engineering*, 135: 101106.

Brewster, A.A. and Irvine, T.F., Jr. 1987. Similitude considerations in laminar flow of power-law fluids in circular ducts, *Wärme und Stoffübertragung*, 21: 83–86.

Capobianchi, M. 2008. Pressure drop predictions for laminar flows of extended modified power law fluids in rectangular ducts, *International Journal of Heat and Mass Transfer*, 51: 1393–1401.

Capobianchi, M. and Aziz, A., Jr. 2012. Laminar natural convection from an isothermal vertical surface to pseudoplastic and dilatant fluids, *ASME Journal of Heat Transfer*, 134: 122502.

Capobianchi, M. and Irvine, T.F., Jr. 1992. Predictions of pressure drop and heat transfer in concentric annular ducts with modified power-law fluids, *Wärme und Stoffübertragung*, 27: 209–215.

Chhabra, R.P. 2003. Fluid mechanics and heat transfer with non-Newtonian liquids in mechanically agitated vessels, *Advances in Heat Transfer*, 37: 77–178.

Chhabra, R.P. 2006. *Bubbles, Drops, and Particles in Non-Newtonian Fluids*, 2nd ed., CRC Press, Boca Raton, FL.

Chhabra, R.P. 2011. Fluid flow and heat transfer from circular and non-circular cylinders submerged in non-Newtonian liquids, *Advances in Heat Transfer*, 43: 289–417.

Chhabra, R.P. 2016. Rheology: From simple fluids to complex suspensions, in *Lingnocellulosic Fibers and Wood Handbook*, Belgacem, N. and Pizzi, A., eds., pp. 407–438, Scrivener, New York.

Chhabra, R.P. and Richardson, J.F. 2008. *Non-Newtonian Flow and Applied Rheology*, 2nd ed., Butterworth-Heinemann, Oxford.

Cho, Y.I. and Hartnett, J.P. 1982. Non-Newtonian fluids in circular pipe flow, *Advances in Heat Transfer*, 15: 59–141.

Coussot, P. 2005. *Rheometry of Pastes, Suspensions and Granular Materials*, Wiley, New York.

Cross, M.M. 1965. Rheology of non-Newtonian fluids: A new flow equation for pseudoplastic systems, *Journal of Colloid Science*, 20: 417–437.

Darby, R. 1976. *Viscoelastic Fluids*: *An Introduction to Their Properties and Behaviour*, Marcel Dekker, New York.

Darby, R., Mun, R. and Boger, D.V. 1992. Predict friction loss in slurry pipes, *Chemical Engineering,* 99 (No. 9): 116–119.

Delplace, F. and Leuliet, J.C. 1995. Generalized Reynolds number of the flow of power law fluids in cylindrical ducts of arbitrary cross-section, *Chemical Engineering Journal,* 56: 33–37.

Dodge, D.W. and Metzner, A.B. 1959. Turbulent flow of non-Newtonian systems, *AIChE Journal,* 5: 189–204; corrections ibid 8: 143 (1962).

Dullaert, K. and Mewis, J. 2005. Thixotropy: Build up and breakdown curves during flow, *Journal of Rheology*, 49: 1213–1230.

Dullaert, K. and Mewis, J. 2006. A structural kinetics model for thixotropy, *Journal of Non-Newtonian Fluid Mechanics*, 139: 21–30.

Dunleavy, J.E. Jr. and Middleman, S. 1966. Correlation of shear behavior of solutions of polyisobutylene, *Transactions of The Society of Rheology*, 10: 157–168.

Ghosh, U.K., Upadhyay, S.N. and Chhabra, R.P. 1994. Heat and mass transfer from immersed bodies to non-Newtonian fluids, *Advances in Heat Transfer*, 25: 251–319.

Gnambode, P.S., Orlandi, P., Ould-Rouiss, M. and Nicolas, X. 2015. Large-eddy simulation of turbulent pipe flow of power-law fluids, *International Journal of Heat and Fluid Flow,* 54: 196–210.

Govier, G.W. and Aziz, K. 1982. *The Flow of Complex Mixtures in Pipes*, 2nd ed., Krieger, Malabar, FL.

Gyr, A. and Bewersdorff, H.-W. 1995. *Drag Reduction of Turbulent Flows by Additions*, Springer, Munich.

Hayes, J.W. and Hutton, J.F. 1972. The effect of very dilute polymer solutions on the formation of Taylor vortices: Comparison of theory with experiment, *Progress in Heat and Mass Transfer*, 5: 195–209.

Haynes, W.M. 2015. *CRC Handbook of Chemistry and Physics*, 96th ed., CRC Press, Boca Raton, FL.

Heywood, N.I. and Cheng, D.C.-H. 1984. Comparison of methods for predicting head loss in turbulent pipe flow of non-Newtonian fluids, *Transactions of the Institution of Measurement and Control*, 6: 33–45.

Ihle, C.H. and Tamburrino, A. 2012. A note on the Buckingham equation, *Canadian Journal of Chemical Engineering*, 90: 944–945.

Irvine, T.F. 1988. A generalized Blasius equation for power-law fluids, *Chemical Engineering Communications*, 65: 39–47.

Irvine, T.F., Jr. and Karni, J. 1987. Non-Newtonian fluid flow and heat transfer, in *Handbook of Single Phase Convective Heat Transfer*, John Wiley & Sons, New York.

Kozicki, W., Chou, C.H. and Tiu, C. 1967. Non-Newtonian flow in ducts of arbitrary cross-sectional shape, *Canadian Journal of Chemical Engineering*, 45: 127–134.

Lawal, A. and Mujumdar, A.S. 1989. Laminar duct flow and heat transfer to purely viscous non-Newtonian fluids, *Advances in Transport Processes*, 5: 352–443.

Macosko, C.W. 1994. *Rheology: Principles, Measurements and Applications*, Wiley, New York.

Mewis, J. and Wagner, N.J. 2012. *Colloidal Suspension Rheology*, Cambridge University Press, New York.

Mujumdar, A., Beris, A.N. and Metzner, A.B. 2002. Transient phenomena in thixotropic systems, *Journal of Non-Newtonian Fluid Mechanics*, 102: 157–178.

Muzychka, Y.S. and Edge, J. 2008. Laminar non-Newtonian fluid flow in noncircular ducts and microchannels, *Journal of Fluids Engineering*, 130: 111201.

Ngugen, Q.D. and Uhlherr, P.H.T. 1983. Thixotropic behavior of concentrated red mud suspensions, *Proceedings of 3rd National Conference Rheology,* Melbourne, Australia, pp. 63–68.

Park, S., Irvine, T.F., Jr. and Capobianchi, M. 1993. Experimental and numerical study of friction factor for a modified power-law fluid in a rectangular duct, *Proceedings of 3rd World Conference Heat Transfer, Fluid Mechanics and Thermodynamics*, Vol. 1, pp. 900–908, Elsevier, New York.

Poling, B.E., Prausnitz, J.M. and O' Connell, J.P. 2000. *The Properties of Gases and Liquids*, 5th ed., McGraw Hill, New York.

Rudman, M., Blackburn, H.M., Graham, L.J.W. and Pullum, L. 2004. Turbulent pipe flow of shear-thinning fluids, *Journal of Non-Newtonian Fluid Mechanics*, 118: 33–48.

Sellin, R.H.J., Hoyt, J.W. and Scrivener, O. 1982. The effect of drag reducing additives on fluid flows and their industrial applications. Part I: Basic concepts, *Journal of Hydraulic Research*, 20: 29–68; Part II: Present applications and future proposals, *Journal of Hydraulic Research*, 20: 235–292.

Skelland, A.H.P. 1967. *Non-Newtonian Flow and Heat Transfer*, Wiley, New York.

Steffe, J.F. 1996. *Rheological Methods in Food Process Engineering*, Freeman Press, East Lansing, MI.

Steg, I. and Katz, D. 1965. Rheopexy in some polar fluids and in their concentrated solutions in slightly polar solvents, *Journal of Applied Polymer Science*, 9: 3177–3193.

Van den Heever, E.M., Sutherland, A.P.N. and Haldenwang, R. 2014. Influence of the rheological model used in pipe-flow prediction techniques for homogeneous non-Newtonian fluids, *Journal of Hydraulic Engineering*, 140: 04014059.
Whorlow, R.W. 1992. *Rheological Techniques*, 2nd ed., Ellis Horwood, London.

Further Information

While it is not possible to include all of the interesting topics relating to non-Newtonian fluids in a Handbook like this one, some relevant references are indicated here. Cho and Hartnett (1982), Irvine and Karni (1987), and Lawal and Mujumdar (1989) have provided overviews of duct flows with and without heat transfer. External boundary layer type flows have been treated by Chhabra (2006, 2011) and by Ghosh et al. (1994). Momentum and heat-transfer characteristics in mechanically agitated vessels has been dealt with by Chhabra (2003). Non-Newtonian effects in porous media and particulate flows are reviewed by Chhabra (2006).

3 Heat and Mass Transfer

Robert F. Boehm, Swati A. Patel, Raj P. Chhabra, George D. Raithby, K.G. Terry Hollands, Anoop K. Gupta, N.V. Suryanarayana, Thomas F. Irvine, Jr., Massimo Capobianchi, Michael F. Modest, Van P. Carey, John C. Chen, Vasilios Alexiades, Jan Kośny, and Anthony F. Mills

CONTENTS

3.1 Conduction Heat Transfer 250
Introduction 250
Fourier's Law 251
Insulations 251
Plane Wall at Steady State 252
Long, Cylindrical Systems at Steady State 253
Overall Heat Transfer Coefficient 253
Critical Thickness of Insulation 254
Internal Heat Generation 254
Fins 255
Transient Systems Negligible Internal Resistance 258
Finite Difference Analysis of Conduction 259
Defining Terms 262
References 262
3.2 Convection Heat Transfer 263
3.2.1 Natural Convection 263
References 280
3.2.2 Forced Convection: External Flows 283
Acknowledgment 303
Bibliography 303
References 304
3.2.3 Forced Convection: Internal Flows 305
Acknowledgment 314
References 314
3.2.4 Convection Heat Transfer in Non-Newtonian Fluids 315
Acknowledgments 338
References 338
3.3 Radiative Heat Transfer 341
Nature of Thermal Radiation 341
Blackbody Radiation 341
Radiative Exchange between Opaque Surfaces 343
Radiative Exchange between Opaque Surfaces (Net Radiation Method) 352
Radiative Exchange within Participating Media 357

References 367
3.4 Phase-Change 368
3.4.1 Boiling and Condensation 368
References 385
3.4.2 Particle Gas Convection 388
References 392
3.4.3 Melting and Freezing 393
References 413
3.5 Mass Transfer 417
Introduction 417
Concentrations, Velocities, and Fluxes 417
Mechanisms of Diffusion 422
Species Conservation Equation 426
Transient Diffusion in Slabs, Cylinders, and Spheres 431
Diffusion in a Moving Medium 434
Diffusion with One Component Stationary 434
Heterogeneous Combustion 434
Droplet Evaporation 435
Droplet Combustion 436
Mass Convection 437
Mass and Mole Transfer Conductances 438
Dimensionless Groups 439
Analogy between Convective Heat and Mass Transfer 439
Simultaneous Heat and Mass Transfer 441
The Wet- and Dry-Bulb Psychrometer 443
High Mass Transfer Rate Theory 444
Variable Property Effects of High Mass Transfer Rates 446
References 449

3.1 CONDUCTION HEAT TRANSFER

Robert F. Boehm

Introduction

Conduction heat-transfer phenomena are found virtually throughout the physical world and the industrial domain. The analytical description of this heat-transfer mode is one of the best understood processes. Some of the bases of the understanding of conduction date back to early history. It was recognized that by invoking certain relatively minor simplifications, mathematical solutions resulted directly. Some of these were easily formulated. What transpired over the years was a vigorous development of applications to a broad range of processes. Perhaps no single work better summarizes the wealth of these studies than does the book by Carslaw and Jaeger (1959). It provides solutions to a broad range of problems, ranging from topics related to the cooling of the earth to the current-carrying capacities of wires. The general analyses given there have been applied to a range of modern-day problems from laser heating to temperature control systems.

Today, conduction heat transfer is still an active area of research and application. A great deal of interest has been shown in recent years in topics like contact resistance—where a temperature difference develops between two solids that do not have perfect contact with each other. Additional issues of current interest include non-Fourier conduction, where the processes occur so fast that the equations described later do not apply. Also, the problems related to transport in miniaturized

systems are garnering a great deal of interest. Increased interest has also been focused on ways of handling composite materials, where the ability to conduct heat is very directional.

Much of the work in conduction analysis is now accomplished by the use of sophisticated computer codes. These tools have given the heat-transfer analyst the capability to solve problems in inhomogeneous media with very complicated geometries and with very involved boundary conditions. It is still important to understand analytical methods for determining the performance of conducting systems. At the minimum, these can be used as calibrations for numerical codes.

Fourier's Law

The basis of conduction heat transfer is *Fourier's law.* This law involves the idea that the heat flux is proportional to the temperature gradient in any direction *n*. *Thermal conductivity*, *k*, a property of materials that is temperature dependent, is the constant of proportionality:

$$q_k = -kA\frac{\partial T}{\partial n} \tag{3.1.1}$$

In many systems, the area A is a function of the distance in the direction n. One important extension is that this can be combined with the first law of thermodynamics to yield the *heat conduction equation*. For constant thermal conductivity, this is given as follows:

$$\nabla^2 T + \frac{\dot{q}_G}{k} = \frac{1}{\alpha}\ \frac{\partial T}{\partial t} \tag{3.1.2}$$

where α is the thermal diffusivity and $\dot{q}_G$ is the internal heat generation per unit volume. Some problems, typically steady-state, one-dimensional formulations where only the heat flux is desired, can be easily solved using Equation 3.1.1. Most conduction analyses are performed with Equation 3.1.2. In the latter—a more general approach—the temperature distribution is found from this equation and the appropriate boundary conditions. Then the heat flux, if desired, is found at any location using Equation 3.1.1. Normally, it is the temperature distribution that is of most importance. For example, it may be desirable to determine through analysis if a material will reach some critical temperature, like its melting point. Less frequently, the heat flux is the desired objective.

While there are times when one needs only to understand the temperature response of a structure, often the engineer is faced with a need to increase or decrease heat transfer to some specific level. Examination of the thermal conductivity of materials gives some insight to the range of possibilities that exist through simple conduction.

Of the more common engineering materials, pure copper exhibits one of the highest abilities to conduct heat, with a thermal conductivity approaching 400 W/m^2 K. Aluminum, also considered to be a good conductor, has a thermal conductivity a little over half that of copper. In order to increase the heat transfer above the values possible through simple conduction, more sophisticated designs are necessary that incorporate a variety of other heat-transfer modes like convection and phase change.

Decreasing the rate of heat transfer is accomplished with the use of insulation. A discussion of this follows.

Insulations

Insulating materials are used to decrease heat flow and thus decrease surface temperatures. These materials are found in a variety of forms, typically *loose fill, batt*, and *rigid*. Even a gas, like air, can be a good insulator if it can be kept from moving when it is heated or cooled. A vacuum is an excellent insulator. Usually, though, the engineering approach to insulation is the addition of a

low-conducting material to the surface. While there are many chemical forms, costs, and maximum operating temperatures of common types of insulators, it seems that when higher operating temperatures are required, the thermal conductivity and cost of the insulation will often also be higher.

Loose-fill insulation includes such materials as milled alumina-silica (maximum operating temperature of 1260°C and thermal conductivities in the range of 0.1–0.2 W/m^2 K) and perlite (maximum operating temperature of 980°C and thermal conductivities in the range of 0.05–1.5 $W/m^2 \cdot K$). Batt-type insulation includes one of the more common types—glass fiber. This type of insulation comes in a variety of densities which, in turn, have a profound effect on the thermal conductivity. Thermal conductivities for glass fiber insulation can range from about 0.03 to 0.06 W/m^2 K. Rigid insulations show a very wide range of types and performance characteristics. For example, a rigid insulation in foam form, polyurethane, is light in weight, shows a very low thermal conductivity (about 0.02 W/m^2 K), but has a maximum operating temperature only up to about 120°C. Rigid insulations in refractory form show quite different characteristics. For example, alumina brick is quite dense, has a thermal conductivity of about 2 $W/m^2 \cdot K$, but can remain operational to temperatures around 1760°C. Many insulations are characterized in the book edited by Guyer (1989).

Often, commercial insulation systems designed for high-temperature operation use a layered approach. Temperature tolerance may be critical. Perhaps a refractory material is applied in the highest temperature region, an intermediate-temperature foam insulation is used in the middle section, and a high-performance, lower operating-temperature insulation is used on the outer side near ambient conditions.

Analyses can be performed showing the effects of temperature variations of thermal conductivity. However, the most frequent approach is to assume that the thermal conductivity is constant at some temperature between the two extremes experienced by the insulation.

Plane Wall at Steady State

Consider steady-state heat transfer in a plane wall of thickness L, but of very large extent in both other directions. The wall has temperature T_1 on one side and T_2 on the other. If the thermal conductivity is considered to be constant, then Equation 3.1.1 can be integrated directly to give the following result:

$$q_k = \frac{kA}{L}\left(T_1 - T_2\right) \tag{3.1.3}$$

This can be used to determine the steady-state heat transfer through slabs.

An electrical circuit analog is widely used in one dimensional conduction analyses. This is accomplished by considering the temperature difference to be analogous to a voltage difference, the heat flux to be like current flow, and the remainder of Equation 3.1.3 to be like a thermal resistance. The latter is seen to be

$$R_k = \frac{L}{kA} \tag{3.1.4}$$

Heat transfer through walls made of layers of different types of materials can be easily found by summing the resistances in series or parallel form, as appropriate.

In the design of systems, seldom is a surface temperature specified or known. More often, the surface is in contact with a bulk fluid, whose temperature is known at some distance from the surface. Convection from the surface is then represented by Newton's law of cooling:

$$q = \overline{h}_c A\left(T_s - T_\infty\right) \tag{3.1.5}$$

This equation can also be represented as a temperature difference divided by a thermal resistance, which is $1/\bar{h}_c A$. It can be shown that a very low surface resistance, as might be represented by phase-change phenomena, effectively imposes the fluid temperature directly on the surface. Hence, usually a *known* surface temperature results from a fluid temperature being imposed directly on the surface through a very high heat-transfer coefficient. For this reason, in the later results given here, particularly those for transient systems, a convective boundary will be assumed. For steady-state results, this is less important because of the ability to add resistances via the circuit analogy.

Long, Cylindrical Systems at Steady State

For long (L), annular systems at steady-state conditions with constant thermal conductivities, the following two equations are the appropriate counterparts to Equations 3.1.3 and 3.1.4. The heat transfer can be expressed as follows:

$$q_k = \frac{2\pi Lk}{\ln\left[r_2/r_1\right]}(T_1 - T_2) \tag{3.1.6}$$

Here, r_1 and r_2 represent the radii of the annular sections. A thermal resistance for this case is as shown:

$$R_k = \frac{\ln[r_2 / r_1]}{2\pi Lk} \tag{3.1.7}$$

Overall Heat Transfer Coefficient

The *overall heat-transfer coefficient* concept is valuable in several aspects of heat transfer. It involves a modified form of Newton's law of cooling, as noted earlier, and it is written as follows:

$$Q = \bar{U}A\Delta T \tag{3.1.8}$$

In this formulation, $\bar{U}$ is the overall heat-transfer coefficient based upon the area A. Because the area for heat transfer in a problem can vary (as with a cylindrical geometry), it is important to note that the $\bar{U}$ is dependent upon which area is selected. The overall heat-transfer coefficient is usually found from a combination of thermal resistances. Hence, for a common-series combination-circuit analog, the $\bar{U}A$ product is taken as the sum of the resistances:

$$\bar{U}A = \frac{1}{\sum_{i=1}^{n} R_i} = \frac{1}{R_{\text{total}}} \tag{3.1.9}$$

For an example of the use of this concept, see Figure 3.1.1.

For steady-state conditions, the product $\bar{U}A$ remains constant for a given heat transfer and overall temperature difference. This can be written as follows:

$$\bar{U}_1 A_1 = \bar{U}_2 A_2 = \bar{U}_3 A_3 = \bar{U}A \tag{3.1.10}$$

If the inside area, A_1, is chosen as the basis, the overall heat-transfer coefficient can then be expressed as follows:

$$\bar{U}_1 = \frac{1}{\frac{1}{\bar{h}_{c,i}} + \frac{r_1 \ln\left(r_2/r_1\right)}{k_{\text{pipe}}} + \frac{r_1 \ln\left(r_3/r_2\right)}{k_{\text{ins}}} + \frac{r_1}{r_3 \bar{h}_{c,o}}} \tag{3.1.11}$$

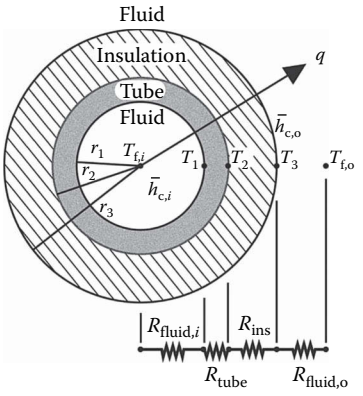


FIGURE 3.1.1 An insulated tube with convective environments on both sides shown with a resistance analog representation.

Critical Thickness of Insulation

Sometimes, adding insulation can cause an increase in heat transfer. This circumstance should be noted so that one can apply it when desired and design around it when an insulating effect is needed. Consider the circumstance shown in Figure 3.1.1. Assume that the temperature on the outside of the tube (inside of the insulation) is known. This could be determined if the inner heat-transfer coefficient is very large and the thermal conductivity of the tube is large. In this case, the inner fluid temperature will be almost the same temperature as the inner surface of the insulation. Alternatively, this could be applied to a coating (say an electrical insulation) on the outside of a wire. By forming the expression for the heat transfer in terms of the variables shown in Equation 3.1.11, and examining the change of heat transfer with variations in r_3 (i.e., the thickness of the insulation); the critical thickness of insulation can be found. While simple results are given in many texts (showing the critical radius as the ratio of the insulation thermal conductivity to the heat-transfer coefficient on the outside), Sparrow (1970) considered a heat-transfer coefficient that varies as follows:

$$\bar{h}_{c,o} \sim r_3^{-m} \left| T_3 - T_{f,o} \right|^n$$

For this case, it is found that the heat transfer is maximized at

$$r_3 = r_{crit} = \left[(1-m)/(1+n) \right] \frac{k_{ins}}{\bar{h}_{c,o}} \tag{3.1.12}$$

By examining the orders of magnitude of m, n, k_{ins}, and $\bar{h}_{c,o}$, the critical radius is often found to be on the order of a *few millimeters*. Hence, additional insulation on small-diameter cylinders such as narrow-gauge electrical wires could actually increase the heat dissipation. On the other hand, the addition of insulation to large-diameter pipes and ducts will almost always decrease the heat-transfer rate.

Internal Heat Generation

The analysis of temperature distributions and the resulting heat transfer in the presence of volume heat sources is required in some circumstances. These include phenomena such as nuclear fission processes, Joule heating, and microwave energy deposition. Consider first a slab of material $2L$ thick but otherwise very large, with internal generation. The outside of the slab is kept at temperature T_1.

To find the temperature distribution within the slab, the thermal conductivity is assumed to be constant. Equation 3.1.2 reduces to the following:

$$\frac{\mathrm{d}^2T}{\mathrm{d}x^2}+\frac{\dot{q}_\mathrm{G}}{k}=0 \tag{3.1.13}$$

Solving this equation by separating the variables, integrating twice, and applying boundary conditions gives

$$T(x)-T_1=\frac{\dot{q}_\mathrm{G}L^2}{2k}\left[1-\left(\frac{x}{L}\right)^2\right] \tag{3.1.14}$$

A similar type of analysis for a long cylindrical element of radius r_1 gives

$$T(r)-T_1=\frac{\dot{q}_\mathrm{G}r_1^2}{4k} \tag{3.1.15}$$

Two additional cases will be given. Both involve the situation when the rate of heat generation is dependent upon the local temperature, according to the following relationship:

$$\dot{q}_\mathrm{G}=\dot{q}_{\mathrm{G,o}}\left[1+\beta\left(T-T_0\right)\right] \tag{3.1.16}$$

For a plane wall of $2L$ thickness and a temperature of T_1 specified on each surface

$$\frac{T(x)-T_\mathrm{o}+1/\beta}{T_1-T_\mathrm{o}+1/\beta}=\frac{\cos\mu x}{\cos\mu L} \tag{3.1.17}$$

For a similar situation in a long cylinder with a temperature of T_1 specified on the outside radius r_1

$$\frac{T(r)-T_\mathrm{o}+1/\beta}{T_1-T_\mathrm{o}+1/\beta}=\frac{J_\mathrm{o}(\mu r)}{J_\mathrm{o}(\mu r_1)} \tag{3.1.18}$$

In Equation 3.1.18, J_o is the typical notation for the Bessel function. Variations of this function are tabulated in Abramowitz and Stegun (1964) or are available on many computer systems. In both of the cases shown here, the following holds:

$$\mu\equiv\sqrt{\frac{\beta\,\dot{q}_{\mathrm{G,o}}}{k}}$$

Fins

Fins are widely used to enhance the heat transfer (usually convective, but also radiative) from a surface. This is particularly true when the surface is in contact with a gas. Fins are used on air-cooled engines, electronic cooling forms, as well as in a number of other applications. Since the heat-transfer coefficient tends to be low in gas convection, area is added to the surface in the form of fins to decrease the convective thermal resistance.

The simplest fins to analyze are those usually found in practice—one-dimensional and constant in cross section. In simple terms, to be one-dimensional, the fins have to be long compared to a transverse dimension. Three cases are normally considered for analysis, and these are shown in Figure 3.1.2. They are the insulated-tip fin, the very long fin, and the convecting-tip fin.

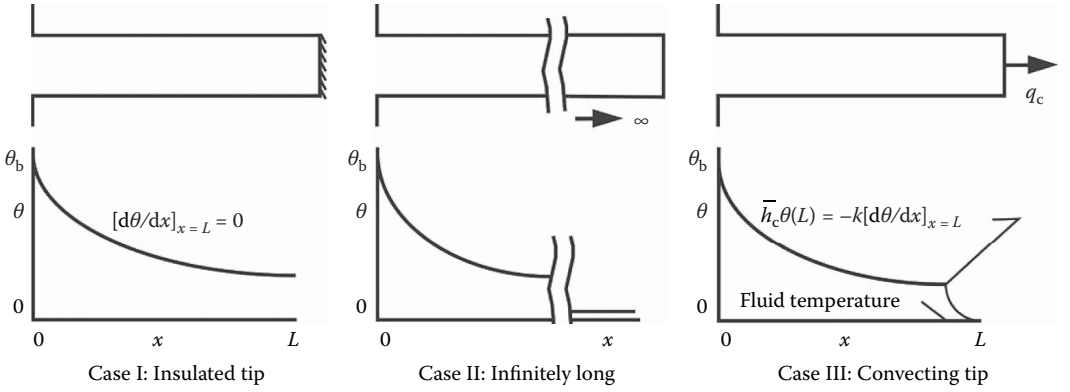


FIGURE 3.1.2 Three typical cases for one-dimensional, constant cross-section fins are shown.

For Case I, the solution to the governing equation and the application of the boundary conditions of the known temperatures at the base and the insulated tip yields

$$\text{Case I:} \qquad \theta = \theta_b = \frac{\cosh m(L-x)}{\cosh mL} \tag{3.1.19}$$

For the infinitely long case, the following simple form results:

$$\text{Case II:} \qquad \theta(x) = \theta_b e^{-mx} \tag{3.1.20}$$

The final case yields the following result:

$$\text{Case III:} \qquad \theta(x) = \theta_b \quad \frac{mL\cosh m(L-x) + \text{Bi}\sinh m(L-x)}{mL\cosh mL + \text{Bi}\sinh mL} \tag{3.1.21}$$

where $\text{Bi} \equiv \bar{h}_c\, L/k$.

In all three of the cases given, the following definitions apply:

$$\theta \equiv T(x) - T_\infty, \theta_b \equiv T(x=0) - T_\infty, \text{and } m^2 \equiv \frac{\bar{h}_c P}{kA}$$

Here, A is the cross section of the fin parallel to the wall. P is the perimeter around that area.

To find the amount of heat removed in any of these cases, the temperature distribution is used in Fourier's law, Equation 3.1.1. For most fins that truly fit the one-dimensional assumption (i.e., long compared to their transverse dimensions), all three equations will yield results that don't differ significantly.

Two performance indicators are found in the fin literature. The *fin efficiency* is defined as the ratio of the actual heat-transfer rate from a fin to the heat-transfer rate from an ideal fin:

$$\eta \equiv \frac{q_{\text{actual}}}{q_{\text{ideal}}} \tag{3.1.22}$$

The ideal heat transfer is found from convective gain or loss from an area the same size as the fin surface area, all at a temperature T_b. Fin efficiency is normally used to tabulate heat-transfer results for various types of fins, including those with nonconstant area or not meeting the one-dimensional

assumption. An example of the former can be developed from a result given by Arpaci (1966). Consider a straight fin of triangular profile, as shown in Figure 3.1.3. The solution is found in terms of modified Bessel functions of the first kind. Tabulations are given in Abramowitz and Stegun (1964):

$$\eta = \frac{I_1\left(2mL^{1/2}\right)}{mL^{1/2} I_o\left(2mL^{1/2}\right)} \tag{3.1.23}$$

Here, $m \equiv \sqrt{2\bar{h}_c L / kb}$.

The *fin effectiveness*, ε, is defined as the heat transfer from the fin compared to the bare-surface transfer through the same base area:

$$\varepsilon = \frac{q_{\text{actual}}}{q_{\text{bare base}}} = \frac{q_f}{\bar{h}_c A\left(T_b - T_\infty\right)} \tag{3.1.24}$$

Carslaw and Jaeger (1959) give an expression for the effectiveness of a fin of constant thickness around a tube (see Figure 3.1.3b). This is given as $\mu \equiv \sqrt{2\bar{h}_c / kb}$

$$\varepsilon = \frac{2}{\mu b} \frac{I_1(\mu r_2) K_1(\mu r_1) - K_1(\mu r_2) I_1(\mu r_1)}{I_o(\mu r_1) K_1(\mu r_2) + K_o(\mu r_1) I_1(\mu r_2)} \tag{3.1.25}$$

Here, the notations *I* and *K* denote Bessel functions that are given in Abramowitz and Stegun (1964).

Fin effectiveness can be used as one indication whether or not fins should be added. A rule of thumb indicates that if the effectiveness is less than about three, fins should not be added to the surface.

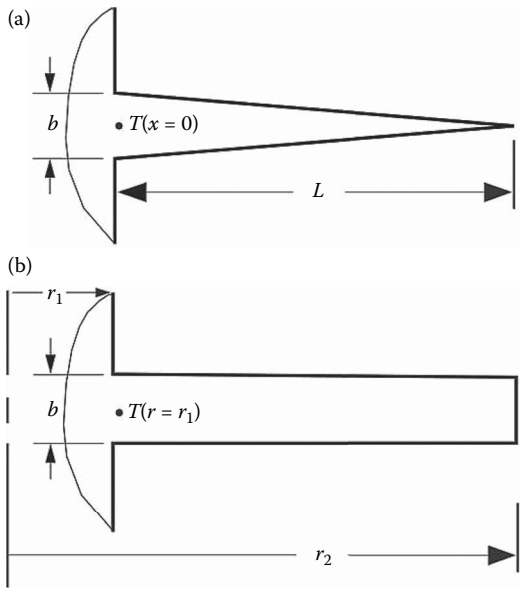


FIGURE 3.1.3 Two fin examples with a cross-sectional area that varies with distance from the base. (a) Straight triangular fin. (b) Annular fin of constant thickness.

Transient Systems with Negligible Internal Resistance

Consider the transient cooling or heating of a body with surface area A and volume V. This is taking place by convection through a heat-transfer coefficient $\bar{h}_c$ to an ambient temperature of T_∞. Assume the thermal resistance to conduction inside the body is significantly less than the thermal resistance to external convection (as represented by Newton's law of cooling) on the surface of the body. This ratio is denoted by the *Biot number*, Bi:

$$\mathrm{Bi} = \frac{R_k}{R_c} = \frac{\bar{h}_c (V/A)}{k} \tag{3.1.26}$$

The temperature (which will be uniform throughout the body at any time for this situation) response with time for this system is given by the following relationship. Note that the shape of the body is not important—only the ratio of its volume to its area matters:

$$\frac{T(t) - T_\infty}{T_o - T_\infty} = \mathrm{e}^{-\frac{\bar{h}_c A t}{\rho V c}} \tag{3.1.27}$$

Typically, this will hold for the Biot number being less than (about) 0.1.

Bodies with Significant Internal Resistance

When a body is being heated or cooled transiently in a convective environment, but the internal thermal resistance of the body cannot be neglected, the analysis becomes more complicated. Only simple geometries (a symmetrical plane wall, a long cylinder, a composite of geometrical intersections of these geometries, or a sphere) with an imposed step change in ambient temperature are addressed here.

The first geometry considered is a large slab of minor dimension $2L$. If the temperature is initially uniform at T_o, and at time 0+, it begins convecting through a heat-transfer coefficient to a fluid at T_∞, the temperature response is given by

$$\theta = 2\sum_{n=1}^{\infty}\left(\frac{\sin\lambda_n L}{\lambda_n L + \sin\lambda_n L\cos\lambda_n L}\right)\exp\left(-\lambda_n^2 L^2 \mathrm{Fo}\right)\cos(\lambda_n x) \tag{3.1.28}$$

and λ_n are the roots of the transcendental equation: $\lambda_n L \tan \lambda_n L = \mathrm{Bi}$. The following definitions hold:

$$\mathrm{Bi} \equiv \frac{\bar{h}_c L}{k} \qquad \mathrm{Fo} \equiv \frac{\alpha t}{L^2} \qquad \theta \equiv \frac{T - T_\infty}{T_o - T_\infty}$$

The second geometry considered is a very long cylinder of diameter $2R$. The temperature response for this situation is

$$\theta = 2\mathrm{Bi}\sum_{n=1}^{\infty}\frac{\exp\left(-\lambda_n^2 R^2 \mathrm{Fo}\right) J_o\left(\lambda_n r\right)}{\left(\lambda_n^2 R^2 + \mathrm{Bi}^2\right) J_o\left(\lambda_n R\right)} \tag{3.1.29}$$

Now, λ_n are the roots of $\lambda_n R\, J_1(\lambda_n R) - \mathrm{Bi}\, J_o(\lambda_n R) = 0$, and

$$\mathrm{Bi} = \frac{\bar{h}_c R}{k} \qquad \mathrm{Fo} = \frac{\alpha t}{R^2} \qquad \theta = \frac{T - T_\infty}{T_o - T_\infty}$$

The common definition of Bessel's functions applies here.

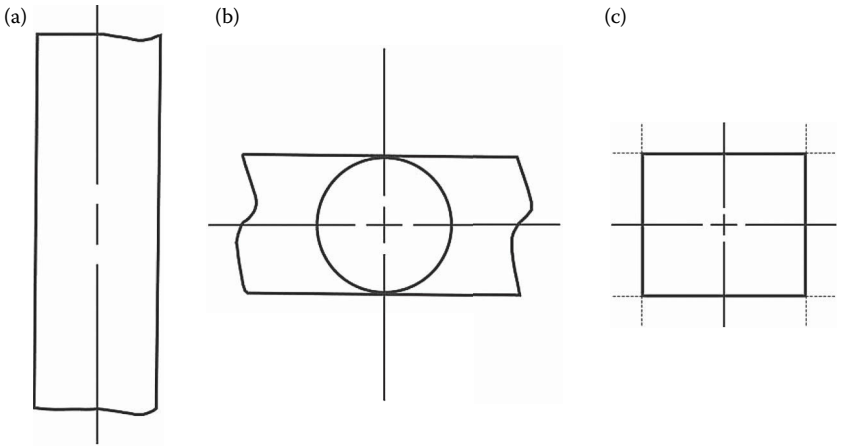


FIGURE 3.1.4 Three types of bodies that can be analyzed with methods given in this section. (a) Large plane wall of $2L$ thickness. (b) Long cylinder with $2R$ diameter. (c) Composite intersection.

For the similar situation involving a solid sphere, the following holds:

$$\theta = 2\sum_{n=1}^{\infty} \frac{\sin(\lambda_n R) - \lambda_n R\cos(\lambda_n R)}{\lambda_n R - \sin(\lambda_n R)\cos(\lambda_n R)} \exp\left(-\lambda_n^2 R^2 \mathrm{Fo}\right) \frac{\sin(\lambda_n r)}{\lambda_n r} \tag{3.1.30}$$

and λ_n are found as the roots of $\lambda_n R\cos\lambda_n R = (1 - \mathrm{Bi})\sin\lambda_n R$. Otherwise, the same definitions hold as were given for the cylinder.

Solids that can be envisioned as the geometrical intersection of the simple shapes described earlier can be analyzed with a simple product of the individual shape solutions. For these cases, the solution is found as the product of the dimensionless temperature functions for each of the simple shapes, with appropriate distance variables taken in each solution. This is illustrated in the right-hand diagram in Figure 3.1.4. For example, a very long rod of rectangular cross section can be seen as the intersection of two large plates. A short cylinder represents the intersection of an infinitely long cylinder and a plate. The temperature at any location within the short cylinder is

$$\theta_{2R,2L\ \mathrm{Rod}} = \theta_{\mathrm{Infinite}\ 2R\ \mathrm{Rod}}\theta_{2L\ \mathrm{Plate}} \tag{3.1.31}$$

Details of the formulation and solution of the partial differential equations in heat conduction are found in the text by Arpaci (1966).

Finite Difference Analysis of Conduction

Today, numerical solution of conduction problems is the most used analysis approach. Two general techniques are applied for this: those based upon finite-difference ideas and those based upon finite-element concepts. General numerical formulations are introduced in Chapter 5 of this text. Here, we briefly introduce some finite-difference concepts as they might be applied in heat conduction analysis.

First, consider formulation ideas. Consider the two-dimensional form of Equation 3.1.2, neglecting heat generation. This is given by

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = \frac{1}{\alpha}\frac{\partial T}{\partial t} \tag{3.1.32}$$

By using finite-difference approximations to the first and second derivatives, the following results:

$$\frac{T_{i+1,j} - 2T_{i,j} - T_{i-1,j}}{\Delta x^2} + \frac{T_{i,j+1} - 2T_{i,j} - T_{i,j-1}}{\Delta y^2} = \frac{T'_{i,j} - T_{i,j}}{\alpha \Delta t} \tag{3.1.33}$$

In this notation, i denotes node center locations in the x-direction, j indicates node center locations in the y-direction, and $T_{i,j} = T(i, j, t)$, and $T'_{i,j} = T(i, j, t+\Delta t)$. As written, the form is denoted as the *explicit* formulation because the forward difference is used for the time derivative (a new time-step temperature is found in terms of all of the old time-step temperatures). If the time derivative is taken as the backward difference, then the new (and, hence, unknown) temperatures at all surrounding nodes are used with the known (old) temperature at each node to find the temperatures at the new time. Because each equation contains a number of unknowns that must be found simultaneously, the result is termed the *implicit* formulation. This is shown below for a general interior node:

$$\frac{T'_{i+1,j} - 2T'_{i,j} - T'_{i-1,j}}{\Delta x^2} + \frac{T'_{i,j+1} - 2T'_{i,j} - T'_{i,j-1}}{\Delta y^2} = \frac{T'_{i,j} - T_{i,j}}{\alpha \Delta t} \tag{3.1.34}$$

A formal differencing approach, where finite differences are generated directly from partial derivatives, can be carried out for each application. However, the possibility of unusual geometries and mixing of boundary conditions often results in situations that are not easily differenced. For this reason, an energy balance method often yields an approach that can be more easily applied physically.

Attention is drawn to a one-dimensional slab (very large in two directions compared to the thickness). The slab is divided across the thickness into smaller subslabs (denoted as nodes in what follows), and this is shown in Figure 3.1.5. All nodes are of thickness Δx except for the two boundaries where the thickness is $\Delta x/2$. A characteristic temperature for each node is assumed to be represented by the temperature at the node center. Of course this assumption becomes more accurate as the size of the node becomes smaller. On the periphery of the object being analyzed, the location representing the temperature of the node is located on the boundary of the object, and this results in fractionally sized nodes at these positions. Hence, with the two boundary node centers located exactly on the boundary, a total of n nodes are used ($n-2$ full, interior nodes and half-nodes on each of the two boundaries).

In the analysis, a general interior node i (this applies to all nodes 2 through $n-1$) is considered for an overall energy balance. Conduction in from node $i-1$ and from node $i+1$, as well as any heat generation present, is assumed to be energy per unit time flowing into the node. This is then equated

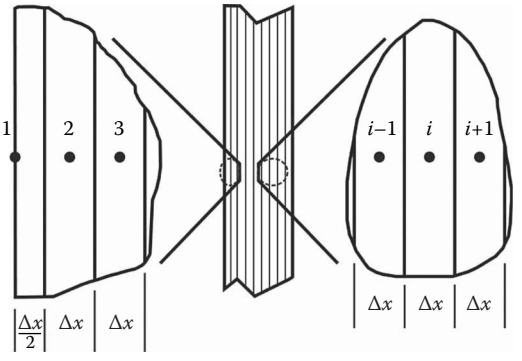


FIGURE 3.1.5 Shown are details of the one-dimensional finite differencing of a slab for a general interior node as well as one surface node.

to the time rate of change of energy within the node. A backward difference on the time derivative is applied here, and the notation $T_i' \equiv T_i(t+\Delta t)$ is used. The balance gives the following on a per unit area basis:

$$\frac{T'_{i-1}-T'_i}{\Delta x/k_-}+\frac{T'_{i+1}-T'_i}{\Delta x/k_+}+\dot{q}_{\mathrm{G},i}\Delta x=\rho\ \Delta x\ c_\mathrm{p}\frac{T'-T'_i}{\Delta t} \tag{3.1.35}$$

In this equation, different thermal conductivities have been used to allow for possible variations in properties throughout the solid.

The analysis of the boundary nodes will depend upon the nature of the conditions there. For the purposes of illustration, convection will be assumed to be occurring off of the boundary at node 1. A balance similar to Equation 3.1.35, but now for node 1, gives the following:

$$\frac{T'_\infty-T'_1}{1/h_\mathrm{c}}+\frac{T'_2-T'_1}{\Delta x/k_+}+\dot{q}_{\mathrm{G},1}\frac{\Delta x}{2}=\rho\ \frac{\Delta x}{2}\ c_\mathrm{p}\frac{T'_1-T_1}{\Delta t} \tag{3.1.36}$$

If the heat-transfer situation in the slab is symmetrical about the slab center line, consideration of this aspect allows only half of the slab to be analyzed. Consideration of the half-node located on the left-hand side of the centerline yields

$$\frac{T'_{n-1}-T'_n}{\Delta x/k_+}+0+\dot{q}_{\mathrm{G},n}\frac{\Delta x}{2}=\rho\ \frac{\Delta x}{2}\ c_\mathrm{p}\frac{T'_n-T_n}{\Delta t} \tag{3.1.37}$$

Here, the zero shown as the second term denotes the fact that no energy flows across the centerline.

After all n equations are written, it can be seen that there are n unknowns represented in these equations: the temperature at all nodes. If one or both of the boundary conditions is known in terms of specified temperatures, this will decrease the number of equations and unknowns by one or two, respectively. To determine the temperature as a function of time, the time step is arbitrarily set, and all the temperatures are found by simultaneous solution at $t=0+\Delta t$. This is denoted by the matrix system shown below. For the situation shown here, the thermal conductivity is taken to be constant throughout the slab, and no internal heat generation is considered. With these simplifications, the system of equations to be solved (using Equations 3.1.35 through 3.1.37) is

$$\begin{bmatrix}
(1+2\mathrm{Fo}+2\mathrm{Fo}Bi) & -\mathrm{Fo} & 0 & 0 & \cdots & 0 \\
-\mathrm{Fo} & (1+2\mathrm{Fo}) & -\mathrm{Fo} & 0 & \cdots & 0 \\
\cdots & -\mathrm{Fo} & (1+2\mathrm{Fo}) & -\mathrm{Fo} & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & -\mathrm{Fo} & (1+2\mathrm{Fo}) & -\mathrm{Fo} \\
\cdots & \cdots & \cdots & \cdots & -2\mathrm{Fo} & (1+2\mathrm{Fo})
\end{bmatrix}
\begin{bmatrix} T'_1 \\ T'_2 \\ T'_3 \\ \cdots \\ T'_{n-1} \\ T'_n \end{bmatrix}
=
\begin{bmatrix} 2\mathrm{Fo}BiT'_\infty+T_1 \\ T_2 \\ T_3 \\ \cdots \\ T_{n-1} \\ T_n \end{bmatrix} \tag{3.1.38}$$

The time is then advanced by Δt and the temperatures are again found by simultaneous solution.

The finite-difference approach just outlined using the backward difference for the time derivative is termed the *implicit* technique, and it results in an $n \times n$ system of linear simultaneous equations. If the forward difference is used for the time derivative, then only one unknown will exist in each equation. This gives rise to what is called an *explicit* or "marching" solution. While this type of system is more straightforward to solve because it deals with only one equation at a time with one unknown, a *stability criterion* must be considered which limits the time step relative to the distance step.

Two- and three-dimensional problems are handled in conceptually the same manner. One-dimensional heat fluxes between adjoining nodes are again considered. Now there are contributions from each of the dimensions represented. Details are outlined in the book by Jaluria and Torrance (2003).

Defining Terms

Biot number: Ratio of the internal (conductive) resistance to the external (convective) resistance from a solid exchanging heat with a fluid.

Explicit formulation: All temperatures in the differenced form of the transient heat conduction equation are at the current time step except for in the time derivative, where the future time value is given.

Fin: Additions of material to a surface to increase area and thus decrease the external thermal resistance from convecting and/or radiating solids.

Fin effectiveness: Ratio of the actual heat transfer from a fin to the heat transfer from the same cross-sectional area of the wall without the fin.

Fin efficiency: Ratio of the actual heat transfer from a fin to the heat transfer from a fin with the same geometry but completely at the base temperature.

Fourier's law: The fundamental law of heat conduction. Relates the local temperature gradient to the local heat flux, both in the same direction.

Heat conduction equation: A partial differential equation in temperature, spatial variables, time, and properties that, when solved with appropriate boundary and initial conditions, describes the variation of temperature in a conducting medium.

Implicit formulation: All temperatures in the differenced form of the transient heat conduction equation are at the new time step except for in the time derivative, where the current time value is given.

Overall heat-transfer coefficient: The analogous quantity to the heat transfer coefficient found in convection (Newton's law of cooling) that represents the overall combination of several thermal resistances, both conductive and convective.

Thermal conductivity: The temperature dependent property of a material that relates a temperature gradient to a heat flux.

REFERENCES

Abramowitz, M. and Stegun, I., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Applied Mathematics Series 55, U.S. Government Printing Office, Washington, DC, 1964.

Arpaci, V., *Conduction Heat Transfer*, Addison-Wesley, Reading, MA, 1966.

Carslaw, H. S. and Jaeger, J. C., *Conduction of Heat in Solids*, 2nd ed., Oxford University Press, London, 1959.

Guyer, E. (Ed.), Part 3, Thermal insulations, in *Handbook of Applied Thermal Design*, McGraw-Hill, New York, 1989.

Jaluria, Y. and Torrance, K., *Computational Heat Transfer*, 2nd ed., Hemisphere, New York, 2003.

Sparrow, E., Reexamination and correction of the critical radius for radial heat conduction, *AIChE J.*, 16(1), 149, 1970.

For Further Information

The references listed earlier will give the reader an excellent introduction to analytical formulation and solution (Arpaci), material properties (Guyer), and numerical formulation and solution (Jaluria and Torrance). Current developments in conduction heat transfer appear in several publications, including the *Journal of Heat Transfer, International Journal of Heat and Mass Transfer*, and *Numerical Heat Transfer*.

3.2 CONVECTION HEAT TRANSFER

3.2.1 Natural Convection

Swati A. Patel, Raj P. Chhabra, George D. Raithby, and K.G. Terry Hollands

The presence of temperature gradient in a fluid coupled with the fluid motion results in heat transfer from the region of high temperature to that of lower temperature principally by convection. It is customary to classify the mode of heat transfer as free or natural convection when the fluid motion is solely due to the density difference which itself stems from the temperature variation in the fluid. On the other hand, when the fluid motion is induced by a mechanical device (pump, fan, blower, mechanical stirrer, rotation, etc.), heat transfer occurs in the forced convection regime. In practice, both these modes contribute in varying proportions to the overall heat transfer depending upon the strength of the buoyancy-induced flow vis-à-vis that of the externally imposed flow. The relative importance of the two modes is ascertained via the use of the familiar Richardson number, $\mathrm{Ri}=\mathrm{Gr}/\mathrm{Re}^2$, which is really a measure of the ratio of the buoyancy to inertial forces. Naturally, vanishingly small values of the Richardson number ($\mathrm{Ri} \to 0$) correspond to the forced convection regime, whereas large values of the Richardson number ($\mathrm{Ri} \to \infty$) denote the other limit of the pure natural convection. Indeed, the values of the Richardson number on the order of unity, $\mathrm{Ri} \sim \mathrm{O}(1)$, denote the case when the two flow velocities are of comparable magnitude. Thus, heat transfer occurs in the mixed convection regime for the intermediate values of the Richardson number. Depending upon the orientations of these two velocities, it is customary to further classify the mixed-convection regime into three types, namely aiding-buoyancy, opposing-buoyancy, and cross-buoyancy, as shown schematically in Figure 3.2.1 for the case of a heated horizontal cylinder. The limiting case of the natural convection is dealt with in this section, whereas that of the forced and mixed convection is treated in Section 3.2.2, and the analogous information for non-Newtonian fluids is presented in Section 3.2.4.

As noted earlier, natural convection heat transfer occurs when the fluid motion is induced solely by density differences that are themselves caused by heating or by the concentration gradient of a solute. The schematics of the free convection is shown in Figure 3.2.2 along with the nomenclature used in this section, where a body at surface temperature T_s is immersed in a quiescent fluid at temperature $T_\infty < T_s$, thereby resulting in heat transfer at the rate of q from the heated object to the ambient fluid.

Notwithstanding the fundamental significance of the temperature and velocity profiles especially in the processing of temperature-sensitive materials (such as that encountered in food-, pharma- and personal care product sectors), it is readily conceded that the need to estimate the overall heat-transfer coefficient to size the process equipment and/or to modulate the rate of heating or cooling in a given application frequently arises in process engineering calculations. This information is conveniently expressed in terms of the dimensionless correlations for Nusselt number. Over the years, based on a combination of the boundary layer approximations, numerical solutions of the coupled momentum and energy equations, and experimental results aided by dimensional considerations, reliable correlations have been developed which span wide ranges of conditions, in terms of the range of Grashof number and Prandtl number or Rayleigh numbers for scores of body shapes and type of boundary conditions (constant temperature or constant heat flux or mixed), etc. An excellent review of the pertinent literature in this field is available (Martynenko and Khramtsov, 2005).

In this section, some of the widely used correlations for the average Nusselt number are provided from which the heat-transfer rate q from surface area A_s can be estimated for a range of shapes for which reliable results have accrued in the literature. The Nusselt number is defined as

$$\mathrm{Nu} = \frac{\bar{h}_c L_c}{k} = \frac{q L_c}{A_s \Delta T k}, \tag{3.2.1}$$

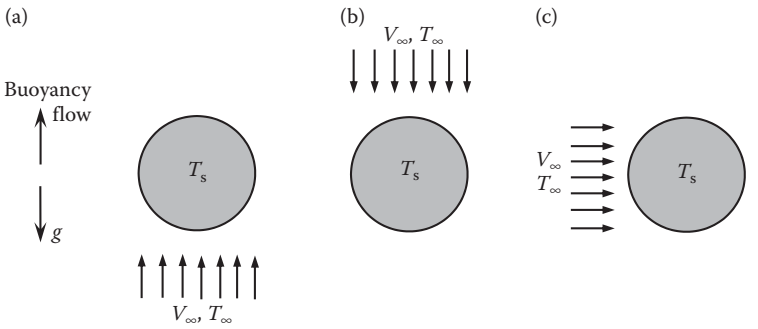


FIGURE 3.2.1 Schematics of mixed-convection regimes (a) aiding-buoyancy, (b) opposing-buoyancy, and (c) cross-buoyancy.

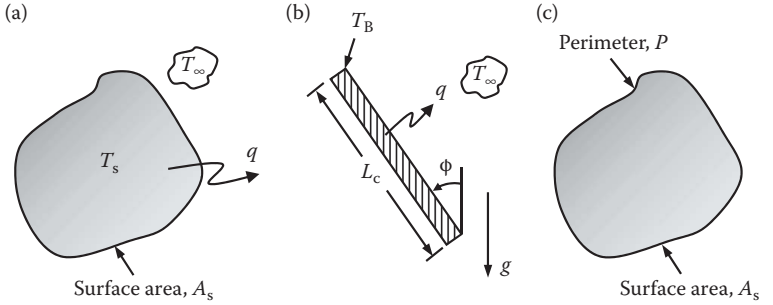


FIGURE 3.2.2 Nomenclature for external heat transfer (a) general sketch, (b) for a tilted flat plate, and (c) defines the length scale for horizontal surfaces.

where $\Delta T = T_s - T_\infty$ is the temperature difference driving the heat transfer. For the simplest case of the constant thermophysical properties (except density), infinite extent of fluid and negligible viscous dissipation, a dimensional analysis of such problems leads to the following general functional relationship:

$$\text{Nu} = f(\text{Ra or Gr, Pr, geometric shape, boundary conditions}) \tag{3.2.2}$$

For given thermal boundary conditions (e.g., isothermal wall and uniform T_∞), and for a given geometry (e.g., a cube or a sphere or a cylinder), Equation 3.2.2 states that the value of the average Nusselt number, Nu, depends only on the Rayleigh number, Ra, and Prandtl number, Pr, or combinations thereof. The length scales that appear in the definitions of the Nusselt number, Nu, and Rayleigh number, Ra, are defined, for each geometry considered, in a separate figure. For small values of ΔT, i.e., $(T_s - T_\infty)$, the fluid properties are generally evaluated at the mean film temperature, i.e., $T_f = (T_s + T_\infty)/2$. The exception is that β, the temperature coefficient of volume expansion, is evaluated at T_∞ for external natural convection (Figures 3.2.2 through 3.2.4) in a gaseous medium. Additional correction is needed when the temperature dependence of the density cannot be well approximated by the familiar Boussinesq relation (Kays et al., 2004).

The functional dependence of the average Nusselt number on the Prandtl number, Pr, is approximately independent of the geometry, and the following Pr-dependent function has been found to be useful for consolidating results for a range of shapes, at least in the laminar heat-transfer regime (Churchill and Usagi, 1972):

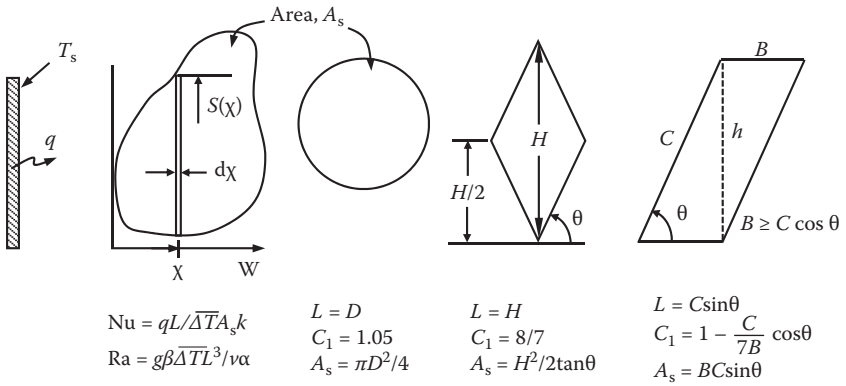


FIGURE 3.2.3 Nomenclature for heat transfer from planar surfaces of different shapes.

$$\bar{C}_\ell = \frac{0.671}{\left(1+\left(\frac{0.492}{\mathrm{Pr}}\right)^{9/16}\right)^{4/9}} \tag{3.2.3}$$

C_t^V and C_t^H functions, defined in Equations 3.2.4 and 3.2.5, will be useful for turbulent heat transfer:

$$C_t^V = \frac{0.13\mathrm{Pr}^{0.22}}{\left(1+0.61\mathrm{Pr}^{0.81}\right)^{0.42}} \tag{3.2.4}$$

$$C_t^H = 0.14\left(\frac{1+0.0107\mathrm{Pr}}{1+0.01\mathrm{Pr}}\right). \tag{3.2.5}$$

The superscripts V and H refer to the vertical and horizontal surface orientation.

The Nusselt numbers for fully laminar and fully turbulent heat-transfer regimes are denoted by Nu_ℓ and Nu_t, respectively. Once known, these two extreme values are blended (Churchill and Usagi, 1972) as follows to obtain the equation for the Nusselt number, Nu spanning the full range of laminar and turbulent flow regimes as

$$\mathrm{Nu} = \left(\left(\mathrm{Nu}_\ell\right)^m + \left(\mathrm{Nu}_t\right)^m\right)^{1/m}. \tag{3.2.6}$$

The value of the blending parameter m in Equation 3.2.6 depends on the body shape and orientation.

The equation for Nu_ℓ in this section is usually expressed in terms of Nu^T, the Nusselt number that would be valid if the thermal boundary layer were thin ($\mathrm{Pr} \to \infty$ or $\mathrm{Gr} \to \infty$). The difference between Nu_ℓ and Nu^T accounts for the effect of the large boundary layer thicknesses encountered in natural convection.

It is assumed that the wall temperature of a body exceeds the ambient fluid temperature ($T_s > T_\infty$). For $T_s < T_\infty$, the same correlations apply with $(T_\infty - T_s)$ replacing $(T_s - T_\infty)$ for a geometry that is rotated 180° relative to the gravitational vector; for example, the correlations for a horizontal heated upward-facing flat plate apply to a cooled downward-facing flat plate of the same planform.

Correlations for External Natural Convection

This section deals with problems where the body shapes shown in Figures 3.2.2 through 3.2.4 are heated while immersed in a quiescent cold fluid. Different cases are detailed as follows:

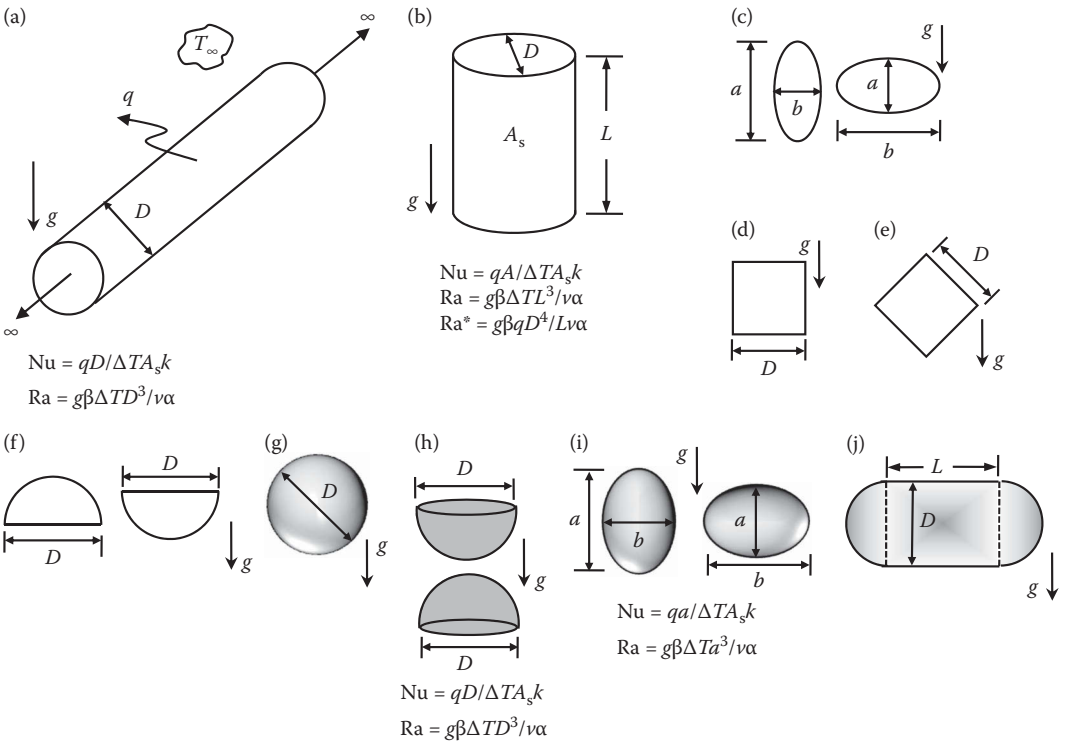


FIGURE 3.2.4 Definitions for computing heat transfer from a long circular cylinder (a), from the lateral surface of a vertical circular cylinder (b), from a long elliptical cylinder (c), square cylinder (d), inclined square cylinder (e), semicircular cylinder (f), from a sphere (g), from a hemisphere (h), spheroid (i), and from a compound body (j).

1. *Isothermal vertical (ϕ=0) flat plate, Figure 3.2.2b.* For heat transfer from a vertical plate (Figure 3.2.2b), for $1 < \mathrm{Ra} < 10^{12}$, the relevant expressions for the Nusselt number are written as

$$\mathrm{Nu}^{\mathrm{T}} = \bar{C}_\ell \mathrm{Ra}^{1/4}$$

$$\mathrm{Nu}_\ell = \frac{2.0}{\ln\left(1 + \dfrac{2.0}{\mathrm{Nu}^{\mathrm{T}}}\right)} \tag{3.2.7}$$

$$\mathrm{Nu}_{\mathrm{t}} = \frac{C_{\mathrm{t}}^{\mathrm{V}} \mathrm{Ra}^{1/3}}{\left(1 + 1.4 \times 10^9 \dfrac{\mathrm{Pr}}{\mathrm{Ra}}\right)}$$

$\bar{C}_\ell$ and $C_{\mathrm{t}}^{\mathrm{V}}$ are given by Equations 3.2.3 and 3.2.4. In a given situation, the value of the Nusselt number, Nu, is obtained by substituting the expressions of the Nu_ℓ and Nu_{t} from Equation 3.2.7 into Equation 3.2.6 with m=6.

2. *Vertical flat plate with uniform heat flux, Figure 3.2.2b.* If the plate surface has a constant (known) heat flux (q) prescribed on its surface, rather than being isothermal, the objective is to calculate the average temperature difference, $\overline{\Delta T}$, between the plate and the fluid. For this situation, and for $15 < \mathrm{Ra}^* < 10^5$,

$$\mathrm{Nu}^{\mathrm{T}} = \bar{G}_{\mathrm{l}}(\mathrm{Ra}^{*})^{1/5}$$

$$\mathrm{Nu}_{\ell} = \frac{1.83}{\ln\left(1+\dfrac{1.83}{\mathrm{Nu}^{\mathrm{T}}}\right)} \tag{3.2.8a}$$

$$\mathrm{Nu}_{\mathrm{t}} = (C_{\mathrm{t}}^{\mathrm{V}})^{3/4}(\mathrm{Ra}^{*})^{1/4}$$

$$\bar{G}_{\ell} = \frac{6}{5}\left(\frac{\mathrm{Pr}}{4+9\sqrt{\mathrm{Pr}}+10\mathrm{Pr}}\right)^{1/5}. \tag{3.2.8b}$$

The modified Rayleigh number, Ra*, is defined as $\mathrm{Ra}^{*} = \mathrm{Nu}\times\mathrm{Ra} = g\beta q L^{4}/\left(A_{\mathrm{s}}\nu\alpha k\right)$, and $C_{\mathrm{t}}^{\mathrm{V}}$ is given by Equation 3.2.4. Find Nu by inserting these expressions for Nu_{ℓ} and Nu_{t} into Equation 3.2.6 with $m=6$. The $\bar{G}_{\ell}$ expression given by Equation 3.2.8a is due to Fujii and Fujii (1976).

3. *Horizontal upward-facing ($\phi=90°$) plates, Figure 3.2.2b.* For horizontal isothermal surfaces of various planforms, correlations are given in terms of a length scale L^{*} (Goldstein et al., 1973), defined in Figure 3.2.2c. For $\mathrm{Ra}\geq 1$,

$$\mathrm{Nu}^{\mathrm{T}} = 0.835\bar{C}_{\ell}\mathrm{Ra}^{1/4}$$

$$\mathrm{Nu} = \frac{1.4}{\ln\left(1+\dfrac{1.4}{\mathrm{Nu}^{\mathrm{T}}}\right)} \tag{3.2.9}$$

$$\mathrm{Nu}_{\mathrm{t}} = C_{\mathrm{t}}^{\mathrm{H}}\mathrm{Ra}^{1/3}.$$

The value of the average Nusselt number, Nu, is obtained by substituting Nu_{ℓ} and Nu_{t} from Equation 3.2.9 into Equation 3.2.6 with $m=10$. For nonisothermal surfaces, replace ΔT by $\overline{\Delta T}$.

4. *Horizontal downward-facing ($\phi=-90°$) plates, Figure 3.2.2c.* For horizontal downward-facing plates of various planforms, the main buoyancy force is into the plate so that only a very weak force drives the fluid along the plate; for this reason, only laminar flows have been investigated. For this case, the following equation applies for $\mathrm{Ra} < 10^{10}$, $\mathrm{Pr}\geq 0.7$:

$$\mathrm{Nu}^{\mathrm{T}} = H_{\ell}\mathrm{Ra}^{1/5}$$

$$H_{\ell} = \frac{0.527}{\left[1+\left(\dfrac{1.9}{\mathrm{Pr}}\right)^{9/10}\right]^{2/9}} \tag{3.2.10}$$

$$\mathrm{Nu}_{\ell} = \frac{2.45}{\ln\left(1+\dfrac{2.45}{\mathrm{Nu}^{\mathrm{T}}}\right)}.$$

The expression for H_{ℓ} is based on the results of Fujii et al. (1973).

5. *Inclined plates, downward facing ($-90°\leq\phi\leq 0$), Figure 3.2.2b.* First calculate q from *Case 1* with g replaced by $g\cos\phi$; then calculate q from *Case 4* (horizontal plate) with g replaced by $g\sin(-\phi)$, and use the maximum of these two values of q for further calculations.
6. *Inclined plates, upward facing ($0\leq\phi\leq 90$), Figure 3.2.2b.* First calculate q from *Case 1* with g replaced by $g\cos\phi$; then calculate q from *Case 3* with g replaced by $g\sin\phi$, and use the maximum of these two values of q for further calculations.

7. *Vertical and tilted isothermal plates of various planforms, Figure 3.2.3.* The line of constant χ in Figure 3.2.3 is the line of steepest ascent on the plate. Provided all such lines intersect the plate edges just twice, as shown in the figure, the thin-layer (Nu^T) heat transfer can be found by subdividing the body into strips of width $\Delta\chi$, calculating the heat transfer from each strip, and adding them up to obtain the overall heat transfer. For laminar flow from an isothermal vertical plate, this results in the following expression:

$$Nu^T = C_1 \bar{C}_\ell Ra^{1/4}, \tag{3.2.11}$$

where

$$C_1 \equiv \frac{L^{1/4}}{A} \int_0^W S^{3/4}\, d\chi.$$

Symbols are defined in Figure 3.2.3, along with L and calculated C_1 values for some plate shapes. If the plate is vertical, follow the procedure in *Case 1* earlier (isothermal vertical flat plate) except replacing the expression for Nu^T in Equation 3.2.7 by Equation 3.2.11. If the plate is tilted, follow the procedure described in *Case 5* or *6* (as appropriate) but again use Equation 3.2.11 for Nu^T in Equation 3.2.7.

8. *Horizontal cylinders, Figure 3.2.4a.* For a long, horizontal circular cylinder, one can use the following expressions for Nu_ℓ and Nu_t:

$$\begin{aligned} Nu^T &= 0.772 \bar{C}_\ell Ra^{1/4} \\ Nu_\ell &= \frac{2f}{\left(1 + \dfrac{2f}{Nu^T}\right)} \\ Nu_t &= \bar{C}_t Ra^{1/3}. \end{aligned} \tag{3.2.12}$$

$\bar{C}_t$ is given in Table 3.2.1. For $Ra > 10^{-2}$, $f = 0.8$ can be used, but for the range $10^{-10} < Ra < 10^{-2}$, use $f = 1 - 0.13/(Nu^T)^{0.16}$. To find Nu, the values of Nu_ℓ and Nu_t from Equation 3.2.12 are substituted into Equation 3.2.6 with $m = 15$ (Clemes et al., 1994).

The following correlation due to Churchill and Chu (1975) is convenient and fairly reliable ($Ra > 10^{-6}$) over the laminar and turbulent flow regimes for an isothermal horizontal cylinder:

$$Nu = 0.60 + \frac{0.387\, Ra^{1/6}}{\left\{\left[1 + \left(\dfrac{0.559}{Pr}\right)^{9/16}\right]^{16/9}\right\}^{1/6}}. \tag{3.2.13}$$

TABLE 3.2.1
Values of $\bar{C}_t$ for Various Shapes and Prandtl Numbers

Pr→	0.01	0.022	0.10	0.71	2.0	6.0	50	100	2000
Horizontal cylinder	0.077	0.81	0.90	0.103	0.108	0.109	0.100	0.097	0.088
Sphere	0.074	0.078	0.088	0.104	0.110	0.111	0.101	0.97	0.086

The corresponding expression for the constant heat flux condition spanning laminar and turbulent flow regimes and $0.022 \le \mathrm{Pr} \le 7640$, $\mathrm{Ra} > 1$ is given as (Churchill and Thelen, 1975)

$$\sqrt{\mathrm{Nu}} = 0.60 + \frac{\left(0.0033\mathrm{Ra}\right)^{1/6}}{\left\{\left[1+\left(\frac{0.44}{\mathrm{Pr}}\right)^{9/16}\right]^{16/9}\right\}^{1/6}}. \tag{3.2.14}$$

The physical properties are evaluated at the mean film temperature.

9. *Vertical cylinders (*$\phi = 90°$*), Figure 3.2.4b.* For high values of the Rayleigh number and cylinders of large diameter, the heat transfer from a vertical cylinder approaches that for a vertical flat plate. This is so due to the fact that the boundary layer thickness is much smaller than the diameter of the cylinder and one can thus neglect the curvature effects. Let the Nu^{T} and Nu_{ℓ} equations for a vertical flat plate of height L, Equation 3.2.7, be rewritten here as $\mathrm{Nu}_{\mathrm{p}}^{\mathrm{T}}$ and Nu_{p}, respectively. At smaller Ra and diameter, transverse curvature plays a role, which is accounted for in the following equations:

$$\mathrm{Nu}_{\ell} = \frac{0.9\xi\mathrm{Nu}_{\mathrm{p}}}{\ln\left(1+0.9\xi\right)}, \tag{3.2.15}$$

where

$$\xi = \frac{2L/D}{\mathrm{Nu}_{\mathrm{p}}^{\mathrm{T}}}.$$

These equations are valid for purely laminar flow. To obtain Nu, blend Equation 3.2.15 for Nu_{ℓ} with Equation 3.2.7 for Nu_{t} using Equation 3.2.6 with $m = 10$.

10. *Horizontal elliptical cylinder*, *Figure 3.2.4c*. For a long isothermal elliptical cylinder with an aspect ratio E (=a/b) (axis along the flow to the axis normal to the flow), the following expression based on $L_{\mathrm{c}} = 2b$ (characteristic linear scale) can be used to evaluate the average heat-transfer coefficient in the fully laminar flow regime (Sasmal and Chhabra, 2012a):

$$\mathrm{Nu} = 0.833(\mathrm{Ra}^{1/4})^{0.89} E^{a_0}. \tag{3.2.16}$$

The values of coefficient, $a_0 = -0.116$ $(0.2 \le E < 1)$ and $a_0 = -0.321$ $(1 < E \le 5)$ in this expression fit the numerical data in the range of Rayleigh number, $7.2 \le \mathrm{Ra} \le 10^7$ and Prandtl number, $0.72 \le \mathrm{Pr} \le 100$.

11. *Horizontal square cylinder*, *Figure 3.2.4d*. For a long isothermal square cylinder of side D, one can use the following expression for Nu (Sasmal and Chhabra, 2011):

$$\mathrm{Nu} = 0.605(\mathrm{Ra}^{1/4})^{0.924}. \tag{3.2.17}$$

Equation 3.2.17 approximates the numerical results for the average heat-transfer coefficient in the range, $7.2 \le \mathrm{Ra} \le 10^7$ and Prandtl number, $0.72 \le \mathrm{Pr} \le 100$ with an error of 11%.

12. *Horizontal inclined square cylinder*, *Figure 3.2.4e*. For a long tilted isothermal square cylinder of side length D, the heat-transfer coefficient in the laminar regime is given by the following expression (based on the characteristic linear scale, $L_{\mathrm{c}} = D$) over the range of Rayleigh number, $7.2 \le \mathrm{Ra} \le 10^7$ and Prandtl number, $0.72 \le \mathrm{Pr} \le 100$ (Sasmal and Chhabra, 2012b):

$$\mathrm{Nu} = 0.761(\mathrm{Ra}^{1/4})^{0.924}. \tag{3.2.18}$$

13. *Horizontal semicircular cylinder, Figure 3.2.4f.* For a long isothermal semicircular cylinder with its flat surface oriented in the upward direction, one can use the following expression for Nu (Chandra and Chhabra, 2012):

$$\mathrm{Nu} = 0.93(\mathrm{Ra}^{1/4})^{0.79}. \tag{3.2.19}$$

In the case of the opposite configuration of the flat surface oriented downward, the average heat-transfer coefficient can be calculated from the following expression (Tiwari and Chhabra, 2013):

$$\mathrm{Nu} = 0.72(\mathrm{Ra}^{1/4})^{0.90}. \tag{3.2.20}$$

Equations 3.2.19 and 3.2.20 are based on the numerical results over the range of Rayleigh number as $7.2 \le \mathrm{Ra} \le 10^7$ and Prandtl number as $0.72 \le \mathrm{Pr} \le 100$ and use the diameter as the characteristic linear scale in the definitions of the Nu and Ra.

14. *Spheres, Figure 3.2.4g.* For spheres use Equation 3.2.6, with $m=6$, and with

$$\mathrm{Nu}_\ell = 2 + 0.878\bar{C}_\ell \mathrm{Ra}^{1/4} \qquad \text{and} \qquad \mathrm{Nu}_\mathrm{t} = \bar{C}_\mathrm{t}\mathrm{Ra}^{1/3}. \tag{3.2.21}$$

The values of $\bar{C}_\mathrm{t}$ are presented in Table 3.2.1.
An alternative correlation of the average Nusselt number for an isothermal sphere (Jafarpur and Yovanovich, 1992) is given as

$$\mathrm{Nu} = \frac{0.589\mathrm{Ra}^{1/4}}{\left[1+\left(\dfrac{0.5}{\mathrm{Pr}}\right)^{9/16}\right]^{4/9}}, \tag{3.2.22}$$

which asymptotically reduces to the following limiting results:

$$\mathrm{Pr} \to 0 : \mathrm{Nu} = 0.714\mathrm{Ra}^{1/4}\mathrm{Pr}^{1/4}. \tag{3.2.23}$$

$$\mathrm{Pr} \to \infty : \mathrm{Nu} = 0.6\mathrm{Ra}^{1/4}. \tag{3.2.24}$$

15. *Hemisphere, Figure 3.2.4h.* The heat-transfer coefficient for an inverted or upright hemisphere can be evaluated from the following expressions (Sasmal and Chhabra, 2014):

$$\text{for } \mathrm{Ra} < 10^4, \mathrm{Nu} = 2.255 + a_1(\mathrm{Ra}^{1/4})^{b_1}. \tag{3.2.25a}$$

$$\text{for } 10^4 < \mathrm{Ra} \le 10^7, \mathrm{Nu} = a_2(\mathrm{Ra}^{1/4})^{b_2}. \tag{3.2.25b}$$

The values of the fitting coefficients a_1, b_1, a_2, and b_2 for the inverted configuration, respectively, are 0.332, 1.2, 1.102, and 0.815 and for an upright hemisphere, the corresponding values are 0.343, 1.18, 1.14, and 0.8, respectively.

16. *Spheroids, Figure 3.2.4i.* For spheroids of aspect ratio, $E=a/b$ (ratio of the axis along the direction of the gravity to the axis transverse to the direction of the gravity), one can use the following expressions for Nu (Gupta et al., 2014):

$$\text{for } \mathrm{Ra} < 10^4, \mathrm{Nu} = \mathrm{Nu}_\infty + 0.706(\mathrm{Ra}^{1/4})^{1.109}E^{-0.166}. \tag{3.2.26a}$$

$$\text{for } 10^4 < \mathrm{Ra} \le 10^7, \mathrm{Nu} = 1.32(\mathrm{Ra}^{1/4})^{0.813}(2.34 - E^{0.368}). \tag{3.2.26b}$$

TABLE 3.2.2
Values of Nusselt Number in the Conduction Limit (Nu_∞)

E	0.2	0.5	1	2	5
Nu_∞	5.37	3.08	2	1.2	0.52

The values of Nu_∞ (conduction limit) are tabulated in Table 3.2.2. Equations 3.2.26a and 3.2.26b fit the numerical results with average errors of ~6% and ~11%, respectively. The characteristic linear dimension in this case is 2*b*.

17. *Combined shapes, Figure 3.2.4j.* For combined shapes, such as the cylinder in Figure 3.2.4j with spherical end caps, calculate the heat transfer from the cylinder of length L (*Case 8*), the heat transfer from a sphere of diameter D (*Case 14*) and add to obtain the total transfer. Other shapes can be treated in a similar manner.

    The effect of viscous dissipation on heat transfer from a horizontal cylinder has been studied by Fand et al. (1977), whereas the influence of very small Grashof numbers (microgravity situations) has been studied by Mograbi et al. (2002). In view of the fact that it is unlikely that the numerical and/or experimental studies will ever span all possible body shapes of interest, some attempts have also been made to introduce shape factors, characteristic linear dimension (like $\sqrt{A_s}$) combined with the consideration of the orientation to develop an approach with universal appeal. This body of literature has been reviewed by Lee et al. (1991) and Eslami and Jafarpur (2012).

Correlations for Open Cavities

There are situations where the fluid motion is constrained by the presence of solid boundaries, which themselves may be at different temperatures both of which, in turn, may deviate from the ambient temperature. A representative selection of examples of this class of problems is shown in Figure 3.2.5. The solid walls partially enclose a fluid region (cavity), where boundary openings permit the fluid to enter and leave the system. Faraway upstream from its point of entry, the fluid is at the ambient temperature, T_∞. Since access of the ambient fluid to the heated surfaces is restricted, some of the heated surface is not in contact with the cool ambient to which heat can be transferred. Therefore, this part of the surface does not contribute directly to the overall heat transfer. As the sizes of the boundary openings are increased, these situations approach the problems described in Figure 3.2.4 (e.g., value of L in Figure 3.2.5a or the value of b in Figures 3.2.5c and 3.2.5d, etc.), the previous class of problems is approached; for example, when the plate spacing in Figure 3.2.5a (Case 18) becomes very large, the heat transfer from each vertical surface is given by *Case 1*.

18. *Isothermal vertical channels, Figures 3.2.5a and b.* Figure 3.2.5a shows an open cavity bounded by two vertical walls and open at the top and bottom. The large opposing plates are isothermal, maintained at temperatures T_1 and T_2, respectively, and the spacing between these plates is small. $\overline{\Delta T}$ is the average temperature difference between the plates and the fluid temperature T_∞ as shown in Figure 3.2.4a, but T_1 and T_2 must not straddle T_∞. For this case,

$$\mathrm{Nu} = \left(\left(\frac{\mathrm{Ra}}{f\,\mathrm{Re}} \right)^m + \left(C_1 \bar{C}_\ell \mathrm{Ra}^{1/4} \right)^m \right)^{1/m} \qquad \mathrm{Ra} \le 10^5, \tag{3.2.27}$$

where f Re is the product of friction factor and Reynolds number for fully developed flow through the gap between the two plates, and C_1 is a constant that accounts for the

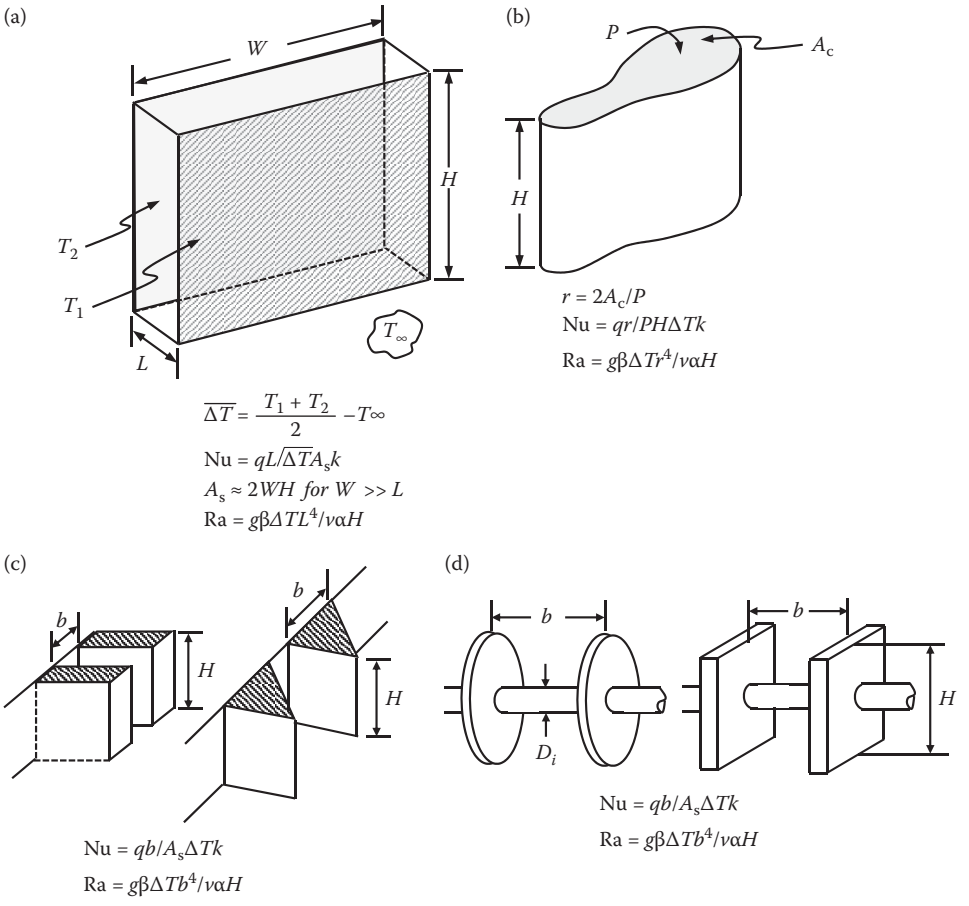


FIGURE 3.2.5 Nomenclature for various open-cavity problems.

augmentation of heat transfer, relative to a vertical flat plate (*Case 1*), due to the chimney effect. The f Re factor accounts for the cross-sectional shape (Elenbaas, 1942a). Symbols are defined in Figure 3.2.5a and b; in the Nu equation, q is the total heat transferred to the ambient fluid from all heated surfaces.

For the parallel plate channel shown in Figure 3.2.5a, use f Re = 24, $m = -1.9$, and for gases $C_1 \approx 1.2$. It should be noted, however, that C_1 must approach 1.0 as Pr increases or as the plate spacing increases. For channels of circular cross section (Figure 3.2.5b), f Re = 16, $m = -1.03$, and for gases $C_1 \approx 1.17$. For other cross-sectional shapes like the square (f Re = 14.23), hexagonal (f Re = 15.05), or equilateral triangle (f Re = 13.3), use Equation 3.2.27 with the appropriate f Re, and with $m = -1.5$, and $C_1 \approx 1.2$ for gases.

The heat-transfer per unit cross-sectional area, q/A_c, for a given channel length H and temperature difference, passes through a maximum at approximately Ra_{max}, where

$$\mathrm{Ra}_{max} = \left(\frac{f\,\mathrm{Re} C_1 \bar{C}_\ell}{2^{1/m}} \right)^{4/3}. \tag{3.2.28}$$

The limiting values of Ra_{max} provide the value of hydraulic radius $r = 2A_c/P$ at this maximum.

19. *Isothermal triangular fins, Figure 3.2.5c*. For a large array of triangular fins (Karagiozis et al., 1994) in air, for $0.4 < \mathrm{Ra} < 5 \times 10^5$:

$$\mathrm{Nu} = \bar{C}_\ell \mathrm{Ra}^{1/4}\left[1+\left(\frac{3.26}{\mathrm{Ra}^{0.21}}\right)^3\right]^{-1/3}. \tag{3.2.29}$$

In this equation, b is the average fin spacing (Figure 3.2.5c), defined such that bL is the cross-sectional flow area between two adjacent fin surfaces up to the plane of the fin tips. For Ra < 0.4, Equation 3.2.29 underestimates the convective heat transfer. When such fins are mounted horizontally (vertical baseplate, but the fin tips are horizontal), there is a substantial reduction in the convective heat transfer (Karagiozis et al., 1994).

20. *U-channel fins, Figure 3.2.5c*. For the fins most often used as heat sinks, there is some uncertainty about the heat transfer at low Rayleigh numbers. By using a conservative approximation applying for Ra < 100 (that underestimates the real heat transfer), the following equation can be used:

$$\mathrm{Nu} = \left[\left(\frac{\mathrm{Ra}}{24}\right)^{-2} + \left(C_1 \bar{C}_\ell \mathrm{Ra}\right)^{-2}\right]^{-0.5}. \tag{3.2.30}$$

For air, the value of C_1 depends on the aspect ratio of the fin as follows (Karagiozis, 1991):

$$C_1 = \left[1+\left(\frac{H}{b}\right), 1.16\right]_{\min}. \tag{3.2.31}$$

Equation 3.2.30 agrees well with measurements for Ra > 200, but for smaller Ra it severely underpredicts the value of the Nusselt number well below data because the leading term does not account for heat transfer from the fin edges and for three-dimensional conduction from the entire array.

21. *Circular fins on a horizontal tube, Figure 3.2.5d*. For heat transfer from an array of circular fins (Edwards and Chaddock, 1963), for $H/D_i = 1.94$, $5 < \mathrm{Ra} < 10^4$, and for air,

$$\mathrm{Nu} = 0.125\mathrm{Ra}^{0.55}\left[1-\exp\left(-\frac{137}{\mathrm{Ra}}\right)\right]^{0.294}. \tag{3.2.32}$$

A more general, but also more complex, relation for this case is reported by Raithby and Hollands (1985).

22. *Square fins on a horizontal tube, Figure 3.2.5d*. Heat transfer (Elenbaas, 1942b) from square fins (excluding the cylinder that connects them) is correlated for gases by

$$\mathrm{Nu} = \left[\left(\frac{\mathrm{Ra}^{0.89}}{18}\right)^m + \left(0.62\mathrm{Ra}^{1/4}\right)^m\right]^{1/m} \quad \text{where} \quad m = -2.7. \tag{3.2.33}$$

Heat Transfer in Enclosures

This section deals with cavities where the bounding walls are entirely closed, so that no mass can enter or leave the cavity. The fluid motion inside the cavity is driven by natural convection, which enhances the heat transfer among the interior surfaces that bound the cavity and the resulting value of the Nusselt number is in excess of its conduction limit.

23. *Extensive horizontal layers, Figure 3.2.6a*. If the heated plate, in a horizontal parallel plate cavity, is on the top ($\theta = 180°$), heat transfer occurs by conduction alone, so that Nu = 1. For heat transfer from below ($\theta = 0°$) (Hollands, 1984):

$$\mathrm{Nu} = 1 + \left[1 - \frac{1708}{\mathrm{Ra}}\right]^{\bullet}\left[k_1 + 2\left(\frac{\mathrm{Ra}^{1/3}}{k_2}\right)^{1-\ln\left(\mathrm{Ra}^{1/3}/k_2\right)}\right] + \left[\left(\frac{\mathrm{Ra}}{5830}\right)^{1/3} - 1\right]^{\bullet}, \qquad (3.2.34)$$

where

$$[x]^{\bullet} = (x,0)_{\max}; \qquad k_1 = \frac{1.44}{1 + \left(\frac{0.018}{\mathrm{Pr}}\right) + \left(\frac{0.00136}{\mathrm{Pr}^2}\right)}; \qquad k_2 = 75\exp(1.5\mathrm{Pr}^{-1/2}). \quad (3.2.35)$$

The equation has been validated for $\mathrm{Ra} < 10^{11}$ for water, up to $\mathrm{Ra} < 10^{8}$ for air, and over a smaller Ra range for other fluids. Equation 3.2.34 applies to extensive layers: $W/L \geq 5$. Correlations for nonextensive layers are provided by Raithby and Hollands (1985).

24. *Vertical layers, Figure 3.2.6a, with* $\theta = 90°$. $W/L > 5$. For a vertical, gas-filled ($\mathrm{Pr} \approx 0.7$) cavity with $H/L \geq 5$, the following equation closely fits the data, for example, that of Shewen et al. (1996) for Ra $(H/L)^3 \leq 5 \times 10^{10}$ and $H/L \geq 40$:

$$\mathrm{Nu}_1 = \left[1 + \left(\frac{0.0665\mathrm{Ra}^{1/3}}{1 + \left(\frac{9000}{\mathrm{Ra}}\right)^{1.4}}\right)^2\right]^{1/2}$$

$$\mathrm{Nu}_2 = 0.242\left(\mathrm{Ra}\frac{L}{H}\right)^{0.273} \qquad (3.2.36)$$

$$\mathrm{Nu} = [\mathrm{Nu}_1, \mathrm{Nu}_2]_{\max}.$$

For $\mathrm{Pr} \geq 4$, the following equation is recommended (Seki et al., 1978) for Ra $(H/L)^3 < 4 \times 10^{12}$:

$$\mathrm{Nu} = \left[1, 0.36\mathrm{Pr}^{0.051}\left(\frac{L}{H}\right)^{0.36}\mathrm{Ra}^{0.25}, 0.084\mathrm{Pr}^{0.051}\left(\frac{L}{H}\right)^{0.1}\mathrm{Ra}^{0.3}\right]_{\max} \qquad (3.2.37a)$$

And for Ra $(H/L)^3 > 4 \times 10^{12}$

$$\mathrm{Nu} = 0.039\mathrm{Ra}^{1/3}. \qquad (3.2.37b)$$

25. *Tilted layers, Figure 3.2.6a, with* $0 \leq \theta \leq 90°$, $W/L > 8$. For gases ($\mathrm{Pr} \approx 0.7$), $0 \leq \theta \leq 60°$, and $\mathrm{Ra} \leq 10^5$ (Hollands et al., 1976), one can use

$$\mathrm{Nu} = 1 + 1.44\left[1 - \frac{1708}{\mathrm{Ra}\cos\theta}\right]^{\bullet}\left[1 - \frac{1708(\sin 1.8\theta)^{1.6}}{\mathrm{Ra}\cos\theta}\right] + \left[\left(\frac{\mathrm{Ra}\cos\theta}{5830}\right)^{1/3} - 1\right]^{\bullet}. \quad (3.2.38)$$

See Equation 3.2.35 for definition of [x]. For $60° \leq \theta \leq 90°$, linear interpolation is recommended using Equations 3.2.36 for $\theta = 90°$ and 3.2.38 for $\theta = 60°$.

26. *Concentric cylinders, Figure 3.2.6b.* For heat transfer across the gap between horizontal concentric cylinders, the Nusselt number is defined as $\mathrm{Nu} = q' \ln(D_o/D_i)/(2\pi k \Delta T)$ where q' is the heat-transfer per unit length of cylinder. For $\mathrm{Ra} \leq 8 \times 10^7$, $0.7 \leq \mathrm{Pr} \leq 6000$, $1.15 \leq D_o/D_i \leq 8$ (Raithby and Hollands, 1975),

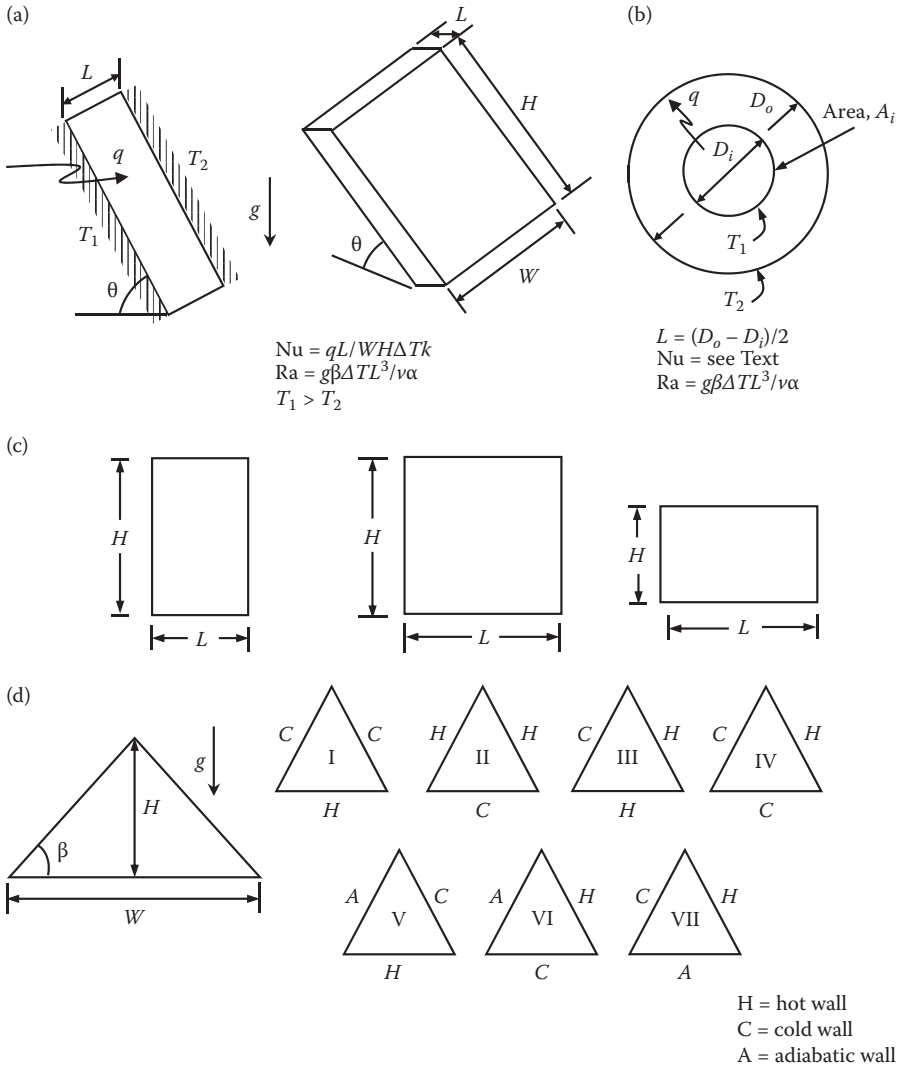


FIGURE 3.2.6 Nomenclature for enclosure problems.

$$\mathrm{Nu} = \left[0.603\bar{C}_\ell \frac{\ln\left(D_o/D_i\right)\mathrm{Ra}^{1/4}}{\left[\left(L/D_i\right)^{3/5} + \left(L/D_o\right)^{3/5}\right]^{5/4}}, 1 \right]_{\max} . \tag{3.2.39}$$

For eccentric cylinders, see Raithby and Hollands (1985).

27. *Concentric spheres, Figure 3.2.6b.* The heat transfer between concentric spheres is given by the following equation (Raithby and Hollands, 1975) for $\mathrm{Ra} \le 6 \times 10^8$, $5 \le \mathrm{Pr} \le 4000$, $1.25 < D_o/D_i \le 2.5$:

$$\mathrm{Nu} = \frac{qL}{D_i D_o k \Delta T} = \left[1.16\bar{C}_\ell \left(\frac{L}{D_i} \right)^{1/4} \frac{\mathrm{Ra}^{1/4}}{\left[\left(D_i/D_o\right)^{3/5} + \left(D_i/D_o\right)^{5/4}\right]}, 1 \right]_{\max} . \tag{3.2.40}$$

For eccentric spheres, see Raithby and Hollands (1985).

28. *Rectangular enclosures*, *Figure 3.2.6c*. A comprehensive review of the buoyancy-driven flow in a rectangular cavity is presented in Bejan (2013) and Martynenko and Khramtsov (2005) spanning laminar and turbulent regimes. From various different configurations of the buoyancy-driven flow in an enclosure of height H and horizontal length L, one of the most studied cases involves a two-dimensional rectangular cavity with differentially heated vertical side walls while the other two walls are adiabatic. Other configurations found in the literature consist of differentially heated horizontal surfaces with insulated vertical walls, sidewalls with uniform heat flux, a heat flux given on the one vertical surface and the temperature on the other, partially divided enclosures, etc. One must resort to the original publications for these cases.
29. *Triangular enclosures*, *Figure 3.2.6d*. The heat-transfer coefficient for isosceles triangular enclosures (aspect ratio, $AR=H/W$) with possible nontrivial combinations of isothermal and adiabatic side walls and bottom can be evaluated by the following equations in laminar regime.

    Configuration VII: For $4.1 \times 10^6 \le \mathrm{Ra}_H \le 12.7 \times 10^6$, Pr=0.71 (Flack and Witt, 1979a, b; Flack et al., 1979)

$$0.29 \le AR \le 0.87, \mathrm{Nu} = 0.269\mathrm{Ra}_{\mathrm{H}}^{0.34} \tag{3.2.41}$$

    Configuration I: For $11.26 \times 10^2 \le \mathrm{Ra}_H \le 9 \times 10^4$, Pr=0.71 (Akinsete and Coleman, 1979a, b, 1982; Coleman, 1980)

$$0.0625 \le AR \le 1, \mathrm{Nu} = 1.122\mathrm{Ra}_{\mathrm{w}}^{0.0535} AR^{-1.19} \tag{3.2.42}$$

    Configuration I, II: For $7.5 \times 10^4 \le \mathrm{Ra}_{\mathrm{H}} \le 2 \times 10^5$, Pr=0.71 (Flack, 1980)

$$AR = 1, \mathrm{Nu} = 0.21\mathrm{Ra}_{\mathrm{H}}^{0.30} \tag{3.2.43}$$

    Configuration I: For $7.5 \times 10^4 \le \mathrm{Gr}_H \le 2 \times 10^5$, Pr=0.72, 6 (Poulikakos and Bejan, 1983)

$$0.2 \le AR \le 1, \mathrm{Nu} = AR^{0.5}\mathrm{Ra}_{\mathrm{H}}^{0.25} \tag{3.2.44}$$

    Configuration I: For $7.1 \times 10^2 \le \mathrm{Ra}_{\mathrm{H}} \le 7.1 \times 10^5$, Pr=0.71 (Hasani and Chung, 1997)

$$0.2 \le AR \le 1, \mathrm{Nu} = 3.778 + 0.332\mathrm{Ra}_{\mathrm{H}}^{0.124} AR^{1.375} \tag{3.2.45}$$

    Configuration II: For $7.1 \times 10^2 \le \mathrm{Ra}_{\mathrm{H}} \le 7.1 \times 10^5$, Pr=0.71 (Hasani and Chung, 1997)

$$0.2 \le AR \le 0.6, \mathrm{Nu} = 3.855 + 0.0043\mathrm{Ra}_{\mathrm{H}}^{0.588} AR^{0.323} \tag{3.2.46a}$$

$$0.6 \le AR \le 1.0, \mathrm{Nu} = 4.342 + 0.0033\mathrm{Ra}_{\mathrm{H}}^{0.6004} AR^{0.699}. \tag{3.2.46b}$$

30. *Array of circular cylinders* (*vertical and horizontal*). For calculating the heat-transfer coefficient for a cylinder array arranged in a vertical line, one can use the following expression (Tokura et al., 1983):

$$\mathrm{Nu}_N = 0.410\mathrm{Ra}^{1/4} \ln\left[\left\{\left(b/d\right)/1.3\right\}^{0.055N}\right] + 0.434. \tag{3.2.47}$$

TABLE 3.2.3
Values of a_3 and b_3 in Equation 3.2.48

s/D→	0.7	0.8333	1.0
a_3	0.5972	0.6186	0.5735
b_3	0.0713	0.0488	0.0301

Equation 3.2.47 fits the experimental data within ±10% in the range of $[(b/d)/1.3]^{0.055N} = 0.7$ to 1.2, $1 < S/d < 40$, $\mathrm{Gr} = 4 \times 10^4$ to 4×10^5 and $\mathrm{Pr} = 0.71$, where *S/D* is the distance between parallel plates and *b/D* is the spacing ratio defined as diameter-to-diameter distance.

For two differentially heated vertically aligned cylinders in an adiabatic circular enclosure with spacing *s/D*, center-to-center distance, the following expression for the average Nusselt number can be used (Ho et al.,1993):

$$\mathrm{Nu} = a_3 \mathrm{Ra}^{b_3}, \tag{3.2.48}$$

where the values of fitted parameters a_3 and b_3 for $10^4 \le \mathrm{Ra} \le 10^7$ and $\mathrm{Pr} = 0.71$ are given in Table 3.2.3.

Example Calculations

Problem 3.2.1: Heat transfer from a vertical plate, Figure 3.2.7a. For the vertical isothermal surface in Figure 3.2.7a with $T_s = 40°\mathrm{C}$, $H_1 = H_2 = W_1 = W_2 = 1$ m and for an ambient air temperature of $T_\infty = 20°\mathrm{C}$ (at 1 atm), find the rate of heat transfer from one side of the plate.

Properties: At $T_f = (T_w + T_\infty)/2 = 30°\mathrm{C}$ and atmospheric pressure for air: $\nu = 1.59 \times 10^{-5}$ m²/s, $\alpha = 2.25 \times 10^{-5}$ m²/s, $\mathrm{Pr} = 0.71$, $k = 0.0263$ W/mK. At T_∞, $\beta \approx 1/T_\infty = 1(273+20) = 0.00341$ K⁻¹.

Solution: For the geometry shown in Figure 3.2.7a

$$A_s = (H_1 + H_2)W_1 + \left(H_1 + \frac{H_2}{2}\right)W_2 = 3.5 \text{ m}^2 \quad \text{(plate surface area)}$$

$$\int_0^{W_1+W_2} S^{3/4}\, d\chi = (H_1+H_2)^{3/4} W_1 + \frac{4}{7}\frac{W_2}{H_2}\left[(H_1+H_2)^{7/4} - H_1^{7/4}\right] = 3.03 \text{ m}^{7/4}$$

$$L^{1/4} = (H_1 + H_2)^{1/4} = 1.19 \text{ m}^{1/4}.$$

Using Equation 3.2.11,

$$C_1 = \frac{L^{1/4} \int_0^{W_1+W_2} S^{3/4}\, d\chi}{A_s} = \frac{1.19 \times 3.03}{3.5} = 1.03$$

$$\mathrm{Ra} = \frac{g\beta_\infty L^3 (T_w - T_\infty)}{\upsilon\alpha} = \frac{9.81 \times 0.00341 \times 2^3 \times (40-20)}{1.59 \times 10^{-5} \times 2.25 \times 10^{-5}} = 1.50 \times 10^{10}$$

$\bar{C}_\ell = 0.514$ from Equation 3.2.3; $C_t = C_t^V = 0.103$ from Equation 3.2.4. $\mathrm{Nu}^T = C_1 \bar{C}_\ell \mathrm{Ra}^{1/4} = 185$ from Equation 3.2.11.

$$\left.\begin{aligned} \mathrm{Nu_t} &= \frac{2.0}{\ln\left(1+2.0/\mathrm{Nu^T}\right)} = 186 \\ \mathrm{Nu_t} &= C_t^{V}\mathrm{Ra}^{1/3}\Big/\left(1+1.4\times10^{9}\,\mathrm{Pr}/\mathrm{Ra}\right) = 238 \end{aligned}\right\} \text{(from Equation (3.2.7))}$$

$$Nu = \frac{qL}{A\Delta Tk} = \left(Nu_\ell^6 + Nu_t^6\right)^{1/6} = 246$$

from Equation 3.2.6 with $m=6$.

$$q = \frac{A_s\Delta Tk\mathrm{Nu}}{L} = \frac{3.5\times20\times0.0263\times246}{2} = 226W.$$

Comments: Since $\mathrm{Nu}_\ell < \mathrm{Nu_t}$, the heat transfer is primarily in the turbulent regime. One must ascertain if it is justified to neglect radiation in this case. Had the surface been specified to be at constant heat flux, rather than isothermal, the equations in this section can be used to find the approximate average temperature difference between the plate and fluid.

Problem 3.2.2: Heat transfer from horizontal strip, Figure 3.2.7b. Find the rate of heat loss per unit length from a very long strip of width $W=0.1$ m with a surface temperature of $T_s=70°C$ in water at $T_\infty=30°C$.

Properties: At $T_f=(T_s+T_\infty)/2=50°C$

$$\nu = 5.35\times10^{-7}\mathrm{m^2/s} \quad \alpha = 1.56\times10^{-7}\mathrm{m^2/s} \quad \mathrm{Pr} = 3.42$$

$$k = 0.645\ \mathrm{W/mK} \quad \beta = 2.76\times10^{-4}K^{-1}$$

Solution: This problem corresponds to *Case 3* and Figure 3.2.2c.

$$C_t^{H} = 0.14.$$

From Equation 3.2.5 and $\bar{C}_l = 0.563$ from Equation 3.2.3,

$$L^* = \lim_{H\to\infty}\left(\frac{WH}{2W+2H}\right) = \frac{W}{2} = 0.05\ \mathrm{m}.$$

From Figure 3.2.2c and Equation 3.2.9,

$$\mathrm{Ra} = \frac{g\beta\Delta TL^{*3}}{\upsilon\alpha} = 1.62\times10^{8} \qquad \mathrm{Nu^T} = 0.835\bar{C}_l\mathrm{Ra}^{1/4} = 55.6$$

$$\mathrm{Nu}_\ell = \frac{1.4}{\ln\left(1+1.4/\mathrm{Nu^T}\right)} = 56.3 \qquad \mathrm{Nu_t} = C_t^{H}\mathrm{Ra}^{1/3} = 76.3$$

$$\mathrm{Nu} = \frac{q}{WH\Delta T}\frac{L^*}{k} = \left(\mathrm{Nu}_l^{10} + \mathrm{Nu}_t^{10}\right)^{0.1} = 76.7$$

$$q/H = \frac{W\Delta Tk\mathrm{Nu}}{L^*} = 3960\ \mathrm{W/m}.$$

Comments: Turbulent heat transfer is dominant. Radiation can be ignored (since it lies in the far infrared region where it is not transmitted by the water).

Problem 3.2.3: Heat loss across a window cavity, Figure 3.2.7c. The interior glazing is at temperature $T_1=10°C$, the exterior glazing at $T_2=-10°C$, the window dimensions are $W=1$ m, $H=1.7$ m,

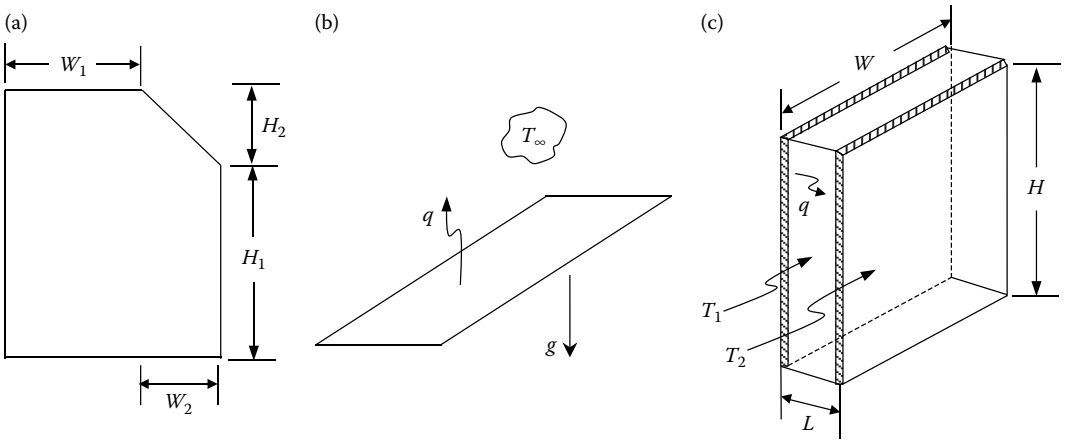


FIGURE 3.2.7 Sketches for example problems.

and the air gap between the glazings is $L = 10$ mm and is at atmospheric pressure. Find the heat loss across the window.

Properties: At $\overline{T} = (T_1 + T_2)/2 = 0^\circ\text{C} = 273\ \text{K}$

$$\nu = 1.35 \times 10^{-5}\ \text{m}^2/\text{s} \quad \alpha = 1.89 \times 10^{-5}\ \text{m}^2/\text{s} \quad \text{Pr} = 0.71$$

$$k = 0.024\ \text{W/mK} \quad \beta = 1/273 = 3.66 \times 10^{-3} \cdot K^{-1}$$

Solution: The appropriate correlations are given in *Case 24* and by Equation 3.2.36:

$$\text{Ra} = \frac{g\beta(T_1 - T_2)L^3}{\upsilon\alpha} = \frac{9.81 \times 3.66 \times 10^{-3} \times 20 \times (0.01)^3}{1.35 \times 10^{-5} \times 1.89 \times 10^{-5}} = 2.81 \times 10^3$$

$$\text{Nu}_1 = \left[1 + \left\{ \frac{0.0665\text{Ra}^{1/3}}{1 + \left(\frac{9000}{\text{Ra}} \right)^{1.4}} \right\}^2 \right]^{1/2} = 1.01$$

$$\text{Nu}_2 = 0.242\left(\text{Ra}\frac{L}{H} \right)^{0.273} = 0.242\left(2.81 \times 10^3 \times \frac{0.01}{1.7} \right)^{0.273} = 0.520$$

$$\text{Nu} = \frac{qL}{WH(T_1 - T_2)k} = (\text{Nu}_1, \text{Nu}_2)_{\max} = 1.01$$

$$q/WH = \frac{\text{Nu}(T_1 - T_2)k}{L} = \frac{1.01 \times 20 \times 0.024}{0.01} = 48.5\ \text{W/m}^2.$$

Comments: For pure conduction across the air layer, Nu = 1.0. For the calculated value of Nu = 1.01, convection must play little role. For standard glass, the heat loss by radiation would be roughly double the natural convection value just calculated.

In this section, the results only for a range of representative configurations have been included and a more exhaustive coverage can be found in the book of Martynenko and Khramtsov (2005).

Special Nomenclature

Note that nomenclature for each geometry considered is provided in the figures that are referred to in the text.

$\overline{C}_{\ell}$ = function of Prandtl number, Equation 3.2.3
C_t^V = function of Prandtl number, Equation 3.2.4
C_t^H = function of Prandtl number, Equation 3.2.5
$\overline{C}_t$ = surface averaged value of C_t, Equation 3.2.12
$\overline{\Delta T}$ = surface averaged value of $T_w - T_\infty$.

REFERENCES

Akinsete, V.A. and Coleman, T.A. 1979a. Heat transfer by steady laminar free convection within triangular enclosures, *Numerical Methods in Thermal Problems*, Pineridge Press, Ltd., Swansea, Wales, 259–268.

Akinsete, V.A. and Coleman, T.A. 1979b. Heat transfer by steady laminar free convection within triangular enclosures. *Proceedings of International Congress Brasileiro De Engenharia Mecanica*, Brazil.

Akinsete, V.A. and Coleman, T.A. 1982. Heat transfer by steady laminar free convection within triangular enclosures, *Int. J. Heat Mass Transfer*, 25, 991–998.

Bairi, A., Zarco-Pernia, E. and Garcia de Maria, J.M. 2014. A review on natural convection in enclosures for engineering applications. The particular case of the parallelogrammic diode cavity, *Appl. Therm. Eng.*, 63, 304–322.

Bejan, A. 2013. *Convection Heat Transfer*, 4th ed., Wiley, New York.

Bergman, T.L., Lavine, A.S., Incropera, F.P. and DeWitt, D.P. 2011. *Introduction to Heat Transfer*, 6th ed., Wiley, New York.

Chandra, A. and Chhabra, R.P. 2012, Laminar free convection from a horizontal semi-circular cylinder to power-law fluids, *Int. J. Heat Mass Transfer*, 55, 2934–2944.

Churchill S.W. and Usagi, R. 1972. A general expression for the correlation of rates of transfer and other phenomena, *AIChE J.*, 18, 1121–1128.

Churchill, S.W. and Chu, H.H.S. 1975. Correlating equations for laminar and turbulent free convection from a vertical plate, *Int. J. Heat Mass Transfer*, 18, 1323–1329.

Churchill, S.W. and Thelen, H.J. 1975. Eine allgemeine korrelationsgleichung für den wärme- und stoffübergang bei freier konvektion, *Chem. Ing. Tech.*, 47, 453.

Clemes, S.B., Hollands, K.G.T. and Brunger, A.P. 1994. Natural convection heat transfer from horizontal isothermal cylinders, *J. Heat Transfer*, 116, 96–104.

Edwards, J.A. and Chaddock, J.B. 1963. An experimental investigation of the radiation and free-convection heat transfer from a cylindrical disk extended surface, *Trans. ASHRAE*, 69, 313–322.

Elenbaas, W. 1942a. The dissipation of heat by free convection: the inner surface of vertical tubes of different shapes of cross-section, *Physica*, 9(8), 865–874.

Elenbaas, W. 1942b. Heat dissipation of parallel plates by free convection, *Physica*, 9(1), 2–28.

Eslami, M. and Jafarpur, K. 2012. Laminar free convection heat transfer from isothermal convex bodies of arbitrary shape: A new dynamic model, *Heat Mass Transfer*, 48, 301–315.

Fand, R.M., Morris, E.W. and Lum, M. 1977. Natural convection heat transfer from horizontal cylinders to air, water and silicone oils for Rayleigh numbers between 3×10^2 and 2×10^7, *Int. J. Heat Mass Transfer*, 20, 1173–1184.

Flack, R.D. 1980. The experimental measurement of natural convection heat transfer in triangular enclosures heated or cooled from below. *J. Heat Transfer*, 102, 770–772.

Flack, R.D. and Witt. C.L. 1979a. Laser velocimeter measurements in natural convection flows in triangular enclosures, *Proceedings 3rd International Workshop on Laser Velocimetry* (*LV-III*), 445–454.

Flack, R.D. and Witt. C.L. 1979b. Velocity measurements in two natural convection air flows using laser velocimeter, *J. Heat Transfer*, 101, 256–260.

Flack, R.D., Konopnicki, T.T. and Rooke, J.H. 1979. The measurement of natural convection heat transfer in triangular enclosures, *J. Heat Transfer*, 101, 648–654.

Fujii, T. and Fujii, M. 1976. The dependence of local Nusselt number on Prandtl number in the case of free convection along a vertical surface with uniform heat flux, *Int. J. Heat Mass Transfer*, 19, 121–122.

Fujii, T., Honda, H. and Morioka, I. 1973. A theoretical study of natural convection heat transfer from downward-facing horizontal surface with uniform heat flux, *Int. J. Heat Mass Transfer*, 16, 611–627.

Goldstein, R.J., Sparrow, E.M. and Jones, D.C. 1973. Natural convection mass transfer adjacent to horizontal plates, *Int. J. Heat Mass Transfer*, 16, 1025–1035.

Gupta, A.K., Sasmal, C., Sairamu, M. and Chhabra, R.P. 2014. Laminar and steady free convection in power-law fluids from a heated spheroidal particle: A numerical study, *Int. J. Heat Mass Transfer*, 75, 592–609.

Hasani, S.M.F. and Chung, B.T.F. 1997. Laminar natural convection in triangular enclosure, *ASME Proceedings of Ocean Engineering Division*, D.T. Valentine and C.C. Jahnke, Eds., ASME, New York, pp. 107–116.

Ho, C.J., Chang, W.S. and Wang, C.C. 1993. Natural convection between two horizontal cylinders in an adiabatic circular enclosure, *J. Heat Transfer*, 115, 158–165.

Hollands, K.G.T. 1984. Multi-Prandtl number correlations equations for natural convection in layers and enclosures, *Int. J. Heat Mass Transfer*, 27, 466–468.

Hollands, K.G.T., Unny, T.E., Raithby, G.D. and Konicek, K. 1976. Free convection heat transfer across inclined air layers, *J. Heat Transfer*, 98, 189–193.

Jafarpur, K. and Yovanovich, M.M. 1992. Laminar free convection heat transfer from isothermal spheres: A new analytical method, *Int. J. Heat Mass Transfer*, 35, 2195–2201.

Kamiyo, O.M., Angeli, E., Barozzi, G.S., Collins, M.W., Olunloyo, V.O.S. and Talabi, S.O. 2010. A comprehensive review of natural convection in triangular enclosures, *Appl. Mech. Rev.*, 63, 060801.

Karagiozis, A. 1991. An investigation of laminar free convection heat transfer from isothermal finned surfaces, Ph.D. Thesis, Department of Mechanical Engineering, University of Waterloo, Waterloo, ON.

Karagiozis, A., Raithby, G.D. and Hollands, K.G.T. 1994. Natural convection heat transfer from arrays of isothermal triangular fins in air, *J. Heat Transfer*, 116, 105–111.

Kays, W.M., Crawford, M.E. and Weigand, B. 2004. *Convective Heat and Mass Transfer*, 4th ed., McGraw-Hill, New York.

Khalifa, A-J. N. 2001a. Natural convective heat transfer coefficient—A review I. Isolated vertical and horizontal surfaces, *Energy. Convers. Manage*., 42, 491–504.

Khalifa, A-J. N. 2001b. Natural convective heat transfer coefficient—A review II. Surfaces in two- and three-dimensional enclosures, *Energy. Convers. Manage*., 42, 505–517.

Kreith, F., Manglik, R.M. and Bohn, M.S. 2011. *Principles of Heat Transfer*, 7th ed., Cengage Learning, Boston, MA.

Lee, S., Yovanovich, M.M. and Jafarpur, K. 1991. Effect of geometry and orientation on laminar natural convection from isothermal bodies, *AIAA J. Thermophys. Heat Transfer*, 5, 208–216.

Martynenko, O.G. and Khramtsov, P. P. 2005. *Free-Convective Heat Transfer*, Springer, New York.

Mograbi, E., Ziskind, G., Katoshevski, D. and Bar-Ziv, E. 2002. Experimental study of the forces associated with mixed convection from a heated sphere at small Reynolds and Grashof numbers. Part II: assisting and opposing flows, *Int. J. Heat Mass Transfer*, 45, 2423–2430.

Poulikakos, D. and Bejan, A. 1983. The fluid dynamics of an attic space, *J. Fluid Mech*., 131, 251–269.

Raithby, G.D. and Hollands, K.G.T. 1975. A general method of obtaining approximate solutions to laminar and turbulent free convection problems, *Advances in Heat Transfer*, 11, 266–315.

Raithby, G.D. and Hollands, K.G.T. 1985. *Handbook of Heat Transfer*, Chap. 6: Natural Convection, W.M. Rohsenow, J.P., Hartnett, and E.H. Ganic, Eds., McGraw-Hill, New York.

Sasmal, C. and Chhabra, R. P. 2011. Laminar natural convection from a heated square cylinder immersed in power-law liquids, *J. Non-Newt. Fluid Mech.,* 166, 811–830.

Sasmal, C. and Chhabra, R. P. 2012a. Effect of aspect ratio on natural convection in power-law fluids from a heated horizontal elliptic cylinder, *Int. J. Heat Mass Transfer*, 55, 4886–4899.

Sasmal, C. and Chhabra, R. P. 2012b. Effect of orientation on laminar natural convection from a heated square cylinder in power-law liquids, *Int. J. Thermal Sci*., 57, 112–125.

Sasmal, C. and Chhabra, R. P. 2014. Laminar free convection in power-law fluids from a heated hemisphere. *AIAA J. Thermophys. Heat Transfer*, 28, 750–763.

Seki, N., Fukusako, S. and Inaba, H. 1978. Heat transfer of natural convection in a rectangular cavity with vertical walls of different temperatures, *Bull. JSME*, 21(152), 246–253.

Shewen, E., Hollands, K.G.T. and Raithby, G.D. 1996. Heat transfer by natural convection across a vertical air cavity of large aspect ratio, *J. Heat Transfer*, 118, 993–995.

Tiwari, A.K. and Chhabra, R.P. 2013. Laminar natural convection in power-law liquids from a heated semicircular cylinder with its flat side oriented downward, *Int. J. Heat Mass Transfer*, 58, 553–567.

Tokura, I., Saito, H., Kishinami, K. and Muramoto, K. 1983. An experimental study of free convection heat transfer from a horizontal cylinder in a vertical array set in free space between parallel walls, *J. Heat Transfer*, 105, 102–107.

Further Information

There are several excellent heat-transfer textbooks that provide fundamental information and correlations for natural convection heat transfer (e.g., Kreith et al., 2011; Bergman et al., 2011). The correlations in this section closely follow the recommendations of Raithby and Hollands (1985), but that reference considers many more problems. Alternative equations are provided by Churchill and coworkers (1972, 1975) and by Martynenko and Khramtsov (2005). Good reviews of free convection in enclosures are available (Khalifa, 2001a, b; Bairi et al., 2014; Kamiya et al., 2010), which should be consulted for further details.

3.2.2 Forced Convection: External Flows

Anoop K. Gupta, Raj P. Chhabra, and N.V. Suryanarayana

Introduction

In this section, we consider forced convection heat transfer between a solid surface and an adjacent fluid, which is in motion relative to the solid surface. If the surface temperature is different from that of the fluid, heat is transferred by forced convection. The velocity and temperature of the fluid far away from the solid surface are at the free-stream velocity (U_∞) and temperature (T_∞) conditions, respectively. Both are usually known or specified in a given situation. We are then required to find the heat flux from or to the surface with specified surface temperature or the surface temperature if the heat flux is specified. The specified temperature or heat flux either may be uniform or may vary along the surface. The convective heat-transfer coefficient h is defined by the equation

$$q'' = h(T_s - T_\infty). \tag{3.2.49}$$

In Equation 3.2.49 with the local heat flux, we obtain the local heat-transfer coefficient, and with the average heat flux with a uniform surface temperature, we get the average heat-transfer coefficient. For a specified heat flux, the local surface temperature is obtained by employing the local convective heat-transfer coefficient in Equation 3.2.49.

Many correlations for estimating the convective heat-transfer coefficient are available, which endeavor to combine the approximate analytical, numerical, and experimental data. Depending upon the idealizations inherent in the analytical and experimental results vis-à-vis the conditions encountered in an application, one should not expect the actual value of the heat-transfer coefficient to be known very precisely, and uncertainties of the order of ±20% or even more are not uncommon.

Many different correlations to determine the convective heat-transfer coefficient have been developed for a range of geometries. In this section, only a selection of widely used correlations for common shapes is given. For other correlations and more details, refer to the books given in the bibliography at the end of this section.

Flat Plate

With a fluid flowing parallel to a flat plate (Figure 3.2.8), changes in velocity and temperature of the fluid are confined to a thin region adjacent to the solid boundary—the boundary layers of thicknesses δ and δ_T, both of which increase with x. Depending upon the conditions, several possibilities exist:

1. Flows without or with pressure gradient
2. Laminar or turbulent boundary layer
3. Negligible or significant viscous dissipation (effect of frictional heating)
4. $Pr \geq 0.7$ (gases and most liquids) or $Pr \ll 1$ (liquid metals)

Each of these is given consideration in the following sections.

Flows with Zero Pressure Gradient and Negligible Viscous Dissipation

When the free-stream pressure is uniform, the free-stream velocity is also uniform. Whether the boundary layer is laminar or turbulent is determined by the value of the local Reynolds number Re_x ($\rho U_\infty x/\mu$) and the shape of the solid surface at the entrance of the flow. With a sharp edge at the leading edge (Figure 3.2.8), the boundary layer is initially laminar, but at some distance downstream there is a transition region where the boundary layer is neither fully laminar nor fully turbulent. Farther downstream of the transition region, the boundary layer becomes fully turbulent. For engineering applications, the existence of the transition region is usually neglected and it is assumed

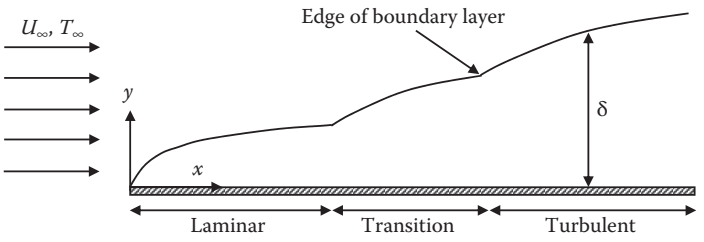


FIGURE 3.2.8 Flow of a fluid over a flat plate with laminar, transition, and turbulent boundary layers.

that the boundary layer becomes turbulent if the Reynolds number, Re_x, is greater than the critical Reynolds number, Re_{cr}. A typical value of 5×10^5 for the critical Reynolds number is generally accepted for a flat plate, but it can be greater if the free-stream turbulence is low and lower if the free-stream turbulence is high, the surface is rough, or the surface does not have a sharp edge at entrance. If the entrance is blunt, the boundary layer may be turbulent from the leading edge itself, albeit a very thin laminar sublayer exists very close to the surface of the plate even under these conditions. The momentum transfer in the y-direction occurs solely by diffusion within the boundary layer. The velocity gradient dU_x/dy gradually decreases with increasing y from its maximum value at $y=0$ to zero at $y=\delta$.

Temperature Boundary Layer

Analogous to the velocity boundary layer, there exists a temperature boundary layer adjacent to a heated (or cooled) plate. The temperature of the fluid changes from the surface temperature (T_s) to the free-stream temperature (T_∞) at the edge of the thermal boundary layer (Figure 3.2.9). By analogy with momentum transfer, heat transfer inside the thermal boundary layer occurs by conduction in the y-direction.

The velocity boundary layer thickness δ depends on the Reynolds number Re_x. The thermal boundary layer thickness δ_T depends both on the Reynolds number (Re_x) and Prandtl number (Pr) as well as on the type of thermal boundary condition prescribed on the surface of the plate as follows:

In the laminar regime, $Re_x < Re_{cr}$:

$$\frac{\delta}{x} = \frac{5}{\sqrt{Re_x}}. \tag{3.2.50a}$$

$$\text{For } Pr > 0.7: \frac{\delta}{\delta_T} = Pr^{1/3}: \tag{3.2.50b}$$

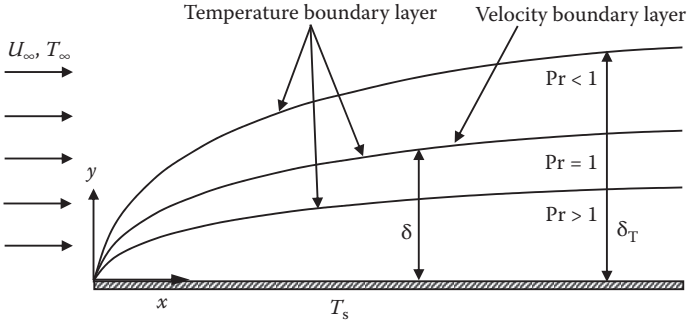


FIGURE 3.2.9 Temperature boundary layer thickness relative to velocity boundary layer thickness.

$$\text{For Pr} << 1 : \frac{\delta}{\delta_T} = \text{Pr}^{1/2} : \tag{3.2.50c}$$

In the turbulent regime, $\text{Re}_{cr} < \text{Re}_x$:

$$\frac{\delta}{x} = \frac{0.37}{\text{Re}_x^{0.2}} \text{ and } \delta \approx \delta_T. \tag{3.2.51}$$

Both these results are based on the assumptions of incompressible flow and constant values of the thermophysical properties of the fluid and negligible viscous dissipation effects. The viscous dissipation and high-speed effects can be neglected if $\text{Pr}^{1/2}\text{Ec}/2 << 1$. These effects are considered in later sections. The Eckert number, Ec, is defined as $\text{Ec} = U_\infty^2 / C_p(T_s - T_\infty)$.

With a rectangular thin plate of length L in the direction of the flow and infinitely wide in the lateral direction, the average heat-transfer coefficient h_L for the case of isothermal surface is given by

$$h_L = \frac{1}{L}\int_0^L h_x \mathrm{d}x. \tag{3.2.52}$$

Laminar Boundary Layer (Figure 3.2.10) ($\text{Re}_x < \text{Re}_{cr}$, $\text{Re}_L < \text{Re}_{cr}$): With heating or cooling starting from the leading edge (at $x=0$), the following correlations are recommended for estimating the value of the Nusselt number. In all equations presented herein, the fluid properties are evaluated at the film temperature defined as the arithmetic mean of the surface and free-stream temperatures, that is, $(T_s + T_\infty)/2$ unless otherwise stated.

Local Heat Transfer Coefficient (Uniform Surface Temperature)

The Nusselt number based on the local convective heat-transfer coefficient is expressed as

$$\text{Nu}_x = f_{\text{Pr}}\text{Re}_x^{1/2} \tag{3.2.53}$$

The classical expression for f_{Pr} is $0.564\text{Pr}^{1/2}$ for liquid metals with very low Prandtl numbers, $0.332\text{Pr}^{1/3}$ for $0.7 < \text{Pr} < 50$ and $0.339\text{Pr}^{1/3}$ for very large values of the Prandtl number. The two widely used correlations (valid for all Prandtl numbers) are due to Churchill (1976), Equation 3.2.54, and Rose (1979), Equation 3.2.55, as follows:

$$\text{Nu}_x = \frac{0.3387\text{Re}_x^{1/2}\text{Pr}^{1/3}}{\left[1+\left(\frac{0.0468}{\text{Pr}}\right)^{2/3}\right]^{1/4}}, \tag{3.2.54}$$

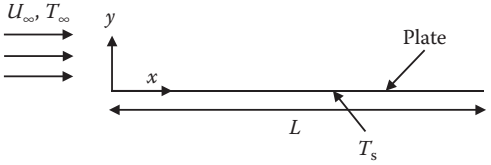


FIGURE 3.2.10 Heated flat plate with heating from the leading edge.

$$\mathrm{Nu}_x = \frac{\mathrm{Re}_x^{1/2}\mathrm{Pr}^{1/2}}{\left[27.8 + 75.9\mathrm{Pr}^{0.306} + 657\mathrm{Pr}\right]^{1/6}}. \tag{3.2.55}$$

In the wide range of Prandtl number as $0.001 < \mathrm{Pr} < 2000$, Equations 3.2.54 and 3.2.55 yield values of the local Nusselt number, which are within ~1.5% and ~0.5% of the exact numerical results based on the solution of the boundary layer energy equations.

Average Heat-Transfer Coefficient

The average heat-transfer coefficient is given by

$$\mathrm{Nu}_L = 2\mathrm{Nu}_{x=L}. \tag{3.2.56}$$

From Equation 3.2.56, it is clear that the average heat-transfer coefficient over a length L is twice the local heat-transfer coefficient at $x=L$.

Uniform Heat Flux

Local Heat Transfer Coefficient

In this case, the need to estimate the unknown temperature frequently arises via the application of Equation 3.2.49. A good summary of the available results is given in Schlichting and Gersten (2011). From an engineering application standpoint, one of the widely used predictive correlation spanning the complete range of Prandtl numbers is that of Churchill and Ozoe (1973), which is written as follows:

$$\mathrm{Nu}_x = \frac{0.886\mathrm{Re}_x^{1/2}\mathrm{Pr}^{1/2}}{\left[1 + \left(\frac{\mathrm{Pr}}{0.0207}\right)^{2/3}\right]^{1/4}}. \tag{3.2.57}$$

Note that for surfaces with uniform heat flux, the local convective heat-transfer coefficient is used to determine the local surface temperature. The total heat-transfer rate being known, an average heat-transfer coefficient is generally not needed and not defined.

Turbulent Boundary Layer ($\mathrm{Re}_x > \mathrm{Re}_{cr}$, $\mathrm{Re}_L > \mathrm{Re}_{cr}$): For turbulent boundary layers with heating or cooling starting from the leading edge, one can use the following correlations:

Local Heat Transfer Coefficient

$$\mathrm{Re}_{cr} < \mathrm{Re}_x < 10^7:$$

$$\mathrm{Nu}_x = 0.0296\mathrm{Re}_x^{4/5}\mathrm{Pr}^{1/3}. \tag{3.2.58}$$

$$10^7 < \mathrm{Re}_x:$$

$$\mathrm{Nu}_x = 1.596\mathrm{Re}_x(\ln \mathrm{Re}_x)^{-2.584}\mathrm{Pr}^{1/3}. \tag{3.2.59}$$

Equation 3.2.59 is obtained by applying Colburn's j-factor in conjunction with the friction factor suggested by Schlichting and Gersten (2011).

In laminar boundary layers, the convective heat-transfer coefficient with uniform heat flux is approximately 36% higher than that for uniform surface temperature conditions. With turbulent boundary layers, the difference is very small and *the correlations for the local convective heat-transfer coefficient can be used for both uniform surface temperature and uniform heat flux with approximately the same level of confidence.*

Average Heat Transfer Coefficient

If the boundary layer is initially laminar over the initial part of the plate followed by the establishment of a turbulent boundary layer at $\mathrm{Re}_x = \mathrm{Re}_{cr}$, the following correlations for $0.7 < \mathrm{Pr} < 60$ are suggested:

$$\mathrm{Re}_{cr} < \mathrm{Re}_L < 10^7:$$

$$\mathrm{Nu}_L = \left[0.664\mathrm{Re}_L^{1/2} + 0.037\left(\mathrm{Re}_L^{4/5} - \mathrm{Re}_{cr}^{4/5}\right)\right]\mathrm{Pr}^{1/3}. \tag{3.2.60}$$

If $\mathrm{Re}_{cr} < \mathrm{Re}_L < 10^7$ and $\mathrm{Re}_{cr} = 10^5$, Equation 3.2.60 simplifies to

$$\mathrm{Nu}_L = \left(0.037\mathrm{Re}_L^{4/5} - 871\right)\mathrm{Pr}^{1/3}, \tag{3.2.61}$$

$10^7 < \mathrm{Re}_L$ and $\mathrm{Re}_{cr} = 5 \times 10^5$, the corresponding result is given as follows:

$$\mathrm{Nu}_L = \left[1.963\mathrm{Re}_L\left(\ln \mathrm{Re}_L\right)^{-2.584} - 871\right]\mathrm{Pr}^{1/3}. \tag{3.2.62}$$

Uniform Surface Temperature—Pr > 0.7: Unheated Starting Length

If heating starts from the leading edge at a distance ($x = x_0$), as shown in Figure 3.2.11, the correlations have to be modified. Correlation for the local convective heat-transfer coefficient for laminar and turbulent boundary layers are given by Equations 3.2.63 and 3.2.64 (Kays et al., 2004)—the constants in Equations 3.2.63 and 3.2.64 have been modified to be consistent with the friction factors. These correlations are also useful as building blocks for finding the heat-transfer rates when the surface temperature varies in a predefined manner. Equations 3.2.65 and 3.2.66, developed by Thomas (1977), provide the average heat-transfer coefficients based on Equations 3.2.63 and 3.2.64.

Local Convective Heat Transfer Coefficient

$$\mathrm{Re}_x < \mathrm{Re}_{cr}:$$

$$\mathrm{Nu}_x = \frac{0.332\mathrm{Re}_x^{1/2}\mathrm{Pr}^{1/3}}{\left[1 - \left(\frac{x_0}{x}\right)^{3/4}\right]^{1/3}} \tag{3.2.63}$$

$$\mathrm{Re}_x > \mathrm{Re}_{cr}:$$

$$\mathrm{Nu}_x = \frac{0.0296\mathrm{Re}_x^{4/5}\mathrm{Pr}^{3/5}}{\left[1 - \left(\frac{x_0}{x}\right)^{9/10}\right]^{1/9}} \tag{3.2.64}$$

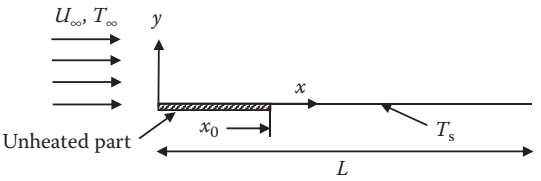


FIGURE 3.2.11 Heated flat plate with unheated starting length ($x \le x_0$).

Average Heat Transfer Coefficient over the Length ($L-x_0$)

$$\mathrm{Re}_L < \mathrm{Re}_{\mathrm{cr}}:$$

$$h_{L-x_0} = \frac{0.664\mathrm{Re}_L^{1/2}\mathrm{Pr}^{1/3}\left[1-\left(\frac{x_0}{L}\right)^{3/4}\right]^{2/3}k}{L-x_0} = 2\frac{1-(x_0/L)^{3/4}}{1-(x_0/L)}h_{x=L}. \tag{3.2.65}$$

In Equation 3.2.65, evaluate $h_{x=L}$ from Equation 3.2.63.

$$\mathrm{Re}_{\mathrm{cr}} = 0:$$

$$h_{L-x_0} = \frac{0.037\mathrm{Re}_L^{4/5}\mathrm{Pr}^{3/5}\left[1-\left(\frac{x_0}{L}\right)^{9/10}\right]^{8/9}k}{L-x_0} = 1.25\frac{1-(x_0/L)^{9/10}}{1-(x_0/L)}h_{x=L}. \tag{3.2.66}$$

In Equation 3.2.66, evaluate $h_{x=L}$ from Equation 3.2.64.

Flat Plate with Prescribed Nonuniform Surface Temperature

In practice, one can encounter the situation where the surface temperature varies in a prescribed manner. In this case, one can apply Equations 3.2.63 through 3.2.66 locally by dividing the total temperature difference into smaller steps. Figure 3.2.12 shows the arbitrarily prescribed surface temperature with a uniform free-stream temperature of the fluid. If the surface temperature is a differentiable function of the coordinate x, the local heat flux can be determined by an expression that involves integration (Kays et al., 2004). If the surface temperature can be approximated as a series of step changes in the surface temperature, the resulting expression for the local heat flux and the total heat-transfer rate is the summation of the individual simple algebraic expressions. Here the method using such an algebraic simplification is presented.

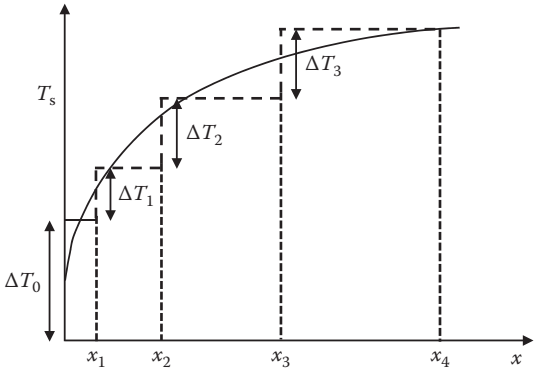


FIGURE 3.2.12 Arbitrary surface temperature approximated as a finite number of step changes.

The local heat flux at a distance x from the leading edge is given by

$$q''_x = \sum_1^n h_{xi} \Delta T_{si}, \tag{3.2.67}$$

where h_{xi} denotes the local heat-transfer coefficient at the position x due to a single step change in the surface temperature ΔT_{si} at location x_i ($x_i < x$). Referring to Figure 3.2.12, the local heat flux at x ($x_3 < x < x_4$) is given by

$$q''_x = h_x(x,0)\Delta T_0 + h_x(x,x_1)\Delta T_1 + h_x(x,x_2)\Delta T_2 + h_x(x,x_3)\Delta T_3, \tag{3.2.68}$$

where $h_x(x, x_1)$ is the local heat-transfer coefficient at x with heating starting from x_1; the local heat transfer is determined from Equation 3.2.63 if the boundary layer is laminar and Equation 3.2.64 if the boundary layer is turbulent from the leading edge. For example, $h_x(x, x_2)$ in the third term is given by

$\mathrm{Re}_x < \mathrm{Re}_{cr}$:

$$h_x(x,x_2) = \frac{0.332\left(\dfrac{\rho U_\infty x}{\mu}\right)^{1/2} \mathrm{Pr}^{1/3}}{\left[1-\left(\dfrac{x_2}{x}\right)^{3/4}\right]^{1/3}} \frac{k}{x}, \tag{3.2.69a}$$

$\mathrm{Re}_x > \mathrm{Re}_{cr}$:

$$h_x(x,x_2) = \frac{0.0296\left(\dfrac{\rho U_\infty x}{\mu}\right)^{4/5} \mathrm{Pr}^{3/5}}{\left[1-\left(\dfrac{x_2}{x}\right)^{9/10}\right]^{1/9}} \frac{k}{x}. \tag{3.2.69b}$$

The procedure for finding the total heat-transfer rate from $x=0$ to $x=L$ is somewhat similar. Denoting the width of the plate by W,

$$\frac{q}{W} = \sum h_{L-x_i} \Delta T_i \left(L - x_i\right), \tag{3.2.70}$$

where h_{L-x_i} is the average heat-transfer coefficient over the length $(L-x_i)$ due to a step change ΔT_i in the surface temperature at x_i. For example, the heat-transfer coefficient in the third term in Equation 3.2.70 is obtained by replacing x_0 by x_2 in Equations 3.2.65 or 3.2.66 depending on whether $\mathrm{Re}_L < \mathrm{Re}_{cr}$ or $\mathrm{Re}_{cr}=0$.

Flows with Pressure Gradient and Negligible Viscous Dissipation

Although most of the correlations available in the literature for flat plates are for a semi-infinite fluid medium adjacent to the plate, most applications of practical interest deal with fluid flowing between two plates. If the spacing between the plates is significantly greater than the maximum boundary layer thickness, the medium can be assumed to approach a semi-infinite medium because the boundary layers along the two plates develop independent of each other. In such a case, if the plates are parallel to each other and if the pressure drop is negligible compared with the absolute pressure of the system, the pressure gradient can be assumed to be negligible. If the plates are nonparallel and if the boundary layer thickness is very much smaller than the spacing between the plates at that

location, the medium can still be considered as approaching a semi-infinite medium with a nonnegligible pressure gradient. In such flows, the free-stream velocity (core velocity outside the boundary layer) is related to the pressure variation by the Bernoulli equation:

$$\frac{p}{\rho}+\frac{U_\infty^2}{2}+zg=\text{constant}. \tag{3.2.71}$$

Another situation where the free-stream velocity varies in the direction of flow giving rise to a pressure gradient is the flow over a wedge. For the family of flows for which the solutions are applicable, the free-stream velocity at the edge of the boundary layer is related to the x-coordinate by a power law, $U_\infty = cx^m$. Flows over semi-infinite wedges (Figure 3.2.13) satisfy that condition. The exponent m is related to the wedge angle $\beta\pi$ as

$$\beta=\frac{2m}{1+m} \quad \text{or} \quad m=\frac{\beta}{2-\beta}. \tag{3.2.72}$$

With laminar boundary layers, the boundary layer thickness, friction factor, and Nusselt numbers are defined by

$$\frac{\delta}{x}=\frac{c_1}{\sqrt{\text{Re}_x}};\ \frac{C_{fx}}{2}=\frac{\tau_w}{\rho U_\infty^2}=\frac{c_2}{\sqrt{\text{Re}_x}};\ \text{Nu}_x=c_3\text{Re}_x^{1/2}. \tag{3.2.73}$$

The values of c_1, c_2, and c_3 are available in Burmeister (1993). For example, for $\beta=0.5$ (wedge angle$=90°$), $m=1/3$, $c_1=3.4$, $c_2=0.7575$, and $c_3=0.384$ for Pr$=0.7$, and $c_3=0.792$ for Pr $= 5$. Re_x is based on $U_\infty=cx^m$; the free-stream velocity is not uniform.

Uniform Temperature: Flat Plate with Injection or Suction with External Flows of a Fluid Parallel to the Surface

Injection (Figure 3.2.14) or suction has engineering applications to control boundary layers and to regulate the rate of heating and cooling. When the free-stream temperature of the fluid is high, as in gas turbines, a cooling fluid is introduced into the mainstream to cool the surface. If the cooling fluid is introduced at discrete locations (either perpendicular to the surface or at an angle), it is known as film cooling. If a fluid is introduced or withdrawn through a porous medium, it is known as transpiration (Figure 3.2.14). An application of suction is to delay or prevent boundary layer separation thereby influencing the hydrodynamic forces.

Analytical solutions for a laminar boundary layer with transpiration suction or blowing are available if the velocity perpendicular to the surface varies in the following manner:

$$v_0=\text{constant}\ x^{(m-1)/2} \tag{3.2.74}$$

Solutions are limited to the cases of the injected fluid being at the same temperature as the surface and the injected fluid being the same as the free-stream fluid. Positive values of v_0 indicate blowing,

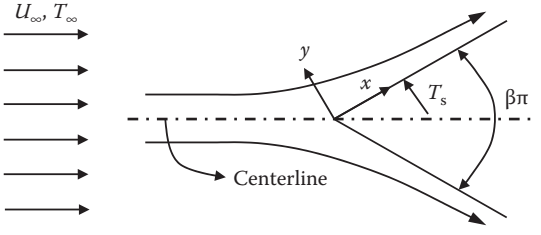


FIGURE 3.2.13 Flow over a wedge. ($\beta\pi$ is the wedge angle.)

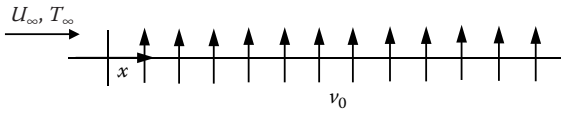


FIGURE 3.2.14 Flat plate with transpiration injection.

and negative values indicate suction. Values of $\mathrm{Nu}_x / \mathrm{Re}_x^{1/2}$ for different values of Pr and for different values of blowing or suction parameter are given in Kays et al. (2004).

For example, for a laminar boundary layer over a flat plate for air (Pr = 0.7), the value of $\mathrm{Nu}_x / \mathrm{Re}_x^{1/2}$ is 0.722 for $(v_0 / U_\infty)\sqrt{\rho U_\infty x / \mu} = -0.75$ (suction), and the corresponding value is 0.166 for $(v_0 / U_\infty)\sqrt{\rho U_\infty x / \mu} = 0.25$ (blowing). Heat transfer coefficient increases with suction which leads to a thinning of the boundary layer. Blowing increases the boundary layer thickness and decreases the heat-transfer coefficient.

For *turbulent boundary layers*, Kays et al. (2004) suggest the following procedure for finding the friction factor and convective heat-transfer coefficient in terms of a blowing parameter for momentum (B_f) and heat transfer (B_h) which are defined as follows:

$$B_\mathrm{f} = \frac{v_0 / U_\infty}{C_\mathrm{f} / 2}, \tag{3.2.75}$$

$$B_\mathrm{h} = \frac{v_0 / U_\infty}{\mathrm{St}} = \frac{\dot{m}'' / G_\infty}{\mathrm{St}}, \tag{3.2.76}$$

where

v_0 = velocity normal to the plate
U_∞ = free-stream velocity
$\dot{m}''$ = mass flux of the injected fluid at the surface (ρv_0)
G_∞ = mass flux in the free stream (ρU_∞)
St = Stanton number = $\mathrm{Nu}_x / \mathrm{Re}_x \mathrm{Pr} = h / \rho U_\infty c_p$

The friction factors, C_f and C_{f_0}, and Stanton numbers, St and St_0, with and without blowing or suction are related by the relations:

$$\frac{C_\mathrm{f}}{C_{\mathrm{f}_0}} = \frac{\ln(1 + B_\mathrm{f})}{B_\mathrm{f}}, \tag{3.2.77}$$

$$\frac{\mathrm{St}}{\mathrm{St}_0} = \frac{\ln(1 + B_\mathrm{h})}{B_\mathrm{h}}. \tag{3.2.78}$$

In Equations 3.2.77 and 3.2.78, C_{f_0} and St_0 are the values of the friction factor and Stanton number with $v_0 = 0$ (no blowing or suction), and C_f and St are the corresponding quantities with blowing or suction at the same values of Re_x ($= \rho U_\infty x / \mu$).

For the more general case of variable free-stream velocity, temperature difference, and transpiration rate, refer to Kays et al. (2004).

Flow over Flat Plate with Zero Pressure Gradient: Effect of High-Speed and Viscous Dissipation

In the boundary layer, the velocity of the fluid is reduced from U_∞ at the edge of the boundary layer to zero at the plate leading to a reduction in the kinetic energy of the fluid. Inside the boundary layer, there is also the work done by viscous forces due to shearing in the fluid; the magnitude of such viscous work is related to the velocity of the fluid, the velocity gradient, and the viscosity of the fluid. The net effect of such a reduction in the kinetic energy and the required viscous work is to increase

the internal energy of the fluid in the boundary layer. The increase in the internal energy may be expected to lead to an increase in the temperature; but because of the heat transfer to the adjacent fluid, the actual increase in the internal energy (and the temperature) will be less than the sum of the decrease in the kinetic energy and viscous work transfer. The actual temperature increase of the fluid depends on the decrease in the kinetic energy, the viscous work transfer, and the heat transfer from the fluid. The maximum temperature in the fluid with an adiabatic plate is known as the adiabatic wall temperature (which occurs at the wall) and is given by

$$T_{aw} = T_\infty + r\frac{U_\infty^2}{2C_p}. \tag{3.2.79}$$

In Equation 3.2.79, r is the (dimensionless) recovery factor and is really ratio of the internal energy to the kinetic energy of the fluid per unit mass. The value of r solely depends on the Prandtl number as follows (Eckert and Drake, 1972):

Laminar boundary layer: $0.6 < \mathrm{Pr} < 15$: $r = \mathrm{Pr}^{1/2}$
Turbulent boundary layer: $r = \mathrm{Pr}^{1/3}$
Equation 3.2.79 can now be recast as

$$\frac{T_{aw} - T_\infty}{T_s - T_\infty} = \frac{r}{2}\frac{U_\infty^2}{C_p(T_s - T_\infty)}. \tag{3.2.80}$$

From Equation 3.2.80, the maximum increase in the fluid temperature as a fraction of the difference between the plate and free-stream temperatures is given by r Ec/2. For instance, with air flowing over a plate at 500 m/s, the increase in the temperature of the air can be as high as 105°C. With T_s=40°C and T_∞=20°C, the temperature of the air close to the plate can be higher than the plate temperature. It is thus possible that although the plate temperature is higher than the free-stream temperature, the heat transfer is from the air to the plate. At a Mach number greater than 0.1 for gases, viscous dissipation becomes significant.

The temperature profiles for high-speed flows for different values of T_s are shown in Figure 3.2.15. In high-speed flows, as heat transfer can be from the fluid to the plate even if the plate temperature is greater than the fluid temperature, the definition of the convective heat-transfer coefficient given in Equation 3.2.49 is not adequate. On the other hand, as the heat transfer is always from the plate to the fluid if $T_s > T_{aw}$, the adiabatic wall temperature (akin to the adiabatic flame temperature in combustion applications) is more appropriate as the reference temperature for defining the heat-transfer coefficient. Thus, in high-speed flows, the definition of the convective heat-transfer coefficient is given by

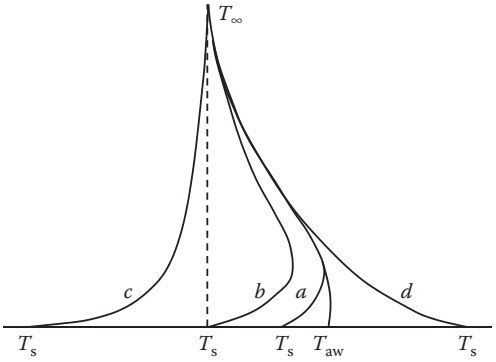


FIGURE 3.2.15 Temperature profiles for high-speed flows: (a) $T_\infty < T_s < T_{aw}$; (b) $T_s = T_\infty$; (c) $T_s << T_\infty$; (d) $T_s > T_{aw}$.

$$q'' = h(T_s - T_{aw}). \tag{3.2.81}$$

Equation 3.2.81 is consistent with Equation 3.2.49 as the adiabatic wall temperature equals the free-stream temperature if the effects of viscous dissipation and reduced kinetic energy in the boundary layer are neglected. With the adiabatic wall temperature as the fluid reference temperature for the definition of the convective heat-transfer coefficient, equations for low speeds can also be used for high-speed flows. Because of the greater variation in the fluid temperature in the boundary layer, the variation of the physical properties of the fluid due to temperature variation becomes important. It is found that the correlations are best approximated if the properties are evaluated at the reference temperature T^* defined by Eckert (1956):

$$T^* = 0.5(T_s + T_\infty) + 0.22(T_{aw} - T_\infty). \tag{3.2.82}$$

With the fluid properties evaluated at the reference temperature given by Equation 3.2.82, Equation 3.2.83 through 3.2.88 are applicable to high-speed flows with Prandtl numbers less than 15. It should be noted that the adiabatic wall temperatures in the laminar and turbulent regions are different affecting both the temperature at which the properties are evaluated and the temperature difference for determining the local heat flux. Therefore, when the boundary layer is partly laminar and partly turbulent, an average value of the heat-transfer coefficient cannot be defined as the adiabatic wall temperatures in the two regions are different. In such cases, the heat-transfer rate in each region is determined separately to find the total heat-transfer rate.

Laminar:

$$\mathrm{Re}_x < \mathrm{Re}_{cr}: \quad \mathrm{Nu}_x = 0.332\mathrm{Re}_x^{1/2}\mathrm{Pr}^{1/3} \tag{3.2.83}$$

$$\mathrm{Re}_L < \mathrm{Re}_{cr}: \quad \mathrm{Nu}_L = 0.664\mathrm{Re}_L^{1/2}\mathrm{Pr}^{1/3} \tag{3.2.84}$$

Turbulent:

$$10^7 > \mathrm{Re}_x > \mathrm{Re}_{cr}: \quad \mathrm{Nu}_x = 0.0296\mathrm{Re}_x^{4/5}\mathrm{Pr}^{1/3}, \tag{3.2.85}$$

$$10^7 < \mathrm{Re}_x < 10^9: \quad \mathrm{Nu}_x = 1.596\mathrm{Re}_x(\ln \mathrm{Re}_x)^{-2.584}\mathrm{Pr}^{1/3}, \tag{3.2.86}$$

$$\mathrm{Re}_{cr} = 0, \mathrm{Re}_L < 10^7: \quad \mathrm{Nu}_L = 0.037\mathrm{Re}_L^{4/5}\mathrm{Pr}^{1/3}, \tag{3.2.87}$$

$$\mathrm{Re}_{cr} = 0, 10^7 < \mathrm{Re}_L < 10^9: \quad \mathrm{Nu}_L = 1.967\mathrm{Re}_L(\ln \mathrm{Re}_L)^{-2.584}\mathrm{Pr}^{1/3}. \tag{3.2.88}$$

When the temperature variation in the boundary layer is large, such that the assumption of constant specific heat is not justified, Eckert (1956) suggests that the properties be evaluated at a reference temperature corresponding to the specific enthalpy of the fluid, i^* given by

$$i^* = 0.5(i_s + i_\infty) + 0.22(i_{aw} - i_\infty), \tag{3.2.89}$$

where i is the specific enthalpy of the fluid evaluated at the temperature corresponding to the subscript. Equation 3.2.89 gives the same values as Equation 3.2.82 if C_p is constant or varies linearly with temperature.

At very high speeds, the gas temperature may reach the values of temperatures that are sufficient to cause disassociation and chemical reactions; these and other effects need to be taken into account in such cases.

Flow over Cylinders, Spheres, and Other Geometries

Flows over a flat plate and wedges were classified as laminar or turbulent, depending on the Reynolds number, and numerous correlations for the local and average convective heat-transfer coefficients for a range of conditions have been developed over the years. But flows over cylinders (perpendicular to the axis) and spheres are more complex. In general, the flow over cylinders and spheres may have a laminar boundary layer followed by a turbulent boundary layer and a wake region depending on the Reynolds number with the diameter as the characteristic length (Figures 3.2.16 and 3.2.17). Because of the complexity of the flow patterns, only correlations for the average heat-transfer coefficients have been developed.

Cylinders: One can use the following correlation proposed by Churchill and Bernstein (1977): $\mathrm{Pe} > 0.2$. Evaluate properties at $(T_s + T_\infty)/2$:

$$\mathrm{Re}_d > 4\times10^5: \quad \mathrm{Nu}_d = 0.3 + \frac{0.62\mathrm{Re}_d^{1/2}\mathrm{Pr}^{1/3}}{\left[1+(0.4/\mathrm{Pr})^{2/3}\right]^{1/4}}\left[1+\left(\frac{\mathrm{Re}_d}{282{,}000}\right)^{5/8}\right]^{4/5}, \tag{3.2.90}$$

$$10^4 < \mathrm{Re}_d < 4\times10^5: \quad \mathrm{Nu}_d = 0.3 + \frac{0.62\mathrm{Re}_d^{1/2}\mathrm{Pr}^{1/3}}{\left[1+(0.4/\mathrm{Pr})^{2/3}\right]^{1/4}}\left[1+\left(\frac{\mathrm{Re}_d}{282{,}000}\right)^{1/2}\right], \tag{3.2.91}$$

$$\mathrm{Re}_d < 10^4: \quad \mathrm{Nu}_d = 0.3 + \frac{0.62\mathrm{Re}_d^{1/2}\mathrm{Pr}^{1/3}}{\left[1+(0.4/\mathrm{Pr})^{2/3}\right]^{1/4}}. \tag{3.2.92}$$

For the flow of liquid metals with very small values of the Prandtl number, one can use the following correlation suggested by Ishiguro et al. (1979):

$$1 < \mathrm{Pe} < 100: \mathrm{Nu}_d = 1.125\mathrm{Pe}^{0.413}. \tag{3.2.93}$$

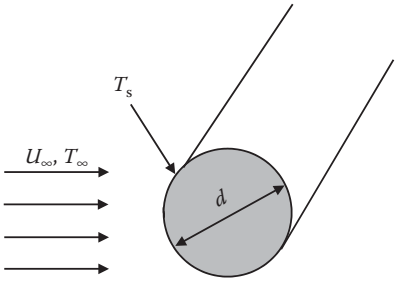


FIGURE 3.2.16 A fluid stream in cross flow over a cylinder.

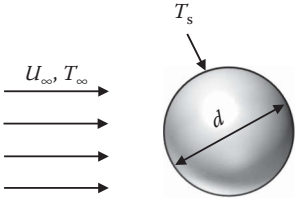


FIGURE 3.2.17 A fluid flowing over a sphere.

For more information on heat transfer with flow over cylinders, refer to Morgan (1975), Zukauskas (1987), Polyanin et al. (2002), Zdravkovich (1997, 2003) and Kakac et al. (2014).

Spheres: For flows over spheres (Figure 3.2.17), one can use any one of the following correlations:

1. Whitaker (1972): Evaluate properties at T_∞ except μ_s at T_s.

$$3.5 < \mathrm{Re}_d < 7.6\times10^4, \quad 0.71 < \mathrm{Pr} < 380, \quad 1 < \mu/\mu_s < 3.2$$

$$\mathrm{Nu}_d = 2.0 + \left(0.4\mathrm{Re}_d^{1/2} + 0.06\mathrm{Re}_d^{2/3}\right)\mathrm{Pr}^{2/5}\left(\frac{\mu}{\mu_s}\right)^{1/4}. \tag{3.2.94}$$

2. Achenbach (1978): For air ($\mathrm{Pr} = 0.71$) only. Evaluate properties at $(T_s + T_\infty)/2$:

$$100 < \mathrm{Re}_d < 2\times10^5$$

$$\mathrm{Nu}_d = 2.0 + \left(0.25\mathrm{Re}_d + 3\times10^{-4}\mathrm{Re}_d^{1.6}\right)^{1/2}, \tag{3.2.95}$$

$$4\times10^5 < \mathrm{Re}_d < 5\times10^6$$

$$\mathrm{Nu}_d = 430 + 5\times10^{-3}\mathrm{Re}_d + 2.5\times10^{-10}\mathrm{Re}_d^2 - 3.1\times10^{-17}\mathrm{Re}_d^3. \tag{3.2.96}$$

3. *Liquid Metals*: From experimental results with liquid sodium, Witte (1968) proposed

$$3.6\times10^4 < \mathrm{Re}_d < 1.5\times10^5: \quad \mathrm{Nu}_d = 2 + 0.386\mathrm{Pe}^{1/2}. \tag{3.2.97}$$

More information on the predictive correlations for spheres can be found in Clift et al. (1978) and Michaelides (2006).

Other Geometries: For geometries other than cylinders and spheres, use Equation 3.2.98 with the characteristic dimensions and values of the constants given in Table 3.2.4.

$$\mathrm{Nu}_D = c\mathrm{Re}_D^m. \tag{3.2.98}$$

Although Equation 3.2.98 is based on experimental data with gases, it can be extended to fluids with moderate Prandtl numbers by multiplying Equation 3.2.98 by the factor $(\mathrm{Pr}/0.7)^{1/3}$.

Heat Transfer across Tube Banks

When tube banks are used in heat exchangers, the flow over the tubes in the second and subsequent rows of tubes is different from the flow over a single tube. Even in the first row, the flow is modified by the presence of the neighboring tubes. The extent of modification depends on the spacing between the tubes and the Reynolds number of the flow. If the spacing is very much greater than the diameter of the tubes, correlations for single tubes can be used due to little or no hydrodynamic interaction between the cylinders. Correlations for flow over tube banks when the spacing between tubes in a row and a column is not much greater than the diameter of the tubes have been developed for use in heat-exchanger applications. Two arrangements of the tubes are considered—aligned and staggered as shown in Figure 3.2.18. The nomenclature used in this section is also shown in this figure.

For the average convective heat-transfer coefficient with tubes at uniform surface temperature, from experimental results, Zukauskas (1987) recommends correlations of the following form:

$$\mathrm{Nu}_d = c\left(\frac{a}{b}\right)^p \mathrm{Re}_d^m \mathrm{Pr}^n \left(\frac{\mathrm{Pr}}{\mathrm{Pr}_s}\right)^{0.25}. \tag{3.2.99}$$

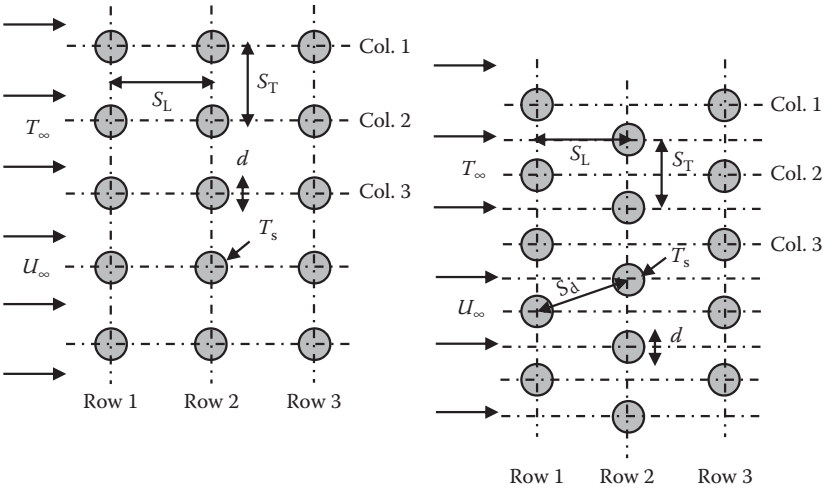


FIGURE 3.2.18 Schematics of in-line (or aligned arrangement) on the left and staggered arrangement on the right. ($a=S_T/d$; $b=S_L/d$.)

In Equation 3.2.99, all physical properties are evaluated at the arithmetic mean of the inlet and exit temperatures of the fluid, except Pr_s which is evaluated at the surface temperature, T_s, of the cylinders. The values of the constants c, p, m, and n are given in Tables 3.2.5 and 3.2.6 for in-line and staggered arrangements, respectively.

The maximum average velocity in between the tubes is used in the calculation of the Reynolds number, Re_d. The maximum velocities for the in-line and staggered arrangements are given by the following expressions:

$$\text{In-line}: U_{\max} = \frac{U_\infty S_T}{S_T - d}, \tag{3.2.100}$$

$$\text{Staggered}: S_d > \frac{S_T + d}{2}.U_{\max} = \frac{U_\infty S_T}{S_T - d}, \tag{3.2.101}$$

$$\text{Staggered}: S_d < \frac{S_T + d}{2}.U_{\max} = \frac{U_\infty S_T}{2(S_d - d)}, \tag{3.2.102}$$

$$S_d = \left[S_L^2 + \left(\frac{S_T}{2} \right)^2 \right]^{1/2}.$$

Equation 3.2.99 is prescribed for tube banks with 16 or more rows. When there are fewer than 16 rows, the heat transfer coefficient given by Equation 3.2.99 is multiplied by the correction factor c_1 defined by Equation 3.2.103 and given in Table 3.2.7.

$$h_N = c_1 h_{16}, \tag{3.2.103}$$

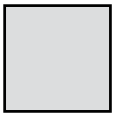
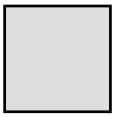
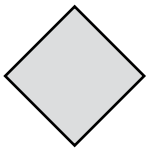
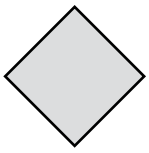
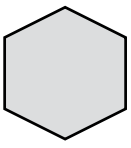
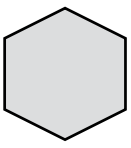
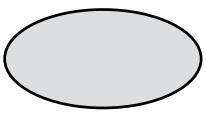
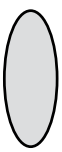
where

h_N=heat transfer coefficient with N rows (fewer than 16)

h_{16}=heat transfer coefficient for $N \geq 16$ (number of rows)

As expected, the correction factor c_1 approaches the value of unity as value of N approaches $N \to 16$.

TABLE 3.2.4
Values of *c* and *m* in Equation 3.2.98 (Jakob, 1949)

Geometry	Re_d	c	m
	$5000–10^5$	0.092	0.675
	2500–8000	0.160	0.699
	$5000–10^5$	0.222	0.588
	2500–7500	0.261	0.624
	$5000–1.95\times10^4$ $1.95\times10^4–10^5$	0.144 0.035	0.638 0.782
	$5000–10^5$	0.138	0.638
	$2500–1.5\times10^4$	0.224	0.612
	$3000–1.5\times10^4$	0.085	0.804
	$4000–1.5\times10^4$	0.205	0.731

Note: Characteristic dimension is the equivalent circular diameter = Perimeter/π. For example, for a square rod with each side a, $D = 4a/\pi$. Flow is from right to left.

TABLE 3.2.5
In-Line Arrangement—Values of Constants in Equation 3.2.99 (p=0 in All Cases)

Re_d	c	m	n
1–100	0.9	0.4	0.36
100–1000	0.52	0.5	0.36
$10^3-2\times10^5$	0.27	0.63	0.36
$2\times10^5-2\times10^6$	0.033	0.8	0.4

TABLE 3.2.6
Staggered Arrangement—Values of Constants in Equation 3.2.99

Re_d	c	p	m	n
1–500	1.04	0	0.4	0.36
500–1000	0.71	0	0.5	0.36
10^3–2×10^5	0.35	0.2	0.6	0.36
2×10^5–2×10^6	0.031	0.2	0.8	0.36

TABLE 3.2.7
Correction Factor c_1 to Be Used with Equation 3.2.103

	Number of Rows (N)							
Tube Arrangement	**1**	**2**	**3**	**4**	**5**	**7**	**10**	**13**
In-line	0.70	0.80	0.86	0.90	0.93	0.96	0.98	0.99
Staggered	0.64	0.76	0.84	0.89	0.93	0.96	0.98	0.99

Pressure Drop: With tube banks, pressure drop is a significant factor, as it determines the fan or pump power required to maintain the required fluid flow. Zukauskas (1987) recommends that the pressure drop be computed from the relation

$$\Delta p = p_i - p_e = N\chi \frac{\rho U_{max}^2}{2} f, \tag{3.2.104}$$

where p_i and p_e are the fluid pressures at the inlet and exit of the tube banks. The values of χ and f are presented in Figure 3.2.19a. In Figure 3.2.19a, the friction factor f for in-line arrangement is presented for different values of b (S_L/d) for $S_L=S_T$. For values of S_L/S_T other than 1, the correction factor χ is given in the inset for different values of $(a-1)/(b-1)$. Similarly, the friction factor for staggered arrangement (for equilateral triangle arrangement) and a correction factor for different values of a/b are also given in Figure 3.2.19b. The value of friction factor f is for one row of tubes; the total pressure drop is obtained by multiplying the pressure drop for one row by the number of rows N in the tube bank.

The temperature of the fluid varies in the direction of flow, and, therefore, the value of the convective heat transfer coefficient (which depends on the temperature-dependent properties of the fluid) also varies in the direction of flow. However, it is common practice to compute the total heat

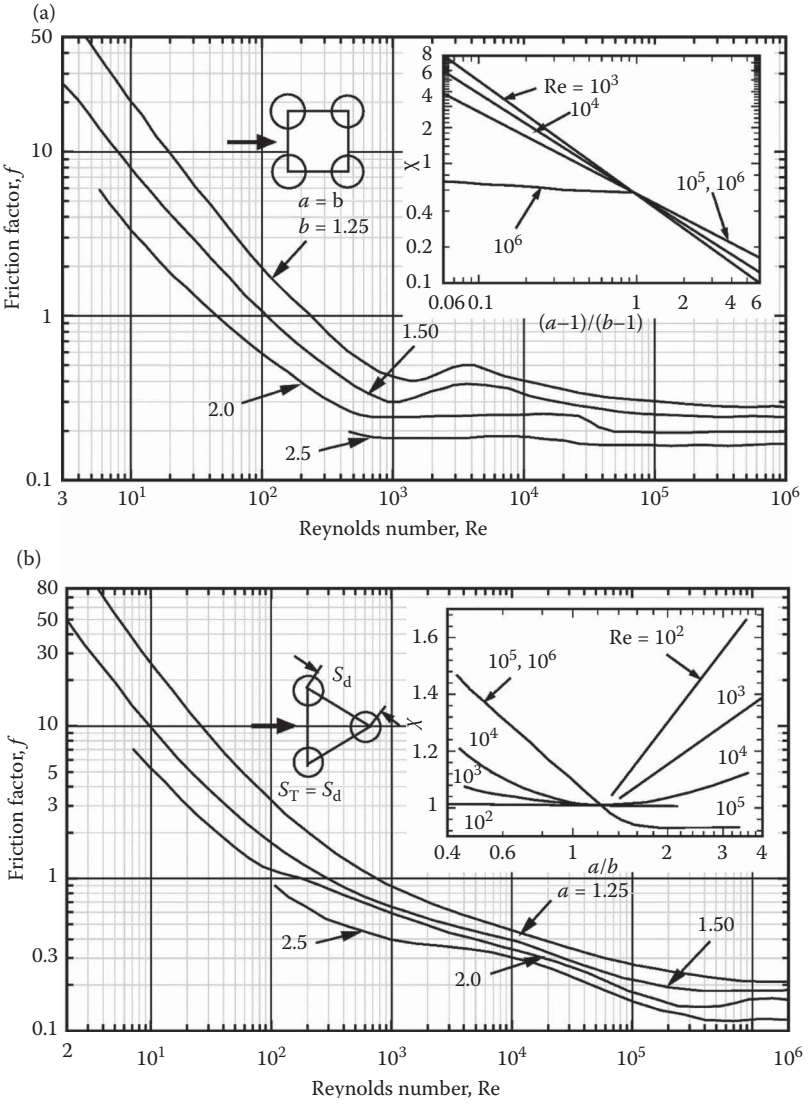


FIGURE 3.2.19 Friction factors for tube banks. (a) In-line arrangement; (b) staggered arrangement (Zukauskas, 1987).

transfer rate with the assumption of uniform convective heat transfer coefficient evaluated at the arithmetic mean of the inlet and exit temperatures of the fluid. With such an assumption of uniform convective heat transfer coefficient, uniform surface temperature and constant specific heat (evaluated at the mean fluid temperature), the inlet and exit fluid temperatures are related by

$$\ln\left(\frac{T_s - T_e}{T_s - T_i}\right) = -\frac{hA_s}{\dot{m}c_p}. \tag{3.2.105}$$

The heat transfer rate to the fluid is given by the equation

$$q = m \cdot c_p \left(T_e - T_i\right). \tag{3.2.106}$$

Example 3.2.4

A heat exchanger with aligned tubes is used to heat 40 kg/s of air from 10°C to 50°C with the tube surfaces maintained at 100°C. Details of the heat exchanger are as follows:

Diameter of tubes: 25 mm
Number of columns: 20
Length of each tube: 3 m
$S_L = S_T = 75$ mm
Determine the number of rows required for this application.

Solution: Average air temperature $= (T_i + T_e)/2 = (10+50)/2 = 30°C$. Properties of atmospheric air (from Suryanarayana, 1995):

$$\rho = 1.165 \text{ kg/m}^3;\ c_p = 1007 \text{ J/kg K}$$

$$\mu = 1.865 \times 10^{-5} \text{Pa} \cdot \text{s};\ k = 0.0264 \text{ W/mK}$$

$$\text{Pr} = 0.712;\ \text{Pr}_s \text{ (at 100°C)} = 0.705$$

To find U_{max}, we need the minimum area of cross section for fluid flow (Figure 3.2.20).
$H = 20 \times 0.075 = 1.5$ m
$A_{min} = 20(0.075 - 0.025) \times 3 = 3 \text{ m}^2$

$$U_{max} = \frac{\dot{m}}{\rho A_{min}} = \frac{40}{1.165 \times 3} = 11.44 \text{ m/s}$$

$$\text{Re}_d = \frac{\rho U_{max} d}{\mu} = \frac{1.165 \times 11.44 \times 0.025}{1.865 \times 10^{-5}} = 17{,}865$$

With values from Table 3.2.5, $c = 0.27$, $m = 0.63$, $n = 0.36$

$$\text{Nu}_d = 0.27 \times 17{,}865^{0.63} \times 0.712^{0.36} \left(\frac{0.712}{0.705}\right)^{0.25} = 114.3$$

$$h = \frac{114.3 \times 0.0264}{0.025} = 120.7 \text{ W/m}^2\text{K}$$

From Equation 3.2.105,

$$\ln\left(\frac{100-50}{100-10}\right) = -\frac{120.7 \times A_s}{40 \times 1007},$$

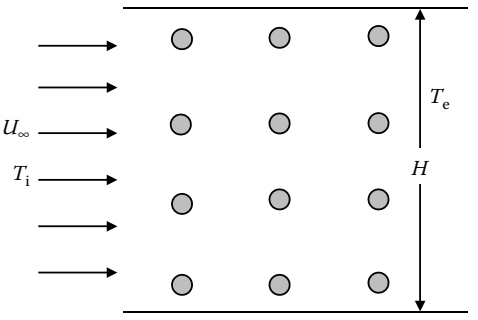


FIGURE 3.2.20 Aligned tube heat exchanger (only a few of the 20 columns and rows are shown).

where $A_s = \pi \times 0.025 \times 3 \times 20 \times N$
N = number of rows = 42
Fan Power: From the first law of thermodynamics (see Chapter 1), the fan power is

$$\dot{W}_F = \dot{m}\left(\frac{p_i}{\rho_i} - \frac{p_e}{\rho_e} + \frac{v_e^2}{2}\right).$$

p_i and p_e are the pressures at inlet and exit of the heat exchanger, and v_e is the fluid velocity at exit. Assuming constant density evaluated at $(T_i + T_e)/2$, the pressure drop is found from Figure 3.2.19a.

$$\mathrm{Re}_d = 17,865: \; a = b = 75/25 = 75/25 = 3.$$

In Figure 3.2.19a, although the friction factor is available for values of b up to 2.5, we will estimate the value of f for $b = 3$. From Figure 3.2.19a, $f \approx 0.11$. The correction factor $c = 1$.

$$p_i - p_e = N\chi \frac{\rho U_{max}^2}{2} f = 42 \times 1 \frac{1.165 \times 11.44^2}{2} \times 0.11 = 352.2 \text{ kPa},$$

$$v_e = \frac{11.44 \times 50}{75} = 7.63 \text{ m/s},$$

$$\dot{W}_F = 40\left(352.2 + \frac{7.63^2}{2}\right) = \underline{15,250 \text{ W}} \text{ or } \underline{15.25 \text{ kW}}.$$

Heat Transfer with Jet Impingement

Jet impingement (Figure 3.2.21) on a heated (or cooled) surface results in high heat transfer rates, and is used in annealing of metals, tempering of glass, cooling of electronic equipment, internal combustion engines, and in a wide variety of industries–textiles, paper, wood, and so on. Usually, the jets are circular, issuing from a round nozzle of diameter d, or rectangular, issuing from a slot of width w. They may be used singly or in an array. The jets may impinge normally to the heated surface or at an angle. If there is no parallel solid surface close to the heated surface, the jet is said to be free; in the presence of a parallel surface close to the heated surface, the jet is termed confined. In this section, only single, free jets (round or rectangular) impinging normally to the heated surface are considered.

Jets may be submerged within the fluid from the nozzle exiting into a body of a fluid (usually the same fluid), for example, air impinging on a surface surrounded by atmospheric air. In submerged jets, entrained fluid (the part of the surrounding fluid dragged by the jet) has a significant effect on the flow and heat-transfer characteristics of the jet, but the effect of gravity is usually negligible. In free-surface jets—a liquid jet in an atmosphere of air is a good approximation to a free-surface jet—the entrainment effect is usually negligible, but the effect of gravity may be significant.

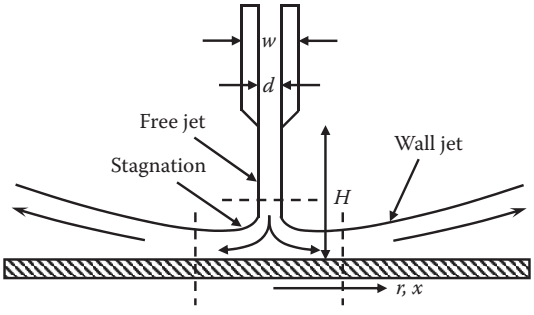


FIGURE 3.2.21 Circular jet of diameter d or a rectangular jet of width w.

A jet is usually divided into three regions: a free-jet region, a stagnation region, and a wall-jet region. In the free-jet region, the effect of the target surface on the flow is negligible. In the stagnation region, the target surface affects the flow field, and the velocity parallel to the surface increases while the velocity component normal to the surface decreases. At the beginning of the stagnation region, the axial velocity of the fluid is very much greater than the radial component (or the x-component) of the velocity. The stagnation region is followed by the wall-jet region where the radial component (or the x-component) of the velocity is much greater than the axial velocity.

The heat-transfer coefficient is a function of H/d (or H/w), Re_d ($\rho v_j d/\mu$) or ($\rho v_j 2w/\mu$), and Pr and depends on the region (stagnation or wall jet), whether it is submerged or non-submerged and whether the flow adjacent to the plate is laminar or turbulent. Some of the heat-transfer correlations suggested by different researchers are given below. All the correlations are for single jets.

Submerged Jets: Single Circular Jets

$$\mathrm{Re}_d = \frac{4\dot{m}}{\pi d\mu}; \quad \mathrm{Nu}_d = \frac{hd}{k},$$

where $\dot{m}$= mass rate of the fluid

Average heat-transfer coefficients up to radius r (Martin, 1990):

$$\mathrm{Nu}_d = 2\left(\frac{d}{r}\right)\frac{1-1.1(d/r)}{1+0.1\{(H/d)-6\}(d/r)}\left[\mathrm{Re}_d\left(1+\frac{\mathrm{Re}_d^{0.55}}{200}\right)\right]^{1/2}\mathrm{Pr}^{0.42}. \tag{3.2.107}$$

Range of validity:

$$2000 \le \mathrm{Re}_d \le 4 \times 10^5; 2.5 \le r/d < 7.5; 2 \le H/d \le 2.$$

Local convective heat-transfer coefficient at radius r (Webb and Ma, 1995):

$$\mathrm{Nu}_d = 1.29\mathrm{Re}_d^{1/2}\mathrm{Pr}^{0.4}\left\{\left[\frac{\tanh(0.88r/d)}{r/d}\right]^{-8.5} + \left[1.69\left(\frac{r}{d}\right)^{-1.07}\right]^{-17}\right\}^{-1/17}. \tag{3.2.108}$$

Submerged Jets: Single Rectangular Jet

$$\mathrm{Re}_w = \frac{\rho v_j(2w)}{\mu} = \frac{2\dot{m}}{\mu},$$

where $\dot{m}$ = mass rate of flow per unit length of jet

$$\mathrm{Nu}_w = \frac{h(2w)}{k}.$$

Average heat-transfer coefficient (Martin, 1990):

$$\mathrm{Nu}_w = \frac{1.53\mathrm{Pr}^{0.42}\mathrm{Re}_w^m}{\frac{x}{2w}+\frac{H}{2w}+1.39}, \tag{3.2.109}$$

where $m = 0.695 - \left[\frac{x}{2w}+\left(\frac{H}{2w}\right)^{1.33}+3.06\right]^{-1}$.

TABLE 3.2.8
Correlations for Free-Surface Jets $r_v/d = 0.1773\ Re_d^{1/3}$

Range of r	Range of Pr	Nu_d
$r/d < 0.787$	$0.15 \le Pr \le 3$	$0.715Re_d^{1/2}Pr^{0.4}$ (3.2.110)
	$Pr > 3$	$0.797Re_d^{1/2}Pr^{0.4}$ (3.2.111)
$0.787 < r/d < r_v/d$		$0.632Re_d^{1/2}Pr^{1/3}\left(\frac{d}{r}\right)^{1/2}$ (3.2.112)
$r_v/d < r/d < r_t/d$		$\frac{0.407Re_d^{1/3}Pr^{1/3}(d/r)^{2/3}}{\left[\frac{0.1713}{(r/d)^2}+\frac{5.147}{Re_d}\frac{r}{d}\right]^{2/3}\left[\frac{(r/d)^2}{2}+C\right]^{1/3}}$, (3.2.113)
		where
		$C = -5.051\times10^{-5}Re_d^{2/:}$
		$\frac{r_t}{d} = \left\{-\frac{s}{2}+\left[\left(\frac{s}{2}\right)^2+\left(\frac{p}{3}\right)^3\right]^{1/2}\right\}^{1/3} + \left\{-\frac{s}{2}+\left[\left(\frac{s}{2}\right)^2-\left(\frac{p}{3}\right)^3\right]^{1/2}\right\}^{1/3}$
		$p = \frac{-2C}{0.2058Pr-1}$ $s = \frac{0.00686Re_dPr}{0.2058Pr-1}$
$r > r_t$	$Pr < 3.9$	$\frac{0.25}{\frac{1}{Re_dPr}\left[1-\left(\frac{r_t}{r}\right)^2\right]\left(\frac{r}{d}\right)^2+0.13\frac{h}{d}+0.0371\frac{h_t}{d}}$, (3.2.114)
		where $h_t = h$ at r_t and $\frac{h}{d} = \frac{0.1713}{r/d}+\frac{5.147}{Re_d}\left(\frac{r}{d}\right)^2$

Free-Surface Jets: Single Circular Jet. Correlations are given in Table 3.2.8 (Liu et al., 1991; Webb and Ma, 1995).

For more information on jet impingement heat transfer, refer to Martin (1977) and Webb and Ma (1995) and the references in these two papers.

ACKNOWLEDGMENT

The preparation of this section was greatly facilitated by the award of the JC Bose Fellowship (Department of Science & Technology, Government of India, New Delhi) to R.P. Chhabra for the period 2015–2020. This support is gratefully acknowledged.

BIBLIOGRAPHY

ASHRAE. 1993. *Handbook of Fundamentals*. American Society of Heating, Ventilating and Air Conditioning Engineers, Atlanta, GA.

Hewitt, G.F. 1990. *Handbook of Heat Exchanger Design*, Hemisphere Publishing, New York.

Bergman, T.L., Lavine, A.S., Incropera, F.P. and Dewitt, D.P. 2011. *Fundamentals of Heat and Mass Transfer*, 7th ed., John Wiley & Sons, New York.

Kakaç, S., Shah, R.K. and Win Aung, Eds. 1987. *Handbook of Single Phase Convective Heat Transfer*, Wiley-Interscience, New York.

Kakac, S., Yener, Y. and Pramuanjaroenkij, A. 2014. *Convective Heat Mass Transfer*, 3rd ed., CRC Press, Boca Raton.

Kreith, F., Manglik, R.M. and Bohn, M.S. 2009. *Principles of Heat Transfer*, 7th ed., PWS, Boston.

Suryanarayana, N.V. 1995. *Engineering Heat Transfer*, PWS, Boston.

REFERENCES

Achenbach, E. 1978. Heat transfer from spheres up to $Re = 6 \times 10^6$, in *Proceedings of the 6th International Heat Transfer Conference*, Vol. 5, Hemisphere Publishing, Washington, DC.

Burmeister, L.C. 1993. *Convective Heat Transfer*, Wiley, New York.

Churchill, S.W. 1976. A comprehensive correlation equation for forced convection from a flat plate, *AIChE J.* 22: 264–268.

Churchill, S.W. and Bernstein, M. 1977. A correlating equation for forced convection from gases and liquids to a circular cylinder in cross flow, *Journal of Heat Transfer* 99: 300–306.

Churchill, S.W. and Ozoe, H. 1973. Correlations for laminar forced convection with uniform heating in flow over a plate and in developing and fully developed flow in a tube, *Journal of Heat Transfer* 95: 78–84.

Clift, R., Grace, J. and Weber, M. 1978. *Bubbles Drops and Particles*, Academic Press, New York.

Eckert, E.R.G. 1956. Engineering relations for heat transfer and friction in high-velocity laminar and turbulent boundary-layer flow over surfaces with constant pressure and temperature, *Transactions of ASME* 78: 1273–1284.

Eckert, E.R.G. and Drake, M., Jr. 1972. *Analysis of Heat and Mass Transfer*, McGraw-Hill, New York.

Ishiguro, R., Sugiyama, K. and Kumada, T. 1979. Heat transfer around a circular cylinder in a liquid-sodium cross flow, *International Journal of Heat and Mass Transfer* 22: 1041–1048.

Jakob, H., 1949. *Heat Transfer*, John Wiley & Sons, London.

Kakac, S., Yener, Y. and Pramuanjaroenkij, A. 2014. *Convective Heat Mass Transfer*, 3rd ed., CRC Press, Boca Raton.

Kays, W.M., Crawford, M.E. and Weigand, B. 2004. *Convective Heat and Mass Transfer*, 4th ed., McGraw-Hill, New York.

Liu, X., Lienhard V, J.H. and Lombara, J.S. 1991. Convective heat transfer by impingement of circular liquid jets, *Journal of Heat Transfer* 113: 571–582.

Martin, H. 1977. Heat and mass transfer between impinging gas jets and solid surfaces, *Advances in Heat Transfer* 13: 1–60.

Martin, H. 1990. Impinging jets, in *Handbook of Heat Exchanger Design*, Hewitt, G.F., Ed., Hemisphere, New York.

Michaelides, E.E. 2006. Particles, *Bubbles and Drops – Their Motion, Heat and Mass Transfer*, World Scientific, Singapore.

Morgan, V.T., 1975. The overall convective heat transfer from smooth circular cylinders, *Advances in Heat Transfer* 11: 199–264.

Polyanin, A.D. Kutepov, A.M., Vyazmin, A.V. and Kazenin, D.A. 2002. *Hydrodynamics, Mass and Heat Transfer in Chemical Engineering*, Taylor and Francis, New York.

Rose, J.W. 1979. Boundary layer flow on a flat plate, *International Journal of Heat and Mass Transfer* 22: 969.

Schlichting, H. and Gersten, K. 2011. *Boundary Layer Theory*, 8th ed., Springer, New York.

Suryanarayana, N.V. 1995. *Engineering Heat Transfer*, West Publishing, Eagan, MN.

Thomas, W.C. 1977. Note on the heat transfer equation for forced-convection flow over a flat plate with an unheated starting length, *Mechanical Engineering News* (*ASEE*) 9(1): 19.

Webb, B.W. and Ma, C.F. 1995. Single-phase liquid jet impingement heat transfer, *Advances in Heat Transfer* 26: 105–217.

Whitaker, S. 1972. Forced convection heat transfer correlations for flow in pipes, past flat plates, single cylinders, single spheres, and for flow in packed beds and tube bundles, *AIChE J.* 18: 361–371.

Witte, L.C. 1968. An experimental study of forced-convection heat transfer from a sphere to liquid sodium, *Journal of Heat Transfer* 90: 9–12.

Zdravkovich, M.M. 1997. *Flow around Circular Cylinders*, *Vol. 1, Fundamentals*, Oxford University Press, New York.

Zdravkovich, M.M. 2003. *Flow around Circular Cylinders*, *Vol. 2, Applications*, Oxford University Press, New York.

Zukauskas, A. 1987. Convective heat transfer in cross flow, in *Handbook of Single-Phase Convective Heat Transfer*, Kakaç, S., Shah, R.K. and Win Aung, Eds., Wiley, New York.

3.2.3 Forced Convection: Internal Flows

Anoop K. Gupta, Raj P. Chhabra, and N.V. Suryanarayana

Introduction

Heat transfer to (or from) a fluid flowing inside a tube or duct is termed *internal forced convection*. The fluid flow may be laminar or turbulent depending upon the value of the duct Reynolds number. If the Reynolds number based on the average velocity of the fluid and diameter of the tube ($\rho v d/\mu$) is less than 2100 (Reynolds numbers in the range of 2000–2300 are cited in different sources), the flow is laminar. If the Reynolds number is >5000 or so, the flow is turbulent. The flow with a Reynolds number in the range 2100–5000 is considered to be in the transitional regime. With heating or cooling of the fluid, there may or may not be a change in the phase of the fluid. Here, only heat transfer to or from a single-phase fluid is considered, i.e., without boiling or condensation.

Fully Developed Velocity and Temperature Profiles

When a fluid enters a tube from a large reservoir, the velocity profile at the entrance is almost uniform as shown in Figure 3.2.22. The fluid in the immediate vicinity of the tube surface is decelerated in order to comply with the no-slip condition at the wall and the velocity increases from zero at the surface to u_c at a distance δ from the surface (i.e., inside the hydrodynamic boundary layer formed on the wall of the duct); in the region $r=0$ to $(R-\delta)$ the velocity is uniform. The value of δ increases in the direction of flow, and with constant fluid density, the value of the uniform velocity u_c increases. At some location downstream, δ reaches its maximum possible value, equal to the radius of the tube, and from that point onward the velocity profile does not change in the axial direction. This condition corresponds to the hydrodynamically developed flow.

The region where δ increases, i.e., where the velocity profile changes, is known as the entrance region or hydrodynamically developing region. The region downstream from the axial location where δ reaches its maximum value and the velocity profile does not change is the fully developed velocity profile or hydrodynamically fully developed region. Similarly, downstream of the location where heating or cooling of the fluid starts, the temperature profile changes in the direction of flow. But beyond a certain distance, the dimensionless temperature profile does not change in the direction of flow. The region where the dimensionless temperature profile changes is the thermally developing region or the thermal entrance region, and the region where the dimensionless temperature profile does not change is the thermally fully developed region. For simultaneously developing velocity and temperature profiles in laminar flows, the hydrodynamic and thermal entrance lengths for a circular pipe are given by

$$\frac{L_e}{d} = 0.0565\mathrm{Re}_d \tag{3.2.115}$$

$$\frac{L_{e,th}}{d} = 0.053\mathrm{Re}_d\mathrm{Pr} \quad \text{(Uniform heat flux)} \tag{3.2.116}$$

$$\frac{L_{e,th}}{d} = 0.037\mathrm{Re}_d\mathrm{Pr} \quad \text{(Uniform surface temperature)}. \tag{3.2.117}$$

In most engineering applications, with turbulent flows, correlations for fully developed conditions can be used after about 10–20 diameters down stream from where the heating starts. More details can be found in Kakac et al. (2014).

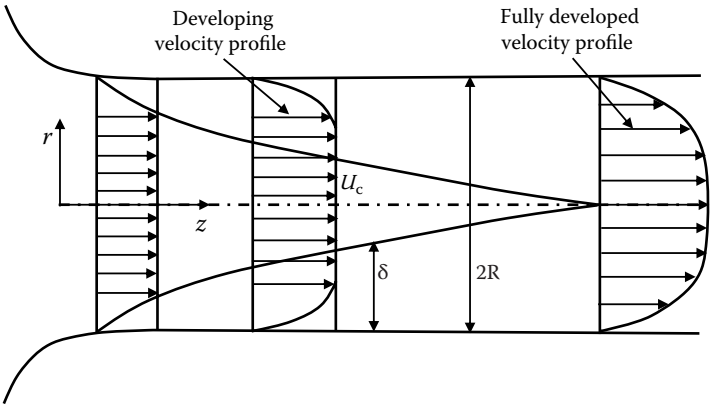


FIGURE 3.2.22 Developing and fully developed velocity profiles.

Convective Heat Transfer Coefficient and Bulk Temperature

The reference temperature for defining the convective heat-transfer coefficient is the bulk temperature T_b, and the convective heat flux is given by

$$q'' = h\left(T_s - T_b\right). \tag{3.2.118}$$

The bulk temperature T_b is determined from the relation

$$T_b = \frac{\int_{A_c} \rho v C_p T \, dA_c}{\int_{A_c} \rho v C_p \, dA_c}, \tag{3.2.119}$$

where A_c is the cross-sectional area perpendicular to the axis of the tube.

If the fluid is drained from the tube at a particular axial location and thoroughly mixed, the temperature of the mixed fluid is the bulk temperature. It is also known as the mixing cup temperature. With heating or cooling of the fluid, the bulk temperature varies in the direction of flow. In some cases, we use the term *mean fluid temperature*, T_m, to represent the arithmetic mean of the fluid bulk temperatures at inlet and exit of the tube.

Heat Transfer Correlations

Laminar Flows: Entrance Region

For laminar flows in a tube with uniform surface temperature, in the entrance region, the correlation of Sieder and Tate (1936) is

$$\overline{\mathrm{Nu}}_d = 1.86\left(\frac{\mathrm{Re}_d \mathrm{Pr}}{L/d}\right)^{1/3}\left(\frac{\mu}{\mu_s}\right)^{0.14} \tag{3.2.120}$$

valid for

$$\frac{L}{d} < \frac{\mathrm{Re}_d \mathrm{Pr}}{8}\left(\frac{\mu}{\mu_s}\right)^{0.42}; \; 0.48 < \mathrm{Pr} < 16{,}700; \; 0.0044 < \frac{\mu}{\mu_s} < 9.75.$$

The overbar in the Nusselt number indicates that it is formed with the average heat-transfer coefficient over the entire length of the tube. Properties of the fluid are evaluated at the arithmetic mean of the inlet and exit bulk temperatures. The (μ / μ_s) term accounts for the variation of the viscosity across the diameter of the tube. In Equation 3.2.120, the heat-transfer coefficient was determined from

$$q = \bar{h}\pi dL\left(T_s - \frac{T_{bi} + T_{be}}{2}\right). \tag{3.2.121}$$

Therefore, to find the total heat-transfer rate with $\bar{h}$ from Equation 3.2.120, one needs to employ Equation 3.2.121 to account for the axial variation in the available driving force for heat transfer.

Alternate empirical correlations have been proposed by Hausen (1943) for a range of conditions in the laminar entrance flow in a tube. For instance, for the case of the constant wall temperature and fully developed velocity profile, Hausen (1943) proposed

$$\mathrm{Nu_d} = 3.66 + \frac{0.0668(d/z)\mathrm{Pe}}{1 + 0.04[(d/z)\mathrm{Pe}]^{2/3}}. \tag{3.2.122}$$

Equation 3.2.122 gives the mean Nusselt number over the length z of the pipe (measured from the inlet). Naturally, for large values of $(z/d)(1/\mathrm{Pe})$, Equation 3.2.122 yields the expected limiting value of $\mathrm{Nu_d} = 3.66$.

On the other hand, when both the velocity and temperature fields are developing simultaneously, Hausen (1943) presented the following approximate expression for Nusselt number for the case of constant heat input and parabolic velocity profile:

$$\mathrm{Nu_d} = 4.36 + \frac{0.023(d/z)\mathrm{Pe}}{1 + 0.0012[(d/z)\mathrm{Pe}]^{0.8}}, \tag{3.2.123}$$

where

$$\mathrm{Pe} = \mathrm{Re}\cdot\mathrm{Pr} = \left(\rho v d C_p / k\right).$$

Laminar Flows: Fully Developed Velocity and Temperature Profiles

Evaluate properties at the bulk temperature.

$$\text{Uniform Surface Temperature}: \mathrm{Nu_d} = 3.66, \tag{3.2.124}$$

$$\text{Uniform Surface Heat Flux}: \ \mathrm{Nu_d} = 4.36. \tag{3.2.125}$$

Turbulent Flows

If the flow is turbulent, the difference between the correlations with uniform surface temperature and uniform surface heat flux is not significant and the correlations can be used for both cases. For turbulent flows, Gnielinski (1976, 1990) recommends the following:

Evaluate properties at the bulk temperature.

$$0.6 < \mathrm{Pr} < 2000; \quad 2300 < \mathrm{Re_d} < 10^6; \quad 0 < d/L < 1$$

$$\mathrm{Nu_d} = \frac{(f/2)\left(\mathrm{Re_d} - 1000\right)\mathrm{Pr}}{1 + 12.7\left(f/2\right)\left(\mathrm{Pr}^{2/3} - 1\right)}\left[1 + \left(\frac{d}{L}\right)^{2/3}\right], \tag{3.2.126}$$

where

$$f = \left[1.58 \ln(Re_d) - 3.28\right]^{-2}. \tag{3.2.127}$$

f = fanning friction factor = $2\tau_w/\rho v^2$.

To reflect the effect of variation of fluid properties with temperature, multiply the Nusselt numbers in Equation 3.2.126 by $(T_b/T_s)^{0.45}$ for gases and $(\mathrm{Pr}/\mathrm{Pr}_s)^{0.11}$ for liquids where the temperatures are absolute, and T and Pr with a subscript s are to be evaluated at the surface temperature. The equations can be used to evaluate the heat transfer coefficient in the developing profile region. To determine the heat-transfer coefficient in the fully developed region, set $d/L = 0$, i.e., $L \to \infty$. A simpler correlation (fully developed region) is the so-called Dittus–Boelter (1930) equation. Evaluate properties at T_b.

$$0.7 \le \mathrm{Pr} \le 160;\ \mathrm{Re}_d > 10{,}000;\ d/L > 10$$

$$\mathrm{Nu}_d = 0.023 \mathrm{Re}_d^{0.8} \mathrm{Pr}^n, \tag{3.2.128}$$

where $n = 0.4$ for heating ($T_s > T_b$) and $n = 0.3$ for cooling ($T_s < T_b$).

For liquid metals with Pr << 1, the following correlations due to Sleicher and Rouse (1975) are recommended:

Uniform surface temperature:

$$\mathrm{Nu}_{d,b} = 4.8 + 0.0156 \mathrm{Re}_{d,f}^{0.85} \mathrm{Pr}_s^{0.93}, \tag{3.2.129}$$

Uniform heat flux:

$$\mathrm{Nu}_{d,b} = 6.3 + 0.0167 \mathrm{Re}_{d,f}^{0.85} \mathrm{Pr}_s^{0.93}. \tag{3.2.130}$$

Subscripts b, f, and s indicate that the variables are to be evaluated at the bulk temperature, film temperature (arithmetic mean of the bulk and surface temperatures), and surface temperature, respectively.

In the computations of the Nusselt number, the thermophysical properties (evaluated at the bulk temperature) vary in the direction of flow and hence give different values of h at different locations. In many cases, a representative average value of the convective heat-transfer coefficient is needed. Such an average value can be obtained either by taking the arithmetic average of the convective heat-transfer coefficients evaluated at the inlet and exit bulk temperatures or the convective heat-transfer coefficient evaluated at the arithmetic mean of the inlet and exit bulk temperatures. If the variation of the convective heat-transfer coefficient is large, it may be appropriate to divide the tube into shorter lengths with smaller variation in the bulk temperatures and evaluating the average heat-transfer coefficient in each section and eventually adding up the individual rate of heat transfer to estimate the overall rate of heat transfer.

Uniform Surface Temperature: Relation between the Convective Heat Transfer Coefficient and the Total Heat Transfer Rate

With a uniform surface temperature, employing an average value of the convective heat-transfer coefficient, the local convective heat flux varies in the direction of flow. To relate the convective heat-transfer coefficient to the temperatures and the surface area, we have, for the elemental length Δz (Figure 3.2.23),

$$\dot{m} C_p \frac{dT_b}{dz} = h \frac{dA_s}{dz} \left(T_s - T_b\right). \tag{3.2.131}$$

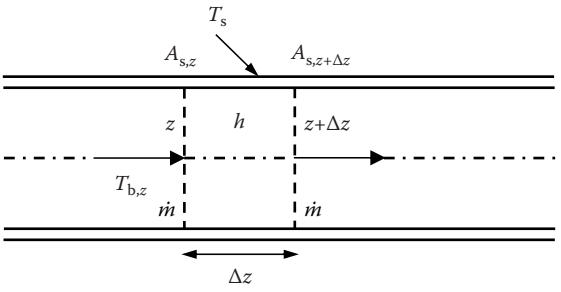


FIGURE 3.2.23 Elemental length of a tube for determining heat-transfer rate.

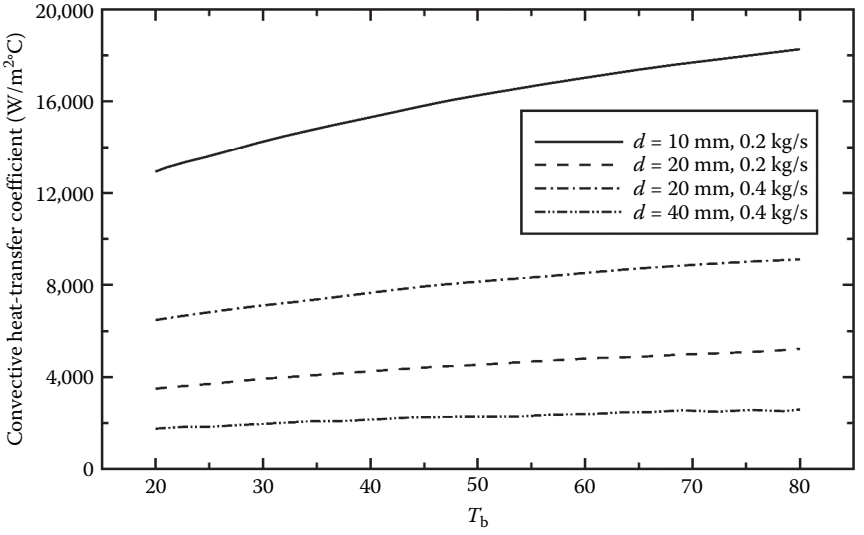


FIGURE 3.2.24 Variation of h with T_b in 1-, 2-, and 4-cm-diameter tubes with water flow rates of 0.2 kg/s and 0.4 kg/s with uniform surface temperature.

Assuming a suitable average convective heat-transfer coefficient over the entire length of the tube, separating the variables, and integrating the equation from $z=0$ to $z=L$, we obtain

$$\ln\frac{T_s - T_{be}}{T_s - T_{bi}} = -\frac{hA_s}{\dot{m}C_p}. \tag{3.2.132}$$

Equation 3.2.132 gives the exit temperature. For a constant-density fluid or an ideal gas, the heat-transfer rate is determined from

$$q = \dot{m}C_p\left(T_{be} - T_{bi}\right). \tag{3.2.133}$$

Equation 3.2.132 was derived on the basis of uniform convective heat-transfer coefficient. However, if the functional relationship between h and T_b is known, Equation 3.2.131 can be integrated by substituting the relationship. The convective heat-transfer coefficient variation with T_b for water in two tubes of different diameters for two different flow rates is shown in Figure 3.2.24. From the figure, it is clear that the heat-transfer coefficient, h, can be very well approximated as a linear function of T. By substituting such a linear function relationship into Equation 3.2.131, it can be shown that

$$\ln \frac{h_i}{h_e} \frac{T_s - T_{be}}{T_s - T_{bi}} = -\frac{h_s A_s}{\dot{m} C_p}, \tag{3.2.134}$$

where h_i, h_e, and h_s are the values of the convective heat-transfer coefficient evaluated at bulk temperatures of T_{bi}, T_{be}, and T_s, respectively. Although it has been demonstrated that h varies approximately linearly with the bulk temperature with water as the fluid, the variation of h with air or oil as the fluid is much smaller and is very well approximated by a linear relationship. For other fluids, it is suggested that the relationship be verified before employing Equation 3.2.134. (*Note:* It is tempting to determine the heat-transfer rate from the relation

$$q = hA_s \frac{\left(T_s - T_{be}\right) + \left(T_s - T_{bi}\right)}{2}. \tag{3.2.135}$$

Replacing q by Equation 3.2.133 and solving for T_{be} for defined values of the mass flow rate and tube surface area, the second law of thermodynamics will be violated if $hA_s/\dot{m}C_p > 2$. Use of Equation 3.2.132 or 3.2.134 ensures that no violation of the second law occurs however large A_s is.)

Uniform Surface Heat Flux

If the imposed heat flux is known, the total heat-transfer rate for a defined length of the tube is also known. From Equation 3.2.133, the exit temperature of the fluid is determined. The fluid temperature at any location in the pipe is known from the heat-transfer rate up to that location ($q = q''A_s$) and Equation 3.2.133. The convective heat-transfer coefficient is used to find the surface temperature of the tube.

Temperature Variation of the Fluid with Uniform Surface Temperature and Uniform Heat Flux

The fluid temperature variations in the two cases are different. With the assumption of uniform heat-transfer coefficient, with a uniform surface temperature the heat flux decreases in the direction of flow leading to a progressively decreasing rate of temperature change in the fluid with axial distance. With uniform heat flux, the surface and fluid temperatures vary linearly except in the entrance region where the higher heat-transfer coefficient leads to a smaller difference between the surface and fluid temperatures. The variation of the fluid temperature in the two cases is shown in Figure 3.2.25.

Convective Heat Transfer in Noncircular Tubes

Laminar flows: The Nusselt numbers for laminar flows have been analytically determined for different noncircular ducts. Some of them can be found in Shah and London (1978), Kakac et al. (1987, 2014), Kays et al. (2004), and Burmeister (1993). A few of the results are given later. The characteristic length for the definition of the Reynolds number and Nusselt number is the hydraulic mean diameter defined as

$$d_h = \frac{4 \times \text{cross-sectional area}}{\text{wetted perimeter}}. \tag{3.2.136}$$

Infinite parallel plates: a = spacing between plates, $d_h = 2a$
Both plates maintained at uniform and equal temperatures: Nu = 7.54
Both plates with imposed uniform and equal heat fluxes: Nu = 8.24
Rectangular ducts: a = longer side, b = shorter side, $d_h = 2ab/(a+b)$

b/a	1	0.7	0.5	0.25	0.125
Uniform surface temperature	2.98	3.08	3.39	4.44	5.6
Uniform heat flux[a]	3.61	3.73	4.12	5.33	6.49

[a] Uniform axial heat flux but circumferentially uniform surface temperature.

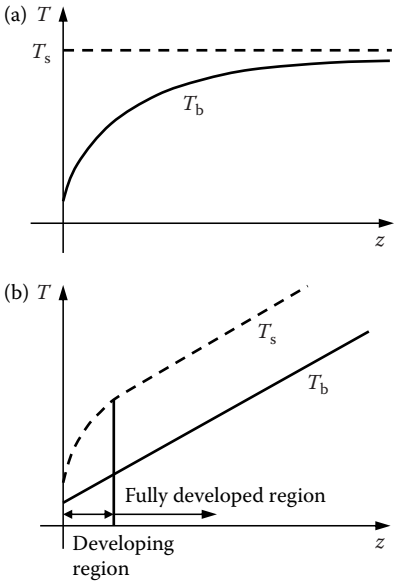


FIGURE 3.2.25 Variation of fluid temperature in a tube with (a) uniform surface temperature and (b) uniform heat flux.

Equilateral triangle: $d_h = a/3^{1/2}$, a = length of each side
Uniform surface temperature: Nu = 2.35
Uniform surface heat flux: *Nu = 3.0

Coaxial tubes: With coaxial tubes many different cases arise—each tube maintained at uniform but different temperatures, each tube subjected to uniform but different heat fluxes (an insulated surface is a special case of imposed heat flux being zero), or a combinations of uniform surface temperature of one tube and heat flux on the other. The manner in which the heat-transfer coefficient is determined for uniform but different heat fluxes on the two tubes is described below. Define:

$$d_h = 2(r_o - r_i); \qquad r^* = r_i/r_o\,,$$

$$q_i'' = h_i(T_i - T_b); \mathrm{Nu}_i = \frac{h_i d_h}{k}; q_o'' = h_o(T_o - T_b); \mathrm{Nu}_o = \frac{h_o d_h}{k},$$

$$q_o'' = 0; \quad \mathrm{Nu}_{ii} = \frac{h_i d_h}{k} \text{ and } q_i'' = 0; \quad \mathrm{Nu}_{oo} = \frac{h_o d_h}{k},$$

then

$$\mathrm{Nu}_i = \frac{\mathrm{Nu}_{ii}}{1 - \dfrac{q_o''}{q_i''}\theta_i^*} \text{ and } \mathrm{Nu}_o = \frac{\mathrm{Nu}_{oo}}{1 - \dfrac{q_i''}{q_o''}\theta_o^*}. \tag{3.2.137}$$

Some of the values needed for the computations of Nu_i and Nu_o (taken from Kays et al., 2004) are given in Table 3.2.9.

TABLE 3.2.9
Values for Use with Equation 3.2.137

r^*	Nu_{ii}	Nu_{oo}	θ_i^*	θ_o^*
0.05	17.81	4.792	2.18	0.0294
0.1	11.91	4.834	1.383	0.0562
0.2	8.499	4.883	0.905	0.1041
0.4	6.583	4.979	0.603	0.1823
0.6	5.912	5.099	0.473	0.2455
0.8	5.58	5.24	0.401	0.299
1.0	5.385	5.385	0.346	0.346

For a more detailed information on heat transfer and friction factors for laminar flows in noncircular tubes, refer to Shah and London (1978) and Kakac et al. (1987, 2014).

Turbulent flows: For noncircular tubes, estimates of the convective heat-transfer coefficient can be obtained by employing equations for circular tubes with d_h replacing d in the computations of the Reynolds and Nusselt numbers. To determine the heat-transfer coefficients in developing regions and for more accurate values with turbulent flows in noncircular tubes refer to Kakac et al. (1987, 2014) and the references in that book.

Mixed Convection

If the fluid velocity is low, the effect of natural convection becomes significant and the heat-transfer rate may be increased or decreased by natural convection. From a review of experimental results, Metais and Eckert (1964) developed maps to delineate the different regimes where one or the other mode is dominant and where both are significant. Figures 3.2.26 and 3.2.27 show the relative significance of natural and forced convection in vertical and horizontal tubes. The maps are applicable for $10^{-2} < \Pr(d/L) < 1$ where d and L are the diameter and the axial length of the tube. The maps show the limits of forced and natural convection regimes. The limits are delineated "in such a way that the actual heat flux under the combined influence of the forces does not deviate by more than 10% from the heat flux that would be caused by the external forces alone or by the body forces alone." The Grashof number is based on the diameter of the tube.

For flows in horizontal tubes, correlations were developed for the mixed convection regime in isothermal tubes by Depew and August (1971) and for uniform heat flux by Morcos and Bergles (1975).

Uniform Surface Temperature

Fully developed velocity profile, developing temperature profile:

$$L/d < 28.4;\ 25 < \mathrm{Gz} < 712;\ 0.7 \times 10^5 < \mathrm{Gr} < 9.9 \times 10^5$$

μ_s = dynamic viscosity, evaluated at the wall temperature

All other properties at the average bulk temperature of the fluid

$$\mathrm{Gz} = \frac{\dot{m}C_p}{kL} \quad \mathrm{Gr} = g\beta\Delta T d^3/\nu^2,$$

$$\mathrm{Nu}_d = 1.75\left[\mathrm{Gz} + 0.12\left(\mathrm{GzGr}^{1/3}\mathrm{Pr}^{0.36}\right)^{0.88}\right]^{1/3}(\mu_b/\mu_s)^{0.14}. \tag{3.2.138}$$

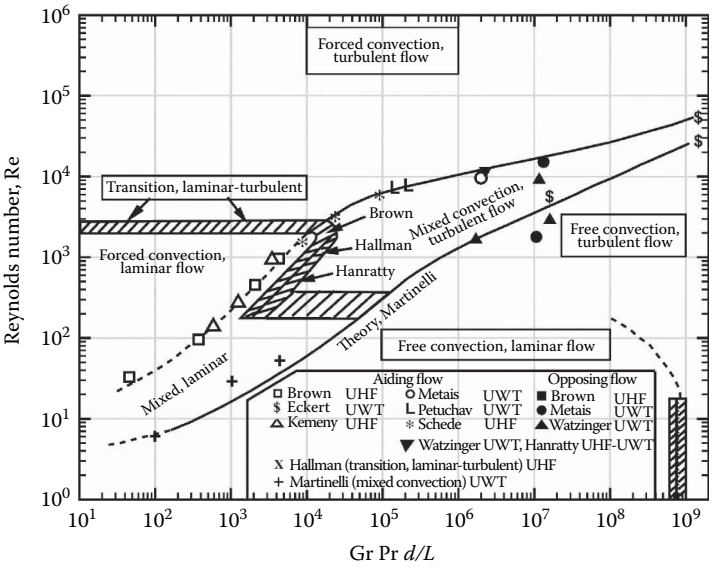


FIGURE 3.2.26 Map delineating forced, mixed, and natural convection—vertical tubes.

Uniform Heat Flux

Properties at $(T_s+T_b)/2$: $3\times10^4 < \mathrm{Ra} < 10^6$, $4 < \mathrm{Pr} < 175$, $2 < hd^2/(k_w t) < 66$, k_w=tube wall thermal conductivity, t=tube wall thickness.

$$\mathrm{Gr}_d^* = g\beta d^4 q_w''/(v^2 k); \qquad P_k = kd/(k_w t); \qquad \mathrm{Ra}_d = g\beta\Delta T d^3 \mathrm{Pr}/v^2,$$

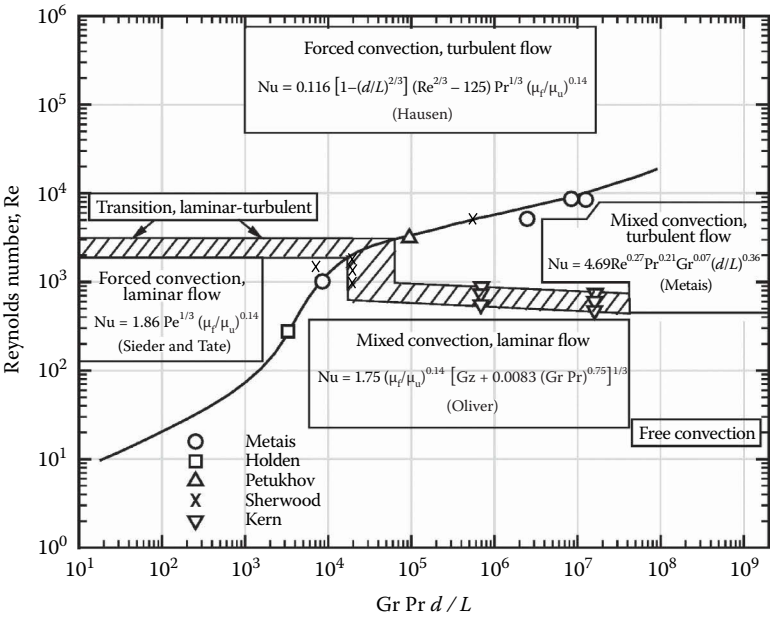


FIGURE 3.2.27 Map delineating forced, mixed, and natural convection—horizontal tubes.

$$\mathrm{Nu_d} = \left\{ 4.36^2 + \left[0.145 \left(\frac{\mathrm{Gr_d^* Pr}^{1.35}}{P_w^{0.25}} \right)^{0.265} \right]^2 \right\}^{0.5} . \tag{3.2.139}$$

In Equations 3.2.138 and 3.2.139, evaluate fluid properties at the arithmetic mean of the bulk and wall temperatures.

ACKNOWLEDGMENT

The preparation of this section was greatly facilitated by the award of the JC Bose Fellowship (Department of Science & Technology, Government of India, New Delhi) to R.P. Chhabra for the period 2015–2020. This support is gratefully acknowledged.

REFERENCES

Burmeister, L.C. 1993. *Convective Heat Transfer*, 2nd ed., Wiley, New York.

Depew, C.A. and August, S.E. 1971. Heat transfer due to combined free and forced convection in a horizontal and isothermal tube, *Transactions of the ASME* 93: 380–384.

Dittus, F.W. and Boelter, L.M.K. 1930. Heat transfer in automobile radiators of the tubular type, *University of California Publications in Engineering* 2: 443–461.

Gnielinski, V. 1976. New equations for heat and mass transfer in turbulent pipe channel flow, *International Chemical Engineering* 16: 359–368.

Gnielinski, V. 1990. Forced convection in ducts, in *Handbook of Heat Exchanger Design*, Hewitt, G.F., Ed., Begell House/Hemisphere, New York.

Hausen, H. 1943. Darstellung des Wärmeiiberganges in Rohren durch verallgemeinerte Potenzbeziehungen, *Z. Ver deutsch, Ing. Beih. Verfahrenstech.*, 4, 91–95.

Kakac, S., Shah, R.K. and Aung, W. 1987. *Handbook of Single-Phase Convective Heat Transfer*, Wiley, New York.

Kakac, S., Yener, Y. and Pramuanjaroenkij, A. 2014. *Convective Heat Transfer*, 3rd ed., CRC Press, Boca Raton.

Kays, W.M., Crawford, M.E. and Weigand, B. 2004. *Convective Heat and Mass Transfer*, 4th ed., McGraw-Hill, New York.

Metais, B. and Eckert, E.R.G. 1964. Forced, mixed, and free convection regimes, *Transactions of the ASME* 86: 295–296.

Morcos, S.M. and Bergles, A.E. 1975. Experimental investigation of combined forced and free laminar convection in a horizontal tube, *Transactions of the ASME* 97: 212–219.

Shah, R.K. and London, A.L. 1978. *Laminar Flow Forced Convection in Ducts*, Advances in Heat Transfer (Supp. 1), Academic Press, New York.

Sieder, E.N. and Tate, C.E. 1936. Heat transfer and pressure drop of liquids in tubes, *Industrial & Engineering Chemistry* 28: 1429–1435.

Sleicher, C.A. and Rouse, M.W. 1975. A convenient correlation for heat transfer to constant and variable property fluids in turbulent pipe flow, *International Journal of Heat and Mass Transfer* 18: 677–683.

3.2.4 Convection Heat Transfer in Non-Newtonian Fluids

Swati A. Patel, Raj P. Chhabra, Thomas F. Irvine, Jr., and Massimo Capobianchi

Introduction

The general characteristics of non-Newtonian fluids are described in Section 2.9 and will not be repeated here. Topics to be included in this section are laminar and turbulent heat transfer in fully developed duct flows, and laminar convection heat transfer in vertical channels and in external boundary layer flows over plates and several other common geometries.

For non-Newtonian flows, except for certain classes of fluids which exhibit a slip phenomenon at solid boundaries, the boundary condition is taken as no-slip or zero (tangential) velocity at all solid surfaces. For heat-transfer analyses, however, the situation is more complicated because there are many different ways to heat a wall, which in turn affects the type of thermal boundary conditions prescribed on solid boundaries.

In general, the rate of heat transfer from a surface, or the temperature difference between the wall and the fluid, is calculated using the equation $q_c = h_c A_q \Delta T$. Since the heat-transfer coefficient can vary considerably for different thermal boundary conditions or due to temperature-dependent properties or due to significant viscous dissipation effects, it is important that the conditions and other restrictions be specified correctly. Although the number of thermal boundary conditions is, in principle, infinite, several classical types have been identified and are in common use. They are usually identified in terms of the Nusselt number, $\mathrm{Nu} = h_c L/k$, with a particular subscript. For example, for duct flow, the symbol Nu_T is used to specify the Nusselt number when the wall temperature is constant in both the flow and peripheral directions. Other thermal boundary conditions are described in Table 3.2.10 for duct heat transfer and will be used throughout this section.

It should be noted that because of the symmetry in circular and parallel plate ducts, Nu_{H1} and Nu_{H2} are identical and are referred to simply as Nu_H. Nu_{H4} with wall conduction is a more complicated problem where the energy equations must be solved simultaneously in both the wall and the fluid. Such problems are called conjugated. In the Nu_{H4} situation, the designer has the flexibility of affecting the heat transfer by varying either or both the characteristics of the duct wall or the convective fluid. In the heat-transfer relations to be considered later, care will be taken to identify the proper thermal boundary conditions using the nomenclature introduced in Table 3.2.10.

Laminar Duct Flow for Purely Viscous, Time-Independent, Non-Newtonian Fluids

Modified Power Law Fluids

As discussed in Section 2.9, a constitutive equation for pseudoplastic fluids (flow index, $n < 1$) is the modified power-law equation:

$$\mu_a = \frac{\mu_o}{1 + \dfrac{\mu_o}{K}\left(\dot{\gamma}\right)^{1-n}}. \tag{3.2.140}$$

Equation 3.2.140 describes the flow curve from the zero shear rate Newtonian region through the power-law region, including the low shear rate transition region. It thus has the characteristic that at low shear rates, the apparent viscosity approaches that of a Newtonian fluid while at large shear rates, it approaches that of a power-law fluid. It is therefore valid for fluids that exhibit no return to the Newtonian behavior at high shear rates. But as indicated in Section 2.9, it is also appropriate for fluids that transition back to Newtonian behavior at high shear rates, provided that the limiting Newtonian viscosities differ by approximately four orders of magnitude or more, and that all the shear rates in the flow field are within the zero shear rate Newtonian region through the power-law region. Solutions that use this constitutive equation entail a shear rate parameter, β, which describes

TABLE 3.2.10
Thermal Boundary Conditions for Duct Heat Transfer

1	Constant wall temperature in both the flow and circumferential direction	Nu_T
2	Constant heat flux in the flow direction and constant temperature in the circumferential direction	Nu_{H1}
3	Constant heat flux in the flow and circumferential directions	Nu_{H2}
4	Constant heat flux per unit volume in the wall with circumferential wall heat conduction	Nu_{H4}

whether any particular system is in the Newtonian, transitional, or power-law region. For duct flow, β is given by

$$\beta = \frac{\mu_o}{K}\left(\frac{\bar{u}}{D_H}\right)^{1-n}. \tag{3.2.141}$$

If $\log_{10}\beta > 2$: Power law region

If $\log_{10}\beta < -2$: Newtonian region

If $-2 < \log_{10}\beta \leq 2$: Transition region

For fully developed flow, the characteristic length is the hydraulic diameter, D_H, and the fluid temperature is the "bulk" temperature defined as

$$T_b = \frac{1}{A_c\bar{u}}\int_{A_c} uT\,dA_c. \tag{3.2.142}$$

Figure 3.2.28 illustrates the values of Nu_T vs. β for a circular duct with the flow index, n, as a parameter. It is seen here that the effect of β on Nu_T is only moderate, but for some applications, it may be important to know at what value of β the system is operating. The situation is similar for boundary condition Nu_H.

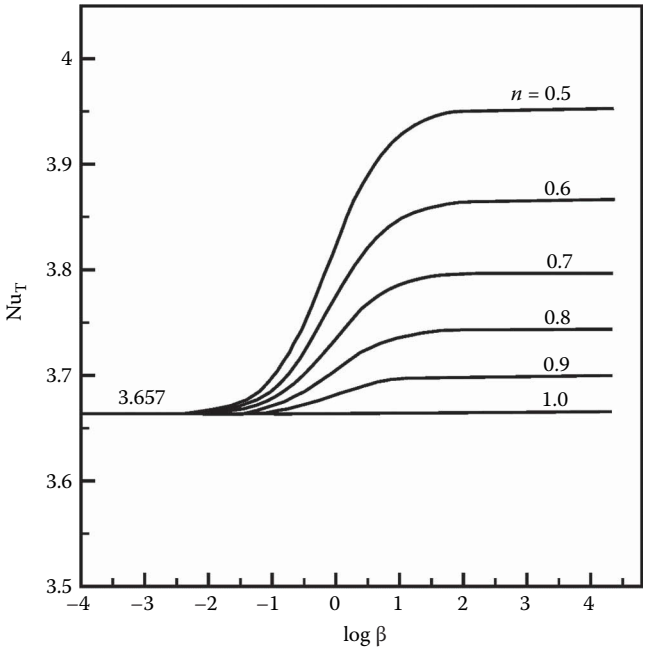


FIGURE 3.2.28 Variation of the fully developed circular duct Nusselt numbers, Nu_T, with the shear rate parameter β and n. (From Irvine, T.F. et al., Graetz problem solutions for a modified power-law fluid, *ASME Symposium on Fundamentals of Forced Convection Heat Transfer*, ASME publ. HTD 101, pp. 123–127, 1988.)

Although Figure 3.2.28 shows the Nusselt number relation graphically, it is convenient to have simple correlation equations to represent the solutions for both boundary conditions. For fully developed Nusselt numbers with values of $0.5 \le n \le 1.0$ and $10^{-4} \le \beta \le 10^4$, Irvine et al. (1988) presented the following equation that represents both solutions with a maximum difference of 1.5%:

$$\mathrm{Nu} = \frac{\mathrm{Nu_N}(1+\beta)}{1+\dfrac{\mathrm{Nu_N}\beta}{\mathrm{Nu_P}}}. \tag{3.2.143}$$

The Newtonian Nusselt numbers are $\mathrm{Nu_N}=3.6568$ for $\mathrm{Nu_T}$, and $\mathrm{Nu_N}=4.3638$ for $\mathrm{Nu_H}$. In addition, Table 3.2.11 lists the power-law Nusselt numbers, $\mathrm{Nu_{TP}}$ and $\mathrm{Nu_{HP}}$, for $\log_{10}\beta=4$, i.e., well into the power law region.

Graetz solutions for the thermal entrance lengths are also available. They assume that the velocity profile is fully developed at the duct entrance and present the duct lengths required for the Nusselt numbers to reach within 1% of the fully developed values. Figure 3.2.29 shows these thermal entrance lengths for $\mathrm{Nu_T}$ thermal boundary condition. The situation is similar for the other boundary condition $\mathrm{Nu_H}$.

A correlation equation for the thermal entrance lengths for both the $\mathrm{Nu_T}$ and $\mathrm{Nu_H}$ boundary conditions by Irvine et al. (1988) represents the numerical solutions within 0.5% for $0.5 \le n \le 1.0$ and $-4 \le \log_{10}\beta \le 4$. Table 3.2.12 lists the power-law thermal entrance lengths, which are needed to evaluate the following correlation equation:

$$x^+_{\mathrm{ent},\beta,n} = \frac{x^+_{\mathrm{ent,N}}(1+\beta)}{1+\dfrac{x^+_{\mathrm{ent,N}}(\beta)}{x^+_{\mathrm{ent,P}}}}, \tag{3.2.144}$$

where $x^+_{\mathrm{ent},\beta,n}$ is the modified power-law dimensionless entrance length defined as $x^+_{\mathrm{ent},\beta,n} = (x_{\mathrm{ent},\beta,n}/D_{\mathrm{H}})/\mathrm{Pe}$, and $x^+_{\mathrm{ent,N}}$ and $x^+_{\mathrm{ent,P}}$ are the corresponding Newtonian and power-law values, respectively. The Newtonian dimensionless entrance lengths are $x^+_{\mathrm{ent,N}} = 0.03347$ for $\mathrm{Nu_T}$ and $x^+_{\mathrm{ent,N}} = 0.04309$ for $\mathrm{Nu_H}$.

Only one noncircular geometry using the modified power-law equation has been studied in the archival literature for laminar fully developed heat transfer (Capobianchi and Irvine, 1992). A correlation equation ($n<1$) for $\mathrm{Nu_{H1}}$ for a concentric annuli with constant heat flux at the inner wall and the outer wall insulated is

$$\mathrm{Nu_{H1}} = \frac{1+\beta}{\dfrac{1}{\mathrm{Nu_{H1,N}}}+\dfrac{\beta}{\mathrm{Nu_{H1,P}}}}. \tag{3.2.145}$$

TABLE 3.2.11
Power Law $\mathrm{Nu_T}$ and $\mathrm{Nu_H}$ Solutions for a Circular Duct ($\log_{10}\beta=4$)

n	$\mathrm{Nu_{TP}}$	$\mathrm{Nu_{HP}}$
1.0 (Newtonian)	3.6568	4.3638
0.9	3.6934	4.4109
0.8	3.7377	4.4679
0.7	3.7921	4.5385
0.6	3.8605	4.6281
0.5	3.9494	4.7456

Source: Irvine, T.F. et al., Graetz problem solutions for a modified power-law fluid, *ASME Symposium on Fundamentals of Forced Convection Heat Transfer*, ASME publ. HTD 101, pp. 123–127, 1988.

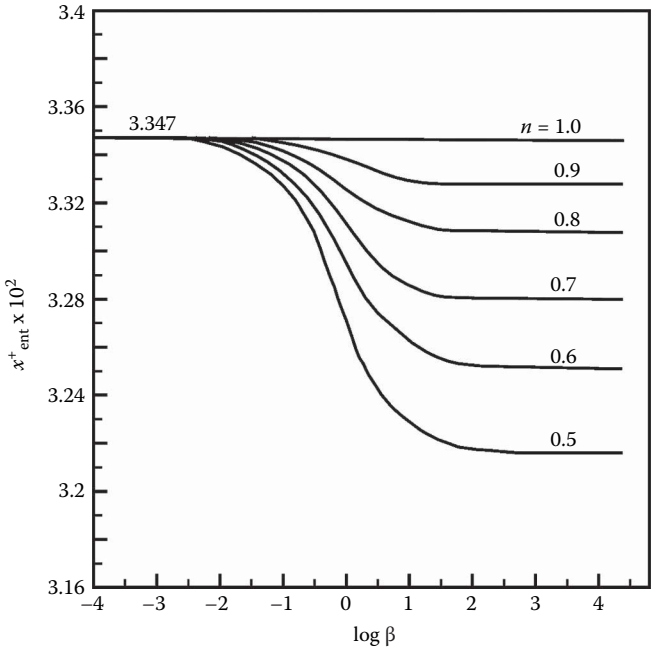


FIGURE 3.2.29 Thermal entrance lengths vs. shear rate parameter β and n for Nu_T in circular ducts. (From Irvine, T.F. et al., Graetz problem solutions for a modified power law fluid, *ASME Symposium on Fundamentals of Forced Convection Heat Transfer*, ASME publ. HTD 101, pp. 123–127, 1988.)

TABLE 3.2.12
Values of Circular Duct Thermal Entrance Lengths for Nu_T and Nu_H for Use in Equation 3.2.144

n	Nu_T, $x^+_{ent,P} \times 10^2$	Nu_H, $x^+_{ent,P} \times 10^2$
1.0 (Newtonian)	3.347	4.309
0.9	3.326	4.281
0.8	3.306	4.248
0.7	3.279	4.210
0.6	3.250	4.166
0.5	3.213	4.114

Source: Irvine, T.F. et al., Graetz problem solutions for a modified power law fluid, *ASME Symposium on Fundamentals of Forced Convection Heat Transfer*, ASME publ. HTD 101, pp. 123–127, 1988.

Extended Modified Power Law Fluids

The extended modified power-law constitutive equations discussed in Section 2.9 describe the flow curve for purely viscous, non-Newtonian fluids with no yield stress that revert back to Newtonian behavior at high shear rates. They span the entire shear rate regime, from zero to the high-shear-rate Newtonian regions, and are valid for any limiting viscosity ratio, $R* = \mu_\infty/\mu_0$. Analyses that use these equations generate a shear rate parameter, β*, whose function is analogous to the parameter β in modified power-law solutions. That is, the value of β* determines the shear rate regime where the system is operating. The definition of β*, its valid range, and the interpretation of its value relative to the shear rate regime is discussed in detail in Section 2.9 and thus no repetition is included here.

Capobianchi and Wagner (2010) obtained numerical solutions for the Nusselt number for fully developed, laminar flows of extended modified power-law fluids in rectangular ducts with n in the range of 0.50–1.50 in increments of 0.10, and $|\log_{10}(R*)| = 1$, 2, 3, and 4. Here, the length scale used in the Nusselt number definition is the hydraulic diameter. They solved both the T and H1 boundary conditions and presented their results graphically for duct aspect ratios, $\alpha*$, of 0.25 and 1.00 (square duct), and in terms of correlation equations with associated constants for $\alpha* = 0.25$, 0.50, 0.75, and 1.00. The latter may be found in Table 3.2.13 and generally overestimate the numerical data by no more than 0.25% for all values of $\alpha*$ when n is restricted to the range ($0.80 \le n \le 1.20$). Outside this range, they overestimate numerical data by more than 20% and are thus not recommended.

Graphical results for the T and H1 boundary conditions for a square duct (i.e., $\alpha* = 1.00$) are shown in Figures 3.2.30 and 3.2.31, respectively. Note that, as in the case of the friction factor—Reynolds number product, the behavior of the Nusselt number follows expected trends; it approaches the Newtonian values at the endpoints of each curve and power-law values at the midpoint of each curve (when plotted in logarithmic coordinates). But power-law values are approached only if, in addition, the limiting viscosities differ sufficiently, by approximately four orders of magnitude or more. Otherwise, the value at the midpoint of each curve is somewhere between the respective power law and Newtonian asymptotes, approaching the latter as $R* \to 1$. This is a result of the effective value of n being closer to 1 than its physical value as the limiting viscosities approach each other. The reader is referred to Section 2.9 for additional details. Finally, it may be noted that, for pseudoplastic fluids, Nusselt numbers are always higher than the Newtonian value and that, referring to Section 2.9, these are associated with pressure drops that are always smaller than Newtonian values. However, the friction factor—Reynolds number product generally departs from Newtonian values far more than the Nusselt number. In fact, the Newtonian Nusselt number was found to underestimate the actual Nusselt number by no more than 7.7% for pseudoplastic fluids, and to overestimate it by no more than 3.3% for dilatant fluids. Thus, in light of the typical uncertainties associated with heat-transfer computations, use of the Newtonian values for calculations involving pseudoplastic or dilatant fluids may be justified depending on the desired solution accuracy.

TABLE 3.2.13
Correlation Equations for the Nusselt Number with Required Constants

Correlation Equations and Required Constants

$\alpha*$	Nu_N	Pseudoplastic fluids: $Nu = \beta^* \times Nu_N \times \left\{ \frac{1+\beta_\infty \times (Nu_N/Nu_g)}{1+\beta_0 \times (Nu_N/Nu_g)} \right\}$					Dilatant fluids: $Nu = \frac{Nu_N}{\beta^*} \times \left\{ \frac{1+\beta_0 \times (Nu_N/Nu_g)}{1+\beta_\infty \times (Nu_N/Nu_g)} \right\}$				
		(Nu_N/Nu_g) for $n =$									
		0.50	0.60	0.70	0.80	0.90	1.10	1.20	1.30	1.40	1.50
Boundary Condition: H1											
1.00	3.60795	0.92345	0.94582	0.96359	0.97802	0.98997	1.00853	1.01587	1.02223	1.02779	1.03269
0.75	3.70052	0.92607	0.94793	0.96518	0.97908	0.99050	1.00801	1.01484	1.02070	1.02577	1.03020
0.50	4.12330	0.93741	0.95697	0.97189	0.98349	0.99266	1.00595	1.01081	1.01483	1.01817	1.02098
0.25	5.33106	0.95777	0.97180	0.98193	0.98950	0.99536	1.00386	1.00706	1.00979	1.01215	1.01421
Boundary Condition: T											
1.00	2.97764	0.92810	0.94884	0.96547	0.97909	0.99042	1.00819	1.01526	1.02143	1.02684	1.03163
0.75	3.05142	0.93054	0.95087	0.96703	0.98015	0.99096	1.00765	1.01418	1.01980	1.02468	1.02894
0.50	3.39243	0.94216	0.96042	0.97430	0.98501	0.99339	1.00526	1.00947	1.01286	1.01562	1.01787
0.25	4.44046	0.97334	0.98420	0.99099	0.99529	0.99809	1.00138	1.00244	1.00330	1.00403	1.00468

Source: Capobianchi, M. and Wagner, D., *Int. J. Heat Mass Transfer*, 53, 558–563, 2010.

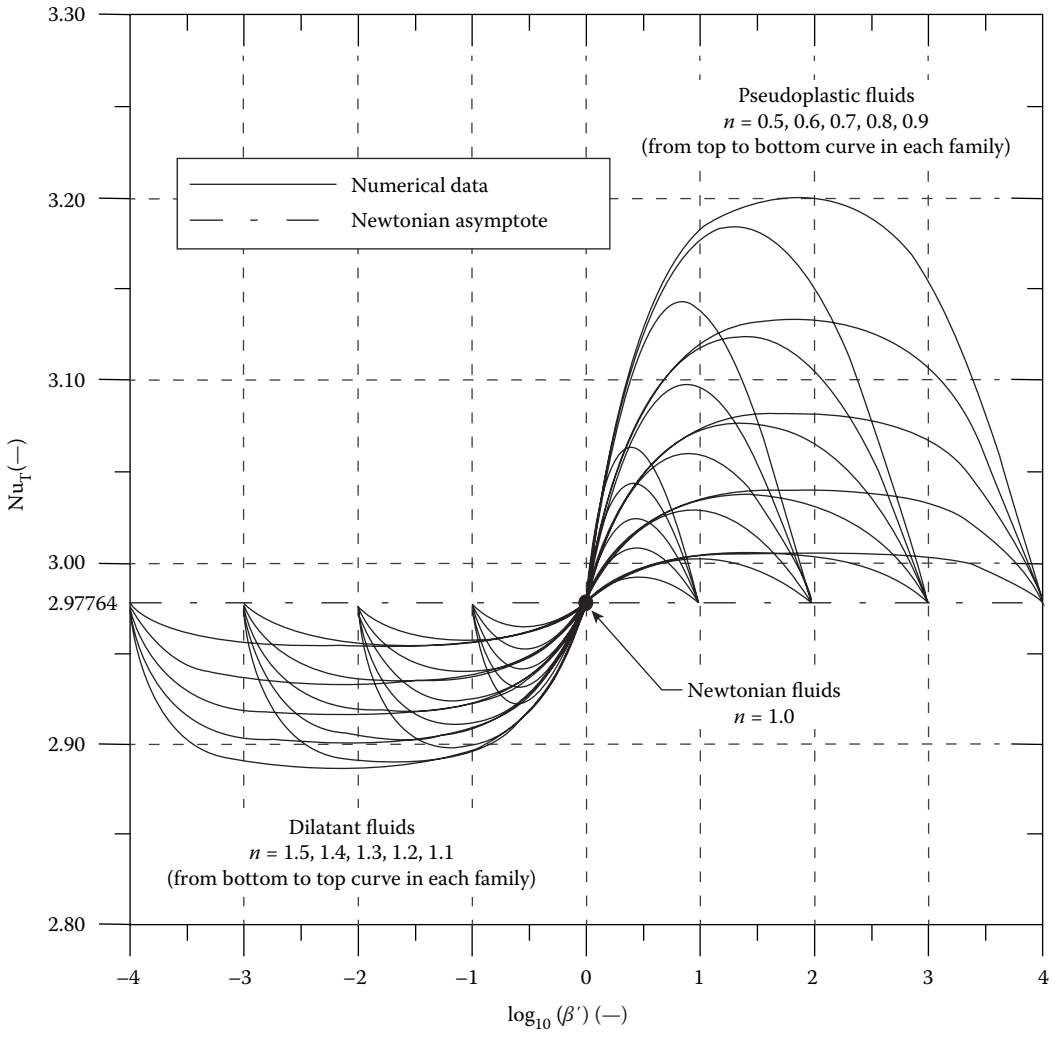


FIGURE 3.2.30 Nusselt number results for a square duct with the T-boundary condition. (Modified from Capobianchi, M. and Wagner, D., *Int. J. Heat Mass Transfer*, 53, 558–563, 2010.)

Power Law Fluids

The mean Nusselt numbers for square ducts and power-law fluids can be found in Chandrupatla and Sastri (1977) and, for isosceles triangular ducts, in Cheng (1984). Thermally developing and thermally developed laminar heat transfer in rectangular channels has been studied by Hartnett and Kostic (1989).

For other cross-sectional shapes, a power-law approximate correlation has been proposed by Cheng (1984):

$$\mathrm{Nu_P} = \mathrm{Nu_N}\left[\frac{(a+bn)}{(a+b)n}\right]^{1/3}, \tag{3.2.146}$$

where a and b are the Kozicki geometric constants listed in Table 2.9.2 in the section on non-Newtonian flows. Equation 3.2.146 applies to any thermal boundary condition. For circular ducts, Equation 3.2.146 predicts the correct solution for both $\mathrm{Nu_T}$ and $\mathrm{Nu_H}$ cases.

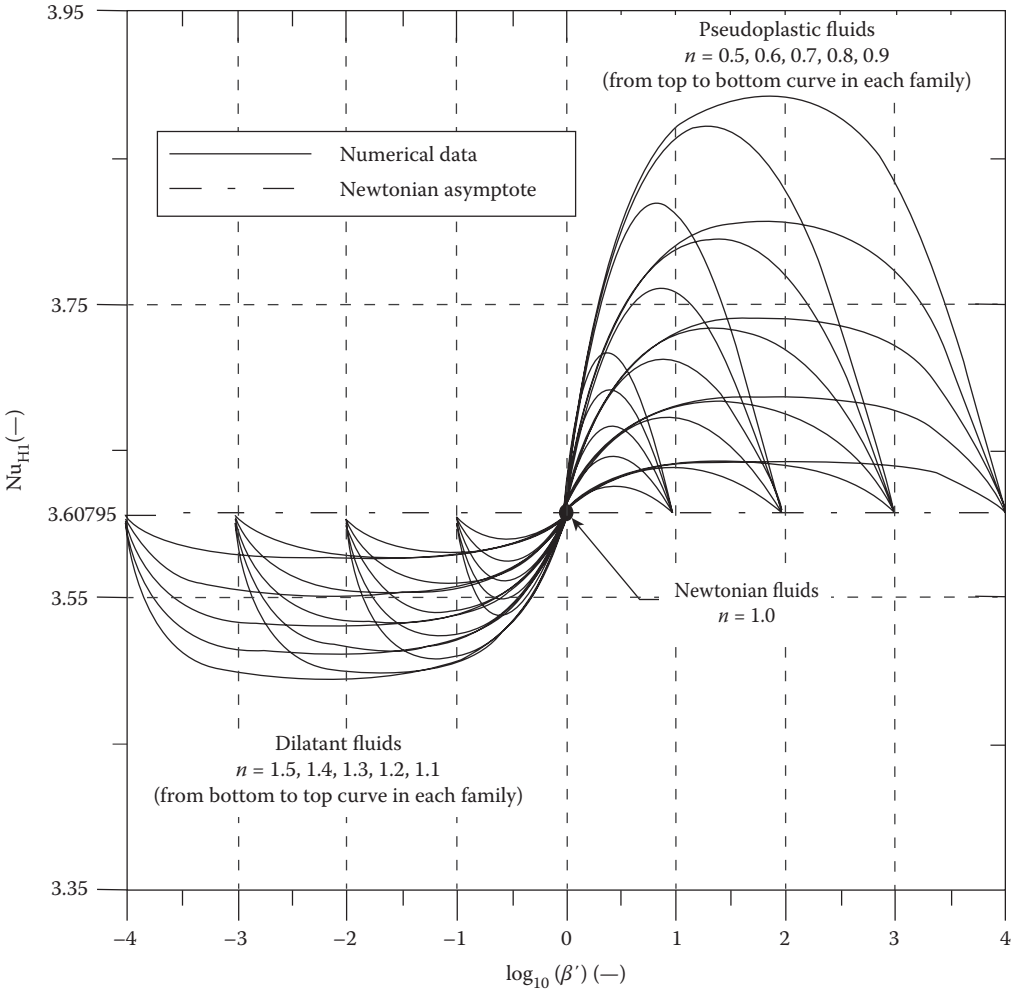


FIGURE 3.2.31 Nusselt number results for a square duct with the H1-boundary condition. (Modified from Capobianchi, M. and Wagner, D., *Int. J. Heat Mass Transfer*, 53, 558–563, 2010.)

Turbulent Duct Flow for Purely Viscous Non-Newtonian Fluids

It is known that in turbulent flow the type of thermal boundary condition has much less effect than that in laminar flow. Therefore, turbulent flow heat-transfer investigations are often reported without specifying the thermal boundary conditions. Yoo (1974) has presented an empirical correlation for turbulent heat transfer in circular ducts for purely viscous power-law fluids.

$$\mathrm{St}\mathrm{Pr}_a^{2/3} = 0.0152\mathrm{Re}_a^{-0.155}. \tag{3.2.147}$$

Equation 3.2.147 describes most of the experimental data available in the literature with a mean deviation of 2.3%. Equation 3.2.147 is recommended in order to predict the turbulent fully developed heat transfer in the ranges $0.2 \le n \le 0.9$ and $3000 \le \mathrm{Re}_a \le 90{,}000$. The Reynolds number and Prandtl numbers in Equation 3.2.147 are based on the apparent viscosity at the wall, μ_a, that is,

$$\mathrm{Re}_a = \frac{\rho \bar{u} D_\mathrm{H}}{\mu_a}, \tag{3.2.148}$$

$$\mathrm{Pr}_a = \frac{\mu_a c_p}{k}. \tag{3.2.149}$$

In order to evaluate Equations 3.2.148 and 3.2.149 in terms of the rheological properties and operating parameters, an expression must be obtained for μ_a in terms of these quantities. The value of μ_a is evaluated by considering that μ_a is determined from fully developed laminar circular tube power-law fluid flow for which it can be shown that (Irvine and Karni, 1987)

$$\mu_a = K\left(\frac{3n+1}{4n}\right)^{n-1}\left(\frac{8\bar{u}}{D_\mathrm{H}}\right)^{n-1}, \tag{3.2.150}$$

assuming that the quantities K, n, c_p, and k are constant. It is also of interest to note here that the Prandtl number is no longer a thermophysical property for power-law fluids but depends upon the average velocity, $\bar{u}$, and the hydraulic diameter, D_H.

Hartnett and Rao (1987) have investigated fully developed turbulent heat transfer for a rectangular duct with a 2:1 aspect ratio and proposed the following equation which generally agreed with their experimental data within ±20%:

$$\mathrm{Nu} = \left(0.0081 + 0.0149n\right)\mathrm{Re}_a^{0.8}\mathrm{Pr}_a^{0.4}. \tag{3.2.151}$$

Viscoelastic Fluids

An important characteristic of viscoelastic fluids is their large hydrodynamic and thermal entrance lengths. Cho and Hartnett (1982) have reported hydrodynamic entrance lengths of up to 100 diameters and thermal entrance lengths up to 200–800 diameters depending upon the Reynolds and Prandtl numbers and the level of fluid elasticity. These can be compared with Newtonian fluids entrance lengths which are of the order of 10–15 diameters. Therefore, care must be used in applying the fully developed relations to practical situations.

Cho et al. (1980) reported heat-transfer measurements in the thermal entrance region and recommend the following empirical equation for saturated aqueous polymer solutions for $6000 \le \mathrm{Re}_a$ and x/D_H values up to 450:

$$J_\mathrm{H} = 0.13\left(x/D_\mathrm{H}\right)^{-0.24}\mathrm{Re}_a^{-0.45}, \tag{3.2.152}$$

where

$$J_\mathrm{H} = \mathrm{St}\,\mathrm{Pr}_a^{2/3} \text{ and } \mathrm{St} = h_c/\rho c_p \bar{u}.$$

All of the reported fully developed turbulent flow heat-transfer measurements have been plagued by solute and solvent, thermal entrance, and degradation effects, and thus there is a considerable scatter in the results. Degradation effects can be reduced or eliminated by using large amounts of polymer (500–10,000 wppm) so that the solution becomes saturated. Cho and Hartnett (1982) attempted to eliminate these effects by using a thermal entrance length of 430 diameters and saturated polymer solutions which should yield maximum heat-transfer reductions. Their experimental results for fully developed heat transfer were correlated for a Reynolds number range $3500 \le \mathrm{Re}_a \le 40{,}000$ and concentration solutions of 500–5000 wppm of polyacrylamide and polyethylene oxide by

$$J_\mathrm{H} = 0.03\mathrm{Re}_a^{-0.45}. \tag{3.2.153}$$

For viscoelastic fluids in fully developed (hydrodynamically and thermally) *laminar flow in circular ducts*, there is no apparent viscoelastic effect. Thus, the heat-transfer relations are the same as

those for time-independent fluids such as power-law or modified power-law fluids. The same situation holds for thermal entrance region heat transfer (Graetz problem). Relations for the laminar Nusselt numbers in thermal entrance regions are presented by Cho and Hartnett (1982).

Free Convection Flows and Heat Transfer

Free convection information available in the heat-transfer literature is concentrated on heat transfer to power-law fluids for vertical plates, spheres and cylinders, and in parallel plate channels. For free convection flows, however, the velocities and thus the shear rates are low and care must be taken that the flow for a particular fluid is in the power-law shear rate region before using power-law solutions or correlations. Comprehensive review articles on free convection with non-Newtonian fluids in external flows have been presented by Shenoy and Mashelkar (1982), Irvine and Karni (1987) and Chhabra (2006).

Internal Flows

Vertical Parallel Plates

Irvine et al. (1982) provided a numerical solution for buoyancy-induced, laminar flows of power-law fluids in vertical, parallel plate channels. Their results are presented graphically in Figure 3.2.32. Of interest are the average Nusselt number $\overline{\mathrm{Nu}}_{\mathrm{T}b}$ and the dimensionless average flow velocity between the plates, U_{o}^{+}. These are shown on the left and right ordinates, respectively, in Figure 3.2.32. The characteristic dimension in the Nusselt and Grashof numbers is the plate spacing, b. The dimensionless quantities used in Figure 3.2.32 are defined as follows:

$$\overline{\mathrm{Nu}}_{\mathrm{T}b} = \frac{\overline{h}_{\mathrm{c}} b}{k} \qquad U_{\mathrm{o}}^{+} = \frac{b u_{\mathrm{o}}}{L u^{*}} \tag{3.2.154}$$

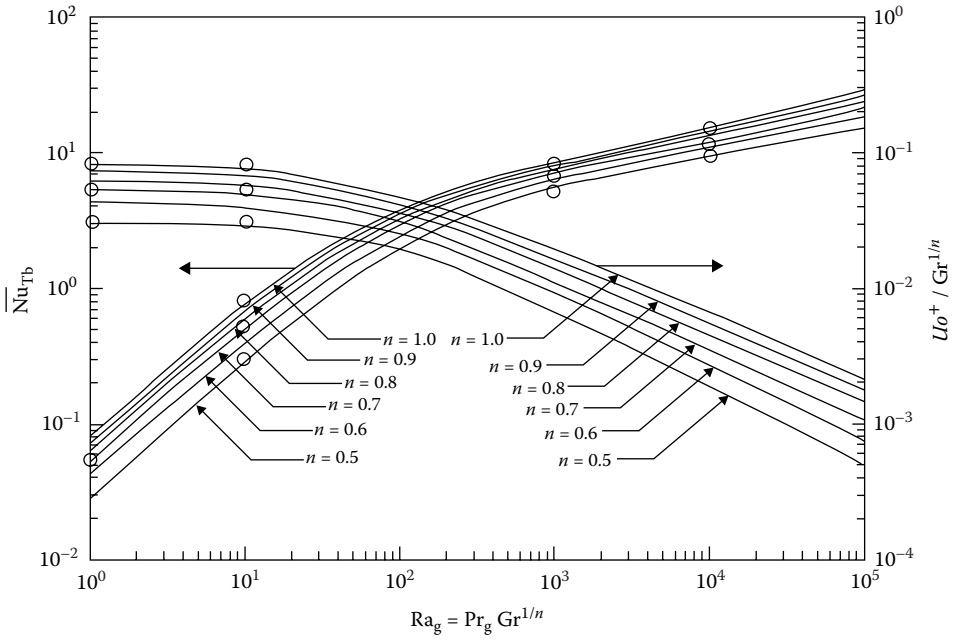


FIGURE 3.2.32 Free convection average Nusselt number, $\overline{\mathrm{Nu}}_{\mathrm{T}b}$, and dimensionless average velocity, U_{o}^{+}, between vertical plates for a power-law fluid vs. generalized Rayleigh number for the T boundary condition. (From Irvine, T.F. et al., Vertical channel free convection to a power-law fluid, ASME Paper 82-WA/HT-69, 1982.)

$$\mathrm{Pr_g} = \frac{\rho c_p}{k}\left[\frac{v_k^{1/(2-n)}}{\left(\frac{L}{b}\right)^{(l-n)/(2-n)} b^{(2n-2)/(2-n)}}\right];\ v_K = \frac{K}{\rho} \tag{3.2.155}$$

$$\mathrm{Gr_g} = \frac{g\alpha\left(T_s - T_\infty\right)b^{(n+2)/(2-n)}}{v_K^{2/(2-n)}\left(\frac{L}{b}\right)^{n/(2-n)}};\qquad u^* = \frac{v_K^{1/(2-n)} b^{(1-2n)/(2-n)}}{L^{(1-n)/(2-n)}} \tag{3.2.156}$$

For vertical parallel plates for the average Nusselt number, $\overline{\mathrm{Nu}}_{\mathrm{H}b}$ and the between plate average velocity, Schneider and Irvine (1984) have presented graphical results similar to Figure 3.2.32.

Lee et al. (1998) have presented a numerical solution for the laminar flow of a *modified power-law fluid* between vertical plates. They have also calculated thermal entrance regions and shown that if a parallel plate system is actually operating in the transition region and if the power-law solution is used, both the total heat transfer and the velocity between plates can differ by over an order of magnitude. It is important to consider the shear rate parameter in order to determine which free convection solution to use.

External Flows

Vertical Plates

For a vertical plate with a modified power-law fluid and a thermal boundary condition $\bar{N}u_{\mathrm{T}}$, in laminar flow, the following relation is recommended by Shenoy and Mashelkar (1982):

$$\bar{\mathrm{N}}\mathrm{u}_{TL} = T(n)\mathrm{Gr}_{TL}^{1/(2n+2)}\,\mathrm{Pr}_{TL}^{n/(3n+1)}, \tag{3.2.157}$$

where $\bar{\mathrm{N}}\mathrm{u}_{TL}$ is the average Nusselt number and

$$\mathrm{Gr}_{TL} = \frac{\rho^2 L^{n+2}}{K^2}\left[g\alpha\left(T_s - T_\infty\right)\right]^{2-n}, \tag{3.2.158}$$

$$\mathrm{Pr}_{TL} = \frac{\rho c_p}{k}\left(\frac{K}{\rho}\right)^{2/(n+1)} L^{(n-1)/(2n+2)}\left[g\alpha\left(T_s - T_\infty\right)\right]^{(3n-3)/(2n+2)}, \tag{3.2.159}$$

where α is the isobaric thermal expansion coefficient.

In the range $0.5 \le n \le 1$, $T(n)$ can be approximated by

$$T\left(n\right) = 0.1636n + \ 0.5139. \tag{3.2.160}$$

The characteristic dimension in the Nusselt and Grashof numbers is the plate height, L.

For thermal boundary conditions $\mathrm{Nu_H}$, the following relation is also recommended by Shenoy and Mashelkar (1982). Since the heat flux, q_w, is specified in this case, the local plate temperature at any x (measured from the bottom of the plate) can be obtained from the local Nusselt number $\mathrm{Nu}_{\mathrm{H}x}$. The heat-transfer coefficient is defined in terms of the difference between the wall and free-stream temperatures.

$$\mathrm{Nu}_{\mathrm{H}x} = 0.619\left[\mathrm{Gr}_{\mathrm{H}x}^{(3n+2)/(n+4)}\mathrm{Pr}_{\mathrm{H}x}^{n}\right]^{0.213}, \tag{3.2.161}$$

where

$$\mathrm{Gr}_{\mathrm{H}x} = \frac{\rho^2 x^4}{k^2}\left(\frac{g\alpha q_w}{k}\right)^{2-n} \tag{3.2.162}$$

$$\mathrm{Pr}_{\mathrm{H}x} = \frac{\rho c_p}{K}\left(\frac{K}{\rho}\right)^{5/(n+4)} x^{(2n-2)/(n+4)}\left(\frac{g\alpha q_w}{k}\right)^{(3n-3)/(n+4)} \tag{3.2.163}$$

Recently, Capobianchi and Aziz (2012, 2014) have solved numerically for the average Nusselt number problem for laminar, natural convection between a vertical surface and pseudoplastic and dilatant fluids for both the constant temperature ($\overline{\mathrm{Nu}}_{TL}$) and the constant heat flux ($\overline{\mathrm{Nu}}_{HL}$) boundary conditions. They used the extended modified power-law equation for the apparent viscosity in their formulation and provided solutions in both graphical and tabular form. Representative plots of the results are shown in Figures 3.2.33 through 3.2.36, and the tabular results are provided in Tables 3.2.14 through 3.2.17. For both boundary conditions, the Nusselt number may be seen to be a function of n, the Boussinesq number, Bo, and the Prandtl number, Pr*, defined as

$$\mathrm{Bo} = \frac{g\alpha L^3 \Delta T^*}{\alpha_{\mathrm{T}}^2} : \mathrm{Pr} = \frac{\mu^*/\rho}{\alpha_{\mathrm{T}}}. \tag{3.2.164}$$

TABLE 3.2.14
Representative Numerical Values of $\overline{\mathrm{Nu}}_{TL}$ for Pseudoplastic Fluids

			$\overline{\mathrm{Nu}}_{TL}$									
			$\log_{10}(R^*) = -4.00$					$\log_{10}(R^*) = -2.00$				
$\log_{10}$ (Bo)	$\log_{10}$ (Pr_0)	n	$\log_{10}(\beta^*)$									
			0.800	1.600	2.000	2.400	3.200	0.400	0.800	1.000	1.200	1.600
4.00	0.00	0.50	7.3910	7.8275	7.9010	7.9360	7.9580	6.7381	7.2738	7.4021	7.4825	7.5646
		0.70	7.0693	7.6795	7.8114	7.8866	7.9482	6.4089	7.0258	7.2195	7.3587	7.5231
		0.90	6.7617	7.4868	7.6774	7.7989	7.9207	6.1482	6.7527	6.9834	7.1717	7.4403
	−1.00	0.50	7.8664	7.9647	7.9796	7.9864	7.9904	7.6694	7.8061	7.8330	7.8489	7.8644
		0.70	7.7399	7.9184	7.9535	7.9729	7.9879	7.4942	7.7148	7.7722	7.8105	7.8526
		0.90	7.5822	7.8427	7.9042	7.9421	7.9791	7.3151	7.5768	7.6643	7.7315	7.8214
6.00	2.00	0.50	13.499	19.396	21.465	22.787	23.827	10.029	12.808	13.919	14.840	16.150
		0.70	11.609	17.070	19.535	21.455	23.487	8.9926	11.432	12.616	13.737	15.630
		0.90	10.409	15.232	17.702	19.882	22.801	8.3371	10.380	11.484	12.625	14.905
	1.00	0.50	20.279	23.612	24.294	24.634	24.852	16.575	19.414	20.230	20.791	21.421
		0.70	18.100	22.313	23.463	24.162	24.756	14.879	17.836	18.979	19.892	21.097
		0.90	16.387	20.820	22.321	23.368	24.493	13.720	16.342	17.562	18.677	20.492
8.00	4.00	0.50	15.553	26.619	33.302	40.015	49.683	10.835	14.413	15.913	17.215	19.243
		0.70	12.705	21.038	26.512	32.833	45.786	9.4785	12.454	13.992	15.524	18.367
		0.90	11.099	17.716	22.187	27.572	40.781	8.6775	11.062	12.426	13.909	17.188
	3.00	0.50	29.076	46.333	54.465	60.780	67.041	20.319	26.329	28.610	30.473	33.155
		0.70	23.079	36.829	44.815	52.792	64.479	17.239	22.494	25.057	27.505	31.749
		0.90	19.663	30.669	37.570	45.178	59.839	15.421	19.582	21.905	24.382	29.653

Source: Capobianchi, M. and Aziz, A., *ASME J. Heat Transfer*, 134(12), 122502, 2012.

TABLE 3.2.15
Representative Numerical Values of $\overline{Nu}_{TL}$ for Dilatant Fluids

			$\overline{Nu}_{TL}$									
			$\log_{10}(R^*)=4.00$					$\log_{10}(R^*)=2.00$				
$\log_{10}$ (Bo)	$\log_{10}$ (Pr_0)	n	$\log_{10}(\beta^*)$									
			−3.200	−2.400	−2.000	−1.600	−0.800	−1.600	−1.200	−1.000	−0.800	−0.400
8.00	0.00	1.50	29.516	19.679	16.068	13.131	8.9261	36.904	30.229	27.690	25.549	22.392
		1.30	32.393	21.488	17.452	14.162	9.4143	40.153	32.817	29.880	27.321	23.272
		1.10	36.124	23.987	19.377	15.601	10.094	43.909	36.257	32.908	29.856	24.644
	−1.00	1.50	44.575	31.834	26.459	21.877	15.178	52.811	45.913	43.107	40.649	36.905
		1.30	48.987	35.032	29.024	23.862	16.155	57.090	49.686	46.393	43.353	38.249
		1.10	54.234	39.334	32.553	26.631	17.552	61.511	54.416	50.799	47.209	40.455
10.00	2.00	1.50	30.872	19.986	16.218	13.205	8.9471	40.016	31.701	28.758	26.356	22.936
		1.30	34.263	21.936	17.674	14.273	9.4445	44.228	34.777	31.292	28.364	23.906
		1.10	38.901	24.701	19.737	15.783	10.142	49.463	39.072	34.935	31.326	25.446
	1.00	1.50	51.790	33.669	27.365	22.328	15.310	67.036	53.569	48.863	45.084	39.908
		1.30	58.283	37.530	30.290	24.504	16.335	74.902	59.344	53.611	48.816	41.641
		1.10	67.190	43.068	34.498	27.634	17.820	84.685	67.527	60.599	54.523	44.592
12.00	4.00	1.50	30.959	20.005	16.227	13.210	8.9484	40.241	31.800	28.830	26.412	22.977
		1.30	34.411	21.969	17.690	14.281	9.4468	44.588	34.936	31.405	28.448	23.960
		1.10	39.178	24.767	19.770	15.799	10.146	50.064	39.355	35.134	31.469	25.524
	3.00	1.50	52.313	33.783	27.420	22.356	15.318	68.359	54.176	49.312	45.435	40.174
		1.30	59.141	37.725	30.386	24.552	16.348	76.936	60.272	54.278	49.315	41.965
		1.10	68.733	43.444	34.686	27.728	17.845	87.965	69.102	61.717	55.328	45.034

Source: Capobianchi, M. and Aziz, A., *ASME J. Heat Transfer*, 134(12), 122502, 2012.

Here, α_T is the thermal diffusivity and μ^* is the extended modified power-law reference viscosity. The definition of the temperature difference, ΔT^*, depends on the boundary condition and equals $(T_s - T_\infty)$ for the constant temperature boundary condition, and is given by qwL/k for the constant heat flux case. Furthermore, Pr^* may be expressed in terms of R^*, β^*, and the Prandtl number based on μ_0, Pr_0. For pseudoplastic and dilatant fluids, respectively, Pr^* becomes

$$Pr^* = \frac{Pr_0}{\beta}, \tag{3.2.165}$$

$$Pr^* = \beta^* R^* Pr_0. \tag{3.2.166}$$

It should be noted that the bulk velocity, V, in the definitions of β and β_∞ (i.e., within the definition of β^*, see Section 2.9) is replaced by a reference velocity, u^*, that depends on the boundary condition. For the constant temperature and constant heat flux boundary conditions, u^* is given, respectively, by

$$u^* = \frac{\alpha_T / L}{\left(\frac{1+Pr^*}{Bo}\right)^{1/2}}, \tag{3.2.167}$$

TABLE 3.2.16
Representative Numerical Values of $\overline{\mathrm{Nu}}_{HL}$ for Pseudoplastic Fluids

			$\overline{\mathrm{Nu}}_{HL}$									
			$\log_{10}(R^*) = -4.00$					$\log_{10}(R^*) = -2.00$				
$\log_{10}$ (Bo)	$\log_{10}$ (Pr_0)	n	$\log_{10}(\beta^*)$									
			0.800	1.600	2.000	2.400	3.200	0.400	0.800	1.0000	1.200	1.600
4.00	0.00	0.50	5.4768	5.8403	5.9067	5.9398	5.9616	5.0185	5.4187	5.5293	5.6028	5.6817
		0.70	5.2942	5.7425	5.8445	5.9038	5.9538	4.8525	5.2724	5.4144	5.5207	5.6520
		0.90	5.1279	5.6268	5.7613	5.8478	5.9353	4.7210	5.1232	5.2806	5.4109	5.6004
	−1.00	0.50	5.8731	5.9605	5.9745	5.9810	5.9850	5.7195	5.8389	5.8647	5.8804	5.8961
		0.70	5.7898	5.9275	5.9552	5.9706	5.9831	5.6129	5.7758	5.8209	5.8519	5.8869
		0.90	5.6950	5.8790	5.9230	5.9502	5.9769	5.5117	5.6919	5.7536	5.8014	5.8662
6.00	2.00	0.50	8.1504	11.183	12.445	13.357	14.180	6.5389	7.9517	8.5741	9.1304	10.008
		0.70	7.4979	10.300	11.624	12.717	13.979	6.1621	7.4374	8.0684	8.6801	9.7690
		0.90	7.0384	9.5925	10.886	12.044	13.646	5.9015	7.0270	7.6250	8.2387	9.4646
	1.00	0.50	11.713	13.828	14.333	14.604	14.792	9.8746	11.435	11.974	12.377	12.869
		0.70	10.871	13.199	13.894	14.335	14.731	9.2690	10.780	11.410	11.941	12.692
		0.90	10.208	12.548	13.369	13.952	14.593	8.8316	10.191	10.830	11.423	12.412
8.00	4.00	0.50	8.8641	13.424	16.159	19.068	24.091	6.8578	8.5764	9.3602	10.084	11.318
		0.70	7.9628	11.808	14.204	16.914	22.637	6.3883	7.8827	8.6589	9.4429	10.956
		0.90	7.3739	10.692	12.800	15.242	20.985	6.0771	7.3597	8.0744	8.8395	10.503
	3.00	0.50	14.456	21.159	24.661	27.834	31.753	11.198	13.853	14.986	15.989	17.584
		0.70	12.763	18.508	21.790	25.150	30.615	10.246	12.600	13.772	14.921	17.022
		0.90	11.640	16.604	19.570	22.769	28.958	9.6168	11.613	12.703	13.852	16.273

Source: Capobianchi, M. and Aziz, A., *ASME J. Heat Transfer*, 136(9), 092501, 2014.

$$u^* = \frac{\alpha_T / L}{\left(\dfrac{1+\mathrm{Pr}^*}{\mathrm{Bo}}\right)^{2/5}}. \tag{3.2.168}$$

Consequently, since there is no imposed velocity in natural convection problems and, instead, a reference velocity is used that depends on both Pr* and Bo, a practical complication exists in evaluating β^* for a given situation. Namely, β^* is only known implicitly and must be evaluated iteratively. The reader is referred to Capobianchi and Aziz (2014) for a description of the recommended iterative technique.

Referring to Figures 3.2.33 through 3.2.36, the results may be seen to group as a function of Bo. The heavy lines in each group are the Newtonian values, the remaining curves being the non-Newtonian solutions. It may be noted that Nusselt numbers for pseudoplastic fluids always exceed their Newtonian counterparts, and that the opposite is true for dilatant fluids. As in the case of the other studies that use the extended modified power-law equations, the endpoints of each curve approach the Newtonian values, while the midpoint (when plotted in logarithmic coordinates) approach the power-law values. And as before, the latter is true only if, in addition, the limiting viscosities, μ_0 and μ_∞, differ by about four orders of magnitude or more. Thus, the behavior in all such studies is analogous.

TABLE 3.2.17
Representative Numerical Values of $\overline{Nu}_{HL}$ for Dilatant Fluids

			$\overline{Nu}_{HL}$									
			$\log_{10}(R^*) = 4.00$					$\log_{10}(R^*) = 2.00$				
$\log_{10}$ (Bo)	$\log_{10}$ (Pr_0)	n	$\log_{10}(\beta^*)$									
			−3.333	−2.667	−2.000	−1.333	−0.667	−1.667	−1.333	−1.000	−0.667	−0.333
8.00	0.00	1.50	18.024	13.714	10.432	7.9377	6.1213	20.966	18.204	16.074	14.376	13.101
		1.30	19.153	14.564	11.018	8.3118	6.3153	22.110	19.255	16.924	14.978	13.412
		1.10	20.490	15.649	11.781	8.8007	6.5705	23.311	20.520	18.017	15.792	13.862
	−1.00	1.50	24.839	19.957	15.619	12.058	9.4228	27.788	25.142	22.947	21.085	19.635
		1.30	26.404	21.289	16.615	12.731	9.7822	29.199	26.555	24.151	21.957	20.086
		1.10	28.118	22.932	17.894	13.613	10.265	30.547	28.156	25.659	23.147	20.767
10.00	2.00	1.50	18.874	13.979	10.516	7.9658	6.1319	22.588	19.085	16.591	14.707	13.342
		1.30	20.227	14.915	11.135	8.3509	6.3297	24.074	20.352	17.571	15.379	13.683
		1.10	21.905	16.144	11.951	8.8594	6.5912	25.724	21.944	18.868	16.308	14.184
	1.00	1.50	28.691	21.325	16.078	12.214	9.4830	34.273	29.093	25.435	22.721	20.827
		1.30	31.031	22.995	17.212	12.938	9.8593	36.797	31.260	27.115	23.864	21.391
		1.10	33.922	25.192	18.720	13.906	10.370	39.579	33.990	29.371	25.491	22.265
12.00	4.00	1.50	18.933	13.995	10.522	7.9675	6.1325	22.718	19.149	16.627	14.730	13.360
		1.30	20.317	14.942	11.143	8.3538	6.3307	24.259	20.445	17.624	15.412	13.707
		1.10	22.050	16.190	11.967	8.8647	6.5931	25.990	22.090	18.953	16.358	14.215
	3.00	1.50	29.010	21.418	16.107	12.223	9.4869	34.945	29.436	25.635	22.854	20.932
		1.30	31.498	23.138	17.258	12.953	9.8652	37.725	31.745	27.395	24.040	21.517
		1.10	34.641	25.430	18.800	13.933	10.380	40.874	34.717	29.794	25.745	22.424

Source: Capobianchi, M. and Aziz, A., *ASME J. Heat Transfer*, 136(9), 092501, 2014.

Sphere and Horizontal Cylinder

For flow over a sphere, the correlation for power-law fluids by Amato and Tien (1976) is

$$\overline{Nu}_{Tr} = c_1 Z^{c_2}, \tag{3.2.169}$$

where

$$Z = Gr_{Tr}^{1/(2n+2)} Pr_{Tr}^{n/(3n+1)} \tag{3.2.170}$$

and

$$c_1 = 0.996 \pm 0.120, c_2 = 0.682 \quad \text{for } Z < 10$$

$$c_1 = 0.489 \pm 0.005, c_2 = 1.10 \qquad \text{for } 10 \le Z \le 40,$$

where the characteristic dimension in all dimensionless variables is the sphere radius, r, and Gr_{Tr} and Pr_{Tr} are defined in Equations 3.2.158 and 3.2.159, respectively.

For pseudoplastic fluids flowing over a cylinder, an experimental correlation proposed by Gentry and Wollersheim (1974) for the average Nusselt number, $\overline{Nu}_{TD}$, is

$$\overline{Nu}_{TD} = \frac{\overline{h}_c D}{k} = 1.19\left(Gr_{TD} Pr_{TD}\right)^{0.2}, \tag{3.2.171}$$

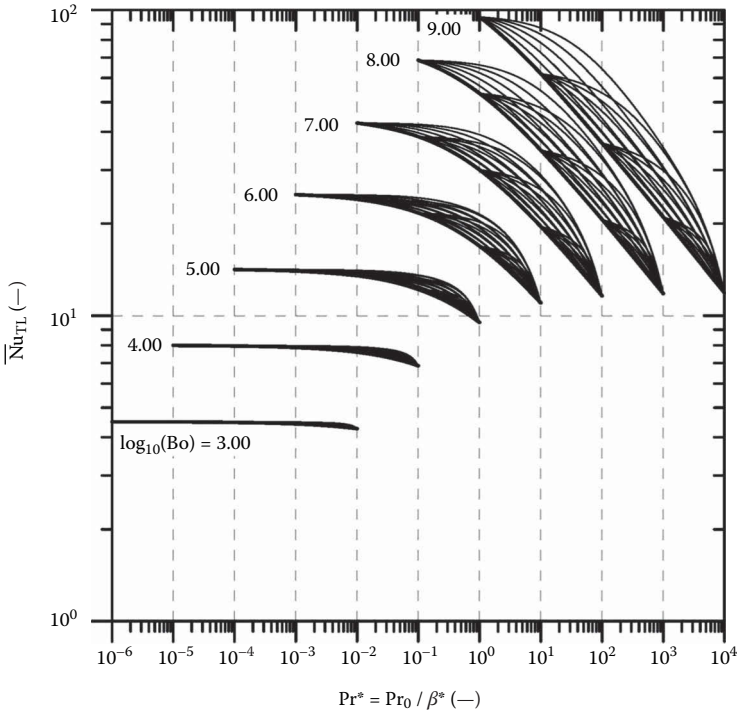


FIGURE 3.2.33 $\overline{\mathrm{Nu}}_{TL}$ for pseudoplastic fluids. The curves in heavy-weight lines are Newtonian fluid (n=1.00) solutions. Curves in normal-weight lines are pseudoplastic fluid solutions for n=0.50, 0.60, 0.70, 0.80, and 0.90, proceeding from the upper to the lower curve in each family. (Modified from Capobianchi, M. and Aziz, A., *ASME J. Heat Transfer*, 134(12), 122502, 2012.)

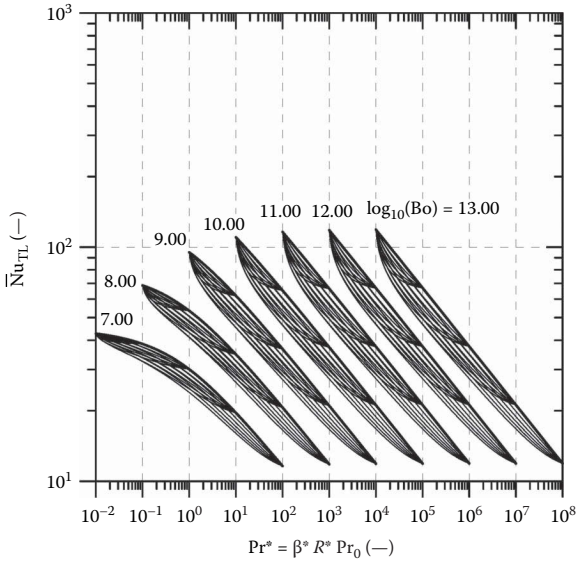


FIGURE 3.2.34 $\overline{\mathrm{Nu}}_{TL}$ for dilatant fluids. The curves in heavy-weight lines are Newtonian fluid (n=1.00) solutions. Curves in normal-weight lines are dilatant fluid solutions for n=1.50, 1.40, 1.30, 1.20, and 1.10, proceeding from the lower to the upper curve in each family. (Modified from Capobianchi, M. and Aziz, A., *ASME J. Heat Transfer*, 134(12), 122502, 2012.)

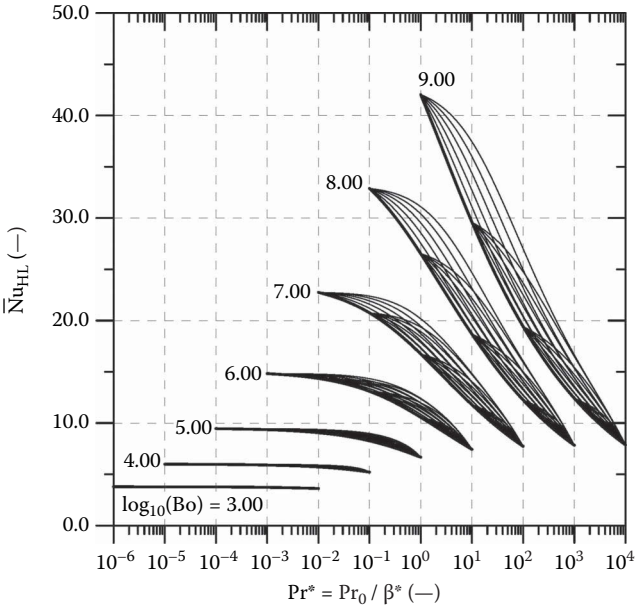


FIGURE 3.2.35 $\overline{Nu}_{HL}$ for pseudoplastic fluids. The curves in heavy-weight lines are Newtonian fluid (n=1.00) solutions. Curves in normal-weight lines are pseudoplastic fluid solutions for n=0.50, 0.60, 0.70, 0.80, and 0.90, proceeding from the upper to the lower curve in each family. (Modified from Capobianchi, M. and Aziz, A., *ASME J. Heat Transfer*, 136(9), 092501, 2014.)

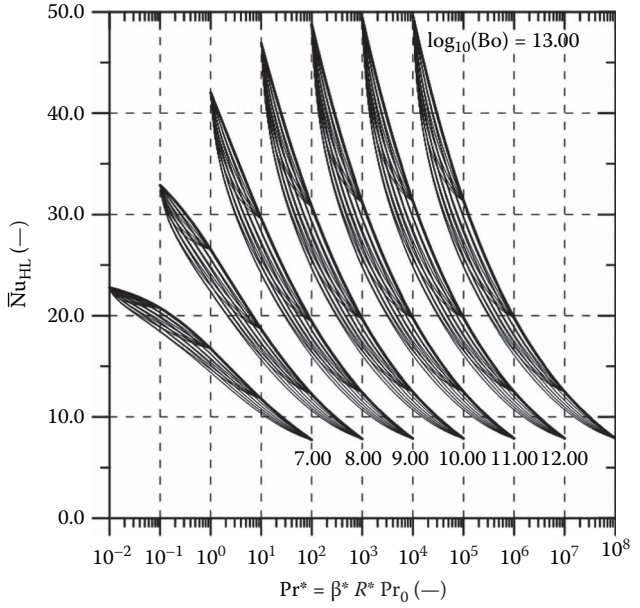


FIGURE 3.2.36 $\overline{Nu}_{HL}$ for dilatant fluids. The curves in heavy-weight lines are Newtonian fluid (n=1.00) solutions. Curves in normal-weight lines are dilatant fluid solutions for n=1.50, 1.40, 1.30, 1.20, and 1.10, proceeding from the lower to the upper curve in each family. (Modified from Capobianchi, M. and Aziz, A., *ASME J. Heat Transfer*, 136(9), 092501, 2014.)

where Gr_{TD} and Pr_{TD} are defined in Equations 3.2.158 and 3.2.159 with the cylinder diameter, D, being used instead of L.

Suffice it to add here that the recent numerical studies of Prhashanna and Chhabra (2010, 2011) for a sphere and cylinder (Table 3.2.18) lend further support to the reliability of Equations 3.2.169 and 3.2.171. The scant results for a range of other shapes including hemispheres, spheroids, and cylinders of various cross sections are also listed in Table 3.2.18 along with the applicable ranges of conditions. Broadly, all else being equal, shear-thinning fluid behavior ($n < 1$) promotes heat transfer over and above that seen in Newtonian fluids. Indeed, it is possible to realize up to 30%–40% augmentation in heat transfer under appropriate conditions. On the other hand, heat-transfer deteriorates in shear-thickening fluids ($n > 1$) by about 15%–20%.

Similarly, the corresponding studies on free convection in Bingham plastic fluids have also been reported for a sphere and a horizontal cylinder, elliptical and semicircular cylinder (Table 3.2.19). The Bingham plastic and other viscoplastic fluids differ from the power law and other generalized Newtonian fluids in a significant manner, which directly impinges on convective heat transport. In a given situation, a Bingham fluid may or may not yield depending upon the magnitude of the prevailing stress *vis-à-vis* the fluid yield stress. Consequently, the flow domain is spanned by the coexistence of solid-like (unyielded) and fluid-like (yielded) regions. Naturally, heat transfer occurs by conduction in solid-like regions and convection is limited only to the fluid-like regions. In steady-state situations, the overall rate of heat transfer is thus limited by conduction. Next, it stands to reason that for a given Bingham number (nondimensionalized yield stress), there exists a critical Rayleigh number denoting the onset of buoyancy-induced flow in these systems. Thus, for a given geometry and boundary conditions, the average Nusselt number decreases, with the increasing Bingham umber, from its maximum value (at $Bn = 0$, i.e., Newtonian fluid) to its minimum value (corresponding to conduction) at high Bingham numbers (Tables 3.2.20 and 3.2.21).

Forced Convection Heat Transfer

Early attempts in this field are based on the standard boundary layer analysis, which is generally valid in the limit of thin boundary layers wherein the curvature effects can also be neglected. This body of knowledge has been summarized by Chhabra (2006). However, over the past 15–20 years, the full momentum and energy equations have been solved, mainly for power law and Bingham plastic fluids, for an isothermal sphere, spheroid, and hemisphere as well as for two-dimensional cylinders of various cross sections. A summary of the predictive correlations based on the numerical results is presented in Table 3.2.22 for power law, values of constants used in Table 3.2.22 is described in Table 3.2.23, and in Table 3.2.24 for Bingham plastic fluids for a range of shapes. In overall terms, shear-thinning viscosity ($n < 1$) promotes heat transfer whereas shear-thickening fluid ($n > 1$) behavior impedes it with reference to that in Newtonian fluids otherwise under identical conditions. Similar enhancement is also possible in Bingham plastic fluids due to the sharpening of the temperature gradients in the thin fluid-like layer adjacent to the heated object (Tables 3.2.25 and 3.2.26).

From the foregoing discussion, it is abundantly clear that it is possible to modulate the rate of heat transfer by tuning the non-Newtonian characteristics (shear-thinning viscosity, yield stress) of the liquid medium via the use of suitable additives like clays, polymers, etc., in small doses. This approach is routinely used in food-, personal care-, and pharma-sectors to achieve improved functionality of a given product. Similarly, it is also possible to achieve some economy in energy consumption in chemical and allied processing industries on account of shear-thinning behavior.

TABLE 3.2.18
Natural Convection Correlations for External Flow over Isothermal Surfaces for Power Law Fluids

Correlation	Range
Sphere (Prhashanna and Chhabra, 2010) $Nu_{Tr}=Z$ for $Z>10$ $Nu_{Tr}=2\,Z^{0.72}$ for $Z\le 10$, where $Z=Gr_{Tr}^{\frac{1}{2(n+1)}}Pr_{Tr}^{\frac{1}{3n+1}}$	$0.72\le Pr_{Tr}\le 100$ $0.4\le n\le 1.8$ $10\le Gr_{Tr}\le 10^7$
Hemisphere (Sasmal and Chhabra, 2014) For $Z\le 10$ Inverted: $Nu_{TD}=2.255+0.332Z^{1.20}$ Upright: $Nu_{TD}=2.255+0.343Z^{1.18}$ For $Z<10$ Inverted: $Nu_{TD}=1.102Z^{0.815}$ Upright: $Nu_{TD}=1.14Z^{0.80}$	$0.72\le Pr_{TD}\le 100$ $0.2\le n\le 1.5$ $10\le Gr_{TD}\le 10^5$
Spheroid (Gupta et al., 2014) $Nu_{Tb}=\left(Nu_{Tb}\right)_\infty+c_1Z^{d_1}$, where $Z=Gr_{Tb}^{\frac{1}{2(n+1)}}Pr_{Tb}^{\frac{1}{3n+1}}$ For constants c_1 and d_1 refer Table 3.2.20	$0.2\le E\le 100$, $E=a/b$ $0.7\le Pr_{Tb}\le 100$ $0.3\le n\le 1.5$ $10\le Gr_{Tb}\le 10^5$, where a=axis normal to the direction of gravity b=axis along the direction of gravity
Horizontal cylinder Circular cylinder (Prhashanna and Chhabra, 2011) $Nu_{Tr}=1.19Z^{0.89}$, where $Z=Gr_{Tr}^{\frac{1}{2(n+1)}}Pr_{Tr}^{\frac{1}{3n+1}}$	$0.72\le Pr_{Tr}\le 100$ $0.3\le n\le 1.8$ $10\le Gr_{Tr}\le 10^5$
Elliptical cylinder (Sasmal and Chhabra, 2012a) $Nu_{Tb}=0.833Z^{0.89}E^{\frac{c_1}{2(n+1)}}$ $E<1$: $c_1=-0.464$ $E>1$: $c_1=-1.285$, where $Z=Gr_{Tb}^{\frac{1}{2(n+1)}}Pr_{Tb}^{\frac{1}{3n+1}}$	$0.2\le E\le 5$, $E=a/b$ $0.72\le Pr_{Tb}\le 100$ $0.3\le n\le 1.5$ $10\le Gr_{Tb}\le 10^5$, where a=axis normal to the direction of gravity b=axis along the direction of gravity
Semicircular cylinder: flat face upward (Chandra and Chhabra, 2012) $Nu_{TD}=0.93Z^{0.79}$, where $Z=Gr_{TD}^{\frac{1}{2(n+1)}}Pr_{TD}^{\frac{1}{3n+1}}$	$0.72\le Pr_{TD}\le 100$ $0.2\le n\le 1.8$ $10\le Gr_{TD}\le 10^5$
Semicircular cylinder: flat face downward (Tiwari and Chhabra, 2013) $Nu_{TD}=0.72Z^{0.90}$, where $Z=Gr_{TD}^{\frac{1}{2(n+1)}}Pr_{TD}^{\frac{1}{3n+1}}$	$0.72\le Pr_{TD}\le 100$ $0.2\le n\le 1.8$ $10\le Gr_{TD}\le 10^5$
Square cylinder (Sasmal and Chhabra, 2011) $Nu_{TB}=0.605Z^{0.924}$, where $Z=Gr_{TB}^{\frac{1}{2(n+1)}}Pr_{TB}^{\frac{1}{3n+1}}$	$0.72\le Pr_{TB}\le 100$ $0.3\le n\le 1.8$ $10\le Gr_{TB}\le 10^5$, where B=side of the cylinder
Square cylinder: tilted by 45° (Sasmal and Chhabra, 2012b) $Nu_{Td}=0.761Z^{0.924}$, where $Z=Gr_{Td}^{\frac{1}{2(n+1)}}Pr_{Td}^{\frac{1}{3n+1}}$	$0.72\le Pr_{Td}\le 100$ $0.3\le n\le 1.8$ $10\le Gr_{Td}\le 10^5$, where d=diagonal of the cylinder

(*Continued*)

TABLE 3.2.18 (*Continued*)
Natural Convection Correlations for External Flow over Isothermal Surfaces for Power Law Fluids

Correlation	Range
Triangular cylinder: Flat face upward (Tiwari and Chhabra, 2014) $Nu_{TB}=0.86Z^{1.089}$ $Nu_{TB}=1.45Z^{0.91}$, where $Z=\mathrm{Gr}_{TB}^{\frac{1}{2(n+1)}}\mathrm{Pr}_{TB}^{\frac{1}{3n+1}}$	$0.72 \le \mathrm{Pr}_{TB} \le 100$ $0.2 \le n \le 1.8$ $10 \le \mathrm{Gr}_{TB} \le 10^5$, where B=side of the cylinder
$\mathrm{Gr}_{TL_c}=\dfrac{\rho\left(L_c\right)^{n+2}\left(g\beta\Delta T\right)^{2-n}}{m^2}$; $\mathrm{Pr}_{TL_c}=\dfrac{\rho C}{k}\left(\dfrac{m}{\rho}\right)^{\left(\frac{2}{1+n}\right)}\left(L_c\right)^{\frac{1-n}{1+n}}\left(L_c g\beta\Delta T\right)^{\frac{3(n-1)}{2(n+1)}}$;	
$\mathrm{Pr}_{TL_c}=\dfrac{\rho C}{k}\left(\dfrac{m}{\rho}\right)^{\left(\frac{2}{1+n}\right)}\left(L_c\right)^{\frac{1-n}{1+n}}\left(L_c g\beta\Delta T\right)^{\frac{3(n-1)}{2(n+1)}}$, where L_c=characteristic length	

TABLE 3.2.19
Natural Convection Correlations for External Flow over Isothermal Surfaces in Bingham Plastic Fluids

Correlation	Range
Sphere (Nirmalkar et al., 2014b) for $\mathrm{Bn}_{TD} < (\mathrm{Bn}_{TD})_{\max}$ $\mathrm{Nu}_{TD}=2+\dfrac{0.55\mathrm{Ra}_{TD}^{1/4}\left(1-\dfrac{\mathrm{Bn}_{TD}}{(\mathrm{Bn}_{TD})_{\max}}\right)^{4.66}}{\left(1+\mathrm{Bn}_{TD}\right)^{0.178}}$ for $\mathrm{Bn}_{TD} \ge (\mathrm{Bn}_{TD})_{\max}$ $\mathrm{Nu}_{TD}=2$	$10 \le \mathrm{Pr}_{TD} \le 100$ $10^2 \le \mathrm{Ra}_{TD} \le 10^6$ $0 \le \mathrm{Bn}_{TD} \le 10^4$
Circular cylinder (Nirmalkar et al., 2014a) $\mathrm{Nu}_{TD}=(\mathrm{Nu}_{TD})_{\infty}+0.242\left(\mathrm{Ra}_{TD}^{*}\right)^{2/5}\left(\mathrm{Pr}_{TD}^{*}\right)^{-0.24}$, where $(\mathrm{Nu}_{TD})_{\infty}=\dfrac{2}{\ln\left(D_{\infty}/D\right)}$ $\mathrm{Ra}_{TD}^{*}=\dfrac{\mathrm{Ra}_{TD}}{1+\mathrm{Bn}_{TD}}$ $\mathrm{Pr}_{TD}^{*}=\mathrm{Pr}_{TD}\left(1+\mathrm{Bn}_{TD}\right)$	$30 \le \mathrm{Pr}_{TD} \le 100$ $10^2 \le \mathrm{Ra}_{TD} \le 10^6$ $0 \le \mathrm{Bn}_{TD} \le 10^4$, where D=diameter of the cylinder D_{∞}=diameter of the material envelope
Elliptical cylinder (Patel and Chhabra, 2016) $\mathrm{Nu}_{Ta}=(\mathrm{Nu}_{Ta})_{\infty}+c_1\left(\mathrm{Ra}_{TD}^{*}\right)^{0.29}\left(1+0.0024\mathrm{Pr}_{Ta}^{*}\right)^{-0.95}\left(1+E\right)^{d_1}$ $E \le 1$: $c_1=0.325$, $d_1=0.35$ $E > 1$: $c_1=0.503$, $d_1=-0.26$ For $(\mathrm{Nu}_{Ta})_{\infty}$ refer Table 3.2.21	$0.1 \le E \le 10$, $E=b/a$ $10 \le \mathrm{Pr}_{Ta} \le 100$ $10^2 \le \mathrm{Ra}_{Ta} \le 10^6$ $\le \mathrm{Bn}_{Ta} \le 10^5$, where a=axis normal to the direction of gravity b=axis along the direction of gravity
Semicircular cylinder: Flat face downward (Tiwari and Chhabra, 2016) $\mathrm{Nu}_{TD}=0.5039+0.51\left(\mathrm{Ra}_{\mathrm{TD}}^{*}\right)^{1/4}\left(1+0.0023\mathrm{Pr}_{\mathrm{TD}}^{*}\right)^{-1.23}$	$10 \le \mathrm{Pr}_{TD} \le 100$ $10^2 \le \mathrm{Ra}_{TD} \le 10^6$ $0.01 \le \mathrm{Bn}_{TD} \le 10^5$
$\mathrm{Pr}_{TL_c}=\dfrac{C\mu_B}{k}$; $\mathrm{Ra}_{TL_c}=\dfrac{\rho_{\infty}^2 Cg\Delta T\left(L_c\right)^3}{\mu_B k}$; $\mathrm{Bn}_{TL_c}=\dfrac{\tau_o L_c}{\mu_B V_c}$, where, L_c=characteristic length	

TABLE 3.2.20
Values of Coefficients c_1 and d_1 Used in Table 3.2.18

E	$(Nu_{Tb})_\infty$	$\mathbb{Z} \le 10$		$\mathbb{Z} > 10$	
		c_1	b_1	c_1	b_1
0.2	5.37	0.81	1.15	1.07	0.99
0.5	3.08	0.77	1.12	1.60	0.83
1	2.00	0.76	1.09	1.39	0.87
2	1.20	0.67	1.12	1.28	0.84
5	0.52	0.56	1.02	1.69	0.57

TABLE 3.2.21
Average Nusselt Number in the Pure Conduction Limit Used in Table 3.2.19

		$(Nu_{Ta})_\infty$
$D_\infty/(b)$	E	Numerical
155	0.1	0.5496
475	0.2	0.4483
550	0.5	0.3934
710	1	0.3048
1100	2	0.1973
2370	5	0.0900
1560	10	0.0547

D_∞, diameter of the hypothetical domain.

TABLE 3.2.22
Forced Convection Correlations for External Flow Over Isothermal Surfaces for Power Law Fluids

Correlation	Range
Sphere (Dhole et al., 2006)	$1 \le Pr_{Tr} \le 400$
Constant wall temperature:	$0.5 \le n \le 2$
$Nu_{TD} = 0.922 + Re_{TD}^{1/(n+2)} Pr_{TD}^{1/3} + 0.1 Re_{TD}^{2/3} Pr_{TD}^{1/3}$	$5 \le Re_{TD} \le 200$
Constant wall heat flux:	
$Nu_{HD} = 1 + Re_{HD}^{1/(n+2)} Pr_{HD}^{1/3} + 0.158 Re_{HD}^{2/3} Pr_{HD}^{1/3}$	
Spheroids (Sreenivasulu et al., 2014)	$0.5 \le E \le 1.5$, $E = b/a$
$Nu = 2E^{0.3} + Pr^{0.37} \left(0.77 Re_{T(2b)}^{\left(\frac{0.766}{n+1.31}\right)} \right) E^{0.07} + 0.02 Re_{2b}^{1.05} E^{0.63}$	$5 \le Pr_{T(2b)} \le 100$
	$0.5 \le n \le 1.5$
	$5 \le Re_{T(2b)} \le 100$
E = aspect ratio of the spheroid	a = semi-axis parallel to the direction of flow
	b = semi-axis normal to the direction of flow

(*Continued*)

TABLE 3.2.22 (*Continued*)
Forced Convection Correlations for External Flow Over Isothermal Surfaces for Power Law Fluids

Correlation	Range
Hemisphere (Sasmal et al., 2013) Constant wall temperature: $\mathrm{Nu}_{TD} = 1.80\mathrm{Re}_{TD}^{1/3}\mathrm{Pr}_{TD}^{1/3}\left(\frac{3n+1}{4n}\right)^{0.21}$ Constant wall heat flux: $\mathrm{Nu}_{TD} = 1.85\mathrm{Re}_{HD}^{1/3}\mathrm{Pr}_{HD}^{1/3}\left(\frac{3n+1}{4n}\right)^{0.24}$	$0.72 \le \mathrm{Pr}_D \le 100$ $0.1 \le n \le 1$ $1 \le \mathrm{Re}_D \le 150$
Circular cylinder (Bharti et al., 2007) Constant wall temperature: $\mathrm{Nu}_{TD} = \phi(n)\mathrm{Re}_{TD}^{a_1}\mathrm{Pr}_{TD}^{a_2}$, where $\phi(n) = 0.052^{(-0.091n+0.246)}\left(\frac{3n+1}{4n}\right)^{2.552}$ $a_1 = \frac{0.495}{0.032n+1}$, $a_2 = \frac{0.721}{0.119n+2}$ Constant heat flux: $\mathrm{Nu}_{HD} = \phi(n)\mathrm{Re}_{HD}^{a_1}\mathrm{Pr}_{HD}^{a_2}$, where $\phi(n) = 0.146^{(-0.104n+0.226)}\left(\frac{3n+1}{4n}\right)^{2.161}$ $a_1 = \frac{0.483}{0.049n+1}$, $a_2 = \frac{0.704}{0.102n+2}$	$1 \le \mathrm{Pr}_{TD} \le 1000$ $0.6 \le n \le 2$ $5 \le \mathrm{Re}_{TD} \le 40$
Elliptical cylinder (Bharti et al., 2008) $\mathrm{Nu} = \lambda + F(n)\mathrm{Re}_{T(2a)}^{\alpha}\mathrm{Pr}_{T(2a)}^{\beta}$, where $F(n) = a^{(-b_1n+c_1)}\left(\frac{3n+1}{4n}\right)^{d_1}$ $\alpha = \frac{e_1}{f_1n+1}$ & $\beta = \frac{g_1}{l_1n+2}$ For values of constants used in above equations, refer Table 3.2.23	$0.2 \le E \le 5$, $E = b/a$ $0.2 \le n \le 1.8$ $0.01 \le \mathrm{Re}_{T(2a)} \le 40$ $1 \le \mathrm{Pr}_{T(2a)} \le 100$ a = semi-axis normal to the direction of flow b = semi-axis along the direction of flow
Semicircular cylinder: curved surface facing flow (Chandra and Chhabra, 2011) $\mathrm{Nu}_{TD} = 0.64\mathrm{Re}_{TD}^{0.49}\mathrm{Pr}_{TD}^{0.36}\left(\frac{3n+1}{4n}\right)^{0.76}$	$1 \le \mathrm{Pr}_{Tr} \le 100$ $0.2 \le n \le 1.8$ $0.01 \le \mathrm{Re}_{TD} \le 30$
Square cylinder (Dhiman et al., 2007) Constant wall temperature: $\mathrm{Nu}_{TD} = 0.635\mathrm{Re}_{TD}^{0.4232}\mathrm{Pr}_{TD}^{1/3}n^{-0.14}$ Constant heat flux: $\mathrm{Nu}_{HD} = 0.6765\mathrm{Re}_{HD}^{0.4413}\mathrm{Pr}_{HD}^{1/3}n^{-0.12}$ $\mathrm{Re}_{L_c} = \frac{\rho(L_c)^n V_c^{2-n}}{m}$; $\mathrm{Pr} = \frac{Cm}{k}\left(\frac{V_c}{L_c}\right)^{n-1}$, where, L_c = characteristic length, V_c = characteristic velocity	$1 \le \mathrm{Pr}_{TD} \le 100$ $0.5 \le n \le 2$ $1 \le \mathrm{Re}_{TD} \le 45$

TABLE 3.2.23
Values of Constants Used in Table 3.2.22

Constants	$E=0.2$	$E=0.5$	$E=1$	$E=2$	$E=5$
λ	0.7864	0.6717	0.4320	0.2094	0.0864
a_1	0.1512	0.1946	0.1102	0.3918	1.6165
b_1	–0.0059	–0.0049	–0.0086	0.0286	–0.1651
c_1	0.4682	0.4654	0.2857	0.7669	–2.3456
d_1	0.4182	0.5135	0.5895	0.6393	0.6362
e_1	0.5614	0.5290	0.4762	0.4541	0.4610
f_1	–0.0074	–0.0127	–0.0286	–0.0104	0.0150
g_1	0.8582	0.8438	0.8151	0.7949	0.7784
l_1	0.1674	0.1913	0.2126	0.2379	0.2245

TABLE 3.2.24
Forced Convection Correlations for External Flow over Isothermal Surfaces for Bingham Plastic Fluids

Correlation	Range
Two-dimensional transverse plate (Raja et al., 2015)	$0.1 \le \mathrm{Re}_L \le 25$ $1 \le \mathrm{Pr}_L \le 100$ $0 \le \mathrm{Bn}_L \le 10^3$ L=length of the plate
Constant wall temperature: $j = \dfrac{1.35}{\mathrm{Re}_{TL}^{*2/3}}$	
Constant wall heat flux: $j = \dfrac{1.23}{\mathrm{Re}_{HL}^{*2/3}}$	
Sphere (Nirmalkar et al., 2013a)	$1 \le \mathrm{Re}_{TD} \le 100$ $1 \le \mathrm{Pr}_{TD} \le 100$ $0 \le \mathrm{Bn}_{TD} \le 10^4$
$j = \dfrac{2.3}{\mathrm{Re}_{TD}^{*2/3}}$ for $10^{-4} \le \mathrm{Re}_{TD}^{*} \le 0.199$	
$j = \dfrac{2.27}{\mathrm{Re}_{TD}^{*2/3}}$ for $0.199 < \mathrm{Re}_{TD}^{*} \le 100$	
Circular cylinder (Nirmalkar and Chhabra, 2014)	$1 \le \mathrm{Re}_{TD} \le 40$ $1 \le \mathrm{Pr}_{TD} \le 100$ $0 \le \mathrm{Bn}_{TD} \le 10^4$
$j = \dfrac{2.37}{\mathrm{Re}_{TD}^{*2/3}}$	
Square cylinder (Nirmalkar et al., 2013b)	$0.1 \le \mathrm{Re}_{Td} \le 40$ $1 \le \mathrm{Pr}_{Td} \le 100$ $0 \le \mathrm{Bn}_{Td} \le 100$
$j = \dfrac{1}{\mathrm{Re}_{Td}^{*2/3}}$,	
where d=side length of the square cylinder	
Semicircular cylinder: Flat face upstream (Tiwari and Chhabra, 2015)	$0.1 \le \mathrm{Re}_{TD} \le 30$ $1 \le \mathrm{Pr}_{TD} \le 100$ $0 \le \mathrm{Bn}_{TD} \le 10^3$
Constant wall temperature : $j = \dfrac{1.56}{\mathrm{Re}_{TD}^{*2/3}}$	
Constant wall heat flux : $j = \dfrac{1.47}{\mathrm{Re}_{HD}^{*2/3}}$	

(*Continued*)

TABLE 3.2.24 (*Continued*)
Forced Convection Correlations for External Flow over Isothermal Surfaces for Bingham Plastic Fluids

Correlation	Range
Elliptical cylinder (Patel and Chhabra, 2014) $0.1 \le E \le 1 : \frac{j}{j_{E=1}} = a' - b' \frac{\mathrm{Bn}_{2b}^{d'}}{E^{e'}}$ $1 < E \le 10 : \frac{j}{j_{E=1}} = \frac{k'}{\left(n' + E\right)^{f'}}$ For values of constants used in above equations, refer Table 3.2.25	$0.1 \le E \le 10$, $E = a/b$ $1 \le \mathrm{Re}_{2b} \le 40$ $1 \le \mathrm{Pr}_{2b} \le 100$ $0.01 \le \mathrm{Bn}_{2b} \le 100$, where a = semi-axis along the direction of flow b = semi-axis normal to the direction of flow
Spheroid (Gupta and Chhabra, 2014) $j = \frac{x_1}{\mathrm{Re}_{\mathrm{T}(2a)}^{*2/3}}$ For values of constant x_1 used in above equation, refer Table 3.2.26	$0.2 \le e \le 5$, $e = b/a$ $0.01 \le \mathrm{Re}_{2a} \le 100$ $1 \le \mathrm{Pr}_{2a} \le 100$ $0 < \mathrm{Bn}_{2a} \le 100$, where a = semi-axis normal the direction of flow b = semi-axis parallel to the direction of flow

$\mathrm{Re}_{L_c}^{*} = \frac{\mathrm{Re}_{L_c}}{1 + \mathrm{Bn}_{L_c}}$; $\mathrm{Pr}_{L_c}^{*} = \mathrm{Pr}_{L_c}\left(1 + \mathrm{Bn}_{L_c}\right)$; $\mathrm{Bn}_{L_c} = \frac{\tau_o \mathrm{L_c}}{\mu_B \mathrm{V}_c}$; $\mathrm{Re}_{L_c} = \frac{V_c \rho L_c}{\mu_B}$; $\mathrm{Pr}_{L_c} = \frac{\mu_B C}{k}$

Colburn j-factor, $j = \frac{Nu}{\mathrm{Re}_{L_c}^{*} \mathrm{Pr}_{L_c}^{*1/3}}$, where L_c = characteristic length, V_c = characteristic velocity

TABLE 3.2.25
Values of Constant a', b', d', e', k', n' and f' Used in Table 3.2.24 for an Elliptical Cylinder

$0.1 \le E \le 1$	CWT	CHF
a'	1.125	1.265
b'	0.048	0.251
d'	0.134	0.053
e'	0.655	0.332
$1 < E \le 10$		
k'	2.582	2.995
n'	2.348	3.120
f'	0.782	0.772

TABLE 3.2.26
Values of Constant x_1 Used in Table 3.2.24 for Spheroids

Range	Aspect ratio (e)	Constant x_1
$9.9\times10^{-3} \le \mathrm{Re}^{*}_{2a} < 0.5$	0.2	2.55
	0.5	2.47
	1	2.18
	2	1.75
	5	1.27
$0.5 \le \mathrm{Re}^{*}_{2a} < 100$	0.2	2.2
	0.5	2.25
	1	2.02
	2	1.66
	5	1.2

ACKNOWLEDGMENTS

The preparation of this section was greatly facilitated by the award of the JC Bose Fellowship (Department of Science and Technology, Government of India, New Delhi) to R.P. Chhabra for the period 2015–2020. This support is gratefully acknowledged.

REFERENCES

Amato, W.S. and Tien, C. 1976. Free convection heat transfer from isothermal spheres in polymer solutions, *Int. J. Heat Mass Transfer*, 19, 1257–1266.

Bharti, R.P., Chhabra, R.P., and Eswaran, V. 2007. Steady forced convection heat transfer from a heated circular cylinder to power-law fluids, *Int. J. Heat Mass Transfer*, 50, 977–990.

Bharti, R.P., Sivakumar, P., and Chhabra, R.P. 2008. Forced convection heat transfer from an elliptical cylinder to power Law fluids, *Int. J. Heat Mass Transfer*, 51, 1838–1853.

Capobianchi, M. and Aziz, A. 2014. Laminar natural convection between a vertical surface with uniform heat flux and pseudoplastic and dilatant fluids, *ASME J. Heat Transfer*, 136(9), 092501.

Capobianchi, M. and Aziz, A. 2012. Laminar natural convection from an isothermal vertical surface to pseudoplastic and dilatant fluids, *ASME J. Heat Transfer*, 134(12), 122502.

Capobianchi, M. and Wagner, D. 2010. Heat transfer in laminar flows of extended modified power law fluids in rectangular ducts, *Int. J. Heat Mass Transfer*, 53, 558–563.

Capobianchi, M. and Irvine, T.F., Jr. 1992. Predictions of pressure drop and heat transfer in concentric annular ducts with modified power law fluids, *Wärme Stoffübertragung*, 27, 209–215.

Chandra, A. and Chhabra, R.P. 2011. Momentum and heat transfer characteristics of a semi-circular cylinder immersed in power-law fluids in the steady flow regime, *Int. J. Heat Mass Transfer*, 54, 2734–2750.

Chandra, A. and Chhabra, R.P. 2012. Laminar free convection from a horizontal semi-circular cylinder to power-law fluids, *Int. J. Heat Mass Transfer*, 55, 2934–2944.

Chandrupatla, A.R. and Sastri, V.M. 1977. Laminar forced convection heat transfer of a non-Newtonian fluid in a square duct, *Int. J. Heat Mass Transfer*, 20, 1315–1324.

Cheng, J.A. 1984. Laminar forced convection heat transfer of power law fluids in isosceles triangular ducts, Ph.D. Thesis, Mechanical Engineering Department, State University of New York at Stony Brook.

Chhabra, R.P. 2006. *Bubbles, Drops and Particles in Non-Newtonian Fluids*, 2nd ed., CRC Press, Boca Raton, FL.

Chhabra, R.P. 2011. Fluid flow and heat transfer from circular and non-circular cylinders submerged in non-Newtonian liquids, *Adv. Heat Transfer*, 43, 289–417.

Cho, Y.I. and Hartnett, J.P. 1982. Non-Newtonian fluids in circular pipe flow, *Adv. Heat Transfer*, 15, 59–141.

Cho, Y.I., Ng, K.S., and Hartnett, J.P. 1980. Viscoelastic fluids in turbulent pipe flow—A new heat transfer correlation, *Lett. Heat Mass Transfer*, 7, 347.

Dhiman, A.K., Chhabra, R.P., and Eswaran, V. 2007. Heat transfer to power-law fluids from a heated square cylinder, *Numer. Heat Transfer* Part A, 52, 185–201.

Dhole, S.D. and Chhabra, R.P. 2006. Forced convection heat transfer from a sphere to non-Newtonian power law fluids, *AIChE J.*, 52, 3658–3667.

Gentry, C.C. and Wollersheim, D.E. 1974. Local free convection to non-Newtonian fluids from a horizontal isothermal cylinder, *ASME J. Heat Transfer*, 96, 3–8.

Gupta, A.K. and Chhabra, R.P. 2014. Spheroids in viscoplastic fluids: Drag and heat transfer, *Ind. Eng. Chem. Res*., 53, 18943–18965.

Gupta, A.K., Sasmal, C., Sairamu, M., and Chhabra, R.P. 2014. Free convection in power-law fluids from a spheroid, *Int. J. Heat Mass Transfer*, 75, 592–609.

Hartnett, J.P. and Kostic, M. 1989. Heat transfer to Newtonian and non-Newtonian fluids in rectangular ducts, *Adv. Heat Transfer*, 19, 247–356.

Hartnett, J.P. and Rao, B.K. 1987. Heat transfer and pressure drop for purely viscous non-Newtonian fluids in turbulent flow through rectangular passages, *Wärme Stoffüberetragung*, 21, 261.

Irvine, T.F., Jr. and Karni, J. 1987. Non-Newtonian flow and heat transfer, in *Handbook of Single Phase Convective Heat Transfer*, Wiley, New York, 20–1–20–57.

Irvine, T.F., Jr., Kim, S.C., and Gui, F.L. 1988. Graetz problem solutions for a modified power law fluid, in *ASME Symposium on Fundamentals of Forced Convection Heat Transfer*, ASME publ. HTD 101, pp. 123–127.

Irvine, T.F., Jr., Wu, K.C., and Schneider, W.J. 1982. Vertical channel free convection to a power law fluid, ASME Paper 82-WA/HT-69.

Lee, S.R., Irvine, T.F., Jr., and Greene, G.A. 1998. A computational analysis of natural convection in a vertical channel with a modified power law fluid, in *Proceedings of 11th International Heat Transfer Conference* (IHTC-11), Kyongju, Korea, 367–372.

Nirmalkar, N., Chhabra, R.P., and Poole, R.J. 2013a. Numerical predictions of momentum and heat transfer characteristics from a heated sphere in yield-stress fluids, *Ind. Eng. Chem. Res*., 52, 6848–6861.

Nirmalkar, N., Chhabra, R.P., and Poole, R.J. 2013b. Laminar forced convection heat transfer from a heated square cylinder in a Bingham plastic fluid, *Int. J. Heat Mass Transfer*, 56, 625–639.

Nirmalkar, N. and Chhabra, R.P. 2014. Momentum and heat transfer from a heated circular cylinder in Bingham plastic fluids, *Int. J. Heat Mass Transfer*, 70, 564–577.

Nirmalkar, N., Bose, A., and Chhabra, R.P. 2014a. Free convection from a heated circular cylinder in Bingham plastic fluids, *Int. J. Therm. Sci*., 83, 33–44.

Nirmalkar, N., Gupta, A.K., and Chhabra, R.P. 2014b. Natural convection from a heated sphere in Bingham plastic fluids, *Ind. Eng. Chem. Res.,* 53, 17818–17832.

Patel, S.A. and Chhabra, R.P. 2014. Heat transfer in Bingham plastic fluids from a heated elliptical cylinder, *Int. J. Heat Mass Transfer*, 73, 671–692.

Patel, S.A. and Chhabra, R.P. 2016. Laminar free convection in Bingham plastic fluids from an isothermal elliptical cylinder, *J. Thermophys. Heat Transfer*, 30, 153–168.

Prhashanna, A. and Chhabra, R.P. 2010. Free convection in power-law fluids from a heated sphere, *Chem. Eng. Sci*., 65, 6190–6205.

Prhashanna, A. and Chhabra, R.P. 2011. Laminar natural convection from a horizontal circular cylinder in power-law fluids, *Ind. Eng. Chem. Res*., 50, 2424–2440.

Raja, A.H., Patel, S.A., and Chhabra, R.P. 2015. Laminar forced convection heat transfer from a two-dimensional transverse plate in Bingham plastic fluids, *Int. J. Heat Mass Transfer*, 83, 690–709.

Sasmal, C. and Chhabra, R.P. 2011. Laminar natural convection from a heated square cylinder immersed in power-law liquids, *J. Non-Newtonian Fluid Mech*., 166, 811–830.

Sasmal, C. and Chhabra, R.P. 2012a. Effect of orientation on laminar natural convection from a heated square cylinder in power-law liquids, *Int. J. Therm. Sci*., 57, 112–125.

Sasmal, C. and Chhabra, R.P. 2012b. Effect of aspect ratio on natural convection in power-law liquids from a heated horizontal elliptical cylinder, *Int. J. Heat Mass Transfer*, 55, 4886–4899.

Sasmal, C. and Chhabra, R.P. 2014. Laminar free convection in power-law fluids from a heated hemisphere, *J. Thermophys. Heat Transfer*, 28, 750–763.

Sasmal, C., Shyam, R. and Chhabra, R.P. 2013. Laminar flow of power-law fluids past a hemisphere: momentum and forced convection heat transfer characteristics, *Int. J. Heat Mass Transfer*, 63, 51–64.

Schneider, W.J. and Irvine, T.F., Jr. 1984. Vertical channel free convection for a power law fluid with constant heat flux, ASME Paper 84-HT-16.

Shenoy, A.V. and Mashelkar, R.A. 1982. Thermal convection in non-Newtonian fluids, *Adv. Heat Transfer*, 15, 143–225.

Sreenivasulu, B., Srinivas, B. and Ramesh, K.V. 2014. Forced convection heat transfer from a spheroid to a power-law fluid, *Int. J. Heat Mass Transfer*, 70, 71–80.

Tiwari, A.K. and Chhabra, R.P. 2013. Laminar natural convection in power-law liquids from a heated semi-circular cylinder with its flat side oriented downward, *Int. J. Heat Mass Transfer*, 58, 553–567.
Tiwari, A.K. and Chhabra, R.P. 2014. Effect of orientation on the steady laminar free convection heat transfer in power law fluids from a heated triangular cylinder, *Numer. Heat Transfer* Part A, 65, 780–801.
Tiwari, A.K. and Chhabra, R.P. 2015. Momentum and heat transfer from a semi-circular cylinder in Bingham plastic fluids, *Appl. Math. Model*., 39, 7045–7064.
Tiwari, A.K. and Chhabra, R.P. 2016. Laminar free convection in Bingham plastic fluids from an isothermal semicircular cylinder, *J. Thermophys. Heat Transfer*, 30, 369–378.
Yoo, S.S. 1974. Heat transfer and friction factors for non-Newtonian fluids in turbulent pipe flow, Ph.D. Thesis, University of Illinois at Chicago.

Further Information

Other sources which may be consulted for more detailed information are Cho and Hartnett (1982), Shenoy and Mashelkar (1982), Irvine and Karni (1987), Hartnett and Kostic (1989), and Chhabra (2006, 2011).

3.3 RADIATIVE HEAT TRANSFER

Michael F. Modest

Nature of Thermal Radiation

All materials continuously emit and absorb radiative energy by lowering or raising their molecular energy levels. This thermal radiative energy may be viewed as consisting of electromagnetic waves or of massless energy parcels, called *photons*. Electromagnetic waves travel through any medium at the speed of light c, which is $c_0 = 2.998 \times 10^8$ m/s in vacuum and approximately the same in most gases such as air and combustion products. They are characterized by their wavelength λ (usually measured in $\mu m = 10^{-6}$ m) or frequency ν (usually measured in cycles/s = Hz), which are related by

$$\nu = \frac{c}{\lambda} \tag{3.3.1}$$

The strength and wavelengths of *emission* and *absorption* depend on the temperature and nature of the material.

The ability of photons to travel unimpeded through vacuum and gases makes thermal radiation the dominant mode of heat transfer in vacuum, low-pressure environments, and outer space applications (due to the near absence of conduction and convection). Its temperature dependence (as given by Equation 3.3.3), on the other hand, guarantees that radiative heat transfer is of utmost importance in high-temperature applications (including solar radiation: with the sun being a high-temperature heat source at an effective temperature of $T_{sun} = 5777\,\mathrm{K}$).

When an electromagnetic wave traveling through a gas (or vacuum) strikes the surface of a medium, the wave may be partly or totally reflected and any nonreflected part will penetrate into the medium. If a wave passes through a medium without any attenuation, the material is called *transparent*. A body with partial attenuation is known as *semitransparent*, and a body through which none of the incoming radiation penetrates is called *opaque*. Most gases are rather transparent to radiation (except for narrow spectral regions, called absorption bands), while most solids tend to be strong absorbers for most wavelengths, making them opaque over a distance of a few nanometers (electrical conductors, i.e., metals) to a few micrometers (ceramics and semiconductors), or more (dielectrics).

Blackbody Radiation

The total amount of radiative energy emitted from a surface into all directions above it is termed emissive power; we distinguish between spectral (at a given wavelength λ, per unit wavelength) and total (encompassing all wavelengths) emissive power. The magnitude of emissive power depends on wavelength λ, temperature T, and a surface property, called emittance ε, which relates the ability of a surface to emit radiative energy to that of an ideal surface, which emits the maximum possible energy (at a given wavelength and temperature). Such an ideal surface is known as a "blackbody" or "black surface," since it absorbs all incoming radiation; that is, it reflects no radiation and is, therefore, invisible ("black") to the human eye. The spectral distribution of the emissive power of a black surface (i.e., its dependence on the spectral variable, say λ) is given by Planck's law as

$$E_{b\lambda} = \frac{C_1}{\lambda^5 \left[e^{C_2/\lambda T} - 1 \right]}, \quad C_1 = 3.7419 \times 10^{-16}\,\mathrm{Wm}^2, \quad C_2 = 14{,}388\,\mu\mathrm{mK} \tag{3.3.2}$$

where C_1 and C_2 are sometimes called Planck function constants, or the first and second radiation constant, respectively. The total emissive power of a blackbody is given by

$$E_b = \int_0^\infty E_{b\lambda}\, d\lambda = \sigma T^4, \quad \sigma = 5.670 \times 10^{-8}\ \mathrm{W/m^2\,K^4} \tag{3.3.3}$$

with σ known as the Stefan–Boltzmann constant. Figure 3.3.1 shows the spectral solar irradiation that impinges on Earth, which closely resembles the spectrum of a blackbody at 5777 K. The general behavior of Planck's law is depicted in Figure 3.3.2, together with the fractional emissive power, $f(\lambda T)$, defined as

$$f(\lambda T) = \frac{1}{E_b}\int_0^\lambda E_{b\lambda}(\lambda, T)\ d\lambda \tag{3.3.4}$$

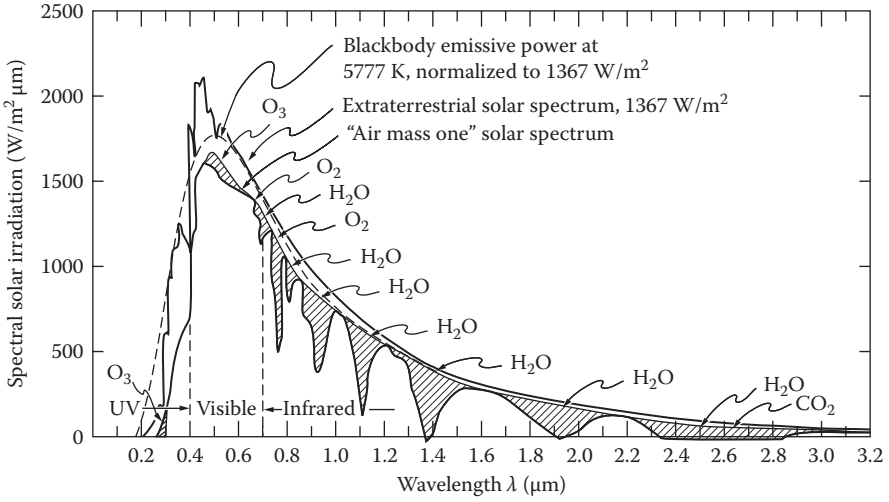


FIGURE 3.3.1 Solar irradiation onto Earth.

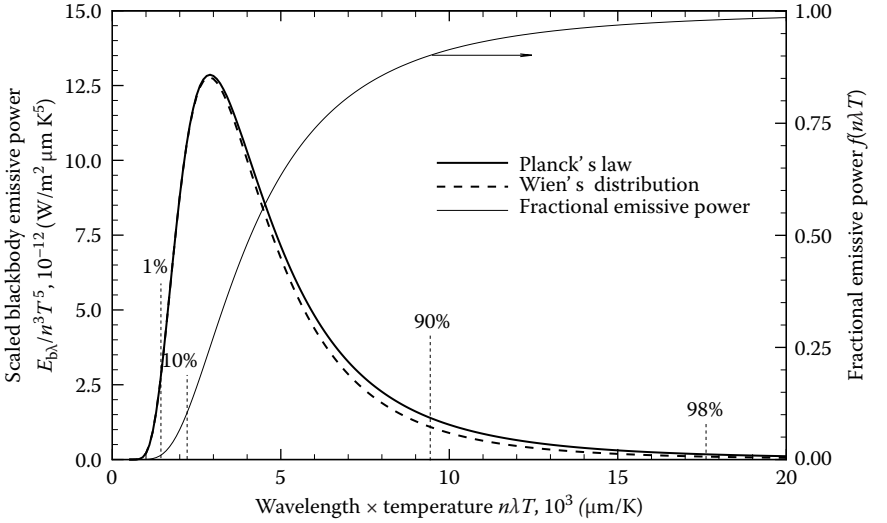


FIGURE 3.3.2 Normalized blackbody emissive power spectrum.

Note that 90% of all blackbody emission takes place between the wavelengths of $\lambda T > 2200\,\mu\text{m K}$ and $\lambda T < 9400\,\mu\text{m K}$. This implies that—for typical high-temperature heat-transfer applications in the range between 1000 and 2000 K—infrared wavelengths in the range $1\,\mu\text{m} < \lambda < 10\,\mu\text{m}$ govern the heat-transfer rates. For solar applications shorter wavelengths, down to $\lambda \cong 0.4\,\mu\text{m}$ are also important. Also shown in Figure 3.3.2 is Wien's law:

$$E_{b\lambda} = \frac{C_1}{\lambda^5} e^{-C_2/\lambda T} \tag{3.3.5}$$

which approximates Planck's law accurately over the part of the spectrum that is important to heat transfer, and that is easier to manipulate mathematically.

Example 3.3.1

What fraction of total solar emission falls into the visible spectrum (0.4 – 0.7 μm)?

Solution: With a solar temperature of 5777 K, it follows that for

$$\lambda_1 = 0.4\,\mu\text{m}, \quad \lambda_1 T_{\text{sun}} = 0.4 \times 5777 = 2311\ \mu\text{mK}$$

and for

$$\lambda_2 = 0.7\,\mu\text{m}, \quad \lambda_2 T_{\text{sun}} = 0.7 \times 5777 = 4044\ \mu\text{mK}$$

From Figure 3.3.2, we can estimate $f(\lambda_1 T_{\text{sun}}) \cong 12\%$ and $f(\lambda_2 T_{\text{sun}}) \cong 49\%$. Thus, the visible fraction of sunlight is $49 - 12 \cong 37\%$: With a bandwidth of only 0.3 μm, the human eye responds to approximately 37% of all emitted sunlight!

Radiative Exchange between Opaque Surfaces

Radiative Properties of Surfaces

Strictly speaking, the surface of an enclosure wall can only reflect radiative energy and allow a part of it to penetrate into the substrate. A surface cannot absorb or emit photons: Attenuation takes place inside the solid, as does emission of radiative energy (with some of the emitted energy escaping through the surface into the enclosure). In practical systems, the thickness of the surface layer over which absorption of *irradiation* from inside the enclosure occurs is very small compared with the overall dimension of an enclosure—usually a few nanometers for metals and a few micrometers for most nonmetals. The same may be said about emission from within the walls that escapes into the enclosure. Thus, in the case of opaque walls it is customary to speak of absorption by and emission from a "surface," although a thin surface layer is implied. Four fundamental radiative properties are defined:

$$\text{Reflectance}, \rho \equiv \frac{\text{Reflected part of incoming radiation}}{\text{Total incoming radiation}} \tag{3.3.6a}$$

$$\text{Absorptance}, \rho \equiv \frac{\text{Absorbed part of incoming radiation}}{\text{Total incoming radiation}} \tag{3.3.6b}$$

$$\text{Transmittance}, \tau \equiv \frac{\text{Transmitted part of incoming radiation}}{\text{Total incoming radiation}} \tag{3.3.6c}$$

$$\text{Emittance}, \varepsilon \equiv \frac{\text{Energy emitted from a surface}}{\text{Energy emitted by a black surface at same temperature}} \tag{3.3.6d}$$

Since all incoming radiation must be reflected, absorbed, or transmitted, it follows that

$$\rho + \alpha + \tau = 1 \tag{3.3.7}$$

In most practical applications, surface layers are thick enough to be opaque ($\tau = 0$)leading to $\rho + \alpha = 1$. All four properties may be functions of wavelength, temperature, incoming direction (except emittance), and outgoing direction (except absorptance).

Directional behavior. For heat-transfer applications, the dependence on incoming direction for absorptance (as well as ρ and τ) and outgoing direction for emittance is generally weak and is commonly neglected; that is, it is assumed that the surface absorbs and emits *diffusely.* Then, for an opaque surface, at any given wavelength

$$\varepsilon_\lambda = \alpha_\lambda = 1 - \rho_\lambda \tag{3.3.8}$$

Published values of emittance are generally either "normal emittance" (the directional value of ε_λ in the direction perpendicular to the surface) or "hemispherical emittance" (an average value over all outgoing directions). The difference between these two values is often smaller than experimental accuracy and/or repeatability.

Reflected energy (due to a single, distinct incoming direction) may leave the surface into a single direction ("specular" reflection, similar to reflection from a mirror for visible light), or the reflection may spread out over all possible outgoing directions. In the extreme case of equal amounts going into all directions, we talk about "diffuse" reflection. Smooth surfaces (as compared with the wavelength of radiation) tend to be specular reflectors, while rough surfaces tend to be more or less diffusely reflecting. Analysis is vastly simplified if diffuse reflections are assumed. Research has shown that—except for some extreme geometries and irradiation conditions susceptible to beam channeling (irradiated open cavities, channels with large aspect ratios)—radiative heat-transfer rates are only weakly affected by the directional distribution of reflections. Therefore, it is common practice to carry out radiative heat-transfer calculations assuming only diffuse reflections.

Spectral dependence. The emittance of a surface generally varies strongly and in complex ways with wavelength, depending on the material, surface layer composition, and surface structure (roughness). Therefore, unlike bulk material properties (such as thermal conductivity) the surface emittance may display significant differences between two ostensibly identical samples, and even for one and the same sample measured at different times (due to surface roughness and contamination). Despite these difficulties, surfaces may be loosely grouped into two categories—metals and nonconductors (dielectrics), and some generalizations can be made.

Polished metals. Smooth, purely metallic surfaces (i.e., without any nonmetallic surface contamination, such as metal oxides) tend to have very low emittances in the infrared. For many clean metals $\varepsilon_\lambda < 0.1$ for $\lambda > 2\,\mu\text{m}$, and spectral as well as temperature dependence are generally well approximated by the proportionality $\varepsilon_\lambda \propto \sqrt{T/\lambda}$ in the infrared. However, for shorter wavelengths ($\lambda < 1\,\mu\text{m}$), emittance values may become quite substantial, and temperature dependence is usually reversed (decreasing, rather than increasing, with temperature). Typical room temperature behavior of several metals is shown in Figure 3.3.3. Caution needs to be exercised when choosing an emittance value for a metal surface: Unless extraordinary care is taken to keep a polished metal clean (i.e., free from oxidation and/or surface contamination), its emittance may soon become several times the value of the original, polished specimen (e.g., consider the formation of aluminum oxide on top of aluminum, Figure 3.3.3).

Ceramics and refractories. Smooth ceramics tend to have fairly constant and intermediate emittance over the near-to mid-infrared, followed by a sharp increase somewhere between 4 and 10 μm. At short wavelengths, these materials display strong decreases in emittance, so that a number of them may appear white to the human eye even though they are fairly black in the infrared. The temperature dependence of the emittance of ceramics is rather weak; generally a slight increase with temperature is observed in the infrared. The spectral emittance of a few ceramics is also shown in Figure 3.3.3.

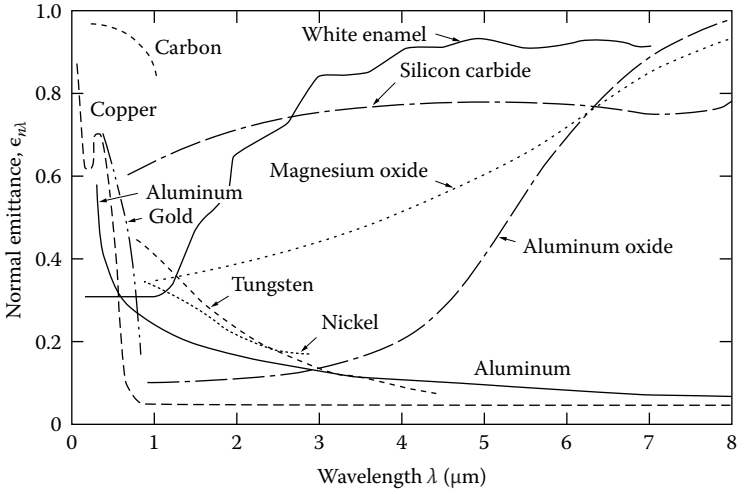


FIGURE 3.3.3 Normal, spectral emittances for selected materials.

Other nonconductors. The behavior of most electrically nonconducting materials is governed by surface structure, nonhomogeneity, dopants, porosity, flaws, surface films, etc. The emittance may vary irregularly across the spectrum because of various emission bands, influence of flaws, etc., making any generalization impossible. This irregularity may be exploited to obtain surfaces of desired spectral behavior, so-called selective surfaces. Some selective surfaces (as compared with a common steel) are depicted in Figure 3.3.4. For a solar collector, it is desirable to have a high spectral emittance for short wavelengths $\lambda < 2.5\mu m$ (strong absorption of solar irradiation), and a low value for $\lambda > 2.5\mu m$ (to minimize reemission from the collector). The opposite is true for a spacecraft radiator panel used to reject heat into space.

It is clear that (1) values of spectral surface emittance are subject to great uncertainty and (2) only a relatively small range of infrared wavelengths are of importance. Therefore, it is often assumed that the surfaces are "gray", that is, the emittance is constant across (the important fraction of) the

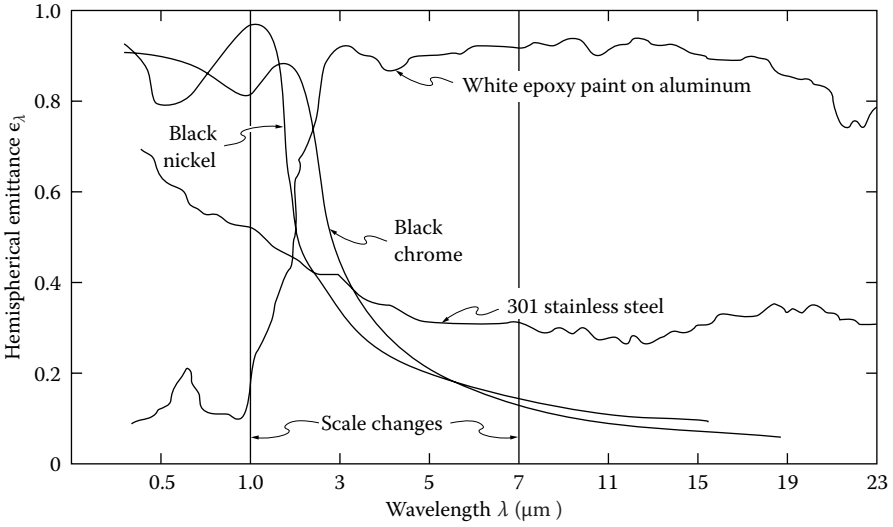


FIGURE 3.3.4 Spectral, hemispherical reflectance of several spectrally selective surfaces.

TABLE 3.3.1
Total Emittance and Solar Absorptance of Selected Surfaces

	Temperature (°C)	Total Normal Emittance	Extraterrestrial Solar Absorptance
Alumina, flame-sprayed	–25	0.80	0.28
Aluminum foil			
As received	20	0.04	
Bright dipped	20	0.025	0.10
Aluminum, vacuum deposited	20	0.025	0.10
Hard anodized	–25	0.84	0.92
Highly polished plate, 98.3% pure	225–575	0.039–0.057	
Commercial sheet	100	0.09	
Rough polish	100	0.18	
Rough plate	40	0.055–0.07	
Oxidized at 600°C	200–600	0.11–0.19	
Heavily oxidized	95–500	0.20–0.31	
Antimony, polished	35–260	0.28–0.31	
Asbestos	35–370	0.93–0.94	
Beryllium	150	0.18	0.77
	370	0.21	
	600	0.30	
Beryllium, anodized	150	0.90	
	370	0.88	
	600	0.82	
Bismuth, bright	75	0.34	
Black paint			
Parson's optical black	–25	0.95	0.975
Black silicone	–25–750	0.93	0.94
Black epoxy paint	–25	0.89	0.95
Black enamel paint	95–425	0.81–0.80	
Brass, polished	40–315	0.10	
Rolled plate, natural surface	22	0.06	
Dull plate	50–350	0.22	
Oxidized by heating at 600°C	200–600	0.61–0.59	
Carbon, graphitized	100–320	0.76–0.75	
	320–500	0.75–0.71	
Candle soot	95–270	0.952	
Graphite, pressed, filed surface	250–510	0.98	
Chromium, polished	40–1100	0.08–0.36	
Copper, electroplated	20	0.03	0.47
Carefully polished electrolytic copper	80	0.018	
Polished	115	0.023	
Plate heated at 600°C	200–600	0.57	
Cuprous oxide	800–1100	0.66–0.54	
Molten copper	1075–1275	0.16–0.13	
Glass, Pyrex, lead, and soda	260–540	0.95–0.85	
Gypsum	20	0.903	
Gold, pure, highly polished	225–625	0.018–0.035	

(*Continued*)

TABLE 3.3.1 (*Continued*)
Total Emittance and Solar Absorptance of Selected Surfaces

	Temperature (°C)	Total Normal Emittance	Extraterrestrial Solar Absorptance
Inconel X, oxidized	–25	0.71	0.90
Lead, pure (99.96%), unoxidized	125–225	0.057–0.075	
Gray oxidized	25	0.28	
Oxidized at 150°C	200	0.63	
Magnesium oxide	275–825	0.55–0.20	
	900–1705	0.20	
Magnesium, polished	35–260	0.07–0.13	
Mercury	0–100	0.09–0.12	
Molybdenum, polished	35–260	0.05–0.08	
	540–1370	0.10–0.18	
	2750	0.29	
Nickel, electroplated	20	0.03	0.22
Polished	100	0.072	
Platinum, pure, polished	225–625	0.054–0.104	
Silica, sintered, powdered, fused silica	35	0.84	0.08
Silicon carbide	150–650	0.83–0.96	
Silver, polished, pure	40–625	0.020–0.032	
Stainless steel			
Type 312, heated 300 h at 260°C	95–425	0.27–0.32	
Type 301 with Armco black oxide	–25	0.75	0.89
Type 410, heated to 700°C in air	35	0.13	0.76
Type 303, sandblasted	95	0.42	0.68
Titanium, 75A	95–425	0.10–0.19	
75A, oxidized 300 h at 450°C	35–425	0.21–0.25	0.80
Anodized	–25	0.73	0.51
Tungsten, filament, aged	27–3300	0.032–0.35	
Zinc, pure, polished	225–325	0.045–0.053	
Galvanized sheet	100	0.21	

spectrum, $\varepsilon_\lambda \neq \varepsilon_\lambda(\lambda)$, since this assumption also vastly simplifies analysis. Table 3.3.1 gives a fairly detailed listing of total emittances of various materials, defined as

$$\varepsilon(T) = \frac{1}{E_b(T)} \int_0^\infty \varepsilon_\lambda(\lambda, T) E_{b\lambda}(T)\ d\lambda \tag{3.3.9}$$

which may be enlisted for a gray analysis.

View Factors

In many engineering applications, the exchange of radiative energy between surfaces is virtually unaffected by the medium that separates them. Such (radiatively) *nonparticipating media* include vacuum as well as monatomic and most diatomic gases (including air) at low to moderate temperature levels (i.e., before ionization and dissociation occurs). Examples include spacecraft heat rejection systems, solar collector systems, radiative space heaters, illumination problems, and so on. It is common practice to simplify the analysis by making the assumption of an *idealized enclosure* and/or of *ideal surface properties*. The greatest simplification arises if all surfaces are black: For such a situation no

reflected radiation needs to be accounted for, and all emitted radiation is diffuse (i.e., the radiative energy leaving a surface does not depend on direction). The next level of difficulty arises if surfaces are assumed to be gray, diffuse emitters (and, thus, absorbers) as well as gray, diffuse reflectors. The vast majority of engineering calculations are limited to such ideal surfaces, since, particularly, the effects of nondiffuse reflections are usually weak (see discussion in previous section).

Thermal radiation is generally a long-range phenomenon. This is always the case in the absence of a participating medium, since photons will travel unimpeded from surface to surface. Therefore, performing a thermal radiation analysis for one surface implies that all surfaces, no matter how far removed, that can exchange radiative energy with one another must be considered simultaneously. How much energy any two surfaces exchange depends in part on their size, separation, distance, and orientation, leading to geometric functions known as *view factors*, defined as

$$F_{i-j} = \frac{\text{Diffuse energy leaving } A_i \text{ directly toward and intercepted by } A_j}{\text{Total diffuse energy leaving } A_i} \tag{3.3.10}$$

In order to make a radiative energy balance, we always need to consider an entire *enclosure* rather than an infinitesimal control volume (as is normally done for other modes of heat transfer, i.e., conduction or convection). The enclosure must be closed so that irradiation from all possible directions can be accounted for, and the enclosure surfaces must be *opaque* so that all irradiation is accounted for, for each direction. In practice, an incomplete enclosure may be closed by introducing artificial surfaces. An enclosure may be idealized in two ways, as indicated in Figure 3.3.5: by replacing a complex geometric shape with a few simple surfaces, and by assuming surfaces to be isothermal with constant (i.e., average) heat flux values across them. Obviously, the idealized enclosure approaches the real enclosure for sufficiently small isothermal subsurfaces.

Mathematically, the view factor needs to be determined from a double surface integral, that is,

$$F_{i-j} = \frac{1}{A_i} \int_{A_i} \int_{A_j} \frac{\cos\theta_i \cos\theta_j}{\pi S_{ij}^2} \, dA_j \, dA_i \tag{3.3.11}$$

where θ_i and θ_j are the angles between the surface normals on A_i and A_j, respectively, and the line (of length S_{ij}) connecting two points on the two surfaces. Analytical solutions to Equation 3.3.11 may be found for relatively simple geometries. A few graphical results for important geometries are shown in Figures 3.3.6 through 3.3.8. More extensive tabulations as well as analytical expressions may be found in textbooks on the subject area (Modest, 2013; Howell et al., 2015) as well as view factor catalogs (Howell, 1982). For nontrivial geometries view factors must be calculated numerically, either (1) by numerical quadrature of the double surface integral in Equation 3.3.11, or (2) by converting Equation 3.3.11 into a double-line integral, followed by numerical quadrature, or (3) by a Monte Carlo method (statistical sampling and tracing of selected light rays).

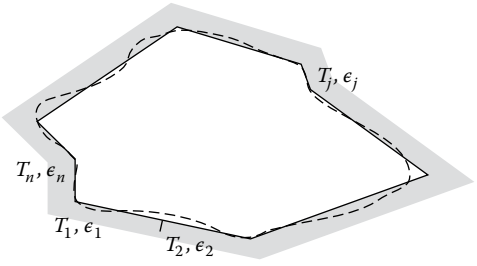


FIGURE 3.3.5 Real and ideal enclosures for radiative transfer calculations.

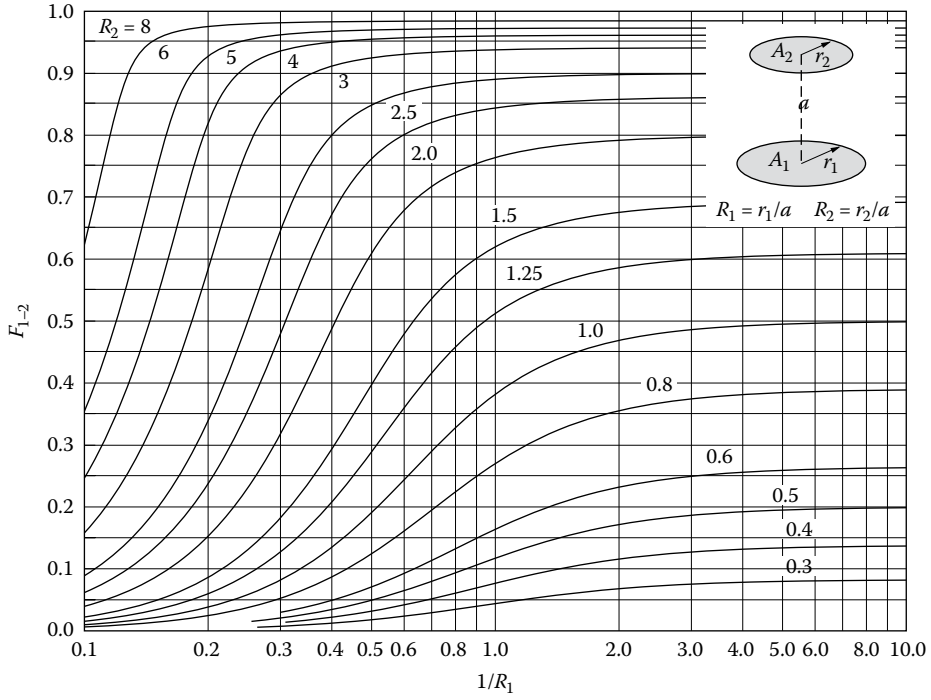


FIGURE 3.3.6 View factor between parallel, coaxial disks of unequal radius.

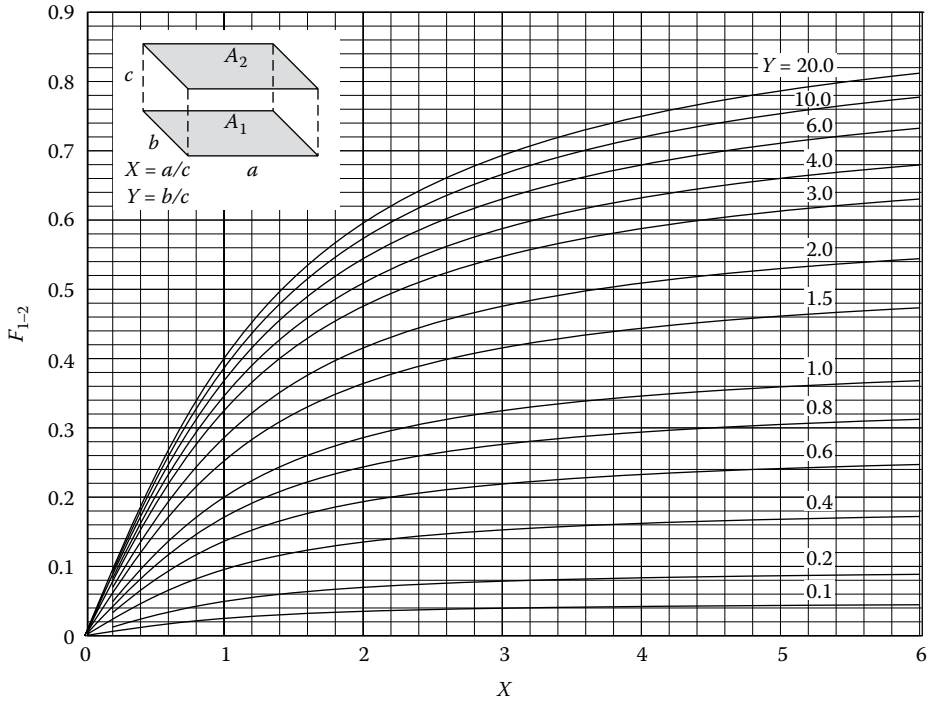


FIGURE 3.3.7 View factor between identical, parallel, directly opposed rectangles.

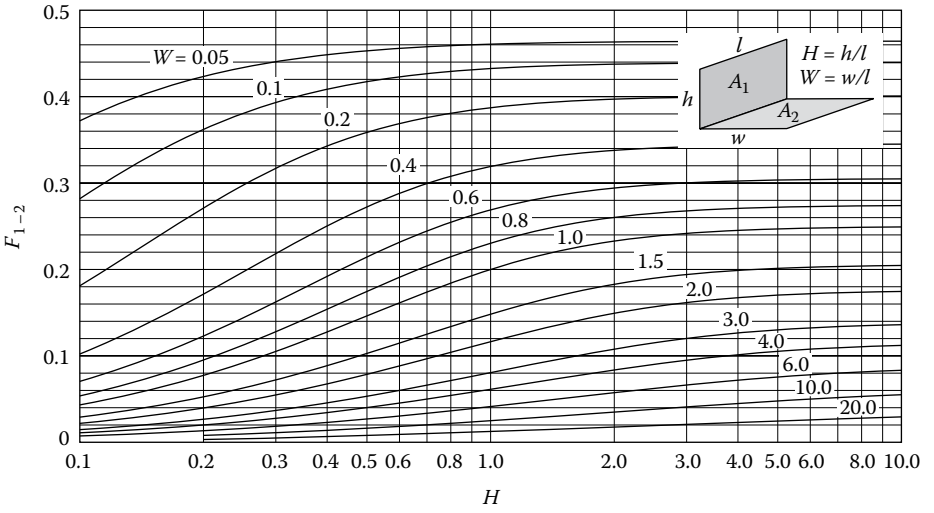


FIGURE 3.3.8 View factor between perpendicular rectangles with common edge.

View Factor Algebra

For simple geometries, analytical values can often be found by expressing the desired view factor in terms of other, known ones. This method is known as view factor algebra, by manipulating the two relations,

$$\text{Reciprocity rule}: A_i F_{i-j} = A_j F_{j-i} \tag{3.3.12}$$

$$\text{Summation rule}: \sum_{j=1}^{N} F_{i-j} = 1, \quad i = 1, N \tag{3.3.13}$$

assuming that the (closed) configuration consists of N surfaces. The reciprocity rule follows immediately from Equation 3.3.11, while the summation rule simply states that the fractional energies leaving surface A_i must add up to a whole.

Example 3.3.2

Assuming the view factor for a finite corner (as shown in Figure 3.3.8) is known, determine the view factor F_{3-4}, between the two perpendicular strips as shown in Figure 3.3.9.

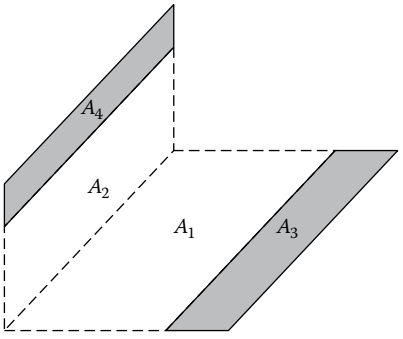


FIGURE 3.3.9 Configuration for Example 3.3.2 (strips on a corner piece).

Solution. From the definition of the view factor, and since the energy traveling to A_4 is the energy going to A_2 and A_4 minus the one going to A_2, it follows that

$$F_{3-4} = F_{3-(2+4)} - F_{3-2}$$

and, using reciprocity,

$$F_{3-4} = \frac{1}{A_3}\left[\left(A_2 + A_4\right)F_{(2+4)-3} - A_2F_{2-3}\right]$$

Similarly, we find

$$F_{3-4} = \frac{A_2 + A_4}{A_3}\left(F_{(2+4)-(1+3)} - F_{(2+4)-1}\right) - \frac{A_2}{A_3}\left(F_{2-(1+3)} - F_{2-1}\right)$$

All view factors on the right-hand side are corner pieces and, thus, are known from Figure 3.3.8.

Crossed-Strings Method: A special type of view factor algebra may be used to determine all the view factors in long enclosures with constant cross section. The method is called the crossed-strings method since the view factors can be determined experimentally with four pins, a roll of string, and a yardstick. Consider the configuration in Figure 3.3.10, which shows the cross section of an infinitely long enclosure, continuing into and out of the plane of the figure. Repeatedly applying reciprocity and summation rules allows the evaluation of F_{1-2} as

$$F_{1-2} = \frac{\left(A_{bc} + A_{ad}\right) - \left(A_{ac} + A_{bd}\right)}{2A_1} \tag{3.3.14}$$

where A_{ab} is the area (per unit depth) defined by the length of the string between points a and b, etc. This formula is easily memorized by looking at the configuration between any two surfaces as a generalized "rectangle," consisting of A_1, A_2, and the two sides A_{ac} and A_{bd}. Then

$$F_{1-2} = \frac{\text{diagonals} - \text{sides}}{2 \times \text{originating area}} \tag{3.3.15}$$

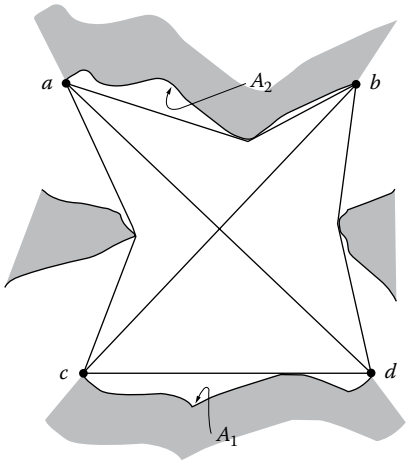


FIGURE 3.3.10 The crossed-strings method for arbitrary two-dimensional configurations.

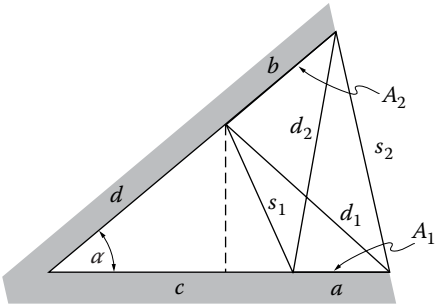


FIGURE 3.3.11 Infinitely long wedge-shaped groove for Example 3.3.3.

Example 3.3.3

Calculate F_{1-2} for the configuration shown in Figure 3.3.11.

Solution. From the figure, it is obvious that

$$s_1^2 = (c - d\cos\alpha)^2 + d^2\sin^2\alpha = c^2 + d^2 - 2cd\cos\alpha$$

Similarly, we have

$$s_2^2 = (a+c)^2 + (b+d)^2 - 2(a+c)(b+d)\cos\alpha$$

$$d_1^2 = (a+c)^2 + d^2 - 2(a+c)d\cos\alpha$$

and

$$d_2^2 = c^2 + (b+d)^2 - 2c(b+d)\cos\alpha$$

$$F_{1-2} = \frac{d_1 + d_2 - (s_1 + s_2)}{2a}$$

Radiative Exchange between Opaque Surfaces (Net Radiation Method)

Consider an enclosure consisting of N opaque surfaces. The enclosure is closed, or, if not, no surface external to the surface reflects or emits radiation into the enclosure (i.e., the open configuration may be artificially closed by replacing openings with cold, black surfaces); any external radiation entering the enclosure is dealt with individually for each surface [see Equation 3.3.17]. All surfaces are assumed to be gray, and emit and reflect diffusely. Traditionally, the *radiosity* J of the surfaces is determined, defined as the total diffuse radiative energy leaving a surface (by emission and reflection),

$$J_i = \varepsilon_i E_{bi} + \rho_i H_i, \quad i = 1, N \tag{3.3.16}$$

where H_i is the incoming radiative flux per unit area (irradiation) onto surface A_i. This leads to N simultaneous equations for the unknown radiosities, specifically,

$$J_i = \varepsilon_i E_{bi} + (1-\varepsilon_i)\left[\sum_{j=1}^{N} J_j F_{i-j} + H_{oi}\right] \tag{3.3.17a}$$

or

$$J_i = q_i + \sum_{j=1}^{N} J_j F_{i-j} + H_{oi} \tag{3.3.17b}$$

depending on whether surface temperature or surface flux are known on surface A_i. In Equation 3.3.17, H_{oi} is irradiation onto surface A_i from outside the enclosure, if any; H_{oi} is always zero for closed configurations, but is useful in the presence of external light sources (such as solar energy, lasers, etc.). The radiosity neither is a useful quantity to determine, nor is there a need to determine it. Eliminating the radiosities from Equations 3.3.17a and 3.3.17b leads to N simultaneous equations in temperature (E_{bi}) and heat flux (q_i):

$$\frac{q_i}{\varepsilon_i} - \sum_{j=1}^{N} \left(\frac{1}{\varepsilon_j} - 1 \right) F_{i-j} q_j + H_{oi} = E_{bi} - \sum_{j=1}^{N} F_{i-j} E_{bj} \tag{3.3.18}$$

Note that no artificial closing surfaces ($j > N$) appear in Equation 3.3.18, since for these surfaces $\varepsilon_j = 1$ and $E_{bj} = 0$. Thus, such closing surfaces may simply be ignored in the analysis.

Since Equation 3.3.18 is a set of N equations, this requires that N values of emissive power E_{bi} and/or flux q_i must be given as boundary conditions, in order to solve for the remaining N unknowns. For computer calculations, Equation 3.3.18 may be recast in matrix form

$$\mathbf{C} \cdot \mathbf{q} = \mathbf{A} \cdot \mathbf{e_b} - \mathbf{h_o} \tag{†(3.3.19a)}$$

where

$$C_{ij} = \frac{\delta_{ij}}{\varepsilon_j} - \left(\frac{1}{\varepsilon_j} - 1 \right) F_{i-j} \tag{3.3.19b}$$

$$A_{ij} = \delta_{ij} - F_{i-j} \tag{3.3.19c}$$

δ_{ij} is Kronecker's delta, that is,

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \tag{3.3.20}$$

and $\mathbf{q}$, $\mathbf{e_b}$, and $\mathbf{h_o}$ are vectors of the surface heat fluxes q_i, emissive powers E_{bi}, and external irradiations H_{oi} (if any). For example, if the temperatures are given for all the surfaces, and the heat fluxes are to be determined, Equation (3.3.19) is solved by matrix inversion, and

$$\mathbf{q} = \left(\mathbf{C^{-1}} \cdot \mathbf{A} \right) \cdot \mathbf{e_b} - \left(\mathbf{C^{-1}} \cdot \mathbf{h_o} \right) \tag{3.3.21}$$

Example 3.3.4

A right-angled groove, consisting of two long black surfaces of width a, is exposed to solar radiation q_{sol} (Figure 3.3.12). The entire groove surface is kept isothermal at temperature T. Determine the net radiative heat-transfer rate from the groove.

Solution. We may employ Equation 3.3.19. However, the enclosure is not closed, and we must close it artificially. We note that any radiation leaving the cavity will not come back (barring any reflection from other surfaces nearby). Thus, our artificial surface should be black. We also assume that, with the exception of the (parallel) solar irradiation, no external radiation enters the cavity.

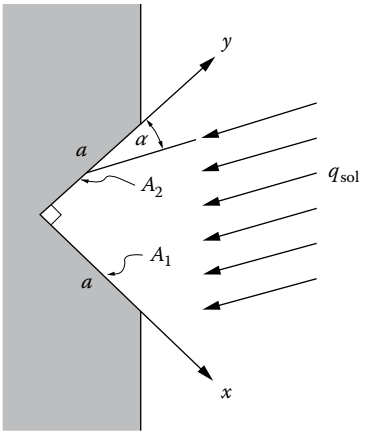


FIGURE 3.3.12 Right-angled groove exposed to solar irradiation, Example 3.3.4.

Since the solar irradiation is best treated separately through the external irradiation term H_o, our artificial surface is nonemitting. Both criteria are satisfied by covering the groove with a black surface at 0 K (A_3). Even though we now have three surfaces, the last one does not really appear in Equation 3.3.18 (since $E_{b3} = 0$ and $1/\varepsilon_3 - 1 = 0$):

$$q_1 = E_{b1} - F_{1-2}E_{b2} - H_{o1} = \sigma T^4(1 - F_{1-2}) - q_{sol}\cos\alpha$$

$$q_2 = E_{b2} - F_{2-1}E_{b1} - H_{o2} = \sigma T^4(1 - F_{2-1}) - q_{sol}\sin\alpha$$

From the crossed-strings method, Equation 3.3.15, we find

$$F_{1-2} = \frac{a + a - \left(\sqrt{2}a + 0\right)}{2a} = \frac{1}{2}\left(2 - \sqrt{2}\right) = 0.293 = F_{2-1}$$

and

$$Q' = a(q_1 + q_2) = a\left[\sqrt{2}\sigma T^4 - q_{sol}(\cos\alpha + \sin\alpha)\right]$$

Example 3.3.5

Consider a very long duct as shown in Figure 3.3.13. The duct is 30×40 cm in cross section, and all surfaces are covered with a gray, diffuse surface material. Top and bottom walls are at

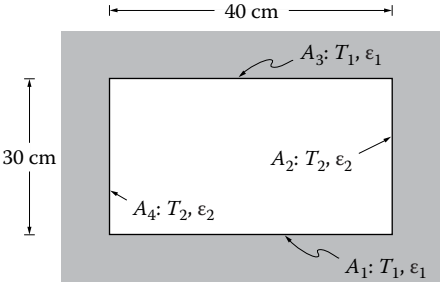


FIGURE 3.3.13 Two-dimensional gray, diffuse duct for Example 3.3.5.

$T_1 = T_3 = 1000$ K with $\varepsilon_1 = \varepsilon_3 = 0.3$, while the side walls are at $T_2 = T_4 = 600$ K with $\varepsilon_2 = \varepsilon_4 = 0.8$ as shown. Determine the net radiative heat-transfer rates for each surface.

Solution. Using Equation 3.3.18 for $i = 1$ and $i = 2$ and noting that $F_{1-2} = F_{1-4}$ and $F_{2-1} = F_{2-3}$, as well as $q_1 = q_3$ and $q_2 = q_4$,

$$i = 1: \quad \frac{q_1}{\varepsilon_1} - 2\left(\frac{1}{\varepsilon_2} - 1\right)F_{1-2}q_2 - \left(\frac{1}{\varepsilon_1} - 1\right)F_{1-3}q_1 = 2F_{1-2}(E_{b1} - E_{b2})$$

$$i = 2: \quad \frac{q_2}{\varepsilon_2} - 2\left(\frac{1}{\varepsilon_1} - 1\right)F_{2-1}q_1 - \left(\frac{1}{\varepsilon_2} - 1\right)F_{2-4}q_2 = 2F_{2-1}(E_{b2} - E_{b1})$$

The view factors are readily evaluated from the crossed-strings method as $F_{1-2} = \frac{1}{4}$, $F_{1-3} = 1 - 2F_{1-2} = \frac{1}{2}$, $F_{2-1} = \frac{4}{3}F_{1-2} = \frac{1}{3}$, and $F_{2-4} = 1 - 2F_{2-1} = \frac{1}{3}$. Substituting these, as well as emittance values, into the relations reduces them to the simpler form of

$$\left[\frac{1}{0.3} - \left(\frac{1}{0.3} - 1\right)\frac{1}{2}\right]q_1 - 2\left(\frac{1}{0.8} - 1\right)\frac{1}{4}q_2 = 2 \times \frac{1}{4}(E_{b1} - E_{b2})$$

$$-2\left(\frac{1}{0.3} - 1\right)\frac{1}{3}q_1 + \left[\frac{1}{0.8} - \left(\frac{1}{0.8} - 1\right)\frac{1}{3}\right]q_2 = 2 \times \frac{1}{3}(E_{b2} - E_{b1})$$

or

$$\frac{13}{6}q_1 - \frac{1}{8}q_2 = \frac{1}{2}(E_{b1} - E_{b2})$$

$$-\frac{14}{9}q_1 + \frac{7}{6}q_2 = -\frac{2}{3}(E_{b1} - E_{b2})$$

Thus,

$$\left(\frac{13}{6} \times \frac{7}{6} - \frac{14}{9} \times \frac{1}{8}\right)q_1 = \left(\frac{1}{2} \times \frac{7}{6} - \frac{2}{3} \times \frac{1}{8}\right)(E_{b1} - E_{b2})$$

$$q_1 = \frac{3}{7} \times \frac{1}{2}(E_{b1} - E_{b2}) = \frac{3}{14}\sigma(T_1^4 - T_2^4)$$

and

$$\left(-\frac{1}{8} \times \frac{14}{9} + \frac{7}{6} \times \frac{13}{6}\right)q_2 = \left(\frac{1}{2} \times \frac{14}{9} - \frac{2}{3} \times \frac{13}{6}\right)(E_{b1} - E_{b2})$$

$$q_2 = -\frac{3}{7} \times \frac{2}{3}(E_{b1} - E_{b2}) = -\frac{2}{7}\sigma(T_1^4 - T_2^4)$$

Finally, substituting values for temperatures,

$$Q_1' = 0.4\,\text{m} \times \frac{3}{14} \times 5.670 \times 10^{-8}\ \text{W/m}^2\text{K}^4(1000^4 - 600^4)\text{K}^4 = 4230\ \text{W/m}$$

$$Q_2' = -0.3\,\text{m} \times \frac{2}{7} \times 5.670 \times 10^{-8}\ \text{W / m}^2\text{K}^4(1000^4 - 600^4)\text{K}^4 = -4230\ \text{W/m}$$

Note that, for conservation of energy, both heat-transfer rates must add up to zero.

Small body inside isothermal enclosure. An especially simple—but important—case occurs if a small, convex body A_1 (i.e., a surface that cannot "see" itself, or $F_{1-1} = 0$) is totally enclosed by an isothermal enclosure A_2. Then, with $N = 2$ and $F_{1-2} = 1$, Equation 3.3.18 reduces to

$$q_1 = \frac{E_{b1} - E_{b2}}{\dfrac{1}{\varepsilon_1} + \dfrac{A_1}{A_2}\left(\dfrac{1}{\varepsilon_2} - 1\right)} = \frac{\sigma(T_1^4 - T_2^4)}{\dfrac{1}{\varepsilon_1} + \dfrac{A_1}{A_2}\left(\dfrac{1}{\varepsilon_2} - 1\right)} \tag{3.3.22}$$

If the enclosure is large, that is, $A_1 \ll A_2$, then Equation 3.3.22 simplifies further to

$$q_1 = \varepsilon_1 \sigma(T_1^4 - T_2^4) \tag{3.3.23}$$

Radiation shields. If it is desired to minimize radiative heat transfer between two surfaces, it is common practice to place one or more radiation shields between them (usually thin metallic sheets of low emittance). If two surfaces A_i and A_j are close together, so that $A_i \cong A_j$ and $F_{i-j} \cong 1$, then the radiative exchange between them is, from Equation 3.3.22,

$$q = \frac{E_{bi} - E_{bj}}{R_{ij}}, \quad R_{ij} = \frac{1}{\varepsilon_i} + \frac{1}{\varepsilon_j} - 1 \tag{3.3.24}$$

where R_{ij} is termed the *radiative resistance*. Equation 3.3.24 is seen to be analogous to an electrical circuit with "current" q and "voltage potential" $E_{bi} - E_{bj}$. Therefore, expressing radiative fluxes in terms of radiative resistances is commonly known as *network analogy*. The network analogy is a very powerful method of solving one-dimensional problems (i.e., whenever only two isothermal surfaces see each other, such as infinite parallel plates, or when one surface totally encloses another). Consider, for example, two large parallel plates, A_1 and A_N, separated by $N - 2$ radiation shields, as shown in Figure 3.3.14. Let each shield have an emittance ε_S on both sides. Then, by applying Equation 3.3.24 to any two consecutive surfaces and using the fact that q remains constant throughout the gap,

$$q = \frac{E_{b1} - E_{b2}}{R_{12}} = \cdots = \frac{E_{bk-1} - E_{bk}}{R_{k-1,k}} = \cdots = \frac{E_{bN-1} - E_{bN}}{R_{N-1,N}} = \frac{E_{b1} - E_{bN}}{\displaystyle\sum_{j-2}^{N} R_{j-1,j}} \tag{3.3.25}$$

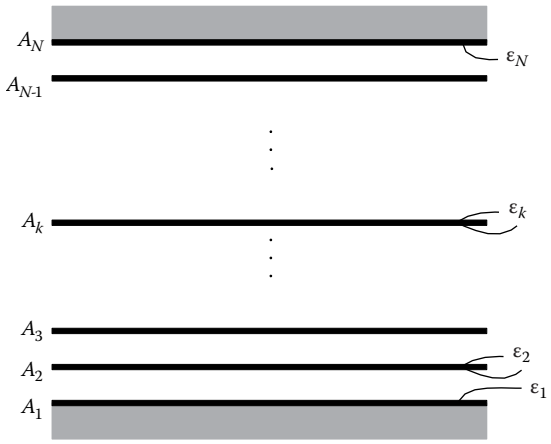


FIGURE 3.3.14 Placement of radiation shields between two large, parallel plates.

where

$$R_{j-1,j} = \frac{1}{\varepsilon_{j-1}} + \frac{1}{\varepsilon_j} - 1 \tag{3.3.26}$$

and, if $= \varepsilon_2 = \varepsilon_3 = \cdots = \varepsilon_{N-1} = \varepsilon_S$,

$$\sum_{j=2}^{N} R_{j-1,j} = \frac{1}{\varepsilon_1} + \frac{1}{\varepsilon_N} - 1 + (N-2)\left(\frac{2}{\varepsilon_S} - 1\right) \tag{3.3.27}$$

Equations 3.3.24 through 3.3.27 are also valid for concentric cylinders, concentric spheres, and similar configurations, as long as $r_N - r_1 \ll r_1$. Also, the relations are readily extended to shields with nonidentical emittance.

While the network analogy can (and has been) applied to configurations with more than two surfaces seeing each other, this leads to very complicated circuits (since there is one resistance between any two surfaces). For such problems, the network analogy is not recommended, and the net radiation method, Equation 3.3.18, should be employed.

Radiative Exchange within Participating Media

In many high-temperature applications, when radiative heat transfer is important, the medium between surfaces is not transparent, but is "participating, " that is, it absorbs, emits, and (possibly) scatters radiation. In a typical combustion process this interaction results in (1) continuum radiation due to tiny, burning soot particles (of dimension $< 1\,\mu\text{m}$) and also due to larger suspended particles, such as coal particles, oil droplets, fly ash; (2) banded radiation in the infrared due to emission and absorption by molecular gaseous combustion products, mostly water vapor and carbon dioxide; and (3) chemiluminescence due to the combustion reaction itself. While chemiluminescence may normally be neglected, particulates as well as gas radiation generally must be accounted for.

Radiative Properties of Molecular Gases

When a photon (or an electromagnetic wave) interacts with a gas molecule, it may be absorbed, raising the energy level of the molecule. Conversely, a gas molecule may spontaneously lower its energy level by the emission of an appropriate photon. This leads to large numbers of narrow spectral lines, which partially overlap and together form so-called vibration–rotation bands. As such, gases tend to be transparent over most of the spectrum, but may be almost opaque over the spectral range of a band. The *absorption coefficient* κ_λ is defined as a measure of how strongly radiation is absorbed or emitted along a path of length, L, leading to the spectral absorptivity and emissivity for this path, or

$$\alpha_\lambda = \varepsilon_\lambda = 1 - e^{-\kappa_\lambda L} \tag{3.3.28}$$

This exponential decay of incident radiation is often called *Beer's law*. Although gases are distinctly nongray, for simple heat-transfer calculations it is often sufficient to determine the total emissivity for an isothermal path [compare Equation 3.3.9]

$$\varepsilon = \frac{1}{E_b} \int_0^\infty (1 - e^{-\kappa_\lambda L}) E_{b\lambda}(T_g)\, d\lambda \tag{3.3.29}$$

For a mixture of gases the total emissivity is a function of path length L, gas temperature T_g, partial pressure (s) of the absorbing gas(es) p_a, and total pressure p. For the—in combustion applications

most important—mixture of nitrogen with water vapor and/or carbon dioxide, the total emissivity may be calculated from Leckner (1972). First, the individual emissivities for water vapor and carbon dioxide, respectively, are calculated separately from

$$\varepsilon(p_a L, p, T_g) = \varepsilon_0(p_a L, T_g)\left(\frac{\varepsilon}{\varepsilon_0}\right)(p_a L, p, T_g) \tag{3.3.30a}$$

$$\left(\frac{\varepsilon}{\varepsilon_0}\right)(p_a L, p, T_g) = \left[1 - \frac{(a-1)(1-P_E)}{a+b-1+P_E}\exp\left(-c\left[\log_{10}\frac{(p_a L)_m}{p_a L}\right]^2\right)\right] \tag{3.3.30b}$$

$$\varepsilon_0(p_a L, T_g) = \exp\left[\sum_{i=0}^{N}\sum_{j=0}^{N} c_{ji}\left(\frac{T_g}{T_0}\right)^j\left(\log_{10}\frac{p_a L}{(p_a L)_0}\right)^i\right] \tag{3.3.30c}$$

Here ε_0 is the total emissivity of a reference state, that is, for the case of $p = 1$ bar and $p_a \to 0$ (but $p_a L > 0$), and the correlation constants a, b, c, c_{ji}, P_E, $(p_a L)_0$, $(p_a L)_m$, and $(p_a L)_m$ are given in Table 3.3.2 for water vapor and carbon dioxide. (For convenience, plots of ε_0 are given in Figures 3.3.15 for CO_2 and 3.3.16 for H_2O.) The total emissivity of a mixture of nitrogen with both water vapor and carbon dioxide is calculated from

$$\varepsilon_{CO_2+H_2O} = \varepsilon_{CO_2} + \varepsilon_{H_2O} - \Delta\varepsilon \tag{3.3.31}$$

$$\Delta\varepsilon = \left(\frac{\zeta}{10.7+101\zeta} - 0.0089\zeta^{10.4}\right)\left(\log_{10}\frac{(p_{H_2O}+p_{CO_2})L}{(p_a L)_0}\right)^{2.76} \tag{3.3.32a}$$

TABLE 3.3.2
Correlation Constants for the Determination of the Total Emissivity for Water Vapor and Carbon Dioxide

Gas	Water Vapor			Carbon Dioxide			
M, N	2, 2			2, 3			
c_{00} ⋯ c_{N0}	–2.2118	–1.1987	0.035596	–3.9893	2.7669	–2.1081	0.39163
⋮ ⋱ ⋮	–0.10838	–0.17156	0.045915	–0.23678	0.19731	–0.19544	0.044644
c_{0M} ⋯ c_{NM}	–0.85667	–0.93048	–0.14391	1.2710	–1.1090	1.0195	–0.21897
P_E	$(p+2.56p_a/\sqrt{t})/p_0$			$(p+0.28p_a)/p_0$			
$(p_a L)_m/(p_a L)_0$	$13.2t^2$			$0.054/t^2$, $t<0.7$ $0.225t^2$, $t>0.7$			
a	2.144, $t<0.75$ $1.88-2.053\log_{10}t$, $t>0.75$			$1+0.1/t^{1.45}$			
b	$1.10/t^{1.4}$			0.23			
c	0.5			1.47			

Note: $T_0 = 1000\,\text{K}, p_0 = 1\,\text{bar}, t = T/T_0, (p_a L)_0 = 1\,\text{bar cm}.$

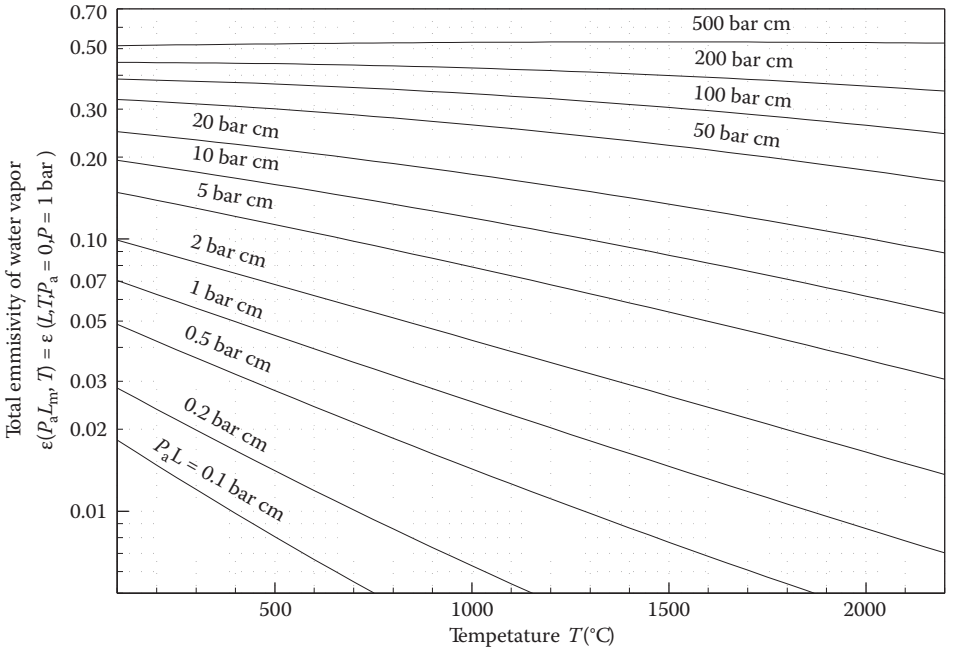


FIGURE 3.3.15 Total emissivity of water vapor at reference state (total gas pressure $p = 1$ bar, partial pressure of H_2O $p_a \to 0$).

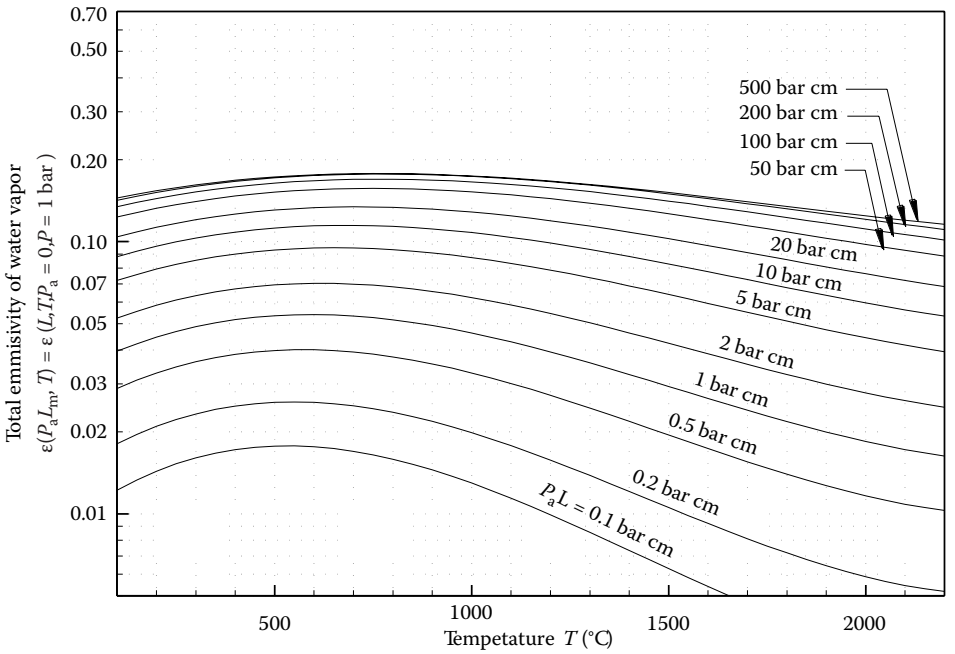


FIGURE 3.3.16 Total emissivity of carbon dioxide at reference state (total gas pressure $p = 1$ bar, partial pressure of CO_2 $p_a \to 0$).

$$\zeta = \frac{p_{H_2O}}{p_{H_2O} + p_{CO_2}} \tag{3.3.32b}$$

where the $\Delta\varepsilon$ compensates for overlap effects between H_2O and CO_2 bands, and the ε_{CO_2} and ε_{H_2O} are calculated from Equation 3.3.30.

If radiation emitted externally to the gas (e.g., by emission from an adjacent wall at temperature T_s) travels through the gas, the total amount absorbed by the gas is of interest. This leads to the absorptivity of a gas path at T_g with a source at T_s:

$$\alpha(p_a L, p, T_g, T_s) = \frac{1}{E_b(T_s)} \int_0^\infty \left(1 - e^{-\kappa_\lambda(T_g)L}\right) E_{b\lambda}(T_s)\, d\lambda \tag{3.3.33}$$

which for water vapor or carbon dioxide may be estimated from

$$\alpha(p_a L, p, T_g, T_s) = \left(\frac{T_g}{T_s}\right)^{1/2} \varepsilon\left(p_a L \frac{T_s}{T_g}, p, T_s\right) \tag{3.3.34}$$

where ε is the emissivity calculated from Equation 3.3.30 evaluated at the temperature of the surface T_s, and using an adjusted pressure path length, $p_a L T_s / T_g$. For mixtures of water vapor and carbon dioxide, band overlap is again accounted for by taking

$$\alpha_{CO_2+H_2O} = \alpha_{CO_2} + \alpha_{H_2O} - \Delta\varepsilon \tag{3.3.35}$$

with $\Delta\varepsilon$ evaluated for a pressure path length of $p_a L T_s / T_g$.

Example 3.3.6

Consider a layer of a gas mixture at 1000 K and 5 bar that consists of 10% carbon dioxide and 70% nitrogen. What is its emissivity for a path length of 1.76 m, and its absorptivity (for the same path) if the layer is irradiated by a source at 1500 K?

Solution. First, we calculate the total emissivity of the CO_2 at the reference state ($p = 1$ bar, $p_a \to 0$) for a length of 1.76 m from Equation 3.3.30c or Figure 3.3.15. With

$$T_g = 1000\,\text{K} = 727^\circ\text{C} \text{ and } p_a L = 0.1 \times 5\,\text{bar} \times 1.76\,\text{m} = 88\,\text{bar cm}$$

one gets, interpolating Figure 3.3.15, $\varepsilon_0 \cong 0.15$. The correction factor in Equation 3.3.30b is calculated from Table 3.3.2 with $P_E = 5 + 0.28 \times 0.5 = 5.14$, $a = 1.1$, $b = 0.23$, $c = 1.47$, and $(p_a L)_m = 0.225$ bar cm. Thus,

$$\frac{\varepsilon}{\varepsilon_0} = 1 - \frac{0.1 \times (-4.14)}{0.33 + 5.14} \exp\left(-1.47\left(\log_{10}\frac{0.225}{88}\right)^2\right) \cong 1$$

and

$$\varepsilon \cong 0.15$$

To calculate the absorptivity ε_0 must be found for a temperature of

$$T_s = 1500\,\text{K} = 1227^\circ\text{C} \text{ and } p_a L \frac{T_s}{T_g} = 88 \times 1500 / 1000 = 132\,\text{bar cm}$$

From Figure 3.3.15, it follows that $\varepsilon_0 \cong 0.15$ again and, with $\varepsilon / \varepsilon_0$ pretty much unchanged, from Equation 3.3.34,

$$\alpha \cong \left(\frac{1000}{1500}\right)^{1/2} \times 0.15 \times 1.00 = 0.122$$

Radiative Properties of Particle Clouds

Nearly all flames are visible to the human eye and are, therefore, called *luminous* (sending out light). Apparently, there is some radiative emission from within the flame at wavelengths where there are no vibration–rotation bands for any combustion gases. This luminous emission is known today to come from tiny char (almost pure carbon) particles, called soot, which are generated during the combustion process. The "dirtier" the flame is (i.e., the higher the soot content), the more luminous it is.

Radiative properties of soot. Soot particles are produced in fuel-rich flames, or fuel-rich parts of flames, as a result of incomplete combustion of hydrocarbon fuels. As shown by electron microscopy, soot particles are generally small and spherical, ranging in size between approximately 50 and 800 Å (0.005 – 0.08 μm), and up to about 3000 Å in extreme cases. While mostly spherical in shape, soot particles may also appear in agglomerated chunks and even as long agglomerated filaments. It has been determined experimentally in typical diffusion flames of hydrocarbon fuels that the volume percentage of soot generally lies in the range between $10^{-4}\%$ to $10^{-6}\%$

Since soot particles are very small, they are generally at the same temperature as the flame and strongly emit thermal radiation in a continuous spectrum over the infrared region. Experiments have shown that soot emission often is considerably stronger than the emission from the combustion gases.

For a simplified heat-transfer analysis, it is desirable to use suitably defined mean absorption coefficients and emissivities. If the soot volume fraction, f_v, is known as well as an appropriate spectral average of the complex index of refraction of the soot, $m = n - ik$, one may approximate the spectral absorption coefficient by (Felske and Tien, 1977) using

$$\kappa_\lambda = C_0 \frac{f_v}{\lambda} \quad C_0 = \frac{36\pi nk}{\left(n^2 - k^2 + 2\right)^2 + 4n^2k^2} = 6\pi E(m) \tag{3.3.36}$$

where the "refractive index function" $E(m)$ is commonly used in the soot literature. A total, or spectral average absorption coefficient may be taken as

$$\kappa_m = 3.72 f_v C_0 T / C_2 \tag{3.3.37}$$

where $C_2 = 1.4388$ mK is the second Planck function constant. Substituting Equation 3.3.37 into Equation 3.3.29 gives a total soot cloud emissivity of

$$\varepsilon(f_v TL) = 1 - e^{-\kappa_m L} = 1 - e^{-3.72 C_0 f_v TL / C_2} \tag{3.3.38}$$

Pulverized coal and fly ash dispersions. To calculate the radiative properties of arbitrary size distributions of coal and ash particles, one must have knowledge of their complex index of refraction as a function of wavelength and temperature. Data for carbon and different types of coal indicate that its real part, n, varies little over the infrared and is relatively insensitive to the type of coal (e.g., anthracite, lignite, bituminous), while the absorptive index, k, may vary strongly over the spectrum and from coal to coal. If the number and sizes of particles are known and if a suitable average value for the complex index of refraction can be found, then the spectral absorption coefficient of the

dispersion may be estimated by a correlation given by Buckius and Hwang (1980). Substitution into Equation 3.3.29 can then provide an estimate of the total emissivity. If both soot as well as larger particles are present in the dispersion, the absorption coefficients of all constituents must be added before applying Equation 3.3.29.

Mixtures of molecular gases and particulates. To determine the total emissivity of a mixture, it is generally necessary to find the spectral absorption coefficient κ_λ of the mixture (i.e., the sum of the absorption coefficient of all contributors), followed by numerical integration of Equation 3.3.29. However, since the molecular gases tend to absorb only over a small part of the spectrum, to some degree of accuracy

$$\varepsilon_{\text{mix}} \cong \varepsilon_{\text{gas}} + \varepsilon_{\text{particulates}} \tag{3.3.39}$$

Equation 3.3.39 gives an upper estimate since overlap effects result in lower emissivity [compare Equation 3.3.31 for gas mixtures].

Heat Exchange in the Presence of a Participating Medium

The calculation of radiative heat-transfer rates through an enclosure filled with a participating medium is a challenging task, to say the least. High-accuracy calculations are rare and a topic of ongoing research. There are, however, several simplistic models available that allow the estimation of radiative heat-transfer rates, and relatively accurate calculations for some simple cases.

Diffusion approximation. A medium through which a photon can only travel a short distance without being absorbed is known as *optically thick*. Mathematically, this implies that $\kappa_\lambda L \gg 1$ for a characteristic dimension L, across which the temperature does not vary substantially. For such an optically thick, nonscattering medium the spectral radiative flux may be calculated from

$$\mathbf{q} = -\frac{4}{3\kappa_\lambda} \nabla E_{\text{b}\lambda} \tag{3.3.40}$$

similar to Fourier's diffusion law for heat conduction. Note that a medium may be optically thick at some wavelengths, but thin ($\kappa_\lambda L \ll 1$) at others (e.g., molecular gases!). For a medium that is optically thick for all wavelengths, Equation 3.3.40 may be integrated over the spectrum, yielding the total radiative flux

$$\mathbf{q} = -\frac{4}{3\kappa_{\text{R}}} \nabla E_{\text{b}} = -\frac{4}{3\kappa_{\text{R}}} \nabla(\sigma T^4) = -\frac{16\sigma T^3}{3\kappa_{\text{R}}} \nabla T \tag{3.3.41}$$

where κ_{R} is the suitably averaged absorption coefficient, termed the *Rosseland-mean absorption coefficient*. For a cloud of soot particles, $\kappa_{\text{R}} \cong \kappa_m$ from Equation 3.3.37 is a reasonable approximation. Equation 3.3.41 may be rewritten by defining a "radiative conductivity" k_{R},

$$\mathbf{q} = -k_{\text{R}} \nabla T \qquad k_{\text{R}} = \frac{16\sigma T^3}{3\kappa_{\text{R}}} \tag{3.3.42}$$

This form shows that the diffusion approximation is mathematically equivalent to conductive heat transfer with a (strongly) temperature-dependent conductivity.

Note: More accurate calculations show that, in the absence of other modes of heat transfer (conduction, convection), there is generally a temperature discontinuity near the boundaries ($T_{\text{surface}} \neq T_{\text{adjacentmedium}}$), and, unless boundary conditions that allow such temperature discontinuities are chosen, the diffusion approximation will do very poorly in the vicinity of bounding surfaces.

Example 3.3.7

A soot cloud is contained between two walls at $T_1 = 1000$ K and $T_2 = 2000$ K, spaced 1 m apart. The effective absorption coefficient of the medium is $\kappa_R = 10\ \text{m}^{-1}$ and the effective thermal conductivity is $k_c = 0.1$W/mK. Estimate the total heat flux between the plates (ignoring convection effects).

Solution. For simplicity, we may want to assume a constant total conductivity $k = k_c + k_R$, leading to

$$q = -k\frac{\mathrm{d}T}{\mathrm{d}x} = k\frac{T_2 - T_1}{L}$$

where k_R must be evaluated at some effective temperature. Choosing, based on its temperature dependence,

$$k_R \cong \frac{8\sigma}{3\kappa_R}(T_1^3 + T_2^3) = \frac{8\times 5.670\times 10^{-8}\ \text{W/m}^2\text{K}^4}{3\times 10/\text{m}}(1000^3 + 2000^3)\text{K}^3 = 136\,\text{W/mK}$$

gives

$$q = (0.1 + 136)\frac{2000 - 1000}{1}\,\text{W/m}^2 = 136\,\text{kW/m}^2\text{K}$$

Note that (1) conduction is negligible in this example and (2) the surface emittances do not enter the diffusion approximation. While a more accurate answer can be obtained by taking the temperature dependence of k_R into account, the method itself should be understood as a relatively crude approximation.

Mean beam length method. Relatively accurate yet simple heat-transfer calculations can be carried out if an isothermal, absorbing–emitting, but not scattering medium is contained in an isothermal, black-walled enclosure. While these conditions are, of course, very restrictive, they are met to some degree by conditions inside furnaces. For such cases, the local heat flux on a point of the surface may be calculated from

$$q = \left[1 - \alpha(L_m)\right]E_{bw} - \varepsilon(L_m)E_{bg} \tag{3.3.43}$$

where E_{bw} and E_{bg} are blackbody emissive powers for the walls and medium (gas and/or particulates), respectively, and $\alpha(L_m)$ and $\varepsilon(L_m)$ are the total absorptivity and emissivity of the medium for a path length L_m through the medium. The length L_m, known as the average mean beam length, is a directional average of the thickness of the medium as seen from the point on the surface. On a spectral basis, Equation 3.3.43 is exact, provided the previous conditions are met and provided an accurate value of the (spectral) mean beam length is known. It has been shown that spectral dependence of the mean beam length is weak (generally less than ±5% from the mean). Consequently, total radiative heat flux at the surface may be calculated very accurately from Equation 3.3.43, provided the emissivity and absorptivity of the medium are also known accurately. The mean beam lengths for many important geometries have been calculated and are collected in Table 3.3.3. In this table, L_0 is known as the geometric mean beam length, which is the mean beam length for the optically thin limit ($\kappa_\lambda \to 0$), and L_m is a spectral average of the mean beam length. For geometries not listed in Table 3.3.3, the mean beam length may be estimated from

$$L_0 \cong 4\frac{V}{A} \quad L_m \cong 0.9L_0 \cong 3.6\frac{V}{A} \tag{3.3.44}$$

where V is the volume of the participating medium and A is its entire bounding surface area.

TABLE 3.3.3
Mean Beam Lengths for Radiation from a Gas Volume to a Surface on Its Boundary

Geometry of Gas Volume	Characterizing Dimension L	Geometric Mean Beam Length L_0/L	Average Mean Beam Length L_m/L	L_m/L_0
Sphere radiating to its surface	Diameter, $L=D$	0.67	0.65	0.97
Infinite circular cylinder to bounding surface	Diameter, $L=D$	1.00	0.94	0.94
Semi-infinite circular cylinder to:	Diameter, $L=D$			
Element at center of base		1.00	0.90	0.90
Entire base		0.81	0.65	0.80
Circular cylinder (height/diameter=1) to:	Diameter, $L=D$			
Element at center of base		0.76	0.71	0.92
Entire surface		0.67	0.60	0.90
Circular cylinder (height/diameter=2) to:	Diameter, $L=D$			
Plane base		0.73	0.60	0.82
Concave surface		0.82	0.76	0.93
Entire surface		0.80	0.73	0.91
Circular cylinder (height/diameter=0.5) to:	Diameter, $L=D$			
Plane base		0.48	0.43	0.90
Concave surface		0.53	0.46	0.88
Entire surface		0.50	0.45	0.90
Infinite semicircular cylinder to center of plane rectangular face	Radius, $L=R$	—	1.26	—
Infinite slab to its surface	Slab thickness, L	2.00	1.76	0.88
Cube to a face	Edge L	0.67	0.6	0.90
Rectangular 1×1×4 parallelepipeds:	Shortest edge, L			
To 1×4 face		0.90	0.82	0.91
To 1×1 face		0.86	0.71	0.83
To all faces		0.89	0.81	0.91

Example 3.3.8

An isothermal mixture of 10% CO_2 and 90% nitrogen at 1000 K and 5 bar is contained between two large, parallel, black plates, which are both isothermal at 1500 K. Estimate the net radiative heat loss from the surfaces.

Solution. The heat loss may be calculated from Equation 3.3.43, after determining the mean beam length, followed by evaluation of $\varepsilon(L_m)$ and $\alpha(L_m)$. From Table 3.3.3, it is clear that $L_m = 1.76 \times$ thickness of slab = 1.76 m. It turns out that the necessary $\varepsilon(L_m) = 0.15$ and $\alpha(L_m) = 0.122$ have already been calculated in Example 3.3.6. Thus, the heat flux is immediately calculated from Equation 3.3.43 as

$$q = (1-0.122)5.670\times10^{-8}\times1500^4 - 0.15\times5.670\times10^{-8}\times1000^4$$

$$= 2.44\times10^5\ \text{W/m}^2 = 244\ \text{kW/m}^2$$

Advanced Concepts

Radiative transfer in a participating medium is governed by the so-called radiative transfer equation (RTE), an energy balance for spectral radiative intensity I_λ traveling in a specific direction $\hat{s}$,

$$\frac{dI_\lambda}{ds} = \kappa_\lambda I_{b\lambda} - \beta_\lambda I_\lambda(\hat{s}) + \frac{\sigma_{s\lambda}}{4\pi}\int_{4\pi} I_\lambda(\hat{s})\Phi_\lambda(\hat{s},\hat{s}')\,d\Omega', \tag{3.3.45}$$

in which $I_{b\lambda} = E_{b\lambda}/\pi$ is the blackbody intensity, κ_λ is the *absorption coefficient* [the ability of a medium to absorb and emit radiation, as given in Equation 3.3.28], $\sigma_{s\lambda}$ is the *scattering coefficient* (the ability of a medium to scatter radiation), $\beta_\lambda = \kappa_\lambda + \sigma_{s\lambda}$ is called the *extinction coefficient*, and $\Phi_\lambda(\hat{s},\hat{s}')$ the scattering phase function (the probability that radiation traveling in the direction of $\hat{s}$ is scattered, i.e., redirected, into the direction of $\hat{s}'$). The RTE states that the strength of a light beam traveling along a path s is augmented by emission from the medium (first term on the RHS), is diminished by absorption and outscattering, that is, scattering *away* from $\hat{s}$ (second term), and increased by inscattering, that is, scattering from all other directions *into* direction $\hat{s}$ (last term). The RTE is a five-dimensional (three space coordinates, and two direction coordinates with local origins) integro-differential equation, which is extremely difficult to solve. The problem is compounded by the complicated spectral behavior usually exhibited by the medium's radiation properties (κ_λ, $\sigma_{s\lambda}$ and Φ_λ). Therefore, numerous approximate and/or numerical techniques have been developed to solve the RTE and to deal with the complicated spectral behavior.

RTE solution methods. All RTE solution methods attempt to either approximate the directional dependence of radiative intensity or express it numerically. The next level of sophistication beyond the diffusion and mean-beam-length approximations are the so-called flux methods. The two-flux model, assuming the intensity to have two constant but different values across two hemispheres of directions, has been successfully applied to one-dimensional problems (Modest, 2013); the six-flux method can be applied to more general geometries, but must be tailored toward each individual problem. The most popular numerical methods are the *spherical harmonics* and the *discrete ordinates* methods and, more recently, the statistical (photon) *Monte Carlo* method. In the first, the directional dependence of intensity is expanded into a series of spherical harmonics, truncated at the *N*th level and resulting in the so-called P-N methods. While higher-order P-N methods are difficult to implement, its lowest order, the *P*–1 approximation, is very popular because of its simplicity and surprising accuracy. Reducing the RTE into a single elliptical partial differential equation, it is easy to implement and has been incorporated into most modern commercial CFD codes. The *P*–1 approximation assumes a near-isotropic intensity distribution and is, therefore, accurate in emission-dominated applications (such as combustion of fuels and high-temperature combustion products) and optically thick systems (large κ_λ); the method fails for predicting heat exchange between hot and cold surfaces separated by an optically thin medium (small κ_λ). The *discrete ordinates* method (DOM) finite differences the directional dependence of intensity, requiring the simultaneous solution of many hyperbolic partial differential equations (Modest, 2013). Because it can be carried to arbitrary levels of accuracy (albeit at a steep computational price) it, and its modern finite-volume implementation (FVM) (Modest, 2013) are also incorporated into most CFD codes. In the photon Monte Carlo (PMC) method, statistically meaningful photon bundles are traced from their point of emission until they are absorbed or leave the computational domain (Modest, 2013). Unlike conventional RTE solvers, the PMC method is easily applied to problems of great complexity (in geometry and/or radiative properties), and is the only method that can produce "exact" results. Because of the independent sampling of photons, the method is also ideal for parallel computing. Their disadvantages are that Monte Carlo methods require the tracing of millions of photon bundles to obtain accurate results, that is, they require vast computer resources, and furthermore, like those from any statistical method, their results display statistical scatter, which can make it difficult to combine them with conventional flow solvers. Some commercial codes also offer the older *discrete transfer* method, which is today known to be less accurate and less efficient (Modest, 2013), and is, thus, not recommended.

Spectral Models. As indicated earlier, the spectral variations of radiation properties of gases and particles are extremely difficult to predict and to deal with. While particles tend to have a distribution of particle sizes and shapes, thus smoothing out spectral variations and allowing the assumption of gray

or near-gray behavior, this is almost never the case for molecular gases (but may result in errors of hundreds, and even thousands, of percent). This has prompted the development of a number of approximate models, which may be loosely grouped into four groups: (1) line-by-line calculations, (2) narrow band models, (3) wide band models, and (4) global models. In line-by-line calculations (LBL), the spectral variation of the gas absorption coefficient is followed in small steps across the entire spectrum, requiring 1,000,000 RTE evaluations or more, making this method suitable only for bench marking. Narrow band models take advantage of the fact that the gas absorption coefficient varies much more rapidly across the spectrum than other quantities, such as emissive power or surface emittance. The models predict gas transmissivities averaged across small spectral ranges, the most successful ones being the so-called statistical narrow band models (SNB), in particular the Malkmus model (Modest, 2013). Wide band models (WBM) are similar to narrow band models, but average over an entire vibration–rotation band and usually employ correlations of somewhat dated experimental data. Finally, global models look at the entire spectrum, the most successful one being the *weighted sum of gray gases* (WSGG) first introduced by Hottel and Sarofim (1967). The most modern, accurate and efficient approach is the method of *full-spectrum correlated k-distributions* (FSCK), in which the absorption coefficient is reordered across the spectrum into a monotonically increasing function. This method also exists in narrow band and wide band forms, and can result in (near-) exact answers at a tiny fraction of the costs for LBL calculations. If the statistical photon Monte Carlo method is employed, spectral variations of radiative properties are readily incorporated into the sampling and LBL-accurate PMC solutions can be obtained. While for conventional RTE solvers LBL accuracy increases computational effort by 6 orders of magnitude over gray media, for PMC calculations the increase in effort is only a few tens of percent. Therefore, if accurate spectral modeling is required, LBL PMC solvers are computationally competitive with the more traditional *P-N* and DOM methods. For more detailed information on spectral modeling, the reader is referred to Modest (2013). Apparently, as of 2015 none of the better known CFD packages includes a credible nongray spectral model.

Defining Terms

Absorptance: The ability of a solid material to absorb (i.e., trap and convert to other forms of energy) incoming radiation; gives the fraction of incoming radiation that is absorbed by the medium.

Absorptivity: The fraction of incoming radiation that is absorbed by a column of a radiatively participating medium (e.g., a gas or a particle cloud).

Absorption coefficient: The ability of a medium to absorb (i.e., trap and convert to other forms of energy) radiative energy over a unit path length; the reciprocal of the mean distance a photon travels before being absorbed.

Blackbody: Any material or configuration that absorbs all incoming radiation completely. A blackbody also emits the maximum possible amount of radiation as described by Planck's law.

Diffuse surface: A surface that emits and/or reflects equal amounts of radiative energy (photons) into all directions. Or a surface that absorbs and/or reflects equal amounts of radiation independent of incoming direction.

Emissive power: The rate of radiative energy leaving a surface through emission. The maximum amount of emissive power is emitted by a blackbody with a spectral strength described by Planck's law.

Emittance: The ability of a solid material to emit (i.e., convert internal energy into electromagnetic waves or photons) thermal radiation; gives the fraction of emission as compared with a blackbody.

Emissivity: The ability of a column of a radiatively participating medium (e.g., a gas or a particle cloud) to emit thermal radiation; gives the fraction of emission as compared with an opaque (i.e., infinitely long) column.

Gray: A medium whose radiative properties (such as absorptance, emittance, reflectance, absorption coefficient) do not vary with wavelength.

Irradiation: Incoming radiative flux onto a surface from outside it.

Network analogy: Expressing radiative heat exchange between surfaces in terms of an electrical network, with heat flux as "current," differences in emissive power as "potentials," and defining radiative resistances.

Opaque medium: A medium of sufficient thickness that absorbs all nonreflected irradiation; no radiation is transmitted through the medium.

Photon: A massless particle carrying energy in the amount of $h\nu$; the quantum mechanical alternative view of an electromagnetic wave carrying radiative energy.

Planck's law: The law describing the spectral distribution of the radiative energy emitted (emissive power) of a blackbody.

Radiosity: Total radiative flux leaving a surface (diffusely), consisting of emitted as well as reflected radiation.

Reflectance: The ability of an interface, or of a medium or of a composite with a number of interfaces, to reflect incoming radiation back into the irradiating medium.

Semitransparent: See transparent.

Spectral value: The value of a quantity that varies with wavelength at a given wavelength; for dimensional quantities the amount per unit wavelength.

Transmittance: The fraction of incoming radiation that is transmitted through a layer of a solid material.

Transmissivity: The fraction of incoming radiation that is transmitted by a column of a radiatively participating medium (e.g., a gas or a particle cloud).

Transparent: The ability of a medium to let incoming radiation pass through it. A medium that lets all radiation pass through it is called transparent, a medium that only allows a part to pass through it is called semitransparent.

View factor: The fraction of diffuse radiant energy leaving one surface that is intercepted by another surface.

REFERENCES

Brewster, M.Q. 1992. *Thermal Radiative Transfer & Properties*, John Wiley & Sons, New York.

Buckius, R.O. and Hwang, D.C. 1980. Radiation properties for polydispersions: application to coal, *J. Heat Transfer*, 102, 99–103.

Felske, J.D. and Tien, C.L. 1977. The use of the Milne-Eddington absorption coefficient for radiative heat transfer in combustion systems, *J. Heat Transfer*, 99(3), 458–465.

Hottel, H.C. and Sarofim, A.F. 1967. *Radiation Transfer*, McGraw-Hill, New York.

Howell, J.R. 1982. *Catalog of Radiation Configuration Factors*, McGraw-Hill, New York. http://www.thermalradiation.net/indexCat.html.

Howell, J.R., Mengüç, P. and Siegel, R. 2015. *Thermal Radiation Heat Transfer*, 4th ed., Taylor & Francis, New York.

Leckner, B. 1972. Spectral and total emissivity of water vapor and carbon dioxide, *Combust. Flame*, 19, 33–48.

Modest, M.F. 2013. *Radiative Heat Transfer*, 3rd ed., Academic Press, New York.

Özisik, M.N. 1973. *Radiative Transfer and Interactions with Conduction and Convection*, John Wiley & Sons, New York.

Sparrow, E.M. and Cess, R.D. 1978. *Radiation Heat Transfer*, Hemisphere, New York.

3.4 PHASE-CHANGE

3.4.1 Boiling and Condensation

Van P. Carey

Introduction

Liquid–vapor phase-change processes play an important role in many technological applications. The virtually isothermal heat transfer associated with boiling and condensation processes makes their inclusion in power and refrigeration processes highly advantageous from a thermodynamic efficiency standpoint. In addition, the high heat-transfer coefficients associated with boiling and condensation have made the use of these processes increasingly attractive in the thermal control of compact devices that have high heat dissipation rates. Applications of this type include the use of boiling heat transfer to cool electronic components in computers and the use of compact evaporators and condensers for thermal control of aircraft avionics and spacecraft environments. Liquid–vapor phase-change processes are also of critical importance to nuclear power plant design, both because they are important in normal operating circumstances, and because they dominate many of the accident scenarios that are studied as part of design evaluation.

The heat transfer and fluid flow associated with liquid–vapor phase-change processes are typically among the more complex transport circumstances encountered in engineering applications. These processes have all the complexity of single-phase convective transport, plus additional elements resulting from motion of the interface, nonequilibrium effects, and dynamic interactions between the phases. Due to the highly complex nature of these processes, development of methods to accurately predict the associated heat and mass transfer is often a formidable task.

In this section, commonly used variables not defined in the nomenclature are as follows: q'' = surface heat flux, μ_l = liquid viscosity, μ_v = vapor viscosity, Pr_l = liquid Prandtl number, T_w = wall surface temperature, T_{sat} = saturation temperature, c_{pl} = liquid specific heat, k_v = vapor thermal conductivity, g = gravitational acceleration, G = mass flow per unit of cross-sectional area, h_{lv} = latent heat of vaporization, and x = mass quality.

Thermophysics of Boiling Processes

Four thermophysical mechanisms that play important roles in boiling processes are (1) surface tension effects, (2) surface wetting characteristics of the liquid, (3) metastable phase stability, and (4) surface morphology effects.

Anyone who has watched small bubbles rise in a carbonated beverage or a pot of boiling water has undoubtedly noticed that the bubbles are almost perfectly spherical, as if an elastic membrane were present at the interface to pull the vapor into a spherical shape. This apparent interfacial tension or *surface tension* σ is equivalent to an energy stored in the interface region per unit area. The excess energy in this region is due to the slightly larger separation of the liquid phase molecules adjacent to the gas phase.

The magnitude of the surface tension for a substance is directly linked to the strength of intermolecular forces in the material. Nonpolar liquids typically have the lowest surface tension. Water and other *polar molecules* have somewhat higher surface tension, and liquid metals, which exhibit metallic bond attraction, have very high surface tension. The surface tension of water at 20°C is 0.0728 N/m, whereas liquid mercury has a surface tension of 0.484 N/m at the same temperature. The surface tension for any pure liquid varies with temperature. It decreases almost linearly with increasing temperature, vanishing altogether at the critical point where the distinction between the phases disappears.

As a result of the surface tension at the interface, the pressure inside a spherical bubble of radius r must exceed that in the surrounding liquid by $2\sigma/r$:

$$P_v = P_l + \frac{2\sigma}{r}. \tag{3.4.1}$$

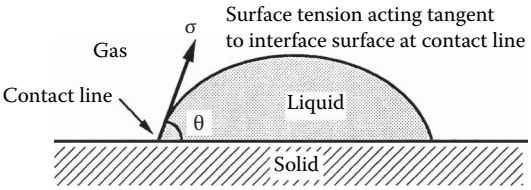


FIGURE 3.4.1 Definition of the contact angle θ.

By using the relation 3.4.1 between the pressures in the two phases, it can be shown that for the bubble to be in equilibrium with the surrounding liquid, the liquid must actually be superheated above the saturation temperature for the ambient liquid pressure. The amount of required superheating increases as the radius of curvature of the bubble interface decreases.

The wetting characteristics of the liquid are generally quantified in terms of a *contact angle* between the solid surface and the tangent to the interface at the point where it contacts the solid. This angle is measured through the liquid phase, as shown in Figure 3.4.1. In some systems, the wetting angle established at equilibrium may depend on the fluid motion history. For some systems, the contact angle established by liquid advancing over a solid surface is larger than that established when a liquid front recedes over the surface. This behavior is referred to as *contact angle hysteresis*. Contact angle hysteresis can have an important effect on boiling and condensation processes, particularly those involving water.

For a bubble with a specified vapor volume, the contact angle will dictate the radius of curvature of the bubble interface. The wetting behavior in combination with the surface tension effect thus determines the level of superheat required for the bubble to be in equilibrium with the surrounding liquid. The liquid must be heated above this superheat level for the bubble to grow. A steady boiling process can be sustained only if the liquid is heated above this threshold superheat level.

It can be shown from basic thermodynamic analysis that a necessary and sufficient condition for phase stability is that

$$\left(\frac{\partial P}{\partial v}\right)_T < 0, \tag{3.4.2}$$

where v is the specific volume. Below the critical temperature, extrapolation of the isotherms for the liquid and vapor phases consistent with an equation of state like the van der Waals equation results in an isotherm shape similar to that shown in Figure 3.4.2.

The locus of points where $(\partial P/\partial v)_T = 0$ are termed *spinodal curves*. Regions of metastable vapor and liquid exist between the saturation curve and the spinodal curves. The effects of surface tension discussed earlier require that fluid surrounding a vapor bubble be in the metastable superheated liquid region. Predictions of statistical thermodynamics imply that as $(\partial P/\partial v)_T$ approaches zero, the level of fluctuations in a fluid system increases. This, in turn, increases the probability that an embryonic new phase will form as a result of density fluctuations. Initiation of a phase change in this manner is termed *homogeneous nucleation*. Generally, a pure liquid must be heated to nearly 90% of its absolute critical temperature before homogeneous nucleation of vapor bubbles occurs.

In most physical systems of engineering interest, the bulk phase is in contact with solid walls of the containing structures or solid particulate contaminants. These solid phases may provide nucleation sites where phase change may occur if the system state is driven into the metastable range. Nucleation of vapor bubbles may preferentially occur at low-liquid superheat levels in crevices in the solid surface where gas is trapped. This type of nucleation at the solid surface of a containment wall is categorized as *heterogeneous nucleation*. Because solid containment walls usually contain microscopic crevice-type imperfections, heterogeneous nucleation is more common than homogeneous nucleation in systems where boiling occurs.

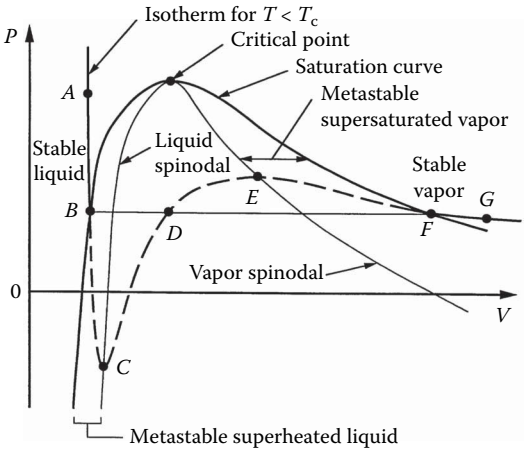


FIGURE 3.4.2 Spinodal lines and metastable regions on a *P–v* diagram.

Surface morphology affects boiling processes by facilitating entrapment in crevices of the heated walls of evaporator heat exchangers, which usually makes it easier to initiate the vaporization process. Vapor bubbles grow from these crevices until buoyancy or drag on the bubbles exceeds the surface tension force holding the bubble to the solid surface. The bubble then releases into the bulk liquid. A small remnant of vapor remains in the crevice after a bubble releases, and this remnant grows in size as further vaporization occurs until the bubble grows out of the crevice again. The result is a cyclic process of bubble growth and release known as the *ebullition cycle*. Crevices at which the ebullition cycle is sustained are said to be active nucleation sites. When the ebullition process occurs at many sites over a heated surface, the overall process is referred to as *nucleate boiling*, which is one possible mode of *pool boiling*. Note that contact angle requirements together with the surface morphology are primary factors that affect the shapes of the liquid–vapor interface and liquid regions in contact with the heated surface, which, in turn, strongly affect the heat transfer processes at the heated surface that are associated with the vaporization process.

Pool Boiling

Vaporization of liquid at the surface of a body immersed in an extensive pool of motionless liquid is generally referred to as pool boiling. The nature of the pool boiling process varies considerably depending on the conditions at which boiling occurs. The level of heat flux, the thermophysical properties of the liquid and vapor, the surface material and finish, and the physical size of the heated surface all may have an effect on the boiling process.

The regimes of pool boiling are most easily understood in terms of the so-called boiling curve: a plot of heat flux q'' vs. wall superheat $T_w - T_{sat}$ for the circumstances of interest. Many of the features of the classic pool boiling curve were determined in the early investigations of pool boiling conducted by Nukiyama (1934). Strictly speaking, the classic pool boiling curve defined by the work of this investigator and others applies to well-wetted surfaces for which the characteristic physical dimension L is large compared to the bubble or capillary length scale L_b defined as

$$L_b = \sqrt{\frac{\sigma}{g(\rho_l - \rho_v)}}. \tag{3.4.3}$$

The discussion in this section is limited to pool boiling of wetting liquids on surfaces with dimensions large compared with L_b. Additional information on features of the boiling curve when the liquid poorly wets the surface or when L/L_b is small can be found in Carey (2008). To make this

discussion concrete, we will assume that the ambient liquid surrounding the immersed body is at the saturation temperature for the ambient pressure. If the surface temperature of the immersed body is controlled and slowly increased, the boiling curve will look similar to that shown in Figure 3.4.3. The axes in this plot are logarithmic scales. The regimes of pool boiling encountered for an upward-facing horizontal flat surface as its temperature is increased are also indicated in Figure 3.4.3. The lateral extent of the surface is presumed to be much larger than L_b. At very low wall superheat levels, no nucleation sites may be active and heat may be transferred from the surface to the ambient liquid by natural convection alone and q'' increases slowly with $T_w - T_{sat}$.

Eventually, the superheat becomes large enough to initiate nucleation at some of the cavities on the surface. This *onset of nucleate boiling* (ONB) condition occurs at point *c* in Figure 3.4.3. Once nucleate boiling is initiated, any further increase in wall temperature causes the system operating point to move upward along section *d–f* of the curve in Figure 3.4.3. This portion of the curve corresponds to the nucleate boiling regime. The active sites are few and widely separated at low wall superheat levels. This range of conditions, corresponding to segment *d–e* of the curve, is sometimes referred to as the *isolated bubble regime*.

With increasing surface superheat, more sites become active, and the bubble frequency at each site generally increases. Eventually, the active sites are spaced so closely that bubbles from adjacent sites merge together during the final stages of growth and release. Vapor is being produced so rapidly that bubbles merging together form columns of vapor slugs that rise upward in the liquid pool toward its free surface. This higher range of wall superheat, corresponding to segment *e–f* of the boiling curve in Figure 3.4.3, is referred to as the *regime of slugs and columns*.

Increasing the wall superheat and heat flux within the regime of slugs and columns produces an increase in the flow rate of vapor away from the surface. Eventually, the volume fraction of vapor becomes so large, and the resulting vapor drag on the liquid moving toward the surface becomes so severe that liquid is unable to reach the surface fast enough to keep the surface completely wetted with liquid. Vapor patches accumulate at some locations, and evaporation of the liquid between the surface and some of these patches dries out portions of the surface.

If the surface temperature is held constant and uniform, dry portions of the surface covered with a vapor film will locally transfer a much lower heat flux than wetted portions of the surface

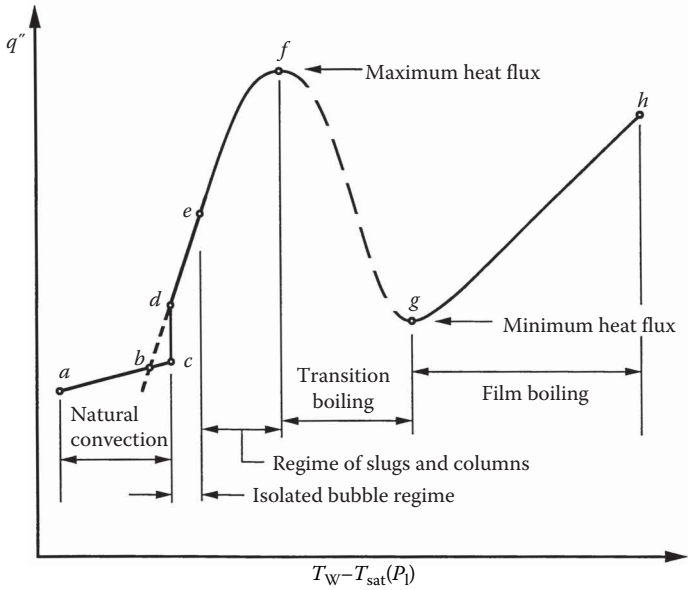


FIGURE 3.4.3 Pool boiling regimes for an independently controlled surface temperature.

where nucleate boiling is occurring. Because of the reduction in heat flux from intermittently dry portions of the surface, the mean overall heat flux from the surface is reduced. Thus, increasing the wall temperature within the slugs and columns region ultimately results in a peaking and rollover of the heat flux. The peak value of heat flux is called the *critical heat flux* (CHF), designated as point *f* in Figure 3.4.3

If the wall temperature is increased beyond the critical heat flux condition, a regime is encountered in which the mean overall heat flux decreases as the wall superheat increases. This regime, which is usually referred to as the *transition boiling* regime, corresponds to segment *f–g* on the boiling curve shown in Figure 3.4.3. The transition boiling regime is typically characterized by rapid and severe fluctuations in the local surface heat flux and/or temperature values (depending on the imposed boundary condition). These fluctuations occur because the dry regions are generally unstable, existing momentarily at a given location before collapsing and allowing the surface to be rewetted.

The vapor film generated during transition boiling can be sustained for longer intervals at higher wall temperatures. Because the intermittent insulating effect of the vapor blanketing is maintained longer, the time-averaged contributions of the blanketed locations to the overall mean heat flux are reduced. The mean heat flux from the surface thus decreases as the wall superheat is increased in the transition regime. As this trend continues, eventually a point is reached at which the surface is hot enough to sustain a stable vapor film on the surface for an indefinite period of time. The entire surface then becomes blanketed with a vapor film, thus making the transition to the *film boiling* regime. This transition occurs at point *g* in Figure 3.4.3.

Within the film boiling regime, the heat flux monotonically increases as the superheat increases. This trend is a consequence of the increased conduction and/or convection transport due to the increased driving temperature difference across the vapor film. Radiative transport across the vapor layer may also become important at higher wall temperatures.

Once a surface is heated to a superheat level in the film boiling regime, if the surface temperature is slowly decreased, in general the system will progress through each of the regimes described earlier in reverse order. However, the path of the boiling curve may differ significantly from that observed for increasing wall superheat, depending on whether the surface heat flux or temperature is controlled.

Experimental evidence summarized by Witte and Lienhard (1982) implies that the path of the transition boiling curve is determined, to a large degree, by the wetting characteristics of the liquid on the solid surface. For a given wall superheat level in the transition boiling regime, a higher heat flux is generally obtained if the liquid wets the surface than if it poorly wets the surface. For systems that exhibit contact angle hysteresis, the transition boiling curves obtained for decreasing and increasing wall superheat may therefore be somewhat different. The transition boiling curve for decreasing wall superheat may be significantly below that for increasing superheat for such circumstances, as indicated in Figure 3.4.4.

For an electrically heated surface, the rise in temperature associated with the jump from nucleate to film boiling at the critical heat flux is very often large enough to melt component materials and burn out the component. As a result, the critical heat flux is often referred to as the *burnout heat flux* to acknowledge the potentially damaging effects of applying this heat flux level to components cooled by nucleate boiling. Once the jump to film boiling has been made, any further increase in applied heat flux increases the wall superheat, and the system follows basically the same film boiling curve as in the temperature-controlled case.

Correlations of nucleate pool boiling heat-transfer data have typically been used as tools to predict nucleate boiling heat transfer in engineering systems and heat exchangers. Many investigators have proposed methods of correlating data of this type; so many, in fact, that a complete discussion of them all could easily fill a major portion of this section. In this section, three of the more commonly used correlation methods will be mentioned. However, before proceeding, two aspects of the interpretation of such correlations are worth noting. First, experimental data indicate that

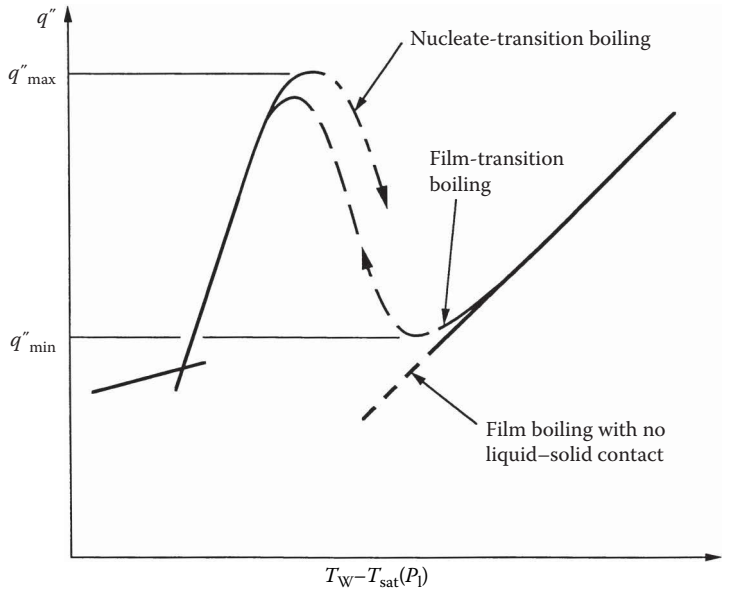


FIGURE 3.4.4 Relative locations of the nucleate transition and film transition portions of the pool boiling curve.

the subcooling of the liquid pool has a negligible effect on the nucleate boiling heat-transfer rate. Consequently, the pool boiling correlations are generally regarded as being valid for both subcooled and saturated nucleate boiling. Second, it has also been observed that at moderate to high heat flux levels, a pool boiling heat-transfer correlation developed for one heated surface geometry in one specific orientation often works reasonably well for other geometries and/or other orientations. Hence, although a correlation was developed for a specific geometry and orientation, it may often be used for other geometries, at least at moderate to high heat flux levels.

Having taken note of the earlier points, a commonly used correlation for nucleate boiling heat transfer developed by Rohsenow (1962) is

$$\frac{q''}{\mu_l h_{fg}}\left[\frac{\sigma}{g(\rho_l-\rho_v)}\right]^{1/2}=\left(\frac{1}{C_{sf}}\right)^{1/r}\mathrm{Pr}_l^{-s/r}\left[\frac{c_{pl}\left[T_w-T_{sat}(P_l)\right]}{h_{fg}}\right]^{1/r}. \tag{3.4.4}$$

Values of $r=0.33$ and $s=1.7$ are recommended for this correlation, but for water s should be changed to 1.0. The values of C_{sf} in this correlation vary with the type of solid surface and the type of fluid in the system. This empirically accounts for material property and/or wetting angle effects. Recommended values of C_{sf} for specific liquid–solid combinations are given by Rohsenow (1962), but whenever possible, an experiment should be conducted to determine the appropriate value of C_{sf} for the particular solid–liquid combination of interest. If this is not possible, a value of $C_{sf}=0.013$ is recommended as a first approximation.

In a more recent study, Stephan and Abdelsalam (1980) proposed the following correlations based on dimensional analysis and an optimal fit to experimental data for water:

$$q''=\{C_1[T_w-T_{sat}(P_l)]\}^{1/0.327}. \tag{3.4.5}$$

In this correlation, C_1 is a constant that varies with pressure, and these investigators provided a plot of C_1 versus pressure that can be used to obtain the value of C_1 for a pressure of interest. The value for boiling of water at atmospheric pressure is $C_1=3.75$. These investigators proposed similar

dimensional correlations for nucleate boiling of hydrocarbons, cryogenic fluids, and refrigerants, and they developed specific dimensionless correlations for nucleate boiling in each of these categories, in addition to the one above for water. Stephan and Abdelsalam (1980) also developed the following correlation that can be applied, with more limited accuracy, to any liquid:

$$\frac{q''D_b}{k_l(T_w - T_{sat})} = 0.23\left(\frac{q''D_b}{k_l T_{sat}}\right)^{0.674}\left(\frac{\rho_v}{\rho_l}\right)^{0.297}\left(\frac{h_{lv}D_b^2}{\alpha_l^2}\right)^{0.371}\left(\frac{\rho_l - \rho_v}{\rho_l}\right)^{-1.73}\left(\frac{\alpha_l^2\rho_l}{\sigma D_b}\right)^{0.35} \quad (3.4.6)$$

In this equation, D_b is the bubble diameter predicted by the bubble departure diameter correlation (6.98) developed by Fritz (1935): $D_d = 0.0208\theta\sqrt{\sigma/[g(\rho_l - \rho_v)]}$, where θ is the contact angle in degrees. Further information on these correlations for specific fluid categories can be obtained in the references by Stephan and Abdelsalam (1980) and Carey (2008).

As noted previously, the pool boiling curve generally exhibits a maximum heat flux or CHF at the transition between nucleate and transition boiling. This peak value is the maximum level of heat flux from the surface, which the system can provide in a nonfilm-boiling mode at a given pressure. The mechanism responsible for the CHF has been the subject of considerable investigation and debate over the past seven decades. As the heat flux increases, bubbles generated at the surface coalesce to form vapor columns or jets. Perhaps the most widely cited CHF model postulates that the CHF condition occurs when Helmholtz instability of the large vapor jets leaving the surface distorts the jets, blocking liquid flow to portions of the heated surface. Continued vaporization of liquid at locations on the surface that are starved of replacement liquid then leads to formation of a vapor blanket over part or all of the surface. According to Zuber (1959), for a flat horizontal surface, the predicted maximum heat flux q''_{max} is

$$q''_{max} = 0.131\rho_v h_{fg}\left[\frac{\sigma_g(\rho_l - \rho_v)}{\rho_v^2}\right]^{1/4}, \quad (3.4.7)$$

but Lienhard and Dhir (1973a) recommend that the constant 0.131 in this relation be replaced with 0.141. Other geometries are treated by Lienhard et al. (1973) and Lienhard and Dhir (1973b). An alternative model has been proposed by Haramura and Katto (1983). Lienhard and Witte (1985) discuss the development and limitations of hydrodynamic CHF theories.

In a more recent study, Kandlikar (2001) developed a model of the critical heat flux mechanism that is not based on hydrodynamic wave instability and includes the effects of contact angle and surface inclination. The model postulates that the critical heat flux condition is attained when the forces acting at the contact line of a bubble on the surface become unbalanced in a way that causes the bubble to spread laterally and blanket the surface.

Kandlikar (2001) argued that components of the surface tension, hydrostatic pressure forces, and evaporation change of momentum forces act parallel to the surface of the heater. Liquid moves toward the interface to replace mass lost by vaporization. The vapor that emerges from the interface on the inside of the bubble near the contact line leaves with a much higher velocity because the high density liquid must accelerate to sustain the same mass flux as low density vapor. The thrust from this accelerating flow acts to push the contact line and sidewall of the bubble laterally outward. The surface tension and hydrostatic pressure force components parallel to the surface act to resist any outward motion of the bubble sidewall and contact line. Kandlikar (2001) developed model relations for the forces on the side of the bubble parallel to the surface and interpreted the point at which they just balance as the incipient CHF condition. Solving for the heat flux at the force balance condition led to the following relation for the critical heat flux

$$q''_{max} = \rho_v h_{lv}\frac{1+\cos\theta}{16}\left[\frac{2}{\pi} + \frac{\pi}{4}(1+\cos\theta)\cos\Omega\right]^{1/2}\left[\frac{\sigma(\rho_l - \rho_v)g}{\rho_v^2}\right]^{1/4}. \quad (3.4.8)$$

The angle of the surface relative to the horizontal Ω was taken into account when constructing the force balance, making this relation applicable to horizontal or inclined surfaces. If both sides of Equation 3.4.8 are divided by the Zuber critical heat flux prediction, and Equation 3.4.7 is used to simplify the result, the following relation is obtained:

$$\frac{q''_{\max}}{q''_{\max,Z}} = \frac{1+\cos\theta}{2.096}\left[\frac{2}{\pi}+\frac{\pi}{4}(1+\cos\theta)\cos\Omega\right]^{1/2}. \tag{3.4.9}$$

Kandlikar (2001) concluded that the dynamic receding contact angle is the appropriate contact to use in this relation. This correlation indicates decreasing critical heat flux as the contact angle increases, which is consistent with the trend in available data. While there are aspects of contact angle effects that warrant further exploration, the modeling used in Kandlikar's (2001) analysis appears to be a promising approach that may lead to better understanding of the effects of contact angle on pool boiling critical heat flux.

As shown in Figure 3.4.3, the boundary between the transition boiling regime and the film boiling regime corresponds to a minimum in the heat flux vs. superheat curve. This condition is referred to as the *minimum heat flux* condition, referred to as the *Leidenfront point*. The minimum heat flux corresponds approximately to the lowest heat flux that will sustain stable film boiling.

For an infinite flat (upward-facing) heated surface, vapor generated at the interface during stable film boiling is released as bubbles at the nodes of a standing two-dimensional Taylor wave pattern. The following relation for the minimum heat flux $q''_{\min}$ was derived by Zuber (1959) and Berenson (1961):

$$q''_{\min} = 0.09\rho_v h_{fg}\left[\frac{\sigma_g(\rho_l-\rho_v)}{(\rho_l+\rho_v)^2}\right]^{1/4}. \tag{3.4.10}$$

$q''_{\min}$ correlations have been developed by Lienhard and Wong (1964) for horizontal cylinders and Gunnerson and Cronenberg (1980) for spheres.

In film boiling, transport of heat across the vapor film from the wall to the interface may occur by convection, conduction, and radiation. The radiation contribution depends on the nature of the solid surface, but when the radiation effect is small, the heat transfer for film boiling is independent of the material properties and finish of the surface. For buoyancy-driven laminar film boiling over a vertical flat isothermal surface in a pool of saturated liquid, the local heat-transfer coefficient from the surface can be obtained from the following relation:

$$h = \left[\frac{k_v^3 g\rho_v(\rho_l-\rho_v)h_{fg}}{4\mu_v(T_w-T_{sat})x}\right]^{1/4}. \tag{3.4.11}$$

At low surface temperatures, radiation effects are negligible and consideration of convective transport alone is sufficient to predict the heat transfer. At higher temperatures, radiation effects must also be included. If the vapor in the film absorbs and emits radiation at infrared wavelengths, a detailed treatment of the radiation interaction with the vapor may be necessary to accurately predict the film boiling heat transfer. Additional information on mechanisms such as interfacial waves, turbulence, and variable properties is summarized in Carey (2008).

Transition pool boiling has traditionally been interpreted as a combination of nucleate and film boiling alternately occurring over the heated surface, and a model of transition boiling that accounts for contact angle effects has been proposed by Ramilison and Lienhard (1987).

Internal Convective Boiling

Flow boiling in tubes is perhaps the most complex convective process encountered in applications. In most evaporator and boiler applications, the flow is either horizontal or vertically upward. Figure 3.4.5 schematically depicts a typical low-flux vaporization process in a horizontal round tube. In this example, liquid enters as subcooled liquid and leaves as superheated vapor. As indicated in Figure 3.4.5, the flow undergoes transitions in the boiling regime and the two-phase flow regime as it proceeds down the tubes. The regimes encountered depend on the entrance conditions and the thermal boundary conditions at the tube wall. At low quality, the vaporization process is dominated by nucleate boiling, with convective effects being relatively weak. As the quality increases, the flow quickly enters the annular film flow regime in which convective evaporation of the annular liquid film is the dominant heat-transfer mechanism. Often the conditions are such that liquid droplets are often entrained in the core vapor flow during annular flow evaporation. Eventually, the annular film evaporates away, leaving the wall dry. Mist-flow evaporation of entrained liquid droplets continues in the post-dryout regime until only vapor remains. Similar, sequences of flow and boiling regimes occurring in vertical upward flow, as indicated in Figure 3.4.6.

The boiling regime trends shown in Figures 3.4.5 and 3.4.6 are typical for low heat flux vaporization processes. At high wall superheat levels, transition boiling or film boiling can also occur. The transition from nucleate boiling to one of these regimes is termed a *departure from nucleate boiling* (DNB) or the CHF condition. However, the heat-transfer performance of an evaporator under transition or film boiling conditions is so poor that equipment is not usually designed to operate under such conditions.

Because low-quality or subcooled flow boiling are nucleate boiling dominated, the heat-transfer coefficient in these regimes is often predicted using a nucleate boiling correlation developed as a fit to pool boiling data. The usefulness of such an approach is a consequence of the fact that for most conditions of practical interest, nucleate boiling heat transfer is only weakly affected by liquid subcooling or liquid bulk convection.

For saturated convective boiling prior to dryout, relations to predict the heat-transfer coefficient have typically been formulated to impose a gradual suppression of nucleate boiling and gradual

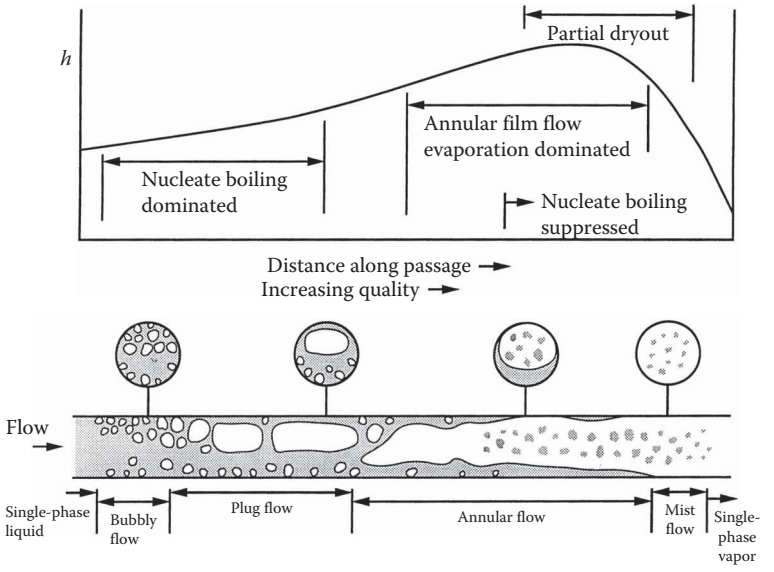


FIGURE 3.4.5 Qualitative variation of the heat-transfer coefficient h and flow regime with quality for internal convective boiling in a horizontal tube at moderate wall superheat.

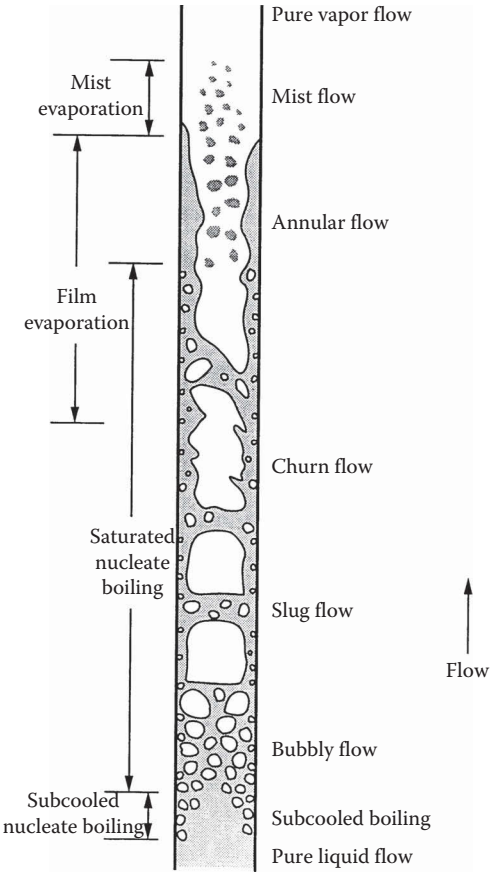


FIGURE 3.4.6 Flow regimes and boiling mechanisms for upflow convective boiling in a vertical tube at moderate wall superheat.

increase in liquid film evaporation heat transfer as the quality increases. A number of correlations based on such an approach have been developed. An early correlation of this type developed by Chen (1966) has been widely used. One of the better tools of this type for predicting the heat-transfer coefficient h is the correlation method proposed by Kandlikar (1991), which was developed to fit a very broad spectrum of heat-transfer data for flow boiling in vertical and horizontal tubes. For this method, the heat-transfer coefficient is computed as follows:

$$h = \text{the maximum of} \begin{Bmatrix} h_{\text{NBD}} \\ h_{\text{CBD}} \end{Bmatrix}, \tag{3.4.12}$$

where for the nucleate boiling dominant regime

$$h_{\text{NBD}} = 0.6683 \left(\frac{\rho_l}{\rho_v} \right)^{0.1} x^{0.16} (1-x)^{0.64} f_2 (\text{Fr}_{\text{le}}) h_{\text{le}} + 1058.0\ \text{Bo}^{0.7} F_{\text{K}} (1-x)^{0.8} h_{\text{le}} \tag{3.4.13}$$

and for the convective boiling dominant regime

$$h_{\mathrm{CBD}} = 1.1360\left(\frac{\rho_{\mathrm{l}}}{\rho_{\mathrm{v}}}\right)^{0.45} x^{0.72}\left(1-x\right)^{0.08} f_2\left(\mathrm{Fr}_{\mathrm{le}}\right) h_{\mathrm{le}} + 667.2\ \mathrm{Bo}^{0.7} F_{\mathrm{K}}\left(1-x\right)^{0.8} h_{\mathrm{le}}. \tag{3.4.14}$$

In this correlation, h_{le} is the single-phase heat-transfer coefficient for the liquid phase flowing alone, and Bo is the boiling number defined as:

$$\mathrm{Bo} = \frac{q''}{G h_{\mathrm{lv}}}. \tag{3.4.15}$$

Kandlikar (1991) later proposed use of the Petukhov–Popov (1963) and Gnielinski (1976) correlations to predict h_{le} because these appeared to better account for the Prandtl number effect for different fluids. For $0.5 \le \mathrm{Pr_l} \le 2000$ and $2300 \le \mathrm{Re_{le}} < 10^4$, h_{le} is therefore computed using the correlation of Gnielinski (1976)

$$h_{\mathrm{le}} = \left(\frac{k_{\mathrm{l}}}{D}\right)\frac{\left(\mathrm{Re_{le}} - 1000\right)\mathrm{Pr_l}\left(f/2\right)}{1 + 12.7\left(\mathrm{Pr_l}^{2/3} - 1\right)\left(f/2\right)^{0.5}}, \tag{3.4.16}$$

and for $0.5 \le \mathrm{Pr_l} \le 2000$ and $10^4 \le \mathrm{Re_{le}} \le 5 \times 10^6$, h_{le} is computed using the correlation of Petukhov and Popov (1963)

$$h_{\mathrm{le}} = \left(\frac{k_{\mathrm{l}}}{D}\right)\frac{\mathrm{Re_{le}}\mathrm{Pr_l}\left(f/2\right)}{1.07 + 12.7\left(\mathrm{Pr_l}^{2/3} - 1\right)\left(f/2\right)^{0.5}}. \tag{3.4.17}$$

In the aforementioned relations for h_{le}, f is the friction factor computed as

$$f = \left[1.58\ \ln\left(\mathrm{Re_{le}}\right) - 3.28\right]^{-2}. \tag{3.4.18}$$

$$\mathrm{Re_{le}} = \frac{GD}{\mu_{\mathrm{l}}}, \tag{3.4.19}$$

and the function $f_2(\mathrm{Fr_{le}})$ is defined as

$$f_2\left(\mathrm{Fr_{le}}\right) = \begin{cases} \left(25\mathrm{Fr_{le}}\right)^{0.3} & \text{for } \mathrm{Fr_{le}} < 0.04 \text{ for horizontal tubes} \\ 1 & \text{for } \mathrm{Fr_{le}} > 0.04 \text{ for horizontal tubes and for vertical tubes.} \end{cases} \tag{3.4.20}$$

Note that on the right side of Equations 3.4.13 and 3.4.14, the first term accounts for the convective effect, and the second term accounts for the nucleate boiling effect. The factor F_{K} is a fluid-dependent parameter, values of which are listed for various fluids in Table 3.4.1.

Methods for predicting the conditions at which dryout or a DNB transition occurs have typically been empirical in nature. Based on fits to extensive data, Levitan and Lantsman (1975) recommended the following relations for the DNB heat flux and the quality at which dryout occurs during flow boiling of water in a tube with an 8 mm diameter.

$$q''_{\mathrm{crit}} = \left[10.3 - 7.8\left(\frac{P}{98}\right) + 1.6\left(\frac{P}{98}\right)^2\right]\left(\frac{G}{1000}\right)^{1.2\left\{\left[0.25(P-98)/98\right]-x\right\}} e^{-1.5x}, \tag{3.4.21}$$

TABLE 3.4.1
Fluid Constant Values for the Kandlikar (1991) Correlation

Fluid	F_K
Water	1.00
R-11	1.30
R-12	1.50
R-13B1	1.31
R-22	2.20
R-113	1.30
R-114	1.24
R-134a	1.63
R-152a	1.10
Nitrogen	4.70
Neon	3.50

$$x_{\text{crit}} = \left[0.39 + 1.57\left(\frac{P}{98}\right) - 2.04\left(\frac{P}{98}\right)^2 + 0.68\left(\frac{P}{98}\right)^3\right]\left(\frac{G}{1000}\right)^{-0.5}. \tag{3.4.22}$$

In these relations, q''_{crit} is in MW/m^2, P is the pressure in bar, and G is in kg/m^2·s. To obtain values of q''_{crit} and x_{crit} for diameters other than 8 mm, Levitan and Lantsman (1975) recommended that the 8 mm values from these relations be corrected as follows:

$$q''_{\text{crit}} = \left(q''_{\text{crit}}\right)_{8\,\text{mm}}\left(\frac{8}{D}\right)^{1/2}, \tag{3.4.23}$$

$$x_{\text{crit}} = \left(x_{\text{crit}}\right)_{8\,\text{mm}}\left(\frac{8}{D}\right)^{0.15}, \tag{3.4.24}$$

where D is the diameter in millimeters. A good generalized empirical correlation for predicting dryout or CHF conditions in vertical uniformly heated tubes is that recently proposed by Katto and Ohno (1984).

In many cases, post-dryout mist flow evaporation is driven primarily by convective transport from the tube wall to the gas and then to the entrained droplets. In some circumstances, impingement of droplets onto the heat surface and radiation interactions may also be important. In cases where convection is dominant, predictions of the heat-transfer coefficient have been developed by modifying a single-phase correlation for the entire flow as vapor with a correction factor, which accounts for the presence of the entrained droplets. Often this correction factor has been presumed to be a function of property ratios. An example of such an approach is the correlation of Dougall and Rohsenow (1963) for which the heat-transfer coefficient h is given by

$$\frac{hD}{k_v} = 0.023\left[\left(\frac{GD}{\mu_v}\right)\left(x + \frac{\rho_v}{\rho_l}(1-x)\right)\right]^{0.8} \text{Pr}_{v,\text{sat}}^{0.4}. \tag{3.4.25}$$

For further information on mechanisms of convective boiling, see the texts of Collier and Thome (2002), Stephan (1992), and Carey (2008).

Thermophysics of Condensation Processes

As in the case of boiling, surface tension effects, surface wetting characteristics, metastable phase stability, and surface morphology also can play important roles in condensation processes. As a result of interfacial tension, the pressure inside a spherical liquid droplet of radius r must exceed that in the surrounding liquid by $2\sigma/r$, so that the liquid and surrounding vapor pressures at equilibrium are related as

$$P_l = P_v + \frac{2\sigma}{r} \tag{3.4.26}$$

A consequence of this and basic thermodynamics is that at equilibrium the surrounding vapor must actually be slightly supersaturated. The amount of supersaturation required at equilibrium increases as the radius of curvature of the bubble interface decreases.

For a liquid droplet on a solid surface with a specified volume, the wetting contact angle dictates the radius of curvature of the droplet interface. Because of the linkage between the interface curvature and the required equilibrium supersaturation, the wetting behavior thus determines the level above which the vapor supersaturation must be raised for the droplet to grow. Steady condensation on the droplet interface can be sustained only if the vapor is driven beyond this supersaturation level by cooling or pressurization. For such conditions, the vapor is in the metastable supersaturated range indicated in Figure 3.4.2.

Condensation on external surfaces of a body immersed in a gas phase generally falls into one or two categories: *dropwise condensation* or *film condensation*. In dropwise condensation, the liquid-phase condensate collects as individual droplets, which grow in size with time on the cold surface. This mode of condensation is most likely when the liquid poorly wets the solid surface. When the condensation rate is high or the liquid readily wets the surface, a film of liquid condensate covers the solid surface, and the process is referred to as film condensation.

Dropwise Condensation

Dropwise condensation may occur on a solid surface cooled below the saturation temperature of a surrounding vapor when the surface is poorly wetted except at locations where well-wetted contaminant nuclei exist. The poorly wetted surface condition can result from contamination or coating of the surface with a substance, which is poorly wetted by the liquid phase of the surrounding vapor. In practice, this can be achieved for steam condensation by (1) injecting a nonwetting chemical into the vapor, which subsequently deposits on the surface, (2) introducing a substance such as a fatty (i.e., oleic) acid or wax onto the solid surface, or (3) by permanently coating the surface with a low-surface-energy polymer or a noble metal. The effects of the first two methods are generally temporary, since the resulting surface films eventually are dissolved or eroded away. More recently, researchers have been exploring ways to microstructure or nanostructure the surface to promote dropwise condensation and enhance heat transfer. Examples of this type of research are described in publications by Ma et al. (2000), Varanasi et al. (2009), Dietz et al. (2010), and Lee et al. (2013).

During dropwise condensation, the condensate is usually observed to appear in the form of droplets, which grow on the surface and coalesce with adjacent droplets. When droplets become large enough, they are generally removed from the surface by the action of gravity or drag forces resulting from the motion of the surrounding gas. As the drops roll or fall from the surface, they merge with droplets in their path, effectively sweeping the surface clean of droplets. Droplets then begin to grow anew on the freshly exposed solid surface. This sweeping and renewal of the droplet growth process is responsible for the high heat transfer coefficients associated with dropwise condensation. Theoretical aspects of dropwise condensation are described in publications by Tanaka (1975, 1979), Wu et al. (2001), Vemuri and Kim (2006), and Lee et al. (2013). A discussion of correlations for the heat transfer coefficient associated with dropwise condensation is provided in the review article by Merte (1973) and the more recent work by Bonner (2013).

External Film Condensation

If the liquid phase fully wets a cold surface in contact with a vapor near saturation conditions, the conversion of vapor to liquid will take the form of film condensation. As the name implies, the condensation takes place at the interface of a liquid film covering the solid surface. Because the latent heat of vaporization must be removed at the interface to sustain the process, the rate of condensation is directly linked to the rate at which heat is transported across the liquid film from the interface to the surface.

The classic integral analysis of Nusselt (1916) for laminar falling-film condensation on a vertical surface considers the physical circumstances shown in Figure 3.4.7. The surface exposed to a motionless ambient of saturated vapor is taken to be isothermal with a temperature below the saturation temperature. Note that although a vertical surface is considered here, the analysis is identical for an inclined surface, except that the gravitational acceleration g is replaced by g sin Ω, with Ω being the angle between the surface and the horizontal. Because the liquid film flows down the surface because of gravity, this situation is sometimes referred to as *falling-film condensation*.

In its simplest form, the classic Nusselt analysis incorporates the following idealizations: (1) laminar flow, (2) constant properties, (3) that subcooling of liquid is negligible in the energy balance, (4) that inertia effects are negligible in the momentum balance, (5) that the vapor is stationary and exerts no drag, (6) that the liquid–vapor interface is smooth, and (7) that heat transfer across film is only by conduction (convection is neglected). With these idealizations, the following relation for the local heat transfer coefficient h can be obtained

$$\frac{hz}{k_l} = \left[\frac{\rho_l (\rho_l - \rho_v) g h_{fg} z^3}{4 k_l \mu_l (T_{sat} - T_w)} \right]^{1/4} . \tag{3.4.27}$$

Modified versions of this analysis have been subsequently developed that relax many of these assumptions. Laminar film condensation on a vertical surface can also be analyzed with a full boundary layer formulation. An example of this type of approach is the analysis presented by Sparrow and Gregg (1959).

The analyses described earlier do not include two physical mechanisms that can significantly affect the transport: (1) the effects of waves on the liquid–vapor interface and (2) interfacial vapor shear drag on the interface. The effects of interfacial shear have been studied analytically by numerous investigators. The effects of surface waves on laminar film condensation are more difficult to incorporate into theoretical analyses. In general, interfacial waves are expected to

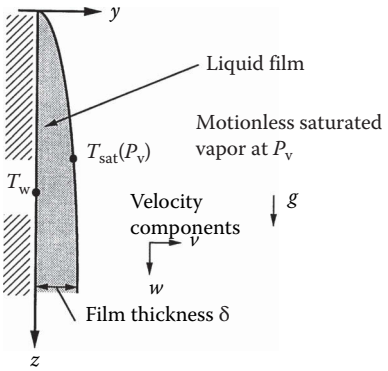


FIGURE 3.4.7 System model for the Nusselt analysis of falling-film condensation.

enhance convective heat transport in the film since it intermittently thins the film, increases the interfacial area, and induces mixing. Because of these effects, laminar film condensation heat transfer data are often significantly higher than the values predicted by simple boundary layer models.

As for any boundary layer flow, when the film Reynolds number becomes large enough, it is expected that a transition to turbulent flow will occur. Eddy diffusivity models of the resulting turbulent transport have been developed by Seban (1954), Dukler (1960), and others. This methodology was later extended to evaporation of a falling liquid film (see, e.g., Mills and Chung, 1973).

Subsequent studies (see, e.g., Mills and Chung, 1973) have suggested that the presence of the interface tends to damp larger turbulent eddies near the interface in the liquid film. This implies that a viscous sublayer exists at the interface as well as at the wall. Recent efforts to model turbulent falling-film evaporation and condensation processes have therefore included a variation of the eddy viscosity in which it goes to zero at both the wall and the interface. The analysis tools and correlations described earlier work reasonably well for values of liquid Prandtl number above 1. However, deviation of the predictions using these methods from heat transfer data for liquid metals can be quite significant.

Because of its importance to the design of tube-and-shell condensers, condensation on the outside of horizontal tubes has been the subject of numerous studies. The length of the tube perimeter over which the condensate flows is usually small for commonly used tubes. Consequently, the film Reynolds number is usually low and the flow in the liquid film is laminar.

With slight modification, the Nusselt (1916) analysis of laminar falling-film condensation over a flat plate can be adapted to film condensation on an isothermal horizontal cylinder. Doing so yields the following relation for the mean heat transfer coefficient:

$$\frac{\bar{h}D}{k_l} = 0.728\left[\frac{(\rho_l - \rho_v)gh_{fg}D^3\,\mathrm{Pr}_l}{\rho_l \nu_l^2 c_{pl}(T_{sat} - T_w)}\right]^{1/4}. \tag{3.4.28}$$

Selin (1961) found that better agreement with film condensation data for horizontal tubes was obtained by replacing the constant factor in Equation 3.4.28 by 0.61. Correlations similar to this single-tube relation have also been developed for the average condensation heat transfer coefficient for banks of round tubes.

Analytical treatment of laminar film condensation on a sphere is virtually the same as that for a horizontal cylinder. The only differences result from the angular variation of the body perimeter for a spherical geometry. A general analytical prediction of the local heat transfer coefficient for laminar film condensation on arbitrary axisymmetric bodies has been developed by Dhir and Lienhard (1971).

Condensation in the Presence of a Noncondensable Gas

In nature and in a number of technological applications, condensation of one component vapor in a mixture may occur in the presence of other noncondensable components. The most common example is the condensation of water vapor in the air on a cold solid surface. If the component gases are considered to be a mixture of independent substances, condensation of one component vapor will occur if the temperature of the surface is below the saturation temperature of the pure vapor at its partial pressure in the mixture. This temperature threshold is referred to as the *dew point* of the mixture.

Because only the vapor is condensed, the concentration of the noncondensable gas at the interface is higher than its value in the far ambient. This, in turn, decreases the partial pressure of the vapor at the interface below its ambient value. The corresponding saturation temperature at the interface is therefore lower than the bulk temperature. The resulting depression of the

interface temperature generally reduces the condensation heat transfer rate below that, which would result for pure vapor alone under the same conditions. Space limitations here preclude a detailed discussion of the effects of noncondensable gases. The interested reader may find more extensive discussions of this topic in the references by Collier and Thome (2002) and Carey (2008).

Internal Convective Condensation

In most power and refrigeration systems, the flow in the condenser is either horizontal or vertically downward. Figure 3.4.8 schematically depicts a typical condensation process in a horizontal round tube. Superheated vapor enters the tube and at the exit end the liquid is subcooled. At a point some distance downstream of the entrance, vapor begins to condense on the walls of the tube. The location at which this occurs is at or slightly before the bulk flow reaches the equilibrium saturation condition. In most condensers, the liquid readily wets the interior of the tube and at high vapor volume fractions the liquid forms a thin liquid film on the interior wall of the tube.

The vapor velocity is generally high at the inlet end of the condenser tube, and the liquid film is driven along the tube by strong vapor shear on the film. At low vapor flow rates, some stratification may occur and the film may be thicker on the bottom of the horizontal tube. At high vapor flow rates, turbulent stresses acting on the liquid may tend to keep the thickness of the liquid film nominally uniform over the perimeter of the tube.

In most condenser applications, shear-dominated annular flow persists to very low qualities and the overwhelming majority of the heat transfer occurs in this regime. The very last stage of the condensation process, corresponding to qualities less than a few percent, may occur in slug, plug, or bubbly two-phase flow. Generally, these regimes represent such a small portion of the overall heat transfer in the condenser that some inaccuracy in estimating the heat transfer coefficient for them is tolerated. As a first estimate, the heat transfer coefficient may be predicted using a correlation for pure single-phase liquid flow in the tube at the same total flow rate, or a correlation for annular flow condensation may simply be extrapolated to zero quality.

Because most of the heat duty occurs in the annular flow regime, accurate prediction of the overall heat transfer performance of the condenser requires a predictive methodology that accurately treats the transport in this regime. For this reason, the form of most correlation methods for predicting local convective condensation heat transfer coefficients is optimized to match data in the annular flow regime. One example of such a correlation is the following relation for the local heat transfer coefficient for annular flow condensation proposed by Traviss et al. (1973):

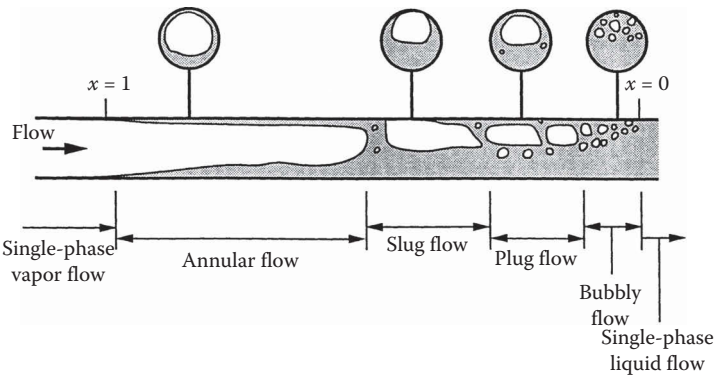


FIGURE 3.4.8 Flow regimes during horizontal cocurrent flow with condensation.

$$\frac{hD}{k_l} = \frac{0.15 \mathrm{Pr}_l \mathrm{Re}_l^{0.9}}{F_T} \left[\frac{1}{X_{tt}} + \frac{2.85}{X_{tt}^{0.476}} \right], \tag{3.4.29}$$

where

$$\mathrm{Re}_l = \frac{G(1-x)D}{\mu_l}, \quad X_{tt} = \left(\frac{1-x}{x} \right)^{0.9} \left(\frac{\rho_v}{\rho_l} \right)^{0.5} \left(\frac{\mu_l}{\mu_v} \right)^{0.1}, \tag{3.4.30}$$

and F_T is given by

$$\begin{aligned} F_T &= 5\mathrm{Pr}_l + 5\ln\{1 + 5\mathrm{Pr}_l\} + 2.5\ln\{0.0031\mathrm{Re}_l^{0.812}\} && \text{for} \quad \mathrm{Re}_l > 1125 \\ &= 5\mathrm{Pr}_l + 5\ln\{1 + \mathrm{Pr}_l(0.096\mathrm{Re}_l^{0.585} - 1)\} && \text{for} \quad 50 < \mathrm{Re}_l < 1125 \\ &= 0.707\mathrm{Pr}_l\mathrm{Re}_l^{0.5} && \text{for} \quad \mathrm{Re}_l < 50 \end{aligned} \tag{3.4.31}$$

Carey (2008) has shown that the generic form of this correlation can be derived from a theoretical model of annular flow condensation in a round tube. Several correlations of this general type have been developed as fits to experimental data; see Carey (2008) for a summary.

Correlation methods of this form have also been proposed for microchannels and channels of noncircular cross section. Wang and Rose (2005) noted that the data from several experimental investigations (see, e.g., Garimella, 2003) indicate that annular flow spans an increasing portion of the condensation process as the hydraulic diameter of the passage decreases. As a consequence, annular flow has been a main focus of efforts to empirically and theoretically predict convective condensation heat transfer in microchannels, and correlations of the form described earlier may be expected to work well for this type of passage. For annular flow condensation in rectangular microchannels, Wang et al. (2002) found that the following predictive correlation provided a best fit to 700 annular flow data for convective condensation heat transfer:

$$\frac{hD_h}{k_l} = 0.027\mathrm{Pr}_l\mathrm{Re}_l^{0.6792} x^{0.2208} \left(\frac{1.376 + 8X_{tt}^{1.655}}{X_{tt}^2} \right)^{0.5}. \tag{3.4.32}$$

In this equation, D_h is the hydraulic diameter of the rectangular microchannel, and the Reynolds number $\mathrm{Re}_l = G(1-x)D_h/\mu_l$ for the liquid flowing alone is based on D_h. At specified local conditions, this empirical correlation generally predicts a local heat transfer coefficient value that is lower than that predicted by the large round tube correlations discussed earlier. As noted by Wang et al. (2002), this suggests that the annular film in a small rectangular tube has a different configuration than that in round tubes for the same conditions.

A correlation methodology that can be used to predict internal convective condensation heat transfer for slug, plug, or wavy stratified flow has also been proposed by Rossen and Meyers (1965).

The predictions of the correlations discussed in this section may vary significantly for a given set of conditions. When possible, a correlation should be selected, which has been tested against data for conditions close to those for the application of interest. To predict the overall heat transfer performance of a condenser, methods to predict the local heat transfer coefficient must be combined with a scheme to numerically integrate finite-difference forms of the energy, mass, and momentum balances in the tube. For further information on such schemes, see the references by Collier and Thome (2002) and Carey (2008) (Figure 3.4.9).

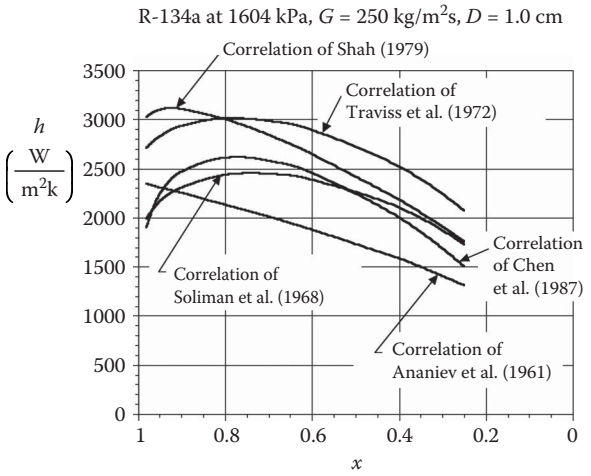


FIGURE 3.4.9 Comparison of the variation of h with x predicted by four correlation methods for internal convective condensation. References cited in this figure are listed in Chapter 11 of Carey (2008).

Defining Terms

Critical heat flux (CHF)	A maximum heat flux condition that characterizes the transition between nucleate boiling and transition boiling or film boiling.
Dropwise condensation	Condensation of vapor into liquid in discrete droplets, usually attained when a cold surface is poorly wetted by the liquid phase.
Film boiling	Generation of vapor at the interface of a vapor film entirely covers the hot surface.
Film condensation	Condensation of vapor onto the interface of a liquid film that completely covers a cold surface.
Minimum heat flux	A minimum heat flux condition on the classic boiling curve that characterizes the transition between film boiling and transition boiling. Also, sometimes referred to as the Leidenfrost point, it is a lower bound for heat flux values at which stable film boiling may occur.
Nucleate boiling	Generation of vapor at a hot surface by formation of bubbles at discrete nucleation sites with full liquid wetting of the surface.
Polar molecules	Molecules that have a permanent electric dipole moment. Examples include water and ammonia.
Pool boiling	Generation of vapor at the surface of a hot body immersed in an extensive liquid pool.
Transition boiling	Generation of vapor at a hot surface with intermittent or partial liquid wetting of the surface.

REFERENCES

Annaiev, E.P., Boyko, L.D. and Kruzhilin, G.N. 1961. Heat transfer in the presence of steam condensation in a horizontal tube. *Proc. 1st Int. Heat Transfer Conference*, part II, 290.

Berenson, P.J. 1961. Film boiling heat transfer from a horizontal surface. *J. Heat Transfer*, 83, 351–356.

Bonner, R.W. III, 2013. Correlation for dropwise condensation heat transfer: Water, organic fluids, and inclination. *Int. J. Heat Mass Transfer*, 61, 245–253.

Carey, V.P. 2008. *Liquid–Vapor Phase Change Phenomena*, 2nd Edition, Taylor and Francis, Washington, DC.

Chen, S.L., Gerner, F.M. and Tien, C.L. 1987. General film condensation correlations. *Experimental Heat Transfer*, 1, 93–107.

Chen, J.C. 1966. Correlation for boiling heat transfer to saturated fluids in convective flow. *Ind. Eng. Chem. Proc. Des. Dev.*, 5(3), 322–339.

Collier, J.G. and Thome, J.R. 2002. *Convective Boiling and Condensation*, 3rd Edition, Oxford University Press, Cary, NC.

Dhir, V.K. and Lienhard, J. 1971. Laminar film condensation on plane and axisymmetric bodies in nonuniform gravity. *J. Heat Transfer,* 93, 97–100.

Dietz, C., Rykaczewski, K., Fedorov, A. G., and Joshi, Y. 2010. Visualization of droplet departure on a superhydrophobic surface and implications to heat transfer enhancement during dropwise condensation. *Appl. Phys. Lett.*, 97, 033104.

Dougall, R.S. and Rohsenow, W.M. 1963. Film boiling on the inside of vertical tubes with upward flow of the fluid at low qualities. MIT, Cambridge, MA, MIT Report No. 9079-26.

Dukler, A.E. 1960. Fluid mechanics and heat transfer in vertical falling film systems. *Chem. Eng. Prog. Symp. Ser.*, 56(30), 1–10.

Fritz, W. 1935. Berechnung des Maximalvolume von Dampfblasen. *Phys. Z.*, 36, 379–388.

Garimella, S. 2003. Condensation flow mechanisms in microchannels: Basis for pressure drop and heat transfer models. *Proceedings 1st International Conference on Microchannels and Minichannels,* ASME, Rochester, NY, 181–192.

Gnielinski, V. 1976. New equations for heat and mass transfer in turbulent pipe and channel flow. *Int. Chem. Eng.*, 16, 359–368.

Gunnerson, F.S. and Cronenberg, A.W. 1980. On the minimum film boiling conditions for spherical geometries. *J. Heat Transfer*, 102, 335–341.

Haramura, Y. and Katto, Y. 1983. A new hydrodynamic model of the critical heat flux, applicable widely to both pool and forced convective boiling on submerged bodies in saturated liquids. *Int. J. Heat Mass Transfer*, 26, 389–399.

Kandlikar, S.G. 1991. Development of a flow boiling map for subcooled and saturated flow boiling of different fluids in circular tubes. *J. Heat Transfer*, 113, 190–200.

Kandlikar, S.G. 2001. A theoretical model to predict pool boiling CHF incorporating effects of contact angle and orientation. *J. Heat Transfer*, 123, 1071–1079.

Katto, Y. and Ohno, H. 1984. An improved version of the generalized correlation of critical heat flux for the forced convective boiling in uniformly heated vertical tubes. *Int. Heat Mass Transfer*, 21, 1527–1542.

Lee, S., Yoon, H.K., Kim, K.J., Kim, S., Kennedy, M., and Zhang, B.J. 2013. A dropwise condensation model using a nano-scale, pin structured surface. *Int. J. Heat Mass Transfer*, 60, 664–671.

Levitan, L.L. and Lantsman, F.P. 1975. Investigating burnout with flow of a steam-water mixture in a round tube. *Therm. Eng.* (USSR). English trans., 22, 102–105.

Lienhard, J.H., and Dhir, V.K. 1973a. Extended hydrodynamic theory of the peak and minimum pool boiling heat fluxes. NASA CR-2270.

Lienhard, J.H., and Dhir, V.K. 1973b. Hydrodynamic prediction of peak pool-boiling heat fluxes from finite bodies. *ASME J. Heat Transfer*, 95, 152–158.

Lienhard, J.H., Dhir, V.K., and Riherd, D.M. 1973. Peak pool boiling heat flux measurements on finite horizontal flat plates. *ASME J. Heat Transfer*, 95, 477–482.

Lienhard, J.H. and Witte, L.C. 1985. A historical review of the hydrodynamic theory of boiling. *Rev. Chem. Eng.*, 3, 187–277.

Lienhard, J.H. and Wong, P.T.Y. 1964. The dominant unstable wavelength and minimum heat flux during film boiling on a horizontal cylinder. *J. Heat Transfer*, 86, 220–226.

Ma, X., Rose, J.W., Xu, D., Lin, J., and Wang, B. 2000. Advances in dropwise condensation heat transfer: Chinese research. *Chem. Eng. J.*, 78, 87–93.

Merte, H. 1973. Condensation heat transfer. *Adv. Heat Transfer*, 9, 181–272.

Mills, A.F. and Chung, D.K. 1973. Heat transfer across turbulent falling films. *Int. J. Heat Mass Transfer*, 16, 694–696.

Nukiyama, S. 1934. The maximum and minimum values of Q transmitted from metal to boiling water under atmospheric pressure. *J. Jpn. Soc. Mech. Eng.*, 37, 367–374.

Nusselt, W. 1916. Die Oberflachenkondensation des Wasser dampfes. Z. *Ver. Dtsch. Ininuere*, 60, 541–575.

Petukhov, B.S. and Popov, V.N. 1963. Theoretical calculation of heat exchange in turbulent flow in tubes of an incompressible fluid with variable physical properties. *High Temp.*, 1, 69–83.

Ramilison, J.M. and Lienhard, J.H. 1987. Transition boiling heat transfer and the film transition regime. *J. Heat Transfer*, 109, 746–752.

Rohsenow, W.M. 1962. A method of correlating heat transfer data for surface boiling of liquids. *Trans. ASME*, 84, 969–975.

Rossen, H.F. and Meyers, J.A. 1965. Point values of condensing film coefficients inside a horizontal tube. *Chem. Eng. Prog. Symp. Ser.*, 61(59), 190–199.

Seban, R. 1954. Remarks on film condensation with turbulent flow. *Trans. ASME*, 76, 299–303.

Selin, G. 1961. Heat transfer by condensing pure vapors outside inclined tubes. *Proceedings First International Heat Transfer Conference*, University of Colorado, Boulder, Part II, 279–289.
Shah, M.M. 1979. A general correlation for heat transfer during film condensation inside pipes. *Int. J. Heat and Mass Transfer*, 22, 547–556.
Soliman, M., Schuster, J.R., and Berenson, P.J. 1968. A general heat transfer correlation for annular flow condensation. *J. Heat Transfer*, 90, 267–276.
Sparrow, E.M. and Gregg, J.L. 1959. A boundary-layer treatment of laminar film condensation. *J. Heat Transfer*, 81, 13–23.
Stephan, K. 1992. *Heat Transfer in Condensation and Boiling.* Springer-Verlag, New York.
Stephan, K. and Abdelsalam, M. 1980. Heat-transfer correlations for natural convection boiling. *Int. J. Heat Mass Transfer*, 23, 73–87.
Tanaka, H. 1975. A theoretical study of dropwise condensation. *J. Heat Transfer*, 97, 72–78.
Tanaka, H. 1979. Further developments of dropwise condensation theory. *J. Heat Transfer*, 101, 603–611.
Traviss, D.P., Rohsenow, W.M. and Baron, A.B. 1972. Forced convection condensation in tubes: A heat transfer correlation for condenser design. *ASHRAE Transactions*, 79(I), 157–165.
Traviss, D.P., Rohsenow, W.M., and Baron, A.B. 1973. Forced convection condensation in tubes: A heat transfer correlation for condenser design. *ASHRAE Trans*., 79(I), 157–165.
Varanasi, K.K., Hsu, M., Bhate, N., Yang, W., and Deng, T. 2009. Spatial control in the heterogeneous nucleation of water. *Appl. Phys. Lett*., 95, 094101.
Vemuri, S. and Kim, K.J. 2006. An experimental and theoretical study on the concept of dropwise condensation. *Int. J. Heat Mass Transfer*, 49, 649–657.
Wang, H.S. and Rose, J.W. 2005. A theory of film condensation in horizontal noncircular section microchannels. *J. Heat Transfer,* 127, 1096–1105.
Wang, W.W.W., Radcliff, T.D., and Christensen, R.N. 2002. A condensation heat transfer correlation for millimeter-scale tubing with flow regime transition. *Exp. Therm. Fluid Sci.,* 26, 473–485.
Witte, L.C. and Lienhard, J.H. 1982. On the existence of two "transition" boiling curves. *Int. J. Heat Mass Transfer*, 25, 771–779.
Wu, Y-T., Yang, C-X., and Yuan, X-G. 2001. Drop distributions and numerical simulation of dropwise condensation heat transfer. *Int. J. Heat Mass Transfer*, 44, 4455–4464.
Zuber, N. 1959. Hydrodynamic aspects of boiling heat transfer. AEC Report AECU-4439.

Further Information

The texts *Heat Transfer in Condensation and Boiling* by K. Stephan (Springer-Verlag, New York, 1992) and *Liquid–Vapor Phase Change Phenomena* by V.P. Carey (Taylor and Francis, Washington, DC, 2008 provide an introduction to the physics of boiling and condensation processes. The text by J.G. Collier and J.R. Thome, *Convective Boiling and Condensation* (3rd ed., McGraw-Hill, New York, 2002), summarizes more-advanced elements of convective boiling and condensation processes. The *ASHRAE Handbook of Fundamentals* (American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA, 2013) provides some information on boiling and condensation heat transfer and is a good source of thermophysical property data needed to analyze boiling and condensation processes.

3.4.2 Particle Gas Convection

John C. Chen

Introduction

Heat transfer in two-phase systems involving gas and solid particles is encountered in several types of operations important in chemical, power, and environmental technologies. Chief among these are gas fluidized beds which are widely used to achieve either physical processing or chemical reactions that require interfacial contact between gas and solid particles. Currently, fluidized beds operate in either the *bubbling regime* or the *fast-circulating regime*. In the first case, particles are retained in the fluidized bed while the gas passes upward past the particles, partially as rising bubbles. In the second case, gas velocities exceed terminal velocity for the individual particles and the two phases flow through the fluidized bed in cocurrent upward flow. For those applications which require thermal control, convective heat transfer between the fluidized medium and heat transfer surfaces (either immersed tubes or the vessel walls) is an essential element of the process design.

Bubbling Fluidized Beds

Bubbling fluidization occurs when the superficial gas velocity exceeds a critical value wherein the gravitational body force on the solid particles is balanced by the shear force between particles and flowing gas. The superficial gas velocity at this condition, commonly called the minimum fluidization velocity (U_{mf}), marks the boundary between gas flow through packed beds and gas flow in fluidized beds. Wen and Yu (1966) derived the following general equation to estimate U_{mf} for spherical particles:

$$\mathrm{Re}_{mf} = \left[(33.7)^2 + 0.0408\,\mathrm{Ar} \right]^{1/2} - 33.7 \tag{3.4.33}$$

where

$$\mathrm{Re}_{mf} = \text{particle Reynolds number at } U_{mf} = \frac{U_{mf} d_p \rho_g}{\mu_g}$$

$$\mathrm{Ar} = \text{Archimedes number} = \frac{d_p^3 \rho_g (\rho_s - \rho_g) g}{\mu_g^2}$$

Increasing gas velocity beyond minimum fluidization causes the excess gas to collect into discrete bubbles that grow and rise through the fluidized matrix of solid particles. In this bubbling fluidization regime, the total pressure drop over the height of the fluidized bed, H, is equal to the hydrostatic pressure of the solid mass,

$$\Delta P = g \rho_s (1 - \varepsilon) H \tag{3.4.34}$$

where ε = volume fraction of gas (void fraction).

Tubes carrying cooling or heating fluids are often immersed in bubbling fluidized beds to extract or add thermal energy. The effective heat transfer coefficient at the surface of such tubes has been the objective of numerous experimental and analytical investigations. Data for the circumferentially averaged heat transfer coefficient for horizontal tubes are shown in Figure 3.4.10 for various types of solid particles. Characteristics representative of such systems are

- The heat transfer coefficient increases sharply as the gas velocity exceeds minimum fluidization velocity.
- After the initial increase, the heat transfer coefficient remains fairly constant over a significant range of the gas velocity beyond minimum fluidization velocity.

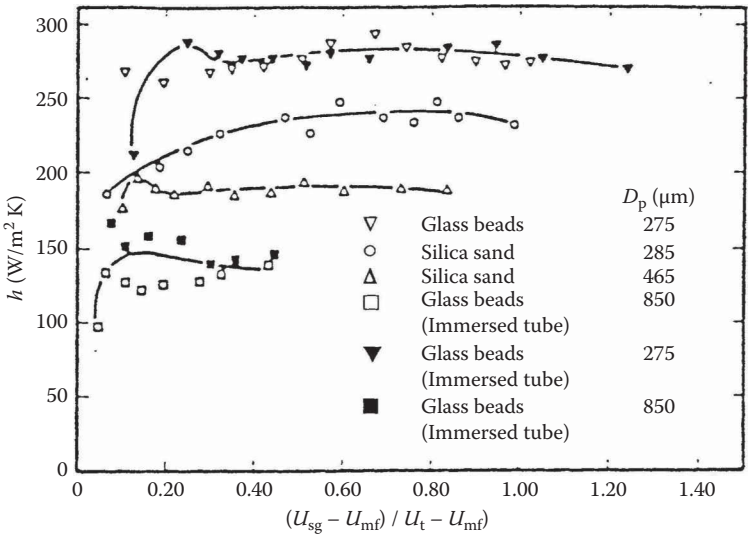


FIGURE 3.4.10 Average heat transfer coefficients for horizontal tubes immersed in bubbling fluidized beds. (From Biyikli, K. et al., *AIChE J.*, 29(5), 712–716, 1983. With permission.)

- The absolute magnitude of the heat transfer coefficient is severalfold greater than single-phase gas convection at the same superficial velocity.
- The heat transfer coefficient increases as particle size decreases.

Kunii and Levenspiel (1991) have shown that increasing gas pressure and density significantly increases the magnitude of the heat transfer coefficient as well as promoting the occurrence of minimum fluidization at a lower value of superficial gas velocity. The effect of bundle spacing is insignificant at 1-atm pressure but becomes increasingly more important as gas pressure and density increase. The data of Jacob and Osberg (1957) indicate that the convective heat transfer coefficient in fluidized beds increases with increasing thermal conductivity of the gas phase, for any given particle size.

Several different types of correlations have been suggested for predicting convective heat transfer coefficients at submerged surfaces in bubbling fluidized beds. The first type attributes the enhancement of heat transfer to the scouring action of solid particles on the gas boundary layer, thus decreasing the effective film thickness. These models generally correlate a heat transfer Nusselt number in terms of the fluid Prandtl number and a modified Reynolds number with either the particle diameter or the tube diameter as the characteristic length scale. Examples are

Leva's correlation for vertical surfaces and larger particles (Leva and Grummer, 1952);

$$\mathrm{Nu}_{\mathrm{dp}} = \frac{h_c d_p}{k_g} = 0.525\left(\mathrm{Re}_p\right)^{0.75} \tag{3.4.35}$$

where

$$\mathrm{Re}_p = \frac{d_p \rho_g U}{\mu_g}$$

Vreedenberg's (1958) correlation for horizontal tubes refers to the particle of diameter D_t.

$$\mathrm{Nu}_{\mathrm{Dt}} = \frac{h_c D_t}{k_g} = 420\left(\frac{\rho_s}{\rho_g}\mathrm{Re}_t\right)^{0.3}\left(\frac{\mu_g^2}{g\rho_s^2 d_p^3}\right)^{0.3}\left(\mathrm{Pr}_g\right)^{0.3} \tag{3.4.36}$$

for

$$\left(\frac{\rho_s}{\rho_g}\mathrm{Re}_t\right) > 2250$$

where

$$\mathrm{Re}_t = \frac{D_t \rho_g U}{\mu_g}$$

Molerus and Scheinzer (1989) developed an alternative type of correlation based on the supposition that the heat transfer is dominated by gas convection through the matrix of particles in the vicinity of the heat transfer surface. Their correlation takes the form:

$$\mathrm{Nu} = \frac{h_c d_p}{k_g} = 0.0247(\mathrm{Ar})^{0.4304}(\mathrm{Pr})^{0.33} \tag{3.4.37}$$

Figure 3.4.11 shows comparison of this model with experimental data obtained at three different pressures. The solid curve represents the relationship for fixed beds, while the dashed lines represent the behavior for fluidized beds (i.e., Equation 3.4.37) upon exceeding minimum fluidization.

A third type of model considers the heat transfer surface to be contacted alternately by gas bubbles and packets of packed particles, leading to a surface renewal process for heat transfer. Mickley and Fairbanks (1955) provided the first analysis of this renewal mechanism. Ozkaynak and Chen (1980) showed that if experimentally measured values of the packet contact time and residence times are used in the packet model analysis, excellent agreement is obtained.

Fast-Circulating Fluidized Beds

Fast fluidization occurs when the superficial gas velocity exceeds the terminal velocity of the solid particles, causing the particles to be suspended in cocurrent upward flow with the gas. This upward flow

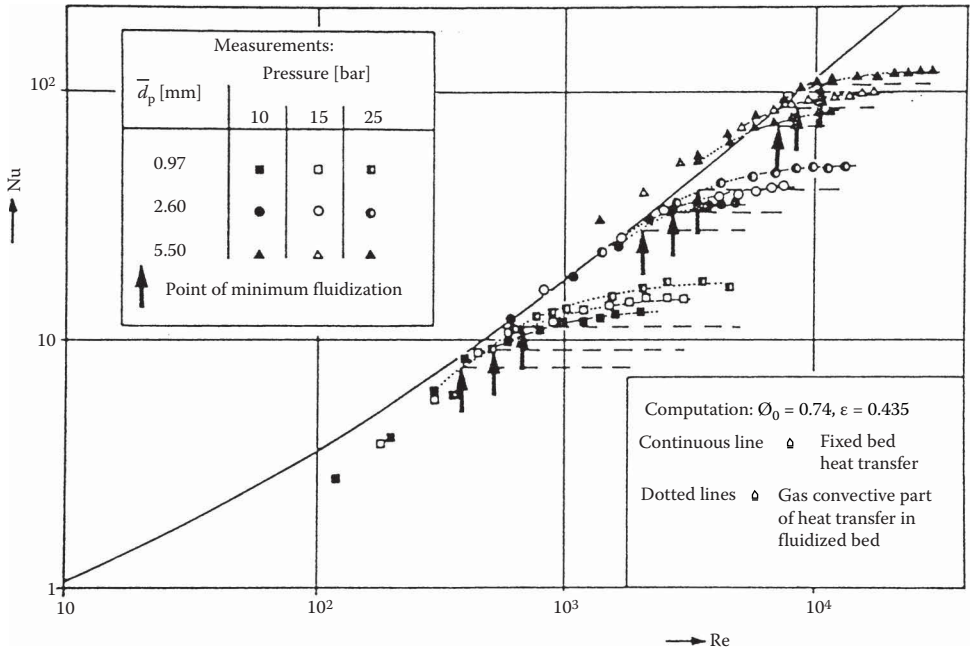


FIGURE 3.4.11 Correlation of Molerus and Scheinzer compared with experimental data (1989).

occurs in "rise reactors" wherein desired physical or chemical reactions occur. In most applications, the two-phase flow exits the top of the riser into a cyclone where the gas phase is separated and exhausted while the solid particles are captured and returned for reinjection at the bottom of the riser. The volumetric concentration of solid particles in these fast fluidized beds (FFBs) tend to be fairly dilute, often with average concentrations of less than 2%. Heat exchange with the particle/gas suspension is usually accomplished through the vertical wall surfaces or through vertical tubes immersed in the duct.

The heat transfer coefficient at vertical surfaces FFBs has been found to increase with increasing solid concentration, aside from other second-order parametric effects. Figure 3.4.12 shows heat transfer coefficients experimentally measured by Dou et al. (1994) for an FFB operating with sand particles of 124 μm mean diameter. Figure 3.4.12b shows that the heat transfer coefficient increased with solid mass flux, for a constant superficial gas velocity. Figure 3.4.12a shows that the heat transfer coefficient decreased parametrically with superficial gas velocity for a constant solid mass flux. Both figures indicate that heat transfer coefficients decrease with increasing elevation in the riser duct. These three parametric trends are all consistent with the hypothesis that heat transfer in FFBs increases with increasing concentration of the solid phase.

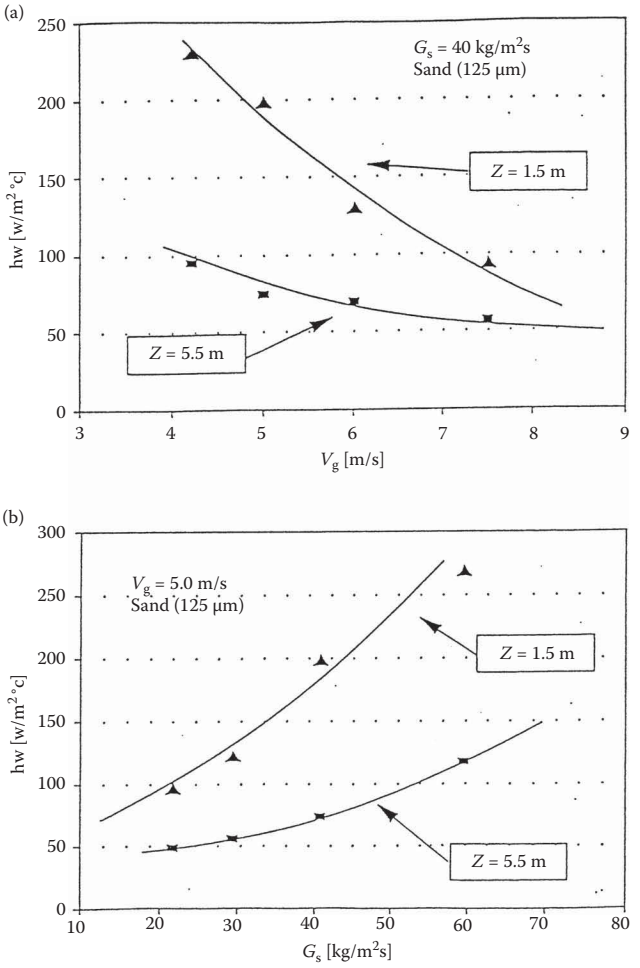


FIGURE 3.4.12 Heat transfer coefficients in fast fluidized beds; V_g is superficial gas velocity, G_s is mass flux of particles, and Z is elevation in FFB. (From Dou, S. et al., 1992. *Fluidization VII,* Engineering Foundation, 793–802. With permission.)

It is generally accepted that the effective heat transfer coefficient for surfaces in FFBs have contributions for gas-phase convection, particle-induced convection, and radiation:

$$h = h_g + h_p + h_r \tag{3.4.38}$$

In contrast to the situation in dense-bubbling fluidized beds, the relatively dilute concentration of solid particles in FFBs often results in significant contributions from all three heat transfer mechanisms. The radiation coefficient can be obtained by a gray body model suggested by Grace (1985). The contribution of the gas phase convection (h_g) is commonly estimated based on correlations for gas flow alone at the same superficial gas velocity. Although the presence of particles may alter the turbulence characteristic of this gas flow, any errors caused by this procedure are usually small since h_g is generally smaller than the particle-phase convective coefficient h_p.

For most FFBs, the particle convective contribution to heat transfer is most important and the prediction of h_p is the major concern in thermal design. Unfortunately, mechanistically based models are still lacking and most design methods rely on empirical correlations which often combine the contributions of gas and particle phases into a single convective heat transfer coefficient (h_c). One such correlation proposed by Wen and Miller (1961) is

$$\mathrm{Nu}_{dp} = \frac{h_c d_p}{k_g} = \left(\frac{C_{pp}}{C_{pg}}\right)\left(\frac{\rho_{susp}}{\rho_p}\right)^{0.3}\left(\frac{V_t}{g d_p}\right)^{0.21} \mathrm{Pr}_g \tag{3.4.39}$$

where V_t=terminal velocity of particle.

Other correlations have been proposed by Fraley (1983) and Martin (1984). These correlations are useful as a starting point but have not yet been verified over wide parametric ranges. Large deviations can occur when compared with measurements obtained outside of the experimental parametric ranges.

REFERENCES

Biyikli, K., Tuzla, K., and Chen, J.C. 1983. Heat transfer around a horizontal tube in freeboard region of fluidized beds, *AIChE J.*, 29(5), 712–716.

Dou, S., Herb, B., Tuzla, K., and Chen, J.C. 1992. Dynamic variation of solid concentration and heat transfer coefficient at wall of circulating fluidized bed, in *Fluidization VII*, Eds. Potter, O.E., and Nicklin, D.J., Engineering Foundation, 793–802.

Fraley, L.D., Lin, Y.Y., Hsiao, K.H., and Solbakken, A. 1983. ASME Paper 83-HT-92, National Heat Transfer Conference, Seattle, ASME, New York.

Grace, J.R. 1985. Heat transfer in circulating fluidized beds, *Circulating Fluidized Bed Technology I*, Ed. Basu, P., Pergamon Press, New York, 63–81.

Jacob, A. and Osberg, G.L. 1957. Effect of gas thermal conductivity on local heat transfer in a fluidized bed, *Can. J. Chem. Eng.*, 35(6), 5–9.

Kunii, D. and Levenspiel, O. 1991. *Fluidization Engineering*, 2nd ed., Butterworth-Heinemann, Boston.

Leva, M. and Grummer, M. 1952. A correlation of solids turnovers in fluidized systems, *Chem. Eng. Prog.*, 48(6), 307–313.

Martin, H. 1984. Heat transfer between gas fluidized beds of solid particles and the surfaces of immersed heat exchanger elements. *Chem. Eng. Process.*, 18, 157–169.

Mickley, H.S. and Fairbanks, D.F. 1955. Mechanism of heat transfer to fluidized beds, *AIChE J.*, 1(3), 374–384.

Molerus, O. and Scheinzer, J. 1989. Prediction of gas convective part of the heat transfer to fluidized beds, in *Fluidization IV*, Engineering Foundation, New York, 685–693.

Ozkaynak, T.F. and Chen, J.C. 1980. Emulsion phase residence time and its use in heat transfer models in fluidized bed, *AIChE J.*, 26(4), 544–550.

Vreedenberg, H.A. 1958. Heat transfer between a fluidized bed and a horizontal tube, *Chem. Eng. Sci.*, 9(1), 52–60.

Wen, C.Y. and Miller, E.N. 1961. Heat transfer in solid-gas transport lines, *Ind. Eng. Chem.*, 53, 51–53.

Wen, C.Y. and Yu, Y.H. 1966. A generalized method for predicting the minimum fluidization velocity, *AIChE J.*, 12(2), 610–612.

3.4.3 Melting and Freezing

Vasilios Alexiades and Jan Kośny

Introduction

Phase-change processes and latent heat storage play an important role in many aspects of our lives, including materials manufacturing (casting of metals and plastics, crystal growth), food preservation (freezing and thawing, freeze dehydration), medical technologies (cryosurgery, cryopreservation of tissues), transportation (aircraft icing), powerline and pipeline icing, and energy conservation (latent heat storage), among many others.

Melting and freezing are among the most familiar processes in everyday life. Melting occurs when a solid is heated above its melting point T_m and changes phase from solid to liquid. At T_m, the solid and liquid phases coexist in thermodynamic equilibrium. The melting point of a pure substance depends on pressure; its value is usually specified at standard (atmospheric) pressure. Conversely, lowering the temperature of a liquid below T_m normally results in formation of a solid—the process is called freezing, solidification, or crystallization.

The primary quantities involved in describing melting/freezing are: thermal energy (enthalpy), which consists of sensible heat and latent heat, along with temperature, composition, and pressure in some cases.

In a solid, atoms are held together by strong intermolecular forces, and vibrate about stable positions. Atoms in a liquid have higher (thermal) energy, are only loosely bound, and vibrate much more. It takes a certain amount of energy, L Joules per gram, to break the bonds of a solid and render it a liquid. This energy L is called *latent heat* of melting.

Physically, a change of phase occurs due to loss of thermodynamic stability in favor of another, more favorable, phase. Below T_m, the solid configuration has lower free energy, so it is more stable. At T_m, solid and liquid have the same free energy, but their enthalpies differ by L. Above T_m, liquid has lower free energy, so it is favored thermodynamically and thus, a solid melts by absorbing the latent heat. The most prominent feature is the absorption or release of latent heat. The region where this occurs is the *interface*, where the phases coexist and across which the enthalpy jumps by L. Its thickness may be a few Angstroms or several centimeters, and its structure may be very complex, depending on several factors (composition, rate of cooling, temperature gradients, surface tension, supercooling, etc.).

Mathematically, problems involving unknown interfaces are called moving boundary problems (often called free boundary problems when time-independent). The location of the interface, being unknown, renders such problems de facto *nonlinear*, the major source of difficulty in treating them.

The latent heat concept was introduced by Joseph Black (1728–1799) who found it experimentally in 1758–1762. Lame-Clapeyron in 1831 incorporated latent heat into heat conduction, and in the early 1860s, Franz Neumann (1798–1895) found the similarity solution that bears his name ("The Stefan Problem" section). Jozef Stefan (1835–1893) clearly formulated the Stefan problem in 1889 while studying freezing of soil (discussed later in "The Stefan Problem" section).

The melting point T_m and the latent heat L are found experimentally, usually with differential scanning calorimetry (DSC). Pure crystalline solids melt at a well-determined temperature T_m, whereas amorphous solids and mixtures or alloys typically melt over a range of temperatures.

Solidification and Supercooling

Freezing, or solidification, the reverse of melting, is a phase transition process in which a liquid turns into a solid when its temperature is lowered below its freezing point. Ideally, the freezing point and heat of fusion of a melt would be the same as the melting point and latent heat of melting of the material. However, creating a stable crystalline structure is much harder than taking one apart,

so the phenomenon of *supercooling* often arises (sometimes also referred to as "undercooling"), whereas superheating is very rare.

At the *microscopic level*, crystallization proceeds by nucleation and subsequent crystal growth. Thus, the first step in the process of crystallization is *nucleation*, in which molecules start to form nano-sized clusters, arranging themselves in the lattice pattern that defines the crystal structure of the material being solidified. There is an energy barrier for (homogeneous) nucleation: if a nucleus is too small, the amount of energy bound in its volume may not be sufficient to create its surface. Thus, the liquid may supercool significantly below T_m without crystalizing. However, this is a thermodynamically *metastable* state; small perturbations will induce rapid nucleation. The presence of seed crystals or foreign particles, even dust, provides sites for (heterogeneous) nucleation, reducing the degree of supercooling. Following nucleation, growth of stable nuclei is governed by free energy minimization considerations, involving competition between volume and surface terms and other effects (surface tension, curvature, capillarity, Marangoni effects, etc.). The structure of the interface may be planar, columnar, dendritic, or amorphous, and difficult to model. The study of microstructural evolution and morphological instabilities is a large and still active field of research in materials science, as they impact the strength of the resulting solid and several of its properties (Kim, 1978; Zaeema and Mesarovicb, 2011; Colin, 2014). In the following sections, we concentrate on *macroscopic aspects* and do not discuss microstructural/morphological aspects any further.

Figure 3.4.13 depicts DSC measured data for *n*-octadecane paraffin. A small hysteresis between melting and freezing curves can be seen, the melting point is 27°C and the freezing point 26°C. Measuring the melting point T_m is more reliable and reproducible, whereas the freezing point is not considered to be a characteristic property of the material.

There are several inorganic substances (hydrated salts), called congruents, which melt and solidify cleanly, with the liquid and solid phases having the same chemical composition. In semi-congruent substances, a salt hydrate and an aqueous solution are formed during the fusion process. This reaction needs to be overturned before freezing to avoid volume reduction of the active phase changing material, which compromises the overall system thermal performance. Usually, semi-congruent melting substances can be modified with additives to render them congruent. Some combinations of an anhydrous salt and an aqueous solution can become incongruent during melting. Unless such material is modified, the anhydrous salt will settle down due to gravity, resulting in a gradual loss in heat storage performance. The most suitable methods of preventing this process are either through thickening the salt solution or through gelling. Thickening means addition of a material to the salt hydrate that increases the viscosity and hereby holds the salt hydrate together (Farid et al., 2004; Mehling and Cabeza, 2008). Gelling means adding a cross-linked material (e.g., polymer) to the salt

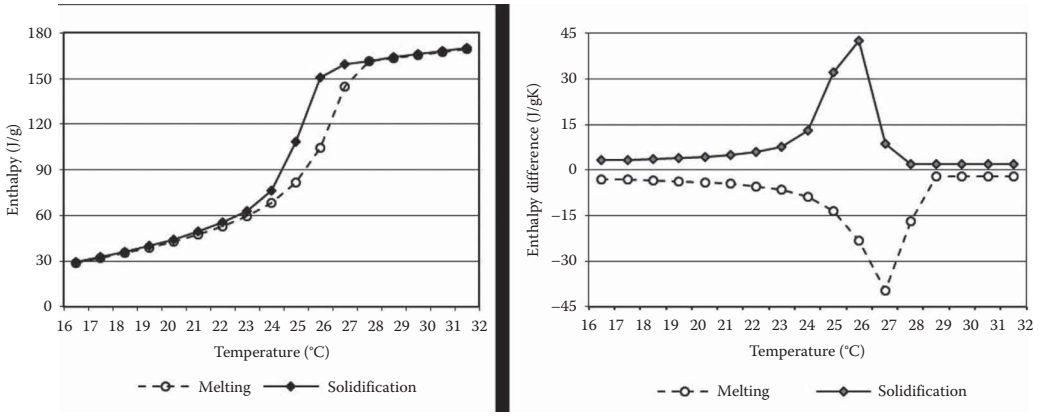


FIGURE 3.4.13 Differential scanning calorimeter (DSC) data for *n*-octadecane. Left: enthalpy vs. temperature. Right: enthalpy difference curves for melting and freezing.

to create a three-dimensional network that holds the salt hydrate together; hence, when anhydrous salt crystals are formed, phase separation does not occur.

Phase Diagrams

The thermodynamic state of a material is determined by its temperature *T*, pressure *P*, and composition *C*. Equality of temperatures ensures thermal equilibrium, equality of pressures ensures mechanical equilibrium, and equality of chemical potentials ensures chemical equilibrium. Phases are regions in (*T*, *P*, *C*)-space, where the material is thermodynamically stable. Phase transitions occur along boundaries between different phases. They are represented commonly by two-dimensional (2D) phase diagrams that are projections on the (*T, P*)-plane, or (*C, T*)-plane, and so on.

Figure 3.4.14 depicts a phase diagram that summarizes the effect of temperature and pressure on a pure substance in a closed container. The melting point of a substance depends on pressure and is usually specified at standard pressure. The phase diagram shown in the figure is divided into three main areas, which represent the solid, liquid, and gaseous states of the substance. The point *B*, called the triple point, represents the only combination of temperature and pressure at which a pure substance can exist simultaneously as a solid, a liquid, and a gas phase. The line between points *B* and *D*, representing solid–liquid equilibria, has, for most materials, a small positive slope. This is because melting points of most solids do not depend very much on changes in pressure. However, for water, the slope of this line is slightly negative because water can melt at temperatures near its freezing point when subjected to increased pressure. The line connecting points *A* and *B* represents solid–gas equilibria. At these temperatures and pressures, the rate at which the solid sublimes to form a gas is equal to the rate at which the gas turns into a solid. Similarly, the solid line between points *B* and *C* shows for gases and liquids the temperature dependence of the vapor pressure. Besides, in the top right corner of the figure, above the critical point, the supercritical fluid (SCF) area is shown. The critical point *C* represents the end point of a phase equilibrium curve. At the critical point, phase boundaries disappear. SCF is a phase at a temperature and pressure above the critical point, where distinct liquid and gas phases do not exist. This is because there is no surface tension in an SCF, as there is no liquid–gas phase boundary. That is why SCF can easily effuse through solids like a gas, and dissolve materials like a liquid.

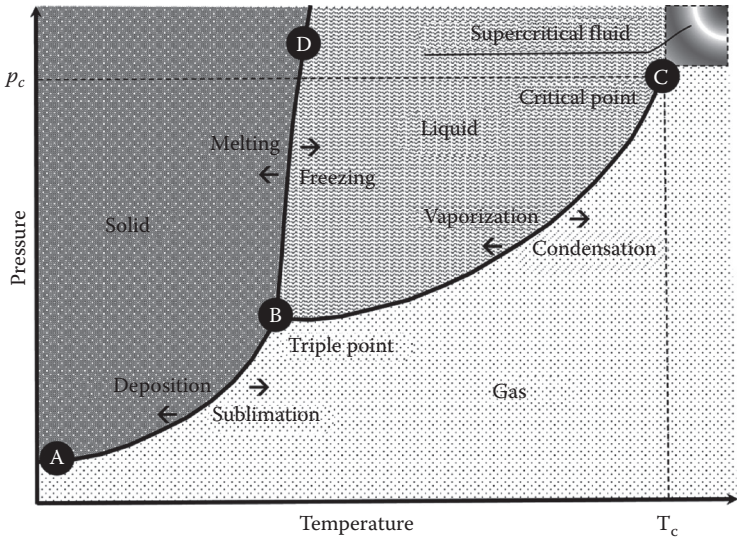


FIGURE 3.4.14 Schematic phase diagram showing the phases of a pure substance as a function of temperature and pressure.

In the general case of a multicomponent, multiphase system, the number of variables, f=degrees of freedom, needed to specify the thermodynamic state is given by the *Gibbs phase rule:* $f=2+N_{components}-N_{phases}$, (Lupis, 1983). Thus, for a pure material, $N_{components}=1$, in the liquid or solid phase: $f=2+1-1=2$, namely T and P are needed to specify the thermodynamic state, whereas along a phase coexistence curve: $f=2+1-2=1$, only one, either T or P suffices. On the other hand, for a binary system $N_{components}=2$ in the liquid or solid phase: $f=2+2-1=3$, namely T, P, C are needed, whereas in the mushy region of the phase diagram $N_{phases}=2, f=2$.

For multicomponent materials (alloys, blends, mixtures), the C–T phase diagram (at a fixed pressure) can be very complex, and the effect of composition is much more dramatic than that of pressure.

The simplest phase diagram for a binary alloy arises when the two components A and B are soluble in all proportions in each other, as is the case for the Si–Ge alloy shown in Figure 3.4.15.

The lens-shaped region between the solidus and the liquidus curves is a mushy region, containing states (C, T) that are mixtures of solid and liquid of different compositions coexisting at equilibrium. This miscibility gap arises from equality of chemical potentials of coexisting phases for each component, as required for a thermodynamic equilibrium (Lupis, 1983).

When cooling down a liquid at some composition C_o (% fraction of component B), solidification begins at $T_o=T^{liquidus}(C_o)$. But at that temperature, liquid of composition C_o is at equilibrium with solid at composition $C^{solidus}(T_o)$, which is what will precipitate out of solution at first. This enriches B in liquid and creates segregation of components, which will set up mass diffusion in the liquid; concurrently the released latent heat must be transferred away before the temperature can be reduced further. It is a coupled conduction–diffusion process. The solid forming will be progressively richer in B, of compositions determined by the lever rule. The mean composition remains at C_o, and the mushy material is *constitutionally supercooled*. This will continue till the temperature decreases down to $T^{solidus}(C_o)$. Thus, solidification occurs over the extended temperature range from $T^{liquidus}(C_o)$ down to $T^{solidus}(C_o)$. A more detailed discussion can be found in (Alexiades and Solomon, 1993, §2.5).

Other Effects

Many different physical phenomena *may* influence a melting process, and especially a freezing process, that can greatly complicate the situation. They include the following:

- Heat and mass transfer by conduction/convection/radiation,
- Possible gravitational, chemical, elastic, electromagnetic effects,

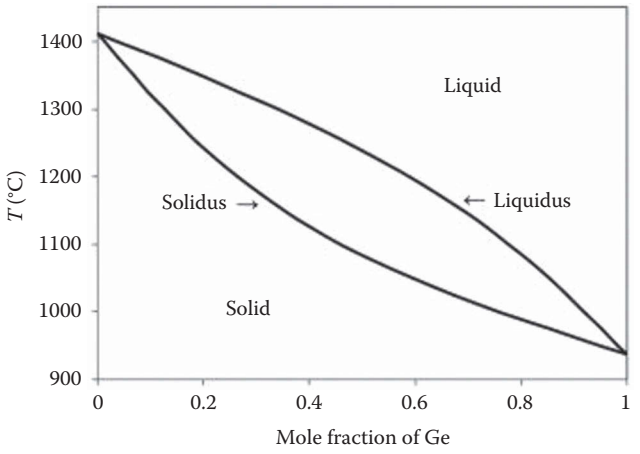


FIGURE 3.4.15 Phase diagram of a Si–Ge binary alloy.

- Variation of thermophysical properties,
- Variation of phase-change temperature,
- Density changes,
- Latent heat evolution over a range of temperatures (rather than at a single T_m),
- Nucleation difficulties, supercooling,
- Surface tension/capillarity/Marangoni, curvature effects.

This is a daunting list. The last two operate at microscopic scales and pertain only to the microstructure during solidification. The rest operate at macroscopic scales for any phase-change process. It is always necessary to make reasonable simplifying assumptions and focus only on essentials relevant to a specific investigation. It is better to start with basic ingredients and progressively incorporate more effects, if deemed necessary.

The core ingredients which give rise to the Stefan problem (discussed in the next section) and which are widely applicable in many situations are the following:

- Heat transfer by conduction only,
- Pure material,
- Constant thermophysical properties in each phase,
- Phase-change temperature is fixed at some known T_m,
- Density is constant and the same in both phases (else there will be movement, and hence convection, bringing in the Navier–Stokes equations!),
- Latent heat is released or absorbed at T_m,
- Microscopic aspects are ignorable.

With this in mind, in the next section, we will discuss the formulation and (exact) solution of the one-dimensional Stefan problem, which in fact serves as prototype of all phase-change problems. In "Analytical Approximations" section, some analytical approximations (quasi-stationary) and melt-time estimates are presented. Finally, in "Numerical Approaches" section, numerical methods are discussed, concentrating on the enthalpy formulation, its discretization, and algorithm.

The Stefan Problem

The prototype of all phase-change problems is the so-called *Stefan problem*, named after Jozef Stefan (1835–1893), even though it had been studied by Lame-Clapeyron in 1831, and an explicit solution had been found by Franz Neumann (1798–1895) in the 1860s.

The Stefan problem is a mathematical model for melting a solid, or freezing a liquid by heat conduction, assuming that the phases are separated by a (locally planar) *moving* front, whose position at time t, denoted by $x=X(t)$, is to be found along with the temperature field $T(x, t)$. The phase-change front (interface) is assumed to be sharp, a surface of zero thickness, along which the temperature is the melting temperature T_m and where the latent heat is absorbed or released.

As mentioned at the end of "The Enthalpy Method" section, the major physical simplifying assumptions are: pure material, constant density (and same in both phases), changing phase at a known T_m, with sharp interface.

Analytical solutions for phase-change processes in pure substances are discussed in Evans (1951), Carslaw and Jaeger (1959), Rubinstein (1971), Crank (1984), Hill (1987), Alexiades and Solomon (1993), among others. Today, the solution of the Stefan problem is widely considered as one of the most used analytical solutions for one-dimensional solid–liquid phase transition.

One-Phase Stefan Problem

The simplest possible phase-change problem involves melting of a phase-change material (PCM) that is initially *solid at its melting temperature* T_m, by imposing a temperature $T_L > T_m$ on its surface, in one space dimension (a rod or a slab, insulated laterally).

Physical Problem: Slab at T_m Melting from the Left

Consider a slab $0 \le x \le \ell$ of phase-change material whose melting temperature is T_m and latent heat of melting L. Initially, the PCM is *solid at temperature* T_m. Melting is induced by imposing a temperature $T_L > T_m$ at the left face $x=0$ (while keeping the right face $x=\ell$ at T_m).

A melt front $x=X(t)$ appears at $x=0$, which advances into the solid while the liquid is getting hotter. At any $t > 0$, the region $0 \le x \le X(t)$ is liquid at temperature $T(x, t)$; the rest remains solid at $T=T_m$. We assume constant thermophysical properties: density ρ, latent heat L, specific heat c_L, and thermal conductivity k_L (hence constant thermal diffusivity $\alpha_L = k_L/\rho c_L$).

The mathematical problem modeling this process is the following.

One-Phase Stefan Problem (for a Slab Initially at T_m Melting from the Left)

Find the interface location $X(t)$, $t > 0$, and the temperature $T(x, t)$, $0 \le x \le X(t) t > 0$, such that:

$$T_t = \alpha_L T_{xx} \quad \text{for } 0 < x < X(t),\ t > 0 \quad \text{(in liquid)} \tag{3.4.40}$$

$$T(X(t),t) = T_m,\ t > 0, \quad \text{(interface temperature } = T_m) \tag{3.4.41}$$

$$\rho L X'(t) = -k_L T_x(X(t),t) t > 0, \quad \text{(Stefan condition)} \tag{3.4.42}$$

$$X(0) = 0 \quad \text{(initially all solid)} \tag{3.4.43}$$

$$T(0,t) = T_L > T_m, \quad t > 0 \quad \text{(imposed temperature } = T_L) \tag{3.4.44}$$

Subscripts "t" and "x" denote partial derivative with respect to t and x, and $X'(t) = \mathrm{d}X/\mathrm{d}t$ in Equation 3.4.42 is the speed of the interface. The *Stefan condition* (Equation 3.4.42), expressing conservation of energy *across* the interface, is the additional condition needed to find the additional unknown $X(t)$. It says that the latent heat absorbed at the front equals the jump of heat flux $[[q]]_S^L \equiv q_L - q_S = -k_L T_x - 0$ across the front (here heat flux in solid $q_S = 0$ since $T \equiv T_m$); see Equation 3.4.51 discussed later.

Note that since the solid remains at T_m, the back face $x=\ell$ plays no role here, so the slab may be considered as extending to infinity. The problem admits a (similarity) solution, known as the *Neumann solution*, provided all parameters and data are constants (Alexiades and Solomon, 1993, p. 35). It is expressed in terms of a single parameter λ, which is the root of Equation 3.4.47.

Interface location:

$$X(t) = 2\lambda\sqrt{\alpha_L t}, \quad t > 0 \tag{3.4.45}$$

Temperature in liquid region $0 < x < X(t)$, $t > 0$:

$$T(x,t) = T_L - (T_L - T_m)\frac{\operatorname{erf}\left(\dfrac{x}{2\sqrt{\alpha_L t}}\right)}{\operatorname{erf}(\lambda)}. \tag{3.4.46}$$

The parameter λ is the unique solution to the transcendental equation

$$\lambda e^{\lambda^2} \operatorname{erf}(\lambda) = \frac{\mathrm{St}_L}{\sqrt{\pi}}, \text{ where } \mathrm{St}_L = \frac{c_L(T_L - T_m)}{L} = \text{Stefan number}, \tag{3.4.47}$$

with erf() as the error function.

For each $\mathrm{St}_L > 0$, the transcendental equation (Equation 3.4.47) has unique solution $\lambda > 0$, which depends only on the (dimensionless) parameter St_L. It can be computed easily by any root finder, for example, bisection or the Newton–Raphson rule.

The *Stefan number* St is the ratio of *sensible heat* $c\Delta T$ to *latent heat* L, and it completely characterizes the one-phase Stefan problem (as can be seen by its undimensionalization, (Alexiades and Solomon, 1993, p. 37). Small St signifies that latent heat dominates, whereas large St indicates that sensible heat dominates. Water and certain waxes have high latent heat, while metals have low latent heat. In particular, water has very high $L=333.4$ J/g, so it is an excellent heat storage material. For example, in the melting of ice at $T_\mathrm{m}=0°\mathrm{C}$ by holding it at body temperature $T_L=37°\mathrm{C}$, we have $\mathrm{St}_L=0.46$ (for water, $c_L=4.2$ J/g K). Similarly, in freezing food from $T_\mathrm{m}=0$ to $T_S=-20°\mathrm{C}$, $c_S\approx 2$ J/g K and so $\mathrm{St}_S=c_S(T_\mathrm{m}-T_S)/L=0.12$.

The Neumann solution (Equations 3.4.45 through 3.4.47) is in fact the *unique* solution of the one-phase Stefan problem. It shows that the front location is proportional to the square root of time, so it starts with infinite speed at $t=0$ (to accommodate the initial discontinuity at $x=0$). Solving Equation 3.4.45 for t gives the time for the melt front to reach a specific location $x=x_\mathrm{o}$: $t_\mathrm{melt}(x_\mathrm{o}) = x_\mathrm{o}^2 / (4\lambda^2\alpha_L)$.

For small $S_L\approx 0$, the root λ is also small and to first order $\lambda \approx \sqrt{(\mathrm{St}_L / 2)}$. This provides a good starting value for the Newton–Raphson iteration. Moreover, the ratio of erf()s in Equation 3.4.46 is $\approx x/X(t)$, so the Neumann temperature becomes $T(x,\, t)\approx T_L-(T_L-T_\mathrm{m})x/X(t)$. Thus, at each time t, the temperature is a line joining the point $(x=0,\, T=T_L)$ with $(x=X(t),\, T=T_\mathrm{m})$.

The analogous *freezing* problem and its solution can be obtained by simply replacing the subscript "L" by "S" and the latent heat L by $-L$. In this case, the imposed temperature will be $T_S < T_\mathrm{m}$ and the Stefan number for the solid will be $\mathrm{St}_S=(T_\mathrm{m}-T_S)/L$.

Two-Phase Stefan Problem

A more realistic case is the melting of a solid initially colder than T_m. In such a situation, there will be heat transfer in both phases.

Physical Problem: Slab Melting from the Left

Consider a slab $0\le x\le \ell$ of phase-change material whose melting temperature is T_m and latent heat of melting is L. Initially, the material is solid at some temperature $T_S < T_\mathrm{m}$. Melting is induced by imposing an elevated temperature $T_L > T_\mathrm{m}$ at the left face ($x=0$), while keeping the right face ($x=\ell$) insulated. We assume constant thermophysical properties: density ρ, latent heat L, specific heats c_L, c_S, thermal conductivities k_L, k_S (hence, constant thermal diffusivities $\alpha_L=k_L/\rho c_L$ and $\alpha_S=k_S/\rho c_S$, and a sharp interface $x=X(t)$.

At each time t, liquid occupies $[0,\, X(t))$ and solid $(X(t),\, \ell]$, separated by the melt front $x=X(t)$. Conservation of (thermal) energy *in each phase* is expressed by the heat conduction equation for temperature $T(x,\, t)$; the Stefan condition expresses conservation of energy *across* the interface. The mathematical problem modeling this process is the following.

Two-Phase Stefan Problem (for a Slab Melting from the Left)

Find the temperature $T(x,\, t)$, $0\le x\le \ell$, $t > 0$, and interface location $X(t)$, $t > 0$, such that:

$$T_t = \alpha_L T_{xx} \quad \text{for } 0 < x < X(t), t > 0 \quad \text{(in liquid)} \tag{3.4.48}$$

$$T_t = \alpha_S T_{xx} \quad \text{for } X(t) < x < \ell, t > 0 \quad \text{(in solid)} \tag{3.4.49}$$

$$T(X(t),t) = T_\mathrm{m}, \quad t > 0, \quad \text{(interface temperature} = T_\mathrm{m}) \tag{3.4.50}$$

$$\rho L X'(t) = -k_L T_x(X(t),t) + k_S T_x(X(t),t), \quad t > 0, \quad \text{(Stefan condition)} \tag{3.4.51}$$

$$X(0) = 0 \quad \text{(initially all solid)} \tag{3.4.52}$$

$$T(x,0) = T_S < T_\text{m}, \quad 0 \le x \le \ell \quad \text{(initial temperature } = T_S\text{)} \tag{3.4.53}$$

$$T(0,t) = T_L > T_\text{m}, \quad t > 0 \quad \text{(imposed temperature} = T_L\text{)} \tag{3.4.54}$$

$$-k_S T_x(\ell,t) = 0, \quad t > 0 \quad \text{(insulated boundary)} \tag{3.4.55}$$

This problem (Equations 3.4.48 through 3.4.56) does *not* admit an exact analytical solution. However, considering it on the semi-infinite slab [0, ∞) (with the boundary condition (Equation 3.4.55) replaced by $\lim_{x\to\infty} T(x, t) = T_S$), the problem *does* admit a (similarity) solution, provided all parameters and data are constants (Alexiades and Solomon, 1993, p. 47).

Neumann Solution of the Two-Phase Stefan Problem
Interface location:

$$X(t) = 2\lambda\sqrt{\alpha_L t}, \quad t > 0 \tag{3.4.56}$$

Temperature in liquid region $0 < x < X(t)$, $t > 0$:

$$T(x,t) = T_L - (T_L - T_\text{m})\frac{\text{erf}\left(\dfrac{x}{2\sqrt{\alpha_L t}}\right)}{\text{erf}(\lambda)}, \tag{3.4.57}$$

Temperature in solid region $x > X(t)$, $t > 0$:

$$T(x,t) = T_S + (T_\text{m} - T_S)\frac{\text{erf}\left(\dfrac{x}{2\sqrt{\alpha_S t}}\right)}{\text{erfc}\left(\lambda\sqrt{\alpha_L / \alpha_S}\right)}. \tag{3.4.58}$$

The parameter λ is the solution to the transcendental equation

$$\frac{\text{St}_L}{\exp(\lambda^2)\text{erf}(\lambda)} - \frac{\text{St}_S}{\nu\exp(\nu^2\lambda^2)\text{erfc}(\nu\lambda)} = \lambda\sqrt{\pi}, \tag{3.4.59}$$

where

$$\text{St}_L = \frac{c_L(T_L - T_\text{m})}{L}, \quad \text{St}_S = \frac{c_S(T_\text{m} - T_S)}{L}, \quad \nu = \sqrt{\frac{\alpha_L}{\alpha_S}}, \tag{3.4.60}$$

and erfc() = 1 − erf(), the complimentary error function.

Note that the value of λ depends only on the two *Stefan numbers* St_L, St_S, and the diffusivity ratio ν. The (dimensionless) transcendental equation (Equation 3.4.59) has exactly one root λ for each $\text{St}_L > 0$, $\text{St}_S \ge 0$, $\nu > 0$ (Alexiades and Solomon, 1993, p. 48)), which can easily be found using a root finder.

Small $\text{St}_L \approx 0$ implies small λ, and then to first-order Equation 3.4.59, it is approximately $\text{St}_L / (2\lambda^2) - \text{St}_S / (\nu\lambda\sqrt{\pi}) = 1$, whence

$$\lambda \approx \frac{1}{2}\left[-\frac{\text{St}_S}{\nu\sqrt{\pi}} + \sqrt{2\text{St}_L + \left(\frac{\text{St}_S}{\nu\sqrt{\pi}}\right)^2}\right] \quad \text{for } \text{St}_L \approx 0. \tag{3.4.61}$$

If, in addition, $St_S \gg St_L \approx 0$, Equation 3.4.61 simplifies to $\lambda \approx (\nu\sqrt{\pi}/2)(St_L / St_S)$.

Setting $T_S = T_m$, whence $St_S = 0$, the 2-phase Neumann solution reduces to the one-phase solution.

The analogous *freezing* problem and its solution is obtained by simply exchanging the subscripts "*L*" and "S" and replacing the latent heat *L* by –*L*.

Analytical Approximations

Exact explicit solutions, presented in "The Stefan Problem" section, provide complete information on how parameters influence the solution and what one is to expect, but exists only for the simplest Stefan problems, only in semi-infinite problems, and only for parameters constant in each phase. Moreover, very few explicit solutions are known in cylindrical or spherical geometries, and none for finite domains and higher dimensions. Thus, for more realistic problems, one must seek approximate solutions, of which there are two classes: *analytical approximations*, discussed in this section, and *numerical approximations*, discussed in the next section.

The four main types of analytical approximations for phase-change problems in use are the *quasi-stationary approximation*, the *Megerlin method*, the *heat balance integral*, and the broad class of *perturbation expansions*. The last one is a general classical methodology of applied mathematics that seeks solutions in the form of a series expansion in some small parameter entering the problem.

The applicability of the other methods relies on being able to simplify the problem to a form that fits the method, which then can generate solvable problems. The resulting solution may turn out to be quite accurate, at least for standard, simple processes. However, there is no way of checking the validity of physical simplifications made, and there are no a priori error estimates for the mathematical approximations. The validity of an analytical approximation may be judged only by comparing it with some other independently validated method and only for a class of very similar problems. Thus, at best, they can provide only qualitative "order of magnitude" information rather than quantitative results, and they are limited to only one-phase problems in one dimension. Nevertheless, when they work, they are very useful for rough sizing estimates of melt depth, melt time, heat flux, etc.

It should be noted that nowadays, in view of the versatility and relative ease of applying direct computational methods, only simple, easy to apply and easy to evaluate analytical solutions are really useful. But in fact, being analytic expressions, they can reveal qualitative behavior, e.g., dependence on parameters, in ways that numerical calculations cannot.

A word of warning on terminology: sometimes in the literature, analytical approximations are advertised as "exact solutions" because an approximate problem is solved "exactly."

We will present only the basics of the quasi-stationary approximation and some useful melt time estimates. The interested reader can refer to (Ozisik, 1993; Alexiades and Solomon, 1993, ch. 3) for more details and other methods.

The quasi-stationary (or quasi-static) approximation is the simplest of the analytical approximation, well suited for "back of the envelope" sizing estimates. It can be applied only to one-phase Stefan problems, but at any of the standard boundary conditions at $x = 0$.

The approximation is based on the physical assumption that *the sensible heat in the process is negligible compared to the latent heat*. This amounts to assuming a small Stefan number: $St \approx 0$

The quasi-stationary approximation consists of replacing the heat equation by the steady-state equation $T_{xx} = 0$, while the front still varies in time. Thus, $T(x, t)$ is linear in x and the temperature profile at any time can be found using the boundary conditions. Then the front location is found from the Stefan condition. We present the approximations for the standard boundary conditions here.

Quasi-Stationary Approximation for Imposed Temperature

The quasi-stationary approximation of the one-phase Stefan problem with imposed temperature at $x = 0$, which may be varying in time: $T(0, t) = T_L(t)$, is the following (Alexiades and Solomon, 1993, p. 128):

$$X^{\mathrm{qs}}(t)=\left(2\frac{k_L}{\rho L}\int_0^t [T_L(s)-T_{\mathrm{m}}]\,\mathrm{d}s\right)^{1/2},\quad t\ge 0, \tag{3.4.62}$$

$$T^{\mathrm{qs}}(x,t)=T_L(t)-[T_L(t)-T_{\mathrm{m}}]\frac{x}{X^{\mathrm{qs}}(t)},\quad 0\le x\le X^{\mathrm{qs}}(t), t\ge 0. \tag{3.4.63}$$

When $T_L(t)\equiv T_L$=constant, this simplifies to

$$X^{\mathrm{qs}}(t)=2\sqrt{\mathrm{St}_L/2}\sqrt{\alpha_L t},\quad t\ge 0, \tag{3.4.64}$$

$$T^{\mathrm{qs}}(x,t)=T_L(t)-[T_L(t)-T_{\mathrm{m}}]\frac{\dfrac{x}{2\sqrt{\alpha_L t}}}{\dfrac{}{2\sqrt{\mathrm{St}_L/2}}},\quad 0\le x\le X^{\mathrm{qs}}(t), t\ge 0, \tag{3.4.65}$$

which is in fact the Neumann solution for the case $\mathrm{St}_L\approx 0$. It follows that the quasi-stationary front $X^{\mathrm{qs}}(t)$ *overestimates* the actual front $X(t)$ and also $T^{\mathrm{qs}}(x, t) > T(x, t)$.

Quasi-Stationary Approximation for Imposed Flux

For imposed flux boundary conditions at $x=0$, namely, $-k_L\,\partial T/\partial x(0,t)=q_L(t)>0$, for a given $q_L(t)$, there is no explicit solution available, even for constant q_L. The quasi-stationary solution is easily found to be (Alexiades and Solomon, 1993, pp. 134–136)

$$X^{\mathrm{qs}}(t)=\frac{1}{\rho L}\int_0^t q_L(s)\mathrm{d}s,\quad t\ge 0 \tag{3.4.66}$$

$$T^{\mathrm{qs}}(x,t)=T_{\mathrm{m}}+\frac{q_L(t)}{k_L}\left[X^{\mathrm{qs}}(t)-x\right],\quad 0\le x\le X^{\mathrm{qs}}(t),\quad t\ge 0 \tag{3.4.67}$$

Example: *Constant flux* $q_L(t)\equiv q_L$. For constant imposed flux, the quasi-stationary temperature is linear in both x and t, the face temperature $T^{\mathrm{qs}}(0, t)$ grows linearly in time, and the melt time of a location x is $t^{\mathrm{qs}}_{\mathrm{melt}}(x)=\rho L\,x/q_L$.

Example: *Sinusoidal flux* $q_L(t)=q_{\max}\sin\left(\pi t/2t_{\mathrm{o}}\right)$. The solution is easily found, and one can check that after a full cycle $2t_{\mathrm{o}}$,

$$X^{\mathrm{qs}}(2t_{\mathrm{o}})=\frac{4t_{\mathrm{o}}q_{\max}}{\rho L\pi},\quad T^{\mathrm{qs}}(x,2t_{\mathrm{o}})=T_{\mathrm{m}}\ \text{(but liquid)}. \tag{3.4.68}$$

Quasi-Stationary Approximation for Convective Boundary Condition

The boundary condition for convective heating at $x=0$ by an ambient temperature $T_L(t) > T_{\mathrm{m}}$, with heat transfer coefficient h, is

$$-k_L\frac{\partial T}{\partial x}(0,t)=h[T_L(t)-T(0,t)],\quad t>0. \tag{3.4.69}$$

The quasi-stationary solution of the Stefan problem turns out to be (Alexiades and Solomon, 1993, pp. 137–139)

$$X^{\mathrm{qs}}(t)=\frac{-k_L}{h}+\left(\frac{k_L^2}{h^2}+2\frac{k_L}{\rho L}\int_0^t [T_L(s)-T_{\mathrm{m}}]\,\mathrm{d}s\right)^{1/2},\quad t\ge 0 \tag{3.4.70}$$

$$T^{\mathrm{qs}}(x,t) = T_{\mathrm{m}} + [T_L(t) - T_{\mathrm{m}}]\frac{h[X^{\mathrm{qs}}(t) - x]}{hX^{\mathrm{qs}}(t) + k_L}, \quad 0 \le x \le X^{\mathrm{qs}}(t), \quad t \ge 0 \tag{3.4.71}$$

and the temperature and flux at the face $x=0$ are

$$T^{\mathrm{qs}}_{\mathrm{face}}(t) = \frac{T_{\mathrm{m}} + \dfrac{hX^{\mathrm{qs}}(t)}{k_L}T_L(t)}{1 + \dfrac{hX^{\mathrm{qs}}(t)}{k_L}}, \quad t \ge 0, \tag{3.4.72}$$

$$q^{\mathrm{qs}}_{\mathrm{face}}(t) = \frac{k_L}{X^{\mathrm{qs}}(t)}[T^{\mathrm{qs}}_{\mathrm{face}}(t) - T_{\mathrm{m}}], \quad t \ge 0. \tag{3.4.73}$$

Solutions for axially or radially symmetric processes has been found by Alexiades and Solomon (1993, pp. 144–152).

Melt Time for a Simple PCM Body with Imposed Temperature

Consider a PCM initially solid at T_{m}, being melted due to a high temperature $T_L > T_{\mathrm{m}}$ imposed on its boundary. Assume that the melting process can be described in terms of a single *effective length* r, $0 \le r \le \ell$. If A is the surface area across which heat is transferred into the PCM and V its volume, define the *shape factor* ω by

$$1 + \omega = \ell\, A / V\,. \tag{3.4.74}$$

Then, $\omega=0$ for a PCM slab or rod of length ℓ insulated at one end, $\omega=1$ for a PCM cylinder of radius ℓ heated radially, and $\omega=2$ for a PCM sphere of radius ℓ. Note that $0 \le \omega \le 2$ always.

The one-phase (inward) melting process, for any $0 \le \omega \le 2$, may be formulated as follows:

$$\frac{\partial T}{\partial t} = \frac{\alpha_L}{r^{\omega}}\frac{\partial}{\partial r}\left(r^{\omega}\frac{\partial T}{\partial r}\right), \quad R(t) \le r \le \ell, \quad t > 0, \tag{3.4.75}$$

$$T(R(t),t) = T_{\mathrm{m}}, \quad \rho L = -k_L(R(t),t), \quad t > 0 \tag{3.4.76}$$

$$R(0) = \ell, \quad T(\ell,t) = T_L, \quad t > 0. \tag{3.4.77}$$

The PCM melts completely when the melt front $R(t)$ reaches $r=0$. Letting

$$\Delta T_L = T_L - T_{\mathrm{m}} \quad \text{and} \quad \mathrm{St}_L = c_L\,\Delta T_L / L, \tag{3.4.78}$$

the melt time of the PCM may be approximated by Alexiades and Solomon (1993, p. 169)

$$t_{\mathrm{melt}} \approx \frac{\ell^2}{2\alpha_L(1+\omega)\mathrm{St}_L}\left[1 + \left(025 + 017\omega^{0.7}\right)\mathrm{St}_L\right], \text{ valid for } 0 \le \mathrm{St}_L \le 4, \tag{3.4.79}$$

and the *average* heat flux at $r=\ell$ during the entire melting process by

$$\overline{q} \approx \frac{-2k_L\Delta T_L}{\ell}\left[1 + (0121 + 0.0424\,\omega)\mathrm{St}_L^{07645-02022\,\omega}\right]. \tag{3.4.80}$$

The *freezing time* of the analogous process is obtained by replacing the subscript L by S and the latent heat L by $-L$.

Note that these expressions, easily computable on a calculator or a spreadsheet, allow one to estimate correlations among parameters. Also note that the melt depth can be found by solving Equation 3.4.79 for ℓ.

Melt Time for a Simple PCM Body with Convective Boundary Conditions

For convective heating with heat transfer coefficient h and ambient temperature T_L, the boundary condition in Equation 3.4.77 reads as

$$-k_L \frac{\partial T}{\partial x} = h[T_L - T(\ell,t)], \quad t > 0, \tag{3.4.81}$$

and the melt time estimate (Equation 3.4.79) becomes

$$t_{\text{melt}} \approx \frac{\ell^2}{2\alpha_L(1+\omega)\text{St}_L}\left[1 + \frac{2}{\text{Bi}} + \left(025 + 017\,\omega^{0.7}\right)\text{St}_L\right], \tag{3.4.82}$$

valid for $0 \le \text{St}_L \le 4$, $\text{Bi} \ge 0.1$, where $\text{Bi} = h\ell/k_L$ is the Biot number of the process.

A very useful aspect of these melt time estimates is that we can explore the influence of parameters, especially of the shape and size of the PCM, by varying ω and ℓ.

Numerical Approaches

The literature on numerical methods for phase-change problems is huge. Currently, a Google search produces 864,000 results!

Phase-change processes are quite complex due to the presence of a moving boundary or region, in which heat and mass balance conditions have to be met—this renders the problem non-linear, and, in addition, the simulated material often has different thermophysical characteristics in different phases (Lamberg et al., 2004).

Numerical methods for addressing phase transition problems have been reviewed in many publications, including: Basu and Date (1988), Idelsohn et al. (1994), Hu and Argyropoulos (1996), Voller et al. (2006), Dutil et al. (2011), AL-Saadi and Zhai (2013). The earliest computational techniques for solving phase transition problems are an enthalpy-type method (Dusinberre, 1945) and a heat source method (Eyres et al., 1946).

Presently used direct numerical simulations methods of phase transition processes can be classified into two categories. The first category, often referred to as *front tracking or multi-domain methods*, uses independent conservation equations in each phase region and couples them with appropriate interface conditions at the (isothermal) moving front. The second category, known as *front capturing or fixed-domain methods*, uses a single region (continuum) formulation in which the interface does not appear explicitly, but can be "captured" from the solution. Over the years, various hybrid methods have also been developed that attempt to combine aspects of both categories in order to take advantage of the best features of each methodology.

Front Tracking Methods

Front tracking methods, also called interface tracing, apply to problems that (are assumed to) have a sharp interface whose precise location is considered crucial and should be tracked. This situation arises in pure materials without internal heating (classical Stefan problems), and in (microscopic) studies of morphological instabilities, among others.

Front tracking techniques follow the dynamic evolution of a moving interface across which physical variables are discontinuous. Usually they employ finite difference or finite-volume discretization of the heat equation; the numerically defined interface is propagated by a set of topologically connected marker points so as to satisfy an explicitly imposed interface (Stefan) condition. However, it is difficult to handle topological changes such as merging fronts or appearing and disappearing phases.

Various types of tracking methods have been developed over the years, such as the following:

a. Fixed grid (Lazaridis, 1969; Rao and Sastri, 1984);
b. Variable time step (Douglas and Gallie, 1955);

c. Variable space grid (Murray and Landis, 1959);
d. Front-fixing by coordinate transformations (Crank and Gupta, 1975; Hsiao, 1985);
e. Boundary integral formulations (Chuang and Szekely, 1972; ONeill, 1983);
f. Lagrangian-type adaptive meshes (Bonnerot and Janet, 1977; Lynch, 1982; Albert and ONeill, 1986; Zabaras and Ruan, 1990);
g. Front tracking for shocks and flames (Glimm et al., 1998)
h. Shyy's ELAFINT (Eulerian–Lagrangian algorithm for interface tracking (Shyy et al., 1996).

It should be noted that front tracking faces several limitations that complicate the solution procedure and restrict its applicability. These include the following: (a) The very formulation of the problem *assumes a priori* that there is a sharp front to be tracked, which may or may not be the case physically. (b) When the interface forms a singularity or changes its topology, tracking becomes difficult to continue beyond the singularity time. (c) It is computationally demanding and limited to modest deformation, especially when interface motion also includes expansion and contraction. The effort for significant topological changes in Lagrangian schemes seems to have reached its limit. To overcome such limitations, alternative methods arose from front capturing techniques, including *immersed boundary* (Uverdi and Tryggvason, 1992), and *level-set-type methods*, discussed in the next section.

Front Capturing Methods

In contrast to front tracking, front capturing methods solve the conservation law(s) over the entire region (posed in a weak sense mathematically), irrespective of phase, and the interface does not appear explicitly in the formulation. Its location can be "captured" from the solution in various ways. Field variables and parameters may change discontinuously between phases, and interfaces may be either sharp or entire (mushy) regions. In this approach, general hydrodynamics (Navier–Stokes + energy balance equations) can be solved on a *fixed Eulerian grid* rather efficiently by treating both phases as a single fluid.

Front capturing numerical approaches fall into two main classes:

1. *Volume of fluid (VoF)-type methods*, primarily the *enthalpy method* (presented in detail in "The Enthalpy Method" section). They arose from, and are very similar to, shock capturing schemes in computational fluid dynamics. The interface does not appear explicitly in the formulation, so no particular structure is assumed about it a priori. It may be a sharp front or a mushy region of any extent. The volume fraction of liquid (or solid), found from conservation laws and appropriate equations of state, is used as a phase indicator to capture the location and extent of phases at each time step. The formulations are general and applicable to complex multiphysics problems, such as multiphase flows, bubble flows, etc.
2. *Level set and phase field methods* (Sethian, 1999; Osher and Fedkiw, 2003). They are concerned with precise evolution and structure of interfaces, so they are appropriate in microscopic settings when the evolution of the front needs to be followed. They can handle topological changes easily on fixed Eulerian grids. The interface is represented as the (zero) level set of a smooth auxiliary function $\phi(x, t)$, so it is possible to compute the normal direction: $\vec{n} = \nabla\phi / |\nabla\phi|$, normal velocity: $\vec{v} = d\vec{x}/dt \cdot \vec{n}$, and mean curvature $\kappa = \nabla \cdot \vec{n}$ from ϕ. Thus, surface tension and curvature effects can be represented in detail.

The phase field and level set methods differ primarily in the interpretation and determination of the phase function ϕ.

In the *level set method*, ϕ is usually taken as distance from the front and it is updated from a kinematic advection equation (but then needs to be re-initialized, which is costly).

In the *phase field approach*, which arose from material science (Cahn–Hilliard equations) considerations, the function ϕ is an order parameter, normalized to be $=\pm 1$ in the bulk phases away from the front. The front is "diffuse": $-1 < \phi < 1$, and its width is controlled by a parameter in the model. The phase field ϕ can be chosen in various ways, usually as a solution of a Cahn–Hilliard-type equation, namely, an Euler–Lagrange equation for the extrema of a (configurational) free energy $J(\phi)$ containing surface and volume terms, and into which various desirable physical effects can be incorporated. Many variants and improvements exist to overcome various issues. Notable are "sharp interface" versions, where ϕ is chosen, e.g., as a hyperbolic tangent of the normal vector $\vec{n}$ (Sun and Beckermann, 2007). For a review, see Qin and Bhadeshia (2010). A comparison of some moving grid, level set, and phase field methods for the Stefan problem has been reported by Javierre et al. (2006).

The Enthalpy Method

The so-called *enthalpy* method is the simplest and most physical, versatile, adaptable, and robust numerical method available for phase-change problems in one, two, or three space dimensions (Alexiades and Solomon, 1993, Chap. 4). It is the basis on which most front capturing and many front tracking schemes are based. It is similar to the VoF and shock capturing schemes in computational fluid dynamics. It is based directly on the energy conservation law

$$\frac{\partial E}{\partial t} + \nabla \cdot Q = S\,, \tag{3.4.83}$$

with $E = \rho e$ the thermal energy (enthalpy per unit volume), Q the (heat) flux, and S any internal (volumetric) energy source or sink. For heat conduction, Q is given by Fourier's law

$$Q = -k\,\nabla T\,, \tag{3.4.84}$$

with T the temperature and k the thermal conductivity, which may be different in different phases and may also depend on T. The enthalpy formulation is based on two important facts:

- The Stefan condition is a "natural" interface condition: it is automatically satisfied by a (weak) solution of the energy conservation law (Equation 3.4.83) along any surface (or region) across which E experiences a jump.
- The recognition that temperature is a poor descriptor of phase, it only captures *sensible heat* but not latent heat, and it stays at T_m during absorption or release of latent heat. The natural and best descriptor is (thermal) energy itself.

By solving the partial differential equation (Equation 3.4.83) *throughout the region* (irrespective of phase), the energy is updated; we then need to find the phase and the temperature T, using the appropriate *equation of state* (EoS) relating energy and temperature. The location of the (moving) phase-change front does *not* appear explicitly in the formulation; it is found (captured) from the solution as the region where the energy jumps by an amount equal to the latent heat. It may be a sharp surface or a region; no a priori assumption about its structure is made. This is the major advantage of the enthalpy formulation over front tracking methods.

The difference between heat transfer without and with phase change is that in the latter case, the energy jumps by L (= latent heat) at T_m, whereas it is continuous in the former case. In heat conduction, the EoS is $de/dT = c_p =$ specific heat, or per unit volume $dE/dT = C_p := \rho c_p$, i.e., $E(T) =$ integral of the volumetric heat capacity $C_p(T)$. But during a change of phase, the latent heat is evolved at $T = T_m$ and the energy jumps by ρL at T_m. The EoS now is (referring to the energy of the *solid* at T_m):

$$E(T)=\begin{cases}\displaystyle\int_{T_{\mathrm{m}}}^{T} C_p^S(\tau)\,\mathrm{d}\tau \approx C_p^S[T-T_{\mathrm{m}}], & T<T_{\mathrm{m}} \quad \text{(solid)},\\ [0,\rho L], & T=T_{\mathrm{m}} \quad \text{(interface)},\\ \rho L+\displaystyle\int_{T_{\mathrm{m}}}^{T} C_p^L(\tau)\,\mathrm{d}\tau \approx \rho L + C_p^L[T-T_{\mathrm{m}}], & T>T_{\mathrm{m}} \quad \text{(liquid)}.\end{cases} \tag{3.4.85}$$

In the region where $0\le E\le \rho L$, the phases coexist; it is where the latent energy is absorbed or released. This is the "interface," which is not necessarily a sharp surface; it may be a region, often called the *mushy region*.

Note that since C_p^S and C_p^L are strictly positive, $E(T)$ depends monotonically on T, so it is invertible, i.e., it can be solved for T in terms of E.

The *liquid fraction* $\lambda=0$ in solid, $\lambda=E/\rho L$ in mushy, $\lambda=1$ in liquid is often used as a convenient phase indicator. Thus, we have the following characterizations of phases:

$$\begin{array}{lcccccl}\text{Solid} & \Leftrightarrow & E<0 & \Leftrightarrow & \lambda=0 & \Leftrightarrow & T<T_{\mathrm{m}},\\ \text{Mushy} & \Leftrightarrow & 0\le E\le \rho L & \Leftrightarrow & \lambda=E/\rho L & \Leftrightarrow & T=T_{\mathrm{m}},\\ \text{Liquid} & \Leftrightarrow & \rho L<E & \Leftrightarrow & \lambda=1 & \Leftrightarrow & T>T_{\mathrm{m}}.\end{array} \tag{3.4.86}$$

Once E is updated from the energy equation, the phase can be found from Equation 3.4.86, and then T can be found from Equation 3.4.85.

Remark 1: The case of *non-isothermal* phase change occurring over an extended temperature range $[T^{\text{solidus}}, T^{\text{liquidus}}]$ may be accommodated by replacing $T=T_{\mathrm{m}}$ in Equation 3.4.86 by $T^{\text{solidus}}\le T\le T^{\text{liquidus}}$.

In fact, instead of the simple EoS (Equation 3.4.85), more general equations of state may be used, see Alexiades and Autrique (2010).

Remark 2: We emphasize again that the front is *not* being tracked; it does not even appear explicitly in the formulation. It can be *captured* from the solution as the region where $0\le E\le \rho L$, see Equation 3.4.108.

Remark 3: The meaning of Equation 3.4.83, when E and Q are discontinuous, needs to be clarified. In fact, it must be interpreted in a *weak* sense mathematically, see Alexiades and Solomon (1993, Chap. 4).

Enthalpy Scheme

The energy conservation law (Equation 3.4.83) is valid in any number of (space) dimensions, for any appropriate constitutive law for flux Q, interpreted in a weak sense mathematically. It may be discretized by finite differences or finite volumes or finite elements. The finite-volume discretization is the most physical and easiest to implement computationally, so we present in the following an algorithm for the *finite-volume enthalpy scheme*, for a typical heat transfer problem, in one space dimension for simplicity. It can be extended to two or three dimensions easily.

To be specific, consider heat transfer in a slab occupying the region $\Omega=[a, b]$, initially at temperature $T=T_{\text{init}}(x)$, with convective heating at the left wall $x=a$, with heat transfer coefficient h and ambient $T_{\text{amb}}(t)\ge T_{\mathrm{m}}$, while keeping the back end $x=b$ insulated. If $T_{\text{init}}(x)\le T_{\mathrm{m}}$ and $T_{\text{amb}}(t)\ge T_{\mathrm{m}}$, then this would be a two-phase Stefan problem, more general than in "Two-Phase Stefan Problem" section. The material parameters involved are: $\rho=$const., c_p^L, c_p^S, k_L, k_S, h (which may depend on T), along with T_{m} and L.

The enthalpy formulation of this problem is as follows:

$$\text{Energy conservation } E_t + Q_x = 0, \quad a < x < b, 0 < t \le t_{\text{end}}, \tag{3.4.87}$$

$$\text{Constitutive law: } Q = -kT_x, \tag{3.4.88}$$

$$\text{Equation of state (3.4.85)}$$

$$\text{Initial condition: } T(x,0) = T_{\text{init}}(x), \quad a \le x \le b, \tag{3.4.89}$$

$$\text{Boundary conditions: } Q(a,t) = h[T_{\text{amb}}(t) - T(a,t)], Q(b,t) = 0, \quad t > 0, \tag{3.4.90}$$

for the given $T_{\text{init}}(x)$ and $T_{\text{amb}}(t)$. The only difference with plain heat conduction is the jump of energy in the EoS (Equation 3.4.85), which encodes the latent heat effect.

Space mesh: Choose an integer M and partition the interval $\Omega=[a, b]$ into M control volumes $V_i=[x_{i-1/2}, x_{i+1/2}]$, $i=1, 2, \ldots, M$, of width $\Delta x_i=x_{i+1/2}-x_{i-1/2}$. Place a node point x_i in V_i, and define the array of nodes $x_0=a, x_1, x_2, \ldots, x_M, x_{M+1}=b$. For a *uniform mesh*, $\Delta x=(b-a)/M$, and the nodes array is: $x_0=a, x_1=\Delta x/2, x_i=x_1+(i-1)\Delta x$ for $i=2, \ldots, M, x_{M+1}=b$. Note that x_0 and x_{M+1} are the boundaries, included into the mesh array for convenience.

Time steps: Choose time steps $\Delta t_n > 0$, small enough for accuracy and stability (see later text), and define the discrete times $t_0=0$, $t_{n+1}=t_n+\Delta t_n$, $n=0, 1,\ldots$. If $\Delta t_n=\Delta t$ for all n, then $t_n=n\Delta t$, and the number of time steps to get to the desired time t_{end} will be $N=t_{\text{end}}/\Delta t$.

Finite volume discretization of Equation 3.4.87: Integrating Equation 3.4.87 over the ith control volume $V_i=[x_{i-1/2}, x_{i+1/2}]$, $i=1, \ldots, M$ and over the time interval $[t_n, t_n+\Delta t_n]$ gives

$$\int_{t_n}^{t_{n+1}} \frac{\partial}{\partial t}\left(\int_{x_{i-1/2}}^{x_{i+1/2}} E(x,t)\,\mathrm{d}x\right)\mathrm{d}t + \int_{t_n}^{t_{n+1}}\int_{x_{i-1/2}}^{x_{i+1/2}} \frac{\partial}{\partial x} Q(x,t)\,\mathrm{d}x\,\mathrm{d}t = 0, \tag{3.4.91}$$

and carrying out integration of the derivatives leads to

$$\int_{x_{i-1/2}}^{x_{i+1/2}} E(x,t)\,\mathrm{d}x\Big|_{t_n}^{t_{n+1}} + \int_{t_n}^{t_{n+1}} [Q(x_{i+1/2},t) - Q(x_{i-1/2},t)]\,\mathrm{d}t = 0, \tag{3.4.92}$$

which expresses *exact discrete conservation*. Now define the discrete mean energy and mean flux as

$$E_i^n = \frac{1}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}} E(x,t_n)\,\mathrm{d}x = \text{ mean value of } E \text{ over } V_i \text{ at } t_n, i = 1,\ldots,M, \tag{3.4.93}$$

$$Q_{i-1/2}^{n+\theta} = \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} Q(x_{i-1/2},t)\,\mathrm{d}t = \text{ mean flux during } [t_n, t_{n+1}] \text{ through face at } x_{i-1/2}, i = 1,\ldots,M+1, \tag{3.4.94}$$

with *implicitness parameter* $0 \le \theta \le 1$. The superscript $^{n+\theta}$ denotes a time between t_n and t_{n+1} at which the mean value of flux is attained. With this notation, Equation 3.4.92 takes the form

$$E_i^{n+1} = E_i^n + \frac{\Delta t_n}{\Delta x}\left[Q_{i-1/2}^{n+\theta} - Q_{i+1/2}^{n+\theta}\right], \quad i = 1,\ldots,M, n = 0,1,\ldots \tag{3.4.95}$$

Note that no approximations have been made so far; this is still the *exact discrete conservation* of energy. It will update E_i to new time t_{n+1} once we specify how to compute the discrete fluxes. Also note that the updating scheme involves only the fluxes across the left and right faces of each control volume.

The first approximation comes from choosing a mean time for flux. For any $0 \le \theta \le 1$, we interpret the *mean time* $t_{n+\theta}$ as $t_{n+\theta} = (1-\theta)t_n + \theta\, t_{n+1}$, and assume that $Q^{n+\theta}_{i-1/2}$ is a convex combination of old $Q^{n}_{i-1/2}$ and new $Q^{n+1}_{i-1/2}$ values:

$$Q^{n+\theta}_{i-1/2} = (1-\theta)\, Q^{n}_{i-1/2} + \theta\, Q^{n+1}_{i-1/2}. \tag{3.4.96}$$

The standard choices for θ are 0, 1/2, 1:

- $\theta = 0$: Then the scheme is *explicit* (forward Euler time discretization). The right-hand side of Equation 3.4.95 is known at t_n, so the update requires only evaluation. However, the time step needs to be restricted to satisfy the Courant–Friedrichs–Lewy (CFL) stability condition, see Equations 3.4.102 and 3.4.103.
- $\theta = 1/2$: *Crank–Nicolson* scheme. It is implicit, so a system of equations needs to be solved at each time step. It is unconditionally stable, but not positivity-preserving unless Δt is restricted, see Equation 3.4.101.
- $\theta = 1$: *Backward Euler* scheme. It is implicit, so a system of equations needs to be solved at each time step. It is unconditionally stable, and positivity-preserving, so the time step needs to be restricted *only* for accuracy.

Discretization of Equation 3.4.88: The second approximation comes from discretization of flux, which involves the gradient of temperature. Let T_i denote the mean temperature of control volume V_i, $i = 1, \ldots, M$, and T_0, T_{M+1} the temperatures at the end points, consistently with our mesh. However, in order to approximate the gradient, the mean temperature T_i needs to be assigned to a specific location, which we take to be the node x_i of V_i. Then the natural discretization of Fourier's law (Equation 3.4.88) for heat flux across face $x_{i-1/2}$, i.e., from x_{i-1} to x_i, is

$$Q_{i-1/2} = -k_{i-1/2}\, \frac{T_i - T_{i-1}}{x_i - x_{i-1}}, \quad i = 1, \ldots, M+1, \tag{3.4.97}$$

but we need to clarify the value of $k_{i-1/2}$. It represents an "effective" conductivity for heat transfer from node x_{i-1} to node x_i. But this involves two, possibly different, conductivities k_{i-1} and k_i. Instead, let us view Fourier's law in terms of *resistivity* $= R = \Delta x/k$ (which is additive), see Alexiades and Solomon (1993, p. 188), so we define

$$\text{Resistivity of the path } [x_{i-1}, x_i] : R_{i-1/2} = \frac{x_{i-1/2} - x_{i-1}}{k_{i-1}} + \frac{x_i - x_{i-1/2}}{k_i}, \tag{3.4.98}$$

and write Equation 3.4.97 as

$$Q_{i-1/2} = -\frac{T_i - T_{i-1}}{R_{i-1/2}}. \tag{3.4.99}$$

Apart from boundary conditions, which are discussed later, Equation 3.4.99 constitute a discretization for heat transfer (in one-dimensional), with or without phase change. Phase change enters via the EoS (Equation 3.4.85), in which case, the nodal conductivities entering Equation 3.4.98 can be taken to be

$$\frac{1}{k_i} = \frac{\lambda_i}{k_L} + \frac{1-\lambda_i}{k_S}, \tag{3.4.100}$$

with λ_i the liquid fraction of node $i = 1, \ldots, M$. For other options, see Alexiades and Solomon (1993, p. 215).

Remark 4: Note that Equation 3.4.97 is a centered finite-difference approximation of the derivative, hence, it is second order, provided the face $x_{i-1/2}$ is chosen half-way between the nodes x_{i-1} and x_i, as is the case for a uniform mesh. Then the explicit and the fully implicit schemes are of second order in space and first order in time, whereas the Crank–Nicolson scheme is second order in both space and time.

Remark 5: It is possible to substitute E_i and fluxes in terms of temperature into Equation 3.4.95 to obtain schemes for directly updating the temperature with latent heat as a source term, (Shyy et al., 1996). We do not recommend such formulations, see Remark 6.

The *positive coefficient rule* is a simple and effective way to guarantee (strong) stability: when T_i^{n+1} is written as a linear combination of its neighbors T_{i-1}^n, T_i^n, T_{i+1}^n, all coefficients should be positive, (Patankar, 1980). This yields a restriction on the time step and ensures that the scheme will be *positivity-preserving* (non-negative values at t_n implies non-negative values at t_{n+1}), and thus free of unphysical oscillations. Positivity-preserving is stronger than and implies numerical stability. Applying the rule to the θ scheme (Equations 3.4.95, 3.4.96, 3.4.99, 3.4.98), in the general case of non-uniform mesh and possibly variable properties (temperature-dependent and/or phase-dependent), it yields the

$$\text{Strong CFL condition}: \mu \le \frac{1}{2(1-\theta)}, \text{ with } \mu = \frac{\Delta t_n}{\rho\, c_{pi}\, \Delta x_i}\, \frac{1}{R_{i\pm 1/2}}. \tag{3.4.101}$$

This shows that only the fully implicit scheme ($\theta = 1$) is truly unconditionally stable and positivity-preserving. All others need a restriction on the time step to preserve positivity. Even though implicit schemes ($0 < \theta < 1$) are stable in the von Neumann sense (errors remain bounded), they can produce unphysical oscillations for large time steps. The Crank–Nicolson scheme ($\theta = 1/2$) needs $\mu \le 1$, and the explicit scheme ($\theta = 0$) needs $\mu \le 1/2$.

The *explicit scheme* ($\theta = 0$) is, by far, the simplest and easiest to implement, as it does not involve any system solving. Moreover, it is also far easier to parallelize for concurrent computation on multi-processors, via domain decomposition using the Message Passing Interface (MPI) library, for demanding problems. The only price to be paid for this simplicity and convenience is that the time step Δt must be restricted to satisfy the CFL condition (Equation 3.4.101), which amounts to

$$\Delta t_n \le \frac{\rho\, c_{pi}\, \Delta x_i}{\dfrac{1}{R_{i-1/2}} + \dfrac{1}{R_{i+1/2}}}, \text{ or simply } \Delta t_n \le \frac{\Delta x_{\min}^2}{2\,\alpha_{\max}}, \tag{3.4.102}$$

with $\Delta x_{\min} = \min \Delta x_i$ and $\alpha_{\max} = \max \alpha_i$. For uniform mesh and constant thermophysical properties, it simplifies to the standard

$$\text{CFL condition}: \Delta t \le \frac{\Delta x^2}{2\,\alpha}, \text{ with } \alpha = \frac{k}{\rho c_p}. \tag{3.4.103}$$

Discretization of Boundary Conditions

The standard types of boundary conditions are Dirichlet, Neumann, and convective (robin) type. We discuss each type at $x = a$ for illustration.

1. Imposed temperature (Dirichlet) boundary condition: $T(a, t) = T_a(t) = \text{given}$. Here the boundary temperature $T_0^n = T_a(t_n)$ is specified, so by Equation 3.4.99, the boundary flux at $x = a$ is:

$$Q_{1/2}^n = -\frac{T_1^n - T_0^n}{R_{1/2}}, \text{with } R_{1/2} = \frac{x_1 - x_0}{k_1}. \tag{3.4.104}$$

Note that for a uniform mesh, $x_1 - x_0 = \Delta x/2$.

2. Imposed flux (Neumann) boundary condition: $Q(a, t) = q_a(t) = \text{given}$.

Here the boundary flux $Q_{1/2}^n = q_a(t_n)$ is specified, which is all the scheme needs to proceed. However, the boundary temperature will change and needs to be found. Solving Equation 3.4.104 for T_0, gives

$$T_0^n = T_1^n + R_{1/2}\, Q_{1/2}^n . \tag{3.4.105}$$

This involves the nodal value T_1, which will be updated from Equation 3.4.95.

A very important particular case is that of an *insulated boundary*: $q_a(t)=0$, so then $T_0^n = T_1^n$.

3. Convective boundary condition: $Q(a, t)=h[T_{\text{amb}}(t)-T(a, t)]$,

with $T_{\text{amb}}(t)$, a given ambient temperature. Here neither T nor Q is specified at the boundary, only a relation between them (Newton's law of cooling). Its discretization is $Q_{1/2}^n = h[T_{\text{amb}}(t_n)-T_0^n]$, or, using Equation 3.4.103, $-\dfrac{T_1^n - T_0^n}{R_{1/2}} = h[T_{\text{amb}}(t_n)-T_0^n]$. Solving for T_0^n gives

$$T_0^n = \frac{T_1^n + h\, R_{1/2}\, T_{\text{amb}}^n}{1 + h\, R_{1/2}} . \tag{3.4.106}$$

Substituting this into the expression for $Q_{1/2}^n$ yields

$$Q_{1/2}^n = -\frac{T_1^n - T_{\text{amb}}^n}{R_{1/2} + 1/h} . \tag{3.4.107}$$

Observe that $h \to \infty$ amounts to the imposed temperature case (Equation 3.4.104), as it should.

Algorithm for the Explicit Finite Volume Scheme

We present the complete finite-volume enthalpy algorithm for the melting problem mentioned at the beginning of "Enthalpy Scheme" section. For simplicity, we assume constant c_S, c_S, k_S, k_L, and uniform mesh.

- Read in data: M, a, b, t_{end}
- Set $\Delta x=(b-a)/M$, the time step Δt from Equation 3.4.103, and $N_{\text{steps}}=t_{\text{end}}/\Delta t$.
- Declare arrays $x(0{:}M+1)$, $T(0{:}M+1)$, $E(1{:}M)$, $\lambda(1{:}M)$, $R(1{:}M+1)$, $Q(1{:}M+1)$. Note that there is no reason to index time into the arrays.
- Set initial values: $T_i=T_{\text{init}}(x_i)$, $E_i=E(T_i)$ from Equation 3.4.85, and λ_i from Equation 3.4.86, $i=1{:}M$.
- Time stepping: for $n=1, \ldots, N_{\text{steps}}$, do:
  1. Set the boundary fluxes $Q_{1/2}$, Q_{M+1} from Equations 3.4.104–3.4.107, and internal fluxes from Equation 3.4.99, $i=2, \ldots, M$.
  2. Update energies E_i from Equation 3.4.95 to new time t_{n+1}.
  3. Update λ_i from Equation 3.4.86, and $R_{i-1/2}$ from Equations 3.4.98 and 3.4.100 to time t_{n+1}.
  4. Update T_i from Equation 3.4.85 and the boundary temperatures from Equations 3.4.106 and 3.4.105 to time t_{n+1}.

Now all quantities have been updated to new time and can go to 1 for the next time step.

For a pure material melting at T_{m}, the algorithm will generate only one mushy control volume, say volume $i=m$ at time t_n, and then the front location at that time can be computed for plotting by

$$X^n = x_{m-1/2} + \lambda_m\, \Delta x_m . \tag{3.4.108}$$

Remark 6: The latent heat effect is only felt in mushy nodes so each control volume should become mushy before changing phase. The aforementioned algorithm is robust in this respect, so skipping the transition indicates a bug or too large a time step. Some other implementations or algorithms may miss this, so one must be careful. Especially prone are schemes that directly update temperature either implicitly using too large a time step, or they account for the latent heat via a source term.

Remark 7: The algorithm generalizes to two or three dimensions easily. We only need to define fluxes in the other directions; the energy PDE will contain an additional pair of fluxes for each additional dimension. For example, in a two-dimensional Cartesian mesh of $M_x \times M_y$ nodes, there are two flux arrays: $Qx_{i-1/2,j}$ for $i=1{:}M_x+1$, $j=1{:}M_y$ and $Qy_{i,\,j-1/2}$ for $i=1{:}M_x$, $j=1{:}M_y+1$; the energy update (Equation 3.4.95) becomes:

$$E_{i,j}^{n+1} = E_i^n + \frac{\Delta t_n}{\Delta x \Delta y}\left[Q_{i-1/2,\,j}^n - Q_{i+1/2,\,j}^n + Q_{i,\,j-1/2}^n - Q_{i,\,j+1/2}^n\right], \quad i=1:M_x,\, j=1:M_y. \tag{3.4.109}$$

Heat Capacity Method

The heat capacity method is a fixed-domain, no front tracking, temperature-based method that incorporates the latent heat into the heat capacity $c_p(T)$. Temperature is updated directly by the heat equation. It assumes non-isothermal phase change over some interval $[T_m - \Delta T,\ T_m + \Delta T]$, within which thermal properties are taken to be convex combinations of solid and liquid properties, interpolated by the liquid fraction λ. There are two main approaches: (i) apparent heat capacity and (ii) effective heat capacity.

1. *Apparent heat capacity method* was introduced by Hashemi and Sliepcevich (1967) and has been developed by Idelsohn et al. (1994), Hu and Argyropoulos (1996), and Voller (1997). This method is relatively popular today because it updates temperature directly from the heat equation, and can use experimental results to formulate an empirical expression for heat capacity. Dynamic measurements can be performed for pure and homogenous materials, utilizing differential scanning calorimetry (DSC) with a step testing method, (Castellon et al., 2008; Gŭnther et al., 2009). For example, Fang and Medina (2009) used DSC to develop a mathematical formula for heat capacity of paraffinic products. In cases of non-uniform materials and composite blends, T-history method, (Zhang et al., 1999), or dynamic heat flow meter apparatus measurements need to be applied, (Kosny et al., 2009a, b; Shukla et al., 2013). Apparent heat capacity can also be estimated using numerical methods (Lemmon, 1981; Voller, 1997).
2. *Effective heat capacity method* became known thanks to the publications of Heim and Clark (2004), Minkowycz et al. (2006), Muhieddine et al. (2009), and Poirier (1988). Since the relationship between heat capacity and temperature is highly non-linear, an explicit scheme is often used. This may cause some inaccuracies in situations of sharp heat capacity changes with respect to temperature (Sadasivam et al., 2011). For single-component cases with well-defined melting point T_m, the temperature range $2\Delta T$ should be as small as possible.

Hybrid Methods

There are a large number of modified methods using components from one or both methods described previously. In the following, we mention a couple of such methods.

Quasi-enthalpy method: Pham (1985, 2006) introduced a hybrid scheme that consists of: (i) a temperature prediction step, and (ii) a verification and correction step in which the predicted temperatures are checked for consistency against the enthalpy–temperature curve and corrected if necessary. This method was introduced as one of the alternative numerical algorithms allowing phase-change simulation in ESP-r, a whole-building energy simulation program, (Sadasivam et al., 2011).

TABLE 3.4.2
Numerical Methods Used in Best-Known Numerical Tools Which Can Be Used for Phase-Change Simulations in Buildings

Whole-Building Computer Tool	Numerical Method	Discretization	Time Stepping	References
ESP-r	Effective heat capacity	Finite volume	Explicit	Heim and Clarke (2004)
TRNSYS type 204, plus type 56	Effective heat capacity	Finite element	Crank–Nicolson	Jokisalo et al. (2000)
TRNSYS type 222, plus type 56	Indirect calculations			Ibanez et al. (2005)
TRNSYS type 241, plus type 56	Enthalpy method	Finite difference		Schranzhofer et al. (2006)
TRNSYS type 260, plus type 56	Effective heat capacity	Finite difference	Explicit	Kuznik et al. (2010)
TRNSYS type 399, plus type 56	Two different methods available based on the heat capacity approach	Finite element	Crank–Nicolson	Dentel and Stephan (2013)
MATLAB	Enthalpy method	Finite volume	Crank–Nicolson	Sadasivam et al. (2011)
EnergyPlus	Enthalpy method	Finite difference	Implicit	Pedersen (2007)

Improved temperature-based equivalent heat capacity method: (Cao and Faghri, 1990; Zhanhua and Yuwen, 2006). After a series of enhancements, this numerical model has been used in a number of projects—(Zeng and Faghri, 1994a, b; Zhanhua and Yuwen, 2006). Today, this method is not in common use; however it offers an alternative to the apparent heat capacity method.

In large containers, natural convection in the liquid phase may impact overall system effectiveness. A study by Ahmad et al. (2006) demonstrated that complexity of thermal models does not always yield improvements in simulation accuracy. In this work, numerical predictions generated by a complex thermal model considering a solid–liquid interface and results of a simpler model using the effective heat capacity method were compared. Recorded results showed that a simpler heat capacity approach was more accurate. In another similar study, finite-element computer simulations were performed to compare accuracies of the enthalpy and heat capacity methods, (Lamberg et al., 2004). When compared to experimental data, better agreement was obtained with the heat capacity method.

A list of numerical methods used in the most popular whole-building energy simulation tools, with the capability of modeling phase changes, is given in Table 3.4.2.

REFERENCES

Ahmad, M., Bontemps, A., Salle, H., and Quenard, D. (2006). Thermal testing and numerical simulation of a prototype cell using light wallboards coupling vacuum insulation panels and phase change materials. *Energy Build.* 38: 673–681.

Albert, M.R., and ONeill K. (1986). Moving boundary-moving mesh analysis of phase change using finite elements with transfinite mappings. *Int. J. Numer. Methods Eng.* 23(4): 591–607.

Alexiades, V., and Solomon, A.D. (1993). *Mathematical Modeling of Melting and Freezing Processes*. Hemisphere Publishing Corporation, Washington, DC.

Alexiades, V., and Autrique, D. (2010). Enthalpy model for heating, melting and vaporization in laser ablation. *Electron. J. Differ. Equations*, Conf. 19: 114.

AL-Saadi, S.N., and Zhai, Z. (2013). Modeling phase change materials embedded in building enclosure: A review. *Renewable Sustainable Energy Rev.* 21: 659–673.

Basu, B., and Date, A. (1988). Numerical modeling of melting and solidification problems—A review. *Sadhana Acad. Proc. Eng. Sci.* 13: 169–213.

Bonnerot, R., and Janet, P. (1977). Numerical computation of the free boundary for the two-dimensional Stefan problem by space–time finite elements. *J. Comput. Phys.* 25: 163–181.

Cao, Y., and Faghri, A.A. (1990). Numerical analysis of phase-change problems including natural convection. *J. Heat Transfer* 112: 812–816.

Carslaw, H.S., and Jaeger, J.C. (1959). *Conduction of Heat in Solids*. Clarendon Press, Oxford, UK.

Castellon, C., Gŭnther, E., Mehling, H., Hiebler, S., and Cabeza, L.F. (2008). Determination of the enthalpy of PCM as a function of temperature using a heat-flux DSC—A study of different measurement procedures and their accuracy. *Int. J. Energy Res.* 32(13): 1258–1265.

Chuang, Y.K., and Szekely, J. (1972). On the use of greens functions for solving melting and solidification problems. *Int. J. Heat Mass Transfer* 15: 1171–1174.

Colin, J. (2014). Morphological instability of a stressed solid cylinder in the solidification and melting regimes. *J. Cryst. Growth* 402: 113–118.

Crank, J. (1984). *Free and Moving Boundary Problems*. Clarendon Press, Oxford, UK.

Crank, J., and Gupta, R. (1975). Isotherm migration method in two dimensions. *Int. J. Heat Mass Transfer* 18: 1101–1117.

Dentel, A., and Stephan, W. (2013). TRNSYS TYPE 399, Phase change materials in passive and active wall constructions; Model description and implementing into TRNSYS. Georg Simon Ohm University of Applied Sciences, Institute for Energy and Building. Version 1.5, May 2013, Germany-http://www.transsolar.com/_software/download/de/ts_type_399_en.pdf

Douglas, J., and Gallie, T.M. (1955). On the numerical integration of a parabolic differential equation subject to a moving boundary condition, *Duke Math. J.* 22: 557–571.

Dusinberre, G.M. (1945). Numerical methods for transient heat flow. *Trans. ASME* 67: 703–712.

Dutil, Y., Rousse, D.R., Salahb, N.B., Lassue, S., and Zalewski, S. (2011). A review on phase-change materials: Mathematical modeling and simulations. *Renewable Sustainable Energy Rev.* 15(1): 112–130.

Evans, G.W. (1951). A note on the existence of a solution to a problem of Stefan. *Q. Appl. Math.* 9: 185–193.

Eyres, N.R., Hartree, D.R., Ingham, J., Jackson, R., Sarjant, R.J., and Wagstaff, S.M. (1946). The calculation of variable heat flow in solids. *Philos. Trans. R. Soc. A* 240: 1–57.

Fang, Y., and Medina, M. (2009). Proposed modifications for models of heat transfer problems involving partially melted phase change processes. *J. ASTM Int.* 6(9): 1–20.

Farid, M.M., Khudhair, A.M., Razack, S.A.K., and Al-Hallaj, S. (2004). A review on phase change energy storage: Materials and applications. *Energy Convers. Manage.* 45(9–10): 1597–1615.

Glimm, J., Grove, J.W., Li, X.L., Shyue, K.-M., Zhang, Q., and Zeng, Y. (1998). Three dimensional front tracking. *SIAM J. Sci. Comput.* 19: 703–727.

Gŭnther, E., Hiebler, S., Mehling, H., and Redlich, R. (2009). Enthalpy of phase change materials as a function of temperature: Required accuracy and suitable measurement methods. *Int. J. Thermophys.* 30: 1257–1269.

Hashemi, H.T., and Sliepcevich, C.M. (1967). A numerical method for solving two-dimensional problems of heat conduction with change of phase. *Chem. Eng. Prog. Symp. Ser.* 63: 34–41.

Heim, D., and Clarke, J.A. (2004). Numerical modeling and thermal simulation of PCMGypsum composites with ESP-r. *Energy Build.* 36: 795–805.

Hill, J.M. (1987). *One-Dimensional Stefan Problems: An Introduction*. Longman Scientific Technical, Harlow, UK.

Hsiao, J.S. (1985). An efficient algorithm for finite-difference analysis of heat transfer with melting and solidification. *Numer. Heat Transfer* 8: 653–666.

Hu, H., and Argyropoulos, S.A. (1996). Mathematical modelling of solidification and melting: A review. *Modell. Simul. Mater. Sci. Eng.* 4: 371–396.

Ibanez, M., Lazaro, A., Zalba, B., and Cabeza, L.F. (2005). An approach to the simulation of PCMs in building applications using TRNSYS. *Appl. Therm. Eng.* 25: 1796–1807.

Idelsohn, S., Storti, M., and Crivelli, L. (1994). Numerical methods in phase-change problems. *Arch. Comput. Methods Eng.* 1: 49–74.

Javierre, E., Vuik, C., Vermolen, F.J., and van der Zwaag, S. (2006). A comparison of numerical models for one-dimensional Stefan problems. *J. Comput. Appl. Math.* 192: 445–459.

Jokisalo, J., Lamberg, P., and Siren, K. (2000). Thermal simulation of PCM structures with TRNSYS. *Proceedings of the 8th International Conference on Thermal Energy Storage*. Stuttgart, Germany.

Kim, K.M. (1978). Morphological instability under constitutional supercooling during the crystal growth of InSb from the melt under stabilizing thermal gradient. *J. Cryst. Growth* 44(4): 403–413.

Kośny, J., Stovall, T. K., and Yarbrough, D.W. (2009a). Dynamic heat flow measurements to study the distribution of phase-change material in an insulation matrix. *Proceedings of the 2009 International Thermal Conductivity Conference (ITCC) and the International Thermal Expansion Symposium (ITES)*, Aug. 29–Sep. 2, 2009, Pittsburgh, PA.

Kośny, J., Kossecka, E., and Yarbrough, D.W. (2009b). Use of a heat flow meter to determine active PCM content in an insulation. *Proceedings of the 2009 International Thermal Conductivity Conference (ITCC) and the International Thermal Expansion Symposium (ITES)*, Aug. 29–Sep. 2, 2009, Pittsburgh, PA.

Kuznik, F., Virgone, J., and Johannes, K. (2010). Development and validation of a new TRNSYS type for the simulation of external building walls containing PCM. *Energy Build.* 42 (7): 1004–1009.

Lamberg, P., Lehtiniemi, R., and Henell, A.M. (2004). Numerical and experimental investigation of melting and freezing processes in phase change material storage. *Int. J. Therm. Sci.* 43: 277–287.

Lazaridis, A. (1969). A numerical solution of the solidification (or melting) problem in multidimensional space. Ph.D. dissertation, Columbia University.

Lemmon, E. (1981). Multidimensional integral phase change approximations for finite element conduction codes, pp. 201–213 in *Numerical Methods in Heat Transfer*, Edited by R.W. Lewis, K. Morgan, O.C. Zienkiewicz. Wiley, New York.

Lupis, C.H.P. (1983). *Chemical Thermodynamics of Materials*. Prentice Hall, Englewood Cliffs, NJ.

Lynch, R.D. (1982). Unified approach to simulation on deforming elements with application to phase change problems. *J. Comput. Phys.* 47(3): 387–411.

Mehling, H., and Cabeza, L.F. (2008). *Heat and Cold Storage with PCM*. Springer, Berlin, Germany.

Minkowycz, W.J., Sparrow E.M., and Murthy J.Y. Editors, (2006). *Handbook of Numerical Heat Transfer*. 2nd edition, Wiley, Hoboken, NJ.

Muhieddine, M., Canot, E., and March, R. (2009). Various approaches for solving problems in heat conduction with phase change. *Int. J. Finite Vol.* 6(1): 120.

Murray, W.D., and Landis F. (1959). Numerical and machine solutions of the transient heat conduction problems involving melting or freezing. *J. Heat Transfer* 81: 106–112.

ONeill, K. (1983). Boundary integral equation solution of moving boundary phase change problems. *Int. J. Numer. Methods Eng.* 19: 1825–1850.

Osher, S.J., and Fedkiw, R.P. (2003). *Level Set Methods and Dynamic of Implicit Surfaces*. Springer-Verlag, New York.

Ozisik, M.N. (1993). *Heat Conduction*. 2nd ed. Wiley, New York.

Patankar, S.V. (1980). *Numerical Heat Transfer and Fluid Flow*. Hemisphere, WA.

Pedersen, C.O. (2007). Advanced zone simulation in EnergyPlus: Incorporation of variable properties and phase change materials (PCM) capability, pp. 1341–1345 in *Proceedings of the 10th International IBPSA Conference*. Beijing, China, September 3–6, 2007.

Pham, Q.T. (1985). A fast, unconditionally stable finite-difference scheme for heat conduction with phase change. *Int. J. Heat Mass Transfer* 28: 2079–2084.

Pham, Q.T. (2006). Modeling heat and mass transfer in frozen foods: A review. *Int. J. Refrig.* 29: 876–888.

Poirier, D.J. (1988). On numerical methods used in mathematical modeling of phase change in liquid metals. *J. Heat Transfer* 110: 562–570.

Qin, R.S., and Bhadeshia, H.K. (2010). "Phase Field Method" review. *Mater. Sci. Technol.* 26(7): 803–811.

Rao, P., and Sastri, V.M.K. (1984). Efficient numerical method for two-dimensional phase change problems. *Int. J. Heat Mass Transfer* 27: 2077–2084.

Rubinstein, L.I. (1971). *The Stefan Problem*. Amer. Math. Soc. (Translated from 1967 Russian edition).

Sadasivam, S., Almeida, F., Zhang, D., and Fung, A.S. (2011). An iterative enthalpy method to overcome the limitations in ESP-rs PCM solution algorithm. *ASHRAE Trans.* 117: 100–107.

Schranzhofer, H., Puschnig, P., Heinz, A., and Streicher, W. (2006). Validation of a TRNSYS simulation model for PCM energy storages and PCM wall construction elements. *Proceedings of ECOSTOCK 2006 10th International Conference on Thermal Energy Storage*, Pomona, NJ.

Sethian, J.A. (1999). *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge University Press, Cambridge, UK.

Shukla, N., Cao, P., and Kośny, J. (2013). Lab-scale dynamic thermal testing of PCM-enhanced building materials. *Proceedings of the ASTM Symposium on Next-Generation Thermal Insulation Challenges and Opportunities*, Oct. 23–24, 2013, Jacksonville, FL.

Shyy, W., Vdaykumar, H.S., Rae, M.M., and Smith, R.W. (1996). *Computational Fluid Dynamics with Moving Boundaries*. Hemisphere, Washington, DC.

Sun, Y., and Beckermann, C. (2007). Sharp interface tracking using the phase-field equation. *J. Comput. Phys.* 210: 626–653.

Uverdi, S.O., and Tryggvason, G. (1992). A front tracking method for viscous, incompressible multi-fluid flows. *J. Comput. Phys.* 100: 25–37.

Voller, V.R. (1997). An overview of numerical methods for solving phase change problems. pp. 341–380 in *Advances in Numerical Heat Transfer.* Edited by Minkowycz, W.J., Sparrow, E.M., Taylor & Francis.

Voller, V.R., Swenson, J.B., Kim, W., and Paola, C. (2006). An enthalpy method for moving boundary problems on the Earths surface. *Int. J. Numer. Methods Heat Fluid Flow* 16(6): 41–54.

Zabaras, N., and Ruan, Y. (1990). Moving and deforming finite element simulation of two-dimensional Stefan problems. *Commun. Appl. Numer. Methods* 6(7): 495–506.

Zaeema, M.A., and Mesarovicb, S.D. (2011). Morphological instabilities in thin films: Evolution maps. *Comput. Mater. Sci.* 50(3): 1030–1036.

Zeng, X., and Faghri, A. (1994a). Temperature-transforming model for binary solid-liquid phase–change problems. Part I: Mathematical modeling and numerical methodology. *Numer. Heat Transfer B* 25: 467–480.

Zeng, X., and Faghri, A. (1994b). Temperature-transforming model for binary solid-liquid phase–change problems. Part II: Numerical simulation. *Numer. Heat Transfer B* 25: 481–500.

Zhang, Y., Jiang, Y., and Jiang, Y. (1999). A simple method, the T-history method, of determining the heat of fusion, specific heat and thermal conductivity of phase-change materials. *Meas. Sci. Technol.* 10: 201–205.

Zhanhua, M., and Yuwen, Z. (2006). Solid velocity correction schemes for a temperature transforming model for convection phase change. *Int. J. Numer. Methods Heat Fluid Flow* 16: 204–225.

3.5 MASS TRANSFER

Anthony F. Mills

Introduction

Mass transfer may occur in a gas mixture, a liquid solution, or a solid solution. There are several physical mechanisms that can transport a chemical species through a phase and transfer it across phase boundaries. The two most important mechanisms are ordinary diffusion and convection. Mass diffusion is analogous to heat conduction and occurs whenever there is a gradient in the concentration of a species. Mass convection is essentially identical to heat convection: a fluid flow that transports heat may also transport a chemical species. The similarity of mechanisms of heat transfer and mass transfer results in the mathematics often being identical, a fact that can be exploited to advantage. But there are some significant differences between the subjects of heat and mass transfer. One difference is the much greater variety of physical and chemical processes that require mass transfer analysis. Another difference is the extent to which the essential details of a given process may depend on the particular chemical system involved, and on temperature and pressure.

In the next subsection, concentrations, velocities, and fluxes are defined, and special attention is paid to phase interfaces where the concentration of a chemical species is almost always discontinuous. Fick's law of ordinary diffusion is introduced in the third section, where other diffusion phenomena are also discussed. The fourth section presents various forms of the species conservation equation. Results for diffusion in a stationary medium are given in the fifth section, and include steady diffusion across a plane wall, transient diffusion in a semi-infinite solid, and diffusion in a porous catalyst. Results for diffusion in a moving medium are given in the sixth section, and the Stefan flow is introduced for diffusion with one component stationary. Also considered are particle combustion, droplet evaporation, and combustion of a volatile liquid hydrocarbon fuel droplet. The last section deals with mass convection. Low mass transfer rate theory is presented and how to exploit the analogy between convective heat and mass transfer is shown. Particular attention is given to situations involving simultaneous heat and mass transfer associated with evaporation or condensation. The section closes by presenting high mass transfer rate theory for convection, and gives engineering calculation methods for boundary layer flows that account for variable property effects.

Concentrations, Velocities, and Fluxes

Definitions of Concentrations

In a gas mixture, or liquid or solid solution, the local *concentration* of a mass species can be expressed in a number of ways. The *number density* of species i in a mixture or solution of n species is defined as

$$\begin{aligned}\text{Number density of species } i &\equiv \text{Number of molecules of } i \text{ per unit volume} \\ &\equiv \mathcal{N}_i \text{ molecules}/\text{m}^3\end{aligned} \tag{3.5.1}$$

Alternatively, if the total number of molecules of all species per unit volume is denoted as $\mathcal{N}$, then we define the *number fraction* of species i as

$$n_i \equiv \frac{\mathcal{N}_i}{\mathcal{N}}; \quad \mathcal{N} = \sum \mathcal{N}_i \tag{3.5.2}$$

where the summation is over all species present, $i = 1,2, \ldots, n$. Equations 3.5.1 and 3.5.2 describe *microscopic* concepts and are used, for example, when the kinetic theory of gases is used to describe transfer processes.

Whenever possible, it is more convenient to treat matter as a continuum. Then the smallest volume considered is sufficiently large for macroscopic properties such as pressure and temperature to have their usual meanings. For this purpose we also require *macroscopic* definitions of concentration. First, on a mass basis,

$$\begin{aligned}\text{Mass concentration of species } i &\equiv \text{partial density of species } i \\ &\equiv \rho_i \text{ kg/m}^3\end{aligned} \tag{3.5.3}$$

The total mass concentration is the total mass per unit volume, that is, the density $\rho = \Sigma\rho_i$. The *mass fraction* of species i is defined as

$$m_i = \frac{\rho_i}{\rho} \tag{3.5.4}$$

Second, on a molar basis,

$$\begin{aligned}\text{Molar concentration of species } i &\equiv \text{number of moles of } i \text{ per unit volume} \\ &\equiv c_i \text{ kmol/m}^3\end{aligned} \tag{3.5.5}$$

If M_i (kg/kmol) is the molecular weight of species i, then

$$c_i = \frac{\rho_i}{M_i} \tag{3.5.6}$$

The total molar concentration is the molar density $c = \Sigma c_i$. The *mole fraction* of species i is defined as

$$x_i \equiv \frac{c_i}{c} \tag{3.5.7}$$

A number of important relations follow directly from these definitions. The mean molecular weight of the mixture of solution is denoted by M and may be expressed as

$$M = \frac{\rho}{c} = \sum x_i M_i \tag{3.5.8a}$$

or

$$\frac{1}{M} = \sum \frac{m_i}{M_i} \tag{3.5.8b}$$

There are summation rules

$$\sum m_i = 1 \tag{3.5.9a}$$

$$\sum x_i = 1 \tag{3.5.9b}$$

It is often necessary to have the mass fraction of species i expressed explicitly in terms of mole fractions and molecular weights; this relation is

$$m_i = \frac{x_i M_i}{\sum x_j M_j} = x_i \frac{M_i}{M} \tag{3.5.10a}$$

and the corresponding relation for the mole fraction is

$$x_i = \frac{m_i / M_i}{\sum m_j / M_j} = m_i \frac{M}{M_i} \tag{3.5.10b}$$

Dalton's law of partial pressures for an ideal gas mixture states that

$$P = \sum P_i, \quad \text{where} \quad P_i = \rho_i R_i T \tag{3.5.11}$$

Dividing partial pressure by total pressure and substituting $R_i = \mathcal{R}/M_i$ gives

$$\frac{P_i}{P} = \frac{\rho_i}{M_i}\frac{\mathcal{R}T}{P} = c_i \frac{\mathcal{R}T}{P} = x_i \frac{c\mathcal{R}T}{P} = x_i \tag{3.5.12}$$

Thus, for an ideal gas mixture, the mole fraction and partial pressure are equivalent measures of concentration (as also is the number fraction).

A commonly used specification of the composition of dry air is 78.1% N_2, 20.9% O_2, and 0.9% Ar, by volume. (The next largest component is CO_2, at 0.3%.) Since equal volumes of gases contain the same number of moles, specifying composition on a volume basis is equivalent to specifying mole fractions, namely,

$$x_{N_2} = 0.781; \quad x_{O_2} = 0.209; \quad x_{Ar} = 0.009$$

The corresponding mass fractions are calculated to be

$$m_{N_2} = 0.755; \quad m_{O_2} = 0.231; \quad m_{Ar} = 0.014$$

Concentrations at Interfaces

Although temperature is continuous across a phase interface, concentrations are usually discontinuous. In order to define clearly concentrations at interfaces, we introduce imaginary surfaces, denoted *u* and *s*, on both sides of the real interface, each indefinitely close to the interface, as shown in Figure 3.5.1 for water evaporating into an airstream. Thus, the liquid-phase quantities at the interface are subscripted u, and gas-phase quantities are subscripted s. If we ignore the small amount of air dissolved in the water, $x_{H_2O,u} = 1$. Notice that the subscript preceding the comma denotes the chemical species, and the subscript following the comma denotes location. To determine $x_{H_2O,s}$ we make use of the fact that, except in extreme circumstances, the water vapor and air mixture at the s-surface must be in thermodynamic equilibrium with water at the u-surface. Equilibrium data for this system are found in conventional steam tables: the saturation vapor pressure of steam at the water temperature, T_s, ($T_s = T_u$), is the required partial pressure $P_{H_2O,s}$. With the total pressure P known, $x_{H_2O,s}$ is calculated as $P_{H_2O,s} / P$. If $m_{H_2O,s}$ is required, Equation 3.5.10a is used.

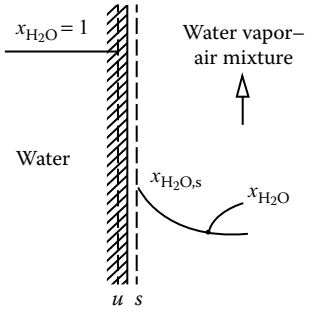


FIGURE 3.5.1 Concentrations at a water-air interface.

For example, at $T_s = 320$ K, the saturation vapor pressure is obtained from steam tables as 0.10535×10^5 Pa. If the total pressure is 1 atm $= 1.0133 \times 10^5$,

$$x_{H_2O,s} = \frac{0.10535 \times 10^5}{1.0133 \times 10^5} = 0.1040$$

$$m_{H_2O,s} = \frac{(0.1040)(18)}{(0.1040)(18) + (1 - 0.1040)(29)} = 0.06720$$

For a gas or solid dissolving in a liquid, equilibrium data are often referred to simply as solubility data, found in chemistry handbooks. Many gases are only sparingly soluble, and for such dilute solutions solubility data are conveniently represented by *Henry's law*, which states that the mole fraction of the gas at the s-surface is proportional to its mole fraction in solution at the u-surface, the constant of proportionality being the *Henry number*, He. For species *i*,

$$x_{i,s} = \mathrm{He}_i x_{i,u} \tag{3.5.13}$$

The Henry number is inversely proportional to total pressure and is also a function of temperature. The product of Henry number and total pressure is the *Henry constant*, C_{He}, and for a given species is a function of temperature only:

$$\mathrm{He}_i P = C_{\mathrm{He}_i}(\mathrm{T}) \tag{3.5.14}$$

Solubility data are given in Table 3.5.1.

For example, consider absorption of carbon dioxide from a stream of pure CO_2 at 2 bar pressure into water at 310 K. From Table 3.5.1, $C_{He} = 2170$ bar; thus

$$\mathrm{He}_{CO_2} = \frac{2170}{2} = 1085; \quad x_{CO_2,u} = \frac{1}{1085} = 9.22 \times 10^{-4}$$

Dissolution of gases into metals is characterized by varied and rather complex interface conditions. Provided temperatures are sufficiently high, hydrogen dissolution is reversible (similar to CO_2 absorption into water); hence, for example, titanium-hydrogen solutions can exist only in contact with a gaseous hydrogen atmosphere. As a result of hydrogen going into solution in atomic form, there is a characteristic square root relation

$$m_{H_2,u} \alpha P_{H_2,s}^{1/2}$$

TABLE 3.5.1
Henry Constants C_{He} for Dilute Aqueous Solutions at Moderate Pressures ($P_{i,s}/x_{i,u}$ in atm, or in bar $= 10^5$ Pa, within the Accuracy of the Data)

Solute	290 K	300 K	310 K	320 K	330 K	340 K
H_2S	440	560	700	830	980	1,140
CO_2	1,280	1,710	2,170	2,720	3,220	—
O_2	38,000	45,000	52,000	57,000	61,000	65,000
H_2	67,000	72,000	75,000	76,000	77,000	76,000
CO	51,000	60,000	67,000	74,000	80,000	84,000
Air	62,000	74,000	84,000	92,000	99,000	104,000
N_2	16,000	89,000	101,000	110,000	118,000	124,000

The constant of proportionality is strongly dependent on temperature, as well as on the particular titanium alloy: for Ti-6Al-4V alloy it is twice that for pure titanium. In contrast to hydrogen, oxygen dissolution in titanium is irreversible and is complicated by the simultaneous formation of a rutile (TiO_2) scale on the surface. Provided some oxygen is present in the gas phase, the titanium-oxygen *phase diagram* (found in a metallurgy handbook) shows that $m_{O_2,u}$ in alpha-titanium is 0.143, a value essentially independent of temperature and O_2 partial pressure. Dissolution of oxygen in zirconium alloys has similar characteristics to those discussed above for titanium.

All the preceding examples of interface concentrations are situations where thermodynamic equilibrium can be assumed to exist at the interface. Sometimes thermodynamic equilibrium does not exist at an interface: a very common example is when a chemical reaction occurs at the interface, and temperatures are not high enough for equilibrium to be attained. Then the concentrations of the reactants and products at the *s*-surface are dependent both on the rate at which the reaction proceeds—that is, the *chemical kinetics*—as well as on mass transfer considerations.

Definitions of Fluxes and Velocities

The mass (or molar) flux of species *i* is a vector quantity giving the mass (or moles) of species *i* that pass per unit time through a unit area perpendicular to the vector (Figure 3.5.2). We denote the absolute mass and molar fluxes of species *i*, that is, relative to stationary coordinate axes, as $\boldsymbol{n}_i$ (kg/m^{2s}) and $\boldsymbol{N}_i$ (kmol/m^{2}s), respectively. The absolute mass flux of the mixture (mass velocity) is

$$\boldsymbol{n} = \sum \boldsymbol{n}_i \tag{3.5.15}$$

and the local mass-average velocity is

$$\boldsymbol{v} = \frac{\boldsymbol{n}}{\rho}\,\text{m/s} \tag{3.5.16}$$

The velocity *v* is the velocity that would be measure by a Pitot tube and corresponds to the velocity used in considering pure fluids. On a molar basis, the absolute molar flux of the mixture is

$$\boldsymbol{N} = \sum \boldsymbol{N}_i \tag{3.5.17}$$

and the local molar-average velocity is

$$\boldsymbol{v}^* = \frac{\boldsymbol{N}}{c}\,\text{m/s} \tag{3.5.18}$$

The absolute fluxes of species *i* have two components. On a mass basis we write

$$\boldsymbol{n}_i = \rho_i \boldsymbol{v} + \boldsymbol{j}_i \tag{3.5.19}$$

where $\rho_i \boldsymbol{n}$ is transport of species *i* by bulk motion of the fluid at velocity $\boldsymbol{v}$ and is the *convective* component. Thus, $\boldsymbol{j}_i$ is transport of species *i* relative to the mass average velocity; it is called the

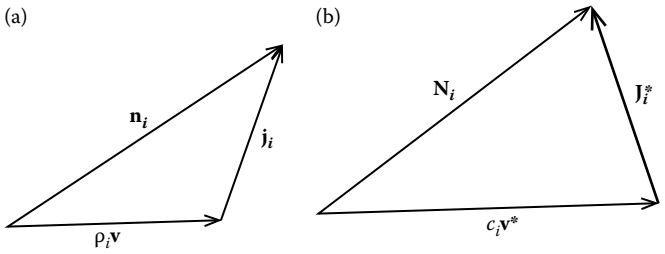


FIGURE 3.5.2 Flux vectors: (a) mass basis and (b) molar basis.

diffusive component because most commonly it is due to ordinary (concentration) diffusion of the species. On a molar basis the corresponding relation is

$$\boldsymbol{N}_i = c_i \boldsymbol{v}^* + \boldsymbol{J}_i^* \tag{3.5.20}$$

Some important relations follow from these definitions:

$$\sum \boldsymbol{j}_i = \sum \boldsymbol{J}_i^* = 0 \tag{3.5.21}$$

$$\boldsymbol{N}_i = \frac{\boldsymbol{n}_i}{M_i} \tag{3.5.22}$$

$$\boldsymbol{n}_i = \rho_i \boldsymbol{v} + \boldsymbol{j}_i = m_i \sum \boldsymbol{n}_i + \boldsymbol{j}_i \tag{3.5.23a}$$

$$\boldsymbol{N}_i = c_i \boldsymbol{v}^* + \boldsymbol{J}_i^* = x_i \sum \boldsymbol{N}_i + \boldsymbol{J}_i^* \tag{3.5.23b}$$

Mechanisms of Diffusion

Ordinary Diffusion

Fick's law of ordinary diffusion is a linear relation between the rate of diffusion of a chemical species and the local concentration gradient of that species. It is exact for a binary gas mixture, for which the kinetic theory of gases gives

$$\boldsymbol{j}_1 = -\rho D_{12} \nabla m_1 \text{ kg/m}^2 \text{ s} \tag{3.5.24a}$$

on a mass basis, and

$$\boldsymbol{J}_1^* = -c D_{12} \nabla x_1 \text{ kg/m}^2 \text{ s} \tag{3.5.24b}$$

on a molar basis; D_{12} (m^2/s) is the binary diffusion coefficient (or mass diffusivity), and $D_{21} = D_{21}$. Equations 3.5.24a and 3.5.24b are mathematically equivalent; however, notice that it is incorrect to write

$$\boldsymbol{j}_1 = -D_{12} \nabla \rho_1 \tag{3.5.25}$$

since $\nabla \rho_1 \neq \rho \nabla m_1$ in general. Fick's law in the form of Equations 3.5.24a and 3.5.24b is also valid for dilute liquid and solid solutions, for which it is often possible to assume ρ (or c) constant, and then Equation 3.5.25 or its molar equivalent are good approximations.

Ordinary diffusion in multicomponent systems is described by the Stefan–Maxwell equations (Hirschfelder et al., 1954). These equations are difficult to use for engineering analysis. In gas mixtures containing species that do not have widely varying molecular weights, it is possible to model approximately the diffusion process by using an effective binary diffusion coefficient in Fick's law. This coefficient is a suitable average over the species in the mixture, and may be calculated from

$$D_{1m} = \frac{(1 - x_1)}{\sum_{i=2}^{n} (x_i / D_{1i})}; \quad x_1 \ll 1 \tag{3.5.26}$$

This equation works well for most mixtures of combustion gases (except those containing appreciable concentrations of H or H_2).

Binary diffusion coefficients at 300 K are of the order of 10^{-5} m^2/s in gases at 1 atm, 10^{-9} m^2/s in aqueous solutions, and 10^{-10} to 10^{-13} m^2/s in solids. However, the product ρD or (cD) is, at most, one order of magnitude different for gases and liquids. Data for diffusion coefficients may be found in Tables 3.5.2 through 3.5.5.

Molecules in a gas mixture, and in a liquid or solid solution, can diffuse by mechanisms other than ordinary diffusion governed by Fick's law. *Thermal diffusion* is diffusion due to a temperature gradient and is often called the *Soret effect*. Thermal diffusion is usually negligible compared with ordinary diffusion, unless the temperature gradient is very large. However, there are some important processes that depend on thermal diffusion, the most well known being the large-scale separation of uranium isotopes. *Pressure diffusion* is diffusion due to a pressure gradient and is also usually negligible unless the pressure gradient is very large. Pressure diffusion is the principle underlying

TABLE 3.5.2
Diffusion Coefficients in Air at 1 atm (1.013×10^5 Pa)

	Binary Diffusion Coefficient ($m^2/s\times10^4$)							
T (K)	O_2	CO_2	CO	C_7H_6	H_2	NO	SO_2	He
200	0.095	0.074	0.098	0.036	0.375	0.088	0.058	0.363
300	0.188	0.157	0.202	0.075	0.777	0.180	0.126	0.713
400	0.325	0.263	0.332	0.128	1.25	0.303	0.214	1.14
500	0.475	0.385	0.485	0.194	1.71	0.443	0.326	1.66
600	0.646	0.537	0.659	0.270	2.44	0.603	0.440	2.26
700	0.838	0.684	0.854	0.364	3.17	0.782	0.576	2.91
800	1.05	0.857	1.06	0.442	3.93	0.978	0.724	3.64
900	1.26	1.05	1.28	0.538	4.77	1.18	0.887	4.42
1000	1.52	1.24	1.54	0.641	5.69	1.41	1.060	5.26
1200	2.06	1.69	2.09	0.881	7.77	1.92	1.440	7.12
1400	2.66	2.17	2.70	1.13	9.90	2.45	1.870	9.20
1600	3.32	2.75	3.37	1.41	12.5	3.04	2.340	11.5
1800	4.03	3.28	4.10	1.72	15.2	3.70	2.850	13.9
2000	4.80	3.94	4.87	2.06	18.0	4.48	3.360	16.6

Note: Owing to the practical importance of water vapor-air mixtures, engineers have used convenient empirical formulas for $D_{H_2O\,air}$. A formula that has been widely used for many years is

$$D_{H_2O\,air} = 1.97\times10^{-5}\left(\frac{P_0}{P}\right)\left(\frac{T}{T_0}\right)^{1.685}\ \text{m}^2/\text{s};\quad 273\,\text{K} < T < 373\,\text{K}$$

where $P_0 = 1$ atm; $T_0 = 256$ K. More recently, the following formula has found increasing use. (Marrero, T.R. and Mason, E.A. 1992. Gaseous diffusion coefficients, *J. Phys. Chem. Ref. Data*, 1, 3–118):

$$D_{H_2O\,air} = 1.87\times10^{-10}\frac{T^{2.072}}{P};\quad 280\,\text{K} < T < 450\,\text{K}$$

$$= 2.75\times10^{-9}\frac{T^{1.632}}{P};\quad 450\,\text{K} < T < 1070\,\text{K}$$

for P in atmospheres and T in kelvins. Over the temperature range 290–330 K, the discrepancy between the two formulas is less than 2.5%. For small concentrations of water vapor in air, the older formula gives a constant value of $Sc_{H_2O\,air} = 0.61$ over the temperature range 273–373 K. On the other hand, the Marrero and Mason formula gives values of $Sc_{H_2O\,air}$ that vary from 0.63 at 280 K to 0.57 at 373 K.

TABLE 3.5.3
Schmidt Number for Vapors in Dilute Mixture in Air at Normal Temperature, Enthalpy of Vaporization, and Boiling Point at 1 atm

Vapor	Chemical Formula	Sc[a]	h_{fg}, J/kg$\times 10^{-6}$	T_{BP} (K)
Acetone	CH_3COCH_3	1.42	0.527	329
Ammonia	NH_3	0.61	1.370	240
Benzene	C_6H_6	1.79	0.395	354
Carbon dioxide	CO_2	1.00	0.398	194
Carbon monoxide	CO	0.77	0.217	81
Chlorine	Cl_2	1.42	0.288	238
Ethanol	CH_3CH_2OH	1.32	0.854	352
Helium	He	0.22	—	4.3
Heptane	C_7H_{16}	2.0	0.340	372
Hydrogen	H_2	0.20	0.454	20.3
Hydrogen sulfide	H_2S	0.94	0.548	213
Methanol	CH_3OH	0.98	1.110	338
Naphthalene	$C_{10}H_8$	2.35[b]	—	491
Nitric oxide	NO	0.87	0.465	121
Octane	C_8H_{18}	2.66	0.303	399
Oxygen	O_2	0.83	0.214	90.6
Pentane	C_5H_{12}	1.49	0.357	309
Sulfur dioxide	SO_2	1.24	0.398	263
Water vapor	H_2O	0.61	2.257	373

Note: With the Clausius–Clapeyron relation, one may estimate vapor pressure as

$$P_{sat} \simeq \exp\left\{-\frac{Mh_{fg}}{\mathcal{R}}\left(\frac{1}{T}-\frac{1}{T_{BP}}\right)\right\} \text{ atm} \quad \text{for } T \sim T_{BP}$$

[a] The Schmidt number is defined as $Sc=\mu/\rho D=\nu/D$. Since the vapors are in small concentrations, values for μ, ρ, and ν can be taken as pure air values.

[b] From a recent study by Cho, C., Irvine, T.F., Jr., and Karni, J. 1992. Measurement of the diffusion coefficient of naphthalene into air, *Int. J. Heat Mass Transfer*, 35, 957–966. Also, $h_{sg}=0.567\times 10^6$ J/kg at 300 K.

TABLE 3.5.4
Schmidt Numbers for Dilute Solution in Water at 300 K

Solute	Sc	M
Helium	120	4.003
Hydrogen	190	2.016
Nitrogen	280	28.02
Water	340	18.016
Nitric oxide	350	30.01
Carbon monoxide	360	28.01
Oxygen	400	32.00
Ammonia	410	17.03
Carbon dioxide	420	44.01
Hydrogen sulfide	430	34.08
Ethylene	450	28.05
Methane	490	16.04
Nitrous oxide	490	44.02

(*Continued*)

TABLE 3.5.4 (*Continued*)
Schmidt Numbers for Dilute Solution in Water at 300 K

Solute	*Sc*	*M*
Sulfur dioxide	520	64.06
Sodium chloride	540	58.45
Sodium hydroxide	490	40.00
Acetic acid	620	60.05
Acetone	630	58.08
Methanol	640	32.04
Ethanol	640	46.07
Chlorine	670	70.90
Benzene	720	78.11
Ethylene glycol	720	62.07
n-Propanol	730	60.09
i-Propanol	730	60.09
Propane	750	44.09
Aniline	800	93.13
Benzoic acid	830	122.12
Glycerol	1040	92.09
Sucrose	1670	342.3

Source: From Spalding, D.B., *Convective Mass Transfer*, McGraw-Hill, New York, 1963. With permission.

Note: Schmidt number $Sc = \mu/\rho D$; since the solutions are dilute, μ and ρ can be taken as pure water values. For other temperatures use $Sc/Sc_{300K} \simeq (\mu^2/\rho T)/(\mu^2\rho T)_{300K}$, where μ and ρ are for water, and T is absolute temperature. For chemically similar solutes of different molecular weights use $Sc_2/Sc_1 \simeq (M_2/M_1)^{0.4}$. A table of $(\mu^2/\rho T)/(\mu^2\rho T)_{300K}$ for water follows.

T (K)	(μ2/ρT)/(μ2/ρT)300K
290	1.66
300	1.00
310	0.623
320	0.429
330	0.296
340	0.221
350	0.167
360	0.123
370	0.097

TABLE 3.5.5
Diffusion Coefficients in Solids, $D = D_0 \exp(-E_a/\mathcal{R}T)$

System	D_0 (m²/s)	E_a (kJ/kmol)
Oxygen-Pyrex glass	6.19×10^{-8}	4.69×10^{4}
Oxygen-fused silica glass	2.61×10^{-9}	3.77×10^{4}
Oxygen-titanium	5.0×10^{-3}	2.13×10^{5}
Oxygen-titanium alloy (Ti-6A1-4V)	5.82×10^{-2}	2.59×10^{5}
Oxygen-zirconium	4.68×10^{-5}	7.06×10^{5}
Hydrogen-iron	7.60×10^{-8}	5.60×10^{3}

(*Continued*)

TABLE 3.5.5 (*Continued*)
Diffusion Coefficients in Solids, $D = D_0 \exp(-E_a/RT)$

System	D_0 (m²/s)	E_a (kJ/kmol)
Hydrogen-α-titanium	1.80×10^{-6}	5.18×10^{4}
Hydrogen-β-titanium	1.95×10^{-7}	2.78×10^{4}
Hydrogen-zirconium	1.09×10^{-7}	4.81×10^{4}
Hydrogen-Zircaloy^{-4}	1.27×10^{-5}	6.05×10^{5}
Deuterium-Pyrex glass	6.19×10^{-8}	4.69×10^{4}
Deuterium-fused silica glass	2.61×10^{-9}	3.77×10^{4}
Helium-Pyrex glass	4.76×10^{-8}	2.72×10^{4}
Helium-fused silica glass	5.29×10^{-8}	2.55×10^{4}
Helium-borosilicate glass	1.94×10^{-9}	2.34×10^{4}
Neon-borosilicate glass	1.02×10^{-10}	3.77×10^{4}
Carbon-FCC iron	2.3×10^{-5}	1.378×10^{5}
Carbon-BCC iron	1.1×10^{-6}	8.75×10^{4}

Source: Various sources.

the operation of a centrifuge. Centrifuges are used to separate liquid solutions and are increasingly being used to separate gaseous isotopes as well. *Forced diffusion* results from an external force field acting on a molecule. Gravitational force fields do not cause separation since the force per unit mass of a molecule is constant. Forced diffusion occurs when an electrical field is imposed on an electrolyte (for example, in charging an automobile battery), on a semiconductor, or on an ionized gas (for example, in a neon tube or metalion laser). Depending on the strength of the electric field, rates of forced diffusion can be very large.

Some interesting diffusion phenomena occur in porous solids. When a gas mixture is in a porous solid, such as a catalyst pellet or silica–gel particle, the pores can be smaller than the mean free path of the molecules. Then, the molecules collide with the wall more often than with other molecules. In the limit of negligible molecule collisions we have *Knudsen diffusion*, also called *free molecule flow* in the fluid mechanics literature. If the pore size approaches the size of a molecule, then Knudsen diffusion becomes negligible and *surface diffusion*, in which adsorbed molecules move along the pore walls, becomes the dominant diffusion mechanism.

Very small particles of 10^{-3} to 10^{-1} μm size—for example, smoke, soot, and mist—behave much like large molecules. Ordinary diffusion of such particles is called *Brownian motion* and is described in most elementary physics texts. Diffusion of particles due to a temperature gradient is called *thermophoresis* and plays an important role for larger particles, typically in the size range 10^{-1} to 1 μm. Diffusion of particles in a gas mixture due to concentration gradients of molecular species is called *diffusiophoresis. Forced diffusion* of a charged particle in an electrical field is similar to that for an ionized molecular species. Thermal and electrostatic precipitators are used to remove particles from power plant and incinerator stack gases, and depend on thermophoresis and forced diffusion, respectively, for their operation. Diffusion phenomena are unimportant for particles of size greater than about 1 μm in air at 1 atm; the motion of such particles is governed by the laws of Newtonian mechanics. Transport of particles is dealt with in the *aerosol science* literature.

Species Conservation Equation

The principle of conservation of a chemical species is used to derive the *species conservation equation*. On a mass basis this equation is

$$\frac{\partial \rho_i}{\partial t} + \nabla \cdot \boldsymbol{n}_i = \dot{r}_i''' \tag{3.5.27}$$

and on a molar basis

$$\frac{\partial c_i}{\partial t} + \nabla \cdot \boldsymbol{N}_i = \dot{R}_i''' \tag{3.5.28}$$

where $\dot{r}_i'''$ and $\dot{R}_i'''$ are the mass and molar rates of production of species i due to chemical reactions. Summing Equation 3.5.27 over all species gives the mass conservation or continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \boldsymbol{v} = 0 \tag{3.5.29}$$

The molar form is

$$\frac{\partial c}{\partial t} + \nabla \cdot c\boldsymbol{v}^* = \sum_i \dot{R}_i''' \tag{3.5.30}$$

since, in general, moles are not conserved in chemical reactions. A useful alternative form to Equation 3.5.27 can be obtained using Equations 3.5.23a and 3.5.29 and is

$$\rho \frac{Dm_i}{Dt} = \nabla \cdot \boldsymbol{j}_i + \dot{r}_i''' \tag{3.5.31}$$

where D/Dt is the substantial derivative operator.

If we consider a binary system of species 1 and 2 and introduce Fick's law, Equation 3.5.24a into Equation 3.5.31, then

$$\rho \frac{Dm_i}{Dt} = \nabla \cdot \left(\rho D_{12} \nabla m_1\right) + \dot{r}_1''' \tag{3.5.32}$$

When working on a mass basis we define a stationary medium as one in which the mass average velocity v is zero everywhere. Substituting in Equation 3.5.32 with no chemical reactions and assuming constant properties,

$$\frac{\partial m_1}{\partial t} = D_{12} \nabla^2 m_1 \tag{3.5.33}$$

which is the diffusion equation, and is the mass transfer analog to Fourier's equation for heat conduction. For steady diffusion, Equation 3.5.33 reduces to Laplace's equation

$$\nabla^2 m_1 = 0 \tag{3.5.34}$$

Notice that since properties have been assumed constant, any measure of concentration can be used in Equations 3.5.33 and 3.5.34, for example ρ_1, c_1, and x_1.

Diffusion in a Stationary Medium

Many problems involving diffusion in a stationary medium are governed by the diffusion equation (Equation 3.5.33). Often solutions may be obtained from their heat conduction analogs. Some important cases follow.

Steady Diffusion through a Plane Wall

The mass flow of species 1 across a plane wall of thickness L and cross-sectional area A is

$$\dot{m}_1 = \frac{\rho D_{12} A}{L}\left(m_{1,u} - m_{1,u'}\right) \text{kg/m}^2\text{s} \tag{3.5.35}$$

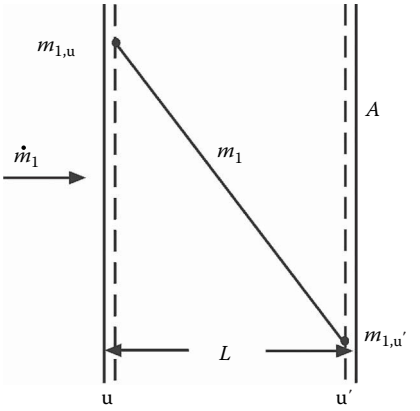


FIGURE 3.5.3 Steady diffusion across a plane wall.

where the u- and u′-surfaces are shown in Figure 3.5.3. Solubility data are required to relate the u- and u′-surface concentrations to s- and s′-surface concentrations. Alternatively for systems that obey Henry's law, a solubility S can be defined as the volume of solute gas (at STP of 0°C and 1 atm) dissolved in unit volume when the gas is at a partial pressure of 1 atm. Then, defining permeability p_{12} as the product $D_{12}S$, the volume flow of species 1 is

$$\dot{V}_1 = \frac{p_{12}A}{L}\left(P_{1,s} - P_{1,s'}\right)\mathrm{m}^3(\mathrm{STP})/\mathrm{s} \tag{3.5.36}$$

where the partial pressures P_1 are in atmospheres. The SI units for permeability are m^3 (STP)/m^2 s (atm/m). Permeability and solubility data are given in Table 3.5.6. For example, consider helium at 10^5 Pa contained in a 7056-glass vessel with a 1-mm-thick wall at 680 K. For a surface area of 0.01 m^2, the leakage rate into ambient air is

$$\dot{V} = \frac{\left(1.0\times10^{-12}\right)\left(0.01\right)}{\left(0.001\right)}\left(10^5 - 0\right) = 1.0\times10^{-6}\ \mathrm{m}^3(\mathrm{STP})/\mathrm{s}$$

where the value p_{12} was obtained from Table 3.5.6.

TABLE 3.5.6
Solubility and Permeability of Gases in Solids

Gas	Solid	Temperature (K)	S (m^3(STP)/m^3 atm) or S'[a]	Permeability[b] m^3(STP)/m^2 s (atm/m)
H_2	Vulcanized rubber	300	S=0.040	0.34×10^{-10}
	Vulcanized neoprene	290	S=0.051	0.053×10^{-10}
	Silicone rubber	300		4.2×10^{-10}
	Natural rubber	300		0.37×10^{-10}
	Polyethylene	300		0.065×10^{-10}
	Polycarbonate	300		0.091×10^{-10}
	Fused silica	400	$S' \simeq 0\ 035$	
		800	$S' \simeq 0.030$	
	Nickel	360	$S' \simeq 0.202$	
		440	$S' \simeq 0.192$	

(*Continued*)

TABLE 3.5.6 (*Continued*)
Solubility and Permeability of Gases in Solids

Gas	Solid	Temperature (K)	S (m^3(STP)/m^3 atm) or S'[a]	Permeability[b] m^3(STP)/m^2 s (atm/m)
He	Silicone rubber	300		2.3×10^{-10}
	Natural rubber	300		0.24×10^{-10}
	Polycarbonate	300		0.11×10^{-10}
	Nylon 66	300		0.0076×10^{-10}
	Teflon	300		0.047×10^{-10}
	Fused silica	300	$S' \simeq 0.018$	
		800	$S' \simeq 0.026$	
	Pyrex glass	300	$S' \simeq 0.006$	
		800	$S' \simeq 0.024$	
	7740 glass	470	$S=0.0084$	4.6×10^{-13}
	(94% $SiO_2+B_2O_3+P_2O_5$,	580	$S=0.0038$	1.6×10^{-12}
	5% $Na_2O+Li_2+K_2O$, 1% other oxides)	720	$S=0.0046$	6.4×10^{-12}
	7056 glass	390	$S'=0.0039$	1.2×10^{-14}
	(90% $SiO_2+B_2O_3+P_2O_5$,	680	$S'=0.0059$	1.0×10^{-12}
	8% $Na_2O+Li_2+K_2O$, 1 % PbO, 5% other oxides)			
O_2	Vulcanized rubber	300	$S=0.070$	0.15×10^{-10}
	Silicone rubber	300		3.8×10^{-10}
	Natural rubber	300		0.18×10^{-10}
	Polyethylene	300		4.2×10^{-12}
	Polycarbonate	300		0.011×10^{-10}
	Silicone-polycarbonate copolymer (57% silicone)	300		1.2×10^{-10}
	Ethyl cellulose	300		0.09×10^{-10}
N_2	Vulcanized rubber	300	$S=0.035$	0.054×10^{-10}
	Silicone rubber	300		1.9×10^{-12}
	Natural rubber	300		0.062×10^{-10}
	Silicone-polycarbonate copolymer (57% silicone)	300		0.53×10^{-10}
	Teflon	300		0.019×10^{-10}
CO_2	Vulcanized rubber	300	$S=0.090$	1.0×10^{-10}
	Silicone rubber	290		21×10^{-10}
	Natural rubber	300		1.0×10^{-10}
	Silicone-polycarbonate copolymer (57% silicone)	300		7.4×10^{-10}
	Nylon 66	300		0.0013×10^{-10}
H_2O	Silicone rubber	310		$0.91–1.8\times10^{-10}$
Ne	Fused silica	300–1200	$S \simeq 0.002$	
Ar	Fused silica	900–1200	$S' \simeq 0.01$	

Source: From various sources, including Geankoplis, C.J., *Transport Processes and Unit Operations*, 3rd ed., Prentice-Hall, Englewood Cliffs, NJ, 1993; Doremus, R.H., *Glass Science*, Wiley, New York, 1973; Alternose, V.O., *J. Appl. Phys*., 32, 1309–1316, 1961. With permission.

[a] Solubility S = volume of solute gas (0°C, 1 atm) dissolved in unit volume of solid when the gas is at 1 atm partial pressure. Solubility coefficient $S'=c_{1,u}/c_{1,s}$.

[b] Permeability $p_{12}=D_{12}S$.

In general, mass fractions are discontinuous across phase interfaces. Hence, Equation 3.5.35 cannot be generalized to a number of walls in series by simply adding diffusion resistances. However, equilibrium partial pressures P_1 are continuous, and for two walls A and B, Equation 3.5.36 becomes

$$\dot{V}_1 = \frac{P_{1,s} - P_{1,s'}}{\dfrac{L_A}{P_{1A}A} + \dfrac{L_B}{P_{1B}A}} \text{ m}^3 \text{ (STP)/s} \tag{3.5.37}$$

Transient Diffusion in a Semi-Infinite Solid

The typically low diffusion coefficients characterizing solids result in many situations where concentration changes are limited to a thin region near the surface (of thickness $\delta_c \sim (D_{12}t)^{1/2}$). Examples include case-hardening of mild steel and coloring of clear sapphires. Details of the geometry are then unimportant and semi-infinite solid model can be used (Figure 3.5.4). For an initial concentration $m_{1,0}$ and a u-surface concentration suddenly changed to $m_{1,u}$ at time $t=0$, the concentration distribution $m_1(z,t)$ is

$$\frac{m_1 - m_{1,0}}{m_{1,u} - m_{1,0}} = \text{erfc}\frac{z}{\left(4D_{12}t\right)^{1/2}} \tag{3.5.38}$$

and the dissolution rate is

$$\dot{m}_1 = j_{1,u}A = \rho A\left(\frac{D_{12}}{\pi t}\right)^{1/2}\left(m_{1,u} - m_{1,0}\right) \text{kg/s} \tag{3.5.39}$$

For example, consider a Pyrex glass slab at 800 K suddenly exposed to helium at 10^4 Pa. The molar equivalent to Equation 3.5.39 for an assumed constant solid phase molar concentration c is

$$\frac{\dot{M}_1}{A} = \left(\frac{D_{12}}{\pi t}\right)^{1/2}\left(c_{1,u} - c_{1,0}\right)$$

From Table 3.5.6, $S' = c_{1,u}/c_{1,s} \cong 0.024$; hence, $c_{1,u} = (0.024)(10^4)/(8314)(800) = 3.61\times10^{-5}$ kmol/m^3. From Table 3.5.4, $D_{12} = 4.76\times10^{-8}\exp[-(2.72\times10^4)(10^3)/(8314)(800)] = 7.97\times10^{-10}$ m^2/s. Hence,

$$\frac{\dot{M}_1}{A} = \left(\frac{7.97\times10^{-10}}{\pi t}\right)^{1/2}\left(3.61\times10^{-5} - 0\right) = 5.75\times10^{-10}/t \text{ kmol/s}$$

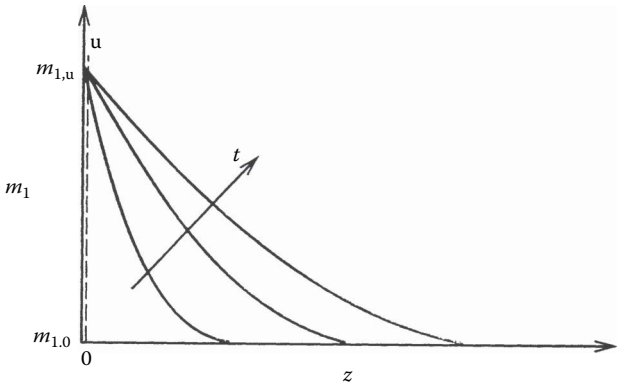


FIGURE 3.5.4 Transient diffusion in a plane slab.

Transient Diffusion in Slabs, Cylinders, and Spheres

Transient heat conduction in slates, cylinders, and spheres with surface convection is dealt with in Section 3.1. The analogous mass diffusion problem for the slab $-L < z < L$ is now considered. On a molar basis the governing differential equation is

$$\frac{\partial x_1}{\partial t} = D_{12} \frac{\partial^2 x_1}{\partial z^2} \tag{3.5.40}$$

with initial condition $x_1 = x_{1,0}$ at $t=0$. Boundary conditions are $\partial x_1/\partial z = 0$ at $z=0$, and at the surface $z=L$,

$$-cD_{12} \left.\frac{\partial x_1}{\partial z}\right|_{z=L} = G_{m1}\left(y_{1,s} - y_{1,e}\right) \tag{3.5.41}$$

The convective boundary condition is of the same form as Newton's law of cooling, and defines the mole transfer conductance G_{m1} (kmol/m^{2}s) (see also the section on mass and mole transfer conductances). Also, we have followed chemical engineering practice and denoted mole fraction x in the solid (or liquid) phase and y in the liquid (or gas) phase, to emphasize that generally mole fraction is not continuous across a phase interface. For example, consider absorption of a sparingly soluble gas into a liquid for which Henry's law, Equation 3.5.13, applies: then $y_{1,s} = \mathrm{He}\, x_{1,u}$.

In using heat conduction charts for mass diffusion problems, particular care must be taken with the evaluation of the Biot number. For heat conduction $\mathrm{Bi} = h_c L/k$, where k is the solid conductivity. For mass diffusion the Biot number accounts for the discontinuity in concentration across the phase interface. Using gas absorption into a plane layer of liquid, for example, when Equation 3.5.41 is put into an appropriate dimensionless form, the mass transfer Biot number is seen to be

$$\mathrm{Bi}_m = \frac{G_{m1}\mathrm{He}L}{cD_{12}} \tag{3.5.42}$$

For sparingly soluble gases, e.g., O_2 or CO_2 in water, He, and hence Bi_m, are very large, and the absorption process is liquid-side controlled; that is, a uniform gas-phase composition can be assumed. Often interface equilibrium data are in graphical or tabular form; then an effective Biot number at the concentration of concern must be used.

For example, consider a 2-mm-diameter droplet of water at 300 K entrained in an air flow at 1 atm pressure containing 1% by volume CO_2. From Table 3.5.5, $\mathrm{He} = C_{\mathrm{He}} = 1710$. The liquid phase molar density can be approximated by the pure water value of $c = \rho/M = 996/18 = 55.3$ kmol/m^3. The liquid phase diffusion coefficient is obtained from Table 3.5.4 as $D_{12} = \nu_{H_2O}/\mathrm{Sc}_{12} = 0.87\times10^{-6}/420 = 2.07\times10^{-9}$ m^2/s. For negligible relative motion between the droplet and gas, the Sherwood number (see the section on dimensionless groups) is approximately 2.0, and hence the gas phase mole transfer conductance is $G_{m1} = 2cD_{12}/D$. For the gas phase, the molar density $c = P/RT = (1.0133\times10^5)/(8314)(300) = 0.0406$ kmol/m^3 and $D_{12} = 0.157\times10^{-4}$ m^2/s from Table 3.5.2. Thus,

$$G_{m1} = \frac{(2)(0.0406)\left(0.157\times10^{-4}\right)}{(0.002)} = 6.37\times10^{-4} \text{ kmol/m}^2\text{s}$$

From Equation 3.5.42 with $L = R$ the droplet radius, the mass transfer Biot number is

$$\mathrm{Bi}_m = \frac{\left(6.37\times10^{-4}\right)(1710)(0.001)}{(55.3)\left(2.07\times10^{-9}\right)} = 9520$$

Thus, even for a small droplet with a relatively large gas-side mole transfer conductance, the absorption process is liquid-side controlled.

Diffusion in a Porous Catalyst

Porous catalysts are used to give a large surface area per unit volume of catalyst surface. Current practice for automobile catalytic convertors is to use a ceramic matrix as a support for a thin porous alumina layer that is impregnated with the catalyst (called a *washcoat*). A typical matrix has passages of hydraulic diameter 1 mm, and the washcoat may be about 20 μm thick. Pore sizes are of the order of 1 μm for which ordinary and Knudsen diffusion resistances are important. A simple model for diffusion in a porous catalyst is

$$J_1 = -cD_{1,\text{eff}} \cdot \nabla x_1 \text{ kmol/m}^2\text{s} \tag{3.5.43}$$

where the subscript eff denotes an effective diffusivity that accounts for the presence of the solid material. Assuming additive resistances,

$$\frac{1}{D_{1,\text{eff}}} = \frac{1}{D_{12,\text{eff}}} + \frac{1}{D_{\text{K1,eff}}} \tag{3.5.44}$$

and

$$D_{12,\text{eff}} = \frac{\varepsilon_v}{\tau} D_{12}; \quad D_{\text{K1,eff}} = \frac{\varepsilon_v}{\tau} D_{\text{K1,eff}} \tag{3.5.45}$$

where ε_v is the volume void fraction and τ is the tortuosity factor (usually between 4 and 8). From the kinetic theory of gases the Knudsen diffusion coefficient is

$$D_{\text{K1}} = 97 r_e \left(T/M_1\right)^{1/2} \text{ m}^2/\text{s} \tag{3.5.46}$$

for effective pore radius r_e in meters and T in kelvins.

When a chemical reaction takes place within a porous layer, a concentration gradient is set up, and surfaces on pores deep within the pellet are exposed to lower reactant concentrations than surfaces near the pore openings. For a first-order reaction, it is straightforward to obtain the concentration distribution. The results of such an analysis are conveniently given in the form of an effectiveness η_p, which is defined as the actual consumption rate of the reactant divided by that for an infinite diffusion coefficient. For a layer of thickness L exposed to reactants on one side, as shown in Figure 3.5.5.

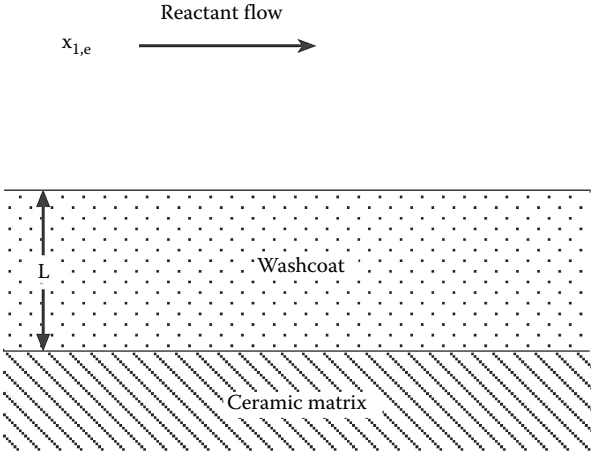


FIGURE 3.5.5 A catalyst layer.

$$\eta_p = \frac{\tanh bL}{bL}; \quad b = \left(\frac{k'' a_p}{D_{1,\text{eff}}} \right)^{1/2} \tag{3.5.47}$$

where k'' (m/s) is the rate constant for a first-order reaction and a_p (m^{-1}) is the catalyst area per unit volume. Notice that this effectiveness is analogous to the efficiency of a heat transfer fin.

For example, consider a 30-μm-thick porous alumina washcoat with a volume void fraction $\varepsilon_v = 0.8$, a tortuosity factor $\tau = 4.0$, average pore radius $r_e = 1$ μm, and catalytic surface area per unit volume $a_p = 7.1 \times 10^5$ cm^2/cm^3. For carbon monoxide oxidation by copper oxide at 800 K, 1 atm, the rate constant is approximately 4.2×10^{-4} m^2/s. To calculate the effectiveness of the washcoat, we first need to calculate the effective diffusion coefficient $D_{1,\text{eff}}$:

$$D_{12,\text{eff}} = \frac{\varepsilon_v}{\tau} D_{12} = \frac{0.8}{4.0}\left(1.06 \times 10^{-4}\right) = 2.12 \times 10^{-5} \text{ m}^2/\text{s}$$

where D_{12} is approximated as the CO-air value from Table 3.5.2.

$$D_{\text{K}1,\text{eff}} = \frac{\varepsilon_v}{\tau} D_{12} = \frac{0.8}{4.0}(97)\left(1 \times 10^{-6}\right)\left(800/28\right)^{1/2} = 1.04 \times 10^{-4} \text{ m}^2/\text{s}$$

$$\frac{1}{D_{1,\text{eff}}} = \frac{1}{2.12 \times 10^{-5}} + \frac{1}{1.04 \times 10^{-4}}; \quad D_{1,\text{eff}} = 1.76 \times 10^{-5} \text{ m}^2/\text{s}$$

$$b = \left[\frac{\left(4.2 \times 10^{-4}\right)\left(7.1 \times 10^5\right)\left(10^2\right)}{1.76 \times 10^{-5}} \right]^{1/2} = 4.2 \times 10^4 \text{ m}^{-1}; \quad bL = \left(4.2 \times 10^4\right)\left(30 \times 10^{-6}\right) = 1.236$$

$$\eta_p = \frac{\tanh 1.236}{1.236} = 68.3\%$$

In an automobile catalytic convertor, Equation 3.5.47 applies to the catalyst washcoat. However, the mass transfer problem also involves a convective process for transport of reactants from the bulk flow. Referring to Figure 3.5.6 there are two mass transfer resistances in series, and the consumption rate of species 1 per unit surface area of the washcoat is

$$J_{1,s} = \frac{-x_{1,e}}{\dfrac{1}{L\eta_p k'' c} + \dfrac{1}{G_{m1}}} \text{ kmol/m}^2\text{ s} \tag{3.5.48}$$

where G_{m1} is the mole transfer conductance describing convective transport to the washcoat surface (see the section on mass and mole transfer conductances). Notice that when $G_{m1} \ll L\eta_p k'' c$ the reaction rate is controlled by mass transfer from the gas stream to the washcoat surface; when

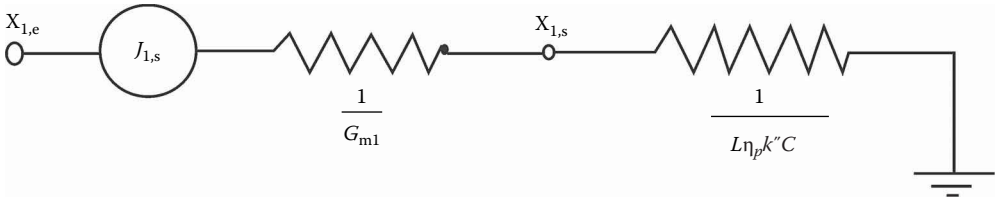


FIGURE 3.5.6 Equivalent circuit for mass transfer in an automobile catalytic converter.

$L\eta_p k'' c <<< G_{m1}$, the reaction rate is controlled by diffusion within the washcoat and the kinetics of the reaction.

Diffusion in a Moving Medium

Net mass transfer across a surface results in a velocity component normal to the surface, and an associated convective flux in the direction of mass transfer. This convective flow is called a *Stefan flow*. The solutions of a number of mass transfer problems, involving a Stefan flow induced by the mass transfer process itself, follow. When necessary to obtain an analytical result, properties are assumed constant. Thus, use of these results requires evaluation of properties at a suitable reference state.

Diffusion with One Component Stationary

As an example, consider the simple heat pipe shown in Figure 3.5.7 with the evaporator and condenser located at the ends only (a bad design!). Then, if the working fluid is species 1, and a noncondensable gas is species 2, the concentration distribution is

$$\left(\frac{1-x_1}{1-x_{1,s}}\right)=\left(\frac{1-x_{1,e}}{1-x_{1,s}}\right)^{z/L} \tag{3.5.49}$$

and the vapor flux along the heat pipe is

$$N_1=\frac{cD_{12}}{L}\ln\frac{1-x_{1,e}}{1-x_{1,s}}\,\text{kmol/m}^2\text{s} \tag{3.5.50}$$

Notice that $N_2=0$; that is, the gas is stationary. The rate of heat flow for a heat pipe of cross-sectional area of A_c is $\dot{Q}=N_1M_1h_{fg}A_c$. Evaluation of the cD product at a reference temperature $T_r=(1/2)(T_s+T_e)$ is adequate for most applications. Equation 3.5.50 applies to any situation where a one-dimensional model of mass transport is appropriate.

Heterogeneous Combustion

As an example, consider a small carbon particle entrained in a high-temperature airstream, as shown in Figure 3.5.8. The surface reaction is $2C+O_2 \rightarrow 2CO$ and there are no reactions in the gas phase. The stoichiometric ratio for the reaction is $r=4/3$ kg oxygen/kg carbon. The reaction is diffusion controlled at the temperatures under consideration, that is, $m_{O_2,s} \simeq 0$. The mass transfer rate is n_s, which we give the distinctive symbol $\dot{m}''$ since it is usually the desired result of an analysis; in this situation $\dot{m}''=n_{C,u}$ is the combustion rate of carbon, and for a spherical particle of radius R is given by

$$\dot{m}''=\frac{\rho D_{O_2,m}}{R}\ln\left[1+\frac{m_{O_2,e}-m_{O_2,s}}{m_{O_2,s}+4/3}\right]=0.160\,\frac{\rho D_{O_2,m}}{R}\,\text{kg/m}^2\text{s} \tag{3.5.51}$$

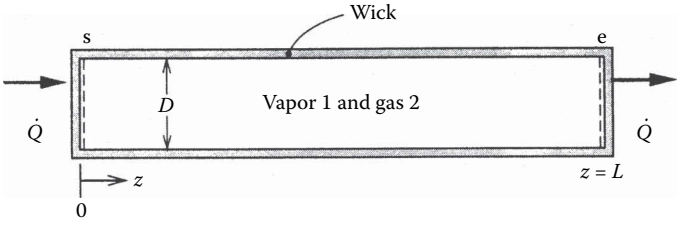


FIGURE 3.5.7 A simple heat pipe with the evaporator and condenser located at its ends.

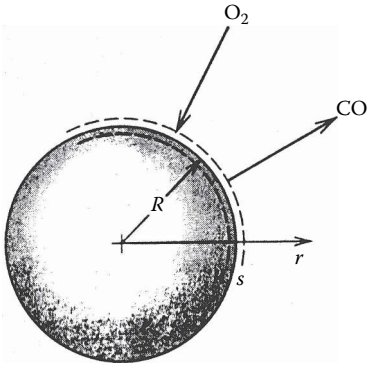


FIGURE 3.5.8 Combustion of a carbon particle in high-temperature air. The surface reaction is $2C+O_2 \rightarrow 2CO$.

The carbon particle temperature depends on its radius, and it is required to evaluate the property product ρD at an appropriate reference temperature: an energy balance on the particle should be performed for this purpose. The resulting particle lifetime τ is

$$\tau = \frac{\rho_{\text{solid}} D_0^2}{1.28\left(\rho D_{O_2,m}\right)_r} \text{ s} \tag{3.5.52}$$

for an initial particle diameter of D_0. Air properties at an average mean film temperature can be used to evaluate $\rho D_{O_2,m}$.

Consider a 10-μm-diameter carbon particle ignited in an airstream at 1500 K and 1 atm. An energy balance on the particle (including radiation to surroundings at 1500 K) shows that the average temperature of the particle is approximately 2550 K, and, thus, $T_r=(1/2)(1500+2550)=2025$ K or $\rho \simeq \rho_{\text{air}}=0.175\,\text{kg/m}^3$ and $D_{O_2,m} \simeq D_{O_2,\text{air}} = 4.89\times10^{-4}\,\text{m}^2/\text{s}$ (from Table 3.5.1). Then

$$\tau = \frac{(1810)\left(10\times10^{-6}\right)^2}{(1.28)(0.175)\left(4.89\times10^{-4}\right)} = 1.65\times10^{-3}\,\text{s}$$

Droplet Evaporation

Consider a small droplet of species 1 entrained in a gas stream, species 2 (Figure 3.5.9). This is a simultaneous heat and mass transfer problem, and the mass transfer rate can be obtained by solving simultaneously

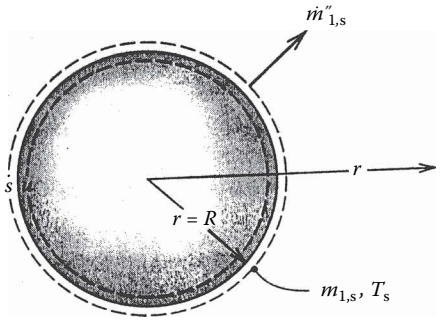


FIGURE 3.5.9 Evaporation of a droplet.

$$\dot{m}'' = \frac{\rho D_{12}}{R} \ln\left(1 + \frac{m_{1,e} - m_{1,s}}{m_{1,s} - 1}\right) = \frac{k/c_{p1}}{R} \ln\left(1 + \frac{c_{p1}(T_e - T_s)}{h_{fg}}\right) \text{kg/m}^2\,\text{s} \tag{3.5.53a}$$

$$m_{1,s} = m_{1,s}(T, P) \quad \text{(from vapor–pressure data)} \tag{3.5.53b}$$

Temperature T_s is the adiabatic vaporization temperature and is essentially the psychrometric wet-bulb temperature. Properties can be evaluated at mean film temperature and composition; alternatively, c_{p1} can be set equal to the reference specific heat and all properties evaluated using Hubbard's 1/3, rule, namely,

$$m_{1,r} = m_{1,s} + (1/3)(m_{1,e} - m_{1,s}) \tag{3.5.54a}$$

$$T_r = T_s + (1/3)(T_e - T_s) \tag{3.5.54b}$$

Droplet Combustion

Figure 3.5.10 shows a schematic of a volatile liquid hydrocarbon fuel droplet burning in air at zero gravity. The flame diameter is typically four to six times the droplet diameter. Heat is transferred from the flame to the droplet and serves to vaporize the fuel. In the flame, the vapor reacts with oxygen to form gaseous products, primarily CO_2 and H_2O. When a fuel droplet ignites, there is a short initial transient during which the droplet heats up, until further conduction into the droplet is negligible and the droplet attains a steady temperature (approximately the wet-bulb temperature, which is very close to the boiling point for a typical hydrocarbon fuel). The reaction in the flame can be modeled as a single-step reaction with a constant stoichiometric ratio, r, and heat of combustion Δh_c J/kg of fuel.

The burning (mass transfer) rate of the droplet is given by the Godsave–Spalding formula,

$$\dot{m}'' = \frac{k/c_p}{R} \ln[1 + B] \text{ kg/m}^2\text{s} \tag{3.5.55}$$

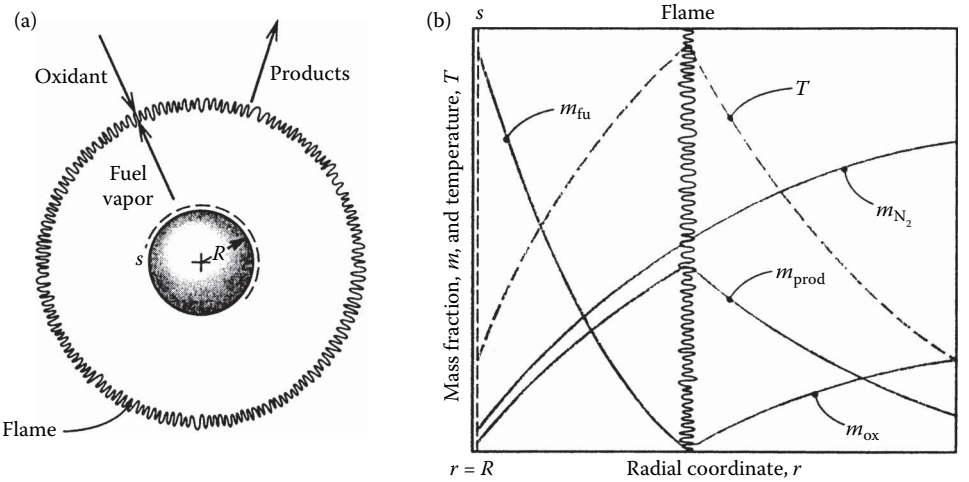


FIGURE 3.5.10 Combustion of a volatile fuel droplet burning in air: (a) schematic showing the flame and (b) concentration and temperature profiles.

where

$$B = \frac{m_{\text{ox,e}}\Delta h_{\text{c}} / r + c_{\text{p}}(T_{\text{e}} - T_{\text{s}})}{h_{\text{fg}}}$$

is the *mass transfer driving force* (or *transfer number*). The droplet lifetime is then

$$\tau = \frac{\rho_l D_0^2}{8(k / c_{\text{p}})\ln(1+B)}(\text{s}) \tag{3.5.56}$$

Based on experimental data for alkane droplets burning in air, Law and Williams (1972) recommend that properties be evaluated at a reference temperature $T_{\text{r}}=(\text{l}/2)(T_{\text{BP}}+T_{\text{flame}})$ where T_{flame} is the adiabatic flame temperature. The reference specific heat is $c_{\text{pr}}=c_{\text{pfu}}$, and the reference thermal conductivity is $k_{\text{r}}=0.4k_{\text{fu}}+0.6k_{\text{air}}$. Radiation has been ignored in the analysis leading to Equation 3.5.55 but is accounted for in using the Law and Williams reference-property scheme.

For example, consider a 1-mm-diameter *n*-octane droplet burning in air at 1 atm and 300 K, at near zero gravity. For *n*-octane (*n*-C_8H_{18}), $\rho_{\text{l}}=611\,\text{kg/m}^3$, $h_{\text{fg}}=3.03\times10^5\,\text{J/kg}$, $\Delta h_{\text{c}}=4.44\times10^7\,\text{J/kg}$, and $T_{\text{BP}}=399$ K. The flame temperature is $T_{\text{flame}}=2320$ K. At a reference temperature of (1/2) $(T_{\text{flame}}+T_{\text{BP}})=1360$ K, property values of *n*-octane vapor include $k=0.113$ W/m K, $c_{\text{p}}=4280\,\text{J/kg K}$. The reaction is

$$C_8H_{18}+12.5O_2 \rightarrow 8CO_2+9H_2O$$

Hence, the stoichiometric ratio $r=400/114.2=3.50$. Also $m_{\text{ox,e}}=0.231$ and $T_{\text{s}} \simeq T_{\text{BP}}=399$ K. Thus, the transfer number is

$$B = \frac{(0.231)\left(4.44\times10^7\right)/(3.50)+4280(300-399)}{3.03\times10^5} = 8.27$$

At $T_{\text{r}}=1360$ K, $k_{\text{air}}=0.085$ W/m K. Hence,

$$k_{\text{r}} = 0.4k_{\text{fu}}+0.6k_{\text{air}} = (0.4)(0.113)+(0.6)(0.085) = 0.096\,\text{W/m K}$$

and the droplet lifetime is

$$\tau = \frac{(611)\left(1\times10^{-3}\right)^2}{(8)(0.096/4280)\ln(1+8.27)} = 1.53\,\text{s}$$

Mass Convection

The terms *mass* convection or *convective mass transfer* are generally used to describe the process of mass transfer between a surface and a moving fluid, as shown in Figure 3.5.11. The surface may be that of a falling water film in an air humidifier, of a coke particle in a gasifier, or of a silica-phenolic heat shield protecting a reentry vehicle. As is the case for heat convection, the flow can be *forced* or *natural, internal* or *external*, and *laminar* or *turbulent*. In addition, the concept of whether the mass transfer rate is *low* or *high* plays an important role: when mass transfer rates are low, there is a simple analogy between heat transfer and mass transfer that can be efficiently exploited in the solution of engineering problems.

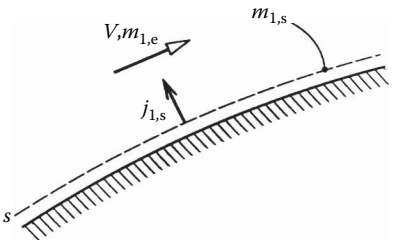


FIGURE 3.5.11 Notation for convective mass transfer into an external flow.

Mass and Mole Transfer Conductances

Analogous to convective heat transfer, the rate of mass transfer by convection is usually a complicated function of surface geometry and s-surface composition, the fluid composition and velocity, and fluid physical properties. For simplicity, we will restrict our attention to fluids that are either binary mixtures or solutions, or situations in which, although more than two species are present, diffusion can be adequately described using effective binary diffusion coefficients, as was discussed in the section on ordinary diffusion. Referring to Figure 3.5.11, we define the *mass transfer conductance* of species 1, g_{m1}, by the relation

$$J_{1,s} = g_{m1} \cdot \Delta m_1;\ \Delta m_1 = m_{1,s} - m_{1,e} \tag{3.5.57}$$

and the units of g_{m1} are seen to be the same as for mass flux (kg/m^{2}s). Equation 3.5.57 is of a similar form to Newton's law of cooling, which defines the heat transfer coefficient h_c. Why we should not use a similar name and notation (e.g., mass transfer coefficient and h_m) will become clear later. On a molar basis, we define the *mole transfer conductance* of species 1, G_{m1}, by a corresponding relation,

$$J_{1,s} = G_{m1} \cdot \Delta x_1;\ \Delta x_1 = x_{1,s} - x_{1,e} \tag{3.5.58}$$

where G_{m1} has units (kmol/m^{2}s).

Low Mass Transfer Rate Theory

Consider, as an example, the evaporation of water into air, as shown in Figure 3.5.12. The water–air interface might be the surface of a water reservoir, or the surface of a falling water film in a cooling tower or humidifier. In such situations, the mass fraction of water vapor in the air is relatively small; the highest value is at the s-surface, but even if the water temperature is as high as 50°C, the corresponding value of $m_{H_2O,s}$ at 1 atm total pressure is only 0.077. From Equation 3.5.54, the driving potential for diffusion of water vapor away from the interface is $\Delta m_1 = m_{1,s} - m_{1,e}$, and is small compared to unity, even if the free-stream air is very dry such that $m_{1,e} \simeq 0$. We then say that the mass transfer rate is *low* and the rate of evaporation (kg/s) of the water can be approximated as $j_{1,s}$; for a surface area A,

$$\dot{m}_1 = \left(m_{1,s} n_s + j_{1,s}\right) A \simeq j_{1,s} A \tag{3.5.59}$$

In contrast, if the water temperature approaches its boiling point, $m_{1,s}$ is no longer small, and of course, in the limit of $T_s = T_{BP}$, $m_{1,s} = 1$. The resulting driving potential for diffusion Δm_1 is then large, and we say that the mass transfer rate is *high*. Then, the evaporation rate cannot be calculated from Equation 3.5.59, as will be explained in the section on high mass transfer rate theory. For water evaporation into air, the error incurred in using low mass transfer rate theory is approximately (1/2) Δm_1, and a suitable criterion for application of the theory to engineering problems is $\Delta m_1 < 0.1$ or 0.2.

A large range of engineering problems can be adequately analyzed assuming low mass transfer rates. These problems include cooling towers and humidifiers as mentioned above, gas absorbers for

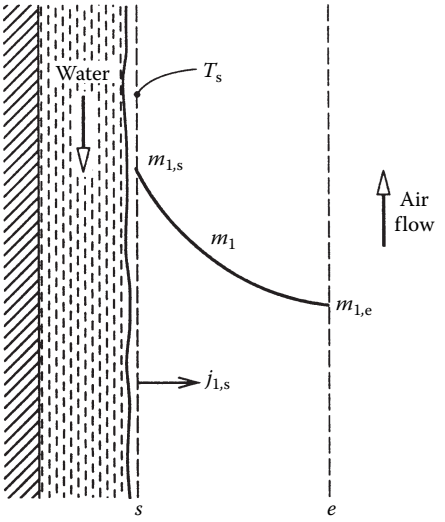


FIGURE 3.5.12 Evaporation of water into an air flow.

sparingly soluble gases, and catalysis. In the case of catalysis, the *net* mass transfer rate is actually zero. Reactants diffuse toward the catalyst surface and the products diffuse away, but the catalyst only promotes the reaction and is not consumed. On the other hand, problems that are characterized by high mass transfer rates include condensation of steam containing a small amount of noncondensable gas, as occurs in most power plant condensers; combustion of volatile liquid hydrocarbon fuel droplets in diesel engines and oil-fired power plants, and ablation of phenolic-based heat shields on reentry vehicles.

Dimensionless Groups

Dimensional analysis of convective mass transfer yields a number of pertinent dimensionless groups that are, in general, analogous to dimensionless groups for convective heat transfer. The most important groups are as follows.

1. The Schmidt number, $\mathrm{Sc}_{12}=\mu/\rho D_{12}$, which is a properties group analogous to the Prandtl number. For gas mixtures, $\mathrm{Sc}_{12}=O(1)$, and for liquid solutions, $\mathrm{Sc}_{12}=O(100)$ to $O(1000)$. There are not fluids for which $\mathrm{Sc}_{12} \ll 1$, as is the case of Prandtl number for liquid metals.
2. The Sherwood number (or mass transfer Nusselt number). $\mathrm{Sh}=g_{m1}L/\rho D_{12}$ $(=G_{m1}L/cD_{12})$ is a dimensionless conductance.
3. The mass transfer Stanton number $\mathrm{St}_m=g_{m1}/\rho V$ $(=G_{m1}/cV)$ is an alternative dimensionless conductance.

As for convective heat transfer, forced convection flows are characterized by a Reynolds number, and natural convection flows are characterized by a Grashof or Rayleigh number. In the case of Gr or Ra it is not possible to replace $\Delta\rho$ by $\beta\Delta T$ since density differences can result from concentration differences (and both concentration and temperature differences for simultaneous heat and mass transfer problems).

Analogy between Convective Heat and Mass Transfer

A close analogy exists between convective heat and convective mass transfer owing to the fact that conduction and diffusion in a fluid are governed by physical laws of identical form, that is, Fourier's

and Fick's laws, respectively. As a result, in many circumstances the Sherwood or mass transfer Stanton number can be obtained in a simple manner from the Nusselt number or heat transfer Stanton number for the same flow conditions. Indeed, in most gas mixtures Sh and St_m are nearly equal to their heat transfer counterparts. For dilute mixtures and solutions and low mass transfer rates, the rule for exploiting the analogy is simple: The *Sherwood or Stanton number is obtained by replacing the Prandtl number by the Schmidt number in the appropriate heat transfer correlation*. For example, in the case of fully developed turbulent flow in a smooth pipe

$$\mathrm{Nu_D} = 0.023\mathrm{Re_D^{0.8}Pr^{0.4}}; \quad \mathrm{Pr} > 0.5 \tag{3.5.60a}$$

which for mass transfer becomes

$$\mathrm{Sh_D} = 0.023\mathrm{Re_D^{0.8}Sc^{0.4}}; \quad \mathrm{Sc} > 0.5 \tag{3.5.60b}$$

Also, for natural convection from a heated horizontal surface facing upward,

$$\overline{\mathrm{Nu}} = 0.54\left(\mathrm{Gr_L Pr}\right)^{1/4}; \quad 10^5 < \mathrm{Gr_L Pr} < 2\times 10^7 \text{ (laminar)} \tag{3.5.61a}$$

$$\overline{\mathrm{Nu}} = 0.14\left(\mathrm{Gr_L Pr}\right)^{1/3}; \quad 2\times 10^7 < \mathrm{Gr_L Pr} < 3\times 10^{10} \text{ (turbulent)} \tag{3.5.61b}$$

which for isothermal mass transfer with $\rho_s < \rho_e$ become

$$\overline{\mathrm{Sh}} = 0.54\left(\mathrm{Gr_L Sc}\right)^{1/4}; \quad 10^5 < \mathrm{Gr_L Sc} < 2\times 10^7 \text{(laminar)} \tag{3.5.62a}$$

$$\overline{\mathrm{Sh}} = 0.14\left(\mathrm{Gr_L Sc}\right)^{1/3}; \quad 2\times 10^7 < \mathrm{Gr_L Sc} < 3\times 10^{10} \text{(turbulent)} \tag{3.5.62b}$$

With evaporation, the condition, $\rho_s < \rho_e$ will be met when the evaporating species has a smaller molecular weight than the ambient species, for example, water evaporating into air. Mass transfer correlations can be written down in a similar manner for almost all the heat transfer correlations given in Section 3.2. There are some exceptions: for example, there are no fluids with a Schmidt number much less than unity, and thus there are no mass transfer correlations corresponding to those given for heat transfer to liquid metals with $\mathrm{Pr} \ll 1$. In most cases it is important for the wall boundary conditions to be of analogous form, for example, laminar flow in ducts. A uniform wall temperature corresponds to a uniform concentration $m_{1,s}$ along the s-surface, whereas a uniform heat flux corresponds to a uniform diffusive flux $j_{1,s}$. In chemical engineering practice, the analogy between convective heat and mass transfer is widely used in a form recommended by Chilton and Colburn in 1934, namely, $\mathrm{St}_m/\mathrm{St} = (\mathrm{Sc/Pr})^{-2/3}$. The Chilton–Colburn form is of adequate accuracy for most external forced flows but is inappropriate for fully developed laminar duct flows.

For example, air at 1 atm and 300 K flows inside a 3-cm-inside-diameter tube at 10 m/s. Using pure-air properties the Reynolds number is $VD/\nu = (10)(0.03)/15.7\times 10^{-6} = 1.911\times 10^4$. The flow is turbulent. Using Equation 3.5.60b with $\mathrm{Sc_{12}} = 0.61$ for small concentrations of H_2O in air,

$$\mathrm{Sh_D} = (0.023)\left(1.911\times 10^4\right)^{0.8}(0.61)^{0.4} = 50.2$$

$$g_{m1} = \rho D_{12}\mathrm{Sh}/D = \rho\nu\mathrm{Sh}/\mathrm{Sc_{12}}D = \frac{(1.177)\left(15.7\times 10^{-6}\right)(50.2)}{(0.61)(0.03)} = 5.07\times 10^{-2} \text{ kg/m}^2\,\text{s}$$

Further insight into this analogy between convective heat and mass transfer can be seen by writing out Equations 3.5.60a and 3.5.60b as, respectively,

$$\frac{(h_c/c_p)D}{k/c_p} = 0.023 Re_D^{0.8} \left(\frac{\mu}{k/c_p} \right)^{0.4} \tag{3.5.63a}$$

$$\frac{g_m D}{\rho D_{12}} = 0.023 Re_D^{0.8} \left(\frac{\mu}{\rho D_{12}} \right)^{0.4} \tag{3.5.63b}$$

When cast in this form, the correlations show that the property combinations k/c_p and ρD_{12} play analogous roles; these are *exchange coefficients* for heat and mass, respectively, both having units kg/m s, which are the same as those for dynamic viscosity μ. Also, it is seen that the ratio of heat transfer coefficient to specific heat plays an analogous role to the mass transfer conductance, and has the same units (kg/m^{2}s). Thus, it is appropriate to refer to the ratio h_c/c_p as the *heat transfer conductance,* g_h, and for this reason we should not refer to g_m as the mass transfer *coefficient*.

Simultaneous Heat and Mass Transfer

Often problems involve simultaneous convective heat and mass transfer, for which the surface energy balance must be carefully formulated. Consider, for example, evaporation of water into air, as shown in Figure 3.5.13. With H_2O denoted as species 1, the steady-flow energy equation applied to a control volume located between the u- and s-surfaces requires that

$$\dot{m}(h_{1,s} - h_{1,u}) = A(q''_{cond} - q''_{conv} - q''_{rad}) \tag{3.5.64}$$

where it has been recognized that only species 1 crosses the u- and s-surfaces. Also, the water has been assumed to be perfectly opaque so that all radiation is emitted or absorbed between the u-surface and the interface.

If we restrict our attention to conditions for which low mass transfer rate theory is valid, we can write $\dot{m}/A \simeq j_{1,s} = g_{m1}(m_{1,s} - m_{1,e})$. Also, we can then calculate the convective heat transfer as if there were no mass transfer, and write $q_{conv} = h_c(T_s - T_e)$. Substituting in Equation 3.5.64 with $q_{conv} = -k\,\partial T / \partial y|_u$, $h_{1,s} - h_{1,u} = h_{fg}$, and rearranging, gives

$$-k \left. \frac{\partial T}{\partial y} \right|_u = h_c(T_s - T_e) + g_{m1}(m_{1,s} - m_{1,e}) h_{fg} + q''_{rad} \quad \text{W/m}^2 \tag{3.5.65}$$

It is common practice to refer to the convective heat flux $h_c(T_s - T_e)$ as the *sensible* heat flux, whereas the term $g_{m1}(m_{1,s} - m_{1,e})\,h_{fg}$ is called the *evaporative* or *latent* heat flux. Each of the terms in Equation 3.5.65 can be positive or negative, depending on the particular situation. Also, the evaluation of the conduction heat flux at the u-surface, $-k\partial T/\partial y|_u$, depends on the particular situation. Four examples are shown in Figure 3.5.13. For a water film flowing down a packing in a cooling tower (Figure 3.5.13b), this heat flux can be expressed in terms of convective heat transfer from the bulk water at temperature T_L to the surface of the film, $-k\partial T/\partial y|_u = h_{cL}(T_L - T_s)$. If the liquid-side heat transfer coefficient h_{cL} is large enough, we can simply set $T_s \simeq T_L$, which eliminates the need to estimate h_{cL}. The evaporation process is then *gas-side controlled*. Figure 3.5.13c shows film condensation from a steam-air mixture on the outside of a vertical tube. In this case we can write $k\partial T/\partial y|_u = U(T_s - T_c)$, where T_c is the coolant bulk temperature. The overall heat transfer coefficient U includes the resistances of the condensate film, the tube wall, and the coolant. Sweat cooling is shown in Figure 3.5.13d, with water from a reservoir (or *plenum chamber*) injected through a porous wall at a rate just sufficient to keep the wall surface wet. In this case, the conduction across the u-surface can be related to the reservoir conditions by application of the steady-flow energy equation to a control

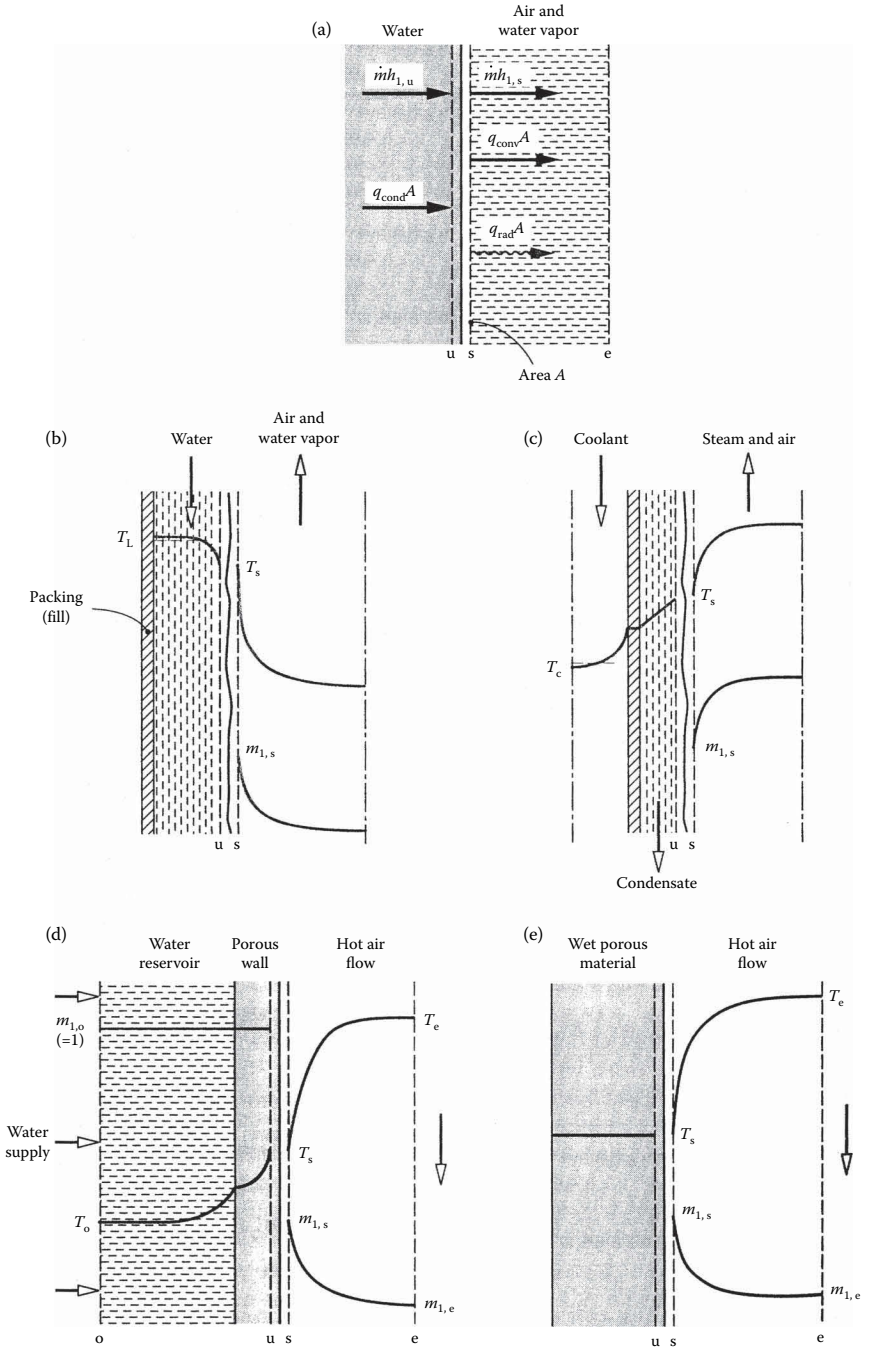


FIGURE 3.5.13 The surface energy balance for evaporation of water into an air stream.

volume located between the o- and u-surfaces. Finally, Figure 3.5.13e shows drying of a wet porous material (e.g., a textile or wood). During the constant-rate period of the process, evaporation takes place from the surface with negligible heat conduction into the solid; then $-k\partial T/\partial y|_u \simeq 0$. The term *adiabatic vaporization* is used to describe evaporation when $q_{cond}=0$; constant-rate drying is one example, and the wet-bulb psychrometer is another.

Consider a 1-m-square wet towel on a washline on a day when there is a low overcast and no wind. The ambient air is at 21°C, 1 atm, and 50.5% RH. In the constant-rate drying period the towel temperature is constant, and $q_{cond}=0$. An iterative calculation is required to obtain the towel temperature using correlations for natural convection on a vertical surface to obtain h_c and g_{m1}; q_{rad} is obtained as $q_{rad}=\sigma\varepsilon\left(T_s^4-T_e^4\right)$ with $\varepsilon=0.90$. The results are $T_s=17.8°C$, $h_c=1.69\,W/m^2\,K$, $g_{m1}=1.82\times10^{-3}\,kg/m^2s$, and the energy balance is

$$q_{cond}=h_c(T_s-T_e)+g_{m1}(m_{1,s}-m_{1,e})h_{fg}+q_{rad}$$

$$0=-5.4+21.7-16.3\,W/m^2$$

Evaluation of composition-dependent properties, in particular the mixture specific heat and Prandtl number, poses a problem. In general, low mass transfer rates imply small composition variations across a boundary layer, and properties can be evaluated for a mixture of the free-stream composition at the mean film temperature. In fact, when dealing with evaporation of water into air, use of the properties of dry air at the mean film temperature gives results of adequate engineering accuracy. If there are large composition variations across the boundary layer, as can occur in some catalysis problems, properties should be evaluated at the mean film composition and temperature.

The Wet- and Dry-Bulb Psychrometer

The wet- and dry-bulb psychrometer is used to measure the moisture content of air. In its simplest form, the air is made to flow over a pair of thermometers, one of which has its bulb covered by a wick whose other end is immersed in a small water reservoir. Evaporation of water from the wick causes the wet bulb to cool and its steady-state temperature is a function of the air temperature measured by the dry bulb and the air humidity. The wet bulb is shown in Figure 3.5.14. In order to determine the water vapor mass fraction $m_{1,e}$, the surface energy balance Equation 3.5.66 is used with conduction into the wick and q''_{rad} set equal to zero. The result is

$$m_{1,e}=m_{1,s}-\frac{c_p}{h_{fg}}\left(\frac{Pr}{Sc_{12}}\right)^{-2/3}(T_e-T_s) \tag{3.5.66}$$

Usually $m_{1,e}$ is small and we can approximate $c_p=c_{p\,air}$ and $(Pr/Sc_{12})^{-2/3}=1/1.08$. Temperatures T_s and T_e are the known measured wet- and dry-bulb temperatures. With T_s known, $m_{1,s}$ can be obtained

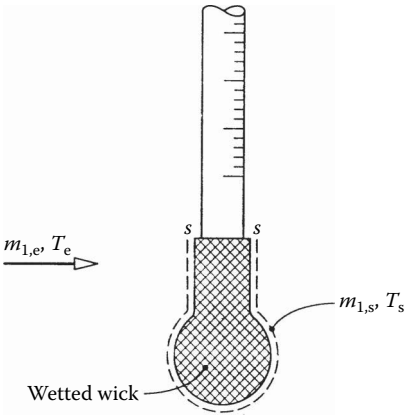


FIGURE 3.5.14 Wet bulb of a wet- and dry-bulb psychrometer.

using steam tables in the usual way. For example, consider an air flow at 1000 mbar with measured wet- and dry-bulb temperatures of 305.0 and 310.0 K, respectively. Then $P_{1,s}=P_{sat}(T_s)=P_{sat}(305.0\text{ K})=4714$ Pa from steam tables. Hence, $x_{1,s}=P_{1,s}/P=4714/10^5=0.04714$, and

$$m_{1,s}=\frac{0.04714}{0.04714+(29/18)(1-0.04714)}=0.02979$$

Also, $h_{fg}(305\text{ K})=2.425\times10^6$ J/kg, and $c_{p\ air}=1005$ J/kg K; thus

$$m_{1,e}=0.02979-\frac{1005}{(1.08)(2.425\times10^6)}(310-305)=0.02787$$

$$x_{1,e}=\frac{0.02787}{0.02787+(18/29)(1-0.02787)}=0.04415$$

$$P_{1,e}=x_{1,e}P=(0.04415)(10^5)=4412\text{ Pa}$$

By definition, the relative humidity is $RH=P_{1,e}/P_{sat}(T_e)$; $RH=4415/6224=70.9\%$.

In the case of other adiabatic vaporization processes, such as constant-rate drying or evaporation of a water droplet, $m_{1,e}$ and T_e are usually known and Equation 3.5.66 must be solved for T_s. However, the thermodynamic wet-bulb temperature obtained from psychrometric charts or software is accurate enough for engineering purposes.

High Mass Transfer Rate Theory

When there is net mass transfer across a phase interface, there is a convective component of the absolute flux (Kg/m^{2}s) of a species across the *s*-surface. From Equation 3.5.23a for species 1,

$$n_{1,s}=\rho_{1,s}v_s+j_{1,s} \tag{3.5.67}$$

During evaporation the convection is directed in the gas phase, with a velocity normal to the surface v_s. When the convective component cannot be neglected, we say that the mass transfer rate is *high*. There are two issues to consider when mass transfer rates are high. First, the rate at which species 1 is transferred across the *s*-surface is not simply the diffusive component $j_{1,s}$ as assumed in low mass transfer rate theory, but is the sum of the convective and diffusive components shown in Equation 3.5.67. Second, the normal velocity component v_s has a *blowing* effect on the concentration profiles, and hence on the Sherwood number. The Sherwood number is no longer analogous to the Nusselt number of conventional heat transfer correlations, because those Nusselt numbers are for situations involving impermeable surfaces, e.g., a metal wall, for which $v_s=0$.

Substituting for $j_{1,s}$ from Equation 3.5.57 into Equation 3.5.67 gives

$$\dot{m}''=g_{m1}\frac{m_{1,e}-m_{1,s}}{m_{1,s}-n_{1,s}/\dot{m}''}=g_{m1}B_{m1} \tag{3.5.68}$$

where $\dot{m}''=n_s$ is the mass transfer rate introduced in the section on heterogeneous combustion and B_{m1} is the *mass transfer driving force*. In the special case where only species 1 is transferred, $n_{1,s}/\dot{m}''=1$, for example, when water evaporates into air, and dissolution of air in the water is neglected. It is convenient to rewrite Equation 3.5.68 as

$$\dot{m}''=g_{m1}^*\left(g_{m1}/g_{m1}^*\right)B_{m1}\ \ (\text{kg/m}^2\text{s}) \tag{3.5.69a}$$

where

$$g_{m1}^{*} = \lim_{B_m \to 0} g_{m1} \tag{3.5.69b}$$

Now g_{m1}^{*} is the limit value of g_{m1} for zero mass transfer (i.e., $v_s = 0$), and Sh* can be obtained from conventional heat transfer Nusselt number correlations for impermeable surfaces. The ratio $\left(g_{m1}/g_{m1}^{*}\right)$ is termed as a *blowing factor* and accounts for the effect of v_s on the concentration profiles. Use of Equation 3.5.69 requires appropriate data for the blowing factor. For the constant-property laminar boundary layer on a flat plate, Figure 3.5.15 shows the effect of the Schmidt number on the blowing factor. The abscissa is a *blowing parameter* $B_m = \dot{m}''/g_m^{*}$.

The blowing velocity also affects the velocity and temperature profiles, and hence the wall shear stress and heat transfer. The curve for Sc = 1 in Figure 3.5.15 also gives the effect of blowing on shear stress as τ_s/τ_s^{*}, and the curve for Sc = 0.7 gives the effect of blowing on heat transfer for air injection into air as h_c/h_c^{*} (since Pr = 0.7 for air).

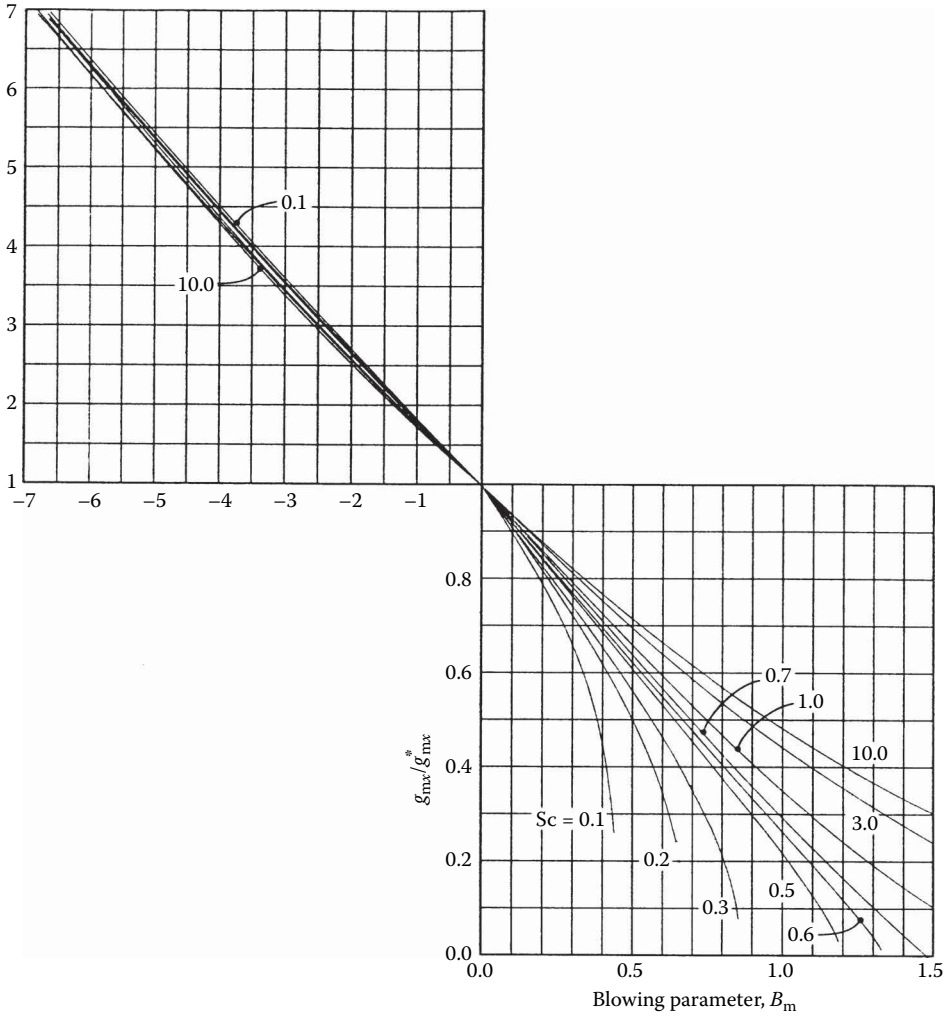


FIGURE 3.5.15 Effect of mass transfer on the mass transfer conductance for a laminar boundary layer on a flat plate: g_m/g_m^{*} vs. blowing parameter $B_m = \dot{m}''/g_m^{*}$.

Variable Property Effects of High Mass Transfer Rates

High mass transfer rate situations are usually characterized by large property variations across the flow, and hence property evaluation for calculating g_m and h_c is not straightforward. An often-encountered situation is transfer of a single species into an inert laminar or turbulent boundary layer flow. The effect of variable properties can be very large as shown in Figure 3.5.16 for laminar boundary layers, and Figure 3.5.17 for turbulent boundary layers.

A simple procedure for correlating the effects of flow type and variable properties is to use weighting factors in the exponential functions suggested by a constant-property Couette-flow model (Mills, 1995). Denoting the injected species as species *i*, we have

$$\frac{g_{m1}}{g_{m1}^*} = \frac{a_{mi}B_{mi}}{\exp(a_{mi}B_{mi})-1}; \quad B_{mi} = \frac{\dot{m}''}{g_{mi}^*} \tag{3.5.70a}$$

or

$$\frac{g_{m1}}{g_{mi}^*} = \frac{\ln(1+a_{mi}B_{mi})}{a_{mi}B_{mi}}; \quad B_{mi} = \frac{\dot{m}''}{g_{mi}} = \frac{m_{i,e}-m_{i,e}}{m_{i,s}-1}$$

$$\frac{\tau_s}{\tau_s^*} = \frac{a_{fi}B_f}{\exp(a_{fi}B_f)-1}; \quad B_f = \frac{\dot{m}''u_e}{\tau_s^*} \tag{3.5.70b}$$

$$\frac{h_c}{h_c^*} = \frac{a_{hi}B_h}{\exp(a_{hi}B_h)-1}; \quad B_h = \frac{\dot{m}''c_{pe}}{h_c^*} \tag{3.5.70c}$$

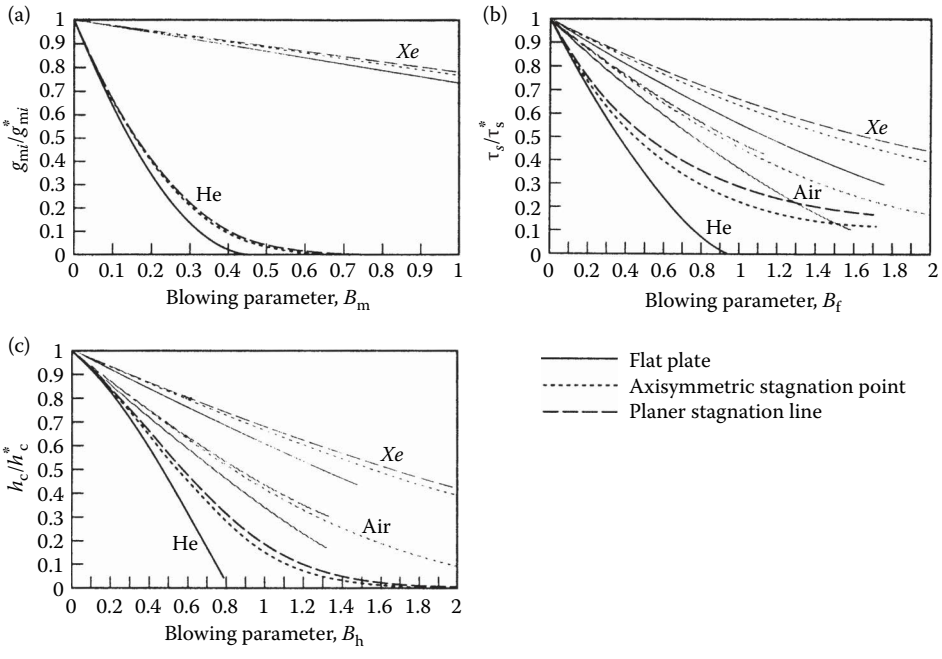


FIGURE 3.5.16 Numerical results for the effect of pressure gradient and variable properties on blowing factors for laminar boundary layers: low-speed air flow over a cold wall (T_s/T_e=0.1) with foreign gas injection: (a) mass transfer conductance, (b) wall shear stress, and (c) heat transfer coefficient. (From Wortman, A., Mass transfer in self-similar boundary-layer flows, Ph.D. dissertation, University of California, 1969. With permission.)

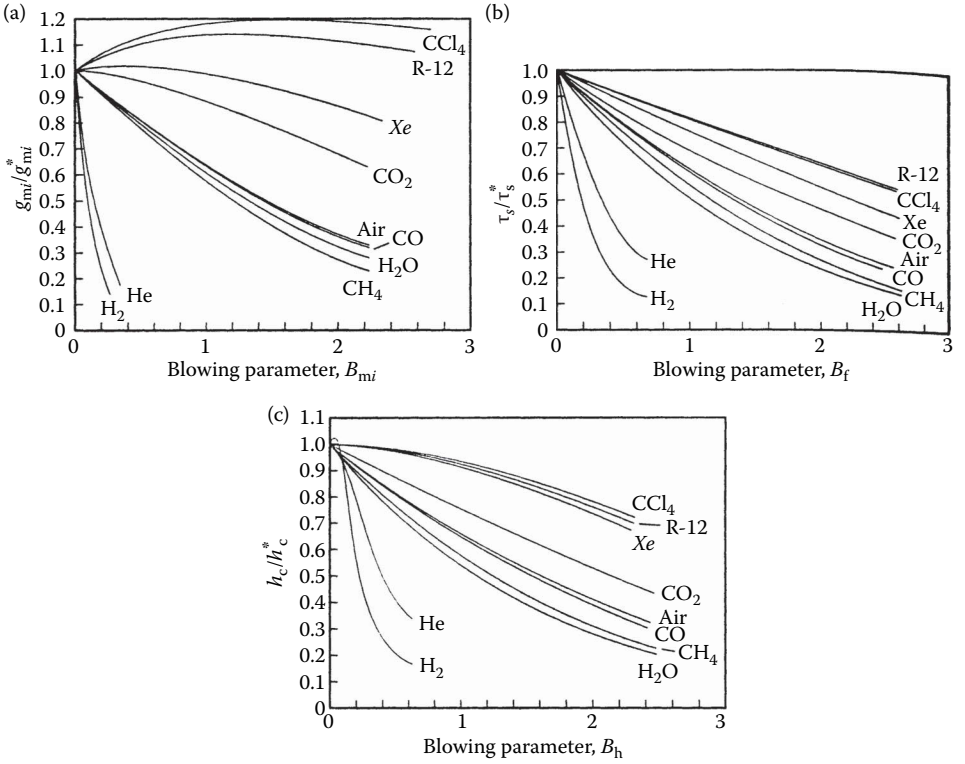


FIGURE 3.5.17 Numerical results for the effect of variable properties on blowing factors for a low-speed turbulent air boundary layer on a cold flat plate ($T_s/T_e=0.2$) with foreign gas injection: (a) mass transfer conductance, (b) wall shear stress, and (c) heat-transfer coefficient. (From Landis, R.B., Numerical solution of variable property turbulent boundary layers with foreign gas injection, Ph.D. dissertation, University of California, 1972. With permission.)

Notice that $g^*_{mi}, \tau^*_s, h^*_c$, and c_{pe} are evaluated using properties of the free-stream gas at the mean film temperature. The weighting factor a may be found from exact numerical solutions of boundary layer equations or from experimental data. Some results for laminar and turbulent boundary layers follow.

1. *Laminar boundary layers.* We will restrict our attention to low-speed air flows, for which viscous dissipation and compressibility effects are negligible, and use exact numerical solutions of the self-similar laminar boundary layer equations (Wortman, 1969). Least-squares curve fits of the numerical data were obtained using Equations 3.5.70c. Then, the weighting factors for axisymmetric stagnation-point flow with a cold wall ($T_s/T_e=0.1$) were correlated as

$$a_{mi} = 1.65\left(\frac{M_{\text{air}}}{M_i}\right)^{10/12} \tag{3.5.71a}$$

$$a_{\text{fi}} = 1.38\left(\frac{M_{\text{air}}}{M_i}\right)^{5/12} \tag{3.5.71b}$$

$$a_{\text{hi}} = 1.30\left(\frac{M_{\text{air}}}{M_i}\right)^{3/12}\left[\frac{c_{pi}}{\left(2.5\mathcal{R}/M_i\right)}\right] \tag{3.5.71c}$$

Notice that $c_{pi}/2.5\,\mathcal{R}/M_i)$ is unity for a monatomic species. For the planar stagnation line and the flat plate, and other values of the temperature ratio T_s/T_e, the values of the species weighting factors are divided by the values given by Equations 3.5.71a through 3.5.71c to give correction factors G_{mi}, G_{fi}, and G_{hi}, respectively. The correction factors are listed in Table 3.5.7.

The exponential relation blowing factors cannot accurately represent some of the more anomalous effects of blowing. For example, when a light gas such as H_2 is injected, Equation 3.5.70c indicates that the effect of blowing is always to reduce heat transfer, due to both the low density and high specific heat of hydrogen. However, at very low injection rates, the heat transfer is actually increased, as a result of the high thermal conductivity of H_2. For a mixture, $k \approx \Sigma x_i k_i$ whereas $c_p = \Sigma m_i c_{pi}$. At low rates of injection, the mole fraction of H_2 near the wall is much larger than its mass fraction; thus, there is a substantial increase in the mixture conductivity near the wall, but only a small change in the mixture specific heat. An increase in heat transfer results. At higher injection rates, the mass fraction of H_2 is also large, and the effect of high mixture specific heat dominates to cause a decrease in heat transfer.

TABLE 3.5.7
Correction Factors for Foreign Gas Injection into Laminar Air Boundary Layers

		$G_{mi}\ T_s/T_e$			$G_{fi}\ T_s/T_e$			$G_{hi}\ T_s/T_e$		
Geometry	**Species**	**0.1**	**0.5**	**0.9**	**0.1**	**0.5**	**0.9**	**0.1**	**0.5**	**0.9**
	H	1.14	1.36	1.47	1.30	1.64	1.79	1.15	1.32	—
	H_2	1.03	1.25	1.36	1.19	1.44	1.49	1.56	1.17	1.32
	He	1.05	1.18	1.25	1.34	1.49	1.56	1.18	1.32	—
	Air	—	—	—	1.21	1.27	1.27	1.17	1.21	—
	Xe	1.21	1.13	1.15	1.38	1.34	1.34	1.19	1.18	—
	$CC1_4$	1.03	0.95	1.00	1.00	1.03	1.03	1.04	1.04	—
Axisymmetric stagnation point	H	1.00	1.04	1.09	1.00	0.62	0.45	1.00	0.94	0.54
	H_2	1.00	1.06	1.06	1.00	0.70	0.62	1.00	1.00	1.01
	He	1.00	1.04	1.03	1.00	0.66	0.56	1.00	1.00	0.95
	C	1.00	1.01	1.00	1.00	0.79	0.69	1.00	0.99	0.87
	CH_4	1.00	1.01	1.00	1.00	0.88	0.84	1.00	1.00	1.00
	O	1.00	0.98	0.97	1.00	0.79	0.70	1.00	0.98	0.95
	H_2O	1.00	1.01	1.00	1.00	0.82	0.73	1.00	1.00	0.99
	Ne	1.00	1.00	0.98	1.00	0.83	0.75	1.00	0.97	0.95
	Air	—	—	—	1.00	0.87	0.82	1.00	0.99	0.97
	A	1.00	0.97	0.94	1.00	0.93	0.91	1.00	0.96	0.95
	CO_2	1.00	0.97	0.95	1.00	0.96	0.94	1.00	0.99	0.97
	Xe	1.00	0.98	0.96	1.00	0.96	1.05	1.00	1.06	0.99
	CCl_4	1.00	0.90	0.83	1.00	1.03	1.07	1.00	0.96	0.93
	I_2	1.00	0.91	0.85	1.00	1.02	1.05	1.00	0.97	0.94
Planar stagnation line	He	0.96	0.98	0.98	0.85	0.53	0.47	0.93	0.91	0.92
	Air	—	—	—	0.94	0.84	0.81	0.94	0.94	—
	Xe	0.92	0.87	0.83	0.90	0.93	0.95	0.93	0.93	—

Source: Based on numerical data of Wortman, A., Mass transfer in self-similar boundary-layer flows, Ph.D. dissertation, School of Engineering and Applied Science, University of California, 1969. Correlations developed by Dr. D.W. Hatfield.

2. *Turbulent boundary layers*. Here we restrict our attention to air flow along a flat plate for Mach numbers up to 6, and use numerical solutions of boundary layer equations with a mixing length turbulence model (Landis, 1972). Appropriate species weighting factors for $0.2 < T_s/T_e < 2$ are

$$a_{mi} = 0.79\left(\frac{M_{\text{air}}}{M_i}\right)^{1.33} \tag{3.5.72a}$$

$$a_{\text{fi}} = 0.91\left(\frac{M_{\text{air}}}{M_i}\right)^{0.76} \tag{3.5.72b}$$

$$a_{\text{hi}} = 0.86\left(\frac{M_{\text{air}}}{M_i}\right)^{0.76} \tag{3.5.72c}$$

In using Equation 3.5.70, the limit values for $\dot{m}''=0$ are elevated at the same location along the plate. Whether the injection rate is constant along the plate or varies as $x^{-0.2}$ to give a self-similar boundary layer has little effect on the blowing factors. Thus, Equation 3.5.72 has quite general applicability. Notice that the effects of injectant molecular weight are greater for turbulent boundary layers than for laminar ones, which is due to the effect of fluid density on turbulent transport. Also, the injectant specific heat does not appear in a_{hi} as it did for laminar flows. In general, c_{pi} decreases with increasing M_i and is adequately accounted for in the molecular weight ratio.

Reference state schemes. The reference state approach, in which constant-property data are used with properties evaluated at some reference state, is an alternative method for handling variable-property effects. In principle, the reference state is independent of the precise property data used and of the combination of injectant and free-stream species. A reference state for a boundary layer on a flat plate that can be used in conjunction with Figure 3.5.14 is (Knuth, 1963)

$$m_{1,\text{r}} = 1 - \frac{M_2}{M_2 - M_1}\frac{\ln\left(M_\text{e}/M_\text{s}\right)}{\ln\left(m_{2,\text{e}}M_\text{e}/m_{2,\text{s}}M_\text{s}\right)} \tag{3.5.73}$$

$$T_\text{r} = 0.5\left(T_\text{e} + T_\text{s}\right) + 0.2r^*\left(u_\text{e}^2/2c_{\text{pr}}\right) + 0.1\left[B_{\text{hr}} + \left(B_{\text{hr}} + B_{\text{mr}}\right)\frac{c_{\text{p1}} - c_{\text{pr}}}{c_{\text{pr}}}\right]\left(T_\text{s} - T_\text{e}\right) \tag{3.5.74}$$

where species 1 is injected into species 2 and r^* is the recovery factor for an impermeable wall. Use of the reference state method is impractical for hand calculations: a computer program should be used to evaluate the required mixture properties.

REFERENCES

Hirschfelder, J.O., Curtiss, C.F., and Bird, R.B. 1954. *Molecular Theory of Gases and Liquids*. John Wiley & Sons, New York.

Knuth, E.L. 1963. Use of reference states and constant property solutions in predicting mass-, momentum-, and energy-transfer rates in high speed laminar flows, *Int. J. Heat Mass Transfer*, 6, 1–22.

Landis, R.B. 1972. Numerical solution of variable property turbulent boundary layers with foreign gas injection, Ph.D. dissertation, School of Engineering and Applied Science, University of California, Los Angeles.

Law, C.K. and Williams, F.A. 1972. Kinetics and convection in the combustion of alkane droplets, *Combust. Flame*, 19, 393–405.

Mills, A.F. 1995. *Heat and Mass Transfer.* Richard D. Irwin, Chicago.
Wortman, A. 1969. Mass transfer in self-similar boundary-layer flows, Ph.D. dissertation, School of Engineering and Applied Science, University of California, Los Angeles.

Further Information

Geankoplis, C.J. 1993. *Transport Processes and Unit Operations*, 3rd ed. Prentice-Hall, Englewood Cliffs, NJ. This text gives a chemical engineering perspective on mass transfer.
Mills, A.F. 1995. *Heat and Mass Transfer.* Richard 101D. Irwin, Chicago. Chapter 11 treats mass transfer equipment relevant to mechanical engineering.
Strumillo, C. and Kudra, T. 1986. *Drying: Principles, Applications and Design*. Gordon and Breach, New York.
Mujamdar, A.S. Ed. 1987. *Handbook of Industrial Drying.* Marcel Dekker, New York.

4 Applications

Joshua D. Ramsey, Ken Bell, Ramesh K. Shah, Bengt Sundén, Zan Wu, Clement Kleinstreuer, Zelin Xu, D. Ian Wilson, Graham T. Polley, John A. Pearce, Kenneth R. Diller, Jonathan W. Valvano, David W. Yarbrough, Moncef Krarti, John Zhai, Jan Kośny, Christian K. Bach, Ian H. Bell, Craig R. Bradshaw, Eckhard A. Groll, Abhinav Krishna, Orkan Kurtulus, Margaret M. Mathison, Bryce Shaffer, Bin Yang, Xinye Zhang, Davide Ziviani, Robert F. Boehm, Anthony F. Mills, Santanu Bandyopadhyay, Shankar Narasimhan, Donald L. Fenton, Raj M. Manglik, Sameer Khandekar, Mario F. Trujillo, Rolf D. Reitz, Milind A. Jog, Prabhat Kumar, K.P. Sandeep, Sanjiv Sinha, Krishna Valavala, Jun Ma, Pradeep Lall, Harold R. Jacobs, Mangesh Chaudhari, Amit Agrawal, Robert J. Moffat, Tadhg O'Donovan, Jungho Kim, S.A. Sherif, Alan T. McDonald, Arturo Pacheco-Vega, Gerardo Diaz, Mihir Sen, K.T. Yang, Martine Rueff, Evelyne Mauret, Pawel Wawrzyniak, Ireneusz Zbicinski, Mariia Sobulska, P.S. Ghoshdastidar, Naveen Tiwari, Rajappa Tadepalli, Raj Ganesh S. Pala, Desh Bandhu Singh, and G. N. Tiwari

CONTENTS

4.1 Heat Exchangers for the Process and Energy Industries 459
Introduction 459
Principles of Heat Exchanger Design 461
Shell-and-Tube Heat Exchangers 471
Air-Cooled Heat Exchangers 484
References 500
Compact Heat Exchangers 501
Types and Description 501
Exchanger Heat Transfer and Pressure Drop Analysis 509
Heat Transfer and Flow Friction Correlations 522
Exchanger Design Methodology 527
Flow Maldistribution 535
Fouling in Heat Exchangers 535
References 542
Further Information 544
4.2 Application of Nanofluids in Heat Exchangers: Performance and Challenges 545

Introduction 545
Nanofluid Preparation and Characterization 547
Properties of Nanofluids 549
Properties of Nanofluids in Heat Exchangers 554
Results of Nanofluids in a Double-Pipe Helical Heat Exchanger 557
Results of Nanofluids in a PHE 560
Results of Hybrid Nanofluids in a PHE 563
Performance of Nanofluids in Heat Pipe Evaporators 566
Challenges of Nanofluids in Heat Exchangers 569
Summary 570
References 571
4.3 Convection Heat Transfer in Conduits with Nanofluids 575
Introduction 575
Preparation of Nanofluids 575
Thermophysical Properties of Nanofluids 578
Heat Transfer in Nanofluid Flow 584
Entropy Generation in Thermal Nanofluid Flow 588
Discrepancies among Enhanced Heat Transfer Studies Using Nanofluids 589
Nanofluid Flow Applications 590
Future Work 597
Conclusions 598
References 598
4.4 Fouling in Crude Oil and Food-Related Heat Transfer Equipment 603
Introduction: The Major Unresolved Problem in Heat Transfer 603
Fouling Mechanisms 609
Designing for Fouling Service 611
Fouling in Food Processing 613
Crude Oil Fouling 616
Closing Remarks 626
References 626
4.5 Bioheat Transfer 629
Abstract 629
Introduction 629
Coupling of Temperature History to Rate Processes 629
Tissue Thermal Transport Properties 630
Background 631
Measurement of Thermal Conductivity and Diffusivity 631
Estimation of the Convective Heat Transfer Coefficient 635
Temperature-Dependent Thermal Properties 637
Effect of Blood Flow on Temperature 641
Limitations of Pennes's Model 643
Continuum Models 644
Vasculature-Based Models 645
Hybrid Models 648
Thermal Measurements of Perfusion 648
Human Thermoregulation 651
Physiological Processes of Thermoregulation 652
Thermoregulatory Processes 652
Wissler Model of Human Thermoregulation 653
Therapeutic Heating 656
Heat Generation Modalities 656

Physiologic Effects of Local Heating 658
Tissue Effects: Elevated Temperatures 658
Theory of Rate Process Descriptions: The Arrhenius Model 659
Application of Kinetic Formulations in Thermal Damage and Cell Death Studies 662
Histologic Markers of Thermal Damage 673
Damage Rate Process Coefficients 680
Summary 683
Tissue Effects: Subzero Temperatures 684
Cryopreservation 684
Cryosurgery 692
Acknowledgments 692
Appendix A 693
Appendix B 696
References 706
4.6 Thermal Insulation 716
Introduction 716
Heat Transfer in Thermal Insulation 716
Insulation Systems (Nonvacuum and Vacuum) 719
Insulation Application 724
Acknowledgment 726
References 726
4.7 Energy Audit for Buildings 728
Introduction 728
Types of Energy Audits 728
General Procedure for a Detailed Energy Audit 730
Common Energy Conservation Measures 731
New Technologies 738
Verification Methods of Energy Savings 746
Summary 747
References 747
4.8 Advanced Energy-Efficient Building Envelope Systems 749
Introduction 749
Dynamic Insulation Materials 749
Variable Reflectivity Coating Roofs 753
Single-Layer Breathing Wall 760
Multilayer Living Walls 768
Summary and Conclusions 776
References 779
4.9 Use of Phase Change Materials in Buildings 781
Introduction and Scope 781
Properties of Commonly Used PCMs 781
Utilization of PCMs in the Building Envelope and Interior Fabric 785
Distributed and Localized PCM Applications in Buildings 787
Utilization of PCMs in Building Space-Conditioning and Water Heating Systems 789
Laboratory Evaluations of PCM Thermal Performance 789
Numerical Tools for PCM Thermal and Energy Performance Analysis 791
Examples of PCM Building Applications 792
References 794
4.10 Thermal Bridges in Building Structures 800
Introduction and Scope 800
Definition of Thermal Bridges and Common Consequences of Thermal Bridges 801

Types of Thermal Bridges and Most Common Locations 802
Parallel Path Method 809
Isothermal Planes (Series–Parallel Path) Method 810
Modified Zone Method 811
Linear Thermal Transmittance Method 811
Numerical Tools Used in Thermal Performance Analysis 813
Building Performance Standards Dealing with Building Thermal Bridging 814
Methods of Thermal Bridge Mitigation 815
References 816
4.11 Compressors 818
Introduction 818
Positive Displacement Compressors and Their Application to Refrigeration, Air Conditioning, and Heat Pumping 818
Modifications to Conventional Vapor Compression Cycles 819
Compressor Type by Construction 822
Compressor Type by Kinematic Linkages 824
Dynamic Compressors 835
Centrifugal Compressors 836
Axial Flow Compressors 841
Air Compressors 843
Compressor Types for Various Applications 844
Performance Considerations 846
Maintenance Requirements 848
Thermal Issues 848
Reduction of Energy Use 850
References 850
4.12 Pumps and Fans 853
Introduction 853
Pumps 853
Centrifugal and Other Velocity Head Pumps 854
Positive-Displacement Pumps 855
Pump/Flow Considerations 856
Fans 858
Vacuum Pumps 861
Defining Terms 862
References 862
Further Information 862
4.13 Cooling Towers 864
Introduction 864
Packing Thermal Performance 864
Thermal-Hydraulic Design of Cooling Towers 869
Cooling Tower Behavior 872
Range and Approach 873
Cooling Demand Curves 873
Legionnaires' Disease 874
References 874
Further Information 874
4.14 Pinch Technology 875
Introduction 875
Motivating Example 875
Pinch Technology 877

Energy Targeting 878
Cost Targeting 884
Targeting for Multiple Utilities 887
Network Design and Evolution 888
Integrating HEN with Other Energy Devices 892
Conclusion 894
References 895
4.15 Air-Conditioning Systems 897
Introduction 897
Properties of Moist Air 897
Thermodynamic Properties of Saturated Water 898
Thermodynamic Properties of Moist Air 899
Psychrometric Chart 901
Thermal Comfort Conditions 907
Load Calculations 908
Cooling Load 909
Heating Load 912
Refrigeration 915
Carnot Refrigeration Cycle 915
Vapor Compression Cycle 916
Compressors 918
Absorption Cycle 922
Refrigerant Selection 923
Exergy Analysis 928
Energy Distribution Systems 928
All-Air Systems 928
Air-and-Water Systems 930
All-Water Systems 931
References 931
4.16 Heat Transfer Enhancement 933
Introduction 933
Single-Phase Free Convection 935
Single-Phase Forced Convection 936
Performance Evaluation Criteria for Single-Phase Forced Convection in Tubes 942
Active and Compound Techniques for Single-Phase Forced Convection 944
Pool Boiling 945
Convective Boiling/Evaporation 947
Vapor-Space Condensation 948
Convective Condensation 949
References 949
Further Information 952
4.17 Heat Pipes 953
Introduction 953
What Is a Conventional Heat Pipe? 954
Brief History of Heat Pipe Development 956
Other Types of Heat Pipes 958
Conventional Heat Pipe Design and Limitations 967
Genealogy of Heat Pipe Family 971
Literature on Heat Pipes 974
Example Problem 975
Thermophysical Property Data 977

Summary and Outlook 979
Acknowledgments 979
Nomenclature 980
References 981
4.18 Liquid Atomization and Spraying 986
Spray Characterization 986
Atomizer Design Considerations 988
Atomizer Types 988
Modeling and Simulating Sprays 994
HRS Strategy 997
References 1000
Further Information 1002
4.19 Heat Transfer in Plasma Sprays 1003
Introduction 1003
Defining Features of Plasma Heat Transfer 1004
Modeling, Data, and Design Correlations 1005
Summary 1014
References 1014
4.20 Thermal Processing and Preservation of Foods 1017
Introduction 1017
Product Considerations 1017
Microorganisms 1021
Kinetics of Reactions 1023
Heating Equipment 1025
Process Considerations 1031
Outlook for the Future 1034
References 1034
4.21 Thermal Conduction in Electronic Microstructures and Nanostructures 1036
Fundamental Heat Conduction Physics 1039
Properties of Electrons in Metals 1042
Properties of Electrons in Semiconductors 1043
Properties of Phonons in Crystals 1045
Theory of Interfacial Thermal Transport 1049
Modeling Techniques for Materials, Devices, and Circuits 1054
Thermometry Techniques 1060
Thermal Conductivity/Conductance of Materials/Interfaces 1065
Power Dissipation and Heat Flow in Devices and Circuits 1071
Summary 1073
Acknowledgments 1074
References 1074
4.22 Role of Cooling in Electronics Reliability 1078
Introduction 1078
Understanding the Role of Temperature in Design 1078
Thermal Characteristics of PCBs 1081
Thermal Characteristics of Electronic Packages 1085
Thermal Interface Materials 1102
Computers 1105
Handheld Communication Devices 1111
Outdoor Telecommunication Electronics 1118
High-Altitude Airborne Electronics 1121
Summary 1124

References 1124
4.23 Direct Contact Heat Transfer 1131
Introduction 1131
Heat Transfer between Continuous Parallel Streams 1131
Heat Transfer to Jets and Sheets 1134
Sensible Heat Transfer to Dispersed Media: Drops, Particles, Bubbles 1136
Direct Contact Heat Transfer with Change of Phase 1140
Direct Contact Condensation 1142
Collapsing Bubble Condensation 1144
Condensation on Jets and Sheets 1145
Direct Contact Condensation Falling Films 1146
Direct Contact Heat Transfer with Solidification 1148
Summary 1148
References 1149
4.24 Heat Transfer in the Presence of Synthetic Jets 1153
Introduction 1153
Flow Characteristics of Synthetic Jet 1154
Heat Transfer Characteristics 1157
References 1161
4.25 Temperature and Heat Transfer Measurements 1162
Temperature Measurement 1162
Heat Flux Measurement 1176
Sensor Environmental Errors 1180
Evaluating the Heat Transfer Coefficient 1182
References 1186
4.26 Flow Measurement 1188
Introduction 1188
Direct Methods 1188
Restriction Flow Meters for Flow in Ducts 1188
Linear Flow Meters 1192
Traversing Methods 1198
Hot-Wire Anemometry 1202
Laser Doppler Anemometry 1206
Defining Terms 1215
References 1215
Additional Reading 1215
4.27 Applications of Artificial Neural Networks and Genetic Methods in Thermal Engineering 1217
Introduction 1217
Artificial Neural Networks 1220
Methodology of ANNs 1221
Application of ANNs to Compact Heat Exchangers 1225
Thermal System Dynamics and Control Using ANNs 1235
Additional Applications of ANNs in Thermal Engineering 1243
General Discussion of ANNs 1245
Genetic Algorithms (GAs) 1248
Methodology of GAs 1248
Applications of GAs to Compact Heat Exchangers 1250
Additional Applications of GAs in Thermal Engineering 1254
Genetic Programming 1255
Applications of GP to Compact Heat Exchangers 1257

Additional Applications of GPs in Thermal Engineering 1260
General Discussion of GAs and GPs 1262
Concluding Remarks 1262
Acknowledgments 1263
References 1263
4.28 Thermal Aspects of Paper Making 1270
Introduction 1270
The Paper Making Process 1270
The Cellulosic Fibers and Their Behavior during the Paper Making Process 1277
Paper: Composition, Grades, and Main Properties 1281
Drying of Paper: General Aspects 1285
The Different Stages of the Drying Process 1286
Global Heat Transfer Coefficient on the Drying Cylinder 1289
Organization of the Cylinder Dryers 1291
Dryer Fabrics 1292
Steam and Condensate Systems 1293
Hood, Ventilation, and Heat Recovery 1295
Physical Properties of Paper Drying 1297
Conclusion 1302
References 1303
4.29 Drying of Materials 1306
Introduction 1306
Principles of Drying 1306
Drying Techniques 1312
Principles of Dryer Design 1330
Safety in Drying 1333
Selection of Drying Method 1333
References 1336
4.30 Heat Transfer in Rotary Kilns 1338
Introduction 1338
Salient Features of a Rotary Kiln 1338
Firing of Rotary Kilns 1338
Phenomenon within the Solid Bed 1339
Transverse Bed Motion 1339
Axial Velocity of Solid Particles 1340
Heat Transfer Processes in a Rotary Kiln 1341
Rotary Kiln Modeling: Various Approaches 1341
Drying of Food Products in a Rotary Kiln 1342
Problem Formulation 1343
Basic Solution Methodology 1349
Results 1350
Summary 1350
Acknowledgments 1351
References 1351
4.31 Heat Transfer in Glass Manufacturing Processes 1352
Introduction 1352
Composition and Its Effect on Properties 1352
Melting Process 1355
Fining and Bubble Dynamics 1359
Heat Treatment for Glass Transformation 1360
Heat Strengthening 1366

Heat Transfer during Tempering 1366
Summary 1367
Acknowledgments 1368
References 1368
4.32 Solar Hydrogen as a "Renewable Reductant": Points and Counterpoints 1369
Introduction and Scope 1369
Physicochemical Principles of PEC Systems 1371
Model Identification and Refinement 1375
Aspects of Efficiency Metrics, Photoreactor Design, and Process Design 1375
Realizable Efficiencies Are via Materials: Some Hope, Some Hype, and the New Hope . 1378
Roads Less Travelled: Some Challenges in Attaining the Holy Grail 1380
Concluding Remarks 1381
Acknowledgments 1382
References 1382
4.33 Passive and Active Solar Distillation 1386
Introduction 1386
Distillation Process 1387
Basic Principle of Solar Distillation 1387
Material Requirement of Solar Still 1389
Classification of Solar Distillation System/Solar Still 1389
Historical Background 1390
Heat Transfer in Solar Still 1394
Thermal Modeling of Solar Stills 1399
Performance Parameters of Solar Still 1405
Passive Solar Stills 1414
Active Solar Still 1426
Climatic, Operational, and Design Parameter 1429
Appendix 1430
References 1434

4.1 HEAT EXCHANGERS FOR THE PROCESS AND ENERGY INDUSTRIES

Joshua D. Ramsey and Ken Bell

Introduction

A heat exchanger is a device to provide for the transfer of internal thermal energy (enthalpy) between two or more fluids, between a solid surface and a fluid, or between solid particulates and a fluid, in thermal contact without external heat interactions. The fluids may be single components or mixtures, and they may be single phase or multiphase. Typical applications involve heating or cooling of a fluid stream of concern, vaporization or condensation of a single- or multicomponent fluid stream, or heat recovery or rejection from a system. In other cases, the objective may be to sterilize, pasteurize, distill, fractionate, concentrate, or crystallize the process fluid. Some heat exchangers use internal mechanical devices such as rotating blades to enhance the heat-transfer process.

In most heat exchangers, the fluids are separated by a solid wall through which heat is transferred by conduction. There may also be unwanted solid deposits on the wall, commonly referred to as "fouling." These deposits can arise from corrosion of the wall material, chemical degradation of the fluid by thermal action, deposition of suspended solid material in the fluid, or freezing of the fluid by excessive cooling. Some heat-transfer devices operate by direct contact between two immiscible fluid streams, which are subsequently separated by gravitational settling or centrifugation; these are not treated in this section.

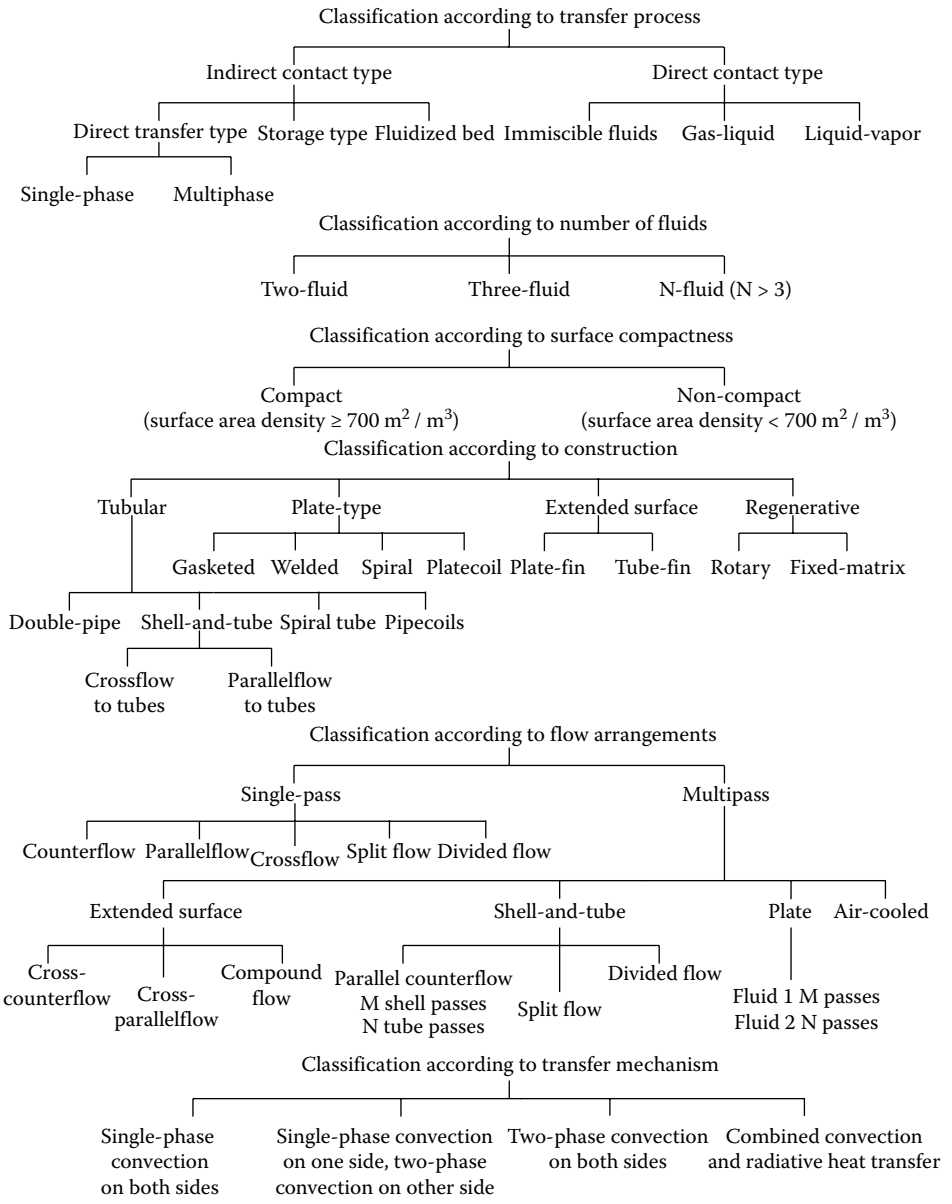


FIGURE 4.1.1 Classification of heat exchangers.

There are several ways to classify heat exchangers—by function, construction, heat-transfer mechanisms, etc. Figure 4.1.1 diagrams several of these classifications. Another important distinction is according to the ultimate market for the exchanger—mass utilization of a given product (as in automotive exchangers, residential and commercial environmental control, refrigeration, etc.), or custom applications (as in process plants, petroleum refineries, and thermal energy plants).

Typically, for mass utilization, a number of feasible configurations are constructed and tested over the plausible range of operating conditions for the application. Based on the test results and the manufacturability analysis, the final selection is made and large numbers of nominally identical units are then produced, with emphasis upon quality control. For most of these applications, a critical consideration in developing the test items is "compactness"—providing a high ratio of

heat-transfer surface area per unit of volume. A "compact" heat exchanger is usually defined as having greater than 700 m^2/m^3 (213 ft^2/ft^3) of surface area on at least one side of the heat-transfer surface, typically in contact with a gas. These heat exchangers are discussed in the second part of this section under the heading of *Compact Heat Exchangers.* The latter section is written by R. K. Shah.

This section deals with custom applications and with three types of heat exchangers commonly used in the process and energy industries: double pipe, shell-and-tube, and air-cooled. The starting point for design is a set of stated requirements, provided by the process engineer, giving the inlet fluid data (composition, flow rate, temperature, pressure) for each stream and the desired exit conditions at least for the process stream. Then an "optimum" design is developed to meet the stated requirements, with due allowance for possible future process changes and for design uncertainties. Construction is typically "one-off," unless multiple units are to be used side-by-side in a given plant. Equal emphasis is given to the "robustness" of the design process, construction quality, and maintainability under often severe operating conditions.

Double pipe exchangers are simple to design and construct from readily available components, and usually adaptable to high pressures and temperatures. However, they are relatively expensive per unit area of heat-transfer surface and are typically employed for smaller heat duties. The subsection titled "Principles of Heat Exchanger Design" introduces the basic design equations for all heat exchangers, using the double pipe as an example.

The most widely used heat exchanger type in the typical process plant is the shell-and-tube exchanger. Rugged, with a wide range of design/construction options and well-developed thermal/hydraulic and mechanical design procedures, these exchangers are described in more detail in the subsection titled, "Shell-and-Tube Heat Exchangers."

Wherever cooling and/or condensation to near atmospheric temperatures are required, the ready-to-hand solution may be an air-cooled exchanger. While air is a poor heat-transfer medium, it is available in generous supply anyplace one is likely to build a process or energy plant. The features and options of air-cooled exchangers are discussed in the subsection titled, "Air-cooled Heat Exchangers."

The most comprehensive literature source on heat exchangers is the five-volume "Heat Exchanger Design Handbook" (Hewitt 2008), available also through online subscription. Several excellent textbooks are available, especially Hewitt et al. (1994) and Serth and Lestina (2014). Many computer-based rating/design programs are available commercially, but no design program output—whether hand-calculated or computer-based—should be accepted until it has been vetted by one skilled in the art.

Principles of Heat Exchanger Design

In general, there are two basic approaches for heat exchanger design. One method is based on the known mean temperature difference (MTD) and is used to calculate the required area for heat transfer. This approach is used for custom design of a heat exchanger and is the typical approach used in the process and energy industries. The second approach is based on the number of transfer units (NTU) and is sometimes referred to as the effectiveness method. The NTU approach is used to determine outlet conditions for a given heat exchanger area. This approach is often thought of in terms of rating rather than designing, but for engineers working in industries where a standard, off-the-shelf configuration is selected or developed for a mass market (e.g., air conditioning systems), the NTU approach is used to determine the suitability of a standard configuration for a newly designed application. The focus here is on the MTD approach, but the reader is directed to other references for comprehensive coverage of the NTU approach (Çengel and Ghajar 2011).

MTD Approach

The MTD approach is used for custom design of a heat exchanger and is specifically used to calculate the area required for heat transfer. The approach requires knowledge of inlet and outlet conditions for each stream, which distinguishes it from the NTU approach. The MTD approach is

based on the local rate equation that relates the differential heat-transfer rate (dq) to a differential heat-transfer area (dA) *at a specific location*,

$$dq = U^* dA^* \left(T_h - T_c\right) \tag{4.1.1}$$

where dq is the differential heat-transfer rate, W (Btu/h); U is the overall heat-transfer coefficient, W/m^2·K (Btu/h·ft^2·°F); dA is the differential area, m^2 (ft^2); and T_h and T_c are the local temperatures of the hot (h) and cold (c) streams, K (°F), respectively. The superscript* on A* and U* simply indicates that a consistent basis must be used when defining each term. For example, for an exchanger with plain tubes, it is customary to use the total outside heat-transfer area of all of the tubes in the exchanger, A_o, as the reference area, and then U_o is the overall heat-transfer coefficient referenced to A_o. If the exchanger has low-finned tubes, A* may refer either to the total outside area including fins or to the inside tube heat-transfer area; the choice is optional, but must be spelled out.

The total required area (A) can be found by rearranging and integrating the local rate Equation 4.1.1,

$$A^* = \int_0^{q_T} \frac{dq}{U^*\left(T_h - T_c\right)} \tag{4.1.2}$$

where q_T is the total heat-transfer rate needed to achieve the desired outlet conditions. Solving the integral expression for the required area is complicated by the fact that the overall convective heat-transfer coefficient and the temperature difference between the hot and cold streams change during flow through the heat exchanger. The expression can be simplified, however, if we assume each term is constant and can be represented by an average value. In the case of the overall convective heat-transfer coefficient, U is calculated using physical properties (e.g., density, viscosity, thermal conductivity, and specific heat capacity) that also vary with temperature (and pressures for gases) and therefore with their position within the heat exchanger. The overall convective heat-transfer coefficient is typically calculated using properties of the streams at their average temperature. The under and overestimation of the physical properties tend to cancel one another leading to a reasonable estimate of U. Similarly, an average temperature difference based on $(T_h - T_c)$ over the length of the heat exchanger is substituted in place of the local temperature difference between the hot and cold streams. The assumption that U and $(T_h - T_c)$ can be represented by a constant, average value leads to

$$A^* = \frac{q_T}{U^* \cdot \mathrm{MTD}} \tag{4.1.3}$$

where MTD is the mean temperature difference between the hot and cold streams, K (°F).

To calculate the MTD, the inlet and outlet conditions of both streams must be known. For one stream, for example, the "hot stream," the inlet and outlet conditions are often set based on process constraints. For this stream, the heat transfer required to meet the outlet conditions can be determined from

$$q_T = \dot{m}_h \, c_{p,h}\left(T_{h,i} - T_{h,o}\right) = \dot{m}_c \, c_{p,c}\left(T_{c,o} - T_{c,i}\right) \tag{4.1.4}$$

where $\dot{m}$ is the mass flow rate kg/h (lb/h), c_p is the specific heat capacity of the respective stream J/kg·K (BTU/lb·°F), and T_i and T_o are the inlet and outlet temperatures, K (°F), respectively. As previously mentioned, the specific heat capacity is typically evaluated at the average temperature based on the inlet and outlet temperatures.

For the cold stream, the inlet and outlet conditions may also be set by process constraints, such as the supply and return temperatures of a utility stream. An overall energy balance requires that the heat-transfer rate for the cold stream equal that calculated for the hot stream (Equation 4.1.4), which in turn requires an appropriate mass flow rate for the given temperature change of the cold stream. Alternatively, it may be the mass flow rate that is fixed, in which case the exit temperature of the cold stream can be calculated using Equation 4.1.4 for the given mass flow rate and heat transfer rate.

The MTD can be calculated once the inlet and outlet conditions of the hot and cold streams are known. For pure countercurrent flow, where the hot and cold streams always move in directions opposite to one another, the MTD is the logarithmic mean temperature difference (LMTD), where

$$\text{MTD} = \text{LMTD}_{\text{countercurrent}} = \frac{\left(T_{\text{h,i}} - T_{\text{c,o}}\right) - \left(T_{\text{h,o}} - T_{\text{c,i}}\right)}{\ln\left(\dfrac{T_{\text{h,i}} - T_{\text{c,o}}}{T_{\text{h,o}} - T_{\text{c,i}}}\right)} \tag{4.1.5}$$

Very occasionally, cocurrent flow is specified, with the two streams flowing in the same direction. Cocurrent flow is often used when one wants to limit the local temperature of the tube wall. For cases with cocurrent flow, the MTD is again the LMTD, but the expression for the LMTD takes on a slightly different form,

$$\text{MTD} = \text{LMTD}_{\text{cocurrent}} = \frac{\left(T_{\text{h,i}} - T_{\text{c,i}}\right) - \left(T_{\text{h,o}} - T_{\text{c,o}}\right)}{\ln\left(\dfrac{T_{\text{h,i}} - T_{\text{c,i}}}{T_{\text{h,o}} - T_{\text{c,o}}}\right)} \tag{4.1.6}$$

Equations 4.1.5 and 4.1.6 apply to heat exchanger configurations with either countercurrent or cocurrent flow, such as that encountered in a double-pipe heat exchanger (Figure 4.1.2). Many heat exchangers, however, have configurations that produce a combination of countercurrent, cocurrent, and/or crossflow within a single heat exchanger. In these situations, the MTD approach must be altered to account for a corrected MTD, where

$$\text{MTD} = F \cdot \text{LMTD}_{\text{countercurrent}} \tag{4.1.7}$$

$\text{LMTD}_{\text{countercurrent}}$ is calculated from Equation 4.1.5, and F is the "configuration correction factor" for the particular flow configuration involved. F can be mathematically derived as a function of dimensionless temperature ratios for most flow configurations of interest and is given for some common cases in the next two subsections. Correction factors for numerous other flow configurations can be found in many heat-transfer references (e.g., TEMA 2007; Hewitt 2008, Vol. 1). F is equal to unity for pure countercurrent flow and is less than unity for all other cases, except for pure cocurrent

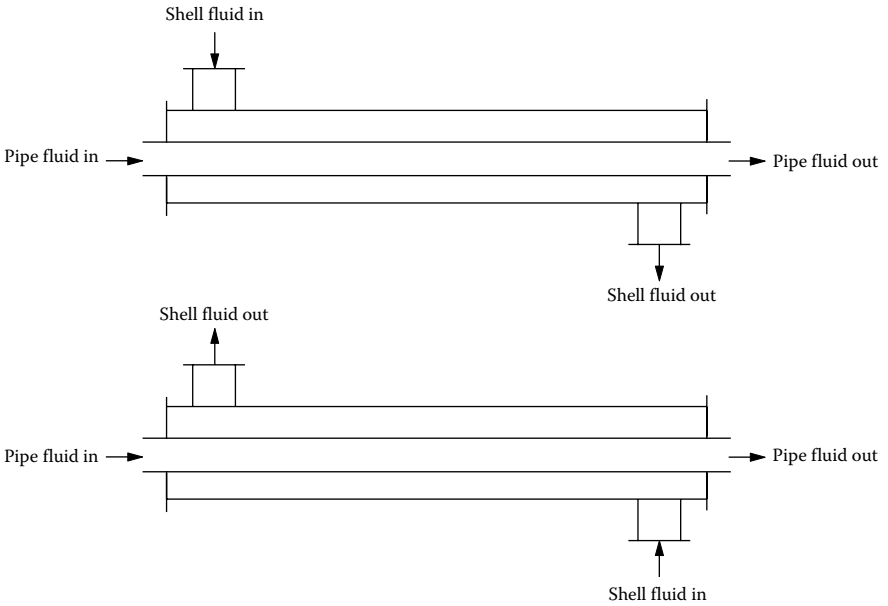


FIGURE 4.1.2 Double pipe heat exchangers with cocurrent flow (top) and countercurrent flow (bottom).

flow, where Equation 4.1.6 applies. Practical considerations usually limit the range of interest to values of F above 0.7 at the lowest and more comfortably to values above 0.8. Values of zero or below indicate conditions that violate the Second Law of Thermodynamics.

The remaining term needed to solve for the required area in Equation 4.1.3 is the overall heat-transfer coefficient, U^*. How the overall heat-transfer coefficient is related to the area, the individual (film) heat-transfer coefficients of the hot and cold streams, the fouling factors associated with both streams, and the construction material is based on the heat exchanger configuration. For example, the overall heat transfer coefficient for heat transfer through a flat wall is related to the other terms by

$$U^* = \frac{1}{\frac{1}{h_h} + R_{fh} + \frac{x_w}{k_w} + R_{fc} + \frac{1}{h_c}} \tag{4.1.8}$$

where h_h and h_c are, respectively, the hot and cold stream film heat transfer coefficients, W/m²·K (Btu/h ft² °F); R_{fh} and R_{fc} are the corresponding fouling resistances, m²·K/W (h ft² °F/Btu); x_w is the wall thickness, m (ft); and k_w is the thermal conductivity of the wall, W/m·K (Btu/h·ft·°F). The overall heat transfer coefficient for heat transfer between two fluids separated by a cylindrical wall has the relationship,

$$U^* = \frac{1}{\frac{A^*}{h_i A_i} + R_{fi}\frac{A^*}{A_i} + \frac{A^* \ln(d_o / d_i)}{2\pi L k_w} + R_{fo}\frac{A^*}{A_o} + \frac{A^*}{h_o A_o}} \tag{4.1.9}$$

where L is the length of the tube, m (ft), and k_w is the thermal conductivity of the tube wall, W/m·K (Btu/h·ft·°F). Notice, it is now more convenient to refer to the *inside* and *outside* of the tube, indicated by subscripts *i* and *o*, respectively, instead of referring to the hot and cold streams.

The equation for the overall heat transfer coefficient has introduced additional terms that are needed before the required area can be calculated. These terms, the film heat transfer coefficients and fouling factors, are addressed in detail in the next two subsections. For estimating purposes, however, it is convenient to assign typical values based on major classes of service (see Table 4.1.1).

TABLE 4.1.1
Typical Film Heat Transfer Coefficients for Shell-and-Tube Heat Exchangers

Fluid Conditions		h, W/m² K[a,b]	Fouling resistance, m² K/W[a]
	Sensible Heat Transfer		
Water[c]	Liquid	5000–7500	$1–2.5\times10^{-4}$
Ammonia	Liquid	6000–8000	$0–1\times10^{-4}$
Light organics[d]	Liquid	1500–2000	$0–2\times10^{-1}$
Medium organics[e]	Liquid	750–1500	$1–4\times10^{-4}$
Heavy organics[f]	Liquid		
	Heating	250–750	$2–10\times10^{-4}$
	Cooling	150–400	$2–10\times10^{-4}$
Very heavy organics[g]	Liquid		
	Heating	100–300	$4–30\times10^{-3}$
	Cooling	60–150	$4–30\times10^{-3}$
Gas[h]	Pressure 100–200 kN/m² abs	80–125	$0–1\times10^{-4}$
Gas[h]	Pressure 1 MN/m² abs	250–400	$0–1\times10^{-4}$
Gas[h]	Pressure 10 MN/m² abs	500–800	$0–1\times10^{-4}$

(Continued)

TABLE 4.1.1 (*Continued*)
Typical Film Heat Transfer Coefficients for Shell-and-Tube Heat Exchangers

Fluid Conditions		h, W/m² K[a,b]	Fouling resistance, m² K/W[a]
	Condensing Heat Transfer		
Steam, ammonia	Pressure 10 kN/m² abs, no noncondensables[i,j]	8000–12000	$0–1\times10^{-4}$
Steam, ammonia	Pressure 10 kN/m² abs, 1% noncondensables[k]	4000–6000	$0–1\times10^{-4}$
Steam, ammonia	Pressure 10 kN/m² abs, 4% noncondensables[k]	2000–3000	$0–1\times10^{-4}$
Steam, ammonia	Pressure 100 kN/m² abs, no noncondensables[i,j,k,l]	10000–15000	$0–1\times10^{-4}$
Steam, ammonia	Pressure 1 MN/m² abs, no noncondensables[i,j,k,l]	15000–25000	$0–1\times10^{-4}$
Light organics[d]	Pure component, pressure 10 kN/m² abs, no noncondensables	1500–2000	$0–1\times10^{-4}$
Light organics[d]	Pressure 10 kN/m² abs, 4% noncondensables[k]	750–1000	$0–1\times10^{-4}$
Light organics[d]	Pure component, pressure 100 kN/m² abs, no noncondensables	2000–1000	$0–1\times10^{-4}$
Light organics[d]	Pure component, pressure 1 MN/m² abs	3000–4000	$0–1\times10^{-4}$
Medium organics[e]	Pure component or narrow condensing range, pressure 100 kN/m² abs[m,n]	1500–4000	$1–3\times10^{-4}$
Heavy organics	Narrow condensing range, pressure 100 kN/m² abs[m,n]	600–2000	$2–5\times10^{-4}$
Light multicomponent mixtures, all condensable[d]	Medium condensing range, pressure 100 kN/m² abs[k,m,o]	1000–2500	$0–2\times10^{-4}$
Medium multicomponent mixtures, all condensable[e]	Medium condensing range, pressure 100 kN/m² abs[k,m,o]	600–1500	$1–4\times10^{-4}$
Heavy multicomponent mixtures, all condensable[f]	Medium condensing range, pressure 100 kN/m² abs[k,m,o]	300–600	$2–8\times10^{-4}$
	Vaporizing Heat Transfer[p,q]		
Water[r]	Pressure < 0.5 MN/m² abs, $\Delta T_{SH,max}=25$ K	3000–10000	$1–2\times10^{-4}$
Water[r]	Pressure > 0.5 MN/m² abs, pressure < 10 MN/m² abs, $\Delta T_{SH,max}=20$ K	4000–15000	$1–2\times10^{-4}$
Ammonia	Pressure < 3 MN/m² abs, $\Delta T_{SH,max}=20$ K	3000–5000	$0–2\times10^{-4}$
Light organics[d]	Pure component, pressure < 2 MN/m² abs, $\Delta T_{SH,max}=20$ K	1000–4000	$1–2\times10^{-4}$
Light organics[d]	Narrow boiling range,[s] pressure < 2 MN/m² abs, $\Delta T_{SH,max}=15$ K	750–3000	$0–2\times10^{-4}$
Medium organics[e]	Pure component, pressure < 2 MN/m² abs, $\Delta T_{SH,max}=20$ K	1000–3500	$1–3\times10^{-4}$
Medium organics[c]	Narrow boiling range,[s] pressure < 2 MN/m² abs, $\Delta T_{SH,max}=15$ K	600–2500	$1–3\times10^{-4}$
Heavy organics[f]	Pure component, pressure < 2 MN/m² abs, $\Delta T_{SH,max}=20$ K	750–2500	$2–5\times10^{-4}$
Heavy organics[g]	Narrow boiling range,[s] pressure < 2 MN/m² abs, $\Delta T_{SH,max}=15$ K	400–1500	$2–8\times10^{-4}$
Very heavy organics[h]	Narrow boiling range,[s] pressure < 2 MN/m² abs, $\Delta T_{SH,max}=15$ K	300–1000	$2–10\times10^{-4}$

Source: Hewitt, G.F., Ed., *Heat Exchanger Design Handbook*, Begell House, New York, 1998. With permission.

[a] Heat transfer coefficients and fouling resistances are based on area in contact with fluid. Ranges shown are typical, not all encompassing. Temperatures are assumed to be in normal processing range; allowances should be made for very high or low temperatures.

[b] Allowable pressure drops on each side are assumed to be about 50–100 kN/m² except for (1) low-pressure gas and two-phase flows, where the pressure drop is assumed to be about 5% of the absolute pressure; and (2) very viscous organics, where the allowable pressure drop is assumed to be about 150–250 kN/m².

[c] Aqueous solutions give approximately the same coefficients as water.

[d] Light organics include fluids with liquid viscosities less than about 0.5×10^{-3} Ns/m², such as hydrocarbons through C_8, gasoline, light alcohols and ketones, etc.

(Continued)

[e] Medium organics include fluids with liquid viscosities between about 0.5×10^{-3} and 2.5×10^{-1} Ns/m^2, such as kerosene, straw oil, hot gas oil, and light crudes.

[f] Heavy organics include fluids with liquid viscosities greater than 2.5×10^{-3} Ns/m^2, but not more than 50×10^{-3} Ns/m^2, such as cold gas oil, lube oils, fuel oils, and heavy and reduced crudes.

[g] Very heavy organics include tars, asphalts, polymer melts, greases, etc., having liquid viscosities greater than about 50×10^{-3} Ns/m^2. Estimation of coefficients for these materials is very uncertain and depends strongly on the temperature difference, because natural convection is often a significant contribution to heat transfer in heating, whereas congelation on the surface and particularly between fins can occur in cooling. Since many of these materials are thermally unstable, high surface temperatures can lead to extremely severe fouling.

[h] Values given for gases apply to such substances as air, nitrogen, carbon dioxide, light hydrocarbon mixtures (no condensation), etc. Because of the very high thermal conductivities and specific heats of hydrogen and helium, gas mixtures containing appreciable fractions of these components will generally have substantially higher heat transfer coefficients.

[i] Superheat of a pure vapor is removed at the same coefficient as for condensation of the saturated vapor if the exit coolant temperature is less than the saturation temperature (at the pressure existing in the vapor phase) and if the (constant) saturation temperature is used in calculating the MTD. But see note k for vapor mixtures with or without noncondensable gas.

[j] Steam is not usually condensed on conventional low-finned tubes; its high surface tension causes bridging and retention of the condensate and a severe reduction of the coefficient below that of the plain tube.

[k] The coefficients cited for condensation in the presence of noncondensable gases or for multicomponent mixtures are only for very rough estimation purposes because of the presence of mass transfer resistances in the vapor (and to some extent, in the liquid) phase. Also, for these cases, the vapor-phase temperature is not constant, and the coefficient given is to be used with the MTD estimated using vapor-phase inlet and exit temperatures, together with the coolant temperatures.

[l] As a rough approximation, the same relative reduction in low-pressure condensing coefficients due to noncondensable gases can also be applied to higher pressures.

[m] Absolute pressure and noncondensables have about the same effect on condensing coefficients for medium and heavy organics as for light organics. For large fractions of noncondensable gas, interpolate between pure component condensation and gas cooling coefficients.

[n] Narrow condensing range implies that the temperature difference between dew point and bubble point is less than the smallest temperature difference between vapor and coolant at any place in the condenser.

[o] Medium condensing range implies that the temperature difference between dew point and bubble point is greater than the smallest temperature difference between vapor and coolant, but less than the temperature difference between inlet vapor and outlet coolant.

[p] Boiling and vaporizing heat transfer coefficients depend very strongly on the nature of the surface and the structure of the two-phase flow past the surface in addition to all of the other variables that are significant for convective heat transfer in other modes. The flow velocity and structure are very much governed by the geometry of the equipment and its connecting piping. Also, there is a maximum heat flux from the surface that can be achieved with reasonable temperature differences between surface and saturation temperatures of the boiling fluid; any attempt to exceed this maximum heat flux by increasing the surface temperature leads to partial or total coverage of the surface by a film of vapor and a sharp decrease in the heat flux.

Therefore, the vaporizing heat transfer coefficients given in this table are only for very rough estimating purposes and assume the use of plain or low-finned tubes without special nucleation enhancement. $\Delta T_{SH,max}$ is the maximum allowable temperature difference between surface and saturation temperature of the boiling liquid. No attempt is made in this table to distinguish among the various types of vapor-generation equipment, since the major heat transfer distinction to be made is the propensity of the process stream to foul. Severely fouling streams will usually call for a vertical thermosiphon or a forced-convection (tube-side) reboiler for ease of cleaning.

[q] Subcooling heat load is transferred at the same coefficient as latent heat load in kettle reboilers, using the saturation temperature in the MTD. For horizontal and vertical thermosiphons and forced-circulation reboilers, a separate calculation is required for the sensible heat transfer area, using appropriate sensible heat transfer coefficients and the liquid temperature profile for the MTD.

[r] Aqueous solutions vaporize with nearly the same coefficient as pure water if attention is given to boiling point elevation, if the solution does not become saturated, and if care is taken to avoid dry wall conditions.

[s] For boiling of mixtures, the saturation temperature (bubble point) of the final liquid phase (after the desired vaporization has taken place) is to be used to calculate the MTD. A narrow-boiling-range mixture is defined as one for which the difference between the bubble point of the incoming liquid and the bubble point of the exit liquid is less than the temperature difference between the exit hot stream and the bubble point of the exit boiling liquid. Wide-boiling-range mixtures require a case-by-case analysis and cannot be reliably estimated by these simple procedures.

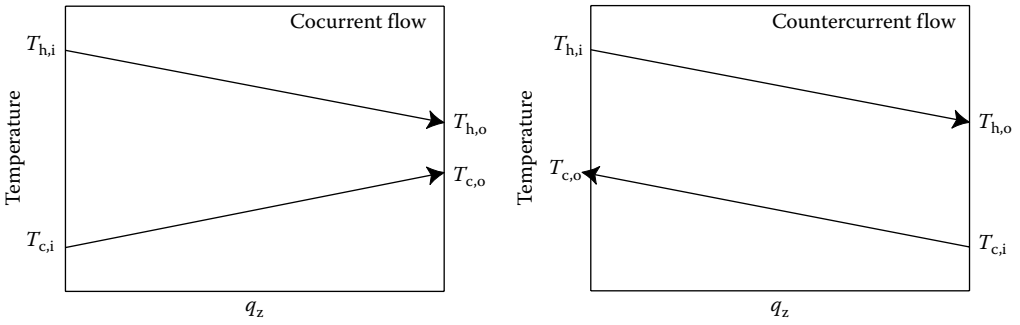


FIGURE 4.1.3 Example of temperature profiles within a cocurrent (left) and countercurrent (right) heat exchanger.

While not elaborated on earlier, it is important to appreciate the assumptions that have gone into deriving the equations so that their suitability can be determined for a particular application. One basic assumption is that heat is not transferred to or from the surroundings, but instead, all of the heat transferred from the hot stream is used to heat the cold stream (Equation 4.1.4). Another assumption related to Equation 4.1.4 is that the specific heat capacities of the streams are relatively constant and can be approximated by the properties of the stream at the average temperature. The assumption of constant specific heat capacity is important for Equation 4.1.4 and is also a major assumption in the derivation of the LMTD equations (Equations 4.1.5 and 4.1.6), which was not shown. The reason the constant specific heat assumption affects the LMTD derivation is because of the additional assumption that the variation in temperature (T) and position (z) within the heat exchanger (or T vs. q_z) is linearly related (Figure 4.1.3). This relationship is affected by the specific heat capacity and can also be affected by assumptions of constant temperature during boiling or condensation. Boiling and condensation, however, do not always occur isothermally due to pressure drop down the length of the heat exchanger or with phase change of multicomponent mixtures. Consequently, significant variations in the specific heat capacity or problems with assumptions regarding the linear relationship between temperature and position (or q_z) will introduce error in the calculation of the MTD.

In addition to constant specific heat capacity, it was also assumed that the overall heat transfer coefficient could be treated as constant if its average value was used in solving Equation 4.1.2. This assumption, however, is often not reasonable, especially when a phase change occurs. In cases where the variation in the overall heat transfer coefficient is a concern, the integral form of the area calculation (Equation 4.1.2) can be used to calculate the total area by dividing the heat exchanger into smaller sections. The iterative approach to finding a solution involves calculating the overall heat transfer coefficient for each section using the properties of the streams at their average temperature for the section. Modern computational tools have greatly simplified this approach.

Considering the limiting assumptions and the approximations often used with the LMTD method, one may wonder how it has been used to successfully design heat exchangers that function quite well in industrial service. The answer is not particularly reassuring. In many cases, the novice process engineer has been fortunate in that overly conservative estimates have offset poor estimates made through inexperience. While this is not meant to discourage the typical process engineer from generating an approximate design, it should serve as a caution to know when one needs to involve a technical expert because of a particular system that pushes the limits of the assumptions that are a part of the MTD approach.

Example of MTD Method Applied to a Double Pipe Heat Exchanger

The MTD method can be better understood by seeing it applied to an example. A countercurrent, double-pipe heat exchanger provides one of the simplest configurations to analyze. Consider the following example.

Problem statement. Determine the length of a double-pipe heat exchanger used to cool 11,905 lb_m/h (1.5 kg/s) kerosene from 140°F to 95°F (60°C–35°C) using cooling water that has a supply temperature of 86°F (30°C) and a maximum return temperature of 122°F (50°C). Cooling water passes through the internal pipe, which is 1.5 in. Schedule 40 carbon steel (ID = 1.61 in. or 40.9 mm; OD = 1.90 in. or 48.3 mm), and kerosene flows countercurrent through the annular region created by a larger pipe that is 2.5 in. Schedule 40 carbon steel (ID = 2.47 in. or 62.2 mm) (Figure 4.1.4). Assume that the thermal conductivity of carbon steel is 31.9 Btu/h ft °F (55 W/m K).

Stream properties

	Kerosene (118°F)	Water (104°F)
Density (lb_m/ft^3)	44.4	61.9
Viscosity (cP)	0.574	0.678
Specific heat (Btu/lb_m °F)	0.548	0.998
Thermal conductivity (Btu/h ft °F)	0.071	0.362

Solution. The general approach is to use Equation 4.1.3 to calculate the required area, from which one can easily determine the length. Before Equation 4.1.3 can be used, however, the heat exchanger duty (q_T), MTD, and overall convective heat transfer coefficient (U) must be determined.

Heat exchanger duty (q_T). The required heat transfer can be determined from Equation 4.1.4, where kerosene has been designated as the hot stream,

$$q_T = (11{,}905\ \text{lb}_m/\text{h})(0.548\ \text{Btu}/\text{lb}_m\ °\text{F})(140°\text{F} - 95°\text{F}) = 293{,}577\ \text{Btu/h} = 86{,}903\ \text{W}$$

Since q_T must be equal for both the kerosene and water streams, Equation 4.1.4 can be rearranged to solve for the mass flow rate of water,

$$\dot{m}_{\text{water}} = \frac{(293{,}577\ \text{Btu}/\text{h})}{(0.998\ \text{Btu}/\text{lb}_m\ °\text{F})(122°\text{F} - 86°\text{F})} = 8{,}171\ \text{lb}_m/\text{h} = 1.03\ \text{kg}/\text{s}$$

Mean temperature difference. The expression to determine the MTD for countercurrent flow was given earlier as Equation 4.1.5. Substituting the inlet and outlet temperatures for the kerosene (hot) and water (cold) streams results in the following MTD,

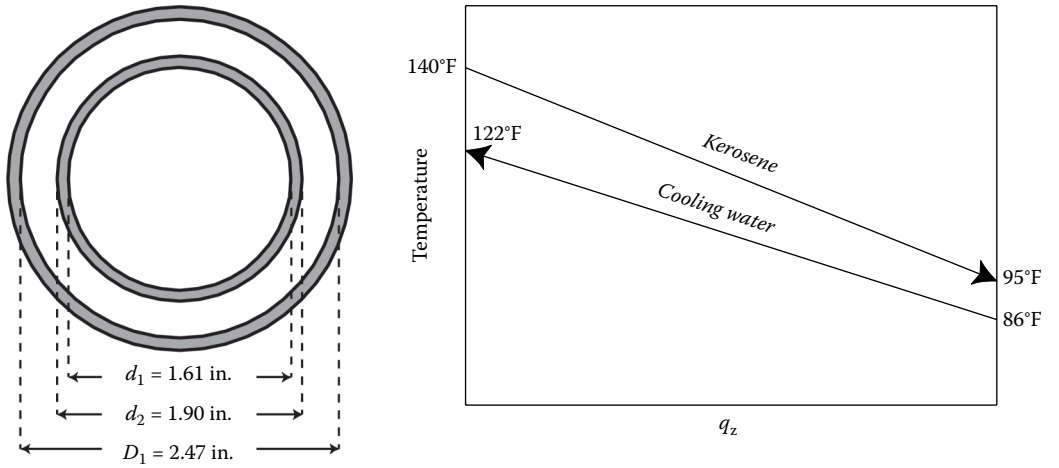


FIGURE 4.1.4 Cross-sectional profile of the double-pipe heat exchanger showing tube diameters (left) and temperature profiles within the heat exchanger.

$$\text{MTD} = \text{LMTD}_{\text{countercurrent}} = \frac{(140°\text{F} - 122°\text{F}) - (95°\text{F} - 86°\text{F})}{\ln\left(\frac{(140°\text{F} - 122°\text{F})}{(95°\text{F} - 86°\text{F})}\right)} = 13°\text{F} = 7.2°\text{C}$$

Notice that no correction factor is needed since the double-pipe configuration does not involve a combination of countercurrent, cocurrent, or crossflow.

Overall convective heat transfer coefficient (U_o). As discussed earlier, the basis of A^* and U^* is optional but must be stated. In this example, the basis will be the outside area ($A^* = A_o$) of the inner pipe. The expression for the overall convective heat transfer coefficient, Equation 4.1.9, can be simplified and written in terms of the provided pipe diameters,

$$U_o = \frac{1}{\frac{d_2}{h_i d_1} + R_{fi}\frac{d_2}{d_1} + \frac{d_2 \ln(d_2/d_1)}{2k_w} + R_{fo} + \frac{1}{h_o}} \tag{4.1.10}$$

Typical fouling factors for water (R_{fi}) and kerosene (R_{fo}) are provided in Table 4.1.1:

Inner fouling R_{fi}	0.0006 h ft² °F/Btu	1.0×10^{-4} m² K/W
Outer fouling R_{fo}	0.0011 h ft² °F/Btu	2.0×10^{-4} m² K/W

The remaining challenge is to calculate the film heat transfer coefficients on the interior (h_i) and exterior (h_o) of the inner pipe. For this example, the Dittus–Boelter correlation will be used,

$$\text{Nu}_D \equiv \frac{hD_h}{k} = 0.023\text{Re}_D^{0.8}\text{Pr}^n \tag{4.1.11}$$

The dimensionless Nusselt number (Nu_D) is defined by the convective heat transfer coefficient, the mean hydraulic diameter (D_h), and the thermal conductivity (k) of the fluid. The Dittus–Boelter correlation (Equation 4.1.11) can be used to calculate the Nusselt number if the Reynolds (Re_D) and Prandtl (Pr) numbers are known. The value of n depends on whether a fluid is being heated ($n = 0.4$) or cooled ($n = 0.3$). The subscript D on the Nusselt and Reynolds numbers indicates that the dimensionless numbers are based on a particular hydraulic diameter, which in this case may be either the hydraulic diameter of the inner pipe ($D_h = d_1$) or the mean hydraulic diameter of the annular region ($D_h = D_1 - d_2$). It is important to note that the Dittus–Boelter correlation is only valid for $\text{Re} \geq 10{,}000$ and $0.6 \leq \text{Pr} \leq 160$. The correlation is also only valid for smooth pipe surfaces, which is not exactly the case in this example. Finally, error is introduced if there are large differences between the pipe wall and the mean temperature of the fluid, which was used to evaluate the fluid properties. Multiple iterations can be used to solve for the wall temperature or other correlations may be useful in minimizing these errors.

For the tube-side fluid,

$$v = \frac{\dot{m}_{\text{water}}}{\rho\left(\frac{\pi d_1^2}{4}\right)} = \frac{(8{,}171\ \text{lb}_m/\text{h})(\text{h}/3600\ \text{s})}{(61.9\ \text{lb}_m/\text{ft}^3)\left(\frac{\pi}{4}\right)(1.61\ \text{in.})^2(\text{ft}/12\ \text{in.})^2} = 2.59\ \text{ft/s} = 0.75\ \text{m/s}$$

$$\text{Re}_{d_1} = \frac{\rho v d_1}{\mu} = \frac{(61.9\ \text{lb}_m/\text{ft}^3)(2.59\ \text{ft/s})(1.61\ \text{in.})(\text{ft}/12\ \text{in.})}{(0.678\ \text{cP})\left(\frac{6.72 \times 10^{-4}\ \text{lb}_m/\text{ft s}}{\text{cP}}\right)} = 47{,}210$$

$$\mathrm{Pr} = \frac{c_p \mu}{k} = \frac{(0.998\ \mathrm{Btu/lb_m\,°F})(0.678\ \mathrm{cP})\left(\frac{6.72 \times 10^{-4}\ \mathrm{lb_m/ft\,s}}{\mathrm{cP}}\right)}{(0.362\ \mathrm{Btu/h\,ft\,°F})(\mathrm{h}/3600\ \mathrm{s})} = 4.52$$

Note that the Reynolds and Prandtl numbers are within the range required for the Dittus–Boelter correlation. Also the tube-side fluid, cooling water, is being heated so $n=0.4$. Using the Dittus–Boelter correlation, the Nusselt number is calculated to be

$$\mathrm{Nu}_{d_1} = 0.023(47{,}210)^{0.8}(4.52)^{0.4} = 230.7$$

The Nusselt number can now be rearranged to solve for the film heat transfer coefficient within the tube,

$$h_i = \mathrm{Nu}_{d_1}\left(\frac{k}{d_1}\right) = (230.7)\left(\frac{(0.362\ \mathrm{Btu/h\,ft\,°F})}{(1.61\ \mathrm{in.})(\mathrm{ft}/12\ \mathrm{in.})}\right) = 622\ \mathrm{Btu/h\,ft^2\,°F} = 3{,}534\ \mathrm{W/m^2\,°C}$$

A similar approach is used for the annular fluid, except the velocity must be calculated using the cross-sectional area of the annular region,

$$v = \frac{\dot{m}_{\mathrm{kerosene}}}{\rho\left(\frac{\pi D_1^2}{4} - \frac{\pi d_2^2}{4}\right)} = \frac{(11{,}905\ \mathrm{lb_m/h})(\mathrm{h}/3600\ \mathrm{s})}{(44.4\ \mathrm{lb_m/ft^3})\left[(\pi/4)(2.47\ \mathrm{in.})^2 - (\pi/4)(1.90\ \mathrm{in.})^2\right](\mathrm{ft}/12\ \mathrm{in.})^2}$$

$$= 5.48\ \mathrm{ft/s} = 1.67\ \mathrm{m/s}$$

Additionally, the mean hydraulic diameter of the annular region is now the difference between the inner diameter of the outer pipe and the outer diameter of the inner pipe,

$$D_h = D_{\mathrm{annular}} = D_1 - d_2 = 2.47\ \mathrm{in.} - 1.90\ \mathrm{in.} = 0.57\ \mathrm{in.}$$

The Reynolds and Prandtl numbers are

$$\mathrm{Re}_{D_{\mathrm{annular}}} = \frac{\rho v D_{\mathrm{annular}}}{\mu} = \frac{(44.4\ \mathrm{lb_m/ft^3})(5.48\ \mathrm{ft/s})(0.57\ \mathrm{in.})(\mathrm{ft}/12\ \mathrm{in.})}{(0.574\ \mathrm{cP})\left(\frac{6.72 \times 10^{-4}\ \mathrm{lb_m/ft\,s}}{\mathrm{cP}}\right)} = 29{,}962$$

and

$$\mathrm{Pr} = \frac{c_p \mu}{k} = \frac{(0.548\ \mathrm{Btu/lb_m\,°F})(0.574\ \mathrm{cP})\left(\frac{6.72 \times 10^{-4}\ \mathrm{lb_m/ft\,s}}{\mathrm{cP}}\right)}{(0.071\ \mathrm{Btu/h\,ft\,°F})(\mathrm{h}/3600\ \mathrm{s})} = 10.7$$

Again, it is important to check that the Reynolds and Prandtl numbers fall within the required range of the Dittus–Boelter correlation, which they do. With this condition met, the Dittus–Boelter correlation can now be used to calculate the Nusselt number for the annular fluid. Since the annular fluid, kerosene, is being cooled, $n=0.3$ and

$$\mathrm{Nu}_{D_{\mathrm{annular}}} = 0.023(29{,}962)^{0.8}(10.7)^{0.3} = 178.6$$

Rearranging the defined expression for the Nusselt number to solve for the film heat transfer coefficient,

$$h_o = \mathrm{Nu}_{D_{\mathrm{annular}}}\left(\frac{k}{D_{\mathrm{annular}}}\right) = (178.6)\left(\frac{(0.071\ \mathrm{Btu/h\,ft\,°F})}{(0.57\ \mathrm{in.})(\mathrm{ft}/12\ \mathrm{in.})}\right)$$

$$h_o = 267 \text{ Btu / h ft}^2 \text{ °F} = 1{,}516 \text{ W/m}^2 \text{ °C}$$

The overall convective heat transfer coefficient can now be calculated,

$$U_o = \left[\left(\frac{1}{622 \text{ Btu / h ft}^2 \text{ °F}} \right) \left(\frac{1.90 \text{ in.}}{1.61 \text{ in.}} \right) + \left(0.0006 \text{ h ft}^2 \text{ °F/ Btu} \right) \left(\frac{1.90 \text{ in.}}{1.61 \text{ in.}} \right) + \left(\frac{(1.90 \text{ in.})(\text{ft} / 12 \text{ in})}{2 \cdot (31.9 \text{ Btu / h ft °F})} \right) \right.$$

$$\left. \ln \left(\frac{1.90 \text{ in.}}{1.61 \text{ in.}} \right) + \left(0.0011 \text{ h ft}^2 \text{ °F / Btu} \right) + \left(\frac{1}{267 \text{ Btu / h ft}^2 \text{ °F}} \right) \right]^{-1}$$

$$U_o = 127 \text{ Btu / h ft}^2 \text{ °F} = 722 \text{ W / m}^2 \text{ °C}$$

Heat exchanger area and length. Equation 4.1.3 can now be used to calculate the required area and the length of the heat exchanger,

$$A_o = \frac{q_T}{U_o \cdot \text{MTD}} = \frac{(293{,}577 \text{ Btu / h})}{(127 \text{ Btu / h ft}^2 \text{ °F})(13 \text{ °F})} = 178 \text{ ft}^2 = 16.5 \text{ m}^2$$

$$L = \frac{A_o}{\pi \cdot d_2} = \frac{(178 \text{ ft}^2)}{\pi (1.90 \text{ in.})(\text{ft} / 12 \text{ in.})} = 358 \text{ ft} = 109 \text{ m}$$

A 358 ft double-pipe heat exchanger would be needed to cool the kerosene stream from 140°F to 95°F using cooling water. The calculated length, however, is much too long for a single double-pipe heat exchanger and would require multiple units connected in series. This example, however, demonstrates how the basic design equations can be used for the custom sizing of a double-pipe heat exchanger.

Shell-and-Tube Heat Exchangers

A shell-and-tube heat exchanger is essentially a bundle of tubes enclosed in a shell and so arranged that one fluid flows through the tubes and another fluid flows across the outside of the tubes, heat being transferred from one fluid to the other through the tube wall. A number of other mechanical components are required to guide the fluids into, through, and out of the exchanger, to prevent the fluids from mixing, and to ensure the mechanical integrity of the heat exchanger. A typical shell-and-tube heat exchanger is shown in Figure 4.1.5 (TEMA 2007), but the basic design allows many modifications and special features, some of which are described later.

Nomenclature of Heat Exchanger Components

For the purpose of establishing standard terminology, Figure 4.1.5 illustrates a typical shell-and-tube heat exchanger. Typical parts and connections, for illustrative purposes only, are numbered for identification:

1. Stationary Head—Channel
2. Stationary Head Bonnet
3. Stationary Head Flange—Channel or Bonnet
4. Channel Cover

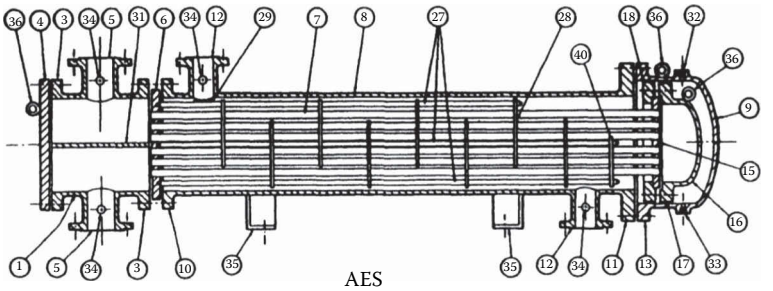


FIGURE 4.1.5 Longitudinal section of a typical shell-and-tube heat exchanger (TEMA AES) with nomenclature. (Modified from TEMA Standards 9th ed., Tubular Exchanger Manufacturers Association, Tarrytown, NY, 2007.)

5. Stationary Head Nozzle
6. Stationary Tubesheet
7. Tubes
8. Shell
9. Shell Cover
10. Shell Flange—Stationary Head End
11. Shell Flange—Rear Head End
12. Shell Nozzle
13. Shell Cover Flange
14. Expansion Joint on Shell (not shown)
15. Floating Tubesheet
16. Floating Head Cover
17. Floating Head Cover Flange
18. Floating Head Backing Device
19. Split Shear Ring
20. Slip-on Backing Flange
21. Floating Head Cover—External
22. Floating Tubesheet Skirt
23. Packing Box
24. Packing
25. Packing Gland
26. Lantern Ring
27. Tierods and Spacers
28. Transverse Baffles or Support Plates
29. Impingement Plates
30. Longitudinal Baffle
31. Pass Partition
32. Vent Connection
33. Drain Connection
34. Instrument Connection
35. Support Saddle
36. Lifting Lug
37. Support Bracket
38. Weir
39. Liquid Level Connection
40. Floating Head Support

Shell-and-tube heat exchangers have been constructed with heat transfer areas from less than 0.1 m^2 (1 ft^2) to over 100,000 m^2 (1,000,000 ft^2), for pressures from deep vacuum to over 1000 bar (15,000 psi), for temperatures from near 0 to over 1400 K (2000°F), and for all fluid services including single-phase heating and cooling and multiphase vaporization and condensation. The key to such flexibility is the wide range of materials of construction, forming and joining methods, and design features that can be built into these exchangers (see Hewitt 2008, Vol. 4; Saunders 1988; and Yokell 1990). Most shell-and-tube heat exchangers are manufactured in conformity with TEMA *Standards* (2007) and the *ASME Boiler and Pressure Vessel Code* (latest edition), but other codes and standards may apply. In the design process, it is important to consider the mechanical integrity under varying operational conditions and the maintainability (especially cleaning) of the exchanger as equally important with the thermal–hydraulic design.

Construction Features

Tubes. Tubes used in shell-and-tube exchangers range from 6.35 mm (1/4 in.) to 50.8 mm (2 in.) and above in outside diameter, with the wall thickness usually being specified by the Birmingham wire gauge (BWG). Tubes are generally available in any desired length up to 30 m (100 ft) or more for plain tubes. While plain tubes are widely used, a variety of internally and/or externally enhanced tubes is available to provide special heat transfer characteristics when economically justified (see subsection on enhancement in Section 4.16). Low-fin tubes having circumferential fins typically 0.8–1.6 mm (0.032–0.062 in.) high, spaced 630–1260 fins/m (16–32 fins/in.) are often employed, especially when the shell-side heat transfer coefficient is substantially smaller than the tube-side coefficient. The outside heat transfer area of a low-fin tube is three to six times the inside area, resulting in a smaller heat exchanger shell for the same service, which may offset the higher cost of the tube per unit length.

The tubes are inserted into slightly oversized holes drilled (or, occasionally, punched) through the tubesheets (items 6 and 15 in Figure 4.1.5). The tubes are secured by several means, depending upon the mechanical severity of the application and the need to avoid leakage between the streams. In some low-severity applications, the tubes are roller-expanded into smooth holes in the tubesheet. For a stronger joint, two shallow circumferential grooves are cut into the wall of the hole in the tubesheet and the tube roller-expanded into the grooves; to eliminate the possibility of leakage, a seal weld can be run between the outer end of the tube and the tubesheet. Alternatively, the tubes may be strength-welded into the tubesheet.

Tube supports. It is essential to provide periodic support along the length of the tubes to prevent sagging and destructive vibration caused by the fluid flowing across the tube bank. A secondary role played by the tube supports is to guide the flow back and forth across the tube bank, increasing the velocity and improving the heat transfer on the shell side (but also increasing the pressure drop). The tube support is usually in the form of single segmental baffles (item 28 in Figure 4.1.5)—circular plates with holes drilled to accommodate the tubes and with a segment sheared off to form a "window" or "turnaround" to allow the shell-side fluid to pass from one crossflow section to the next. The baffles must overlap at least one full row of tubes to give the bundle the necessary rigidity against vibration. When minimizing shell-side pressure drop is not a priority, a baffle cut of 15%–25% of the shell inside diameter is customary. Baffle spacing is determined first by the necessity to avoid vibration and secondarily to approximately match the free crossflow area between adjacent baffles to the flow area in the window; that is, small baffle cuts correspond to closer baffle spacing.

In situations such as low-pressure gas flows on the shell side where pressure drop is severely limited, double segmental and strip baffle arrays can be used. A helical baffle arrangement has been introduced (Kral et al. 1996; Zhang et al. 2010) which causes the shell-side fluid to spiral through the exchanger giving improved heat transfer versus pressure drop characteristics. More recently, tubes with an external longitudinal spiral configuration have been offered. The shell-side fluid flows in a spiral path down the length of the tubes. In addition, the tubes can be bundled in such a way that they are supported by surrounding tubes, thereby eliminating the need for baffles. Where vibration prevention and/or minimum pressure drop are the main concerns, grids of rods or strips can be used

(Gentry et al. 1982). The rod baffle design is also advantageous in shell-side thermosiphon vaporizer applications because it minimizes phase separation, slugging, and vibration.

Shells. The shell is the cylinder which confines the shell-side fluid (item 8 in Figure 4.1.5), fitted with nozzles for fluid inlet and exit. Diameters range from less than 50 mm (2 in.) to 3.05 m (10 ft) commonly, and at least twice that value for special applications. In diameters up to 610 mm (24 in.), shells are usually made from standard pipe or tubular goods by cutting to the desired length; in larger sizes, metal plates are rolled to the desired diameter and welded.

A variety of nozzle arrangements are used for special purposes, and TEMA has a standard code to identify the major types, as well as the various front and rear head configurations on the tube side. Figure 4.1.6 shows these configurations with the corresponding code letters.

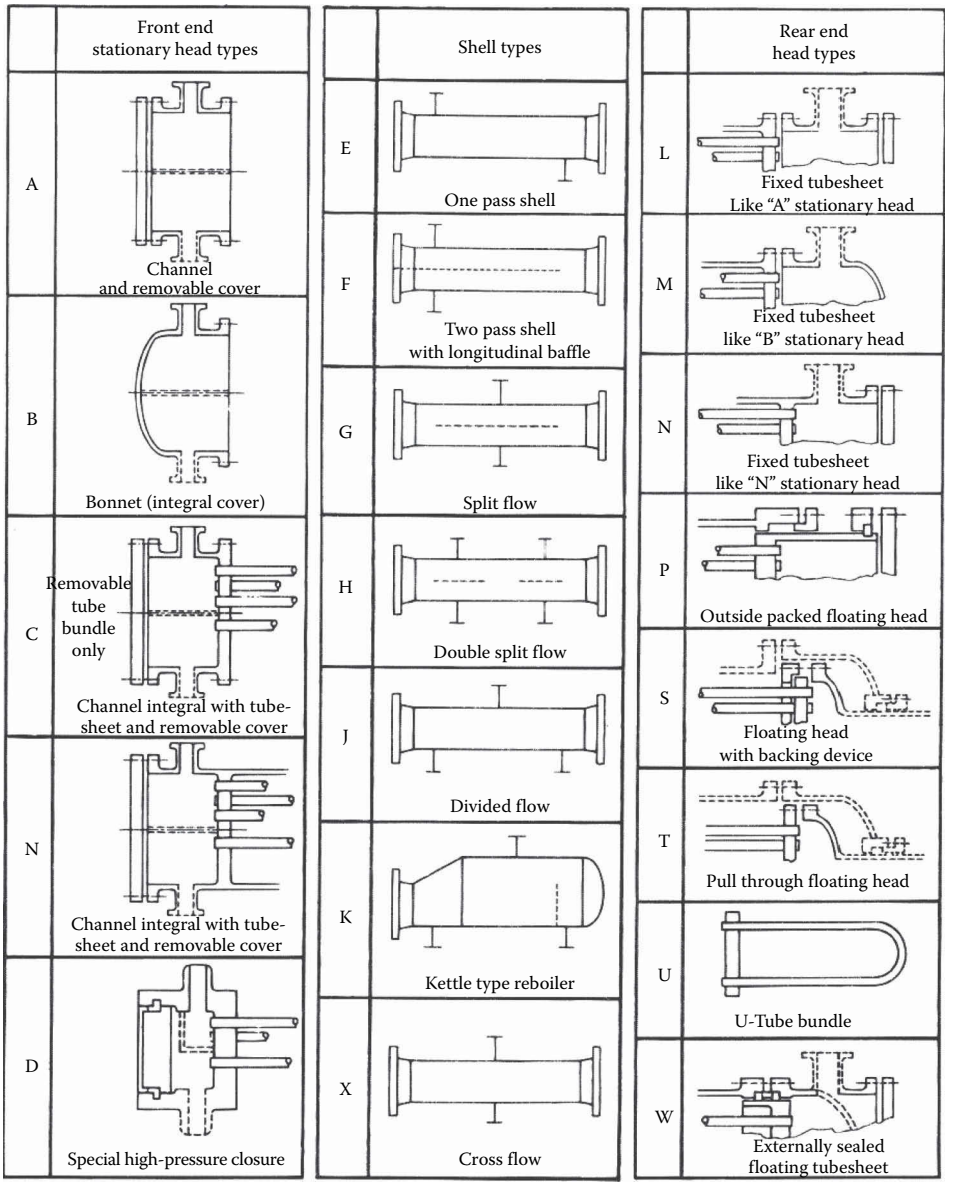


FIGURE 4.1.6 TEMA nomenclature for shell-and-tube configurations. (From TEMA Standards 9th ed., Tubular Exchanger Manufacturers Association, Tarrytown, NY, 2007. With permission.)

The E shell (center column, top) has the nozzles on opposite ends of the shell and is the most common configuration. It is used for any of the thermal services (single-phase heating or cooling, vaporization, and condensation). The nozzles may be on opposite sides of the shell as shown or on the same side; the choice is largely determined by plumbing convenience. The E shell allows countercurrent flow (shown later) of the two streams if there is one tube-side pass (i.e., the tube-side fluid flows through all of the tubes in parallel).

The F shell has both nozzles at one end of the shell and uses a longitudinal baffle on the shell side (shown dashed in the drawing) to force the shell-side fluid to flow to the far end of the heat exchanger and then back to the exit nozzle on the other side of the longitudinal baffle. Ideally, this allows countercurrent flow of the two streams if there are two tube-side passes (i.e., the tube-side fluid flows through half of the tubes in one direction, is turned around in the rear head, and returns through the other half of the tubes—see discussion of head types). However, the longitudinal baffle must be carefully sealed to the shell to prevent leakage of the shell-side fluid across it; this is done by welding the longitudinal baffle to the shell and front tubesheet (which limits some design options) or by using mechanical seals. The F shell is mainly used for sensible heat transfer services.

The G shell has both nozzles at the center of the shell, with a centrally located longitudinal baffle to force the fluid to the ends of the shell before returning. While the G shell is used for all services, its main application is as a shell-side vaporizer with either forced or natural (thermosiphon) convection of the boiling fluid; in the latter service, limited leakage across the baffle generally does not greatly degrade the thermal performance and the longitudinal baffle does not need to be perfectly sealed against the shell.

The H shell is effectively a double G shell and is employed in the same services. It is considered when the calculated shell-side pressure drop for a G arrangement is too high and threatens to limit the circulation rate or cause maldistribution of the flow.

The J shell, with one nozzle on top of the shell and two on the bottom, or vice versa, is commonly used in vacuum-condensing applications because of its low-pressure drop. Two J shells (one inverted) may be mated in series for long-condensing-range mixtures. The nozzles are usually different diameters, with the large diameter accommodating the inlet vapor. The baffles are vertically cut.

The K shell (or kettle reboiler or flooded chiller) is exclusively intended for vaporization of liquid on the shell side, with a condensing vapor (usually steam) or a hot liquid on the tube side as the heating medium. The tubesheet diameter is large enough to accommodate the tube bundle, but the shell transitions to a larger diameter to allow the vapor to disengage from the liquid pool and exit from the top nozzle. A weir or other level control is used to maintain the liquid level, usually just above the top tubes in the bundle.

The X shell is intended to provide a well-distributed crossflow of the shell-side fluid, the fluid usually entering at the top and exiting at the bottom but occasionally used for upflow or horizontal crossflow. To obtain good distribution, multiple nozzles from a properly designed manifold may be required. Alternatively, the upper tubes in the bundle may be omitted to allow internal redistribution, or a large plenum chamber may be welded to the top of the shell ("vapor dome" or "bathtub nozzle"), or a diverging transition section may be placed between the inlet piping and the top of the shell. The tube supports may be complete circles since there is little or no longitudinal shell-side flow. The X shell gives the lowest shell-side pressure drop of any configuration and is often used for low-pressure vapor condensers.

Front head. TEMA recognizes several front head designs as shown in the first column of Figure 4.1.6. Any of these designs will get the tube-side fluid into the tubes, but each has special features which recommend it to meet special needs. In Figure 4.1.6, the dashed lines indicate optional features depending upon need.

The A head, a channel with removable cover, bolts directly to the shell flange as shown in Figure 4.1.5, the tubesheet in that case being held between them and sealed with gaskets. Alternatively, the tubesheet may be integral with the shell (see the L rear head in Figure 4.1.6). A removable channel cover permits inspection, cleaning, removal, and replacement of tubes without disturbing the piping. The dashed lines at the center and the lower nozzle indicate that a pass partition plate may be welded

in the channel (and gasketed against the tubesheet and channel cover) to provide for two tube-side passes (as shown in Figure 4.1.5 and required by the F shell design). Additional pass partitions may be provided to allow four, six, or any even number of tube-side passes. This permits the designer to utilize the available tube-side pressure drop to increase velocity, improve the heat transfer coefficient, and possibly reduce fouling. A second nozzle is required on the channel for multipass designs.

The B, or bonnet, front head reduces the number of gasketed joints and thus the opportunity for leakage, but it does not permit inspection of the tubes without breaking the piping connection. It is generally less expensive than the A head.

C and N heads retain the removable cover feature of the A head but, respectively, replace the channel-to-tubesheet and the tubesheet-to-shell gasketed connections with welds to minimize leakage possibilities. The N head is particularly used in nuclear services.

The D head is mainly used in feedwater heater applications where tube-side pressures are in the 100–400 bar range. The internal partition (optional) need only withstand the 1–2 bar pressure drop through the tubes so it can be of lightweight construction. The high-pressure closure against the atmosphere uses a shear key ring to lock the main closure in place.

Rear head. A variety of rear head designs are used in shell-and-tube exchangers, primarily because of the need to accommodate thermally induced stresses. During operation, the tubes and the shell have different temperatures and therefore will expand (or try to) different amounts, even if there were no residual stresses in the exchanger before start-up and even if the entire exchanger is made out of the same material. The thermal stress problem is exacerbated if there are residual stresses, or if the exchanger is made of different materials, or during transient operation (including start-up and shutdown). If the temperature differences are small, the structure may be able to accommodate the thermal stresses safely; usually, however, it is necessary to make specific provision to allow the shell and the tubes to expand or contract independently. Failure to do so can result in buckling, bending, or even rupture of the shell or the tubes, or destruction of the tube-to-tubesheet joint.

A simple solution is to incorporate an expansion joint or a bellows into the shell (or in certain special applications, into the tube-side piping internal to the shell cover). However, this solution cannot cover the entire range of pressures and temperature differences encountered in practice. Further, it is usually possible to incorporate other desirable features, such as removable bundles, with thermal stress relief in the variety of rear head designs available. These are shown in the last column of Figure 4.1.6.

The L and M rear heads correspond to the A and B front heads previously described. As shown, they require a fixed tubesheet design; that is, the tubesheets are rigidly fastened to the shell, and thermal stress relief, if necessary, must be provided by a shell-side expansion joint or bellows. The tube bundle cannot be removed for inspection or mechanical cleaning on the shell side. However, the outer tube limit—the diameter of the tube field circumscribing the outermost tubes in the bundle—can be as little as 0.4 in. (10 mm) less than the inside diameter of a pipe shell and 0.5 in. (12.7 mm) for a rolled shell. Therefore, the tube field can be very full, giving more tubes and minimizing bypass flow. Similar comments apply to the N rear head, except that more clearance must be left between the outermost tubes and the shell.

The type P head uses packing between the skirt on the rear tubesheet and the shell extension to seal the shell-side fluid against leakage. The compression on the packing has to be adjusted to prevent excessive leakage on the one hand and to allow limited movement of the tube-side head on the other, so the shell-side fluid must be benign and cheap (not surprisingly, it is often cooling water). On the other hand, leakage between the two fluids can occur only through tube hole leaks. Because of the tubesheet skirt, clearance between the outermost tubes and the shell must increase compared with types L or M; accordingly, fewer tubes are possible in a given shell, and sealing strips to partially block the bundle-to-shell bypass stream are recommended. When the floating head cover and packing gland are removed, the tube bundle can be pulled out of the shell for inspection and cleaning.

The TEMA S split-ring floating head design uses a split backing ring to hold the floating head cover and its gasket to the tubesheet. The split backing ring is bolted to the cover with a bolt circle

outside the diameter of the tubesheet. Therefore, when the split ring is removed, the entire tube bundle may be pulled out of the shell. Tube count is similar to type P design and sealing strips are recommended. Usually, the split-ring floating head is used with an even number of tube passes so that a plain bonnet-type shell cover can be used. However, as shown by the dashed lines in Figure 4.1.6, single tube-side pass design (and countercurrent flow) can be achieved by use of a packing gland on the exit piping through the bonnet; alternatively, a deep bonnet can be used together with an expansion joint or bellows on the tube-side exit piping.

The pull-through floating head, type T, uses a floating head cover that flanges directly to the tubesheet, reducing the possibility of internal leakage compared with type S, but also eliminating more tubes around the periphery. Sealing strips are a virtual necessity. Single tube-side pass design is similar to type S, but is rarely used.

TEMA type U uses a bundle of U tubes and hence requires no rear head at all. The U-tube bundle effectively eliminates the thermal stress problem between shell and tubes, because each tube is free to expand or contract independently. The U bundle is also the cheapest construction because the cost of a second tubesheet is avoided. However, there are a number of drawbacks: Designs must have an even number of tube-side passes, mechanical cleaning of the smaller bend radius tubes in the U bend is impossible, individual tubes cannot be replaced except in the outer row, some tube count is lost because of minimum bend limits, and the U bend must be carefully supported against vibration or kept out of the crossflow stream by placing the shell nozzle upstream of the bend. The tube side in the U bend is susceptible to erosion, especially with two-phase or particulate-containing fluids.

Type W uses two sets of packing, often with a lantern ring in between. This construction is generally limited to benign fluids and low to very moderate pressures and temperatures.

Other features. Numerous other components are necessary or optional to construction of shell-and-tube exchangers. Probably the most complete discussion is given by Yokell (1990).

Principles of Design

Design logic. The design of a shell-and-tube exchanger involves the following steps:

1. Selection of a set of design features that are required for mechanical integrity and ease of maintenance, and which will likely lead to satisfying the thermal requirements within the allowable pressure drops, and at lowest cost.
2. Selection of a set of dimensions for the actual exchanger.
3. For the dimensions selected in (2), calculation of the thermal performance of the heat exchanger and both tube-side and shell-side pressure drops, using available rating procedures.
4. Comparison of the thermal performance calculated in (3) with that required and examination of the pressure drops calculated in (3) to ensure that the allowed pressure drops are reasonably used but not exceeded.
5. Adjustment of the dimensions selected in (2) and repetition of steps (3) and (4) until the criteria are satisfied.
6. Completion of the mechanical design to satisfy code requirements.
7. Cost estimation.

Basic design equations. Some of the basic design equations for a shell-and-tube exchanger in steady-state service were introduced earlier in this chapter in the subsection title, "Principles of Heat Exchanger Design." The approach to designing a shell-and-tube exchanger follows the basic approach outlined in this earlier subsection. Equation 4.1.2 relates the differential heat transfer area to the differential heat that is transferred across the area, the temperature difference $(T_h - T_c)$ across the differential area, and the overall convective heat transfer coefficient (U^*). Since T_h and T_c generally vary with the amount of heat transferred (following the first law of thermodynamics, and excepting isobaric phase transition of a pure component) and U^* may vary with local heat transfer conditions, in principle Equation 4.1.2

must be numerically integrated with T_h, T_c, and U^* calculated along the path of integration, and this process is performed by the most-advanced computer-based design methods.

For many applications, certain reasonable assumptions can be made allowing the analytical integration of Equation 4.1.2 to produce Equation 4.1.3. Depending on the flow configuration, the MTD may be calculated using either Equation 4.1.5 or Equation 4.1.6. Most shell-and-tube exchangers, however, have nozzle and tube pass configurations that lead to mixed countercurrent and cocurrent flow regions (as well as crossflow in the X shell). For these cases, the MTD must be calculated using Equation 4.1.7.

The "configuration correction factor" for the flow configuration involved has been found as a function of dimensionless temperature ratios for most flow configurations of interest and is given in analytical and/or graphical form in many heat transfer references (e.g., Hewitt 2008, Vol. 1). As an example, Figure 4.1.7 shows F, the configuration correction factor, for a shell-and-tube heat exchanger with a single shell pass and an even number of tube passes (i.e., 2, 4, 6, etc.). The correction factor is a function of dimensionless temperatures, P and R, where

$$P = \frac{T_{\text{tube, out}} - T_{\text{tube, in}}}{T_{\text{shell, in}} - T_{\text{tube, in}}} \tag{4.1.12}$$

and

$$R = \frac{T_{\text{shell, in}} - T_{\text{shell,out}}}{T_{\text{tube,out}} - T_{\text{tube, in}}} \tag{4.1.13}$$

Overall heat transfer coefficient. The relationship between the overall heat transfer coefficient U^*, referenced to the heat transfer area A^*, and the individual (film) heat transfer coefficients and the fouling resistances for heat transfer across a single cylinder was presented in Equation 4.1.9. This expression can be adjusted to account for heat transfer across a bundle of tubes,

$$U^* = \frac{1}{\dfrac{A^*}{h_i A_i} + R_{fi}\dfrac{A^*}{A_i} + \dfrac{A^* \ln(d_o/d_i)}{2\pi N_t L K_w} + R_{fo}\dfrac{A^*}{A_o} + \dfrac{A^*}{h_o A_o}} \tag{4.1.14}$$

where N_t is the total number of tubes in the heat exchanger. For the special but important case of plain tubes,

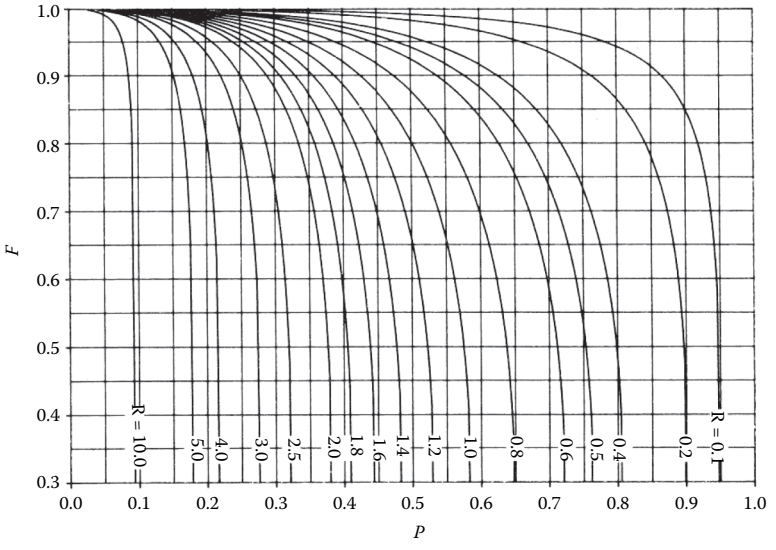


FIGURE 4.1.7 Configuration correction factor for a shell-and-tube heat exchanger with one shell pass and an even number of tube passes. (From Hewitt, G.F., *Heat Exchanger Design Handbook*, Begell House, New York, 2008. With permission.)

$$A^* = A_o = N_t(\pi d_o L) \tag{4.1.15}$$

and Equation 4.1.14 reduces to

$$U_o = \frac{1}{\frac{d_o}{h_i d_i} + R_{fi}\frac{d_o}{d_i} + \frac{d_o \ln(d_o/d_i)}{2K_w} + R_{fo} + \frac{1}{h_o}} \tag{4.1.16}$$

If finned tubes are used, the root diameter d_r of the fins replaces d_o in Equation 4.1.16 and A_o includes the surface area of the fins as well as the bare tube surface between the fins; it is also necessary to include a fin efficiency (typically about 0.8–0.95) multiplier in the numerators of the last two terms on the right side of Equation 4.1.16 to account for resistance to conduction in the fins. The treatment of fin efficiency is fully developed in Kern and Kraus (1972).

Film heat transfer coefficients. Calculation of single-phase tube-side heat transfer coefficients for plain tubes is discussed in Section 3.2.3; special correlations are required for internally enhanced tubes, see discussion of enhancement in Section 4.16. Intube condensation and vaporization are covered in the subsection on boiling and condensation in Section 3.4.1.

Shell-side heat transfer calculations are more complex owing to the large number and range of design variables and process conditions that can occur. The most accurate methods are proprietary and computer based. The best known of these methods are those of Heat Transfer Research, Inc. (HTRI), Navasota, TX, although there are other proprietary software packages available to and used by manufacturers. For single-phase flow, the Delaware method appears to be the best in the open literature, and it is feasible for both hand and computer use; various presentations of the method appear in many references, including Hewitt (2008) Vol. 3 and Hewitt et al. (1994). These references also give methods for shell-side vaporizing and condensing design. An approximate design procedure is given in the next subsection.

Fouling. Fouling is the formation of any undesired deposit on the heat transfer surface, and it presents an additional resistance to the flow of heat. Several different types of fouling are recognized:

Sedimentation: deposition of suspended material on the surface.
Crystallization: precipitation of solute from supersaturated solutions.
Corrosion: formation of corrosion products of the heat exchanger components.
Thermal degradation/polymerization: formation of insoluble products by oxidation, charring, and/or polymerization of a process stream.
Biofouling: growth of large organisms (e.g., barnacles) that interfere with flow to or past a heat transfer surface ("macrobiofouling") or small organisms (e.g., algae) that form a fouling layer on the surface ("microbiofouling").

The effect of fouling on design is twofold: Extra surface must be added to the heat exchanger to overcome the additional thermal resistance, and provision must be made to allow cleaning either by chemical or mechanical means. The fouling resistances included in Equation 4.1.16 result in requiring extra surface by reducing U^* (though they do not properly account for the time-dependent nature of fouling) and should be chosen with care. Prior experience with a given service is the best source of values. Ranges of typical values for major classes of service are included in Table 4.1.1.

Other things being equal, a fouling stream that requires mechanical cleaning should be put in the tubes because it is easier to clean the tube side. If this is not possible or desirable, then a removable bundle with a rotated square tube layout should be chosen to facilitate cleaning by mechanical, hydraulic (typically high-velocity water jets), or chemical means.

Pressure drop. Tube-side pressure drop in plain tubes is discussed in Section 2.4. These calculations are straightforward and quite accurate as long as the tubes are smooth and clean; however, even a small amount of roughening due to corrosion or fouling (sometimes with a significant reduction of flow area) can double or triple tube-side pressure drop. Special correlations are required for internally enhanced tubes.

Calculation of shell-side pressure drop is implicit in the design methods mentioned previously for heat transfer. Roughness has less effect on shell-side pressure drop than on the tube side, but fouling still may have a very substantial effect if the deposits fill up the clearances between the baffles and the shell and between the tubes and the baffles, or if the deposits are thick enough to narrow the clearances between adjacent tubes. Existing design methods can predict these effects if the thickness of the fouling layer can be estimated.

Limitations of design. It should be recognized that even under the best of conditions—new, clean exchangers with conventional construction features—heat exchanger design is not highly accurate. The best methods, when compared with carefully taken test data, show deviations of ±20% on overall heat transfer and ±40% on shell-side pressure drop (e.g., Palen and Taborek 1969). These ranges are considerably worsened in fouling services. In these cases, the thermal *system* should be designed for operational flexibility, including carefully chosen redundancy of key components, and easy maintenance.

Approximate Design Method

Because of the complexity of rigorous design methods, it is useful to have an estimation procedure that can quickly give approximate dimensions of a heat exchanger for a specified service. Such a method is given here for purposes of preliminary cost estimation, plant layout, or checking the results of computer output. This method is based upon Equation 4.1.3 with $A^* = A_o$ and $U^* = U_o$ and depends upon rapidly estimating values for q_T, MTD, and U_o. The procedure is as follows:

Estimation of q_T. For sensible heat transfer, q_T can be found using Equation 4.1.4. For isothermal phase change,

$$q_T = \dot{m} h_{fg} \tag{4.1.17}$$

where $\dot{m}$ is the mass rate of condensation or vaporization and h_{fg} is the latent heat of phase transformation.

For more complex cases, such as partial or multicomponent condensation, more elaborate analyses are required, although this method can still be used with care to give rough estimates.

Estimation of MTD. The first step is to calculate or estimate $\text{LMTD}_{\text{countercurrent}}$ from Equation 4.1.5 and then estimate F as follows:

1. If the two streams are in countercurrent flow, $F = 1$.
2. If the two streams are in a combination of countercurrent and cocurrent flows (i.e., multiple tube passes) and the outlet temperatures of the two streams are equal, $F = 0.8$.
3. If the exchanger has multiple passes and $T_{h,o} > T_{c,o}$, then $0.8 < F < 1.0$, with the actual value depending upon the temperature ranges of the two streams and $(T_{h,o} - T_{c,o})$. It is usually sufficiently accurate to take $F = 0.9$.
4. Design of a multiple tube pass exchanger with $T_{h,o} < T_{c,o}$ (i.e., a temperature cross) leads to $F < 0.8$, which is inefficient, of uncertain inaccuracy, and perhaps even thermodynamically impossible. The problem can be handled with multiple shells in series.
5. Then, $\text{MTD} = F(\text{LMTD})_{\text{countercurrent}}$, (Equation 4.1.7).

Estimation of U_o. The best way to estimate U_o is to use Equation 4.1.16, together with values of h_o, h_i, $R_{f,o}$, and $R_{f,i}$, chosen from Table 4.1.1. This table includes ranges of values that are typical of the fluids and services indicated, assuming normally allowable pressure drops, exchanger construction, and fouling. However, care should be taken in selecting values to consider possible unusual conditions, for example, especially high or low velocities (implying correspondingly high or low allowable pressure drops) and especially fouling. In selecting values from the table, the user should carefully read the footnotes for each entry.

Calculation of A_o. The total outside tube heat transfer area required in the heat exchanger is now found from Equation 4.1.3.

Estimation of exchanger dimensions. Figure 4.1.8 shows the relationship among A_o, effective tube length L, and inside shell diameter for a fully tubed, fixed tubesheet heat exchanger with one tube-side pass, with 3/4 in. (19.05 mm) plain tubes on a 15/16 in. (23.8 mm) pitch equilateral triangular tube layout. These curves are constructed using tube count tables (e.g., Saunders 1988). The dashed lines marked 3:1, 6:1, 8:1, 10:1, and 15:1 indicate ratios of tube length to shell inside diameter for guidance in selection. Exchangers of less than 3:1 ratio are expensive because of the large-diameter shell and tubesheet, with more holes to be drilled and tubes rolled and/or welded, and shell-side flow distribution is likely to be poor and lead to excessive fouling. Exchangers greater than 15:1 ratio are probably beyond the point of saving money by reducing shell diameter and number of tubes and may require excessive clear way for pulling the bundle; the bundles may be springy and difficult to handle during maintenance. Most heat exchangers fall into the 5:1–10:1 range.

Figure 4.1.8 is a very specific case that is used as a reference. In order to extend its usefulness to other tube diameters, layouts, bundle constructions, etc., Equation 4.1.18 is used:

$$A_o' = A_o F_1 F_2 F_3 \tag{4.1.18}$$

where A_o' is the value to be used with Figure 4.1.8, A_o is the required area calculated from Equation 4.1.3, and

F_1 is the correction factor for the tube layout. $F_1 = 1.00$ for 3/4 in. (19.05 mm) outside diameter tubes on a 15/16 in. (23.8 mm) triangular pitch. Values of F_1 for other tube diameters and pitches are given in Table 4.1.2.

F_2 is the correction factor for the number of tube-side passes. $F_2 = 1.00$ for one tube-side pass, and Table 4.1.3 gives values of F_2 for more passes.

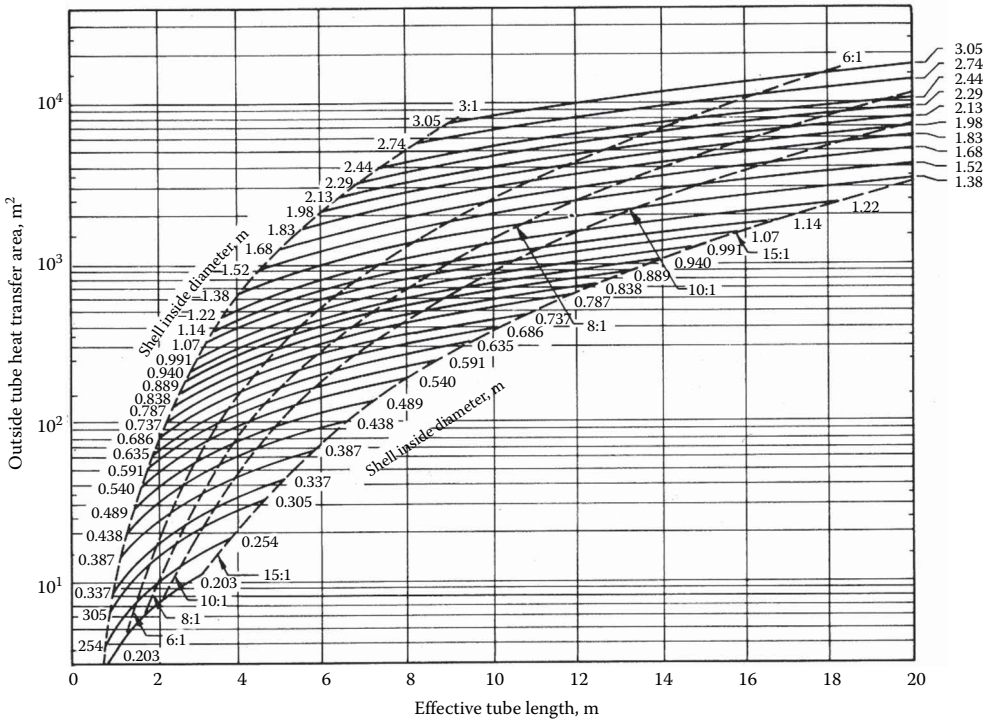
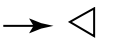


FIGURE 4.1.8 Heat transfer area as a function of shell inside diameter and effective tube length for 19.05 mm (3/4 in.) tubes on a 23.8 mm (15/16 in.) equilateral triangular tube layout, fixed tubesheet, one tube-side pass, fully tubed shell. (From Hewitt, G.F., *Heat Exchanger Design Handbook*, Begell House, New York, 1998. With permission.)

F_3 is the correction factor for shell construction/tube bundle configuration. $F_3 = 1.00$ for fixed tubesheet, fully tubed shells, and Table 4.1.4 gives values of F_3 for the standard TEMA types.

TABLE 4.1.2
Values of F_1 for Various Tube Diameters and Layouts

Tube Outside Diameter, in. (mm)		Tube Pitch, in. (mm)		Layout	F_1
5/8	(15.88)	13/16	(20.6)		0.90
5/8	(15.88)	13/16	(20.6)		1.04
3/4	(19.05)	15/16	(23.8)		1.00
3/4	(19.05)	15/16	(23.8)		1.16
3/4	(19.05)	1	(25.4)		1.14
3/4	(19.05)	1	(25.4)		1.31
1	(25.4)	1 1/4	(31.8)		1.34
1	(25.4)	1 1/4	(31.8)		1.54

$$F_1 = \frac{(\text{Heat transfer area/cross-sectional area of unit cell})\,\text{Reference}}{(\text{Heat transfer area/cross-sectional area of unit cell})\text{New Case}}$$

This table may also be used for low-finned tubing in the following way. The value estimated for h_o from Table 4.1.1 should be multiplied by the fin efficiency (usually between 0.75 and 1 for a good application; 0.85 is a good estimate) and used in Equation 4.1.14 with $A^* = A_o$, the total outside heat transfer area including fins. Then this value of A_o is divided by the ratio of the finned tube heat transfer area to the plain tube area (per unit length). The result of this calculation is used as A_o in Equation 4.1.17 to find A'_o to enter Figure 4.1.8.

Source: Hewitt, G.F., Ed., *Heat Exchanger Design Handbook*, Begell House, New York, 1998. With permission.

TABLE 4.1.3
Values of F_2 for Various Numbers of Tube-Side Passes[a]

	F_2			
	Number of Tube-Side Passes			
Inside Shell Diameter, in. (mm)	**2**	**4**	**6**	**8**
Up to 12 (305)	1.20	1.40	1.80	—
13¼ to 17¼ (337 to 438)	1.06	1.18	1.25	1.50
19¼ to 23¼ (489 to 591)	1.04	1.14	1.19	1.35
25 to 33 (635 to 838)	1.03	1.12	1.16	1.20
35 to 45 (889 to 1143)	1.02	1.08	1.12	1.16
48 to 60 (1219 to 1524)	1.02	1.05	1.08	1.12
Above 60 (above 1524)	1.01	1.03	1.04	1.06

Source: Hewitt, G.F., Ed., *Heat Exchanger Design Handbook*, Begell House, New York, 1998. With permission.

[a] Since U-tube bundles must always have at least two passes, use of this table is essential for U-tube bundle estimation. Most floating head bundles also require an even number of passes.

TABLE 4.1.4
F_3 for Various Tube Bundle Constructions

	F3				
	Inside Shell Diameter, in. (mm)				
Type of Tube Bundle Construction	Up to 12 (305)	13–22 (330–559)	23–36 (584–914)	37–48 (940–1219)	Above 48 (1219)
Split backing ring (TEMA S)	1.30	1.15	1.09	1.06	1.04
Outside packed floating heat (TEMA P)	1.30	1.15	1.09	1.06	1.04
U-Tube[a] (TEMA U)	1.12	1.08	1.03	1.01	1.01
Pull-through floating head (TEMA T)	—	1.40	1.25	1.18	1.15

Source: Hewitt, G.F., Ed., *Heat Exchanger Design Handbook*, Begell House, New York, 1998. With permission.

[a] Since U-tube bundles must always have at least two tube-side passes, it is essential to use Table 4.1.3 also for this configuration.

Once a value of A_o' has been calculated from Equation 4.1.18, enter the ordinate of Figure 4.1.8 at that value and move horizontally, picking off the combinations of shell inside diameter and tube length that meet that requirement. The final choice can then be made from among those possibilities.

Example of the Approximate Design Method

Problem statement. Estimate the dimensions of a shell-and-tube heat exchanger to cool 100,000 lb_m/h (12.6 kg/s) of liquid toluene from 250°F to 110°F (121.1°C–43.3°C) using cooling tower water available at 80°F (26.7°C). Use split-ring floating head construction (TEMA S) with 3/4 in. (19.05 mm) outside diameter × 14 BWG (0.083 in. = 2.11 mm wall) low-carbon steel tubes on 15/16 in. (23.8 mm) equilateral triangular pitch. This construction implies one shell-side pass and an even number of tube-side passes—assume two for the present. Choose cooling water exit temperature of 100°F (37.8°C). Specific heat of toluene is 0.52 Btu/lb_m °F (2177 J/kg K) and viscosity at 180°F (82.2°C) is 0.82 lb_m/ft h (0.34×10^{-3} N·s/m^2 or 0.34 cP).

Solution.

$$q_T = \left(100{,}000 \text{ lb}_m / \text{h}\right)\left(0.52 \text{ Btu} / \text{lb}_m \,^\circ\text{F}\right)\left(250 - 110\right)^\circ\text{F}$$

$$= 7.28 \times 10^6 \text{ Btu} / \text{h} = 2.14 \times 10^6 \text{ W}$$

$$\text{LMTD}_{\text{countercurrent}} = \frac{(250 - 100) - (110 - 80)}{\ln \dfrac{250 - 100}{110 - 80}} = 74.6^\circ\text{F} = 41.4^\circ\text{C}$$

Since there are at least two tube-side passes, flow is not countercurrent, and $T_{h_o} > T_{c_o}$, estimate $F \approx 0.9$. Therefore, MTD = 0.9 (74.6°F) = 67.1°F = 37.3°C.

Estimation of U_o. Light organic liquid cooled by liquid water. (Note that 1 Btu/h ft² °F = 5.678 W/m^2 K.)

Water (in tubes) h_i	1000 Btu/h ft² °F	5700 W/m^2 K
Toluene (in shell) h_o	300 Btu/h ft² °F	1700 W/m^2 K
Tube-side fouling R_{f_T}	0.001 h ft² °F/Btu	1.8×10^{-4} m^2 K/W
Shell-side fouling R_{f_o}	0.0005 h ft² °F/Btu	8.8×10^{-5} m^2 K/W

Tube wall resistance (for estimation purposes, this term can be approximated by x_w/k_w, where x_w is the wall thickness):

$$\frac{x_w}{k_w} = \frac{0.083 \text{ in.}}{(12 \text{ in.} / \text{ft})(26 \text{ Btu} / \text{h ft} \,^\circ\text{F})} = 2.7 \times 10^{-4} \text{ h ft}^2 \,^\circ\text{F} / \text{Btu} = 4.6 \times 10^{-5} \text{ m}^2 \text{ K} / \text{W}$$

Then,

$$U_o = \frac{1}{\frac{0.750}{1000(0.584)} + \frac{0.001(0.750)}{0.584} + 2.7\times10^{-4} + 0.0005 + \frac{1}{300}}$$

$$= 150 \text{ Btu / h ft}^2\,^\circ\text{F} = 848 \text{ W / m}^2 \text{ K}$$

$$A_o = \frac{7.28\times10^6 \text{ Btu/h}}{(150 \text{ Btu /h ft}^2\,^\circ\text{F})\,(67.1^\circ\text{F})} = 723 \text{ ft}^2 = 67.7 \text{ m}^2$$

Correct for changes in construction features (preliminary examination of Figure 4.1.8 indicates shell inside diameter will be in the range of 500 mm, or 20 in.):

F_1:F_1 = 1.00 since the same tube size and layout is used
F_2:F_2 = 1.04, assuming two passes
F_3:F_3 = 1.15, TEMA S construction

$$A_o' = \left(723 \text{ ft}^2\right)(1.00)(1.04)(1.15) = 865 \text{ ft}^2 = 81 \text{ m}^2$$

From Figure 4.1.8, entering at A_o' pick off the following combinations of shell inside diameter and tube length:

Shell Inside Diameter		Effective Tube Length		L/D_s
in.	mm	ft	m	
27	686	6.6	2.0	2.9
25	635	7.5	2.3	3.6
23¼	591	9.2	2.8	4.7
21¼	540	10.8	3.3	6.1
19¼	489	13.1	4.0	8.2
17¼	438	16.7	5.1	11.6

Any of these combinations would supply the desired area; the 21¼ in. (540 mm) and 19¼ in. (489 mm) would appear to be likely choices.

Air-Cooled Heat Exchangers

Chemical process plants as well as steam power plants require the rejection of large quantities of heat at low temperatures. The traditional sink has been supplies of natural water, either on a once-through basis where a large river, lake, or ocean was available, or on a recycle basis through a cooling tower operation. With the decreasing number of suitable sites available for providing cooling water with the increasing restrictions on the thermal and chemical quality of water returned to natural sources, there has been greater interest in the use of air as a cooling medium. In fact, air has always been of interest in special applications where water was very scarce or nonexistent or where the temperature levels of heat rejection were such as to render the use of water inadvisable from a fouling standpoint. However, there has been an enormous growth in the use of air, and particularly in the use of air for rejection of heat at fairly low-temperature levels, on the order of 120°F or so.

The main reason for using air cooling is that air is always available wherever a process plant may be located on the surface of the earth, but one can argue also that air is essentially noncorrosive and nonfouling for the usual range of process applications. There are many reasons against using air. It has poor physical properties from the standpoint of the heat exchanger designer. It has low-thermal conductivity, low density, and a modest specific heat. The low-thermal conductivity and low density combine to produce a low heat transfer coefficient while the low density implies that very large volumes of air will be required to pick up a reasonable amount of heat. The low density and large volume flow rate also mean that, within reasonable pumping power limitations, the equipment available to move the air can only sustain very low pressure drops and, therefore, can only provide low velocities. The low velocity in turn further reduces the possibilities for a high heat transfer coefficient. If large quantities of water were available with no more than reasonable restrictions on the temperature at which that water could be discharged into the environment, air cooling would seldom be used in processing equipment.

Considering the disadvantages of using air, it is not surprising that the equipment designed for using air for cooling purposes has some very special features. Because of the low heat transfer coefficient on the air side, there is great incentive for using high-finned tubes with a ratio of total outside heat transfer area to inside heat transfer area up to 25. This helps to even out the product of heat transfer coefficient and the heat transfer area for the air side as compared to the process (tube) side, that is, $h_o A_o \rightarrow h_i A_i$. However, for a wide variety of processes, such as condensing stream or cooling aqueous streams, the controlling resistance to heat transfer will still be on the air side of the heat exchanger.

The combination of large volume flow rates with low allowable pressure drop combines to fix the geometrical configuration of the air-cooled equipment into the form of large flow area, shallow tube banks. The provision of air from the surrounding atmosphere to these tube banks requires not only large surface areas in the heat exchangers but large, open areas adjacent to the heat exchangers in which atmospheric air may flow in to supply the requirements of the heat exchangers. Quite literally, one may talk in terms of acres of air-cooled heat exchangers being required to provide the heat dissipation requirements of even modest-sized process plants and power plants. The industrial standard for air-cooled heat exchangers is largely set by API 661 (API 2013).

Equipment Configurations

Finned tube type. Various finning constructions of high-finned tubes are diagrammed in Figure 4.1.9. The basic problem is to provide the dense spacing of the fins on the outside of the base tube, to provide reasonable mechanical integrity of those fins, and to provide good thermal contact between the base of the fins and the tube.

1. The cheapest configuration is the so-called edge-wound or tension-wound fin, which is simply a thin strip of metal wound in tension around the outside of the plain base tube. One purpose of the tension is to keep the base of the fin in good contact with the base tube. The fin must be tightly secured to the tube at least at each end or it will tend to coil up like a giant Slinky toy.

   The tube and the fin need not be the same metal, though there may be some accelerated corrosion problems as the result of electrochemical cells being formed in damp atmospheres if the two metals are different. Very commonly, the fins are aluminum and the base tube is low carbon steel. The performance of these tubes tends to deteriorate substantially with time, as differential expansion takes place between the aluminum fins and the base tube as a result of operating temperature transients. The fins, being at a lower temperature, may tend to be expanded by the outward pressure of the tube on the inside, causing the fins to deform plastically and fail to fit as tightly as they did originally. Rough physical handling can accomplish the same thing. Corrosion and/or dirt deposits establishing themselves at the base of the fins may also impede heat flow from the base tube into the fins.

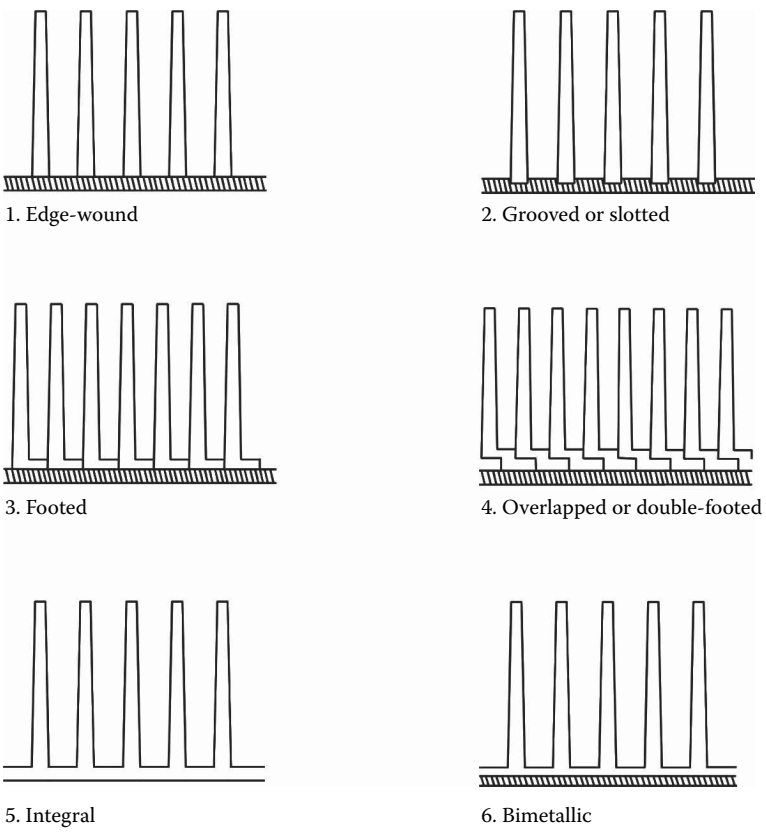


FIGURE 4.1.9 Schematic diagram of various configurations for finned tubes for air-cooled exchangers.

2. An improved version (compared to (1)) is made by tension-winding the finned strip into grooves or slots in the tube, followed by a roller tool to deform the upset metal of the slot back over the base of the fin in order to give an improved mechanical bond between fin and tube. To some extent the groove protects the fin from the atmosphere and reduces the possibilities of accelerated corrosion and/or dirt deposition at the base of the fin as well as making the fin mechanically more strongly bonded to the tube. This construction, of course, costs somewhat more than the simple edge-wound finned tube.
3. The footed fin differs from the above in having a portion of the fin bent at right angles to the main portion of the fin so as to form a small foot which is held in contact upon the base tube by the tension in the fin. If the fin and tube materials are compatible, the fin foot can be spot or continuously welded to the tube. This construction also tends to minimize the possibilities of corrosion and dirt deposition and to provide better thermal contact between the fin and the bare tube. Deterioration over time of the contact between the foot and the tube still occurs as a result of thermal cycling and penetration of atmospheric moisture into the space between the foot of one fin and the adjacent fin.
4. The next improvement is shown in the figure as the overlapped or double-footed fin in which each fin rests upon the foot of the fin immediately adjacent to it, giving further protection against penetration of atmospheric moisture and dirt and providing good thermal contact between the fin and the tube. All of the cases so far discussed may have fins of quite different material from the bare tube (unless welding of fin to tube is intended).
5. The integral fin is formed by a series of machining operations that raise the fin from the original thick-walled plain tube, leading to a fin that is now mechanically integral with the

base tube. This operation can only be performed on fairly soft, mechanically workable material such as aluminum or copper. The advantage is that there is now no possibility of thermal contact resistance developing between the fin and the base tube inasmuch as they are both the same piece of metal. The possibilities of corrosion are very much reduced as well as the penetration of fouling material into the area of the base of the fin. Also, this fin arrangement is much less susceptible to mechanical damage and stripping of the fins than the separately finned tubes mentioned earlier. The disadvantages of this tube are that firstly, it is more expensive, and secondly, the base tube must be of the same metal as the fin and construction is restricted to only a very few metals. If one is trying to cool a substance in the tube that is not compatible with one of the metals from which the integrally finned tube configuration can be made, this configuration cannot be used.

6. The "ultimate" refinement of the finned tube is the bimetallic tube. In this case, a close-fitting liner of essentially any metal required by the corrosive properties of the in-tube fluid may be placed inside the outer tube which is then swaged onto the liner and upset to form the fins. The outer tube is almost always aluminum for its high workability, high thermal conductivity, and low cost. The liner tube may be essentially any metal and may be of whatever wall thickness is required to provide protection against high pressure on the tube side and/or penetration of the tube wall by a pitting-type corrosion action. The disadvantages of the bimetallic tube are that firstly, they are relatively expensive, and secondly, a thermal contact resistance has been introduced between the liner and the finned outer of the tube. The tubes, as originally produced, ordinarily show very low thermal resistance at the contact surface. However, as time goes on, and the tube goes through many thermal cycles, the effect will be generally to develop a possibly very appreciable contact resistance. This problem has been studied at length by Kulkarni and Young (1966).

The choice of which finned tube to select is strongly dependent upon the particular circumstance of the application. In critical or especially severe applications, it will be necessary to go to a better construction at a higher price.

Tube-bank layouts. As shown in Figure 4.1.10, high-finned tubes may be arranged either as equilateral triangular layouts or as rotated square arrangements, though the latter are rather less common than the triangular ones. The tubes are ordinarily spaced very closely together, typically between 1/8 and 1/4 in. being allowed between the outer fins of adjacent tubes. It is necessary to leave some fin-tip-to-fin tip clearance to avoid contact of the fins during the minor vibration that will occur in field applications; yet, one does not wish to place the tubes so far apart that the air can flow through between the fin tips without mixing thoroughly with the air flowing across the fins.

In-line square arrangements are never used with high-finned tubes. It has been conclusively demonstrated (Weierman et al. 1978) that the air will tend to flow in the channels between the fin tips from the inlet to the tube bank to the exit, with little mixing occurring between this air and the air that has actually penetrated in between the fins and taken part in the heat transfer process. The net result is a severe deterioration in the apparent heat transfer coefficient on the fins of the tubes. In fact, this is a classic example of internal flow bypassing (Bell and Kegler 1978).

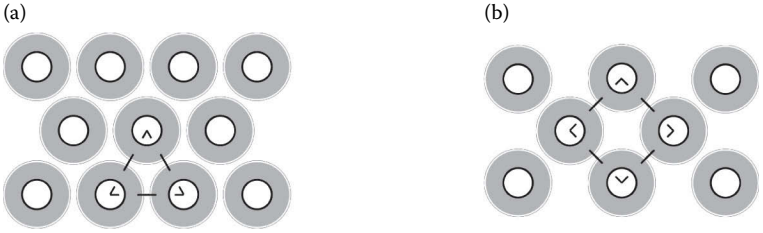


FIGURE 4.1.10 Finned tube unit cell geometries: (a) Equilateral triangular layout and (b) rotated square.

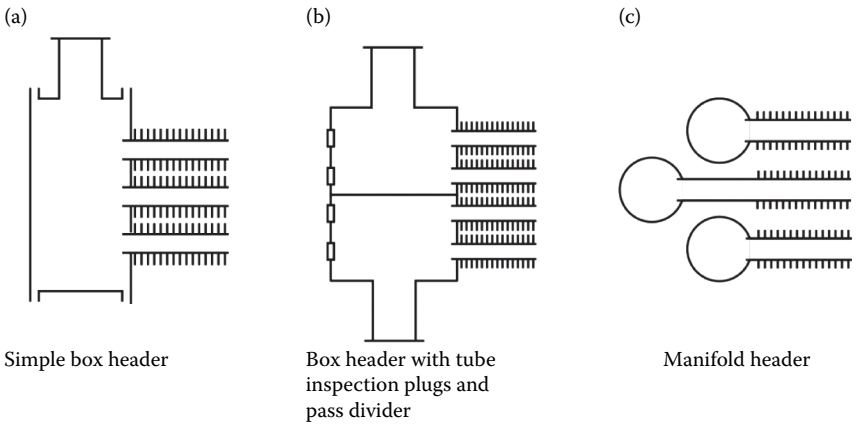


FIGURE 4.1.11 Header configurations for high-finned tubes.

Tube-bank construction. The tubes are commonly inserted into box headers, as shown in Figure 4.1.11. The tubes can be expanded into place if leakage of the tube-side fluid is not serious and if the tube-side fluid is flowing at relatively low pressures. However, the more common construction is for the tubes to be welded to the tubesheet portion of the header. The box header may be closed on the face opposite the tubesheet by a flanged cover with suitable gasketing if the pressure is not too great, or by a welded cover if greater mechanical integrity is required. If a welded cover is used, it is customary to use plugs opposite each tube as shown in Figure 4.1.11, to permit inspection and, if necessary, tube-side cleaning without cutting apart the header.

For high-pressure tube-side fluids, it is common to use a pipe header into which the tubes are inserted and then welded. In this design, inspection of the tube interior is not possible, and tube replacement is considerably more difficult than with the box-header configuration.

The box headers may have one or more pass partition as shown in Figure 4.1.12 in order to provide two or more passes through the tubes in series. When this is done, it is customary for the hot tube-side fluid to enter at the top of the header and flow downward through the consecutive passes, resulting in some approximation to countercurrent flow of air and process fluid and an improved MTD.

Air-cooled heat exchanger arrangement. There are two basic air-cooled heat exchanger arrangements as shown in Figure 4.1.13. The first configuration, called the forced-draft arrangement, has the fan underneath the tube-bank pushing air across the bank of finned tubes and discharging into the atmosphere above. The other arrangement, induced draft, has the fan located above the bank of tubes, drawing air across the tube bank and then discharging it through an upper plenum back into the atmosphere.

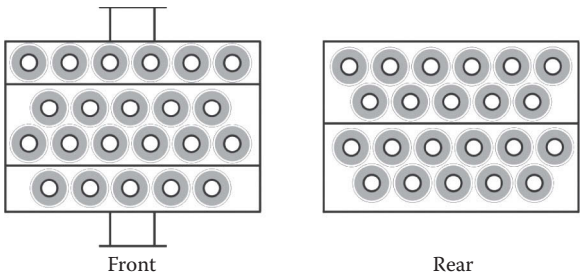


FIGURE 4.1.12 Front and rear header arrangements for a four rows, four pass air cooler.

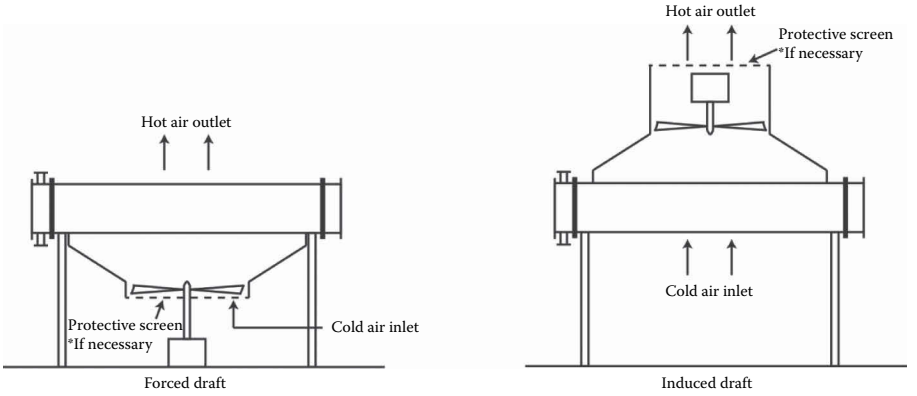


FIGURE 4.1.13 Forced-draft and induced-draft arrangements of air-cooled heat exchangers.

The forced-draft arrangement has the advantage of putting the fan and driver underneath the tube bank where they can be easily reached for maintenance and repair. The driver and fan are both operating in and with atmospheric temperature air, leading to fewer maintenance problems, a higher fan efficiency, and a wider choice of less expensive materials of construction for the fan. Among the disadvantages of this kind of construction are as follows: (1) the fan is exposed to places where operating personnel may walk and there is therefore a greater danger to the personnel; (2) the fan is more likely to ingest trash and debris, damaging the fan; (3) the discharge velocity from the fan is not very uniform and leads to a wide range of actual local air flow velocities across the tube bank, slightly reducing the thermal efficiency, especially in applications for close-temperature approach between the air and the tube-side fluid; (4) the air is discharged from the top of the heat exchanger at fairly low velocities and can mix with the incoming fresh air, giving rise to partial recirculation of the hot air and an increase in the inlet air temperature to the heat transfer surface; and (5) the top of the heat exchanger is exposed to the weather and mechanical damage and may suffer degradation and corrosion.

The induced-draft arrangement has advantages where the forced-draft exchanger has disadvantages and vice versa. That is, the fan and driver on the induced-draft exchanger are located above the heat exchanger where they are difficult to service and where they are exposed to the hot air coming off of the heat exchanger, possibly requiring selection of special materials of construction to satisfy the operating temperature limits. The fan, operating on the heated, less-dense air, tends to be less efficient in moving air. On the other hand, the fan is above the heat exchanger where it is not likely to constitute a danger to even careless personnel. A more nearly uniform air flow occurs across the tube bank, and the discharge plenum can produce a reasonably high-velocity plume of hot air, discharging it high into the atmosphere where there is much less likelihood of recirculation, and consequently deterioration of the inlet temperature difference.

One variant that has its own combination of advantages and disadvantages is to mount the driver for the fan underneath the heat exchanger with a long shaft extending up through the tube bundle to the fan above. The driver is out of the hot air and down where it can be more readily maintained. The fan, however, must be connected to the driver by a long shaft which must be supported against shaft whip and similar instabilities. The penetration of the shaft through the tube field may produce a local inhomogeneity in the tube field (i.e., an internal bypass area), which will at least locally reduce the efficiency of heat transfer in the heat exchanger.

Other configurations of air-cooled exchangers are used for special situations. The A-frame exchanger (Figure 4.1.14) provides more heat transfer surface for a given ground area. For steam and geothermal power plants, the A-frame configuration provides a shorter run for the low-pressure exhaust steam lines. Smaller A-frame exchangers have been used as condensers at the top of distillation columns.

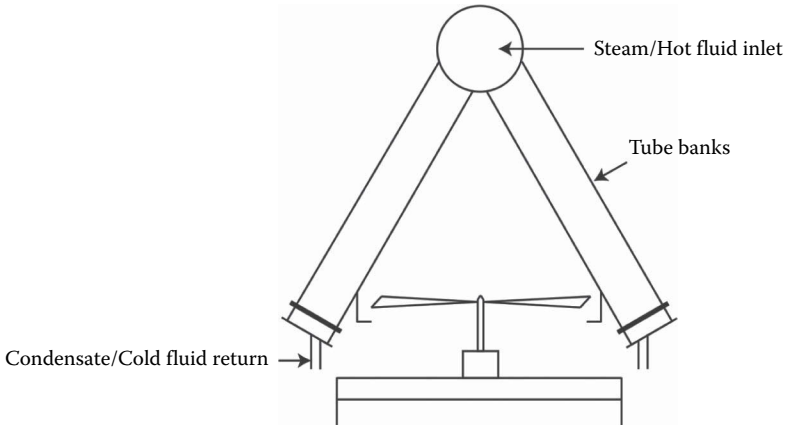


FIGURE 4.1.14 A-frame air-cooled heat exchanger.

Air-cooled exchangers are often used in areas in which the ambient air temperature is well below freezing conditions for the fluid (often water) being cooled and/or condensed. The solution is to provide a means of controlling the inlet air temperature by deliberately and controllably recirculating a portion of the exit hot air from the exchanger. A typical configuration is shown in Figure 4.1.15. By controlling the blowers on the outlet and exit air streams and the recirculating interval flow, the inlet air temperature to the tube bundle can be adjusted to the desired operational condition.

Another variant often used in natural gas production areas is shown in Figure 4.1.16. This unit, typically mounted on a concrete pad with a natural gas fueled diesel engine and a natural gas compressor, provides a variety of cooling services, including condensation of heavy components in the natural gas, interstage cooling of the compressed natural gas, and cooling the gas engine and compressor coolant.

The fan is driven by the gas engine. The several cooling services (1–5) share a common fan but are individually designed as to tube size, number, and tube-bank configuration. Since not all of the tube sections receive the same air throughput (taking into account inherent air flow maldistribution off of the fan and the differences in "exposed" tube length among the different services), allowance must be made in the actual heat transfer surface allocated to each.

Again, the final choice of the proper configuration depends very much upon the particular circumstances operating in each particular case.

Fans for air-cooled heat exchangers. The selection of fans for air-cooled heat exchangers is a rather complex art and the reader is referred to the literature for a more detailed treatment of this

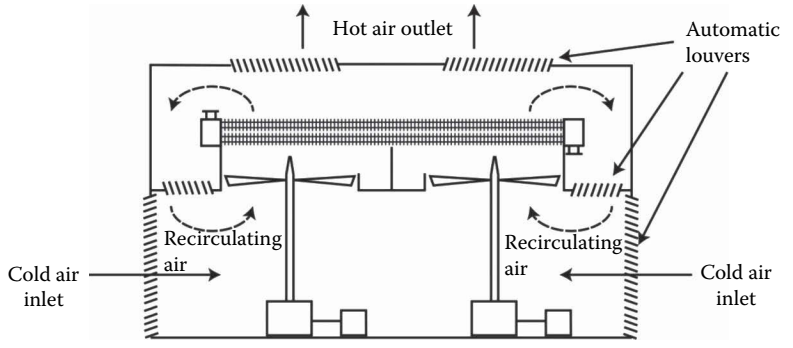


FIGURE 4.1.15 Recirculating air-cooled heat exchanger.

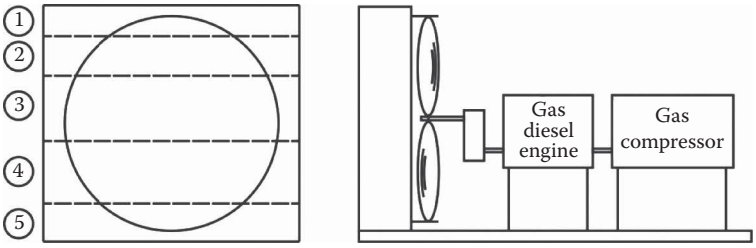


FIGURE 4.1.16 Schematic of skid-mounted well head gas compression unit. (1) Engine jacket water cooler, (2) compressor jacket water cooler, (3) compressor intercooler, (4) compressor aftercooler, and (5) lube oil cooler.

subject (Glass 1978). The fans are rather large diameter, multibladed, fairly low-speed propeller types. The pressure rise capability of these fans is limited to little more than 1 in., which must move the large quantities of air across the banks of finned tubes. A typical design value is on the order of 1/2 in. frictional pressure drop across the tube bank. This limitation allows face velocities (that is the approach velocity of the air to the face of the heat exchanger, assuming a uniform velocity) of about 300–900 ft/min. It may be noted that such air velocities are not greatly different from typical ambient wind velocities at many locations where air-cooled exchangers will be used. (For example, the average annual velocity in Oklahoma City is on the order of 11 or 12 miles/h, or about 1000 ft/min.) This must be taken into account in designing and laying out air-cooled heat exchangers so as to minimize the effect of atmospheric wind currents on the performance of the exchangers.

Basic Equations for Design of Air-Cooled Heat Exchangers

Heat transfer correlations for banks of high-finned tubes. A commonly recommended correlation for heat transfer to banks of high-finned tubes and which is sensitive to the geometrical parameters of the tube is due to Briggs and Young (1963). This equation is:

$$\frac{h_o D_r}{k} = 0.134\left(\frac{D_r \rho V_{max}}{\mu}\right)^{0.68} P_r^{1/3}\left(\frac{H}{S}\right)^{-0.2}\left(\frac{Y}{S}\right)^{-0.12} \tag{4.1.19}$$

The geometrical quantities H, S, and Y are defined in Figure 4.1.17.

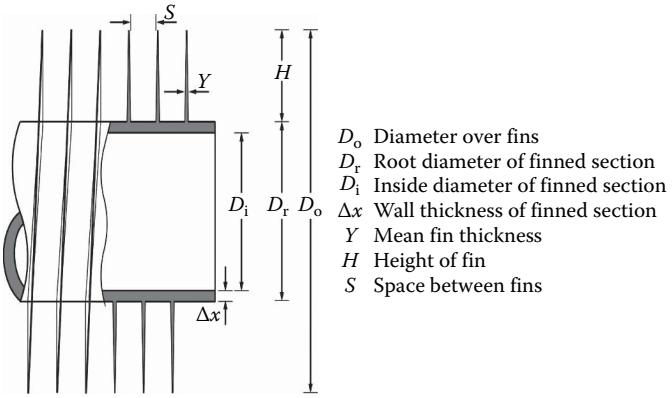


FIGURE 4.1.17 Definition of geometrical parameters for high-finned tubes.

Friction factor correlation for banks of high-finned tubes. The pressure drop for a flow of air across banks of high-finned tubes is usually correlated in terms of a friction factor defined by Equation 4.1.20:

$$f_r = \frac{\Delta P g_c}{n\rho V_{max}^2} \tag{4.1.20}$$

In Equation 4.1.20, n is equal to the number of tube rows crossed by the air.

The best generally available correlation for tubes laid out in an equilateral triangular arrangement is by Robinson and Briggs (1965):

$$f_r = 18.93\left(\frac{D_r \rho V_{max}}{\mu}\right)^{-0.316}\left(\frac{P_t}{D_r}\right)^{-0.927} \tag{4.1.21}$$

where P_t is the transverse pitch between centers of adjacent tubes in the same row. For tubes laid out in an isosceles triangular or rotated square arrangement, the friction factor equation, again from Robinson and Briggs, is

$$f_r = 18.93\left(\frac{D_r \rho V_{max}}{\mu}\right)^{-0.316}\left(\frac{P_t}{D_r}\right)^{-0.927}\left(\frac{P_t}{P_l}\right)^{0.52} \tag{4.1.22}$$

where P_l is the longitudinal pitch between centers of adjacent tubes in different rows measured along the diagonal. It will be noted that when $P_t = P_l$ as it does for an equilateral triangular arrangement, Equation 4.1.22 reduces to Equation 4.1.21.

Fin efficiency. In dealing with heat transfer in finned tubes, it is necessary to take into account the thermal resistance encountered due to flow of the heat through the length of the fin. This results in a reduction of the temperature at the tip of the fin compared to the value at the base of the fin. A sketch of these temperature relationships is shown in Figure 4.1.18.

There are several ways in which this effect may be taken into account, the most common one in the literature being the fin efficiency concept. The fin efficiency, η, is defined as the heat that would be transferred from a real fin (composed of metal having a finite thermal conductivity) divided by the amount of heat that would be transferred by that fin if it were everywhere at its base temperature (i.e., if the fin thermal conductivity were infinite). There has been a great deal of work done on the mathematical analysis of fin efficiency in the literature; much of this work is summarized by Kern

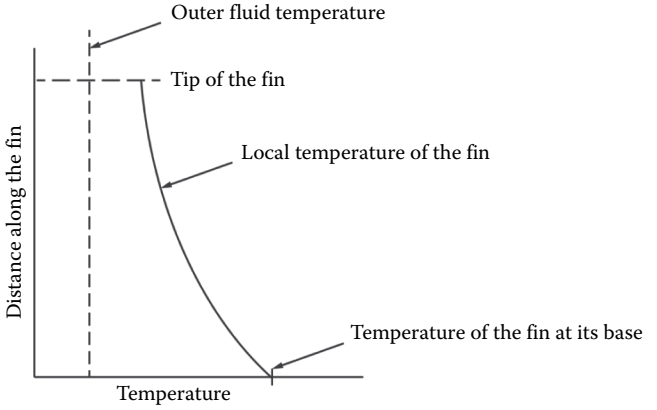


FIGURE 4.1.18 Typical temperature distribution in a fin.

and Kraus (1972). For fins of nearly constant cross section and of dimensions similar to those used in high-fin tubing, there are two approximate expressions sufficient for present purposes. The first set of equations are given as Equations 4.1.23, 4.1.24, and 4.1.25:

$$\eta = \frac{\tanh(H_c B)}{H_c B} \tag{4.1.23}$$

$$B = \left(\frac{2h_o}{k_f Y}\right)^{1/2} \tag{4.1.24}$$

$$H_c = \frac{D_r}{2}\left[\left(\frac{D_o}{D_r} - 1\right)\left(1 + 0.35 \ln \frac{D_o}{D_r}\right)\right] \tag{4.1.25}$$

The other set of equations, which is somewhat more convenient to use, are given in Equations 4.1.26 and 4.1.27:

$$\eta = \frac{1}{1 + \frac{m^2}{3}\sqrt{\frac{D_o}{D_r}}} \tag{4.1.26}$$

$$m = H \sqrt{\frac{2}{\left(\frac{1}{h_o} + R_{fo}\right) k_f Y}} \tag{4.1.27}$$

where R_{fo} is the fouling resistance on the fin surface, usually taken to be zero for air-cooled heat exchangers.

A further concept related to fin efficiency is the fin resistance that incorporates the total thermal resistance of the fin into an equation which can be computed once and for all for a wide variety of cases and used directly in the defining expression for the overall heat transfer coefficient. The fin resistance, R_{fin}, is given by Equation 4.1.28:

$$R_{fin} = \left[\frac{1 - \eta}{\frac{A_{root}}{A_{fin}} + \eta}\right]\left[\frac{1}{h_o} + R_{fo}\right] \tag{4.1.28}$$

where A_{fin} is the heat transfer area of all fins on a unit length of tube, and A_{root} is the surface area of a unit length of plain tube of diameter equal to the root diameter of the finned tube; that is,

$$A_{root} = \pi D_r N_f S = \pi D_r \left(\frac{S}{S + Y}\right) \tag{4.1.29}$$

R_{fin} is shown in Table 4.1.5 as a function of the outside heat transfer coefficient and fouling resistance for a variety of high-finned tubes used in air-cooled tubes used in heat exchangers.

Overall heat transfer coefficient. The overall heat transfer coefficient, based upon the total outside surface area of a heat exchanger with high-finned tubes, is given in terms of the individual coefficients and resistances in Equation 4.1.30.

$$\frac{1}{U_o} = \frac{1}{h_o} + R_{fo} + R_{fin} + \frac{\Delta x}{k_w}\frac{A_o}{A_m} + \left(R_{fi} + \frac{1}{h_i}\right)\frac{A_o}{A_i} \tag{4.1.30}$$

TABLE 4.1.5
Values of ϕ and R_{fin} for Typical High-Finned Tubes

	$h_o = 5$			$h_o = 10$			$h_o = 15$		
Tube Description	m	ϕ	R_{fin}	m	ϕ	R_{fin}	m	ϕ	R_{fin}
$H/R-Al$ $d_r = 1.00$ in., $d_o = 1.875$ in. $H = 0.437$ in. 11 fins/in. $Y = 0.015$ in. $A_{root} = 0.262$ ft²/ft; $A_{fin} = 3.62$ ft²/ft	0.311	0.958	8.19×10^{-3} (122)	0.439	0.919	8.16×10^{-3} (123)	0.538	0.883	8.14×10^{-3} (123)
$H/R-Al$ $d_r = 0.750$ in., $d_o = 1.625$ in. $H = 0.437$ in. 11 fins/in. $Y = 0.015$ in. $A_{root} = 0.196$ ft²/ft; $A_{fin} = 2.99$ ft²/ft	0.311	0.955	8.86×10^{-3} (113)	0.439	0.914	8.83×10^{-3} (113)	0.538	0.876	8.80×10^{-3} (114)
$H/R-Al$ $d_r = 1.00$ in., $d_o = 1.875$ in. $H = 0.437$ in. 5 fins/in. $Y = 0.023$ in. $A_{root} = 0.262$ ft²/ft; $A_{fin} = 1.646$ ft²/ft	0.251	0.972	4.93×10^{-3} (203)	0.355	0.946	4.91×10^{-3} (203)	0.434	0.921	4.90×10^{-3} (204)
$H/R-Al$ $d_r = 0.375$ in., $d_o = 0.9375$ in. $H = 0.281$ in. 11 fins/in. $Y = 0.015$ in. $A_{root} = 0.0982$ ft²/ft; $A_{fin} = 1.063$ ft²/ft	0.1997	0.979	3.84×10^{-3} (260)	0.282	0.960	3.83×10^{-3} (261)	0.346	0.941	3.83×10^{-3} (261)
$H/R-Al$ $d_r = 0.375$ in., $d_o = 0.9375$ in. $H = 0.281$ in. 5 fins/in. $Y = 0.023$ in. $A_{root} = 0.0982$ ft²/ft; $A_{fin} = 0.483$ ft²/ft	0.1613	0.986	2.27×10^{-3} (424)	0.228	0.973	2.27×10^{-3} (441)	0.279	0.961	2.26×10^{-3} (442)
$H/R-Al$ $d_r = 0.750$ in., $d_o = 1.625$ in. $H = 0.437$ in. 5 fins/in. $Y = 0.023$ in. $A_{root} = 0.196$ ft²/ft; $A_{fin} = 1.360$ ft²/ft	0.251	0.970	5.37×10^{-3} (186)	0.355	0.942	5.35×10^{-3} (187)	0.434	0.915	5.33×10^{-3} (188)

Explanation: All of the tubes shown are Wolverine Trufin Type H/R manufactured of 3003 Aluminum. The important geometrical parameters are given. The values of m, ϕ, and R_{fin} are given for $h_o = 5$, 10, and 15 Btu/h ft² °F. The number in brackets in the R_{fin} columns is the reciprocal of R_{fin}, i.e., the effective heat transfer coefficient corresponding to the fin resistance in Btu/h ft² °F.

In addition to the terms shown in Equation 4.1.30, it is necessary to add a term for the contact resistance between the fin and the tube if the fin is not integrally bonded to the tube, and also a contact resistance between the finned tube and a liner if using a bimetallic tube. The contact resistance between the fin and the tube can only be estimated from experience. The contact resistance between a finned tube and a liner has been studied analytically and to a very limited degree experimentally, with the results being reported in quantitative form in Kulkarni and Young (1966).

Mean temperature difference. In crossflow heat exchangers, the MTD is related to the LMTD as calculated for countercurrent flow in the usual way, as shown by Equation 4.1.31:

$$\text{MTD} = F(\text{LMTD})_{\text{countercurrent}} = F\left(\frac{(T_1 - t_2) - (T_2 - t_1)}{\ln\left(\frac{T_1 - t_2}{T_2 - t_1}\right)}\right) \tag{4.1.31}$$

The temperature definitions for a single tube-side pass (all tubes are treated as being in parallel) are shown in Figure 4.1.19. The correction factors for this configuration have been analyzed for the case of a single tube-side pass with the tube-side fluid and the air-side fluid each unmixed as they flow through the tube bank and are shown in Figure 4.1.20. The correction factors for two tube-side passes in series with the general flow arrangement being countercurrent to the flow of the air have

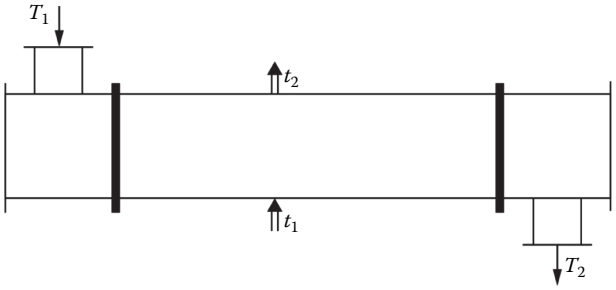


FIGURE 4.1.19 Temperature definitions for a single pass air-cooled heat exchanger.

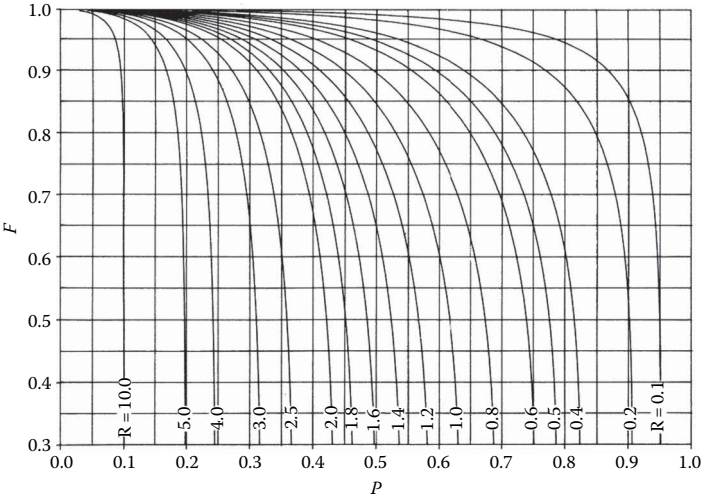


FIGURE 4.1.20 MTD configuration correction factor for crossflow with one tube-side pass. (From Hewitt, G.F., *Heat Exchanger Design Handbook*, Begell House, New York, 2008. With permission.)

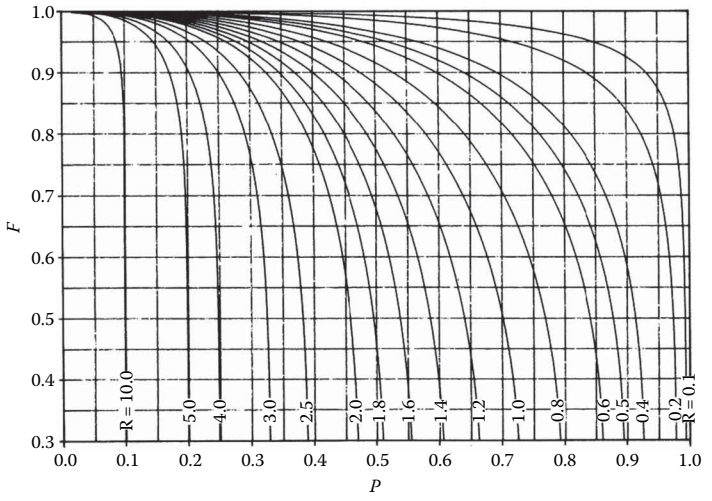


FIGURE 4.1.21 MTD configuration correction factor for crossflow with two tube-side passes. (From Hewitt, G.F., *Heat Exchanger Design Handbook*, Begell House, New York, 2008. With permission.)

been computed as well and are shown in Figure 4.1.21. For more than two tube-side passes, the correction factor F is essentially equal to 1, indicating that this situation results in a flow configuration very close to purely countercurrent flow.

Design Envelope

The possible range of construction and operational parameters for air-cooled heat exchangers is bounded by three limitations which, taken together, constitute the design envelope. The first of these limitations is due to the limited ability of the fan to provide high-pressure drops and will be referred to as the pumping limitation. The second is the thermodynamic limitation, which requires that a certain mass of air be circulated through the heat exchanger in order to remove the heat within the temperature limitations imposed on the air side. The third boundary condition is the rate limitation; that is, the heat transfer area provided must be sufficient that the heat may be transferred into the air with the overall coefficient and the MTD provided for in the design of the heat exchanger. In the following sections, we will examine each of these limitations to see how it tends to fix a fairly narrow range of operating conditions and design parameters for an exchanger for any given service.

Pumping limitation. The maximum design pressure drop obtainable from a typical fan used for air-cooled heat exchangers is on the order of 1 in. H_2O (approximately 0.04 psi). A commonly used design value is on the order of ½ in. H_2O or 0.02 psi. This pressure drop corresponds to a set of values of design face velocities and number of rows of tubes as given in Table 4.1.6.

While the above correspondence is only approximate, it will be found that any attempt to significantly increase the face velocity above the values shown will result in exceeding the allowable pressure that can be developed by the fan. Conversely, a lower velocity than that will tend to result in an excessively large heat exchanger face areas required and less than optimum design of the exchanger from a cost standpoint.

Thermodynamic limitation. The thermodynamic limitation refers to the fact that the weight of the air blown past the heat transfer surface is related to the heat that can be removed and the temperature change on the air side by the following equation:

$$\dot{m}_{\text{air}} = \frac{q_{\text{T}}}{C_{\text{p, air}}\left(t_2 - t_1\right)} \tag{4.1.32}$$

The face area required to satisfy the thermodynamic limitation is given by Equation 4.1.33.

TABLE 4.1.6
Typical Combinations of Number of Tube Rows and Design Face Velocities in Air-Cooled Heat Exchangers

n	V_{face} (ft/min)
3	900
4	800
5	700
6	600
8	500

$$\left(A_{face}\right)_T = \frac{\dot{m}_{air}}{\rho_{air} V_{face}} \tag{4.1.33}$$

in which $\left(A_{face}\right)_T$ is the total face area of the heat exchanger (that is, the length of the tubes×the width of the tube field) required to move the required mass of air at a given face velocity,V_{face}. We can combine Equations 4.1.32 and 4.1.33 to give

$$\frac{\left(A_{face}\right)_T}{q_T} = \frac{1}{c_{p,\,air}\left(t_2 - t_1\right)\rho_{air} V_{face}} \tag{4.1.34}$$

The significance of this equation is illustrated by the following example: 100,000 lb_m/h of water is to be cooled from 180 °F to 120 °F, using air at 90 °F. If the air is to be heated to 140 °F, what face area is required?

Solution:

$$\frac{\left(A_{face}\right)_T}{q_T} = \frac{1}{\left(0.240\ \frac{Btu}{lb_m\,°F}\right)(140-90)°F\left(0.0737\ \frac{lb_m}{ft^3}\right)\left(V_{face,}\frac{ft}{min}\right)\left(60\frac{min}{h}\right)} = \left[\frac{0.01885}{V_{face,}\frac{ft}{min}}\right]\frac{ft^2}{Btu/h}$$

$$q_T = \left(100{,}000\frac{lb_m}{h}\right)\left(1\frac{Btu}{lb_m\,°F}\right)(180-120)°F = 6\times10^6\ Btu/h$$

n	$V_{face}\left(\frac{ft}{min}\right)$	$\frac{\left(A_{face}\right)_T}{q_T}\left(\frac{ft^2}{Btu/\,h}\right)$	$\left(A_{face}\right)_T\left(ft^2\right)$
3	900	2.09×10^{-5}	126
4	800	2.36×10^{-5}	141
5	700	2.69×10^{-5}	162
6	600	3.14×10^{-5}	189
8	500	3.77×10^{-5}	226

Rate limitation. The basic rate equation was presented earlier (Equation 4.1.3). Also, we may write the relationship between the face area required for heat transfer and the number of rows of tubes and the total finned tube heat transfer area:

$$\left(A_{face}\right)_{HT} = \frac{A_o}{nA_{HT}^*} \tag{4.1.35}$$

where A_{HT}^* is the total finned tube heat transfer area per unit face area and per row of tubes. This quantity must be calculated for a given tube geometry and tube layout. Combining Equations 4.1.3 and 4.1.35, we obtain Equation 4.1.36:

$$\frac{\left(A_{\text{face}}\right)_{\text{HT}}}{q_{\text{T}}} = \frac{1}{n\ A^{*}_{\text{HT}} U_{\text{o}}\ (\text{MTD})} \tag{4.1.36}$$

Carrying this out for the example stated in the previous section, we obtain the following results:

n	No. of Tube Passes	F	MTD (°F)	V_{face} (ft/min)	U_{o}	$\frac{(A_{\text{face}})_{\text{HT}}}{q_{\text{T}}}$	$(A_{\text{face}})_{\text{HT}}$ (ft²)
3	1	0.82	28.5	900	12	5.84×10^{-5}	350
4	2	0.91	31.7	800	11	4.29×10^{-5}	258
5	2	0.91	31.7	700	10	3.78×10^{-5}	227
6	3	1.00	34.8	600	9	3.19×10^{-5}	191
8	4	1.00	34.8	500	8	2.69×10^{-5}	161

Resolution of the design. Using the results from the previous three sections, we can determine the combination of number of rows of tubes, the design face velocity, and the required face area for an air-cooled heat exchanger to carry out this service. In Figure 4.1.22, we plot the face area required by the thermodynamic limitation on the left-hand ordinate versus the number of tube rows in the design. On the other ordinate of Figure 4.1.22, we plot the face area required for the heat transfer rate versus the number of tube rows.

The point at which these two curves cross is the design that will satisfy both the thermodynamic limitation and the rate limitation. We can read off directly from Figure 4.1.22 the face area and the number of tube rows required. Note that the pumping limitation is satisfied for all of the points shown by virtue of respecting the relationship between the number of tube rows and the design face velocity as outlined earlier (the subsection *Pumping Limitation*). We see, for the example problem, that the main design features are

$n=6$ rows of tubes
$A_{\text{face}}=190$ ft², or approximately a layout of 10 ft tube field width × 20 ft tube length
$V_{\text{face}}=600$ ft/min

To complete the design, it is, of course, necessary to do a detailed check on the air-side heat transfer coefficient, the fin resistance term, the inside heat transfer coefficient, the overall coefficient, the required heat transfer area, pressure drops on both sides, etc., and rationalize these dimensions to a standard configuration.

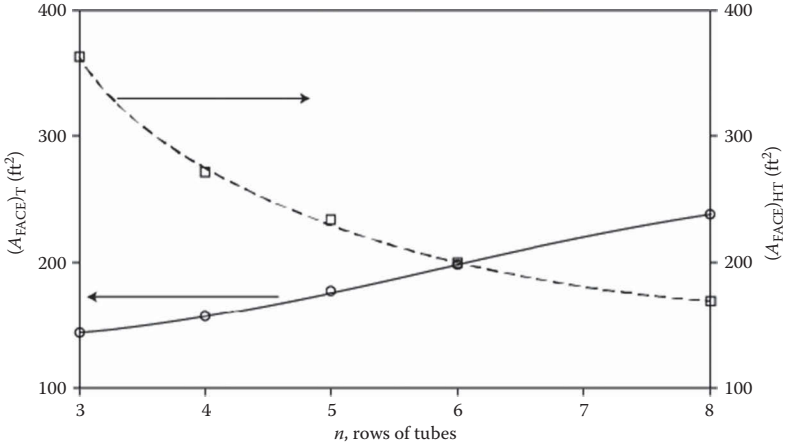


FIGURE 4.1.22 Resolution of the air cooler design.

Nomenclature

A	Total heat transfer area on one fluid side of a heat exchanger, m^2 (ft^2)
$(A_{face})_{HT}$	Total face area of an air-cooled heat exchanger required to satisfy the heat transfer limitation, m^2 (ft^2)
$(A_{face})_T$	Total face area of an air-cooled heat exchanger required to accommodate the air flow rate needed to satisfy the thermodynamic limitation, m^2 (ft^2)
A_{fin}	Heat transfer area of all fins on a unit length of tube, m^2 (ft^2)
A^*_{HT}	Total finned tube heat transfer area per unit face area and per row of tubes
A_m	Logarithmic mean heat transfer area of the bare tube well, m^2 (ft^2)
A'_o	Corrected outside tube heat transfer area used with Figure 4.1.8, ft^2
A_{root}	Surface area of a unit length of plain tube of diameter equal to the root diameter of the finned tube, m^2 (ft^2)
c_p	Specific heat capacity, J/kg·K (Btu/lb·°F)
D_h	Hydraulic diameter, m (in.)
D_I	Inside diameter of fin section, m (in.)
D_o	Diameter over fins, m (in.)
D_r	Root diameter of fin section, m (in.)
d	Tube diameter, m (in.)
d_A	Differential area, m^2 (ft^2)
d_q	Differential heat transfer rate, W (Btu/h)
F	Configuration correction factor, dimensionless
F_1	Correction factor for the tube layout
F_2	Correction factor for the number of tube-side passes
F_3	Correction factor for shell construction/tube bundle configuration
f_r	Friction factor for flow of air across banks of high-finned tubes
g_c	Proportionality constant in Newton's second law of motion, gc = 1 and dimensionless in SI units
H	Height of fin, m (in.)
h	Film heat transfer coefficient, W/m^2·K (Btu/h ft^2 °F)
h_{fg}	Latent heat of phase transformation, J/g (Btu/lb)
K	Fluid thermal conductivity, W/m·K (Btu/h·ft·°F)
k_f	Thermal conductivity of fin material, W/m·K (Btu/h·ft·°F)
k_w	Thermal conductivity of wall material, W/m·K (Btu/h·ft·°F)
L	Tube length, m (ft)
LMTD	Logarithmic mean temperature difference, K (°F)
MTD	Mean temperature difference between the hot and cold streams, K (°F)
$\dot{m}$	Mass flow rate, kg/h (lb/h)
N_f	Number of fins per unit length, m^{-1} (ft^{-1})
N_t	Total number of tubes in the heat exchanger
Nu	Nusselt number, hDh/k, dimensionless
n	Number of tube rows
P	Temperature effectiveness of one fluid, dimensionless
Pr	Prandtl number, μcp/k, dimensionless
P_l	Longitudinal pitch between centers of adjacent tubes in different rows, m (in.)
P_t	Transverse pitch between centers of adjacent tubes in the same row, m (in.)
q_T	Total heat transfer rate, W (Btu/h)
R	Temperature change ratio (also heat capacity rate ratio), dimensionless
Re	Reynolds number, ρvd/μ, dimensionless
R_f	Fouling resistance, m^2·K/W (h ft^2 °F/Btu)
R_{fin}	Fin resistance, m^2·K/W (h ft^2 °F/Btu)
S	Distance between adjacent fins, m (in.)
T_c	Local temperature of the cold stream, K (°F)
T_h	Local temperature of the hot stream, K (°F)
U	Overall heat transfer coefficient, W/m^2·K (Btu/h·ft^2·°F)

V_{face}	Face velocity of air through a air-cooled heat exchanger, m/min (ft/min)
V_{max}	Maximum air velocity across a bank of high-finned tubes, m/min (ft/min)
v	Fluid velocity, m/s (ft/s)
x_w	Wall thickness, m (ft)
Y	Mean fin thickness, m (in.)
ΔP	Pressure drop across a bank of high-finned tubes, Pa
Δx	Wall thickness of finned section, m (in.)
η	Fin efficiency, dimensionless
μ	Fluid dynamic viscosity, Pa·s (cP)
ρ	Fluid density, kg/m^3 (lb/ft^3)
Subscripts	
C	Cold fluid side
H	Hot fluid side
I	Inlet to the exchanger (or inside of tube)
o	Outlet to the exchanger (or outside of tube)
w	Wall or properties at the wall temperature
1	One section (inlet or outlet) of the exchanger
2	Other section (outlet or inlet) of the exchanger

REFERENCES

API (2013). 661: *Petroleum, Petrochemical, and Natural Gas Industries—Air Cooled Heat* Exchangers. American petroleum institute.

Bell, K., and W. Kegler (1978). Analysis of bypass flow effects in tube banks and heat exchangers. *AIChE Symp. Ser.*, 74(174): 47–52.

Briggs, D.E., and E.H. Young (1963). Convection heat transfer and pressure drop of air flowing across triangular pitch banks of finned tubes. *Chem. Eng. Prog. Symp. Ser.*, 59(41): 1–10.

Çengel, Y.A., and A.J. Ghajar (2011). *Heat and Mass Transfer: Fundamentals & Applications*. New York: McGraw-Hill.

Gentry, C.C., R.K. Young and W.M. Small (1982). RODbaffle heat exchanger thermal-hydraulic predictive methods. In *Proceedings of the Seventh International Heat Transfer Conference*, Munich, Gernmany, 6: 197–202.

Glass, J. (1978). Design of air-cooled exchangers; Specifying and rating fans. *Chem. Eng.,* 85(8): 120–124.

Hewitt, G.F. (2008). *Heat Exchanger Design Handbook 2008*. New York: Begell house.

Hewitt, G.F., G.L. Shires and T.R. Bott (1994). *Process Heat Transfer.* Boca Raton, FL: CRC/Begell House.

Kern, D.Q., and A.D. Kraus (1972). *Extended Surface Heat Transfer.* New York: McGraw-Hill.

Kral, D., P. Stehlik, H.J. Van der Ploeg and B.I. Master (1996). Helical baffles in shell and tube heat exchangers. Part I: Experimental verification. *Heat Transfer Eng.,* 17(1): 93–101.

Kulkarni, M., and E. Young (1966). Bimetallic finned tubes. *Chem. Eng. Prog.*, 62(7): 68–71.

Palen, J.W., and J. Taborek (1969). Solution of shell side flow pressure drop and heat transfer by stream analysis method. *Chem. Eng. Prog. Symp. Ser. No. 92, Heat Transfer-Philidelphia* 65: 53–63.

Robinson, K.K., and D.E. Briggs (1965). Pressure drop of air flowing across triangular pitch banks of finned tubes. *Chem. Eng. Prog. Symp. Ser.*, 62(64): 177–184.

Saunders, E.A. D. (1988). *Heat Exchangers: Selection, Design, and Construction*. New York: Longman Scientific & Technical/John Wiley & Sons.

Serth, R.W., and T.G. Lestina (2014). *Process Heat Transfer : Principles, Applications and Rules of Thumb.* Amsterdam: Academic Press.

TEMA (2007). *Standards*, 9th ed. Tarrytown, NY: Tubular Exchanger Manufacturers Association.

Weierman, C., J. Taborek and W. Marner (1978). Comparison of the performance of in-line and staggered banks of tubes with segmented fins. *Am. Inst. Chem. Eng* 74(174): 39–46.

Yokell, S. (1990). *A Working Guide to Shell and Tube Heat Exchangers.* New York: McGraw-Hill.

Zhang, J.-F., Y.-L. He and W.-Q. Tao (2010). A design and rating method for shell-and-tube heat exchangers with helical baffles. *J. Heat Transfer* 132(5): 051802–051802.

COMPACT HEAT EXCHANGERS

Ramesh K. Shah

As noted earlier, a gas-to-fluid heat exchanger is referred to as a compact heat exchanger if it incorporates a heat transfer surface having a surface area density above about 700 m^2/m^3 (213 ft^2/ft^3) on at least one of the fluid sides that usually has gas flow, see Figure 4.1.1. It is referred to as a laminar flow heat exchanger if the surface area density is above about 3000 m^2/m^3 (914 ft^2/ft^3), and as a micro heat exchanger if the surface area density is above about 10,000 m^2/m^3 (3050 ft^2/ft^3). A liquid/two-phase heat exchanger is referred to as a compact heat exchanger if the surface area density on any one fluid side is above about 400 m^2/m^3 (122 ft^2/ft^3). A typical process industry shell-and-tube exchanger has a surface area density of less than 100 m^2/m^3 on one fluid side with plain tubes, and two to three times that with the high-fin density low-finned tubing. Plate-fin, tube-fin, and rotary regenerators are examples of compact heat exchangers for gas flows on one or both fluid sides, and gasketed and welded plate heat exchangers (PHEs) are examples of compact heat exchangers for liquid flows.

Types and Description

Gas-to-Fluid Exchangers

The important design and operating considerations for compact extended surface exchangers are (1) usually at least one of the fluids is a gas or specific liquid that has low h; (2) fluids must be clean and relatively noncorrosive because of small hydraulic diameter (D_h) flow passages and no easy techniques for mechanically cleaning them; (3) the fluid pumping power (i.e., pressure drop) design constraint is often equally as important as the heat transfer rate; (4) operating pressures and temperatures are somewhat limited compared with shell-and-tube exchangers as a result of the joining of the fins to plates or tubes such as brazing, mechanical expansion, etc.; (5) with the use of highly compact surfaces, the resultant shape of a gas-to-fluid exchanger is one having a large frontal area and a short flow length (the header design of a compact heat exchanger is thus important for a uniform flow distribution among the very large number of small flow passages); and (6) the market potential must be large enough to warrant the sizable manufacturing research and tooling costs for new forms to be developed.

Some advantages of plate-fin exchangers over conventional shell-and-tube exchangers are as follows. Compact heat exchangers, generally fabricated from thin metallic plates, yield large heat transfer surface area per unit volume (β), typically up to 10 times greater than the 50–100 m^2/m^3 provided by a shell-and-tube exchanger for general process application and from 1000 to 6000 m^2/m^3 for highly compact gas-side surfaces. Compact liquid or two-phase side surfaces have a β ratio ranging from 500 to 600 m^2/m^3. A compact exchanger provides a tighter temperature control; thus it is useful for heat-sensitive materials, improves the product (e.g., refining fats from edible oil) and its quality (such as a catalyst bed). Also, a compact exchanger could provide rapid heating or cooling of a process stream, thus improving the product quality. The plate-fin exchangers can accommodate multiple (up to 12 or more) fluid streams in one exchanger unit with proper manifolding, thus allowing process integration and cost-effective compact solutions.

Fouling is one of the potential major problems in compact heat exchangers (except for plate-and-frame heat exchangers), particularly those having a variety of fin geometries or very fine circular or noncircular flow passages that cannot be cleaned mechanically. Chemical cleaning may be possible; thermal baking and subsequent rinsing is possible for small-size units. Hence, extended surface compact heat exchangers may not be used in heavy fouling applications.

Liquid-to-Liquid Exchangers

Liquid-to-liquid and phase-change exchangers are plate-and-frame and welded PHEs, spiral plate, and printed circuit exchangers; some of them are described next in some detail along with other compact heat exchangers and their applications.

Plate-Fin Heat Exchangers

This type of exchanger has "corrugated" fins or spacers sandwiched between parallel plates (referred to as plates or parting sheets) as shown in Figure 4.1.23. Sometimes fins are incorporated in a flat tube with rounded corners (referred to as a formed tube), thus eliminating a need for the side bars. If liquid or phase-change fluid flows on the other side, the parting sheet is usually replaced by a flat tube with or without inserts/webs. Other plate-fin constructions include drawn-cup (see Figure 4.1.24) or tube-and-center configurations. Fins are die- or roll-formed and are attached to the plates by brazing, soldering, adhesive bonding, welding, mechanical fit, or extrusion. Fins may be used on both sides in gas-to-gas heat exchangers. In gas-to-liquid applications, fins are usually used

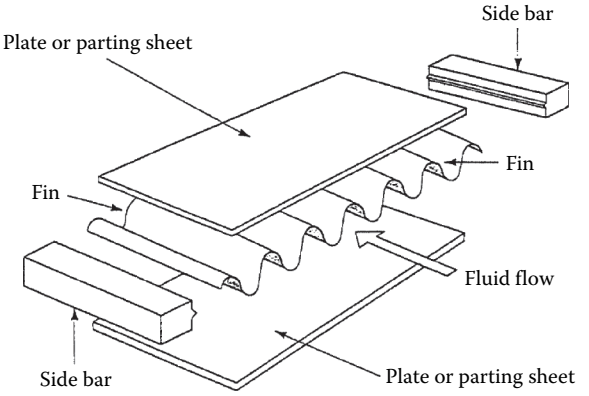


FIGURE 4.1.23 Typical components of a plate-fin exchanger.

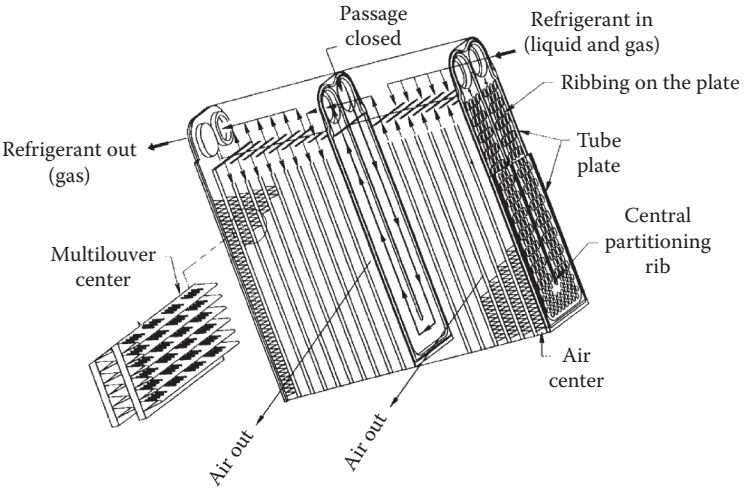


FIGURE 4.1.24 U-channel ribbed plates and multilouver fin automotive evaporator. (Courtesy of Delphi Harrison Thermal Systems, Lockport, NY.)

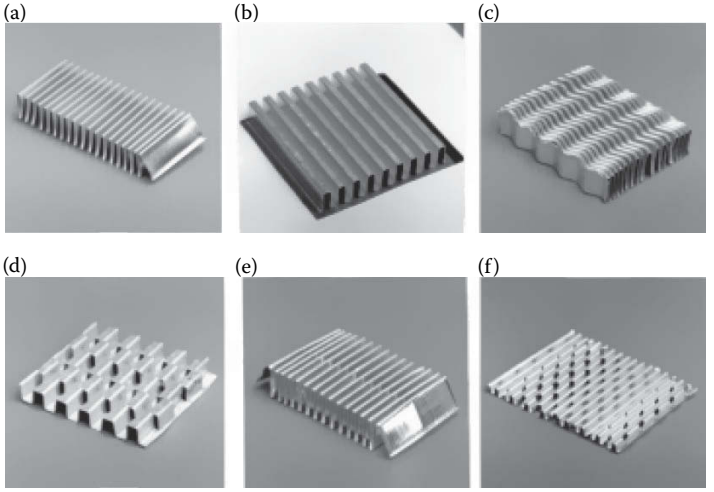


FIGURE 4.1.25 Fin geometries for plate-fin heat exchangers: (a) plain triangular fin, (b) plain rectangular fin, (c) wavy fin, (d) offset strip fin, (e) multilouver fin, and (f) perforated fin.

only on the gas side; if employed on the liquid side, they are used primarily for structural strength and flow-mixing purposes. Fins are also sometimes used for pressure containment and rigidity.

Plate fins are categorized as (1) plain (i.e., uncut) and straight fins, such as plain triangular and rectangular fins; (2) plain but wavy fins (wavy in the main fluid flow direction); and (3) interrupted fins such as offset strip, louver, and perforated. Examples of commonly used fins are shown in Figure 4.1.25.

Plate-fin exchangers have been built with a surface area density of up to about 5900 m^2/m^3 (1800 ft^2/ft^3). There is a total freedom of selecting fin surface area on each fluid side, as required by the design, by varying fin height and fin density. Although typical fin densities are 120–700 fins/m (3–18 fins/in.), applications exist for as many as 2100 fins/m (53 fins/in.). Common fin thicknesses range from 0.05 to 0.25 mm (0.002–0.010 in.). Fin heights range from 2 to 25 mm (0.08–1.0 in.). A plate-fin exchanger with 600 fins/m (15.2 fins/in.) provides about 1300 m^2 (400 ft^2/ft^3) of heat transfer surface area per cubic meter volume occupied by the fins. Plate-fin exchangers are manufactured in virtually all shapes and sizes, and made from a variety of materials.

Tube-Fin Heat Exchangers

In this type of exchanger, round and rectangular tubes are the most common, although elliptical tubes are also used. Fins are generally used on the outside, but they may be used on the inside of the tubes in some applications. They are attached to the tubes by a tight mechanical fit, tension winding, adhesive bonding, soldering, brazing, welding, or extrusion. Fins on the outside of the tubes may be categorized as follows: (1) normal fins on individual tubes, referred to as individually finned tubes or simply as *finned tubes*, as shown in Figures 4.1.26a and 4.1.27; (2) flat or continuous (plain, wavy, or interrupted) external fins on an array of tubes, as shown in Figures 4.1.26b and 4.1.28; (3) longitudinal fins on individual tubes. The exchanger having flat (continuous) fins on tubes has also been referred to as a *plate-fin and tube* exchanger in the literature. In order to avoid confusion with plate-fin surfaces, we will refer to it as a tube-fin exchanger having flat (plain, wavy, or interrupted) fins. Individually finned tubes are probably more rugged and practical in large tube-fin exchangers. Shell-and-tube exchangers sometimes employ low-finned tubes to increase the surface area on the shell side when the shell-side heat transfer coefficient is low compared with the tube-side coefficient. The exchanger with flat fins is usually less expensive on a unit heat transfer surface area basis because of its simple and mass production-type construction features. Longitudinal fins are generally used in condensing applications and for viscous fluids in double-pipe heat exchangers.

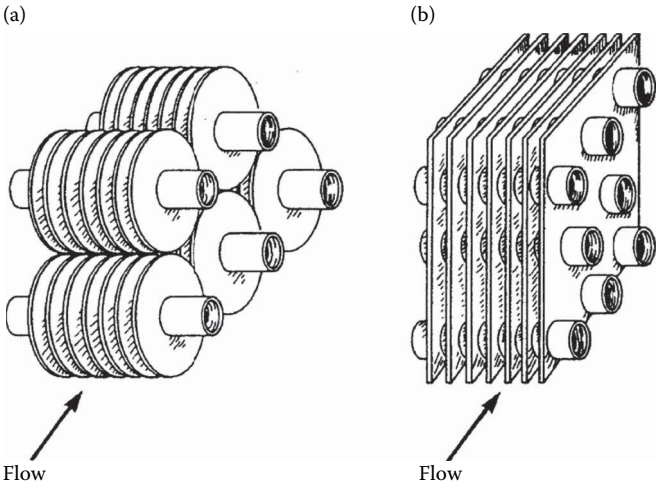


FIGURE 4.1.26 (a) Individually finned tubes, (b) flat or continuous fins on an array of tubes.

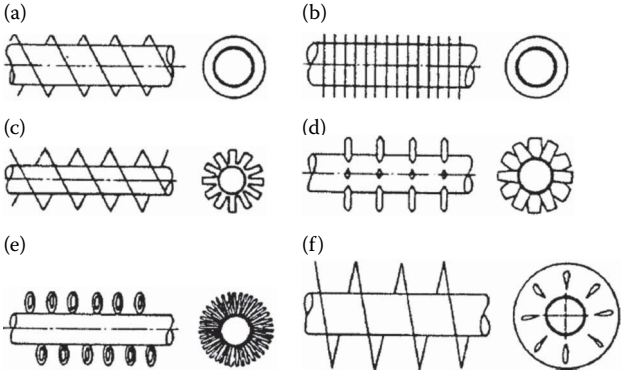


FIGURE 4.1.27 Individually finned tubes: (a) helical, (b) annular disk, (c) segmented, (d) studded, (e) wire loop, and (f) slotted helical.

Tube-fin exchangers can withstand high pressures on the tube side. The highest temperature is again limited by the type of bonding, the materials employed, and the material thickness. Tube-fin exchangers with an area density of about 3300 m^2/m^3 (1000 ft^2/ft^2) are commercially available. On the fin side, the desired surface area can be employed by using the proper fin density and fin geometry. The typical fin densities for flat fins vary from 250 to 800 fins/m (6–20 fins/in.), fin thicknesses vary from 0.08 to 0.25 mm (0.003–0.010 in.), and fin flow lengths from 25 to 250 mm (1–10 in.). A tube-fin exchanger having flat fins with 400 fins/m (10 fins/in.) has a surface area density of about 720 m^2/m^3 (220 ft^2/ft^3). These exchangers are extensively used as condensers and evaporators in air-conditioning and refrigeration applications, as condensers in electric power plants, as oil coolers in propulsive power plants, and as air-cooled exchangers (also referred to as a fin-fan exchanger) in process and power industries.

Regenerators

The regenerator is a storage-type exchanger. The heat transfer surface or elements are usually referred to as a matrix in the regenerator. In order to have continuous operation, either the matrix must be moved periodically into and out of the fixed streams of gases, as in a *rotary* regenerator (Figure 4.1.29a), or the gas flows must be diverted through valves to and from the fixed matrices as in a *fixed-matrix* regenerator (Figure 4.1.29b). The latter is also sometimes referred to as a

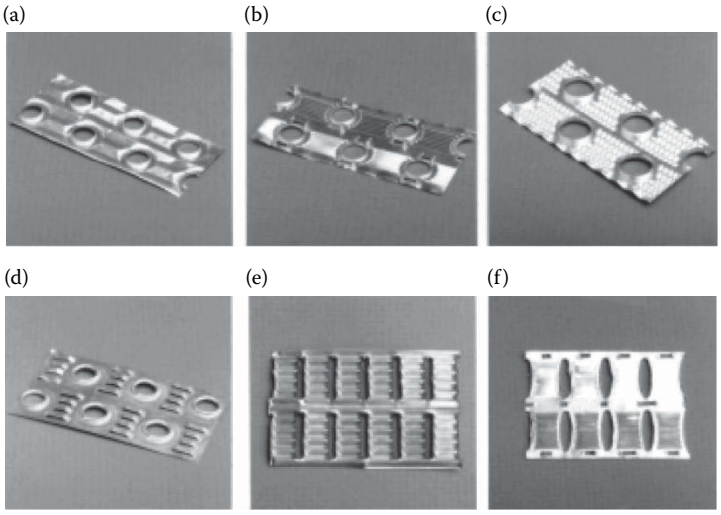


FIGURE 4.1.28 Flat or continuous fins on an array of tubes: On round tubes: (a) wavy fin, (b) multilouver fin, (c) fin with structured surface roughness (dimples), (d) parallel louver fin, (e) louver fin on flat tubes, and (f) multilouver fin on elliptical tubes.

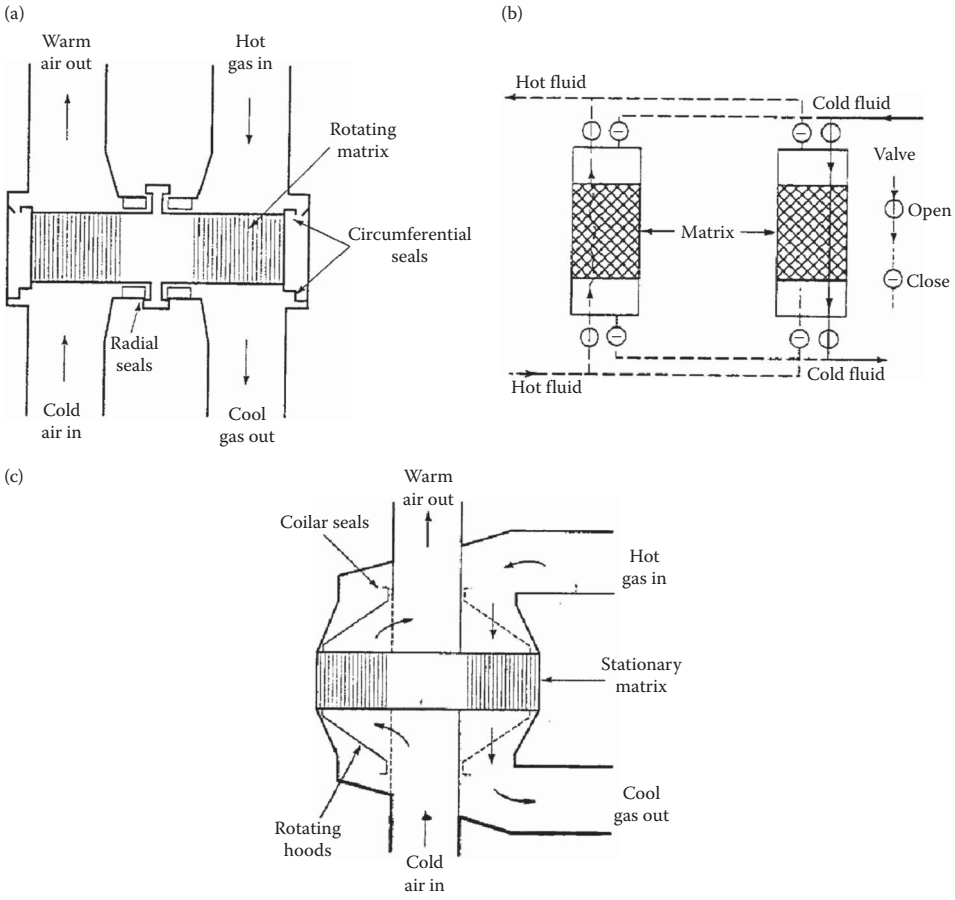


FIGURE 4.1.29 Regenerators: (a) rotary, (b) fixed-matrix, and (c) Rothemuhle.

periodic-flow regenerator or a *reversible heat accumulator.* A third type of regenerator has a fixed matrix (in the disk form) and the fixed stream of gases, but the gases are ducted through rotating hoods (headers) to the matrix as shown in Figure 4.1.29c. This Rothemuhle regenerator is used as an air preheater in some power-generating plants. The thermodynamically superior counterflow arrangement is usually employed in regenerators.

The *rotary regenerator* is usually a disk type in which the matrix (heat transfer surface) is in a disk form and fluids flow axially. It is rotated by a hub shaft or a peripheral ring gear drive. For a rotary regenerator, the design of seals to prevent leakage of hot to cold fluids and vice versa becomes a difficult task, especially if the two fluids are at significantly differing pressures. Rotating drives also pose a challenging mechanical design problem.

Major advantages of rotary regenerators are the following. For a highly compact regenerator, the cost of the regenerator surface per unit of heat transfer area is usually substantially lower than that for the equivalent recuperator. A major disadvantage of a regenerator is an unavoidable carryover of a small fraction of the fluid trapped in the passage to the other fluid stream just after the periodic flow switching. Since fluid contamination (small mixing) is prohibited with liquids, the regenerators are used exclusively for gas-to-gas heat or energy recovery applications. Cross-contamination can be minimized significantly by providing a purge section in the disk and using double-labyrinth seals.

Rotary regenerators have been designed for a surface area density of up to about 6600 m^2/m^3 (2000 ft^2/ft^3), and exchanger effectivenesses exceeding 85% for a number of applications. They can employ thinner stock material, resulting in the lowest amount of material for a given effectiveness and pressure drop of any heat exchanger known today. The metal rotary regenerators have been designed for continuous inlet temperatures up to about 790°C (1450°F) and ceramic matrices for higher-temperature applications; these regenerators are designed up to 400 kPa or 60 psi pressure differences between hot and cold gases. Plastic, paper, and wool are used for regenerators operating below 65°C (150°F) inlet temperature of the hot gas and 1 atm pressure. Typical regenerator rotor diameters and rotational speeds are as follows: up to 10 m (33 ft) and 0.5–3 rpm for power plant regenerators, 0.25–3 m (0.8–9.8 ft) and up to 10 rpm for air-ventilating regenerators, and up to 0.6 m (24 in.) and up to 18 rpm for vehicular regenerators. Refer to Shah (1994) for the description of *fixed-matrix regenerator*, also referred to as a *periodic-flow, fixed bed, valved*, or *stationary* regenerator.

Plate-Type Heat Exchangers

These exchangers are usually built of thin plates (all prime surface). The plates are either smooth or have some form of corrugations, and they are either flat or wound in an exchanger. Generally, these exchangers cannot accommodate very high pressures, temperatures, and pressure and temperature differentials. These exchangers may be further classified as plate, spiral plate, lamella, and plate-coil exchangers as classified in Figure 4.1.1. The PHE, being the most important of these, is described next.

The *plate-and-frame or gasketed PHE* consists of a number of thin rectangular corrugated or embossed metal plates sealed around the edges by gaskets and held together in a frame as shown in Figure 4.1.30. The plate pack with fixed and movable end covers is clamped together by long bolts, thus compressing the gaskets and forming a seal. Sealing between the two fluids is accomplished by elastomeric molded gaskets (typically 5 mm or 0.2 in. thick) that are fitted in peripheral grooves mentioned earlier. The most conventional flow arrangement is one pass to one pass counterflow with all inlet and outlet connections on the fixed end cover. By blocking flow through some ports with proper gasketing, either one or both fluids could have more than one pass. Also more than one exchanger can be accommodated in a single frame with the use of intermediate connector plates such as up to five "exchangers" or sections to heat, cool, and regenerate heat between raw milk and pasteurized milk in a milk pasteurization application.

Typical PHE dimensions and performance parameters are given in Table 4.1.7 (Shah, 1994). Any metal that can be cold-worked is suitable for PHE applications. The most common plate materials are stainless steel (AISI 304 or 316) and titanium. Plates made from Incoloy 825, Inconel 625, and

Hastelloy C-276 are also available. Nickel, cupronickel, and Monel are rarely used. Carbon steel is not used because of low corrosion resistance for thin plates. The heat transfer surface area per unit volume for plate exchangers ranges from 120 to 660 m^2/m^3 (37–200 ft^2/ft^3).

In PHEs, the high turbulence due to plates reduces fouling from about 10%–25% of that of a shell-and-tube exchanger. High thermal performance can be achieved in plate exchangers because the high degree of counterflow in PHEs makes temperature approaches of up to 1°C (2°F) possible. The high thermal effectiveness (up to about 93%) makes low-grade heat recovery economical. PHEs are most suitable for liquid–liquid heat transfer duties.

Welded PHEs. One of the limitations of gasketed PHE is the presence of the gaskets which restricts the use to compatible fluids and which limits operating temperatures and pressures. In order to overcome this limitation, a number of welded PHE designs have surfaced with a welded pair of plates for one or both fluid sides. However, the disadvantage of such design is the loss

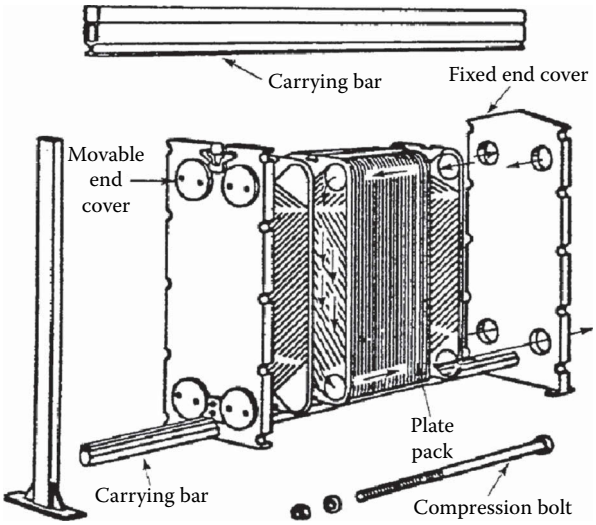


FIGURE 4.1.30 A plate-and-frame or gasketed PHE.

TABLE 4.1.7
Some Geometric and Operating Condition Characteristics of Plate-and-Frame Heat Exchangers

Unit		Operation	
Maximum surface area	2500 m^2	Pressure	0.1–2.5 MPa
Number of plates	3–700	Temperature	–40°C to 260°C
Port size	Up to 400 mm	Maximum port velocity	6 m/s
Plates		Channel flow rates	0.05–12.5 m^3/h
Thickness	0.5–1.2 mm	Maximum unit flow rate	2500 m^3/h
Size	0.03–3.6 m^2	*Performance*	
Spacing	1.5–5 mm	Temperature approach	As low as 1°C
Width	70–1200 mm	Heat exchanger efficiency	Up to 93%
Length	0.6–5 m	Heat transfer coefficients for water–water duties	3000–7000 $W/m^2 K$

Source: Shah, R.K., in *Encyclopedia of Energy Technology and the Environment*, A. Bision and S.G. Boots, Eds., John Wiley & Sons, New York, 1651–1670, 1994. With permission.

of disassembling flexibility on the fluid side where the welding is done. Essentially, welding is done around the complete circumference where the gasket is normally placed. A *stacked plate heat exchanger* is another welded PHE design from Pacinox in which rectangular plates are stacked and welded at the edges. The physical size limitations of PHEs (1.2 m wide × 4 m long max, 4 × 13 ft) are considerably extended to 1.5 m wide × 20 m long (5 × 66 ft) in this exchanger. A maximum surface area of 10,000 m^2 or over 100,000 ft^2 can be accommodated in one unit. The potential maximum operating temperature is 815°C (1500°F) with an operating pressure of up to 20 MPa (3000 psig) when the stacked plate assembly is placed in a cylindrical pressure vessel. For operating pressures below 2 MPa (300 psig) and operating temperatures below 200°C (400°F), the plate bundle is not contained in a pressure vessel, but is bolted between two heavy plates. Some of the applications of this exchanger are catalytic reforming, hydrosulfurization, crude distillation, synthesis converter feed effluent exchanger for methanol, propane condenser, etc.

A number of other PHE constructions have been developed to address some of the limitations of the conventional PHEs. A double-wall PHE is used to avoid mixing of the two fluids. A wide-gap PHE is used for fluids having high fiber content or coarse particles. A graphite PHE is used for highly corrosive fluids. A flow-flex exchanger has plain fins on one side between plates and the other side has conventional plate channels, and is used to handle asymmetric duties (flow rate ratio of 2:1 and higher).

A vacuum *brazed PHE* is a compact PHE for high-temperature and high-pressure duties, and it does not have gaskets, tightening bolts, frame, or carrying and guide bars. It simply consists of stainless-steel plates and two end plates. The brazed unit can be mounted directly on piping without brackets and foundations.

Printed circuit heat exchangers. This exchanger, as shown in Figure 4.1.31, has only primary heat transfer surfaces as PHEs. Fine grooves are made in the plate by using the same techniques as those employed for making printed electrical circuits. High surface area densities (650–1350 m^2/m^3 or 200–400 ft^2/ft^3 for operating pressures of 500–100 bar respectively) are achievable. A variety of materials including stainless-steel, nickel, and titanium alloys can be used. It has been successfully used with relatively clean gases, liquids and phase-change fluids in chemical processing, fuel processing, waste heat recovery, and refrigeration industries. Again, this exchanger is a new construction with limited special applications currently.

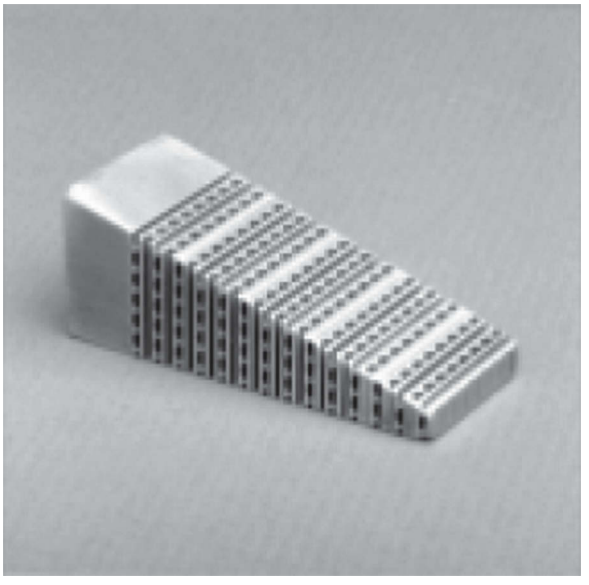


FIGURE 4.1.31 A section of a printed circuit heat exchanger. (Courtesy of Heatric Ltd., Dorset, UK.)

Exchanger Heat Transfer and Pressure Drop Analysis

In this subsection, starting with the thermal circuit associated with a two-fluid exchanger, ε-NTU, P-NTU, and MTD methods used for an exchanger analysis are presented, followed by the fin efficiency concept and various expressions. Finally, pressure drop expressions are outlined for various single-phase exchangers.

Two energy conservation differential equations for a two-fluid exchanger with any flow arrangement are (see Figure 4.1.32 for counterflow) as follows:

$$dq = q''\, dA = -C_h dT_h = \pm C_c dT_c \tag{4.1.37}$$

where the ± sign depends upon whether dT_c is increasing or decreasing with increasing dA or dx. The local overall rate equation is

$$dq = q''\, dA = U\left(T_h - T_c\right)\text{local}\, dA = U\Delta T dA \tag{4.1.38}$$

As shown earlier, integration of Equations 4.1.37 and 4.1.38 across the exchanger surface area results in

$$q = C_h\left(T_{h,i} - T_{h,o}\right) = C_c\left(T_{c,o} - T_{c,i}\right) \tag{4.1.39}$$

and

$$q = UA\Delta T_m = \Delta T_m / R_o \tag{4.1.40}$$

where ΔT_m is the true MTD that depends upon the exchanger flow arrangement and degree of fluid mixing within each fluid stream. The inverse of the overall thermal conductance UA is referred to as the overall thermal resistance R_o as follows (see Figure 4.1.33).

$$R_o = R_h + R_{s,h} + R_w + R_{s,c} + R_c \tag{4.1.41}$$

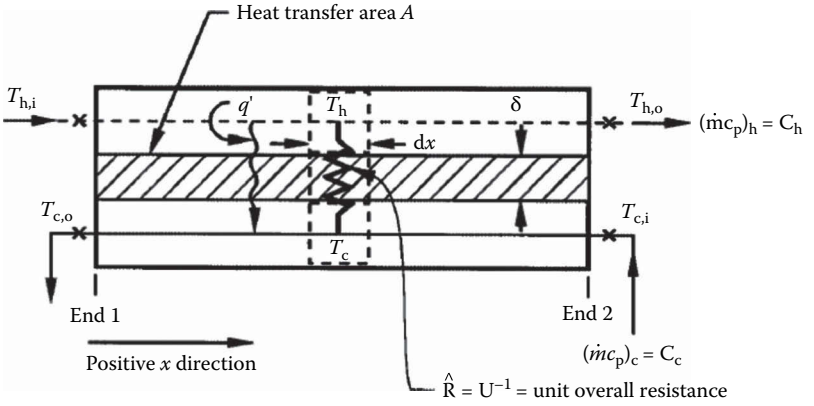


FIGURE 4.1.32 Nomenclature for heat exchanger variables.

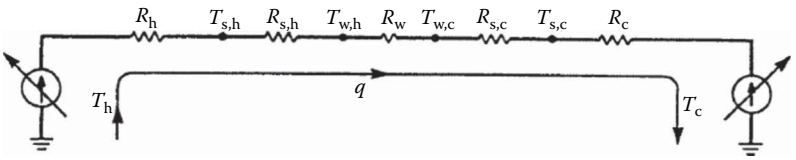


FIGURE 4.1.33 Thermal circuit for heat transfer in an exchanger.

where the subscripts *h*, *c*, *s*, and *w* denote hot, cold, fouling (or scale), and wall, respectively. In terms of the overall and individual heat transfer coefficients, Equation 4.1.41 is represented as

$$\frac{1}{UA}=\frac{1}{\left(\eta_o hA\right)_h}+\frac{1}{\left(\eta_o h_s A\right)_h}+R_w+\frac{1}{\left(\eta_o h_s A\right)_c}+\frac{1}{\left(\eta_o hA\right)_c} \tag{4.1.42}$$

where η_o=the overall surface efficiency of an extended (fin) surface and is related to the fin efficiency η_f, fin surface area A_f, and the total surface area A as follows:

$$\eta_o=1-\frac{A_f}{A}\left(1-\eta_f\right) \tag{4.1.43}$$

The wall thermal resistance R_w of Equation 4.1.41 is given by

$$R_w=\begin{cases}\delta/A_w k_w & \text{for a flat wall}\\ \dfrac{\ln\left(d_o/d_j\right)}{2\pi k_w L N_t} & \text{for a circular tube with a single-layer wall}\\ \dfrac{1}{2\pi L N_t}\left[\displaystyle\sum_j \frac{\ln\left(d_{j+1}/d_j\right)}{k_{wj}}\right] & \text{for a circular tube with a multiple-layer wall}\end{cases} \tag{4.1.44}$$

If one of the resistances on the right-hand side of Equation 4.1.41 or 4.1.42 is significantly higher than the other resistances, it is referred to as the *controlling thermal resistance*. A reduction in the controlling thermal resistance will have much more impact in reducing the exchanger surface area (*A*) requirement compared with the reduction in *A* as a result of the reduction in other thermal resistances.

The overall conductance *UA* of Equation 4.1.42 may be defined in terms of hot or cold fluid side surface area or wall conduction area as

$$UA=U_h A_h=U_c A_c=U_w A_w \tag{4.1.45}$$

When R_w is negligible, $T_{w,h}=T_{w,c}=T_w$ of Figure 4.1.33 is computed from

$$T_w=\frac{T_h+\left[\left(R_h+R_{s,h}\right)/\left(R_c+R_{s,c}\right)\right]T_c}{1+\left[\left(R_h+R_{s,h}\right)/\left(R_c+R_{s,c}\right)\right]} \tag{4.1.46}$$

When $R_{s,h}=R_{s,c}=0$, Equation 4.1.46 reduces to

$$T_w=\frac{T_h/R_h+T_c/R_c}{1/R_h+1/R_c}=\frac{\left(\eta_o hA\right)_h T_h+\left(\eta_o hA\right)_c T_c}{\left(\eta_o hA\right)_h+\left(\eta_o hA\right)_c} \tag{4.1.47}$$

ε-*NTU, P-NTU, and MTD methods*. If we consider the fluid outlet temperatures or heat transfer rate as dependent variables, they are related to independent variable/parameters of Figure 4.1.32 as follows.

$$T_{h,o}, T_{c,o}, \text{ or } q=f\left\{T_{h,j}, T_{c,i}, C_c, C_h, U, A, \text{ flow arrangement}\right\} \tag{4.1.48}$$

Six independent and three dependent variables of Equation 4.1.48 for a given flow arrangement can be transferred into two independent and one dependent dimensionless groups; three different methods are presented in Table 4.1.8 based on the choice of three dimensionless groups. The relationship among

TABLE 4.1.8
General Functional Relationships and Dimensionless Groups for ε-NTU, P-NTU, and MTD Methods

ε-NTU Method	P-NTU Method[a]	MTD Method[a]
$q=\varepsilon C_{\min}(T_{\mathrm{h,i}}-T_{\mathrm{c,i}})$	$q=P_1C_1\lvert T_{1,\mathrm{i}}-T_{2,\mathrm{i}}\rvert$	$q=UAF\Delta T_{\mathrm{lm}}$
$\varepsilon=\phi$ (NTU, C^*, flow arrangement)	$P_1=\phi$ (NTU1, R_1, flow arrangement)	$F=\phi$ (P, R, flow arrangement)[b]
$\varepsilon=\dfrac{C_{\mathrm{h}}\left(T_{\mathrm{h,i}}-T_{\mathrm{h,o}}\right)}{C_{\min}\left(T_{\mathrm{h,i}}-T_{\mathrm{c,i}}\right)}=\dfrac{C_{\mathrm{c}}\left(T_{\mathrm{c,o}}-T_{\mathrm{c,i}}\right)}{C_{\min}\left(T_{\mathrm{h,i}}-T_{\mathrm{c,i}}\right)}$	$P_1=\dfrac{T_{1,\mathrm{o}}-T_{1,\mathrm{i}}}{T_{2,\mathrm{i}}-T_{1,\mathrm{i}}}$	$F=\dfrac{\Delta T_{\mathrm{m}}}{\Delta T_{\mathrm{lm}}}$
$\mathrm{NTU}=\dfrac{UA}{C_{\min}}=\dfrac{1}{C_{\min}}\displaystyle\int_A U\ \mathrm{d}A$	$\mathrm{NTU}_1=\dfrac{UA}{C_1}=\dfrac{\lvert T_{1,\mathrm{o}}-T_{1,\mathrm{i}}\rvert}{\Delta T_{\mathrm{m}}}$	$\mathrm{LMTD}=\Delta T_{\mathrm{lm}}=\dfrac{\Delta T_1-\Delta T_2}{\ln\left(\Delta T_1/\Delta T_2\right)}$
$C^*=\dfrac{C_{\min}}{C_{\max}}=\dfrac{\left(\dot{m}c_{\mathrm{p}}\right)_{\min}}{\left(\dot{m}c_{\mathrm{p}}\right)_{\max}}$	$R_1=\dfrac{C_1}{C_2}=\dfrac{T_{2,\mathrm{i}}-T_{2,\mathrm{o}}}{T_{1,\mathrm{o}}-T_{1,\mathrm{i}}}$	$\Delta T_1=T_{\mathrm{h,i}}-T_{\mathrm{c,o}};\ \Delta T_2=T_{\mathrm{h,o}}-T_{\mathrm{c,i}}$

[a] Although P, R, and NTU are defined on fluid side 1, it must be emphasized that all the results of the P-NTU and MTD methods are valid if the definitions of P, NTU, and R are consistently based on C_c, C_s, C_h, or C.

[b] P and R are defined in the P-NTU method.

three dimensionless groups is derived by integrating Equations 4.1.37 and 4.1.38 across the surface area for a specified exchanger flow arrangement. Such expressions are presented later in Table 4.1.10 for the industrially most important flow arrangements. Now we briefly describe the three methods.

In *the ε-NTU method*, the heat transfer rate from the hot fluid to the cold fluid in the exchanger is expressed as

$$q = eC_{\min}\left(T_{\mathrm{h,j}}-T_{\mathrm{c,j}}\right) \tag{4.1.49}$$

Here, the exchanger effectiveness ε is an efficiency factor. It is a ratio of the actual heat transfer rate from the hot fluid to the cold fluid in a given heat exchanger of any flow arrangement to the maximum possible heat transfer rate $q_{\max}$ thermodynamically permitted. The $q_{\max}$ is obtained in a *counterflow* heat exchanger (recuperator) of *infinite surface area* operating with the fluid flow rates (heat capacity rates) and fluid inlet temperatures equal to those of an actual exchanger (constant fluid properties are idealized). As noted in Table 4.1.7, the exchanger effectiveness ε is a function of NTU and C^* in this method. The NTU is a ratio of the overall conductance UA to the smaller heat capacity rate $C_{\min}$. NTU designates the dimensionless "heat transfer size" or "thermal size" of the exchanger. Other interpretations of NTU are given by Shah (1983). The heat capacity rate ratio C^* is simply a ratio of the smaller to the larger heat capacity rate for the two fluid streams. Note that $0\le\varepsilon\le1$, $0\le\mathrm{NTU}\le\infty$ and $0\le C^*\le1$.

The P-NTU method represents a variant of the ε-NTU method. The ε-NTU relationship is different depending upon whether the shell fluid is the $C_{\min}$ or $C_{\max}$ fluid in the (stream unsymmetric) flow arrangements commonly used for shell-and-tube exchangers. In order to avoid possible errors and to avoid keeping track of the $C_{\min}$ fluid side, an alternative is to present the temperature effectiveness P as a function of NTU and R, where P, NTU, and R are defined consistently either for Fluid 1 side or Fluid 2 side; in Table 4.1.8, they are defined for Fluid 1 side (regardless of whether that side is the hot or cold fluid side), and Fluid 1 side is clearly identified for each flow arrangement in Table 4.1.10; it is the shell side in a shell-and-tube exchanger. Note that

$$q = P_1C_1\left|T_{1,\mathrm{i}}-T_{2,\mathrm{i}}\right| = P_2C_2\left|T_{2,\mathrm{i}}-T_{1,\mathrm{i}}\right| \tag{4.1.50}$$

$$P_1 = P_2 R_2;\ P_2 = P_1 R_1 \tag{4.1.51}$$

$$\mathrm{NTU}_1 = \mathrm{NTU}_2 R_2; \mathrm{NTU}_2 = \mathrm{NTU}_1 R_1 \tag{4.1.52}$$

and

$$R_1 = 1 / R_2 \tag{4.1.53}$$

In *the MTD method*, the heat transfer rate from the hot fluid to the cold fluid in the exchanger is given by

$$q = UA\Delta T_m = UAF\Delta T_{lm} \tag{4.1.54}$$

where ΔT_m the LMTD, and F the LMTD correction factor, a ratio of true (actual) MTD to the LMTD, where

$$\mathrm{LMTD} = \Delta T_{lm} = \frac{\Delta T_1 - \Delta T_2}{\ln\left(\Delta T_1 / \Delta T_2\right)} \tag{4.1.55}$$

Here, ΔT_1 and ΔT_2 are defined as

$$\Delta T_1 = T_{h,i} - T_{c,o} \text{ and } \Delta T_2 = T_{h,o} - T_{c,i} \quad \text{for all flow arrangements except for parallel flow} \tag{4.1.56}$$

$$\Delta T_1 = T_{h,i} - T_{c,j} \quad \Delta T_2 = T_{h,o} - T_{c,o} \quad \text{for parallel flow} \tag{4.1.57}$$

The LMTD represents a true MTD for a counterflow arrangement under the idealizations listed later. Thus, the LMTD correction factor F represents a degree of departure for the MTD from the counterflow LMTD; it does not represent the effectiveness of a heat exchanger. It depends on two dimensionless group P_1 and R_1 or P_2 and R_2 for a given flow arrangement.

The relationship among the dimensionless groups of the ε-NTU, P-NTU, and MTD methods are presented in Table 4.1.9. The closed-form formulas for industrially important exchangers are presented in terms of P_1, NTU_1, and R_1 in Table 4.1.10. These formulas are valid under idealizations that include (1) steady-state conditions; (2) negligible heat losses to the surrounding; (3) no phase changes in the fluid streams flowing through the exchanger, or phase changes (condensation or boiling) occurring at constant temperature and constant effective specific heat; (4) uniform velocity and temperature at the entrance of the heat exchanger on each fluid side; (5) the overall extended surface efficiency η_o is uniform and constant; (6) constant individual and overall heat transfer coefficients; (7) uniformly distributed heat transfer area on each fluid side; (7) the number of baffles as large in shell-and-tube exchangers; (8) no flow maldistribution; and (9) negligible longitudinal heat conduction in the fluid and exchanger wall.

The overall heat transfer coefficient can vary as a result of variations in local heat transfer coefficients due to two effects: (1) change in heat transfer coefficients in the exchanger as a result of changes in the fluid properties or radiation due to rise or drop of fluid temperatures and (2) change in heat transfer coefficients in the exchanger due to developing thermal boundary layers; it is referred to as the *length effect*. The first effect due to fluid property variations (or radiation) consists of two components: (1) distortion of velocity and temperature profiles at a given flow cross section due to fluid property variations—this effect is usually taken into account by the so-called property ratio method, with the correction scheme of Equations 4.1.91 and 4.1.92—and (2) variations in the fluid temperature along the axial and transverse directions in the exchanger depending upon the exchanger flow arrangement—this effect is referred to as the *temperature effect*. The resultant axial

TABLE 4.1.9
Relationships between Dimensionless Groups of the P-NTU and LMTD Methods and Those of the ε-NTU Method

$$P_1 = \frac{C_{\min}}{C_1}\varepsilon = \begin{cases} \varepsilon & \text{for} \quad C_1 = C_{\min} \\ \varepsilon C^* & \text{for} \quad C_1 = C_{\max} \end{cases}$$

$$R_1 = \frac{C_1}{C_2} = \begin{cases} C^* & \text{for} \quad C_1 = C_{\min} \\ 1/C^* & \text{for} \quad C_1 = C_{\max} \end{cases}$$

$$\mathrm{NTU}_1 = \mathrm{NTU}\frac{C_{\min}}{C_1} = \begin{cases} \mathrm{NTU} & \text{for} \quad C_1 = C_{\min} \\ \mathrm{NTU}\, C^* & \text{for} \quad C_1 = C_{\max} \end{cases}$$

$$F = \frac{\mathrm{NTU}_{\mathrm{cf}}}{\mathrm{NTU}} = \frac{1}{\mathrm{NTU}\left(1 - C^*\right)} \ln\left[\frac{1 - C^*\varepsilon}{1 - \varepsilon}\right] \xrightarrow[C^*=1]{} \frac{\varepsilon}{\mathrm{NTU}\left(1 - \varepsilon\right)}$$

$$F = \frac{1}{\mathrm{NTU}_1\left(1 - R_1\right)} \ln\left[\frac{1 - RP_1}{1 - P_1}\right] \xrightarrow[R_1=1]{} \frac{P_1}{\mathrm{NTU}_1\left(1 - P_1\right)}$$

changes in the overall mean heat transfer coefficient can be significant; the variations in U_{local} could be nonlinear, dependent upon the type of the fluid. The effect of varying U_{local} can be taken into account by evaluating U_{local} at a few points in the exchanger and subsequently integrating U_{local} values by the Simpson or Gauss method (Shah, 1993). The temperature effect can increase or decrease mean U slightly or significantly, depending upon the fluids and applications. The length effect is important for developing laminar flows for which high heat transfer coefficients are obtained in the thermal entrance region. However, in general, it will have less impact on the overall heat transfer coefficient because the other thermal resistances in series in an exchanger may be controlling. The length effect reduces the overall heat transfer coefficient compared with the mean value calculated conventionally (assuming uniform mean heat transfer coefficient on each fluid side). It is shown that this reduction is up to about 11% for the worst case (Shah, 1993).

Shah and Pignotti (1997) have shown that the following are the specific number of baffles beyond which the influence of the finite number of baffles on the exchanger effectiveness is not significantly larger than 2%: $N_b \geq 10$ for 1-1 TEMA E counterflow exchanger; $N_b \geq 6$ for 1-2 TEMA E exchanger for $\mathrm{NTU}_s \leq 2$, $R_s \leq 5$; $N_b \geq 9$ for 1-2 TEMA J exchanger for $\mathrm{NTU}_s \leq 2$, $R_s \leq 5$; $N_b \geq 5$ for 1-2 TEMA G exchanger for $\mathrm{NTU}_s \leq 3$, all R_s; $N_b \geq 11$ for 1-2 TEMA H exchanger for $\mathrm{NTU}_s \leq 3$, all R_s. Various shell-and-tube heat exchangers (such as TEMA E, G, H, J, etc.) are classified by the Tubular Exchanger Manufacturers' Association (TEMA, 1988).

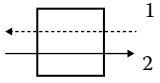
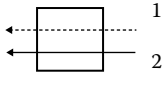
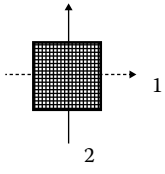
If any of the basic idealizations are not valid for a particular exchanger application, the best solution is to work directly with either Equations 4.1.37 and 4.1.38 or their modified form by including a particular effect, and to integrate them over a small exchanger segment numerically in which all of the idealizations are valid.

Fin Efficiency and Extended Surface Efficiency

Extended surfaces have fins attached to the primary surface on one or both sides of a two-fluid or a multifluid heat exchanger. Fins can be of a variety of geometries—plain, wavy, or interrupted—and can be attached to the inside, outside, or both sides of circular, flat, or oval tubes, or parting sheets. Fins are primarily used to increase the surface area (when the heat transfer coefficient on that fluid side is relatively low) and consequently to increase the total rate of heat transfer. In addition, enhanced fin geometries also increase the heat transfer coefficient compared to that for a plain fin.

TABLE 4.1.10

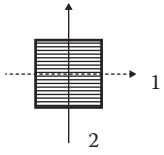
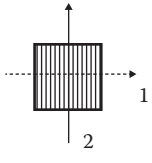
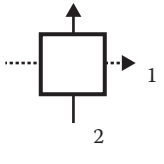
P_1—NTU_1 Formulas and Limiting Values P_1 and $R_1 = 1$ and $NTU_1 \to \infty$ for Various Exchanger Flow Arrangements[a]

Flow Arrangement	Eq. no.	General formula	Value for $R_1 = 1$	Value for $NTU_1 \to \infty$
	I.1.1	$P_1 = \dfrac{1-\exp[-\mathrm{NTU}_1(1-R_1)]}{1-R_1\exp[-\mathrm{NTU}_1(1-R_1)]}$	$P_1 = \dfrac{\mathrm{NTU}_1}{1+\mathrm{NTU}_1}$	$P_1 \to$ for $R_1 \le 1$ $P_1 \to 1/R_1$ for $R_1 \ge 1$
	I.1.2	$\mathrm{NTU}_1 = \dfrac{1}{(1-R_1)}\ln\left[\dfrac{1-R_1P_1}{1-P_1}\right]$	$\mathrm{NTU}_1 = \dfrac{P_1}{1-P_1}$	$\mathrm{NTU}_1 \to \infty$
Counterflow exchanger, stream symmetric	I.1.3	$F = 1$	$F = 1$	$F = 1$
	I.2.1	$P_1 = \dfrac{1-\exp\left[-\mathrm{NTU}_1(1+R_1)\right]}{1+R_1}$	$P_1 = \dfrac{1}{2}\left[1-\exp(-2\mathrm{NTU}_1)\right]$	$P_1 \to \dfrac{1}{1+R_1}$
	I.2.2	$\mathrm{NTU}_1 = \dfrac{1}{1+R_1}\ln\left[\dfrac{1}{1-P_1(1+R_1)}\right]$	$\mathrm{NTU}_1 = \dfrac{1}{2}\ln\left[\dfrac{1}{1-2P_1}\right]$	$\mathrm{NTU}_1 \to \infty$
Parallel flow exchanger, stream symmetric	I.2.3	$F = \dfrac{(R_1+1)\ln\left[\dfrac{1-R_1P_1}{1-P_1}\right]}{(R_1-1)\ln[1-P_1(1+R_1)]}$	$F = \dfrac{2P_1}{(P_1-1)\ln(1-2P_1)}$	$\mathrm{F} \to 0$
	II.1	$P_1 = 1-\exp(\mathrm{NTU}_1) - \exp[-(1+R_1)\mathrm{NTU}_1]\times \sum_{n=1}^{\infty} R_1^n P_n(\mathrm{NTU}_1)$	Same as Equation (II. 1) with $R_1 = 1$	$P_1 \to 1$ for $R_1 \le 1$ $P_1 \to \dfrac{1}{R_1}$ for $R_1 \ge 1$

(*Continued*)

TABLE 4.1.10 (*Continued*)

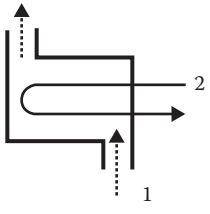
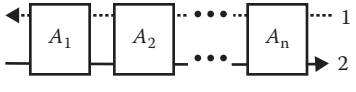
P_1—NTU_1 Formulas and Limiting Values P_1 and $R_1 = 1$ and $NTU_1 \to \infty$ for Various Exchanger Flow Arrangements[a]

Flow Arrangement	Eq. no.	General formula	Value for $R_1 = 1$	Value for $NTU_1 \to \infty$
Single-pass crossflow exchanger, both fluids unmixed, stream symmetric		$P_n(y) = \frac{1}{(n+1)!}\sum_{j=1}^{n}\frac{(n+1-j)}{j!}y^{n+j}$		
	II.2.1	$P_1 = [1-\exp(-KR_1)]/R_1$ $K = 1-\exp(-NTU_1)$	$P_1 = 1-\exp(-K)$	$P_1 \to \frac{1-\exp(-R_1)}{R_1}$
	II.2.2	$NTU = \ln\left[\frac{1}{1+\frac{1}{R_1}\ln(1-R_jP_1)}\right]$	$NTU_1 = \ln\left[\frac{1}{1+\ln(1-P_1)}\right]$	$NTU_1 \to \infty$
Single-pass crossflow exchanger, fluid 1 unmixed, fluid 2 mixed	II.2.3	$F = \frac{\ln[(1-R_1P_1)/(1-P_1)]}{(R_1-1)\ln\left[1+\frac{1}{R_1}\ln(1-R_1P_1)\right]}$	$F = \frac{P_1}{(P_1-1)\ln[1+\ln(1-P_1)]}$	$F \to 0$
	II.3.1	$P = 1-\exp(-K/R_1)$ $K = 1-\exp(-R_1\,NTU_1)$	$P = 1-\exp(-K)$ $K = 1-\exp(-NTU_1)$	$P_1 \to 1-\exp(-1/R_1)$
	II.3.2	$NTU_1 = \frac{1}{R_1}\ln\left[\frac{1}{1+R_1}\ln(1-P_1)\right]$	$NTU_1 = \ln\left[\frac{1}{1+\ln(1-P_1)}\right]$	$NTU_1 \to \infty$
Single-pass crossflow exchanger, fluid 1 mixed, fluid 2 unmixed	II.3.3	$F = \frac{\ln(1-R_1P_1)/(1-P_1)]}{(1-1/R_1)\ln[1+R_1\ln(1-P_1)]}$	$F = \frac{P_1}{(P_1-1)\ln[1+\ln(1-P_1)]}$	$P_t \to \frac{1}{1+R_1}$
	II.4	$P_t = \left[\frac{1}{K_1}+\frac{R_1}{K_2}-\frac{1}{NTU_t}\right]^{-1}$	$P_t = \left[\frac{2}{K_1}-\frac{1}{NTU_t}\right]^{-1}$	$P_1 \to \frac{1}{1+R_t}$

(*Continued*)

TABLE 4.1.10 (*Continued*)

P_1—NTU_1 Formulas and Limiting Values P_1 and $R_1 = 1$ and $NTU_1 \to \infty$ for Various Exchanger Flow Arrangements[a]

Flow Arrangement	Eq. no.	General formula	Value for $R_1 = 1$	Value for $NTU_1 \to \infty$
Single-pass crossflow exchanger, both fluids mixed, stream symmetric 		$K_1 = 1 - \exp(-\mathrm{NTU}_1)$ $K_2 = 1 - \exp(-R_1\,\mathrm{NTU}_1)$		
	III.1.1	$P_1 = \dfrac{2}{1 + R_1 + E\coth(E\,\mathrm{NTU}_1/2)}$ $E = \left[1 + R_1^2\right]^{1/2}$	$P_1 = \dfrac{1}{1 + \coth(\mathrm{NTU}_1/\sqrt{2})/\sqrt{2}}$	$P_1 \to \dfrac{2}{1 + R_1 + E}$
	III. 1.2	$\mathrm{NTU}_1 = \dfrac{1}{E}\ln\left[\dfrac{2 - P_1(1 + R_1 - E)}{2 - P_1(1 + R_1 + E)}\right]$	$\mathrm{NTU}_1 = \ln\left[\dfrac{2 - P_1}{2 - 3P_1}\right]$	$\mathrm{NTU}_1 \to \infty$
1-2 TEMA E shell-and-tube exchanger, shell fluid mixed, stream symmetric	III.1.3	$F = \dfrac{E\ln[(1 - R_1P_1)/(1 - P_1)]}{(1 - R_1)\ln\left[\dfrac{2 - P_1(1 + R_1 - E)}{2 - P_1(1 + R_1 + E)}\right]}$	$F = \dfrac{P_1/(1 - P_1)}{\ln(2 - P_1)/(2 - 3P_1)]}$	$F \to 0$
	IV.1.1	$P_1 = \dfrac{\prod_{i=1}^{n}(1 - R_1P_{1,A_i}) - \prod_{i=1}^{n}(1 - P_{1,A_i})}{\prod_{i=1}^{n}(1 - R_1P_{1,A_i}) - R_1\prod_{i=1}^{n}(1 - P_{1,A_i})}$	$P_1 = \dfrac{\sum_{i=1}^{n}\frac{P_{1,A_i}}{1 - P_{1,A_i}}}{1 + \sum_{i=1}^{n}\frac{P_{1,A_i}}{1 - P_{1,A_i}}}$	Same as Equation (I.1.1) counterflow
	IV.1.2	$R_1 = R_{1,A_i},\ i = 1,\ldots,n$	$1 = R_{1,A_i},\ i = 1,\ldots,n$	Same as Equation (IV.1.2)
Series coupling of n exchangers, overall counterflow arrangement. Stream symmetric if all A_1 are stream symmetric	IV.1.3	$\mathrm{NTU}_1 = \sum_{i=1}^{n}\mathrm{NTU}_{1,A_i}$	Same as Equation (IV.1.3)	Same as Equation (IV.1.3)
	IV.1.4	$F = \dfrac{1}{\mathrm{NTU}_1}\sum_{i=1}^{n}\mathrm{NTU}_{1,A_i}F_{A_i}$	Same as Equation (IV.1.4)	Same as Equation (IV.1.4)

Source: Shah, R.K. and Mueller, A.C., Heat exchange, *Ullmann's Encyclopedia of Industrial Chemistry*, *Unit Operations II*, vol. B3, Chap. 2, 108 p, VCH, Weinheim, Germany, 1988. With permission.

[a] In this table, all variables, except P_1, R_1, NTU, and F, are local or dummy variables not necessarily related to similar ones defined in the nomenclature and the text.

Fins may also be used on the high heat transfer coefficient fluid side in a heat exchanger primarily for structural strength purposes (e.g., for high-pressure water flow through a flat tube) or to provide a thorough mixing of a highly viscous liquid (such as for laminar oil flow in a flat or a round tube). Fins are attached to the primary surface by brazing, soldering, welding, adhesive bonding, or mechanical expansion, or they are extruded or integrally connected to the tubes. Major categories of extended surface heat exchangers are plate-fin (Figures 4.1.23 through 4.1.25) and tube-fin (Figures 4.1.26 through 4.1.28) exchangers. Note that shell-and-tube exchangers sometimes employ individually finned tubes—low-finned tubes (similar to Figure 4.1.26a but with low-height fins) (Shah, 1985).

The concept of fin efficiency accounts for the reduction in temperature potential between the fin and the ambient fluid due to conduction along the fin and convection from or to the fin surface depending upon the fin cooling or heating situation. The fin efficiency is defined as the ratio of the actual heat transfer rate through the fin base divided by the maximum possible heat transfer rate through the fin base which would be obtained if the entire fin were at the base temperature (i.e., its material thermal conductivity were infinite). Since most of the real fins are "thin," they are treated as one-dimensional (1-D) with standard idealizations used for the analysis (Huang and Shah, 1992). This 1-D fin efficiency is a function of the fin geometry, fin material thermal conductivity, heat transfer coefficient at the fin surface, and the fin tip boundary condition; it is not a function of the fin base or fin tip temperature, ambient temperature, and heat flux at the fin base or fin tip in general. Fin efficiency formulas for some common fins are presented in Table 4.1.11 (Shah, 1985). Huang and Shah (1992) also discuss the influence on η_f if any of the basic idealizations used in the fin analysis are violated.

The fin efficiency for flat fins (Figure 4.1.26b) is obtained by a sector method (Shah, 1985). In this method, the rectangular or hexagonal fin around the tube (Figures 4.1.28a, b) or its smallest symmetrical section is divided into n sectors (Figure 4.1.34). Each sector is then considered as a circular fin with the radius $r_{e,i}$ equal to the length of the centerline of the sector. The fin efficiency of each sector is subsequently computed using the circular fin formula of Table 4.1.11. The fin efficiency η_f for the whole fin is then the surface area weighted average of $\eta_{f,i}$ of each sector.

$$\eta_f = \frac{\sum_{i=1}^{n} \eta_{f,i} A_{f,i}}{\sum_{i=1}^{n} A_{f,i}} \tag{4.1.58}$$

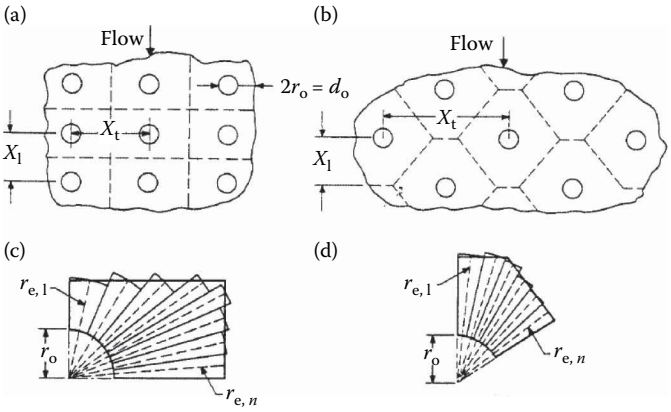
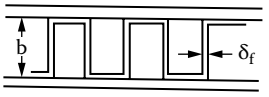


FIGURE 4.1.34 A flat fin over (a) an in-line and (b) staggered tube arrangement; the smallest representative segment of the fin for (c) an in-line and (d) a staggered tube arrangement.

TABLE 4.1.11
Fin Efficiency Expressions for Plat-Fin and Tube-Fin Geometries of Uniform Fin Thickness

Geometry | **Fin Efficiency Formula**

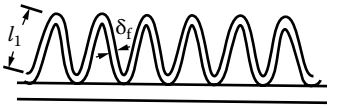


Plain, wavy, or offset strip fin of rectangular cross section

$$m_i = \left[\frac{2h}{k_j\delta_i}\left(1+\frac{\delta_t}{l_f}\right)\right]^{1/2};\ E_1 = \frac{\tanh(m_i l_i)}{m_i t_1}\ i = 1,2$$

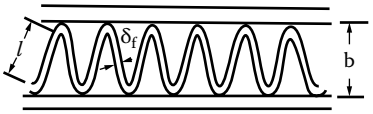
$$\eta_f = E_1$$

$$l_1 = \frac{b}{2} - \delta_1;\ \delta_1 = \delta_f$$



Triangular fin heated from one side

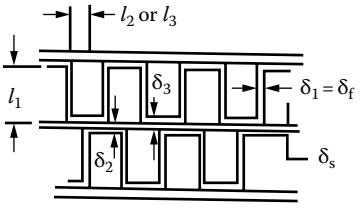
$$\eta_f = \frac{hA_1(T_0 - T_a)\dfrac{\sinh(m_1 l_1)}{m_1 l_1} + q_e}{\cosh(m_1 l_1)\left[hA_1(T_0 - T_a) + q_e\dfrac{T_0 - T_a}{T_1 - T_a}\right]}$$



Plain, wavy, or louver fin of triangular cross section

$$\eta_f = E_1$$

$$l_1 = \frac{l}{2};\ \delta_1 = \delta_f$$

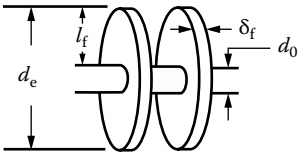


Double sandwich fin

$$\eta_f = \frac{E_1 l_1 + E_2 t_2}{l_1 + l_2}\frac{1}{1 + m_1^2 E_1 E_2 l_1 l_2}$$

$$\delta_1 = \delta_f;\ \delta_2 = \delta_3 = \delta_f + \delta_s$$

$$l_1 = b - \delta_f + \frac{\delta_s}{2};\ l_2 = l_3 = \frac{p_f}{2}$$



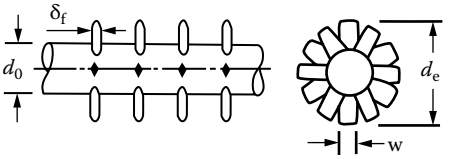
Circular fin

$$\eta_j = \begin{cases} a(ml_e)^{-b} & \text{for } \Phi > 0.6 + 2.257(r^*)^{-0.445} \\ \dfrac{\tanh\Phi}{\Phi} & \text{for } \Phi \le 0.6 + 2.257(r^*)^{-0.445} \end{cases}$$

$$a = (r^*)^{-0.246};\ \Phi = ml_e(r^*)^{\exp(0.13ml_e - 1.3863)}$$

$$b = \begin{cases} 0.9107 + 0.0893r^* & \text{for } r^* \le 2 \\ 0.9706 + 0.17125\ln r^* & \text{for } r^* > 2 \end{cases}$$

$$m = \left(\frac{2h}{k_f\delta_f}\right)^{1/2};\ l_e = l_j + \frac{\delta_f}{2};\ r^* = \frac{d_e}{d_o}$$



Studded fin

$$\eta_j = \frac{\tanh(ml_e)}{ml_e}$$

$$m = \left[\frac{2h}{k_f\delta_f}\left(1+\frac{\delta_f}{w}\right)\right]^{1/2};\ l_e = l_f + \frac{\delta_f}{2};\ l_f = \frac{(d_e - d_o)}{2}$$

Since the heat flow seeks the path of least thermal resistance, actual η_f will be equal to or higher than that calculated by Equation 4.1.58; hence, Equation 4.1.58 yields a somewhat conservative value of η_f.

The η_f values of Table 4.1.11 or Equation 4.1.58 are not valid in general when the fin is thick, when it is subject to variable heat transfer coefficients or variable ambient fluid temperature, or when it has a temperature depression at the fin base. See Huang and Shah (1992) for details. For a thin rectangular fin of constant cross section, the fin efficiency as presented in Table 4.1.11 is given by

$$\eta_f = \frac{\tanh(ml)}{ml} \tag{4.1.59}$$

where $m = \left[2h\left(1+\delta_f l_f\right)/k_f\delta_f\right]^{1/2}$. For a thick rectangular fin of constant cross section, the fin efficiency (a counterpart of Equation 4.1.59) is given by (Huang and Shah, 1992)

$$\eta_f = \frac{\left(\mathrm{Bi}^+\right)^{1/2}}{K\mathrm{Bi}} \tanh\left[K\left(\mathrm{Bi}^+\right)^{1/2}\right] \tag{4.1.60}$$

where $\mathrm{Bi}^+ = \mathrm{Bi}/(1+\mathrm{Bi}/4)$, $\mathrm{Bi} = (h\delta_f/2k_f)^{1/2}$, $K = 2l/\delta_f$. Equation 4.1.59 is accurate (within 0.3%) for a "thick" rectangular fin having $\eta_f > 80\%$; otherwise, use Equation 4.1.60 for a thick fin.

In an extended-surface heat exchanger, heat transfer takes place from both the fins ($\eta_f < 100\%$) and the primary surface ($\eta_f = 100\%$). In that case, the total heat transfer rate is evaluated through a concept of extended surface efficiency η_o defined as

$$\eta_o = \frac{A_p}{A} + \eta_f \frac{A_f}{A} = 1 - \frac{A_f}{A}\left(1-\eta_f\right) \tag{4.1.61}$$

where A_f is the fin surface area, A_p is the primary surface area, and $A = A_f + A_p$. In Equation 4.1.61, heat transfer coefficients over the finned and unfinned surfaces are idealized to be equal. Note that $\eta_o \geq \eta_f$ and η_o is always required for the determination of thermal resistances of Equation 4.1.41 in heat exchanger analysis.

Pressure Drop Analysis

Usually a fan, blower, or pump is used to flow fluid through individual sides of a heat exchanger. Due to potential initial and operating high cost, low fluid pumping power requirement is highly desired for gases and viscous liquids. The fluid pumping power $\wp$ is approximately related to the core pressure drop in the exchanger as (Shah, 1985)

$$\wp = \frac{\dot{m}\Delta p}{\rho} \approx \begin{cases} \dfrac{1}{2g_c}\dfrac{\mu}{\rho^2}\dfrac{4L}{D_h} f\mathrm{Re} & \text{for laminar flow} \quad (4.1.62) \\ \dfrac{0.046}{2g_c}\dfrac{\mu^{0.2}}{\rho^2}\dfrac{4L}{D_h}\dfrac{\dot{m}^{2.8}}{A_0^{1.8}D_h^{0.2}} & \text{for turbulent flow} \quad (4.1.63) \end{cases}$$

It is clear from Equations 4.1.62 and 4.1.63 that the fluid pumping power is strongly dependent upon the fluid density ($\wp \propto 1/\rho^2$) particularly for low-density fluids in laminar and turbulent flows, and upon the viscosity in laminar flow. In addition, the pressure drop itself can be an important consideration when blowers and pumps are used for the fluid flow since they are head limited. Also for condensing and evaporating fluids, the pressure drop affects the heat transfer rate. Hence, the pressure drop determination in the exchanger is important.

The pressure drop associated with a heat exchanger consists of (1) core pressure drop and (2) the pressure drop associated with the fluid distribution devices such as inlet and outlet manifolds, headers, tanks, nozzles, ducting, and so on, which may include bends, valves, and fittings. This second Δp component is determined from Idelchik (1994) and Miller (1990). The core pressure drop may consist of one or more of the following components depending upon the exchanger construction: (1) friction losses associated with fluid flow over heat transfer surface (this usually consists of skin friction, form (profile) drag, and internal contractions and expansions, if any); (2) the momentum effect (pressure drop or rise due to fluid density changes) in the core; (3) pressure drop associated with sudden contraction and expansion at the core inlet and outlet; and (4) the gravity effect due to the change in elevation between the inlet and outlet of the exchanger. The gravity effect is generally negligible for gases. For vertical flow through the exchanger, the pressure drop or rise ("static head") due to the elevation change is given by

$$\Delta p = \pm \frac{\rho_m g L}{g_c} \tag{4.1.64}$$

Here, the "+" sign denotes vertical upflow (i.e., pressure drop), the "–" sign denotes vertical downflow (i.e., pressure rise). The first three components of the core pressure drop are now presented for plate-fin, tube-fin, plate, and regenerative heat exchangers.

Plate-fin heat exchangers. For the plate-fin exchanger (Figure 4.1.23), all three components are considered in the core pressure drop evaluation as follows:

$$\frac{\Delta p}{p_i} = \frac{G^2}{2g_c}\frac{1}{p_i\rho_i}\left[\left(1-\sigma^2+K_c\right)+f\frac{L}{r_h}\rho_i\left(\frac{1}{\rho}\right)_m+2\left(\frac{\rho_i}{\rho_o}-1\right)-\left(1-\sigma^2-K_e\right)\frac{\rho_i}{\rho_o}\right] \tag{4.1.65}$$

where f is the Fanning friction factor, K_c and K_e are flow contraction (entrance) and expansion (exit) pressure loss coefficients (see Figure 4.1.35), and σ is a ratio of minimum free flow area to frontal area. K_c and K_e for four different entrance flow passage geometries are presented by Kays and London (1998). The entrance and exit losses are important at low values of σ and L (short cores), high values of Re, and for gases, they are negligible for liquids. The values of K_c and K_e apply to long tubes for which flow is fully developed at the exit. For partially developed flows, K_c and K_e are higher than that for fully developed flows. For interrupted surfaces, flow is never a fully developed boundary layer type. For highly interrupted fin geometries, the entrance and exit losses are generally small compared to the core pressure drop and the flow is well mixed; hence, K_c and K_e for Re $\rightarrow \infty$ should represent a good approximation. The mean specific volume v_m or $(1/\rho)_m$ in Equation 4.1.65 is given as follows. For liquids with any flow arrangement, or for a perfect gas with $C^*=1$ and any flow arrangement (except for parallel flow),

$$\left(\frac{1}{\rho}\right)_m = v_m = \frac{v_i+v_o}{2} = \frac{1}{2}\left(\frac{1}{\rho_i}+\frac{1}{\rho_o}\right) \tag{4.1.66}$$

where ν is the specific volume in m^3/kg. For a perfect gas with $C^*=0$ and any flow arrangement,

$$\left(\frac{1}{\rho}\right)_m = \frac{\tilde{R}}{p_{ave}}T_{lm} \tag{4.1.67}$$

Here $\tilde{R}$ is the gas constant in J/(kg K), $p_{ave}=(p_i+p_o)/2$, and $T_{lm}=T_{const}+\Delta T_{lm}$, where T_{const} is the mean average temperature of the fluid on the other side of the exchanger; the LMTD ΔT_{lm} is defined in Table 4.1.8. The core frictional pressure drop in Equation 4.1.65 may be approximated as

$$\Delta p \approx \frac{4fLG^2}{2g_cD_h}\left(\frac{1}{\rho}\right)_m \tag{4.1.68}$$

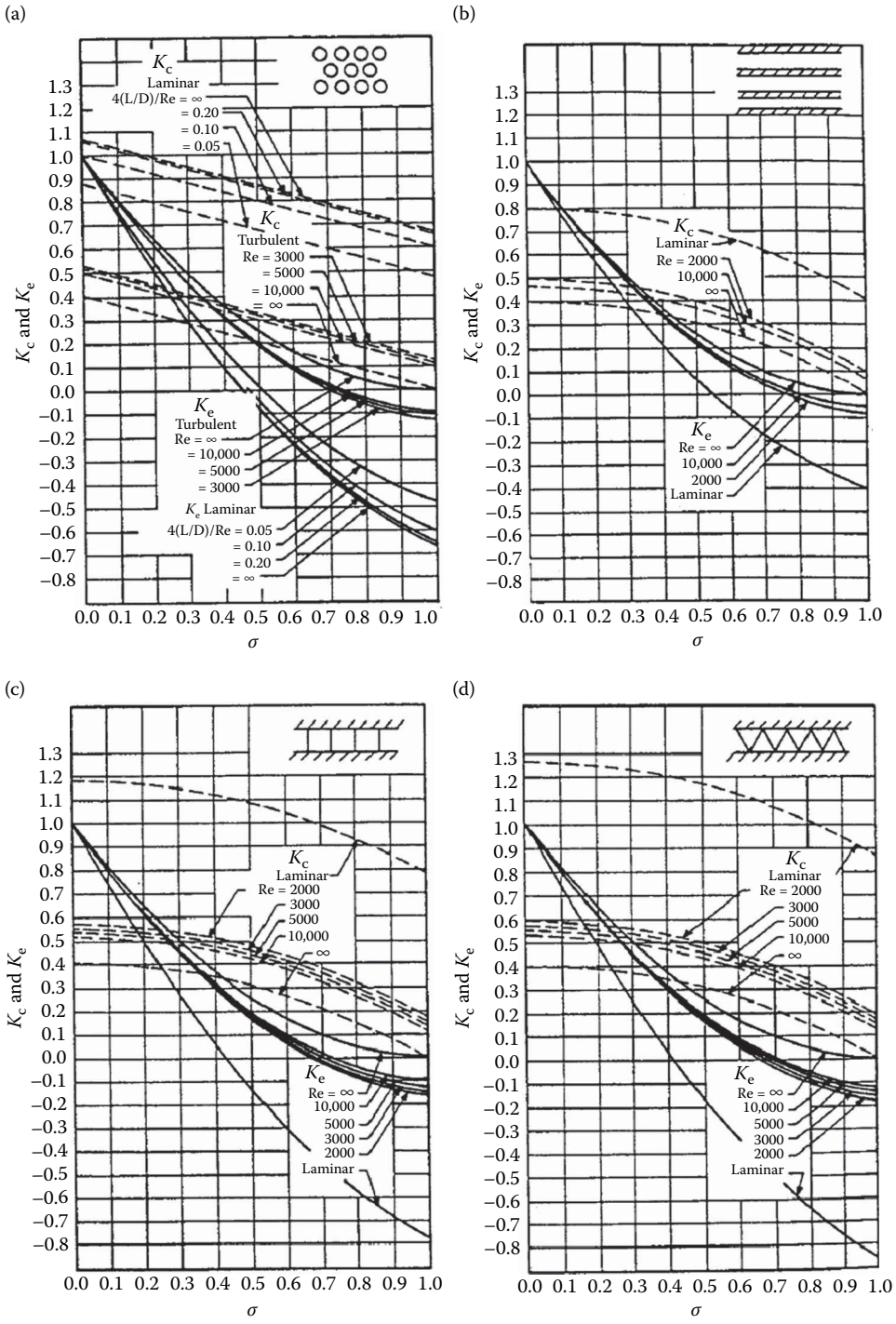


FIGURE 4.1.35 Entrance and exit pressure loss coefficients: (a) circular tubes, (b) parallel plates, (c) square passages, and (d) triangular passages. (From Kays, W.M. and London, A.L., *Compact Heat Exchangers*, 3rd ed., McGraw-Hill, New York, 1998. With permission.) For each of these flow passages, shown in the inset, the fluid flows perpendicular to the plane of the paper into the flow passages.

Tube-fin heat exchangers. The pressure drop inside a circular tube is computed using Equation 4.1.65 with proper values of f factors, and K_c and K_e for circular tubes using the method outlined in section 2.4.

For flat fins on an array of tubes (see Figure 4.1.26b), the components of the core pressure drop (such as those in Equation 4.1.65) are the same with the following exception: The core friction and momentum effect take place within the core with $G = \dot{m} / A_0$, where A_0 is the minimum free-flow area within the core, and the entrance and exit losses occur at the leading and trailing edges of the core with the associated flow area A_0' such that

$$\dot{m} = GA_o = G'A_o' \quad \text{or} \quad G'\sigma' = G\sigma \tag{4.1.69}$$

where σ' is the ratio of free-flow area to frontal area at the fin leading edges. The pressure drop for flow normal to a tube bank with flat fins is then given by

$$\frac{\Delta p}{p_i} = \frac{G^2}{2g_c}\frac{1}{p_i\rho_i}\left[f\frac{L}{r_h}\rho_i\left(\frac{1}{\rho}\right)_m + 2\left(\frac{\rho_i}{\rho_o}-1\right)\right] + \frac{G'^2}{2g_c}\frac{1}{p_i\rho_i}\left[\left(1-\sigma'^2-K_c\right)-\left(1-\sigma'^2-K_e\right)\frac{\rho_i}{\rho_o}\right] \tag{4.1.70}$$

For individually finned tubes as shown in Figure 4.1.26a, flow expansion and contraction take place along each tube row, and the magnitude is of the same order as that at the entrance and exit. Hence, the entrance and exit losses are generally lumped into the core friction factor. Equation 4.1.65 then reduces for individually finned tubes to

$$\frac{\Delta p}{p_i} = \frac{G^2}{2g_c}\frac{1}{p_i\rho_i}\left[f\frac{L}{r_h}\rho_i\left(\frac{1}{\rho}\right)_m + 2\left(\frac{\rho_i}{\rho_o}-1\right)\right] \tag{4.1.71}$$

Regenerators. For regenerator matrices having cylindrical passages, the pressure drop is computed using Equation 4.1.65 with appropriate values of f, K_c, and K_e. For regenerator matrices made up of any porous material (such as checkerwork, wire, mesh, spheres, copper wools, etc.), the pressure drop is calculated using Equation 4.1.71 in which the entrance and exit losses are included in the friction factor f.

Plate heat exchangers. Pressure drop in a PHE consists of three components: (1) pressure drop associated with the inlet and outlet manifolds and ports, (2) pressure drop within the core (plate passages), and (3) pressure drop due to the elevation change. The pressure drop in the manifolds and ports should be kept as low as possible (generally <10%, but it is found as high as 25%–30% or higher in some designs). Empirically, it is calculated as approximately 1.5 times the inlet velocity head per pass. Since the entrance and exit losses in the core (plate passages) cannot be determined experimentally, they are included in the friction factor for the given plate geometry. The pressure drop (rise) caused by the elevation change for liquids is given by Equation 4.1.64. Hence, the pressure drop on one fluid side in a PHE is given by

$$\Delta p = \frac{1.5G^2N_p}{2g_c\rho_i} + \frac{4fLG^2}{2g_cD_e}\left(\frac{1}{\rho}\right)_m + \left(\frac{1}{\rho_o}-\frac{1}{\rho_i}\right)\frac{G^2}{g_c} \pm \frac{\rho_m gL}{g_c} \tag{4.1.72}$$

where N_p is the number of passes on the given fluid side and D_e is the equivalent diameter of flow passages (usually twice the plate spacing). Note that the third term on the right-hand side of the equality sign of Equation 4.1.72 is for the momentum effect which is generally negligible in liquids.

Heat Transfer and Flow Friction Correlations

Accurate and reliable surface heat transfer and flow friction characteristics are a key input to the exchanger heat transfer and pressure drop analyses or to the rating and sizing problems (Shah,

1985). Some important analytical solutions and empirical correlations are presented next for selected exchanger geometries.

The heat transfer rate in laminar duct flow is very sensitive to the thermal boundary condition. Hence, it is essential to identify carefully the thermal boundary condition in laminar flow. The heat transfer rate in turbulent duct flow is insensitive to the thermal boundary condition for most common fluids ($Pr > 0.7$); the exception is liquid metals ($Pr < 0.03$). Hence, there is generally no need to identify the thermal boundary condition in turbulent flow for all fluids except liquid metals.

Fully developed laminar flow analytical solutions for some duct shapes of interest in compact heat exchangers are presented in Table 4.1.12 for three important thermal boundary conditions denoted by the subscripts $H1$, $H2$, and T (Shah and London, 1978; Shah and Bhatti, 1987). Here, $H1$

TABLE 4.1.12
Solutions for Heat Transfer and Friction for Fully Developed Flow-Through Specified Ducts

Geometry ($L/D_h > 100$)	Nu_{H1}	Nu_{H2}	Nu_t	fRe	J_{H1}/f[a]	Nu_{H1}/Nu_t
2b, 2a; $\frac{b}{a} = \frac{\sqrt{3}}{2}$	3.014	1.474	2.39[b]	12.630	0.269	1.26
60°, 2b, 2a; $\frac{b}{a} = \frac{\sqrt{3}}{2}$	3.111	1.892	2.47	13.333	0.263	1.26
2b, 2a; $\frac{b}{a} = 1$	3.608	3.091	2.976	14.227	0.286	1.21
	4.002	3.862	3.34[b]	15.054	0.299	1.20
2b, 2a; $\frac{b}{a} = \frac{1}{2}$	4.123	3.017	3.391	15.548	0.299	1.22
	4.364	4.364	3.657	16.000	0.307	1.19
2b, 2a; $\frac{b}{a} = \frac{1}{4}$	5.331	2.94	4.439	18.233	0.329	1.20
2b, 2a; $\frac{b}{a} = \frac{1}{8}$	6.490	2.94	5.597	20.585	0.355	1.16
$\frac{b}{a} = 0$	8.235	8.235	7.541	24.000	0.386	1.09
$\frac{b}{a} = 0$, Insulated	5.385	—	4.861	24.000	0.253	1.11

[a] This heading is the same as $Nu_{H1}\,Pr^{-1/3}/f\,Re$ with $Pr = 0.7$.

[b] Interpolated values.

denotes constant axial wall heat flux with constant peripheral wall temperature, *H*2 denotes constant axial and peripheral wall heat flux, and *T* denotes constant wall temperature. The entrance effects, flow maldistribution, free convection, property variation, fouling, and surface roughness all affect fully developed analytical solutions. In order to account for these effects in real plate-fin plain fin geometries having fully developed flows, it is best to reduce the magnitude of the analytical Nu by at least 10% and to increase the value of the analytical *f*Re by 10% for design purposes.

The initiation of transition flow, the lower limit of the critical Reynolds number ($\mathrm{Re}_{\mathrm{crit}}$), depends upon the type of entrance (e.g., smooth vs. abrupt configuration at the exchanger flow passage entrance). For a sharp square inlet configuration, $\mathrm{Re}_{\mathrm{crit}}$ is about 10%–15% lower than that for a rounded inlet configuration. For most exchangers, the entrance configuration would be sharp. Some information on $\mathrm{Re}_{\mathrm{crit}}$ is provided by Ghajar and Tam (1994).

Transition flow and fully developed turbulent flow Fanning friction factors (within ±2% accuracy) are given by Bhatti and Shah (1987) as

$$f = A + B\mathrm{Re}^{-1/m} \tag{4.1.73}$$

where

$$A = 0.0054, \quad B = 2.3\times10^{-8}, \quad m = -2/3, \quad \text{for } 2100 \le \mathrm{Re} \le 4000$$

$$A = 0.00128, \quad B = 0.1143, \quad m = 3.2154, \quad \text{for } 4000 \le \mathrm{Re} \le 10^7$$

The transition flow and fully developed turbulent flow Nusselt number correlation for a circular tube is given by Gnielinski as reported in Bhatti and Shah (1987) as

$$\mathrm{Nu} = \frac{(f/2)(\mathrm{Re}-1000)\mathrm{Pr}}{1+12.7(f/2)^{1/2}\left(\mathrm{Pr}^{2/3}-1\right)} \tag{4.1.74}$$

which is accurate within about ±10% with experimental data for $2300 \le \mathrm{Re} \le 5\times10^6$ and $0.5 \le \mathrm{Pr} \le 2000$.

A careful observation of accurate experimental friction factors for all noncircular smooth ducts reveals that ducts with laminar *f*Re < 16 have turbulent *f* factors lower than those for the circular tube, whereas ducts with laminar *f*Re > 16 have turbulent *f* factors higher than those for the circular tube (Shah and Bhatti, 1988). Similar trends are observed for the Nusselt numbers. Within ±15% accuracy, Equations 4.1.73 and 4.1.74 for *f* and Nu can be used for noncircular passages with the hydraulic diameter as the characteristic length in *f*, Nu, and Re; otherwise, refer to Bhatti and Shah (1987) for more accurate results for turbulent flow.

For hydrodynamically and thermally developing flows, the analytical solutions are boundary condition dependent (for laminar flow heat transfer only) and geometry dependent. The hydrodynamic entrance lengths for developing laminar and turbulent flows are given by Shah and Bhatti (1987) and Bhatti and Shah (1987) as

$$\frac{L_{\mathrm{hy}}}{D_{\mathrm{h}}} = \begin{cases} 0.0565\mathrm{Re} & \text{for laminar flow}\left(\mathrm{Re} \le 2100\right) \quad (4.1.75) \\ 1.359\mathrm{Re}^{1/4} & \text{for turbulent flow}\left(\mathrm{Re} \ge 10^4\right) \quad (4.1.76) \end{cases}$$

Analytical results are useful for well-defined constant cross-sectional surfaces with essentially unidirectional flows. The flows encountered in heat exchangers are generally very complex having flow separation, reattachment, recirculation, and vortices. Such flows significantly affect Nu and *f* for the specific exchanger surfaces. Since no analytical or accurate numerical solutions are available, the information is derived experimentally. Kays and London (1998) and Webb (1994) present most of

the experimental results reported in the open literature. In the following, empirical correlations for only some important surfaces are summarized due to space limitations.

Plate-Fin Extended Surfaces

Offset strip fins. This is one of the most widely used enhanced fin geometries (Figure 4.1.36) in aircraft, cryogenics, and many other industries that do not require mass production. This surface has one of the highest heat transfer performances relative to the friction factor. The most comprehensive correlations for j and f factors for the offset strip fin geometry are provided by Manglik and Bergles (1995) as follows:

$$j = 0.6522\mathrm{Re}^{-0.5403}\left(\frac{s}{h'}\right)^{-0.1541}\left(\frac{\delta_f}{l_f}\right)^{0.1499}\left(\frac{\delta_f}{s}\right)^{-0.0678} \times\left[1+5.296\times10^{-5}\mathrm{Re}^{1.340}\left(\frac{s}{h'}\right)^{0.504}\left(\frac{\delta_f}{l_f}\right)^{0.456}\left(\frac{\delta_f}{s}\right)^{-1.055}\right]^{0.1} \tag{4.1.77}$$

$$f = 9.6243\mathrm{Re}^{-0.7422}\left(\frac{s}{h'}\right)^{-0.1856}\left(\frac{\delta_f}{l_f}\right)^{0.3053}\left(\frac{\delta_f}{s}\right)^{-0.2659} \times\left[1+7.669\times10^{-8}\mathrm{Re}^{4.429}\left(\frac{s}{h'}\right)^{0.920}\left(\frac{\delta_f}{l_f}\right)^{3.767}\left(\frac{\delta_f}{s}\right)^{0.236}\right]^{0.1} \tag{4.1.78}$$

where

$$D_h = 4A_o/\left(A/l_f\right) = 4sh'l_f/\left[2\left(sl_f + h'l_f + \delta_f h'\right) + \delta_f s\right] \tag{4.1.79}$$

Geometric symbols in Equation 4.1.79 are shown in Figure 4.1.36.

These correlations predict the experimental data of 18 test cores within ±20% for $120 \le \mathrm{Re} \le 10^4$. Although all the experimental data for these correlations are obtained for air, the j factor takes into consideration minor variations in the Prandtl number, and the above correlations should be valid for $0.5 < \mathrm{Pr} < 15$.

Louver fins. Louver or multilouver fins are extensively used in the auto industry because of their mass production manufacturability and hence lower cost. The louver fin has generally higher j and f factors than those for the offset strip fin geometry, and also the increase in the friction factors is in general higher than the increase in the j factors. However, the exchanger can be designed for higher

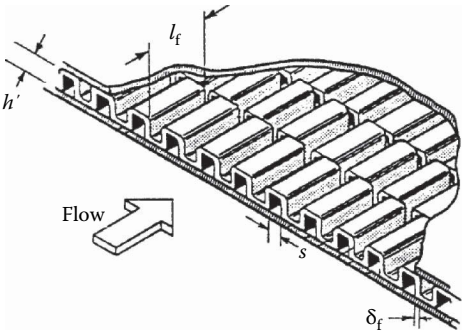


FIGURE 4.1.36 An offset strip fin geometry.

heat transfer and the same pressure drop compared to that with the offset strip fins by a proper selection of exchanger frontal area, core depth, and fin density. Published literature and correlations on the louver fins are summarized by Webb (1994) and Cowell et al. (1995), and the understanding of flow and heat transfer phenomena is summarized by Cowell et al. (1995). Because of the lack of systematic studies reported in the open literature on modern louver fin geometries, no correlation can be recommended for the design purpose.

Tube-Fin Extended Surfaces

Two major types of tube-fin extended surfaces as shown in Figure 4.1.26 are (1) individually finned tubes and (2) flat fins (also sometimes referred to as plate fins) with or without enhancements/interruptions on an array of tubes. An extensive coverage of the published literature and correlations for these extended surfaces are provided by Webb (1994), Kays and London (1998), and Rozenman (1976). Empirical correlations for some important geometries are summarized later.

Individually finned tubes. This fin geometry, helically wrapped (or extruded) circular fins on a circular tube as shown in Figure 4.1.26a, is commonly used in process and waste heat recovery industries. The following correlation for *j* factors is recommended by Briggs and Young (see Webb 1994) for individually finned tubes on staggered tube banks:

$$j = 0.134\mathrm{Re}_d^{-0.319}\left(s/l_f\right)^{0.2}\left(s/\delta_f\right)^{0.11} \tag{4.1.80}$$

where l_f is the radial height of the fin, δ_f the fin thickness, $s = p_f - \delta_f$ is the distance between adjacent fins and p_f is the fin pitch. Equation 4.1.80 is valid for the following ranges: $1100 \le \mathrm{Re}_d \le 18{,}000$, $0.13 \le s/l_f \le 0.63$, $1.01 \le s/\delta_f \le 6.62$, $0.09 \le l_f/d_o \le 0.69$, $0.011 \le \delta_f/d_o \le 0.15$, $1.54 \le X_t/d_o \le 8.23$, fin root diameter d_o between 11.1 and 40.9 mm, and fin density N_f (= $1/p_f$) between 246 and 768 fin/m. The standard deviation of Equation 4.1.80 with experimental results was 5.1%.

For friction factors, Robinson and Briggs (see Webb 1994) recommended the following correlation:

$$f_{tb} = 9.465Re_d^{-0.316}\left(X_t/d_o\right)^{-0.927}\left(X_t/X_d\right)^{0.515} \tag{4.1.81}$$

Here $X_d = \left(X_t^2 + X_l^2\right)^{1/2}$ is the diagonal pitch, and X_t and X_l are the transverse and longitudinal tube pitches, respectively. The correlation is valid for the following ranges: $2000 \le \mathrm{Re}_d \le 50{,}000$, $0.15 \le s/l_f \le 0.19$, $3.75 \le s/\delta_f \le 6.03$, $0.35 \le l_f/d_o \le 0.56$, $0.011 \le \delta_f/d_o \le 0.025$, $1.86 \le X_t/d_o \le 4.60$, $18.6 \le d_o \le 40.9$ mm, and $311 \le N_f \le 431$ fin/m. The standard deviation of Equation 4.1.81 with correlated data was 7.8%

For crossflow over low-height finned tubes, a simple but accurate correlation for heat transfer is given by Ganguli and Yilmaz (1987) as

$$j = 0.255\mathrm{Re}_d^{-0.3}\left(d_e/s\right)^{-0.3} \tag{4.1.82}$$

A more accurate correlation for heat transfer is given by Rabas and Taborek (1987). Chai (1988) provides the best correlation for friction factors:

$$f_{tb} = 1.748\mathrm{Re}_d^{-0.233}\left(\frac{l_f}{s}\right)^{0.552}\left(\frac{d_o}{X_t}\right)^{0.599}\left(\frac{d_o}{X_l}\right)^{0.1738} \tag{4.1.83}$$

This correlation is valid for $895 < \mathrm{Re}_d < 713{,}000$, $20 < \theta < 40°$, $X_t/d_o < 4$, $N \ge 4$, and θ is the tube layout angle. It predicts 89 literature data points within a mean absolute error of 6%; the range of actual error is from –16.7 to 19.9%.

Flat plain fins on a staggered tubebank. This geometry, as shown in Figure 4.1.26b, is used in air-conditioning/refrigeration industry as well as where the pressure drop on the fin side prohibits

the use of enhanced/interrupted flat fins. An inline tubebank is generally not used unless very low fin side pressure drop is the essential requirement. Heat transfer correlation for Figure 4.1.26b flat plain fins on staggered tubebanks is provided by Gray and Webb (see Webb 1994) as follows for four or more tube rows.

$$j_4 = 0.14\mathrm{Re}_\mathrm{d}^{-0.328}\left(X_\mathrm{t}/X_l\right)^{-0.502}\left(s/d_\mathrm{o}\right)^{0.031} \tag{4.1.84}$$

For the number of tube rows N from 1 to 3, the j factor is lower and is given by

$$\frac{j_\mathrm{N}}{j_4} = 0.991\left[2.24\mathrm{Re}_\mathrm{d}^{-0.092}\left(N/4\right)^{-0.031}\right]^{0.607(4-N)} \tag{4.1.85}$$

Gray and Webb (see Webb 1994) hypothesized the friction factor consisting of two components: one associated with the fins and the other associated with the tubes as follows.

$$f = f_\mathrm{f}\frac{A_\mathrm{f}}{A} + f_\mathrm{t}\left(1-\frac{A_\mathrm{f}}{A}\right)\left(1-\frac{\delta_\mathrm{f}}{p_\mathrm{f}}\right) \tag{4.1.86}$$

where

$$f_\mathrm{f} = 0.508\mathrm{Re}_\mathrm{d}^{-0.521}\left(X_\mathrm{t}/d_\mathrm{o}\right)^{1.318} \tag{4.1.87}$$

and f_t (defined the same way as f) is the Fanning friction factor associated with the tube and can be determined using the charts available in Zukauskas (1987) as $f_\mathrm{t}=\mathrm{EuN}(X_\mathrm{t}-d_\mathrm{o})/\pi d_\mathrm{o}$. Equation 4.1.86 correlated 90% of the data for 19 heat exchangers within ±20%. The range of dimensionless variables of Equations 4.1.86 and 4.1.87 are $500 \le \mathrm{Re}_\mathrm{d} \le 24{,}700$, $1.97 \le X_\mathrm{t}/d_\mathrm{o} \le 2.55$, $1.7 \le X_\mathrm{t}/d_\mathrm{o} \le 2.58$, and $0.08 \le s/d_\mathrm{o} \le 0.64$.

Exchanger Design Methodology

The problem of heat exchanger design is complex and multidisciplinary (Shah, 1991). The major design considerations for a new heat exchanger include process/design specifications, thermal and hydraulic design, mechanical design, manufacturing and cost considerations, and trade-offs and system-based optimization, as shown in Figure 4.1.37 with possible strong interactions among these considerations as indicated by double-sided arrows. The thermal and hydraulic design methods are mainly analytical, and the structural design is analytical to some extent. Most of the other major design considerations involve qualitative and experience-based judgments, trade-offs, and compromises. Therefore, there is no unique solution to designing a heat exchanger for given process specifications. Further details on this design methodology are given by Shah (1991).

Two important heat exchanger design problems are the rating and sizing problems. Determination of heat transfer and pressure drop performance of either an existing exchanger or an already sized exchanger is referred to as the rating problem. The objective here is to verify vendor's specifications or to determine the performance at off-design conditions. The rating problem is also sometimes referred to as the performance problem. In contrast, the design of a new or existing type of exchanger is referred to as the sizing problem. In a broad sense, it means the determination of the exchanger construction type, flow arrangement, heat transfer surface geometries and materials, and the physical size of an exchanger to meet the specified heat transfer and pressure drops. However, from the viewpoint of quantitative thermal–hydraulic analysis, we will consider that the selection of the exchanger construction type, flow arrangement, and materials has already been made. Thus, in the sizing problem, we will determine the physical size (length, width, height) and surface areas on each side of the exchanger. The sizing problem is also sometimes referred to as the design problem.

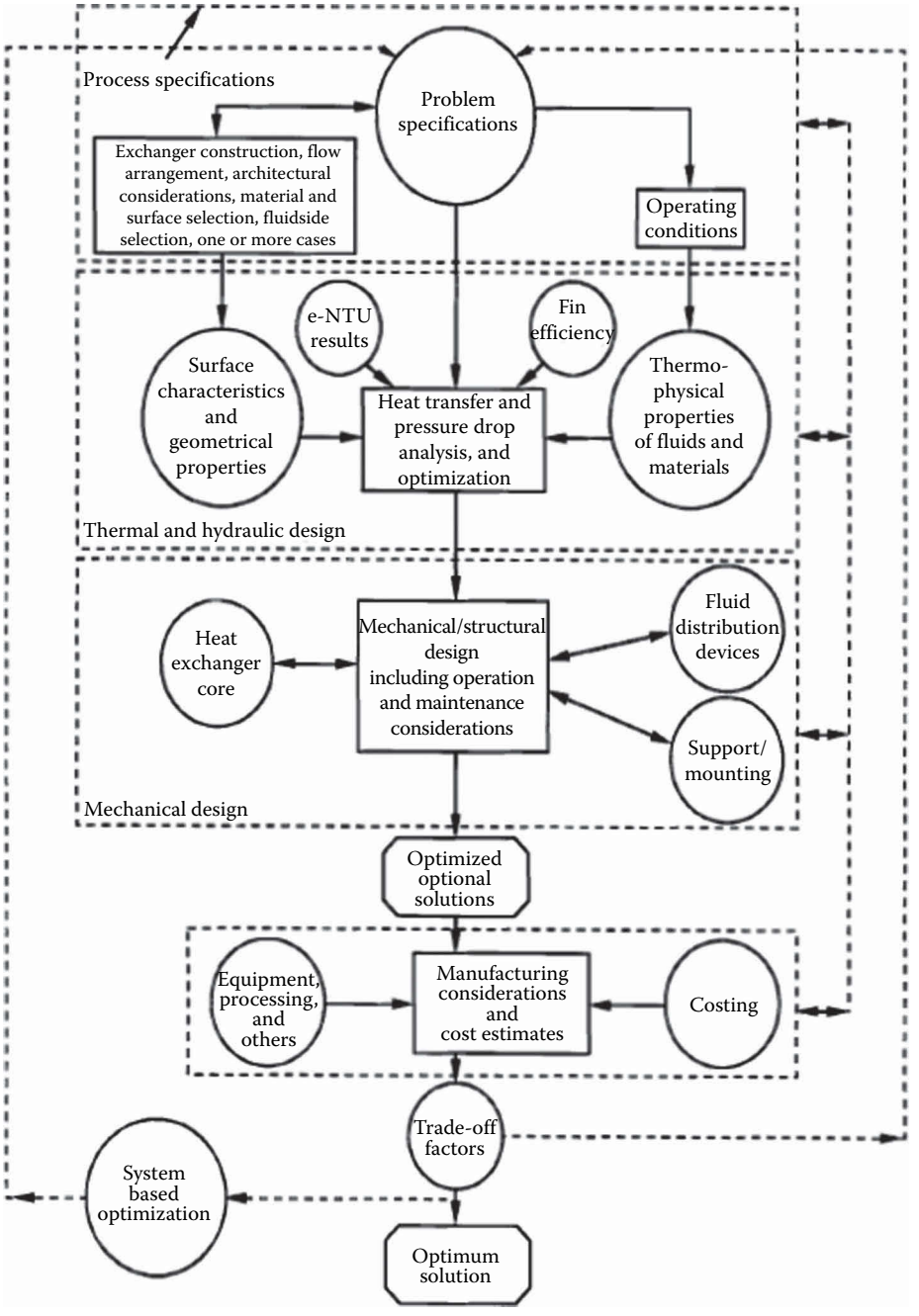


FIGURE 4.1.37 Heat exchanger design methodology.

The step-by-step solution procedures for the rating and sizing problems for counterflow and crossflow single-pass plate-fin heat exchangers have been presented with a detailed illustrative example by Shah (1981). Shah (1988a) presented further refinements in these procedures as well as step-by-step procedures for two-pass cross-counterflow plate-fin exchangers, and single-pass crossflow and two-pass cross-counterflow tube-fin exchangers. Also, step-by-step solution procedures for the rating and sizing problems for rotary regenerators (Shah 1988b), heat pipe heat exchangers (Shah, and

Giovannelli, 1988), and PHEs (Shah and Wanniarachchi, 1991) are available. As an illustration, the step-by-step solution procedures will be covered here for a single-pass crossflow exchanger.

Rating Problem for a Crossflow Plate-Fin Exchanger

Following is a step-by-step procedure for rating a crossflow plate-fin exchanger. Inputs to the rating problem for a two-fluid exchanger are the exchanger construction, flow arrangement and overall dimensions, complete details on the materials and surface geometries on both sides including their nondimensional heat transfer and pressure drop characteristics (j and f vs. Re), fluid flow rates, inlet temperatures, and fouling factors. The fluid outlet temperatures, total heat transfer rate, and pressure drops on each side of the exchanger are then determined as the rating problem solution.

1. Determine the surface geometric properties on each fluid side. This includes the minimum free-flow area A_o, heat transfer surface area A (both primary and secondary), flow lengths L, hydraulic diameter D_h, heat transfer surface area density β, the ratio of minimum free-flow area to frontal area σ, fin length l_f, and fin thickness δ for fin efficiency determination, and any specialized dimensions used for heat transfer and pressure drop correlations.
2. Compute the fluid bulk mean temperature and fluid thermophysical properties on each fluid side. Since the outlet temperatures are not known for the rating problem, they are estimated initially. Unless it is known from past experience, assume an exchanger effectiveness as 60%–75% for most single-pass crossflow exchangers or 80%–85% for single-pass counterflow exchangers. For the assumed effectiveness, calculate the fluid outlet temperatures.

$$T_{h,o} = T_{h,i} - e\left(C_{min} / C_h\right)\left(T_{h,i} - T_{c,i}\right) \tag{4.1.88}$$

$$T_{c,o} = T_{c,i} - e\left(C_{min} / C_c\right)\left(T_{h,i} - T_{c,i}\right) \tag{4.1.89}$$

   Initially, assume $C_c/C_h = \dot{m}_c/\dot{m}_h$ a gas-to-gas exchanger, or $C_c/C_h = \dot{m}_c c_{p,c}/\dot{m}_h c_{p,h}$ for a gas-to-liquid exchanger with very approximate values of c_p for the fluids in question.

   For exchangers with $C^* > 0.5$ (usually gas-to-gas exchangers), the bulk mean temperatures on each fluid side will be the arithmetic mean of the inlet and outlet temperatures on each fluid side (Shah, 1981). For exchangers with $C^* > 0.5$ (usually gas-to-gas exchangers), the bulk mean temperature on the C_{max} side will be the arithmetic mean of inlet and outlet temperatures; the bulk mean temperature on the C_{min} side will be the log-mean average temperature obtained as follows:

$$T_{m,C_{min}} = T_{m,C_{max}} \pm \Delta T_{lm} \tag{4.1.90}$$

   where ΔT_{lm} is the LMTD based on the terminal temperatures (see Equation 4.1.55). Use the plus sign if the C_{min} side is hot; otherwise, use the negative sign.

   Once the bulk mean temperature is obtained on each fluid side, obtain the fluid properties from thermophysical property books or from handbooks. The properties needed for the rating problem are μ, c_p, k, Pr, and ρ. With this c_p, one more iteration may be carried out to determine $T_{h,o}$ or $T_{c,o}$ from Equation 4.1.88 or 4.1.89 on the C_{max} side, and subsequently T_m on the C_{max} side, and refine fluid properties accordingly.

3. Calculate the Reynolds number $\mathrm{Re} = GD_h/\mu$ and/or any other pertinent dimensionless groups (from the basic definitions) needed to determine the nondimensional heat transfer and flow friction characteristics (e.g., j or Nu and f) of heat transfer surfaces on each side of the exchanger. Subsequently, compute j or Nu and f factors. Correct Nu (or j) for variable fluid property effects (Shah, 1981) in the second and subsequent iterations from the following equations.

$$\text{For gases}: \quad \frac{\mathrm{Nu}}{\mathrm{Nu}_{cp}} = \left[\frac{T_w}{T_m}\right]^{n'}; \quad \frac{f}{f_{cp}} = \left[\frac{T_w}{T_m}\right]^{m'} \tag{4.1.91}$$

$$\text{For liquids}: \quad \frac{\text{Nu}}{\text{Nu}_{cp}} = \left[\frac{\mu_w}{\mu_m}\right]^{n'}; \quad \frac{f}{f_{cp}} = \left[\frac{\mu_w}{\mu_m}\right]^{m'} \tag{4.1.92}$$

where the subscript cp denotes constant properties, and m' and n' are empirical constants provided in Table 4.1.13. Note that T_w and T_m in Equations 4.1.91 and 4.1.92 and in Tables 4.1.13a and b and are absolute temperatures.

4. From Nu or j, compute the heat transfer coefficients for both fluid streams.

$$h = \text{Nu}\ k / D_h = jGc_p\text{Pr}^{-2/3} \tag{4.1.93}$$

Subsequently, determine the fin efficiency η_f and the extended surface efficiency η_o

$$\eta_f = \frac{\tanh ml}{ml} \quad \text{where} \quad m^2 = \frac{h\tilde{p}}{k_f A_k} \tag{4.1.94}$$

TABLE 4.1.13A
Property Ratio Method Exponents of Equations 4.3.91 and 4.3.92 for Laminar Flow

Fluid	Heating	Cooling
Gases	$n'=0.00$, $m'=1.00$ for $1 < T_w/T_m < 3$	$n'=0.0$, $m'=0.81$ for $0.5 < T_w/T_m < 1$
Liquids	$n'=-0.14$, $m'=0.58$ for $\mu_w/\mu_m < 1$	$n'=-0.14$, $m'=0.54$ for $\mu_w/\mu_m > 1$

Source: Shah, R.K., Compact heat exchangers, *Heat Exchangers: Thermal-Hydraulic Fundamentals and Design*, S. Kakaç et al., Eds., Hemisphere Publishing, Washington, DC, 1981. With permission.

TABLE 4.1.13B
Property Ratio Method Correlations of Exponents of Equations 4.3.91 and 4.3.92 for Turbulent Flow

Fluid	Heating	Cooling
Gases	$\text{Nu}=5+0.012\ \text{Re}^{0.83}\ (\text{Pr}+0.29)\ (T_w/T_m)^n$	$n'=0$
	$n=-[\log_{10}(T_w/T_m)]^{1/4}+0.3$ for $1 < T_w\ T_m < 5$, $0.6 < \text{Pr} < 0.9$, $10^4 < \text{Re} < 10^6$, and $L/D_h > 40$	
	$m'=-0.1$ for $1 < T_w/T_m < 2.4$	$m'=-0.1$ (tentative)
Liquids	$n'=-0.11$[a] for $0.08 < \mu_w/\mu_m < 1$	$n'=-0.25$[a] for $1 < \mu_w/\mu_m < 40$
	$f/f_{cp}=(7-\mu_m/\mu_w/6$[b] or $m'=0.25$ for $0.35 < \mu_w/\mu_m < 1$	$m'=0.24$[b] or $1 < \mu_w/\mu_m < 2$

Source: Shah, R.K., Compact heat exchangers, *Heat Exchangers: Thermal-Hydraulic Fundamentals and Design*, S. Kakaç et al., Eds., Hemisphere Publishing, Washington, DC, 1981. With permission.

[a] Valid for $2 \le \text{Pr} \le 140$, $10^4 \le \text{Re} \le 1.25 \times 10^5$.

[b] Valid for $1.3 \le \text{Pr} \le 10$, $10^4 \le \text{Re} \le 2.3 \times 10^5$.

where $\tilde{P}$ is the wetted perimeter of the fin surface.

$$\eta_o = 1 - \frac{A_f}{A}\left(1 - \eta_f\right) \tag{4.1.95}$$

Also calculate the wall thermal resistance $R_w = \delta/A_w k_w$. Finally, compute the overall thermal conductance UA from Equation 4.1.42 knowing the individual convective film resistances, wall thermal resistances, and fouling resistances, if any.

5. From the known heat capacity rates on each fluid side, compute $C^* = \boldsymbol{C}_{min}/C_{max}$. From the known UA, determine NTU $= UA/C_{min}$. Also calculate the longitudinal conduction parameter λ. With the known NTU, $\boldsymbol{C}^*$, λ, and the flow arrangement, determine the exchanger effectiveness ε from either closed-form equations of Table 4.1.10 or tabular/graphical results from Kays and London (1998).
6. With this ε, finally compute the outlet temperatures from Equations 4.1.88 and 4.1.89. If these outlet temperatures are significantly different from those assumed in Step 2, use these outlet temperatures in Step 2 and continue iterating Steps 2–6, until the assumed and computed outlet temperatures converge within the desired degree of accuracy. For a gas-to-gas exchanger, most probably one or two iterations will be sufficient.
7. Finally, compute the heat duty from

$$q = \varepsilon C_{min}\left(T_{h,i} - T_{c,i}\right) \tag{4.1.96}$$

8. For the pressure drop calculations, first we need to determine the fluid densities at the exchanger inlet and outlet (ρ_i and ρ_o) for each fluid. The mean specific volume on each fluid side is then computed from Equation 4.1.66.

Next, the entrance and exit loss coefficients, K_c and K_e, are obtained from Figure 4.1.35 for known σ, Re, and the flow passage entrance geometry.

The friction factor on each fluid side is corrected for variable fluid properties using Equation 4.1.91 or 4.1.92. Here, the wall temperature T_w is computed from

$$T_{w,h} = T_{m,h} - (R_h + R_{s,h})\,q \tag{4.1.97}$$

$$T_{w,c} = T_{m,c} + \left(R_c + R_{s,c}\right)q \tag{4.1.98}$$

where the various resistance terms are defined by Equation 4.1.42.

The core pressure drops on each fluid side are then calculated from Equation 4.1.65. This then completes the procedure for solving the rating problem.

Sizing Problem for a Crossflow Plate-Fin Exchanger

As defined earlier, we will concentrate here to determine the physical size (length, width, and height) of a single-pass crossflow exchanger for specified heat duty and pressure drops. More specifically, inputs to the sizing problem are surface geometries (including their nondimensional heat transfer and pressure drop characteristics), fluid flow rates, inlet and outlet fluid temperatures, fouling factors, and pressure drops on each side.

For the solution to this problem, there are four unknowns—two flow rates or Reynolds numbers (to determine correct heat transfer coefficients and friction factors) and two surface areas—for the two-fluid crossflow exchanger. The following four equations—Equations 4.1.99, 4.1.101, 4.1.102 and 4.1.103—are used to solve iteratively the surface areas on each fluid side: UA in Equation 4.1.99 is

determined from NTU computed from the known heat duty or ε and C^*; G in Equation 4.1.101 represents two equations, for Fluids 1 and 2 (Shah, 1988a); and the volume of the exchanger in Equation 4.1.103 is the same based on the surface area density of Fluid 1 or Fluid 2.

$$\frac{1}{UA} \approx \frac{1}{(\eta_o hA)_h} + \frac{1}{(\eta_o hA)_c} \tag{4.1.99}$$

Here, we have neglected the wall and fouling thermal resistances. This equation in nondimensional form is given by

$$\frac{1}{\text{NTU}} = \frac{1}{\text{ntu}_h (C_h/C_{\min})} + \frac{1}{\text{ntu}_c (C_c/C_{\min})} \tag{4.1.100}$$

$$G_i = \left[\frac{2 g_c \Delta p}{\text{Deno}}\right]_i^{1/2} \qquad i = 1,2 \tag{4.1.101}$$

where

$$\text{Deno}_i = \left[\frac{f}{j}\frac{\text{ntu}}{\eta_o}\text{Pr}^{2/3}\left(\frac{1}{\rho}\right)_m + 2\left(\frac{1}{\rho_o} - \frac{1}{\rho_i}\right) + \left(1-\sigma^2+K_c\right)\frac{1}{\rho_i} - \left(1-\sigma^2-K_e\right)\frac{1}{\rho_o}\right]_i \tag{4.1.102}$$

$$V = \frac{A_1}{\alpha_1} = \frac{A_2}{\alpha_2} \tag{4.1.103}$$

In the iterative solutions, the first time one needs ntu_h and ntu_c to start the iterations. These can be either determined from the past experience or by estimations. If both fluids are gases or both fluids are liquid, one could consider that the design is "balanced," that is, that the thermal resistances are distributed approximately equally on the hot and cold sides. In that case, $C_h = C_c$, and

$$\text{ntu}_h \approx \text{ntu}_c \approx 2\,\text{NTU} \tag{4.1.104}$$

Alternatively, if we have liquid on one side and gas on the other side, consider 10% thermal resistance on the liquid side, that is,

$$0.10\left(\frac{1}{UA}\right) = \frac{1}{(\eta_o hA)_{\text{liq}}} \tag{4.1.105}$$

Then, from Equations 4.1.99 and 4.1.100 with $C_{\text{gas}} = C_{\min}$, we can determine the ntu values on each side as follows:

$$\text{ntu}_{\text{gas}} = 1.11\text{NTU}, \qquad \text{ntu}_{\text{liq}} = 10C^*\text{NTU} \tag{4.1.106}$$

Also note that initial guesses of η_o and *j/f* are needed for the first iteration to solve Equation 4.1.102. For a good design, consider $\eta_o = 0.80$ and determine an approximate value of *j/f* from the plot of *j/f* vs. Re curve for the known *j* and *f* vs. Re characteristics of each fluid side surface. The specific step-by-step design procedure is as follows:

1. In order to compute the fluid bulk mean temperature and the fluid thermophysical properties on each fluid side, determine the fluid outlet temperatures from the specified heat duty

$$q = (\dot{m}c_p)_h (T_{h,i} - T_{h,o}) = (\dot{m}c_p)_c (T_{c,o} - T_{c,i}) \tag{4.1.107}$$

or from the specified exchanger effectiveness using Equations 4.1.88 and 4.1.89. For the first time, estimate the values of c_p.

For exchangers with $C^* \geq 0.5$, the bulk mean temperature on each fluid side will be the arithmetic mean of inlet and outlet temperatures on each side. For exchangers with $C^* < 0.5$, the bulk mean temperature on the C_{max} side will be the arithmetic mean of the inlet and outlet temperatures on that side and the bulk mean temperature on the C_{min} side will be the log-mean average as given by Equation 4.1.90. With these bulk mean temperatures, determine c_p and iterate one more time for the outlet temperatures if warranted. Subsequently, determine μ, c_p, k, Pr, and ρ on each fluid side.

2. Calculate C^* and ε (if q is given), and determine NTU from the ε-NTU expression, tables, or graphical results for the selected flow arrangement (in this case, it is unmixed–unmixed crossflow, Table 4.1.10). The influence of longitudinal heat conduction, if any, is ignored in the first iteration since we do not know the exchanger size yet.
3. Determine ntu on each side by the approximations discussed with Equations 4.1.104 and 4.1.106 unless it can be estimated from past experience.
4. For the selected surfaces on each fluid side, plot *j*/*f* vs. Re curve from the given surface characteristics and obtain an approximate value of *j*/*f*. If fins are employed, assume $\eta_o = 0.80$ unless a better value can be estimated.
5. Evaluate G from Equation 4.1.101 on each fluid side using the information from Steps 1–4 and the input value of Δp.
6. Calculate Reynolds number Re, and determine j and f on each fluid side from the given design data for each surface.
7. Compute h, η_f, and η_o using Equations 4.1.93 through 4.1.95. For the first iteration, determine U_1 on Fluid 1 side from the following equation derived from Equations 4.1.42 and 4.1.103.

$$\frac{1}{U_1} = \frac{1}{(\eta_o h)_1} + \frac{1}{(\eta_o h_s)_1} + \frac{\alpha_1/\alpha_2}{(\eta_o h_s)_2} + \frac{\alpha_1/\alpha_2}{(\eta_o h)_2} \tag{4.1.108}$$

where $\alpha_1/\alpha_2 = A_1/A_2$, $\alpha = A/V$, V is the exchanger total volume, and subscripts 1 and 2 denote Fluid 1 and 2 sides. For a plate-fin exchanger, α terms are given by Shah (1981) and Kays and London (1998):

$$\alpha_1 = \frac{b_1\beta_1}{b_1 + b_2 + 2a}; \qquad \alpha_2 = \frac{b_2\beta_2}{b_1 + b_2 + 2a} \tag{4.1.109}$$

Note that the wall thermal resistance in Equation 4.1.108 is ignored in the first iteration. In second and subsequent iterations, compute U_1 from

$$\frac{1}{U_1} = \frac{1}{(\eta_o h)_1} + \frac{1}{(\eta_o h_s)_1} + \frac{\delta A_1}{k_w A_w} + \frac{A_1/A_2}{(\eta_o h_s)_2} + \frac{A_1/A_2}{(\eta_o h_2)_2} \tag{4.1.110}$$

where the necessary geometry information A_1/A_2 and A_1/A_w is determined from the geometry calculated in the previous iteration.

8. Now calculate the core dimensions. In the first iteration, use NTU computed in Step 2. For subsequent iterations, calculate longitudinal conduction parameter λ (and other dimensionless groups for a crossflow exchanger). With known ε, C^*, and λ, determine the correct value of NTU using either a closed-form equation or tabular/graphical results (Kays and London 1998). Determine A_1 from NTU using U_1 from the previous step and known C_{min}.

$$A_1 = \text{NTU}\, C_{min} / U_1 \tag{4.1.111}$$

and hence

$$A_2 = (A_2 / A_1)A_1 = (\alpha_2 / \alpha_1)A_1 \tag{4.1.112}$$

A_o from known $\dot{m}$ and G is given by

$$A_{o,1} = (\dot{m}/G)_1; \quad A_{o,2} = (\dot{m}/G)_2 \tag{4.1.113}$$

so that

$$A_{fr,1} = A_{o,1} / \sigma_1; \quad A_{fr,2} = A_{o,2} / \sigma_2 \tag{4.1.114}$$

where σ_1 and σ_2 are generally specified for the surface or can be computed for plate-fin surfaces from Shah (1981) and Kays and London (1998):

$$\sigma_1 = \frac{b_1\beta_1 D_{h,1}/4}{b_1 + b_2 + 2\delta}; \quad \sigma_2 = \frac{b_2\beta_2 D_{h,2}/4}{b_1 + b_2 + 2\delta} \tag{4.1.115}$$

Now compute the fluid flow lengths on each side (see Figure 4.1.38) from the definition of the hydraulic diameter of the surface employed on each side.

$$L_1 = \left(\frac{D_h A}{4A_o}\right)_1; \quad L_2 = \left(\frac{D_h A}{4A_o}\right)_2 \tag{4.1.116}$$

Since $A_{fr,1} = L_2L_3$ and $A_{fr,2} = L_1L_3$, we can obtain

$$L_3 = \frac{A_{fr,1}}{L_2} \quad \text{or} \quad L_3 = \frac{A_{fr,2}}{L_1} \tag{4.1.117}$$

Theoretically, L_3 calculated from both expressions of Equation 4.1.117 should be identical. In reality, they may differ slightly because of the round-off error. In that case, consider an average value for L_3.

9. Finally, compute the pressure drop on each fluid side, after correcting f factors for variable property effects, in a manner similar to Step 8 of the Rating Problem for a Crossflow Plate Fin Exchanger.

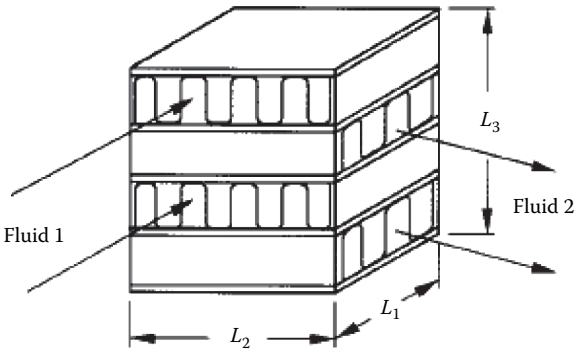


FIGURE 4.1.38 A single-pass crossflow exchanger.

10. If the calculated values of Δp are within and close to input specifications, the solution to the sizing problem is completed. Finer refinements in the core dimensions, such as integer numbers of flow passages, etc., may be carried out at this time. Otherwise, compute the new value of G on each fluid side using Equation 4.1.65 in which Δp is the input-specified value, and f, K_c, K_e, and geometric dimensions are from the previous iteration.
11. Repeat (iterate) Steps 6–10 until both heat transfer and pressure drops are met as specified. It should be emphasized that since we have imposed no constraints on the exchanger dimensions, this procedure will yield L_1, L_2, and L_3 for the selected surfaces such that the design will meet exactly the heat duty and pressure drops on both fluid sides.

Flow Maldistribution

In the previously presented heat transfer (ε-NTU, MTD, etc. methods) and pressure drop analyses, it is presumed that the fluid is uniformly distributed through the core. In practice, flow maldistribution does occur to some extent and often severely, and may result in a significant reduction in exchanger heat transfer performance and an increase in the pressure drop. Hence, it may be necessary for the designer to take into account the effect of flow maldistribution causing undesirable performance deterioration up front while designing a heat exchanger.

Some maldistributions are geometry-induced (i.e., the result of exchanger fabrication conditions, such as header design or manufacturing tolerances, or the duct geometry/structure upstream of the exchanger), and other maldistributions are the result of exchanger operating conditions. Gross, passage-to-passage, and manifold-induced flow maldistributions are examples of the former category, while viscosity, natural convection, and density-difference-induced flow maldistributions are of the latter category. Flow maldistributions associated with two-phase and multiphase flow are too complex, with only limited information available in the literature. The analysis methods and results for some of the above flow maldistributions for single-phase flows are given by Shah (1985), Mueller and Chiou (1987), and Putnam and Rohsenow (1985).

Fouling in Heat Exchangers

Fouling, Its Effect, and Mechanisms

Fouling refers to undesired accumulation of solid material (by-products of the heat transfer processes) on heat exchanger surfaces which results in additional thermal resistance to heat transfer, thus reducing exchanger performance. The fouling layer also blocks the flow passage/area and increases surface roughness, thus either reducing the flow rate in the exchanger or increasing the pressure drop or both. The foulant deposits may be loose such as magnetite particles or hard and tenacious such as calcium carbonate scale; other deposits may be sediment, polymers, coking or corrosion products, inorganic salts, biological growth, etc. Depending upon the fluids, operating conditions, and heat exchanger construction, the maximum fouling layer thickness on the heat transfer surface may result in a few hours to a number of years.

Fouling could be very costly depending upon the nature of fouling and the applications. It increases capital costs: (1) oversurfacing heat exchanger, (2) provisions for cleaning, and (3) use of special materials and constructions/surface features. It increases maintenance costs: (1) cleaning techniques, (2) chemical additives, and (3) troubleshooting. It may cause a loss of production: (1) reduced capacity and (2) shutdown. It increases energy losses: (1) reduced heat transfer, (2) increased pressure drop, and (3) dumping dirty streams. Fouling promotes corrosion, severe plugging, and eventual failure of uncleaned heat exchangers. In a fossil-fired exhaust environment, gas-side fouling produces a potential fire hazard in heat exchangers.

The following are the major fouling mechanisms:

- Crystallization or precipitation fouling results from the deposition/formation of crystals of dissolved substances from the liquid onto heat transfer surface due to solubility changes with temperature beyond the saturation point. If the deposited layer is hard and tenacious, it is often referred to as scaling. If it is porous and mushy, it is called sludge.
- Particulate fouling results from the accumulation of finely divided substances suspended in the fluid stream onto heat transfer surface. If the settling occurs as a result of gravity, it is referred to as sedimentation fouling.
- Chemical reaction fouling is defined as the deposition of material produced by chemical reaction (between reactants contained in the fluid stream) in which the heat transfer surface material does not participate.
- Corrosion fouling results from corrosion of the heat transfer surface that produces products fouling the surface and/or roughens the surface, promoting attachment of other foulants.
- Biological fouling results from the deposition, attachment, and growth of biological organisms from liquid onto a heat transfer surface. Fouling due to microorganisms refers to microbial fouling and fouling due to macroorganisms refers to macrobial fouling.
- Freezing fouling results from the freezing of a single-component liquid or higher-melting-point constituents of a multicomponent liquid onto a subcooled heat transfer surface.

Biological fouling occurs only with liquids since there are no nutrients in gases. Also crystallization fouling is not too common with gases since most gases contain few dissolved salts (mainly in mists) and even fewer inverse-solubility salts. All other types of fouling occur in both liquid and gas. More than one mechanism is usually present in many fouling situations, often with synergetic results. Liquid-side fouling generally occurs on the exchanger side where the liquid is being heated, and gas-side fouling occurs where the gas is being cooled; however, reverse examples can be found.

Importance of Fouling

Fouling in liquids and two-phase flows has a significant detrimental effect on heat transfer with some increase in pressure drop. In contrast, fouling in gases reduces heat transfer somewhat (5%–10% in general) in compact heat exchangers, but increases pressure drop significantly up to several hundred percent. For example, consider $U = 1400\,\mathrm{W/m^2\,K}$ as in a process plant liquid-to-liquid heat exchanger. Hence, $R = 1/U = 0.00072\,\mathrm{m^2\,K/W}$. If the fouling factors $(r_{s,h} + r_{s,c})$ together amount to 0.00036 (considering a typical TEMA value of the fouling factor as 0.00018), 50% of the heat transfer area requirement A for given q is chargeable to fouling. However, for gas flows on both sides of an exchanger, $U \approx 280\,\mathrm{W/m^2\,K}$, and the same fouling factor of 0.00036 would represent only about 10% of the total surface area. Thus, one can see a significant impact on the heat transfer surface area requirement due to fouling in heat exchangers having high $\boldsymbol{U}$ values (such as having liquids or phase-change flows).

Considering the core frictional pressure drop (Equation 4.1.68) as the main pressure drop component, the ratio of pressure drops of fouled and cleaned exchangers is given by

$$\frac{\Delta P_F}{\Delta p_C} = \frac{f_F}{f_C}\left(\frac{D_{h,C}}{D_{h,F}}\right)\left(\frac{u_{m,F}}{u_{m,C}}\right)^2 = \frac{f_F}{f_C}\left(\frac{D_{h,C}}{D_{h,F}}\right)^5 \tag{4.1.118}$$

where the term after the second equality sign is for a circular tube and the mass flow rates under fouled and clean conditions remain the same. Generally, $f_F > f_c$ due to the fouled surface being rough. Thus, although the effect of fouling on the pressure drop is usually neglected, it can be significant, particularly for compact heat exchangers with gas flows. If we consider $f_F = f_c$, and the reduction in the tube inside diameter due to fouling by only 10% and 20%, the resultant pressure drop increase

will be 69% and 205%, respectively, according to Equation 4.1.118 regardless of whether the fluid is liquid or gas!

Accounting of Fouling in Heat Exchangers

Fouling is an extremely complex phenomenon characterized by a combined heat, mass, and momentum transfer under transient condition. Fouling is affected by a large number of variables related to heat exchanger surfaces, operating conditions, and fluids. Fouling is time dependent, zero at $\tau = 0$; after the induction or delay period τ_d, the fouling resistance is either pseudolinear, falling rate, or asymptotic.

Fouling is characterized by all or some of the following sequential events: initiation, transport, attachment, removal, and aging (Epstein, 1983). Research efforts are concentrated on quantifying these events by semitheoretical models (Epstein, 1978) with very limited success on specific fouling situations. Hence, the current heat exchanger design approach is to use a constant (supposedly an asymptotic) value of the fouling factor $r_s = 1/h_s$. Equation 4.1.42 presented earlier includes the fouling resistances on the hot and cold sides for a nontubular extended-surface exchanger. Here $1/h_s = r_s$ is generally referred to as the *fouling factor.* Fouling factors for some common fluids are presented in Tables 4.1.14 and 4.1.15. Also see Section 4.4 in this Handbook.

The specification of fouling effects in a process heat exchanger is usually represented in the following form, wherein the combined fouling factor $r_{s,t}$ is the sum of the fouling factors on the hot and cold sides:

$$\text{Combined fouling factor} \qquad r_{s,t} = \frac{1}{U_C} - \frac{1}{U_F} \tag{4.1.119}$$

$$\text{Cleanliness factor CF} = U_F / U_C \tag{4.1.120}$$

$$\text{Percentage oversurface} \quad \%\text{OS} = \left(\frac{A_F}{A_C} - 1 \right) 100 \tag{4.1.121}$$

Here the subscripts *F* and *C* denote fouled and clean exchanger values. From Equation 4.1.42 with $A_h = A_c = A$, $\eta_o = 1$, $\Delta T_{m,F} = \Delta T_{m,C}$, it can be shown that

$$\frac{A_F}{A_C} = \frac{U_C}{U_F} = 1 + U_C r_{s,t} \tag{4.1.122}$$

where $r_{s,t} = r_{s,h} + r_{s,c}$. In heat exchanger design, constant (supposedly an asymptotic) values of $r_{s,h}$ and $r_{s,c}$ are used. Accordingly, extra heat transfer surface area is provided to take into account the deleterious effect of fouling. Thus, the heat exchanger will be "oversized" for the initial clean condition, "correctly sized" for asymptotic fouling (if it occurs in practice), and "undersized" just before the cleaning operation for nonasymptotic fouling.

Influence of Operating and Design Variables

Based on operational experience and research over the last several decades, many variables have been identified that have a significant influence on fouling. The most important variables are summarized next.

Flow velocity. Flow velocity is one of the most important variables affecting fouling. Higher velocities increase fluid shear stress at the fouling deposit–fluid interface and increase the heat transfer coefficient; but, at the same time, increased pressure drop and fluid pumping power may erode the surface and may accelerate the corrosion of the surface by removing the protective oxide layer. The fouling buildup in general is inversely proportional to $u_m^{1.5}$. For water, the velocity should be kept above 2 m/s to suppress fouling, and the absolute minimum should be above 1 m/s to minimize fouling.

TABLE 4.1.14
Fouling Factors for Various Fluid Streams Used in Heat Exchangers

Water Type	Fouling Factors (m^2 K)/W
Seawater (43°C maximum outlet)	0.000275–0.00035
Brackish water (43°C maximum outlet)	0.00035–0.00053
Treated cooling tower water (49°C maximum outlet)	0.000175–0.00035
Artificial spray pond (49°C maximum outlet)	0.000175–0.00035
Closed-loop treated water	0.000175
River water	0.00035–0.00053
Engine jacket water	0.000175
Distilled water or closed-cycle condensate	0.00009–0.000175
Treated boiler feedwater	0.00009
Boiler blowdown water	0.00035–0.00053
Liquids	
No. 2 fuel oil	0.00035
No. 6 fuel oil	0.0009
Transformer oil	0.000175
Engine lube oil	0.000175
Refrigerants	0.000175
Hydraulic fluid	0.000175
Industrial organic HT fluids	0.000175–0.00035
Ammonia	0.000175
Ammonia (oil bearing)	0.00053
Methanol solutions	0.00035
Ethanol solutions	0.00035
Ethylene glycol solutions	0.00035
MEA and DEA solutions	0.00035
DEG and TEG solutions	0.00035
Stable side draw and bottom products	0.000175–0.00035
Caustic solutions	0.00035
Gas or Vapor	
Steam (nonoil bearing)	0.0009
Exhaust steam (oil bearing)	0.00026–0.00035
Refrigerant (oil bearing)	0.00035
Compressed air	0.000175
Ammonia	0.000175
Carbon dioxide	0.00035
Coal flue gas	0.00175
Natural gas flue gas	0.00090
Acid gas	0.00035–0.00053
Solvent vapor	0.000175
Stable overhead products	0.000175
Natural Gas and Petroleum Streams	
Natural gas	0.000175–0.00035
Overhead products	0.000175–0.00035
Lean oil	0.00035
Rich oil	0.000175–0.00035
Natural gasoline and liquefied petroleum gases	0.000175–0.00035
Oil Refinery Streams	
Crude and vacuum unit gases and vapors	
Atmospheric tower overhead vapors	0.00017
Light naphthas	0.00017
Vacuum overhead vapors	0.00035

(*Continued*)

TABLE 4.1.14 (*Continued*)
Fouling Factors for Various Fluid Streams Used in Heat Exchangers

Water Type	Fouling Factors (m^2 K)/W
Oil Refinery Streams	
Crude and vacuum liquids	
Gasoline	0.00035
Naphtha and light distillates	0.00035–0.00053
Kerosene	0.00035–0.00053
Light gas oil	0.00035–0.00053
Heavy gas oil	0.00053–0.0009
Heavy fuel oil	0.00053–0.00123
Vacuum tower bottoms	0.00176
Atmospheric tower bottoms	0.00123
Cracking and Coking Unit Streams	
Overhead vapors	0.00035
Light cycle oil	0.00035–0.00053
Heavy cycle oil	0.00053–0.0007
Light coker gas oil	0.00053–0.0007
Heavy coker gas oil	0.00070–0.0009
Bottoms slurry oil (1.5 m/s minimum)	0.00053
Light liquid products	0.00035
Catalytic Reforming, Hydrocracking, and Hydrodesulfurization Streams	
Reformer charge	0.00026
Reformer effluent	0.00026
Hydrocharger charge and effluent[a]	0.00035
Recycle gas	0.000175
Liquid product over 50°C (API)[b]	0.000175
Liquid product 30°C–50°C (API)[b]	0.00035
Light ends processing streams	
Overhead vapors and gases	0.000175
Liquid products	0.000175
Absorption oils	0.00035–0.00053
Alkylation trace acid streams	0.00035
Reboiler streams	0.00035–0.00053

Source: Chenoweth, J., Final Report, HTRI/TEMA Joint Committee to Review the Fouling Section of TEMA Standards, HTRI, Alhambra, CA, 1988. With permission.

[a] Depending on charge characteristics and storage history, charge fouling resistance may be many times this value.

[b] American Petroleum Institute.

Surface temperature. Higher surface temperatures promote chemical reaction, corrosion, crystal formation (with inverse solubility salts), and polymerization, but reduce biofouling for temperatures above the optimum growth, avoid potential freezing fouling, and avoid precipitation of normal-solubility salts. It is highly recommended that the surface temperature be maintained below the reaction temperature; it should be kept below 60°C for cooling tower water.

Tube material. The selection of the tube material is important from the corrosion point of view which in turn could increase crystallization and biological fouling. Copper alloys can reduce certain biofouling, but their use is limited by environmental concerns with river, ocean, and lake waters.

There are many other variables that affect fouling. It is beyond the scope here, but the reader may refer to TEMA (1988).

TABLE 4.1.15
Fouling Factors and Design Parameters for Finned Tubes in Fossil Fuel Exhaust Gases

Type of Flue Gas	Fouling Factor, m^2 K/W	Minimum Spacing between Fins, m	Maximum Gas Velocity to Avoid Erosion, m/s
	Clean Gas (Cleaning Devices Not Required)		
Natural Gas	0.0000881–0.000528	0.00127–0.003	30.5–36.6
Propane	0.000176–0.000528	0.00178	—
Butane	0.000176–0.000528	0.00178	—
Gas turbine	0.000176	—	—
	Average Gas (Provisions for Future Installation of Cleaning Devices)		
No. 2 oil	0.000352–0.000704	0.00305–0.00384	25.9–30.5
Gas turbine	0.000264	—	—
Diesel engine	0.000528	—	—
	Dirty Gas (Cleaning Devices Required)		
No. 6 oil	0.000528–0.00123	0.00457–0.00579	18.3–24.4
Crude oil	0.000704–0.00264	0.00508	—
Residual oil	0.000881–0.00352	0.00508	—
Coal	0.000881–0.00881	0.00587–0.00864	15.2–21.3

Source: Weierman, R.C., 1982. Design of Heat Transfer Equipment for Gas-Side Fouling Service, Workshop on an Assessment of Gas-Side Fouling in Fossil Fuel Exhaust Environments, W.J. Marner and R.L. Webb, Eds., JPL Publ. 82–67, Jet Propulsion Laboratory, California Institute of Technology, Pasadena. With permission.

Fouling Control and Cleaning Techniques

Control of fouling should be attempted first before any cleaning method is attempted. For gas-side fouling, one should verify that fouling exists, identify the sequential event that dominates the foulant accumulation, and characterize the deposit. For liquid-side fouling, fouling inhibitors/additives should be employed while the exchanger is in operation; for example, use antidispersant polymers to prevent sedimentation fouling, "stabilizing" compounds to prevent polymerization and chemical reaction fouling, corrosion inhibitors to prevent corrosion fouling, biocide/germicide to prevent biofouling, softeners, acids, and polyphosphates to prevent crystallization fouling.

If the foulant control is not effective, the exchanger must be cleaned either online or off-line. Online cleaning includes flow-driven brushes/sponge balls inside tubes, power-driven rotating brushes inside tubes, acoustic horns/mechanical vibrations for tube banks with gases, soot blowers, and shutting off of the cold gas supply, flowing hot gas, or reversing of the fluids. Off-line cleaning methods, without dismantling the exchanger include chemical cleaning (circulate acid/detergent solutions), circulating of particulate slurry (such as sand and water), and thermal melting of frost layers. Off-line cleaning with a heat exchanger opened includes high-pressure steam or water cleaning, and thermal baking of an exchanger and then rinsing for small heat exchanger modules removed from the container of the modular exchangers.

Nomenclature

A	Total heat transfer area (primary + fin) on one fluid side of a heat exchanger, A_p: primary surface area, A_f: fin surface area, m^2
A_{fr}	Frontal area on one side of an exchanger, m^2
A_k	Total wall cross-sectional area for heat conduction in fin or for longitudinal conduction in the exchanger, m^2

A_o	Minimum free-flow area on one fluid side of a heat exchanger, m^2
b	Plate spacing, $h'+\delta_f$, m
C	Flow stream heat capacity rate with a subscript c or h, $\dot{m}c_p$, W/°C
C^*	Heat capacity rate ratio, C_{min}/C_{max}, dimensionless
c_p	Specific heat of fluid at constant pressure, J/kg K
D_h	Hydraulic diameter of flow passages, $4A_oL/A$, m
d_e	Fin tip diameter of an individually finned tube, m
d_i, d_o	Tube inside and outside diameters, respectively, m
Eu	N-row average Euler number, $\Delta p/(\rho\, \mu_m^2\, N/2g_c)$, $\rho\Delta p g_c/(NG^2/2)$, dimensionless
F	Log-mean temperature difference correction factor, dimensionless
f	Fanning friction factor, $\rho\Delta p g_c D_h/(2LG^2)$, dimensionless
f_{tb}	Fanning friction factor per tube row for crossflow over a tube bank outside, $\rho\Delta p g_c/(2NG^2)$
G	Mass velocity based on the minimum free flow area, $\dot{m}/A_o$, kg/m^2s
g	Gravitational acceleration, m^2/s
g_c	Proportionality constant in Newton's second law of motion, $g_c=1$ and dimensionless in SI units
H	Fin length for heat conduction from primary surface to either fin tip or midpoint between plates for symmetric heating, m
h	Heat transfer coefficient, W/m^2 K
h'	Height of the offset strip fin (see Figure 4.1.36), m
j	Colburn factor, $NuPr^{-1/3}/Re$, $StPr^{2/3}$, dimensionless
k	Fluid thermal conductivity, W/m K
k_f	Thermal conductivity of the fin material, W/m K
k_w	Thermal conductivity of the matrix (wall) material, W/m K
L	Fluid flow (core or tube) length on one side of an exchanger, m
l	Fin length for heat conduction from primary surface to the midpoint between plates for symmetric heating, see Table 4.1.11 for other definitions of l, m
l_f	Offset trip fin length or fin height for individually finned tubes, l_f represents the fin length in the fluid flow direction for an uninterrupted fin with $l_f=L$ in most cases, m
m	Fin parameter, 1/m
N	Number of tube rows, dimensionless
N_f	Number of fins per meter, 1/m
N_t	Total number of tubes in an exchanger, dimensionless
NTU	Number of heat transfer units, UA/C_{min}, it represents the total number of transfer units in a multipass unit, $NTU_s=UA/C_{shell}$, dimensionless
Nu	Nusselt number, hD_h/k, dimensionless
ntu_c	Number of heat transfer units based on the cold side, $(\eta_o hA)_c/C_c$, dimensionless
ntu_h	Number of heat transfer units based on the hot side, $(\eta_o hA)_h/C_h$, dimensionless
$\dot{m}$	Mass flow rate, kg/s
P	Temperature effectiveness of one fluid, dimensionless
$\wp$	Fluid pumping power, W
Pr	Fluid Prandtl number, $\mu c_p/k$, dimensionless
p	Fluid static pressure, Pa
Δp	Fluid static pressure drop on one side of heat exchanger core, Pa
P_f	Fin pitch, m
q	Heat duty, W
$\boldsymbol{q_e}$	Heat transfer rate (leakage) at the fin tip, W
q''	Heat flux, W/m^2
R	Heat capacity rate ratio used in the P-NTU method, $R_1=C_1/C$, $R_2=C_2/C_1$, dimensionless
R	Thermal resistance based on the surface area A, compare Equations 4.1.41 and 4.1.42 for definitions of specific thermal resistances, K/W
Re	Reynolds number, GD_h/μ, dimensionless

Re_d	Reynolds number, $\rho u_m d_o/\mu$, dimensionless
r_h	Hydraulic radius, $D_h/4$, $A_o L/A$, m
r_s	Fouling factor, $1/h_s$ m² K/W
St	Stanton number, h/Gc_p, dimensionless
s	Distance between adjacent fins, $p_f - \delta_f$, m
T	Fluid static temperature to a specified arbitrary datum, °C
T_a	Ambient temperature, °C
T_o	Fin base temperature, °C
T_l	Fin tip temperature, °C
U	Overall heat transfer coefficient, W/m² K
u_m	Mean axial velocity in the minimum free flow area, m/s
V	Heat exchanger total volume, m³
X_d	Diagonal tube pitch, m
X_l	Longitudinal tube pitch, m
X_t	Transverse tube pitch, m
α	Ratio of total heat transfer area on one side of an exchanger to the total volume of an exchanger, A/V, m²/m³
β	Heat transfer surface area density, a ratio of total transfer area on one side of a plate-fin heat exchanger to the volume between the plates on that side, m²/m³
ε	Heat exchanger effectiveness, it represents an overall exchanger effectiveness for a multipass unit, dimensionless
δ	Wall thickness, m
δ_f	Fin thickness, m
η_f	Fin efficiency, dimensionless
η_o	Extended surface efficiency, dimensionless
λ	Longitudinal wall heat conduction parameter based on the total conduction area, $\lambda = k_w A_{k,t}/C_{min}L$; $\lambda_c = k_w A_{k,c}/C_c L$, $\lambda_h = k_w A_{k,h}/C_h L$, dimensionless
μ	Fluid dynamic viscosity, Pa·s
ρ	Fluid density, kg/m³
σ	Ratio of free flow area to frontal area, A_o/A_{fr}, dimensionless

Subscripts

C	Clean surface value
c	Cold fluid side
F	Fouled surface value
f	Fin
h	Hot fluid side
i	Inlet to the exchanger
o	Outlet to the exchanger
s	Scale or fouling
w	Wall or properties at the wall temperature
1	One section (inlet or outlet) of the exchanger
2	Other section (outlet or inlet) of the exchanger

REFERENCES

Bhatti, M.S., and Shah, R.K. 1987. Turbulent and transition flow convective heat transfer in ducts. In *Handbook of Single-Phase Convective Heat Transfer*, eds. S. Kakaç, R. K. Shah, and W. Aung. New York: John Wiley & Sons, Chap. 4, 166 pp.

Chai, H.C. 1988. A simple pressure drop correlation equation for low finned tube crossflow heat exchangers. *Int. Commun. Heat Mass Transfer*, 15, 95–101.

Chenoweth, J. 1988. Final Report, HTRI/TEMA joint committee to review the fouling section of TEMA standards, HTRI, Alhambra, CA.

Cowell, T.A., Heikal, M.R., and Achaichia, A. 1995. Flow and heat transfer in compact louvered fin surfaces. *Exp. Therm. Fluid Sci.*, 10, 192–199.

Epstein, N. 1978. Fouling in heat exchangers. In *Heat Transfer 1978*, Vol. 6. New York: Hemisphere Publishing, 235–254.

Epstein, N. 1983. Thinking about heat transfer fouling: a 5×5 matrix. *Heat Transfer Eng*., 4(1), 43–56.

Foumeny, E.A., and Heggs, P.J. 1991. *Heat Exchange Engineering*, Vol. 2, *Compact Heat Exchangers: Techniques for Size Reduction*. London: Ellis Horwood Ltd.

Ganguli, A., and Yilmaz, S.B. 1987. New heat transfer and pressure drop correlations for crossflow over low-finned tube banks. *AIChE Symp. Ser.*, 257(83), 9–14.

Ghajar, A.J., and Tam, L.M. 1994. Heat transfer measurements and correlations in the transition region for a circular tube with three different inlet configurations. *Exp. Therm. Fluid Sci*., 8, 79–90.

Hesselgreaves, J.E. 2001. *Compact Heat Exchangers: Selection, Design and Operation*. Oxford: Pergamon.

Huang, L.J., and Shah, R.K. 1992. Assessment of calculation methods for efficiency of straight fins of rectangular profiles. *Int. J. Heat Fluid Flow*, 13, 282–293.

Idelchik, I.E. 1994. *Handbook of Hydraulics Resistance*, 3rd ed. Boca Raton, FL: CRC Press.

Kakaç, S., Ed. 1991. *Boilers, Evaporators, and Condensers*. New York: John Wiley & Sons.

Kakaç, S., Bergles, A.E., and Fernandes, E.O. 1988. *Two-Phase Flow Heat Exchangers: Thermal Hydraulic Fundamentals and Design*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Kakaç, S., Bergles, A.E., and Mayinger, F. 1981. *Heat Exchangers: Thermal-Hydraulic Fundamentals and Design*. Washington, DC: Hemisphere Publishing.

Kakac, S., Liu, H., and Pramuanjaroenkij, A. 2012. *Heat Exchangers: Selection, Rating and Thermal Design*, 3rd ed. Boca Raton, FL: CRC Press.

Kakaç, S., Shah, R.K., and Bergles, A.E. 1983. *Low Reynolds Number Flow Heat Exchangers*. Washington, DC: Hemisphere Publishing.

Kays, W.M., and London, A.L. 1998. *Compact Heat Exchangers*, 3rd ed. Malabar, FL: Krieger.

Klemes, J.J., Arsenyeua, O., Kapustenko, P., and Touazhnyanskyy, L. 2015. *Compact Heat Exchangers for Energy Transfer Intensification*. Boca Raton, FL: CRC Press.

Manglik, R.M., and Bergles, A.E. 1995. Heat transfer and pressure drop correlations for the rectangular offset-strip-fin compact heat exchanger. *Exp. Therm. Fluid Sci*., 10, 171–180.

Miller, D.S. 1990. *Internal Flow Systems*, 2nd ed., *BHRA* (*Information Services*). Bedford: Cranfield.

Mueller, A.C., and Chiou, J.P. 1987. *Review of Various Types of Flow Maldistribution in Heat Exchangers*, Book No. H00394, HTD-Vol. 75. New York: ASME, 3–16.

Putnam, G.R., and Rohsenow, W.M. 1985. Viscosity induced nonuniform flow in laminar flow heat exchangers. *Int. J. Heat Mass Transfer*, 28, 1031–1038.

Rabas, T.J., and Taborek, J. 1987. Survey of turbulent forced-convection heat transfer and pressure drop characteristics of low-finned tube banks in cross flow. *Heat Transfer Eng*., 8(2), 49–62.

Roetzel, W., Heggs, P.J., and Butterworth, D., Eds. 1991. *Design and Operation of Heat Exchangers*. Berlin: Springer-Verlag.

Rozenman, T. 1976. Heat transfer and pressure drop characteristics of dry cooling tower extended surfaces, Part I: Heat transfer and pressure drop data, Report BNWL-PFR 7-100; Part II: Data analysis and correlation, Report BNWL-PFR 7-102, Battelle Pacific Northwest Laboratories, Richland, WA.

Shah, R.K. 1981. Compact heat exchangers. In *Heat Exchangers: Thermal-Hydraulic Fundamentals and Design*, S. Kakaç, A.E. Bergles, and F. Mayinger, Eds. Washington, DC: Hemisphere Publishing, 111–151.

Shah, R.K. 1983. Heat exchanger basic design methods. In *Low Reynolds Number Flow Heat Exchanger*, S. Kakaç, R.K. Shah and A.E. Bergles, Eds. Washington, DC: Hemisphere, 21–72.

Shah, R.K. 1985. Compact heat exchangers. In *Handbook of Heat Transfer Applications*, 2nd ed., W.M. Rohsenow, J.P. Hartnett, and E.N. Ganic, Eds., Chap. 4, Part 3. New York: McGraw-Hill.

Shah, R.K. 1988a. Plate-fin and tube-fin heat exchanger design procedures. In *Heat Transfer Equipment Design*, R.K. Shah, E.C. Subbarao, and R.A. Mashelkar, Eds. Washington, DC: Hemisphere Publishing, 255–266.

Shah, R.K. 1988b. Counterflow rotary regenerator thermal design procedures. In *Heat Transfer Equipment Design*, R.K. Shah, E.C. Subbarao, and R.A. Mashelkar, Eds. Washington, DC: Hemisphere Publishing, 267–296.

Shah, R.K. 1991. Multidisciplinary approach to heat exchanger design. In *Industrial Heat Exchangers*, J.-M. Buchlin, Ed., Lecture Series No. 1991–04, von Kármán Institute for Fluid Dynamics, Rhode Saint Genèse, Belgium.

Shah, R.K. 1993. Nonuniform heat transfer coefficients for heat exchanger thermal design. In *Aerospace Heat Exchanger Technology 1993*, R.K. Shah and A. Hashemi, Eds. Amsterdam, Netherlands: Elsevier Science, 417–445.

Shah, R.K. 1994. Heat exchangers. In *Encyclopedia of Energy Technology and The Environment*, A. Bision and S.G. Boots, Eds. New York: John Wiley & Sons, 1651–1670.

Shah, R.K., Bell, K.J., Mochizuki, S., and Wadekar, V.V., Eds., 1997. *Compact Heat Exchangers for the Process Industries*. New York: Begell House.
Shah, R.K., and Bhatti, M.S. 1987. Laminar convective heat transfer in ducts. In *Handbook of Single-Phase Convective Heat Transfer*, S. Kakaç, R.K. Shah, and W. Aung, Eds. New York: John Wiley & Sons, Chap. 3, 137 pp.
Shah, R.K., and Bhatti, M.S. 1988. Assessment of correlations for single-phase heat exchangers. In *Two-Phase Flow Heat Exchangers: Thermal-Hydraulic Fundamentals and Design*, S. Kakaç, A.E. Bergles, and E.O. Fernandes, Eds. Dordrecht: The Netherlands: Kluwer Academic Publishers, 81–122.
Shah, R.K., and Giovannelli, A.D. 1988. Heat pipe heat exchanger design theory. In *Heat Transfer Equipment Design*, R.K. Shah, E.C. Subbarao, and R.A. Mashelkar, Eds. Washington, DC: Hemisphere Publishing, 609–653.
Shah, R.K., and Hashemi, A., Eds. 1993. *Aerospace Heat Exchanger Technology.* Amsterdam: Elsevier Science.
Shah, R.K., Kraus, A.D., and Metzger, D.E., Eds., 1990. *Compact Heat Exchangers—A Festschrift for Professor A.L. London*. Washington, DC: Hemisphere.
Shah, R.K., and London, A.L. 1978. Laminar flow forced convection in ducts, Suppl. 1 to *Advances in Heat Transfer.* New York: Academic Press.
Shah, R.K., and Mueller, A.C. 1988. Heat exchange. In *Ullmann's Encyclopedia of Industrial Chemistry*, *Unit Operations II*, vol. B3, Chap. 2, 108 pages. Weinheim, Germany: VCH.
Shah, R.K., and Pignotti, A. 1997. The influence of a finite number of baffles on the shell-and-tube heat exchanger performance. *Heat Transfer Eng.*, 18, 82–94.
Shah, R. K., and Sekulic, D. P. 2003. *Fundamentals of Heat Exchanger Design*. New York: Wiley.
Shah, R.K., Subbarao, E.C., and Mashelkar, R.A., Eds. 1988. *Heat Transfer Equipment Design*. Washington, DC: Hemisphere Publishing.
Shah, R.K., and Wanniarachchi, A.S. 1991. Plate heat exchanger design theory, in *Industrial Heat Exchangers*, J.-M. Buchlin, Ed., Lecture Series No. 1991–04, von Kármán Institute for Fluid Dynamics, Rhode Saint Genèse, Belgium.
Taylor, M.A. 1987. *Plate-Fin Heat Exchangers: Guide to Their Specificaitons and Use*, 1st ed., HTFS, Harwell Laboratory, Oxon, rev. 1990.
TEMA, 1988. *Standards of the Tubular Exchanger Manufacturers Association*, 7th ed. New York: Tubular Exchanger Manufacturers Association.
Thulukkanam, K. 2013. *Heat Exchanger Design Handbook*, 2nd ed. Boca Raton, FL: CRC Press.
Webb, R.L. 1994. *Principles of Enhanced Heat Transfer.* New York: John Wiley & Sons.
Weierman, R.C. 1982. Design of heat transfer equipment for gas-side fouling service, *Workshop on an Assessment of Gas-Side Fouling in Fossil Fuel Exhaust Environments*, W.J. Marner and R.L. Webb, Eds., JPL Publ. 82–67, Jet Propulsion Laboratory, California Institute of Technology, Pasadena.
Zohuri, B. 2016. *Compact Heat Exchangers: Selection, Application, Design and Evaluation*. New York: Springer.
Zukauskas, A. 1987. Convective heat transfer in cross flow. In *Handbook of Single-Phase Convective Heat Transfer*, S. Kakaç, R.K. Shah, and W. Aung. New York: John Wiley & Sons, Chap. 6.

Further Information

Heat exchangers play a crucial and dominant role in many developments related to energy conservation, recovery, utilization, economic development of new energy sources, and environmental issues such as air and water pollution control, thermal pollution, waste disposal, etc. Many new and innovative heat exchangers have been developed for these and many other applications worldwide. A broad overview is provided for various heat exchangers and basic design theory for single-phase heat exchangers. For further details and study, the reader may refer to the following references: Kakaç et al. (1981, 1983, 1988, 2012), Taylor (1987), Shah et al. (1988, 1990), Shah and Sekulic (2003), Foumeny and Heggs (1991), Kakaç (1991), Roetzel et al. (1991), Shah and Hashemi (1993), Shah et al. (1997), Hesselgreaves (2001), Klemes et al. (2015), Thulukkanam (2013) and Zohuri (2016).

4.2 APPLICATION OF NANOFLUIDS IN HEAT EXCHANGERS: PERFORMANCE AND CHALLENGES

Bengt Sundén and Zan Wu

Introduction

The focus of this section is performance and challenges of nanofluids in heat exchanger applications. Heat exchangers are equipment being used for transfer of heat between two or more fluids at different temperatures (Sundén, 2012). Heat exchangers play an important role in the reduction of greenhouse gas emissions and for achieving clean and sustainable development. They are widely used in many applications including power plants, automotives, space heating, refrigeration and air-conditioning systems, chemical plants, petrochemical processes, electronic cooling, environment engineering, etc. (Sundén and Wu, 2015). Due to advancement in manufacturing processes, various heat-transfer enhancement techniques have been routinely used in different types of heat exchangers to improve their thermal efficiency. Heat-transfer enhancement techniques can be classified either as passive (no external power needed) or as active (external power required). Passive techniques employ special surface geometries (e.g., fins, corrugations, insert devices, curved tubing, and surface coatings) or additives (e.g., particles, phase change materials, and surfactants) for heat-transfer enhancement. Interested readers can refer to Webb and Kim (2005) for more details about these enhancement techniques.

Due to recent progress in nanotechnology, thermal conductivity values can be increased by adding nanometer-sized structures (e.g., nanoparticles, nanofibers, and nanotubes) in conventional heat-transfer fluids to form the so-called nanofluids. As shown in Table 4.2.1, conventional heat-transfer fluids such as water, ethylene glycol, and engine oil have relatively low-thermal conductivity values, which thus limit the heat-transfer rates. The thermal conductivity of these fluids can be enhanced by adding solid particles of high thermal conductivities. There are several advantages by using nanoparticles instead of microparticles. Nanoparticles are more stable and stay suspended much longer than microparticles if well prepared, with relatively little particle settling, channel erosion, and clogging. The nanoparticle settling velocity due to gravity can be obtained from a balance of the buoyancy and viscous forces (Buongiorno, 2006):

$$\frac{\pi}{6} d_p^3 \left(\rho_p - \rho_f\right) g = 3\pi d_p u_p \mu \Rightarrow u_p = \frac{d_p^2 \left(\rho_p - \rho_f\right) g}{18\mu}, \tag{4.2.1}$$

where the viscous resistance is calculated by the Stokes law. The settling velocity u_p is $<1.6 \times 10^{-8}$ m/s for nanoparticles with size <100 nm. Besides, the surface area per unit volume of nanoparticles is much larger than that of microparticles, allowing them to absorb and transfer heat more efficiently. Furthermore, there might be additional nanoparticle/fluid slip mechanisms in nanofluids, such as Brownian motion and thermophoresis, which can lead to further heat-transfer enhancement. Nanofluids have promising potential in nanomedicine, electronics, chemistry, optics, and various transport processes.

Nanoparticle additions increase the thermal conductivity, which is beneficial for convective heat transfer. Viscosity is also increased in nanofluids, which is not favorable especially taking the pumping power into account. In general, the increase in viscosity is higher than the increase in thermal conductivity.

Convective heat-transfer performance of nanofluids in straight smooth tubes has been extensively investigated, for example, Xuan and Li (2003), Ding et al. (2006), Heris et al. (2007), Williams et al. (2008), Yu et al. (2008), Choi (2009), Bianco et al. (2011), Haghighi et al. (2014), and Huang et al. (2015). However, there is considerable disagreement about whether addition of nanoparticles could enhance the convective heat transfer (Sergis and Hardalupas, 2011). On one hand, it is very

TABLE 4.2.1
Thermal Conductivity of Typical Materials at 300 K (Han, 2008; Pop et al., 2012)

Solids/Liquids	Material	Thermal Conductivity (W m^{-1} K^{-1})
Metallic solids	Silver	429
	Copper	401
	Gold	318
	Aluminum	237
Nonmetallic solids	Suspended graphene without substrate	Up to 4000
	Diamond	Up to 3300
	Carbon nanotube	Up to 3000
	Silicon	148
	Alumina (Al_2O_3)	40
Conventional heat-transfer fluids	Water	0.611
	Ethylene glycol (EG)	0.258
	Engine oil	0.145

important to choose appropriate performance comparison criteria (i.e., figures of merit) to compare the performance of nanofluids over their base fluids. Inappropriate comparison criteria such as heat-transfer coefficient ratios based on the constant Reynolds number may give misleading results because the net result for the constant Reynolds number basis is a combination of the nanofluid property effect and the flow velocity effect (Wu et al., 2013; Vanapalli and ter Brake, 2013; Sundén and Wu, 2015). Due to the higher viscosity of the nanofluid, the flow velocity in the nanofluid is generally higher than that of the base fluid at the same Reynolds number, which provides an advantage for the nanofluid over the base fluid. Therefore, more appropriate comparison criterion such as comparison of heat-transfer enhancement based on equal flow velocity and equal pumping power should be used to evaluate the benefit of nanofluids over their base fluids.

On the other hand, as indicated in Prabhat et al. (2011) and Wu and Sundén (2016), there seems no anomalous heat-transfer enhancement in the fully developed laminar flow and turbulent flow. Previous single-phase correlations can accurately reproduce the fully developed convective heat-transfer behavior of nanofluids in tubes by adopting the measured temperature- and loading-dependent thermal conductivities and viscosities of the nanofluids in the analysis. However, anomalous heat transfer could be obtained during the developing laminar flow (Liu, 2012).

So far, there are few studies on heat-transfer characteristics of nanofluids in complex geometries (e.g., helically coiled tubes and corrugated plates) and heat exchangers. Seyf and Mohammadian (2011) numerically simulated laminar nanofluid flow in a counterflow microchannel heat exchanger and suggested that nanofluid can enhance the heat exchanger effectiveness by nanoparticle Brownian motion. A single-phase approach was used for nanofluid modeling, and arbitrary thermal conductivity and viscosity values were adopted in the simulation. Pandey and Nema (2012) experimentally observed higher heat-transfer coefficients of Al_2O_3/water nanofluids than water in a corrugated plate heat exchanger (PHE). Heat-transfer coefficient decreases when the nanoparticle volume concentration increases from 2.0% to 4.0%. The pressure drop of nanofluids increased with increasing nanoparticle volume concentrations and was higher than that of water. Prasad et al. (2015) experimentally studied turbulent convective heat transfer and friction of alumina/water nanofluids flowing through a concentric tube U-bend heat exchanger with and without helical tape inserts. In these works, usually the thermal performance of nanofluids was evaluated by comparing the Nusselt number or the heat-transfer coefficient based on the Reynolds number. However, this comparison criterion is misleading. Besides, property correlations were used in most of these experimental and numerical studies rather than the measured nanofluid properties, which confuses researchers further to get a uniform opinion.

To provide a clearer view of the application of nanofluids in heat exchangers, our previous experimental investigations on the thermal performance of various nanofluids in different types of heat exchangers have been presented in this chapter. As the thermal performance of nanofluids is dependent on the thermophysical properties, nanofluid preparation, characterization, and property correlations have also been reported.

Nanofluid Preparation and Characterization

Preparation of nanofluids has an important influence on nanofluid thermophysical properties and stability. There are two main types of methods to produce nanofluids: single-step method and two-step method. The single-step method is a process that simultaneously makes and disperses the nanoparticles into base fluids, for which the nanoparticles are directly prepared by physical vapor deposition or a liquid chemical method. In the single-step method, the processes of drying, storage, transportation, and dispersion of nanoparticles are avoided, so the agglomeration of nanoparticles is minimized and the stability of fluids is increased (Li et al., 2009). To prevent oxidation of metallic particles, the single-step method is always preferable for synthesizing nanofluids containing metal particles. One disadvantage of the single-step method is that only low-vapor-pressure fluids are compatible with the process. Another disadvantage is that the single-step method is difficult to scale-up due to high cost and its demand for a vacuum.

The two-step method is a process that first generates nanoparticles and subsequently disperses the nanoparticles into base liquids. Nanoparticles, nanofibers, or nanotubes used in the two-step method are produced as a dry powder by suitable techniques such as inert gas condensation, chemical vapor deposition, and mechanical alloying. An advantage of the two-step method is that the inert gas condensation technique has already been scaled up to commercial nano-powder production (Wang and Fan, 2010). Besides, the particle concentration and size distribution could be manipulated. A deficiency of the two-step method is particle agglomeration during drying, storage and transportation of nanoparticles, and dispersion of nanoparticles in the base fluids. The nanoparticles tend to form agglomerations of different size due to the van der Waals attractive forces. Severe agglomeration might not only result in particle settlement and channel clogging, but also alter the thermal properties. Addition of surfactants, pH adjustment, and ultrasonic or microwave vibrations can reduce the size of the agglomerates substantially but are not able to break the agglomerates into primary particles.

Characterization of nanofluids includes colloidal stability, particle agglomerations, size distribution, and nanofluid properties (e.g., thermal conductivity and viscosity). One very easy and rough way to detect nanofluid stability is to see if there is any sedimentation. More accurately, the nanofluid stability can be characterized by the microscale agglomerations, size distributions, or property changes. For example, Figure 4.2.1 shows the change of the averaged thermal conductivity for each day with elapsed time for an alumina/water nanofluid with a nanoparticle volume fraction of 7.82% at 20°C. In general, thermal conductivity decreases with elapsed time. A 7.0% reduction in thermal conductivity averaged for each day was observed after 55 days. Therefore, the nanofluid is unstable with elapsed time even if the nanofluid looks nearly the same to the naked eye. It suggests that the microscale agglomerations in the nanofluid change with elapsed time.

Transmission electron microscopy (TEM) can be used to verify particle dimensions and to identify agglomerations. The electron beam can be used to see features on the nanometer level. One major disadvantage to use TEM is that the nanoparticle samples must be dried out of the nanofluids in order to be attached to the carbon matrix for observation under vacuum conditions. There is always a risk that the nanoparticles might have agglomerated in a different way when it has dried. One alternative is to cryo-plunge nanoparticles with the nanofluids in the aqueous state down to liquid nitrogen temperature. The resolution becomes lower by this method. For the two-step nanofluid preparation, TEM images can be obtained for the dry powder produced by

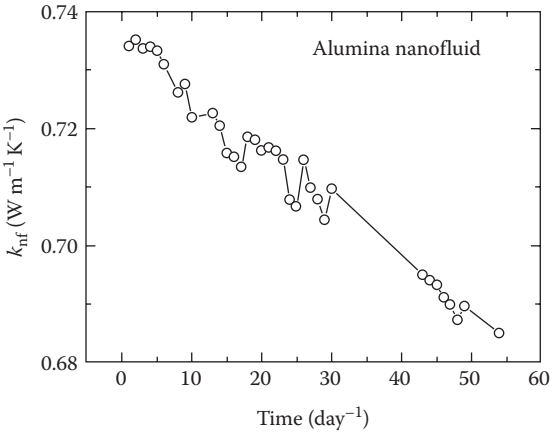


FIGURE 4.2.1 Change in thermal conductivity averaged for each day with elapsed time for an alumina/water nanofluid (7.82% volume fraction) at 20°C. (From Wu, Z. et al., *Front. Heat Mass Transfer*, 5, 18, 2014.)

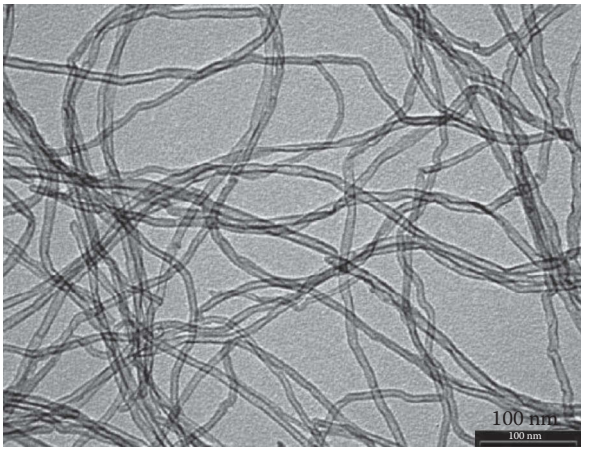


FIGURE 4.2.2 TEM characterization of MWCNTs. (From Wu, Z. et al., *Appl. Therm. Eng.*, 96, 364–371, 2016.)

the first step. Figure 4.2.2 shows an example of TEM characterization of multiwalled carbon nanotubes (MWCNTs). Apparently, the MWCNTs are aggregated and entangled before dispersion into the base fluid. Please note that the microscale structures formed by nanoparticles in the nanofluid will be different from dry powder observations.

The dynamic light scattering (DLS) technique can be used to observe the effective particle size distribution in a dilute nanofluid. When light from a laser passes through the nanofluid, the light will scatter uniformly from particles with a size less than about 250 nm. The particle size can be determined by the time measurements of the scattered light intensity. The time variation of the intensity can be correlated to the hydrodynamic radius of the particles in the nanofluid. For example, Figure 4.2.3 shows the particle size distributions by DLS of two MWCNT/water nanofluids of weight concentrations 0.05 and 0.1 wt%, respectively. The two nanofluids give similar particle size distributions and have a peak, about 200 nm for 0.05 wt% and 220 nm for 0.1 wt%. One drawback of DLS is that it can only be used for dilute nanofluids. Special care must be taken when diluting the nanofluids, for example, to maintain the nanofluid pH stable during dilution. Another drawback of DLS is that the size measurement of very polydisperse systems might be misleading.

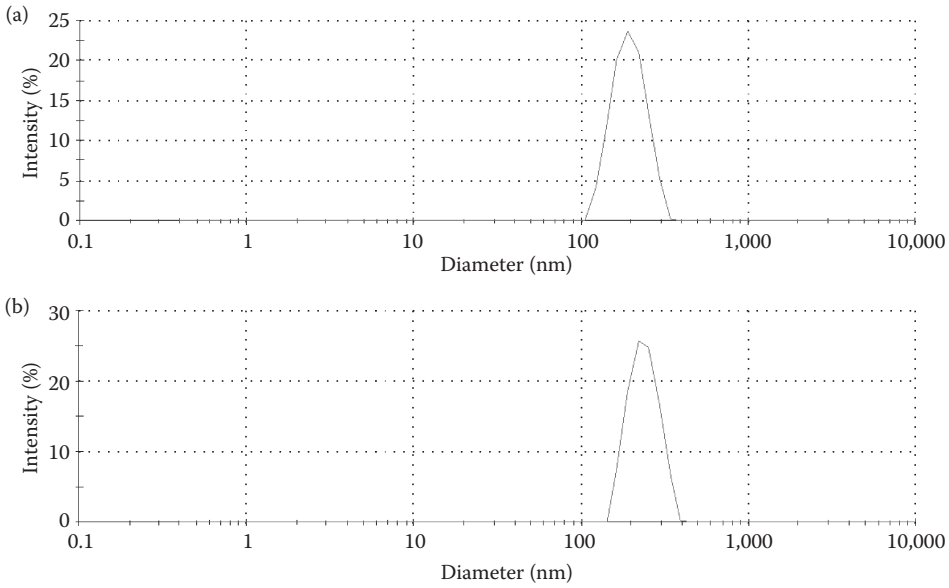


FIGURE 4.2.3 The particle size distributions by DLS of two MWCNT/water nanofluids: (a) 0.05 wt% and (b) 0.1 wt%. (From Wu, Z. et al., *Appl. Therm. Eng.*, 96, 364–371, 2016.)

Shear viscosities of the nanofluid can be measured using rotational viscometers/rheometers of cone and plate geometries. Rotational viscometers use the idea that the force required to turn an object in a fluid can indicate the viscosity of that fluid. Thermal conductivities of the nanofluid can be measured by a transient hot wire method or a transient plane source method. Please refer to Healy et al. (1976) and Gustafsson (1991) for detail descriptions of the transient hot wire technique and the transient plane source technique, respectively.

Properties of Nanofluids

Heat-transfer performance of nanofluids is mainly affected by the physical properties of nanofluids. Physical properties, such as density, specific heat, viscosity, and thermal conductivity, are decided by nanoparticles (e.g., particle size and shape, concentration, and material), the base fluid (e.g., fluid type, additives), and preparation methods, as shown in Figure 4.2.4. Property characterization of nanofluids has received intensive attention, especially viscosity and thermal conductivity. Please refer to reviews such as Fan and Wang (2011), Kleinstreuer and Feng (2011), and Taylor et al. (2013) for more details. It is generally accepted that the density of the nanofluid can be calculated by the mixing theory (volume-weighted average).

$$\rho_{nf} = (1-\varphi)\rho_f + \varphi\rho_p, \tag{4.2.2}$$

where φ is the nanoparticle volume concentration. Subscripts nf, f, and p indicate the nanofluid, the base fluid, and the particle, respectively. Properties of surfactant additives should be considered when evaluating the properties of the base fluid.

The specific heat of the nanofluid cannot be calculated from the volume-weighted average as the volume-weighted average deviated very significantly from the experimental data (O'Hanley et al., 2012). Assuming that the base fluid and the nanoparticles are in thermal equilibrium, the following expression can predict the specific heat of nanofluids accurately:

$$\rho_{nf} c_{p,nf} = (1-\varphi)\rho_f c_{p,f} + \varphi\rho_p c_{p,p}. \tag{4.2.3}$$

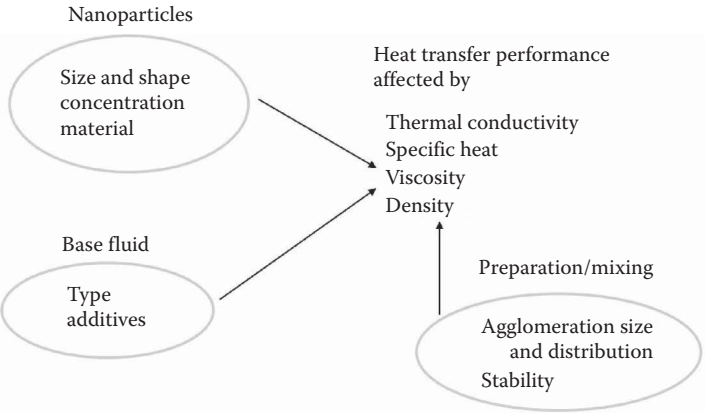


FIGURE 4.2.4 Properties of nanofluids.

Given its sound theoretical basis, this equation is generally applicable. Small discrepancies between the data and predictions might occur because properties (density and specific heat) may differ if the material is in nanoparticle form versus bulk form.

Up to now, there are no general correlations to predict the thermal conductivity and viscosity of nanofluid. Thermal conductivity and viscosity of nanofluids are strongly dependent on parameters such as particle concentration, particle size and shape, the presence of agglomerations (i.e., nanoclusters and aggregates), the nature of the base fluid, temperature, and nanofluid stability. Due to attractive van der Waals forces, nanoparticles tend to agglomerate in nanofluids to form the so-called aggregate-laden nanofluids.

Existence of nanoparticle agglomerations has already been recognized by DLS and SEM/TEM observations in the literature, for example, Keblinski et al. (2002), Hong et al. (2006), Prasher et al. (2006), Karthikeyan et al. (2008), and Anoop et al. (2009). Aggregation affects the nanofluid properties such as thermal conductivity and viscosity and further affects the heat-transfer performance. On one hand, agglomeration tends to enhance nanofluid viscosity due to the immobilized fluid trapped in the particle clusters and thus a higher effective volume fraction than the actual solid volume fraction. Anoop et al. (2009) considered the viscosity increase to be primarily due to the agglomeration of particles in water-based and ethylene-based nanofluids. On the other hand, fractal agglomerations or aggregates can lead to thermal conductivity enhancement due to the ability of the heat to move rapidly along the backbone of the clusters. Note that relatively large packed clusters may deteriorate thermal conductivity as it may induce sedimentation and therefore decrease the particle volume fraction.

Previous correlations for viscosity and thermal conductivity were given in the literature, for example, Chen et al. (2009), Philip and Shima (2012), and Wu et al. (2014). Classical models such as the Einstein model (Einstein, 1906), the Maxwell model (Maxwell, 1881), and the Hamilton–Crosser model (Hamilton and Crosser, 1962) might not be able to predict the nanofluid properties accurately. Most of the developed correlations for nanofluids are based on very limited experimental data, and mainly depend on the properties of the base fluid and the nanoparticle, particle volume concentration, temperature, and aspect ratio for nonspherical particles. Effects of the microscopic aggregates are not considered in most of these correlations. However, the microscale structures in the nanofluid affect macroscopic thermophysical properties, especially viscosity and thermal conductivity. Here we suggest correlations which include effects of aggregation and interfacial thermal resistance. The bulk effective dynamic viscosity for nanofluids including fractal aggregates can be calculated by the modified Krieger–Dougherty model (1959):

$$\frac{\mu_{nf}}{\mu_f} = \left(1 - \frac{\varphi_a}{\varphi_m}\right)^{-[\mu]\varphi_m} = \left(1 - \frac{\varphi/\varphi_{in}}{\varphi_m}\right)^{-[\mu]\varphi_m}, \tag{4.2.4}$$

where φ_m is the maximum particle packing fraction, which is 0.61 for rigid spherical particles at high shear rates. φ_{in} and φ_a denote the volume fraction of the particles in the aggregates and the volume fraction of the aggregates in the entire nanofluid, respectively. φ_{in} can be calculated by

$$\varphi_{in} = \left(\frac{d_a}{d_p}\right)^{D-3}. \tag{4.2.5}$$

The parameter [μ] is the intrinsic viscosity with a typical value of 2.5 for spherical nanoparticles. The fractal index D ranges from 2.4 to 1.8, depending on aggregate formations (Gharagozloo and Goodson, 2010). Normally, for spherical particles, D is 1.8. The normalized viscosity data of ethylene glycol (EG)-based and EG-water based SiO_2 nanofluids with reference to the viscosity of base fluids are presented in Figure 4.2.5. The well-known Einstein equation (Einstein, 1906) and the Batchelor equation (Batchelor, 1977) underestimate the nanofluid viscosity data by a large margin. Feng and Li (2013) showed in Figure 4.2.5 that the Krieger–Dougherty model can estimate their viscosity data relatively well if the values of φ_{in} are in the range of 5.5–16.5%.

The lower and upper limits of the nanofluid thermal conductivity can be completely determined by the serial mode and the parallel mode (Nielsen, 1978) only using volume fractions and thermal conductivities of the two phases, respectively. Hashin and Shtrikman (1963) proposed strict bounds based on the classical effective medium theory, given below for the case of $k_p/k_f > 1$:

$$k_f\left(1+\frac{3\varphi(k_p-k_f)}{3k_f+(1-\varphi)(k_p-k_f)}\right) \le k_{nf} \le k_p\left(1-\frac{3(1-\varphi)(k_p-k_f)}{3k_p-\varphi(k_p-k_f)}\right). \tag{4.2.6}$$

At the lower bound, nanoparticles are well-dispersed in the base fluid. The lower bound is basically applicable for dilute suspensions of well-dispersed noninteracting spherical particles. At the upper bound, the nanoparticles form a continuous phase with linear or chainlike particle morphologies,

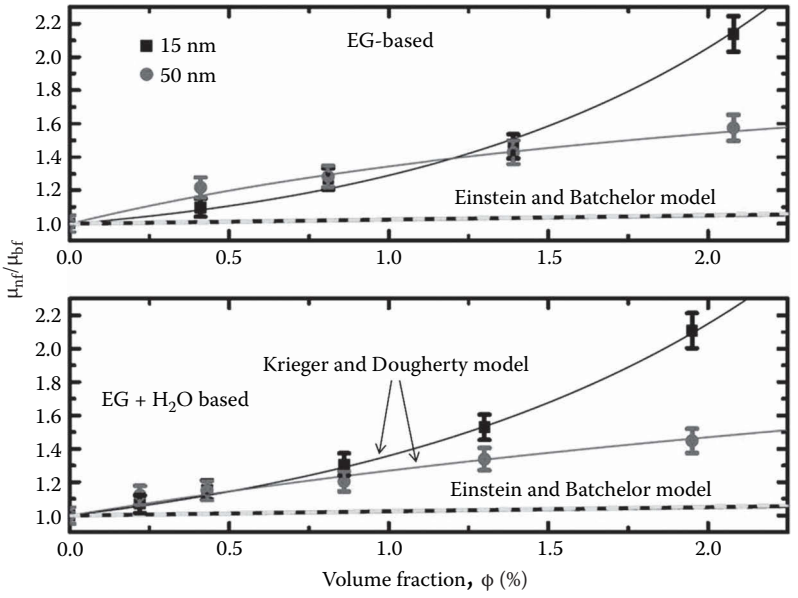


FIGURE 4.2.5 Evaluation of the Krieger–Dougherty model by EG-based and EG-water-based SiO_2 nanofluids at 20°C. The Einstein equation and the Batchelor equation are shown for comparison. (From Feng, Z.Z. and Li, W. *Int. J. Heat Mass Transfer*, 65, 919–927, 2013.)

and the base fluid becomes a dispersed phase. Figure 4.2.6 evaluates the lower and upper bounds and the Bruggeman model (Bruggeman, 1935) by the literature and experimental data of alumina/water nanofluids. The literature data from Kim et al. (2007), Timofeeva et al. (2007), and Williams et al. (2008) are shown for comparison. The primary diameter of the alumina nanoparticles is approximately 40 nm. As shown in Figure 4.2.6, considering the measurement uncertainty range, most of the data points are located near the lower bound. The thermal conductivity enhancement of Timofeeva et al. (2007) is even lower than the lower bound. One of the possible reasons might be particle cluster deposition, which decreases the nanoparticle volume fraction in the nanofluid and accordingly decreases the thermal conductivity. Figure 4.2.6 suggests that there is still a lot of room for thermal conductivity improvement by nanoparticles.

Besides the agglomeration morphology described earlier, liquid layering at the liquid–particle interface, Brownian motion, and coupled transport can also influence the thermal conductivity values (Fan and Wang, 2011). Using X-ray reflectivity, Mo et al. (2005) observed the presence of an ordered liquid layer near a nanoparticle surface by which the atomic structure of the liquid near the surface is significantly more ordered than that of bulk liquid. The thermal conductivity of ordered liquid layers tends to be larger than that of the bulk base fluid. Therefore, ordered liquid layers may enhance the effective thermal conductivity of the nanofluid by augmenting the particle effective volume fraction. Wang and Fan (2011) suggested the liquid layers offer insignificant conductivity enhancement for water-based nanofluids containing spherical particles as the liquid film thickness is only 0.28 nm for aqueous nanofluids. For nanoparticle size of the order of 10 nm, the increase in effective volume fraction induced by ordered liquid layers is only 0.1%, which contributes little. However, their presence may facilitate formation of interconnected particle agglomerations by relaxing the requirement of particle physical contact with each other. In addition, the liquid layering at the liquid–particle interface may present interfacial thermal resistance at the particle–fluid interface and among the particle–particle interface due to different nature of thermal conduction in nanoparticles and the base fluid, and thus decrease thermal conductivity enhancement.

Brownian motion can enable direct interparticle transport of heat from one particle to another and induce surrounding fluid flow and thus so-called micro-convection. The ratio of the contribution to thermal conductivity by Brownian motion and micro-convection (k_{BM}) to the base fluid conductivity (k_f) was estimated based on the kinetic theory by Evans et al. (2006).

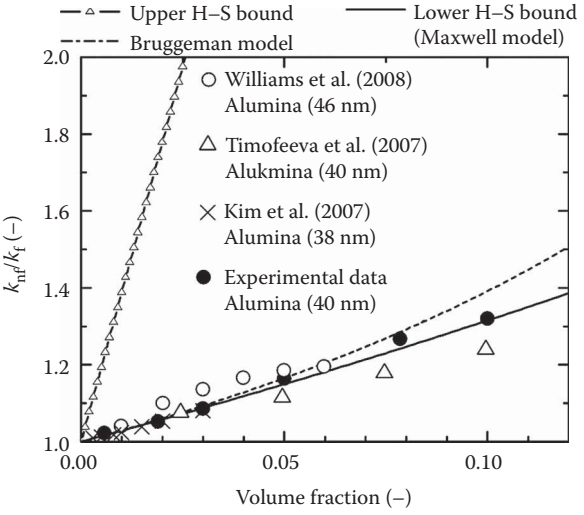


FIGURE 4.2.6 Evaluation of the lower and upper H–S bounds (Hashin and Shtrikman, 1963) and the Bruggeman model (Bruggeman, 1935) by the literature and experimental data (Wu et al., 2014) for alumina/water nanofluids.

$$\frac{k_{BM}}{k_f} = \frac{\rho_p c_{p,f} \varphi k_B T}{3\pi\mu_f d_p k_f} + \frac{k_B T}{3\pi\mu_f d_p \alpha_f}. \tag{4.2.7}$$

The contribution of Brownian motion of nanoparticles is not always important. For example, for alumina/water nanofluids with an average nanoparticle diameter of 40 nm, the contribution of Brownian motion and micro-convection can be negligible at room temperatures according to the abovementioned equation. Brownian motion becomes more important at higher temperatures for smaller particles. However, for nanofluids with surfactant additives, the operating temperature of the nanofluid is generally limited because the surfactants lose their effectiveness at high temperatures (e.g., 60°C).

Eapen et al. (2010) believed that both the Soret effect (also known as thermodiffusion or thermophoresis) and the Dufour effect (an induced heat flow caused by the concentration gradient) do not directly influence the nanofluid thermal conductivity by analyzing the orders of magnitude, but the coupled or crossed transport between them or other transport processes might affect the thermal conductivity. Wang and Fan (2011) stated that the coupled transport could change the nature of heat conduction in nanofluids from a diffusion process to a wave process, thus giving a nanofluid thermal conductivity enhancement as high as 10%.

The nanofluid thermal conductivity can be calculated by the procedure proposed by Evans et al. (2008), as shown in Figure 4.2.7. The Evans et al. procedure was later modified by Wu and Sundén (2016) to consider the effects of interfacial thermal resistance. Figure 4.2.7 briefly illustrates the calculation procedure. Aggregate consists of a few approximately linear chains (black circles) and dead ends (gray circles). The linear chains that span the whole cluster are called the backbone. The other particles are called dead ends. To calculate the nanofluid thermal conductivity, the aggregate is decomposed into dead ends with the fluid and the backbone. Thermal conductivity of the aggregate with only dead end particles, denoted as k_{nc}, can be calculated by the modified Bruggeman model:

$$\begin{aligned}\frac{k_{nc}}{k_f} = {} & \frac{1}{4}\left[\left(-2+3\varphi_{in,nc}\right)\left(\frac{k_p}{k_f\left(1+\mathrm{Bi}\right)}-1\right)+\frac{k_p}{k_f\left(1+\mathrm{Bi}\right)}\right] \\ & +\frac{1}{4}\sqrt{\left[\left(-2+3\varphi_{in,nc}\right)\left(\frac{k_p}{k_f\left(1+\mathrm{Bi}\right)}-1\right)+\frac{k_p}{k_f}\right]^2+8\frac{k_p}{k_f\left(1+\mathrm{Bi}\right)}},\end{aligned} \tag{4.2.8}$$

where the nanoparticle Biot number, $\mathrm{Bi} = 2R_k k_f/d_p$, accounts for the interfacial thermal resistance R_k. $\varphi_{in,nc}$ indicates the volume fraction of the dead end nanoparticles. By assuming that the backbone

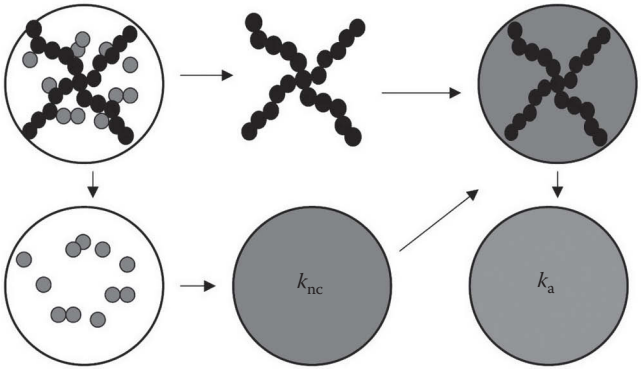


FIGURE 4.2.7 Illustrations of an aggregate and the calculation procedure for thermal conductivity proposed by Evans et al. (2008). (From Evans, W. et al., *Int. J. Heat Mass Transfer*, 51(5), 1431–1438, 2008.)

is embedded in a medium with an effective conductivity of k_{nc}, the effective thermal conductivity of the aggregate including both dead ends and linear chains can be calculated by the Nan et al. model (Nan et al., 2003).

$$\frac{k_a}{k_{nc}} = \frac{3+\varphi_{in,c}\left[2\beta_x\left(1-L_x\right)+\beta_z\left(1-L_z\right)\right]}{3-\varphi_{in,c}\left[2\beta_x L_x+\beta_z L_z\right]}, \tag{4.2.9}$$

where $\varphi_{in,c}$ is the volume fraction of the linear chains. Please refer to Evans et al. (2008) for details to determine $\varphi_{in,nc}$ and $\varphi_{in,c}$. The volume fraction of the nanoparticles in the aggregates (φ_{in}) is the addition of the volume fraction of the dead end particles in the aggregate ($\varphi_{in,nc}$) and the volume fraction of the linear chains in the aggregate ($\varphi_{in,c}$):

$$\varphi_{in} = \varphi_{in,nc} + \varphi_{in,c}. \tag{4.2.10}$$

The parameters β_x and β_z are calculated by

$$\beta_x = \frac{k_x - k_{nc}}{k_{nc} + L_x\left(k_p - k_{nc}\right)}, \tag{4.2.11}$$

$$\beta_z = \frac{k_z - k_{nc}}{k_{nc} + L_z\left(k_p - k_{nc}\right)}, \tag{4.2.12}$$

$$k_x = \frac{k_p}{1+\left(2+1/p\right)\mathrm{Bi}\cdot L_x k_p / k_f}, \tag{4.2.13}$$

$$k_z = \frac{k_p}{1+\left(2+1/p\right)\mathrm{Bi}\cdot L_z k_p / k_f}, \tag{4.2.14}$$

$$L_x = \frac{p^2}{2\left(p^2-1\right)} - \frac{p}{2\left(p^2-1\right)^{3/2}}\cosh^{-1} p \quad \text{for } p>1, \tag{4.2.15}$$

where p is the aspect ratio. For aggregate linear chains, the aspect ratio is given as $p = d_a/d_p$:

$$L_z = 1-2L_x. \tag{4.2.16}$$

Finally, the effective thermal conductivity of the whole nanofluid is calculated by the Maxwell–Garnet model (Levy and Stroud, 1997), as given below:

$$\frac{k_{nf}}{k_f} = \frac{k_a + 2k_f + 2\varphi_a\left[k_a - k_f\right]}{k_a + 2k_f - \varphi_a\left(k_a - k_f\right)}. \tag{4.2.17}$$

Besides, macroscopic properties such as viscosity and thermal conductivity can give some information about the microscopic aggregates. Specifically, from the modified Bruggeman model and the viscosity data, information of φ_{in} (i.e., the volume fraction of the particles in the aggregates) can be obtained. Based on φ_{in}, the modified Evans et al. model, and the thermal conductivity data, the interfacial thermal resistance can be estimated.

Properties of Nanofluids in Heat Exchangers

This part aims to briefly present our experimental investigations on heat transfer and pressure drop characteristics of nanofluids in complex geometries and heat exchangers. A schematic of the experimental setup is shown in Figure 4.2.8. Please refer to Wu et al. (2013) for a detailed description of the

test rig. It consists of two loops, for the cold and hot fluids, respectively. The hot water or nanofluid runs in the hot closed loop, while cold water is forced in the cold open loop. Water or nanofluid is heated in a 50-L reservoir by an imbedded electric heater of 6 kW fixed at the bottom of the reservoir. Each loop has two rotameters of small and large ranges for accurate flow rate measurement. A differential pressure transducer with an accuracy of ±0.075% of the set span was used to measure the pressure drop of the nanofluid side. All rotameters were calibrated for water and nanofluids of different type and concentrations at different temperatures by using a stopwatch and measuring cylinders. The inlet and outlet temperatures of the inner tube and the annulus were measured by four calibrated copper–constantan thermocouples with an accuracy of ±0.1 K, respectively. Three heat exchangers were tested, that is, one double-pipe helical heat exchanger and two chevron PHEs. The heat exchangers were thermally insulated to reduce heat losses to the ambient. Geometrical parameters of the helically coiled tube and the corrugated plate are detailed in Table 4.2.2. Uncertainties of all relevant parameters are listed in Table 4.2.3.

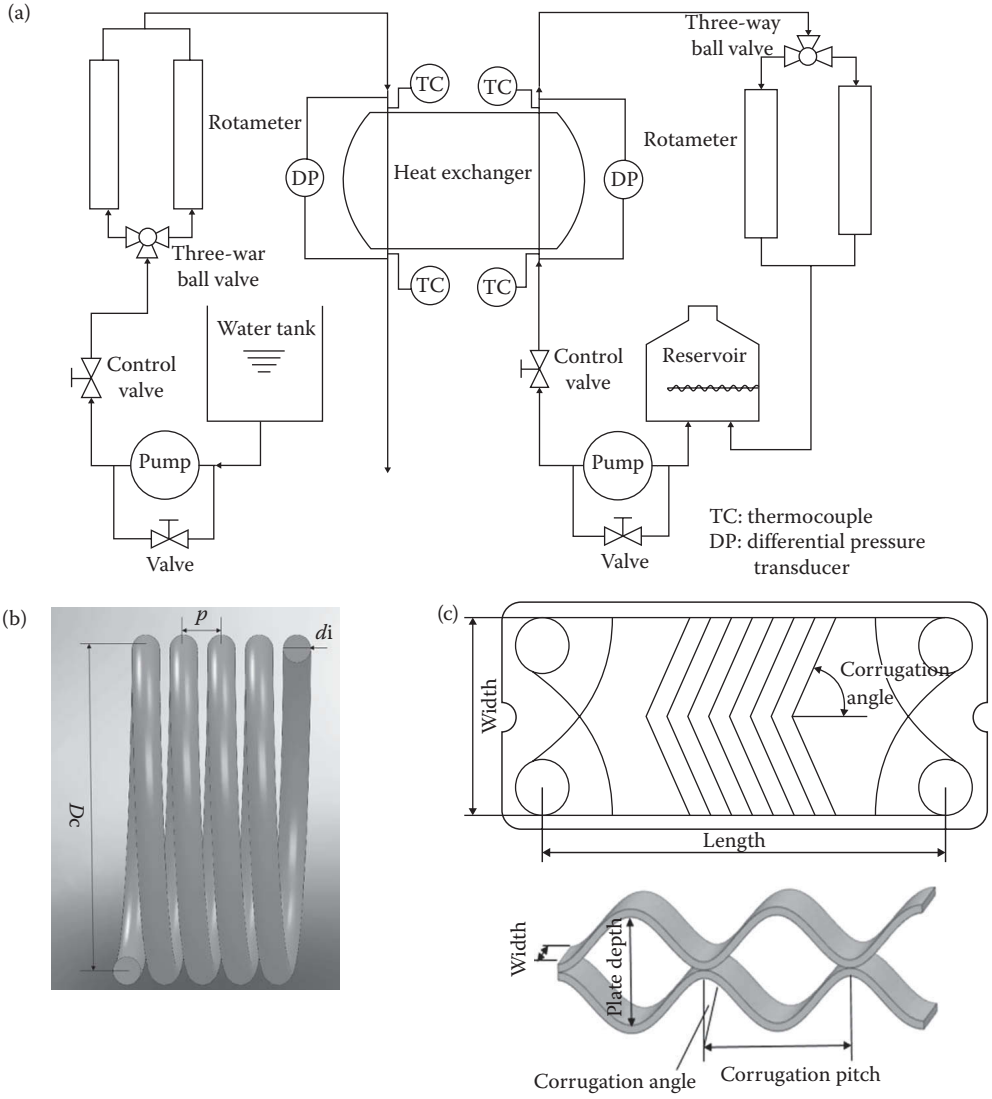


FIGURE 4.2.8 Schematics of (a) the test rig, (b) helically coiled tube, and (c) corrugated plate.

TABLE 4.2.2
Geometrical Parameters of the Tested Double-Pipe Helical Heat Exchanger and the Two Chevron Plate Heat Exchangers

Parameters	Values	
	Helical Heat Exchanger	
Inside diameter of the inner tube	13.28 mm	
Inside diameter of the outer tube	26.0 mm	
Number of turns	4.5	
Coil diameter of curvature (D_c)	254 mm	
Pitch of the helical coil (p)	34.5 mm	
Tube length	3.591 m	
	PHE I	**PHE II**
Plate length, L	0.154 m	0.154 m
Plate width, w	0.07 m	0.07 m
Plate depth, $2b_c$	0.004 m	0.00382 m
Plate thickness, δ	0.0003 m	0.00025 m
Heat-transfer area per plate	0.0132 m^2	0.0132 m^2
Corrugation pitch	0.006 m	0.007 m
Corrugation angle	60°	65°
Number of plates	20	10
Surface enlargement ratio	1.2375	1.22

TABLE 4.2.3
Uncertainties Estimation of Relevant Parameters

Primary measurements	
Diameter	±0.05 mm
Length/Width	±0.2 mm
Temperature	±0.1 K
Dependent quantities	
Mass flow rate m, kg s^{-1}	±2.0%
Heat flux q, W m^{-2}	±5.2%
LMTD, K	±1.5%
Apparent Darcy friction factor f_{app}	±3.5%
Heat-transfer coefficient h, W m^{-2} K^{-1}	±5.4%

Three types of nanofluids were tested, that is, γ-Al_2O_3/water nanofluids of six concentrations, multiwalled carbon nanotube (MWCNT)/water nanofluids of three concentrations, and a hybrid γ-Al_2O_3 + MWCNT/water nanofluid. Untreated concentrated γ-Al_2O_3/water nanofluid with spherical alumina nanoparticles of 40-nm mean diameter was purchased from a commercial company (Nanophase Technologies Corporation, US). No surfactants were added in the nanofluid. Different amounts of concentrated nanofluid were diluted in water to obtain nanofluids with low concentrations. The diluted nanofluid mixture was mechanically stirred for 0.5 h followed by ultrasonic vibration for 3 h. The pH value of the prepared alumina nanofluid is about 3–3.5, which is far away from the isoelectric point of alumina nanofluid to maintain colloidal stability.

An aqueous MWCNT suspension of 1.0% mass fraction was purchased from a commercial company (Nanocyl, Belgium). The suspension consists of thin MWCNTs dispersed in deionized (DI) water (97% mass fraction) and surfactant sodium dodecyl benzene sulfonate (SDBS, 2.0% mass fraction).

TABLE 4.2.4
Measured Thermal Conductivity and Viscosity of the Ten Nanofluids

Nanofluids	Volume Concentration (Weight Concentration)	Thermal Conductivity Enhancement (%)	Viscosity Enhancement (%)
γ-Al_2O_3/water nanofluid	0.2 vol% (0.78 wt%)	0.96	0.15
	0.56 vol% (2.18 wt%)	2.25	0.59
	1.02 vol% (3.89 wt%)	3.25	5.68
	1.5 vol% (5.68 wt%)	4.31	11.11
	1.88 vol% (7.04 wt%)	5.16	13.68
	2.84 vol% (10.56 wt%)	8.1	20
MWCNT/water nanofluid	0.0111 vol% (0.02 wt%)	0.07	12.99
	0.0278 vol% (0.05 wt%)	0.2	15.87
	0.0555 vol% (0.1 wt%)	0.42	20.66
Hybrid Al_2O_3+MWCNT/water nanofluid	1.89 vol% Al_2O_3/water + 0.0111 vol% MWCNT/water at a volume ratio of 1:2.5	4.4	10.8

The MWCNT, produced via the catalytic carbon vapor deposition process, has an average length of 1.5 μm and an average diameter of 9.5 nm, with an average aspect ratio of 158. The surface area of the MWCNT is 250–300 m^2/g. The carbon purity of the MWCNTs is 90%, while the remaining 10% is metal oxide. Similar to alumina nanofluid, different amounts of concentrated nanofluid were diluted in water to obtain MWCNT/water nanofluids with different fractions. The diluted nanofluid mixture was mechanically stirred for 10 min followed by ultrasonic vibration for 1 h. The pH values of the prepared MWCNT/water nanofluids are in the range 7.0–8.0.

The hybrid nanofluid mixture was obtained by mixing the 1.89 vol.% Al_2O_3/water nanofluid and the 0.0111 vol.% MWCNT/water nanofluid at a volume ratio of 1:2.5 and ultrasonically vibrated for 3 h. The equivalent volume concentration of Al_2O_3 nanoparticles in the hybrid nanofluid mixture is 1.48%. Each prepared nanofluid was immediately tested in the heat exchanger within 2 days after the nanofluid preparation.

The thermal conductivity (λ) and viscosity (μ) of the 10 nanofluids were experimentally measured. A thermal constants analyzer (TPS 2500S from Hot Disk AB, Sweden) using the transient plane source method (TPS) was employed to measure the thermal conductivity. The thermal conductivity uncertainty was estimated, from the standard deviations of experimental data and departures from literature data, to be lower than 3.0%. A rotational rheometer HAAKE RS6000 (Thermo Fisher Scientific Inc., US) was used to measure the rheology behavior of those nanofluids. The standard deviation of the measured dynamic viscosity data of water (the base fluid) and departures from the literature data is <3.0%. The measured thermal conductivity and viscosity data of the nanofluids are listed in Table 4.2.4.

Results of Nanofluids in a Double-Pipe Helical Heat Exchanger

Figure 4.2.9 shows the relationship between the apparent Darcy friction factor f_{app} and the Reynolds number Re_b for water and eight nanofluids. The subscript "b" indicates that the properties were calculated at the average bulk temperature. The average bulk temperature was estimated from the inner-tube inlet and outlet temperatures. The apparent Darcy friction factor was calculated by the following equation:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{4.2.18}$$

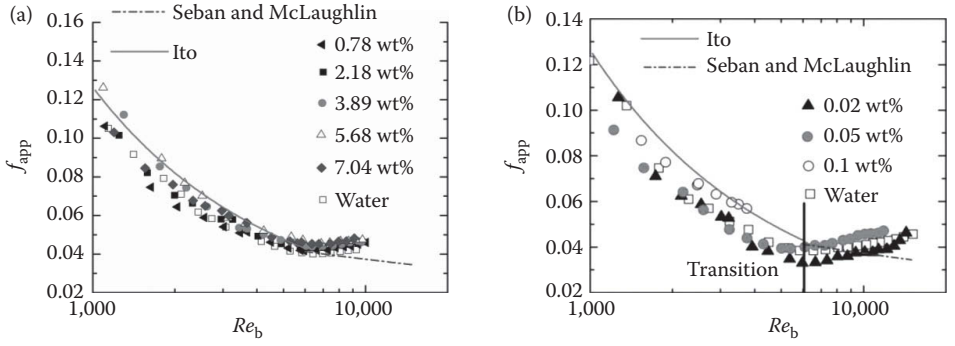


FIGURE 4.2.9 The f_{app}–Re_b relationship for water and the tested alumina/water and MWCNT/water nanofluids. The Ito (1969) equation and the Seban and McLaughlin (1963) equation were drawn for comparison. (a) Alumina nanofluid and (b) MWCNT nanofluid. (Data from Wu, Z. et al., *Appl. Therm. Eng.*, 60(1), 266–274, 2013; Wu, Z. et al., *Appl. Therm. Eng.*, 96, 364–371, 2016.)

The apparent friction factor decreases with Re_b when $Re_b < 6000$, while it increases slowly when $Re_b > 6000$. The transitional Reynolds number from laminar flow to turbulent flow is approximately 6000 for water and the nanofluids. Therefore, the transitional velocity of the nanofluids will be larger than that of water due to the larger viscosity of the former compared to the latter. It suggests that nanoparticles might stabilize the flow in helically coiled tubes. No obvious difference exists among the five alumina nanofluids, especially in laminar flow. For MWCNT/water nanofluid in laminar flow, the apparent friction factor for the 0.1 wt% nanofluid is the largest among the three tested MWCNT nanofluids mainly due to its higher increase in viscosity. As shown in Figure 4.2.9, the Ito (1969) equation and the Seban and McLaughlin (1963) equation can predict the experimental data relatively well for laminar flow and turbulent flow, respectively. For turbulent flow, the Seban and McLaughlin equation predicts a decreasing trend with Reynolds number, while our data points show an increasing trend. A possible reason might be that the data points are still located in the transition region. Besides, the Seban and McLaughlin equation also underestimates the friction factor for water in the previous literature for turbulent flow.

Figure 4.2.10 demonstrates the relationship between $Nu_b(Pr_b)^{-0.4}$ and the inner tube Dean number De_b ($=Re_b(d_i/D_c)^{0.5}$) for laminar flow. All properties used in the dimensionless numbers were calculated at the average bulk temperature. The measured nanofluid properties were used.

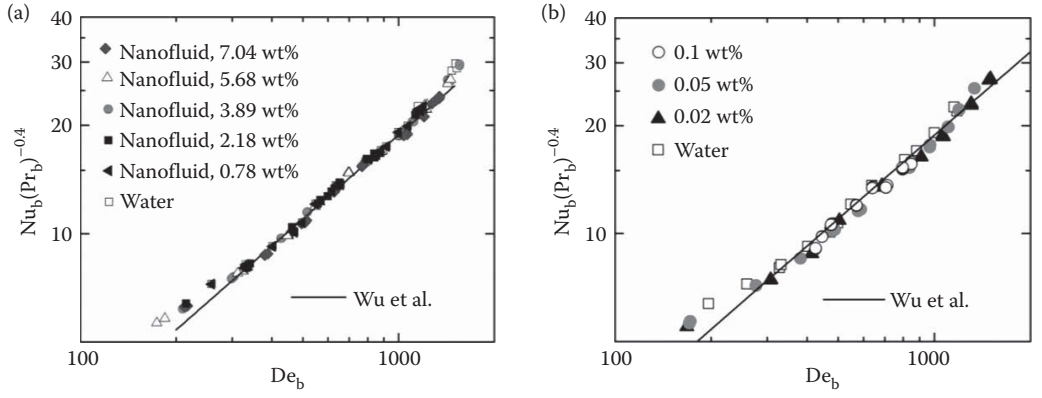


FIGURE 4.2.10 $Nu_b(Pr_b)^{-0.4}$ vs. De_b for laminar flow against the Wu et al. (2013) correlation. (a) Alumina nanofluid and (b) MWCNT nanofluid. (Data from Wu, Z. et al., *Appl. Therm. Eng.*, 60(1), 266–274, 2013; Wu, Z. et al., *Appl. Therm. Eng.*, 96, 364–371, 2016.)

Temperature effects were accounted for in the Prandtl number Pr_b. As shown in Figure 4.2.10, the Nusselt number increases with the Dean number. Water, the five alumina nanofluids, and the three MWCNT nanofluids present very similar heat-transfer characteristics. This indicates that the net effect of nanoparticles (alumina or MWCNT) on the heat-transfer performance in helically coiled tubes is probably insignificant. The thermal conductivity increase by nanoparticles is beneficial for heat transfer, whereas the secondary flow intensity induced by centrifugal forces might be mitigated by nanoparticles due to the larger viscosity and density of the nanofluid compared to those of the base fluid. Besides, the Wu et al. (2013) correlation, originally developed based on water and alumina/water nanofluid data in helically coiled tubes, can also predict the heat-transfer performance of MWCNT/water nanofluid very well.

Figure 4.2.11 presents the relationship between $Nu_b(Pr_b)^{-0.4}$ and the inner tube Reynolds number Re_b for the turbulent flow. It is clear that $Nu_b(Pr_b)^{-0.4}$ increases with Re_b and the eight nanofluids and water show very similar trends. The Seban and McLaughlin (1963) correlation can predict the thermal behavior of water and nanofluids very accurately, with a mean absolute error and a standard deviation of 2.63 and 3.25%, respectively. The existing correlation can accurately reproduce the turbulent convective heat-transfer behavior of nanofluids in helically coiled tubes by adopting the measured properties of the nanofluids in the analysis. Based on our experimental analysis, no anomalous heat-transfer enhancement exists in our cases for both laminar flow and turbulent flow.

Heat-transfer enhancement comparison for nanofluids over their base fluids based on equal Reynolds number can be misleading and tends to overestimate the actual heat-transfer enhancement largely. Figure 4.2.12 demonstrates two performance comparison criteria for the tested alumina nanofluids, that is, based on equal Reynolds number and equal flow velocity. The viscosity in the Reynolds number of Figure 4.2.12a was calculated by the Williams et al. (2008) correlation. Over 40% heat-transfer enhancement can be obtained for the 7.04% wt. nanofluid compared to water for the equal Reynolds number basis, as shown in Figure 4.2.12a. However, from Figure 4.2.12b at the same flow velocity, the heat-transfer enhancement of nanofluids over the base fluid is much less than that based on the equal Reynolds number. Thus, the anomalous heat-transfer enhancement shown in Figure 4.2.12a is just an analysis artifact. The method based on the equal Reynolds number comparison should not be used. Figure 4.2.13 shows heat-transfer coefficient comparisons for water and MWCNT nanofluids using two criteria: the fixed flow velocity and the fixed pumping power. Pumping power is calculated by the product of ΔP and the volume flow rate. The heat transfer was not enhanced by addition of MWCNTs based on the two criteria. On the contrary, the h_{nf} values are $<h_f$ except at very low flow velocities.

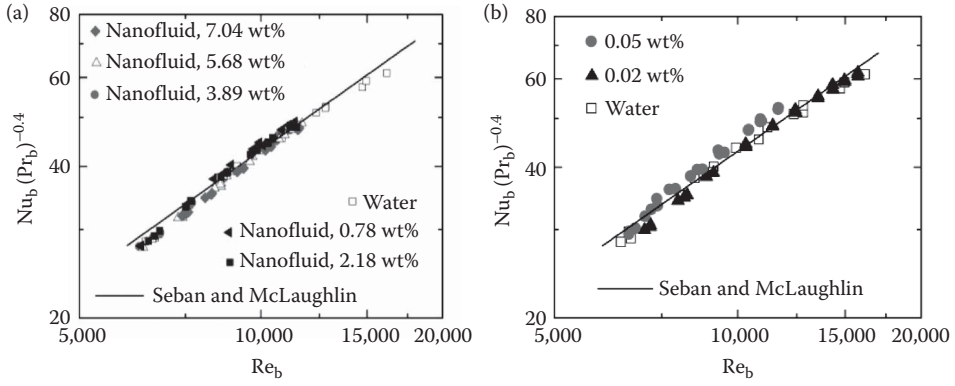


FIGURE 4.2.11 $Nu_b(Pr_b)^{-0.4}$ vs. Re_b for turbulent flow against the Seban and McLaughlin (1963) equation. (a) Alumina nanofluid and (b) MWCNT nanofluid. (Data from Wu, Z. et al., *Appl. Therm. Eng.*, 60(1), 266–274, 2013; Wu, Z. et al., *Appl. Therm. Eng.*, 96, 364–371, 2016.)

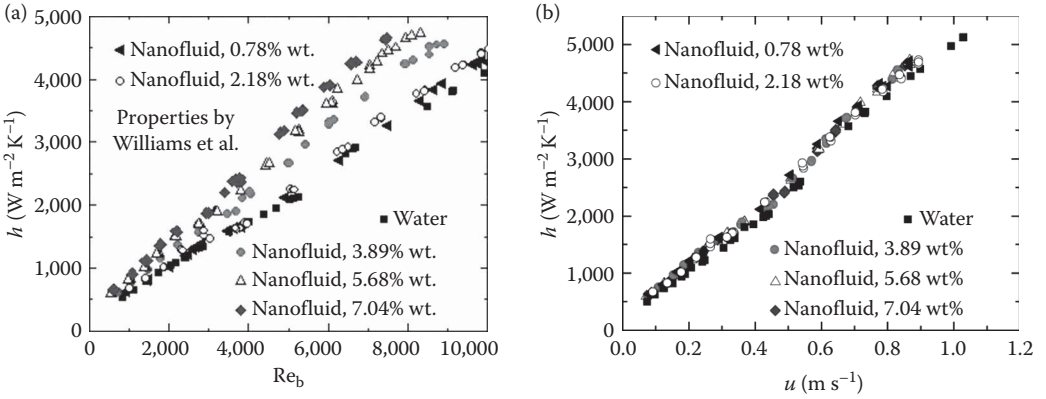


FIGURE 4.2.12 Heat-transfer coefficient comparisons: (a) h vs. Re_b for the constant Reynolds number basis, with dynamic viscosity calculated by the Williams et al. (2008) equation and (b) h vs. u for the constant flow velocity basis. (From Wu, Z. et al., *Appl. Therm. Eng.*, 60(1), 266–274, 2013.)

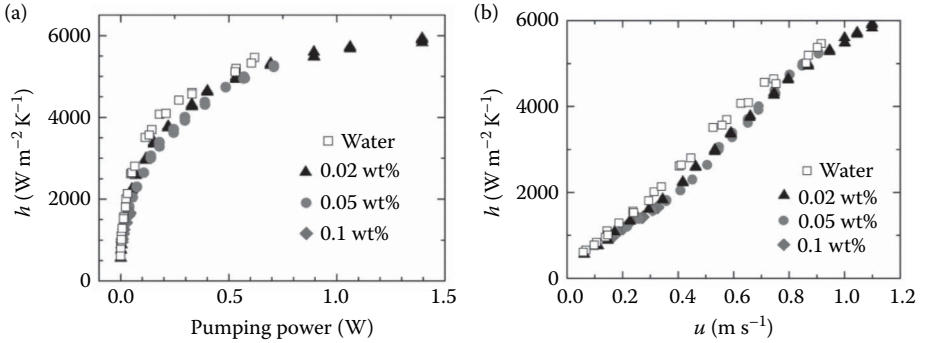


FIGURE 4.2.13 Heat-transfer coefficient comparison for water and MWCNT nanofluids at (a) a fixed flow velocity and (b) a fixed pumping power. (From Wu, Z. et al., *Appl. Therm. Eng.*, 96, 364–371, 2016.)

The heat transfer degradation in the tested MWCNT nanofluids is mainly due to the low enhancement of the thermal conductivity and the high increase of the viscosity.

Results of Nanofluids in a PHE

The relationship between the measured pressure drop (ΔP) across the PHE I and the volume flow rate for nanofluids and water is depicted in Figure 4.2.14. The pressure drop is related to the thermophysical properties, flow velocity, and the geometry of the PHE I. As shown in Figure 4.2.14, the pressure drop of nanofluids is a little bit larger than that of water and seems to increase with increasing particle volume concentrations due to the increase in viscosity. However, there is no much difference between the pressure drop of nanofluids and that of water as the particle volume concentration is low. In Figure 4.2.15, the friction factor f is plotted versus the dimensionless number Re_e for both nanofluids and water. A correlation for predicting the friction factor is found to be: $f = 20.202Re_e^{-0.191}$, where the dimensionless number Re_e is based on the equivalent diameter, $D_e = 2b_c$. It shows that all the experimental data are located within the ±10% error band. The correlation is in good agreement with the experimental data and can predict the friction factor of nanofluids flowing in PHEs. Pantzali et al. (2009) experimentally investigated the heat-transfer performance of CuO–water nanofluids in a chevron-type PHE. The experimental friction factor

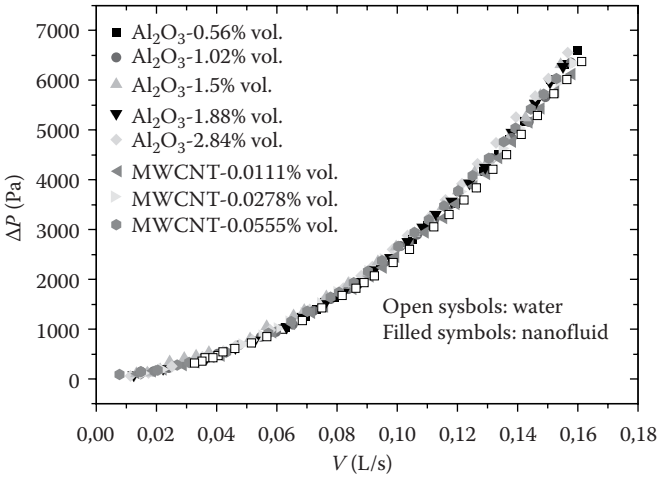


FIGURE 4.2.14 Measured pressure drop vs. the volumetric flow rate. (From Huang, D. et al., *Int. Commun. Heat Mass Transfer*, 63, 54–61, 2015.)

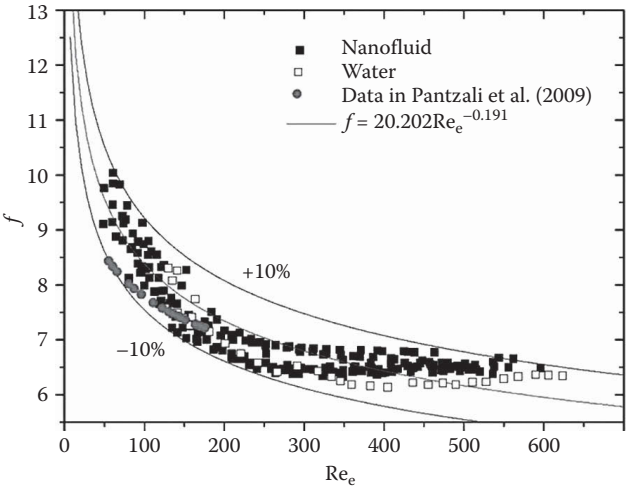


FIGURE 4.2.15 Comparison between experimental and predictive f values. (From Huang, D. et al., *Int. Commun. Heat Mass Transfer*, 63, 54–61, 2015.)

values of Pantzali et al. (2009) are also presented in Figure 4.2.15 for comparison. It shows that data of Pantzali et al. (2009) are also within the ±10% error band.

Figures 4.2.16 and 4.2.17 depict the heat-transfer performances of Al_2O_3–water nanofluids and MWCNT/water nanofluids versus Re_e and u, respectively. The Nusselt number increases with increasing Re_e. The heat-transfer coefficient increases with increasing u. Nanofluids of different volume concentrations show similar heat-transfer performances. When compared based on the equal Reynolds number, Figures 4.2.16a and 4.2.17a show that the nanofluids can always enhance heat transfer. However, the heat transfer is not improved at all if the comparison is based on a constant flow velocity. Besides, the heat transfer deterioration increases as flow velocity increases. One possible reason is that the increase in viscosity by nanoparticle additives will suppress the turbulence at high flow velocity, while the relatively small increase in thermal conductivity is not able to compensate this suppression. Also, the specific heat (c_p) decreases by adding nanoparticles. Thus the heat-transfer coefficient of nanofluids is lower than that of water at a constant flow velocity.

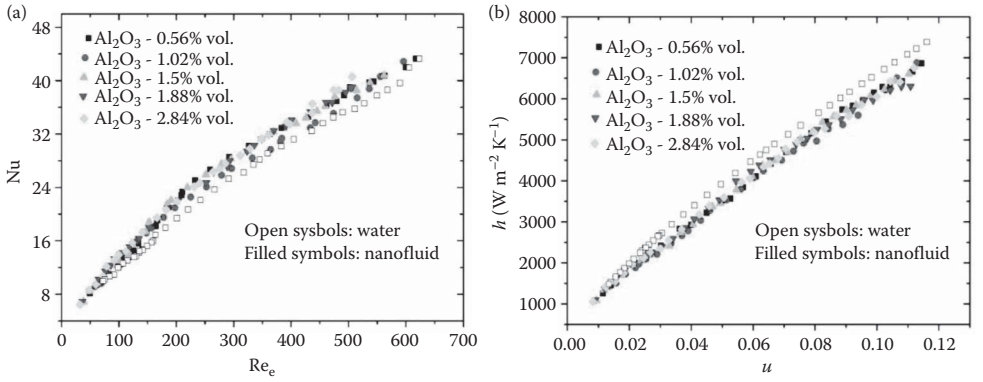


FIGURE 4.2.16 Heat-transfer performance of alumina/water nanofluids in PHE I (a) variation of Nusselt number with Reynolds number (b) variation of Heat-Transfer Coefficient (HTC) with flow velocity. (From Huang, D. et al., *Int. Commun. Heat Mass Transfer*, 63, 54–61, 2015.)

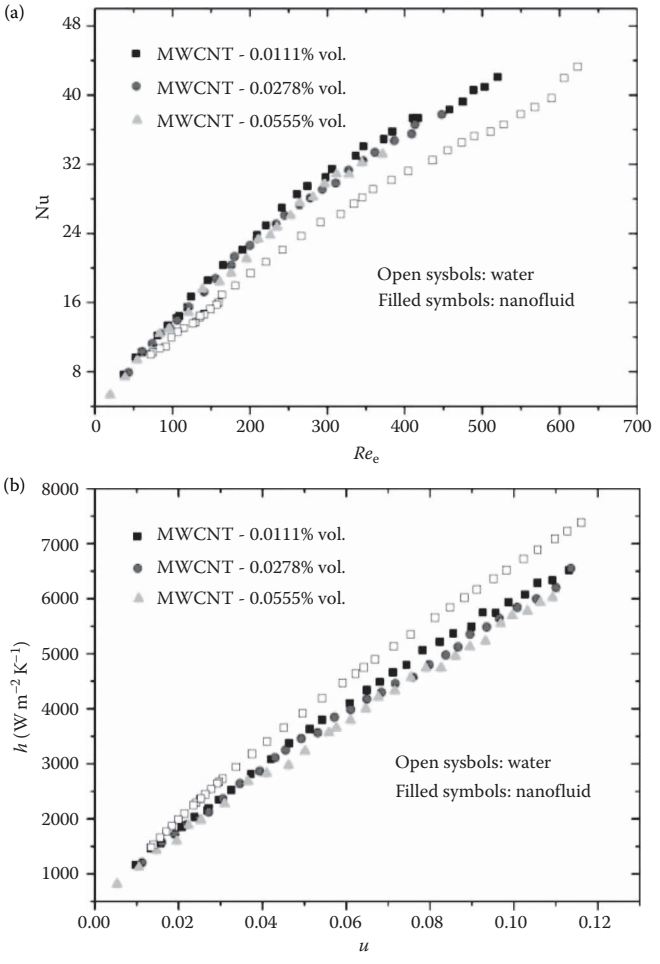


FIGURE 4.2.17 Heat-transfer performance of MWCNT/water nanofluids in PHE I (a) variation of Nusselt number with Reynolds number and (b) variation of heat-transfer coefficient with flow velocity. (From Huang, D. et al., *Int. Commun. Heat Mass Transfer*, 63, 54–61, 2015.)

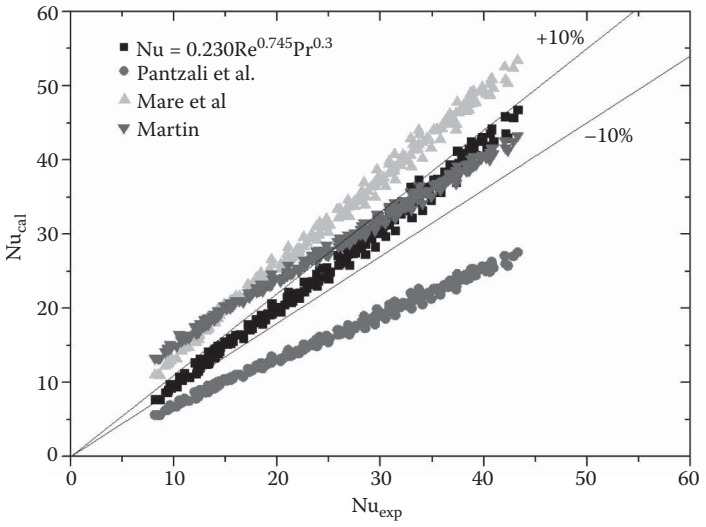


FIGURE 4.2.18 Comparison of the new correlation and three existing correlations. (From Huang, D. et al., *Int. Commun. Heat Mass Transfer*, 63, 54–61, 2015.)

A new correlation was proposed based on the experimental data. Comparison of the experimental Nu values with the predicted Nu values based on the newly proposed correlation (Huang et al., 2015) is given in Figure 4.2.18. Another three empirical correlations, Martin (1996), Pantzali et al. (2009), and Maré et al. (2011) are also shown for comparison. The newly developed correlation provides the best prediction as all the experimental data points are located within the ±10% error band. Martin (1996) predicts the experimental data very well especially at high Nusselt numbers, while the deviation is very large at low Nusselt numbers, as the correlation proposed by Martin (1996) was obtained at a Reynolds number range of $60 < Re < 30000$. The correlations proposed by Pantzali et al. (2009) and Maré et al. (2011) show large deviations from the experimental data, with the Pantzali et al. correlation underestimating the experimental data and the Maré et al. correlation overestimating the experimental data. The over- or underestimation may be caused by the different geometrical parameters and manufacturing variations of PHEs.

Results of Hybrid Nanofluids in a PHE

Hybrid nanofluid is a new kind of nanofluid, including two different types of nanoparticles. Han et al. (2007) proposed a hybrid sphere/carbon nanotube (CNT) particle, consisting of numerous CNTs attached to an alumina/iron oxide sphere, for applications in nanofluids. In such hybrid nanoparticles, heat is expected to transport rapidly from one CNT to another through the center sphere and thus leading to less thermal contact resistance between CNTs when compared to simple CNTs dispersed in fluids. CNTs have an extremely high thermal conductivity, but thermal resistance between the CNTs and the fluid has limited their performance in nanofluids. The proposed hybrid sphere/CNT particles are synthesized by spray pyrolysis followed by catalytic growth of CNTs. The spheres are about 70 nm in diameter on average, and the attached CNTs have a length up to 2 μm. The mass fraction of each component in these hybrid sphere/CNT particles is as follows: CNT ~ 3 wt%, alumina ~ 32 wt%, and iron oxide ~ 65 wt%. These hybrid nanoparticles are dispersed to poly-alpha-olefin with sonication and a small amount of surfactants to form stable nanofluids. The thermal conductivity of the fluids has been measured by a 3ω-wire method over a temperature range of 10–90°C. The results indicate that the effective thermal conductivity of the fluids is increased by about 21% at room temperature for a particle volume fraction of 0.2%.

Suresh et al. (2011) synthesized nanocrystalline alumina–copper hybrid powder by a thermochemical synthesis method. The hybrid nanofluids were prepared by dispersing the hybrid powder into deionized water. Sodium lauryl sulpfate (SLS) was used as a dispersant to get a stable suspension. The Al_2O_3–Cu/water nanofluids behave as Newtonian fluids. Figure 4.2.19 presents the thermal conductivity and viscosity data of the hybrid nanofluids compared to those of alumina nanofluids at the same concentration. Both thermal conductivity and viscosity increase with increasing nanoparticle volume fractions. The hybrid nanofluids have higher thermal conductivity and viscosity values than the alumina nanofluids.

As hybrid nanofluids seem to give a significant enhancement in thermal conductivity, the heat-transfer performance of hybrid nanofluids was experimentally investigated in PHE II. Preparation and properties of the alumina + MWCNT/water hybrid nanofluid were given earlier. As shown in Figure 4.2.20, the heat-transfer coefficient increases with increasing flow velocity. The heat-transfer

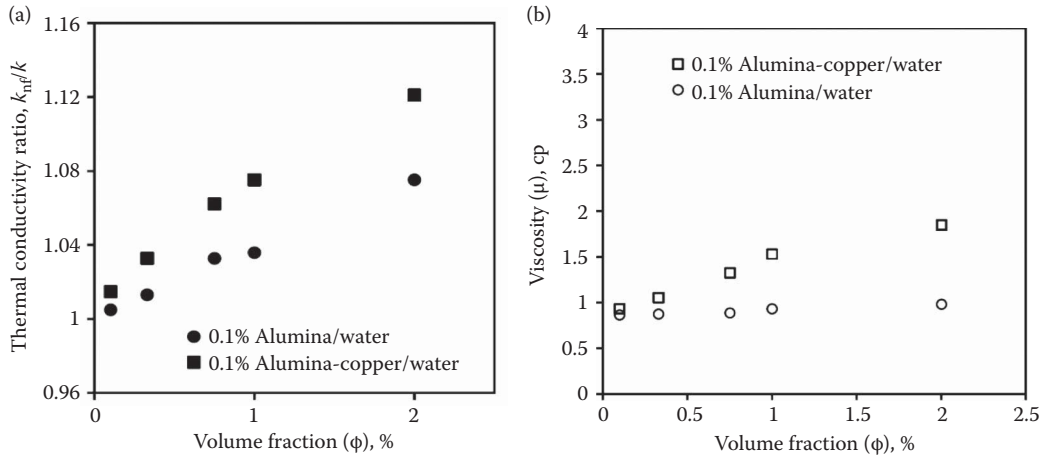


FIGURE 4.2.19 (a) Thermal conductivity and (b) viscosity of the alumina–copper/water hybrid nanofluids as a function of particle volume concentration. (From Suresh, S. et al., *Colloids Surf. A*, 388(1), 41–48, 2011.)

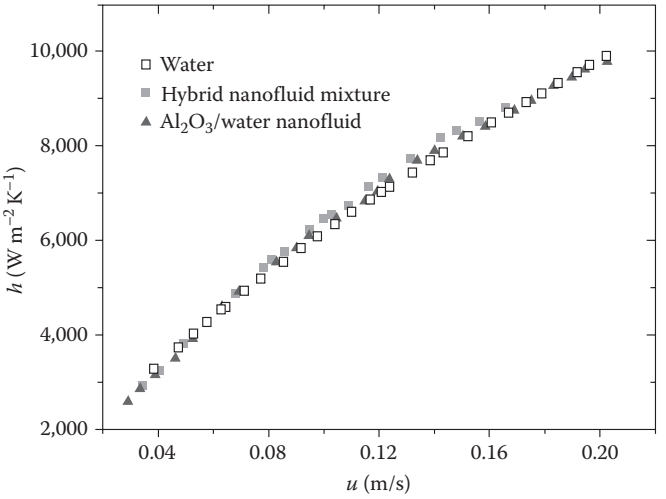


FIGURE 4.2.20 Variation of HTC versus flow velocity for hybrid nanofluids. (From Huang, D. et al., *Exp. Therm. Fluid Sci.*, 72, 190–196, 2016.)

coefficient of the hybrid nanofluid mixture seems to be slightly larger than that of the Al_2O_3/water nanofluid and water, indicating an enhanced heat transfer by using hybrid nanofluid mixture. However, it should be noted that the heat-transfer enhancement is less than that reported in Sundar et al. (2014) and Madhesh et al. (2014) and the increase in heat transfer is within the error range. This might be due to the different preparation methods of nanofluids: the hybrid nanofluid mixture was prepared by mixing the Al_2O_3/water nanofluid and MWCNT/water nanofluid in this work. Thus the alumina particles and MWCNTs might not be able to combine perfectly, and accordingly the thermal resistance between the nanostructures and fluids could not be reduced effectively.

The relationship between the measured pressure drop (ΔP) across the PHE II and the volumetric flow rate for the hybrid nanofluid mixture, Al_2O_3/water nanofluid and water is depicted in Figure 4.2.21. The pressure drop is related to the thermophysical properties, flow velocity, and the geometry of the PHE. As shown in Figure 4.2.21, pressure drop of the hybrid nanofluid is slightly less than that of the Al_2O_3/water nanofluid and slightly higher than that of water. The nanoparticle clusters in the hybrid nanofluid mixture might increase the hydrodynamic diameter of nanoparticles and thus cause the increase in viscosity and the increase in pressure drop. When evaluating the effects of nanofluids in PHEs, a comparison in terms of the pumping power and heat-transfer coefficient is necessary. The relationship between heat-transfer coefficient and pumping power is presented in Figure 4.2.22. It is observed that the hybrid nanofluid mixture seems to exhibit the highest heat-transfer coefficient at a given power consumption. This indicates that the hybrid nanofluid mixture might be a promising working fluid in heat-transfer applications as it has better heat-transfer performance than water and the Al_2O_3/water nanofluid but only with a small pressure drop penalty.

From our experimental studies in heat exchangers, it seems that nanofluid is not favorable for heat-transfer enhancement in heat exchangers when taking the pumping power into account. The primary reason is the significant increase in viscosity and only a moderate increase in thermal conductivity of the nanofluids compared to base fluids. In addition, the intensity of the turbulence produced by the complex or enhanced geometries in heat exchangers might be mitigated by nanoparticles due to their large inertia. Besides, the slip mechanisms such as Brownian diffusion and thermophoresis are relatively unimportant compared to convective transport in these applications. New types of nanofluids such as hybrid nanofluids need to be developed in order to benefit the thermal performance of heat exchangers.

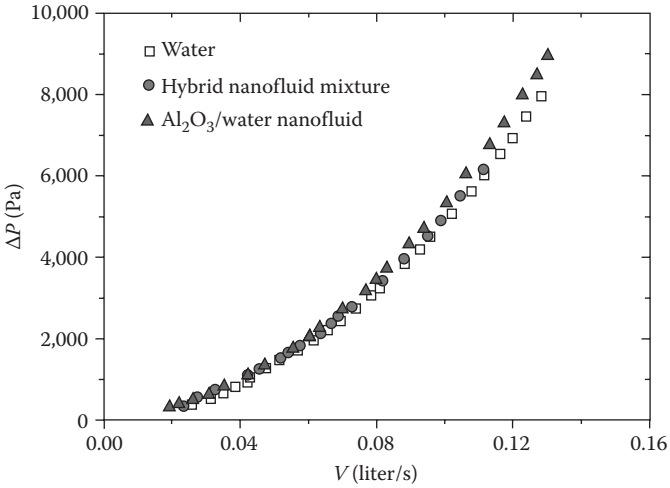


FIGURE 4.2.21 Measured pressure drop vs. the volumetric flow rate for hybrid nanofluids. (From Huang, D. et al., *Exp. Therm. Fluid Sci.*, 72, 190–196, 2016.)

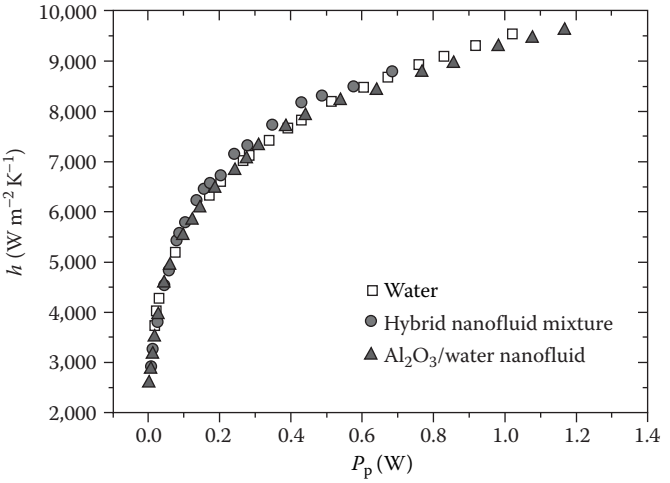


FIGURE 4.2.22 Heat-transfer coefficient vs. pumping power for hybrid nanofluids. (From Huang, D. et al., *Exp. Therm. Fluid Sci.*, 72, 190–196, 2016.)

Performance of Nanofluids in Heat Pipe Evaporators

The earlier results are for nanofluid flow in heat exchangers without phase change. Nanofluid boiling and evaporation has shown potential to increase heat transfer and critical heat flux (Coursey and Kim, 2008; Kim et al., 2010; Kwark et al. 2012; Sayahi and Bahrami, 2016). One likely mechanism for enhancement is the formation of nanoparticle deposition layers during the boiling/evaporation process, which provide more nucleation sites and also modify the surface wettability (Wu and Sundén, 2014). Up to now, there are only very few works on nanofluid boiling/evaporation in heat exchangers. Heat pipes are heat-transfer devices that combine the principles of both thermal conductivity and phase transition to efficiently manage the transfer of heat between two solid interfaces. They can easily be implemented as heat exchangers in many heat-transfer processes. Studies (e.g., Ma et al., 2006; Liu et al., 2007; Kang et al., 2009; Do and Jang, 2010; Hung et al., 2013; Kole and Dey, 2013; Karthikeyan et al., 2014; Putra et al., 2015) suggest that nanofluid can significantly enhance the thermal performance of heat pipes.

Qu et al. (2010) investigated the effects of spherical Al_2O_3 nanoparticles of 56 nm in diameter on the thermal performance of an oscillating heat pipe (OHP) charged with base water. The experimental rig is shown in Figure 4.2.23. The tested OHP was made of a bending stainless steel capillary tube with inner and outer diameters of 2 and 3 mm, respectively. The vertically oriented OHP consists of evaporation, adiabatic, and condensation sections with 50, 105, and 70 mm in length, respectively. The evaporation section (evaporator) was electrically heated, and the condensation section (condenser) was cooled by water which pumped from the cold bath. Figure 4.2.24 shows the thermal resistance (R) of the alumina nanofluid-charged OHP as a function of the power input (Q). The thermal resistance in Figure 4.2.24 is calculated by

$$R = \frac{T_{\text{evap}} - T_{\text{cond}}}{Q}, \tag{4.2.19}$$

where T_{evap} and T_{cond} denote the average wall temperatures of the evaporator and condenser, respectively. It can be seen that: (1) the thermal resistance apparently decreased for nanofluids, that is, the addition of alumina nanoparticles enhanced the heat transport capability of the OHP; (2) the thermal resistance depended greatly on the filling ratio (FR), and the lower FR led to a smaller thermal resistance; and (3) the thermal resistance decreased with increasing weight concentrations

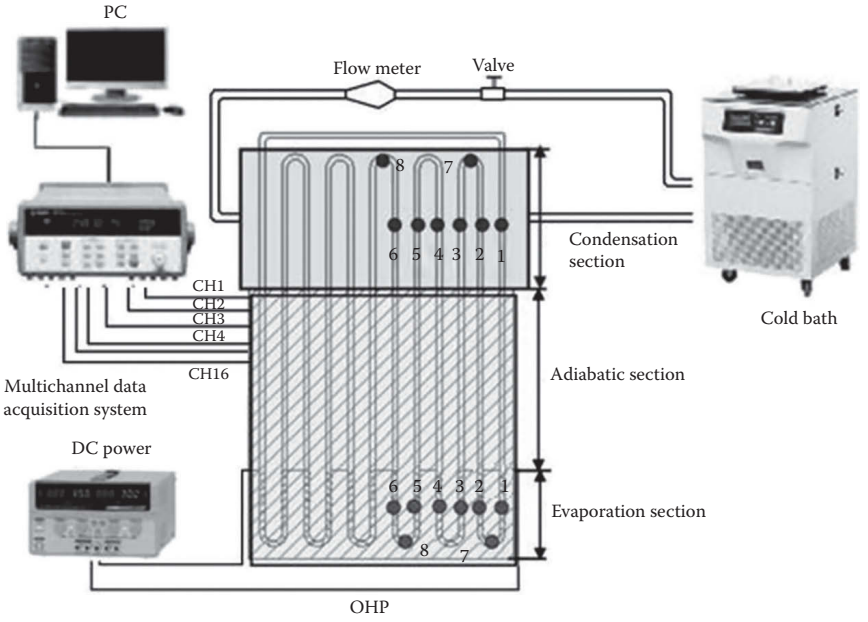


FIGURE 4.2.23 Experimental rig for nanofluids in an OHP. (From Qu, J. et al., *Int. Commun. Heat Mass Transfer*, 37(2), 111–115, 2010.)

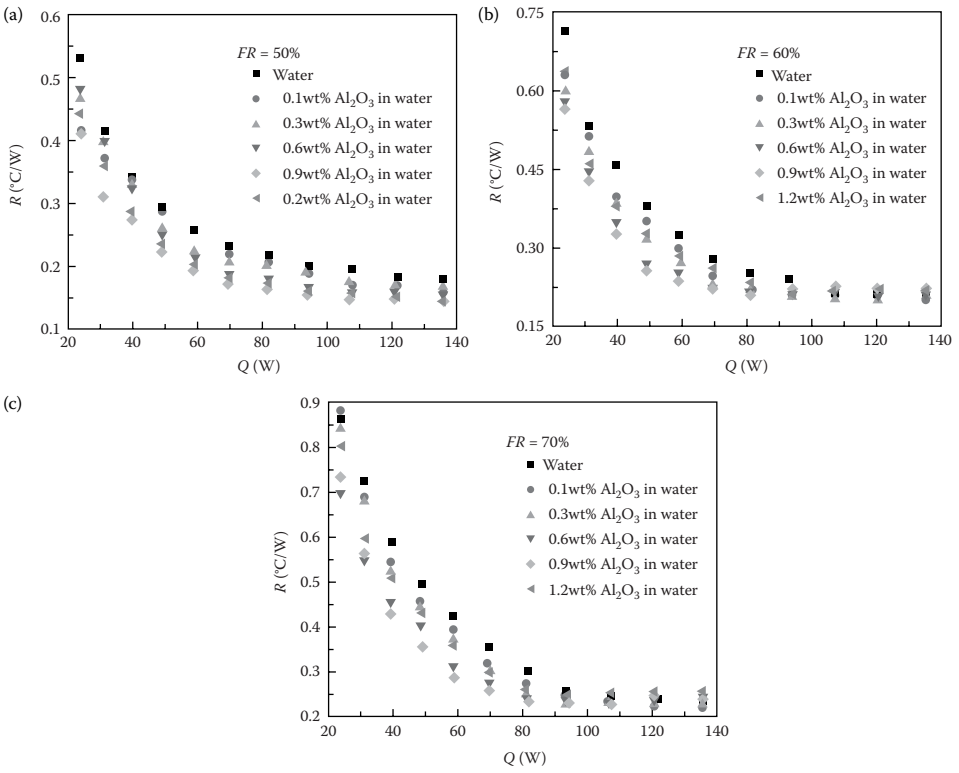


FIGURE 4.2.24 Thermal resistance comparison for the OHP charged with water and alumina nanofluids at FRs of (a) 50%, (b) 60%, and (c) 70%. (From Qu, J. et al., *Int. Commun. Heat Mass Transfer*, 37(2), 111–115, 2010.)

of nanoparticles from 0 wt% (i.e., pure water) to 0.9 wt%, but it increased when the weight concentration reached 1.2 wt%. At the optimal value of 0.9 wt%, the maximal decrease of the thermal resistance was about 0.14°C/W (or 32.5%) as compared with pure water, which occurred at 70% FR when the power input was 58.8 W. Thus, the addition of alumina nanoparticles in the base fluid enhanced the thermal performance of the OHP.

Generally, the active nucleation site density is related to the surface condition and roughness. Figure 4.2.25 shows the 2-D atomic force microscope images of the surface morphology for clean substrate and alumina nanoparticle-deposited substrate of the evaporator, respectively. As can be seen from Figure 4.2.25b, the alumina nanostructure completely modified the surface of the evaporator. The cavities of the clean surface were on the magnitude of about 2–3 μm as shown in Figure 4.2.25a, so they are one or two orders of magnitude larger than the size of nanoparticles deposited on the evaporator surface. As a result, when the smaller nanoparticles sit on nucleation sites, they can create more new active nucleation sites by splitting a single nucleation site into multiple ones, and thus enhance the boiling heat transfer (Das et al., 2008). In addition, the irregular nanopores formed between the deposited alumina nanoparticles which created the nano-roughness within the micrometer-roughened surface would affect the bubble release diameter and frequency. The bubble generation on the nanomicroroughened hierarchical surface was different from that of a clean surface. Within the cavities through the nanopore network, the nanobubbles may be continuously generated and feed the nucleation and growth of larger bubbles at the microscale cavities (Li et al., 2008). This enables stable nucleation with increasing bubble release frequency.

Kole and Dey (2013) added stable copper/water nanofluids instead of water in screen mesh wick heat pipes and found that the average evaporator wall temperature of the heat pipe using nanofluids is reduced compared to those with water. Figure 4.2.26 shows heat pipe thermal resistance as a function of nanoparticle concentration for 10 and 100 W respectively inclined at different angles. As shown in Figure 4.2.26, vertical heat pipes are found to perform better than at other inclinations. It may be noted that the total thermal resistance decreases both with the increasing Cu concentration in nanofluid and also with increase in the input power. High thermal resistance observed for low heat loads is due to the formation of a solid/liquid film on the evaporator section. On the other hand, nucleate boiling which is more important at high heat loads brings in a rapid decrease in the thermal resistance.

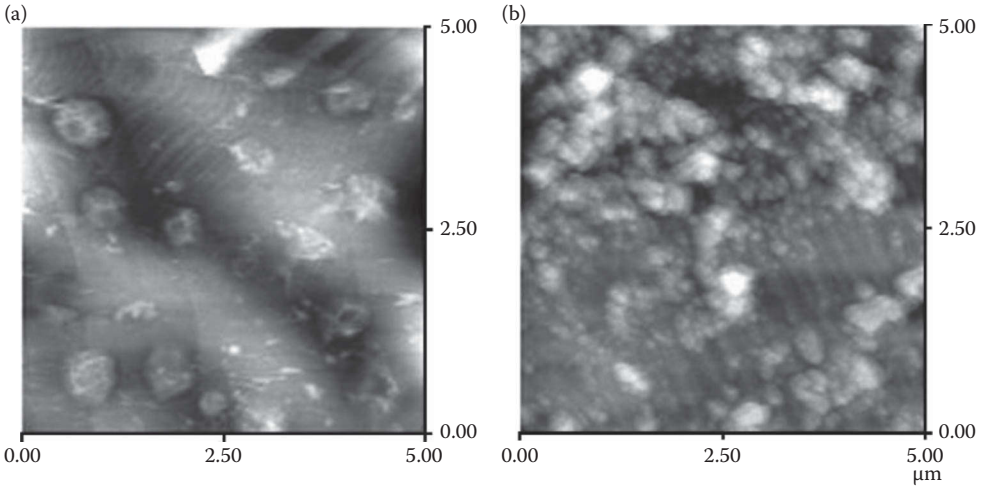


FIGURE 4.2.25 2-D atomic force microscope images of (a) a clean substrate boiled in pure water and (b) a nanoparticle-deposited substrate at the evaporator. (From Qu, J. et al., *Int. Commun. Heat Mass Transfer*, 37(2), 111–115, 2010.)

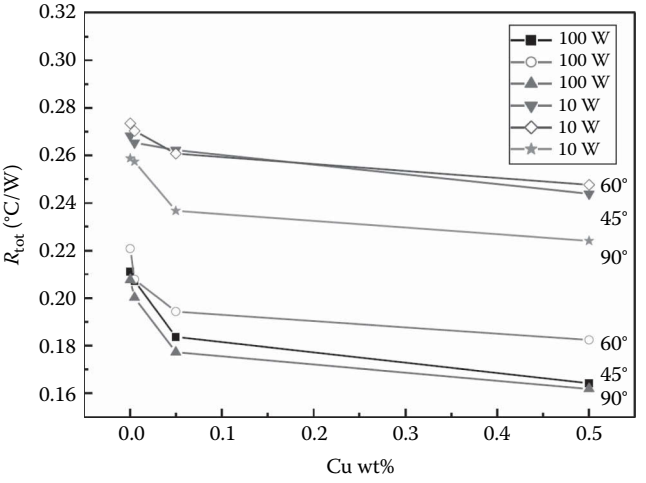


FIGURE 4.2.26 Heat pipe thermal resistance as a function of nanoparticle concentrations for 10 and 100 W, respectively, inclined at different angles. (From Kole, M. and Dey, T.K. *Appl. Therm. Eng.*, 50(1), 763–770, 2013.)

Challenges of Nanofluids in Heat Exchangers

Many efforts are needed to target the following challenges towards application of nanofluids in heat exchangers.

There is still a lot of room for thermal conductivity enhancement of nanofluids. Further thermal conductivity enhancement can be obtained by manipulating the nanostructure morphology in nanofluids to form elongated and/or percolated aggregates. More investigations should be conducted to achieve large thermal conductivity enhancement with relatively low viscosity increase, for example, by modifying the nanostructure morphology in nanofluids. Nanofluids with compact nanoparticles of different size or with hybrid or multicomponent nanoparticles might be promising as they generally present larger thermal conductivities.

New kinds of nanofluids, such as hybrid nanofluids, might be promising for heat-transfer augmentation. However, if not properly synthesized, the hybrid nanofluids will not be superior to common nanofluids. Cost-effective manufacturing and preparation methods need to be developed.

Nanofluid stability is always a problem for application of nanofluids. More investigations on the lift cycle of nanofluids should be performed.

Nanofluids of relatively large particle volume fractions or nanoparticles of nonspherical shape (e.g., nanorod, nanotube) present non-Newtonian behavior, which may cause problems for heat exchangers designed for Newtonian fluids.

For rod-like nanoparticles and nanoparticle agglomerations, the thermal conductivity enhancement at flowing conditions may be lower than that at static conditions. As shown in Figure 4.2.27, the randomly oriented agglomerations (might be percolated) at static conditions tend to be parallel with the flow direction at flowing conditions due to the liquid inertia. The long axis of the agglomeration is prone to be perpendicular to the temperature gradient under flowing conditions, and therefore it degrades the thermal conductivity enhancement.

Nanofluid boiling and evaporation has potential to increase heat transfer and critical heat flux in heat pipes and evaporators. However, it is hard to control the nanoparticle deposition process. Thick nanoparticle deposition is unfavorable as the deposition layer adds an additional resistance to heat transfer. For flow boiling in heat exchangers such as PHEs, the nanoparticle deposition process is even harder to control as the nanoparticle deposition is tempo-spatial variable in the heat exchanger. Besides, some nanofluids use surfactants to increase nanofluid stability. However, these surfactants might also affect the boiling/evaporation process (Feng et al., 2016).

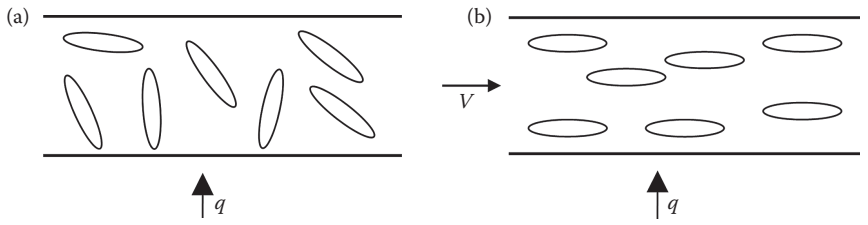


FIGURE 4.2.27 Possible agglomeration morphology at (a) static conditions and (b) flowing conditions. (From Wu, Z. et al., *Front. Heat Mass Transfer*, 5, 2014.)

Summary

Nanofluid and its application in various heat exchangers have been discussed in this chapter. Influential parameters of nanofluids on the thermal performance have been identified such as nanofluid stability, viscosity, and thermal conductivity. Reliable correlations for nanofluid properties (density, specific heat, viscosity, and thermal conductivity) have also been given.

For nanofluid flow without phase change, nanofluids are more favorable for developing flow than developed flow, and more favorable for laminar flow compared to turbulent flow.

Based on our experimental investigations on the thermal performance of alumina/water and MWCNT/water nanofluids and hybrid alumina + MWCNT/water nanofluid mixture in a helical heat exchanger and PHEs, the nanofluids and base fluid show very similar heat-transfer performances, which indicates that the net effect of nanoparticles on the thermal performance in heat exchangers is probably insignificant. When the measured nanofluid properties were used for calculation, conventional heat-transfer correlations can reproduce the thermal behaviors of the tested nanofluids for laminar flow and turbulent flow very well. Heat-transfer comparison with the base fluid at a constant Reynolds number can be misleading and should not be used to evaluate heat-transfer enhancement. Based on the fixed flow velocity and the fixed pumping power, no heat transfer enhancement can be achieved by nanoparticle additions because the increase in viscosity is generally much larger than the increase in thermal conductivity.

Nanofluid additives are very effective to enhance the thermal performance of heat pipe evaporators due to the formation of nanoparticle deposition layers during the boiling or evaporation process, which provide more nucleation sites and modify the surface wettability.

Nomenclature

Bi	Particle Biot number, $2R_k k_f/d_p$
c_p	Specific heat at constant pressure (J kg^{-1} K^{-1})
D_c	Coil diameter of curvature (m)
d	Diameter (m)
De	Dean number, $Re_b(d_i/D_c)^{0.5}$
f	Friction factor
h	Heat-transfer coefficient (W m^{-2} K^{-1})
k	Thermal conductivity (W m^{-1} K^{-1})
L	Length (m)
LMTD	Logarithmic mean temperature difference (K)
Nu	Nusselt number, hd/k
ΔP	Pressure drop (Pa)
p	Pitch of the helical coil (m); aspect ratio
Pr	Prandtl number, $c_p\mu/k$

Q	Power (W)
q	Heat flux (W m^{-2})
R_k	Interfacial thermal resistance (m^2 K W^{-1})
Re	Reynolds number, $\rho ud/\mu$
T	Temperature (K)
u	Velocity (m s^{-1})
V	Volume flow rate (m^3 s^{-1})
w	Plate width (m)
Greek symbols	
μ	Dynamic viscosity (Pa s)
ρ	Density (kg m^{-3})
φ	Volume fraction
Subscripts	
a	Agglomeration
app	Apparent
cond	Condenser
evap	Evaporator
f	Base fluid
Nf	Nanofluid
P	Nanoparticle

REFERENCES

Anoop, K.B., Sundararajan, T., and Das, S.K. Effect of particle size on the convective heat transfer in nanofluid in the developing region. *Int. J. Heat Mass Transfer*, 52(9), 2189–2195, 2009.

Batchelor, G.K. The effect of Brownian motion on the bulk stress in a suspension of spherical particles. *J. Fluid Mech.*, 83(1), 97–117, 1977.

Bianco, V., Manca, O., and Nardini, S. Numerical investigation on nanofluids turbulent convection heat transfer inside a circular tube. *Int. J. Therm. Sci.*, 50(3), 341–349, 2011.

Bruggeman, D.A.G. Dielectric constant and conductivity of mixtures of isotropic materials. *Ann. Phys. (Leipzig)*, 24, 636–679, 1935.

Buongiorno, J. Convective transport in nanofluids. *ASME J. Heat Transfer*, 128(3), 240–250, 2006.

Chen, H., Witharana, S., Jin, Y., Kim, C., and Ding, Y. Predicting thermal conductivity of liquid suspensions of nanoparticles (nanofluids) based on rheology. *Particuology*, 7(2), 151–157, 2009.

Choi, S.U.S. Nanofluids: From vision to reality through research. *ASME J. Heat Transfer*, 131(3), 033106, 2009.

Coursey, J.S. and Kim, J. Nanofluid boiling: The effect of surface wettability. *Int. J. Heat Fluid Flow*, 29(6), 1577–1585, 2008.

Das, S.K., Narayan, G.P., and Baby, A.K. Survey on nucleate pool boiling of nanofluids: The effect of particle size relative to roughness. *J. Nanoparticle Res.*, 10(7), 1099–1108, 2008.

Ding, Y., Alias, H., Wen, D., and Williams, R.A. Heat transfer of aqueous suspensions of carbon nanotubes (CNT nanofluids). *Int. J. Heat Mass Transfer*, 49(1), 240–250, 2006.

Do, K.H. and Jang, S.P. Effect of nanofluids on the thermal performance of a flat micro heat pipe with a rectangular grooved wick. *Int. J. Heat Mass Transfer*, 53(9), 2183–2192, 2010.

Eapen, J., Rusconi, R., Piazza, R., and Yip, S. The classical nature of thermal conduction in nanofluids. *ASME J. Heat Transfer*, 132(10), 102402, 2010.

Einstein, A. Eine neue bestimmung der molekul-dimension (A new determination of the molecular dimensions). *Ann. Phy.*, 19, 289–306, 1906.

Evans, W., Fish, J., and Keblinski, P. Role of Brownian motion hydrodynamics on nanofluid thermal conductivity. *Appl. Phys. Lett.*, 88(9), 093116, 2006.

Evans, W., Prasher, R., Fish, J., Meakin, P., Phelan, P., and Keblinski, P. Effect of aggregation and interfacial thermal resistance on thermal conductivity of nanocomposites and colloidal nanofluids. *Int. J. Heat Mass Transfer*, 51(5), 1431–1438, 2008.

Fan, J. and Wang, L. Review of heat conduction in nanofluids. *ASME J. Heat Transfer*, 133(4), 040801, 2011.

Feng, Z.Z. and Li, W. Laminar mixed convection of large-Prandtl-number in-tube nanofluid flow, Part I: Experimental study. *Int. J. Heat Mass Transfer*, 65, 919–927, 2013.

Feng, Z.Z., Wu, Z., Li, W., and Sundén, B. Effects of surfactant on flow boiling heat transfer of ethylene glycol/water mixtures in a mini-tube. *Heat Transfer Eng.*, 37, 1126–1135, 2016. doi:10.1080/01457632.2015.1111112.

Gharagozloo, P.E. and Goodson, K.E. Aggregate fractal dimensions and thermal conduction in nanofluids. *J. Appl. Phys.*, 108(7), 074309, 2010.

Gustafsson, S.E. Transient plane source techniques for thermal conductivity and thermal diffusivity measurements of solid materials. *Rev. Sci. Instrum.*, 62(3), 797–804, 1991.

Haghighi, E.B., Utomo, A.T., Ghanbarpour, M., Zavareh, A.I., Poth, H., Khodabandeh, R., Pacek, A., and Palm, B.E. Experimental study on convective heat transfer of nanofluids in turbulent flow: Methods of comparison of their performance. *Exp. Therm. Fluid Sci.*, 57, 378–387, 2014.

Hamilton, R.L. and Crosser, O.K. Thermal conductivity of heterogeneous two-component systems, *I&EC Fundam.*, 1(3), 182–191, 1962.

Han, Z.H. Nanofluids with enhanced thermal transport properties, PhD thesis, University of Maryland, 2008.

Han, Z.H., Yang, B., Kim, S.H., and Zachariah, M.R. Application of hybrid sphere/carbon nanotube particles in nanofluids. *Nanotechnology*, 18(10), 105701, 2007.

Hashin, Z. and Shtrikman, S. A variational approach to the theory of the elastic behaviour of multiphase materials. *J. Mech. Phys. Solid.*, 11(2), 127–140, 1963.

Healy, J.J., De Groot, J.J., and Kestin, J. The theory of the transient hot-wire method for measuring thermal conductivity. *Phys. B+C*, 82(2), 392–408, 1976.

Heris, S.Z., Esfahany, M.N., and Etemad, S.G. Experimental investigation of convective heat transfer of Al 2 O 3/water nanofluid in circular tube. *Int. J. Heat Fluid Flow*, 28(2), 203–210, 2007.

Hong, K.S., Hong, T.K., and Yang, H.S. Thermal conductivity of Fe nanofluids depending on the cluster size of nanoparticles. *Appl. Phys. Lett.*, 88(3), 031901, 2006.

Huang, D., Wu, Z., and Sundén, B. Pressure drop and convective heat transfer of Al_2O_3/water and MWCNT/water nanofluids in a chevron plate heat exchanger. *Int. J. Heat Mass Transfer*, 89, 620–626, 2015.

Huang, D., Wu, Z., and Sundén, B. Effects of hybrid nanofluid mixture in plate heat exchangers. *Exp. Therm. Fluid Sci.*, 72, 190–196, 2016.

Huang, D., Wu, X.Y., Wu, Z., Li, W., Zhu, H.T., and Sundén, B. Experimental study on heat transfer of nanofluids in a vertical tube at supercritical pressures. *Int. Commun. Heat Mass Transfer*, 63, 54–61, 2015.

Hung, Y.H., Teng, T.P., and Lin, B.G. Evaluation of the thermal performance of a heat pipe using alumina nanofluids. *Exp. Therm. Fluid Sci.*, 44, 504–511, 2013.

Ito, H. Laminar flow in curved pipes. *J. Appl. Math. Mech.*, 49(11), 653–663, 1969.

Kang, S.W., Wei, W.C., Tsai, S.H., and Huang, C.C. Experimental investigation of nanofluids on sintered heat pipe thermal performance. *Appl. Therm. Eng.*, 29(5), 973–979, 2009.

Karthikeyan, N.R., Philip, J., and Raj, B. Effect of clustering on the thermal conductivity of nanofluids. *Mater. Chem. Phys.*, 109(1), 50–55, 2008.

Karthikeyan, V.K., Ramachandran, K., Pillai, B.C., and Solomon, A.B. Effect of nanofluids on thermal performance of closed loop pulsating heat pipe. *Exp. Therm. Fluid Sci.*, 54, 171–178, 2014.

Keblinski, P., Phillpot, S.R., Choi, S.U.S., and Eastman, J.A. Mechanisms of heat flow in suspensions of nano-sized particles (nanofluids). *Int. J. Heat Mass Transfer*, 45(4), 855–863, 2002.

Kim, S.H., Choi, S.R., and Kim, D. Thermal conductivity of metal-oxide nanofluids: Particle size dependence and effect of laser irradiation. *ASME J. Heat Transfer*, 129(3), 298–307, 2007.

Kim, S.J., McKrell, T., Buongiorno, J., and Hu, L.W. Subcooled flow boiling heat transfer of dilute alumina, zinc oxide, and diamond nanofluids at atmospheric pressure. *Nucl. Eng. Des.*, 240(5), 1186–1194, 2010.

Kleinstreuer, C. and Feng, Y. Experimental and theoretical studies of nanofluid thermal conductivity enhancement: A review. *Nanoscale Res. Lett.*, 6(1), 1–13, 2011.

Kole, M. and Dey, T.K. Thermal performance of screen mesh wick heat pipes using water-based copper nanofluids. *Appl. Therm. Eng.*, 50(1), 763–770, 2013.

Krieger, I.M. and Dougherty, T.J. A mechanism for non-Newtonian flow in suspensions of rigid spheres. *Trans. Society Rheol.*, 3(1), 137–152, 1959.

Kwark, S.M., Kumar, R., Moreno, G., and You, S.M. Transient characteristics of pool boiling heat transfer in nanofluids. *ASME J. Heat Transfer*, 134(5), 051015, 2012.

Levy, O. and Stroud, D. Maxwell Garnett theory for mixtures of anisotropic inclusions: Application to conducting polymers. *Phys. Rev. B*, 56(13), 8035, 1997.

Li, C., Wang, Z., Wang, P.I., Peles, Y., Koratkar, N., and Peterson, G.P. Nanostructured copper interfaces for enhanced boiling. *Small*, 4(8), 1084–1088, 2008.

Li, Y., Tung, S., Schneider, E., and Xi, S. A review on development of nanofluid preparation and characterization. *Powder Technol.*, 196(2), 89–101, 2009.

Liu, J.T.C. On the anomalous laminar heat-transfer intensification in developing region of nanofluid flow in channels or tubes. *Proc. R. Soc.*, A8, 2383–2398, 2012.

Liu, Z.H., Xiong, J.G., and Bao, R. Boiling heat-transfer characteristics of nanofluids in a flat heat pipe evaporator with micro-grooved heating surface. *Int. J. Multiph. Flow*, 33(12), 1284–1295, 2007.

Ma, H.B., Wilson, C., Yu, Q., Park, K., Choi, S.U.S., and Tirumala, M. An experimental investigation of heat transport capability in a nanofluid oscillating heat pipe. *ASME J. Heat Transfer*, 128(11), 1213–1216, 2006.

Madhesh, D., Parameshwaran, R., and Kalaiselvam, S. Experimental investigation on convective heat transfer and rheological characteristics of Cu-TiO_2 hybrid nanofluids. *Exp. Therm. Fluid Sci.*, 52, 104–115, 2014.

Maré, T., Halelfadl, S., Sow, O., Estellé, P., Duret, S., and Bazantay, F. Comparison of the thermal performances of two nanofluids at low temperature in a plate heat exchanger. *Exp. Therm. Fluid Sci.*, 35(8), 1535–1543, 2011.

Martin, H. A theoretical approach to predict the performance of chevron-type plate heat exchangers. *Chem. Eng. Proc. Process Intensif.*, 35(4), 301–310, 1996.

Maxwell, J.C. *A Treatise on Electricity and Magnetism*, Second Edition, Clarendon Press, Oxford, 1881.

Mo, H., Evmenenko, G., and Dutta, P. Ordering of liquid squalane near a solid surface. *Chem. Phys. Lett.*, 415(1), 106–109, 2005.

Nan, C.W., Shi, Z., and Lin, Y. A simple model for thermal conductivity of carbon nanotube-based composites. *Chem. Phys. Lett.*, 375(5), 666–669, 2003.

Nielsen, J.K. Host plant selection of monophagous and oligophagous flea beetles feeding on crucifers. *Entomol. Exp. Appl.*, 24(3), 562–569, 1978.

O'Hanley, H., Buongiorno, J., McKrell, T., and Hu, L.W. Measurement and model validation of nanofluid specific heat capacity with differential scanning calorimetry. *Adv. Mech. Eng.*, 4, 181079, 2012.

Pandey, S.D. and Nema, V.K. Experimental analysis of heat transfer and friction factor of nanofluid as a coolant in a corrugated plate heat exchanger. *Exp. Therm. Fluid Sci.*, 38, 248–256, 2012.

Pantzali, M.N., Mouza, A.A., and Paras, S.V. Investigating the efficacy of nanofluids as coolants in plate heat exchangers (PHE). *Chem. Eng. Sci.*, 64(14), 3290–3300, 2009.

Philip, J. and Shima, P.D. Thermal properties of nanofluids. *Adv. Colloid Interface Sci.*, 183, 30–45, 2012.

Pop, E., Varshney, V., and Roy, A.K. Thermal properties of graphene: Fundamentals and applications. *MRS Bull.*, 37, 1273–1281, 2012.

Prabhat, N., Buongiorno, J., and Hu, L.W. Convective heat-transfer enhancement in nanofluids: Real anomaly or analysis artifact? *ASME/JSME 2011 8th Thermal Engineering Joint Conference*, March 13–17, 2011, Honolulu, Hawaii.

Prasad, P.D., Gupta, A., Sreeramulu, M., Sundar, L.S., Singh, M.K., and Sousa, A.C. Experimental study of heat transfer and friction factor of Al_2O_3 nanofluid in U-tube heat exchanger with helical tape inserts. *Exp. Therm. Fluid Sci.*, 62, 141–150, 2015.

Prasher, R., Phelan, P.E., and Bhattacharya, P. Effect of aggregation kinetics on the thermal conductivity of nanoscale colloidal solutions (nanofluid). *Nano Lett.*, 6(7), 1529–1534, 2006.

Putra, N., Ramadhan, R.S., and Septiadi, W.N. Visualization of the boiling phenomenon inside a heat pipe using neutron radiography. *Exp. Therm. Fluid Sci.*, 66, 13–27, 2015.

Qu, J., Wu, H.Y., and Cheng, P. Thermal performance of an oscillating heat pipe with Al_2O_3-water nanofluids. *Int. Commun. Heat Mass Transfer*, 37(2), 111–115, 2010.

Sayahi, T. and Bahrami, M. Investigation on the effect of type and size of nanoparticles and surfactant on pool boiling heat transfer of nanofluids. *ASME J. Heat Transfer*, 138(3), 031502, 2016.

Seban, R.A. and McLaughlin, E.F. Heat transfer in tube coils with laminar and turbulent flow. *Int. J. Heat Mass Transfer*, 6(5), 387–395, 1963.

Sergis, A. and Hardalupas, Y. Anomalous heat transfer modes of nanofluids: A review based on statistical analysis. *Nanoscale Res. Lett.*, 6(1), 391, 2011.

Seyf, H.R. and Mohammadian, S.K. Thermal and hydraulic performance of counterflow microchannel heat exchangers with and without nanofluids. *ASME J. Heat Transfer*, 133(8), 081801, 2011.

Sundar, L.S., Singh, M.K., and Sousa, A.C. Enhanced heat transfer and friction factor of MWCNT–Fe 3 O 4/ water hybrid nanofluids. *Int. Commun. Heat Mass Transfer*, 52, 73–83, 2014.

Sundén, B. *Introduction to Heat Transfer*, WIT Press, Southampton, 2012.

Sundén, B. and Wu, Z. Performance of heat exchangers using nanofluids. In *Heat-Transfer Enhancement with Nanofluids*, eds V. Bianco, O. Manca, S. Nardini, and K. Vafai, CRC Press, Boca Raton, 2015. Chapter 7, 207–233.

Suresh, S., Venkitaraj, K.P., Selvakumar, P., and Chandrasekar, M. Synthesis of Al_2O_3-Cu/water hybrid nanofluids using two step method and its thermo physical properties. *Colloids Surf. A*, 388(1), 41–48, 2011.

Taylor, R., Coulombe, S., Otanicar, T., Phelan, P., Gunawan, A., Lv, W., Rosengarten, G., Prasher, R., and Tyagi, H. Small particles, big impacts: A review of the diverse applications of nanofluids. *J. Appl. Phys.*, 113(1), 011301, 2013.

Timofeeva, E.V., Gavrilov, A.N., McCloskey, J.M., and Tolmachev, Y.V. Thermal conductivity and particle agglomeration in alumina nanofluids: Experiment and theory. *Phys. Rev. E*, 76(6), 061203, 2007.

Vanapalli, S. and ter Brake, H.J.M. Assessment of thermal conductivity, viscosity and specific heat of nanofluids in single phase laminar internal forced convection. *Int. J. Heat Mass Transfer*, 64, 689–693, 2013.

Wang, L. and Fan, J. Nanofluids research: Key issues. *Nanoscale Res. Lett.*, 5(8), 1241–1252, 2010.

Webb, R.L. and Kim, N.H. *Principle of Enhanced Heat Transfer*, Taylor & Francis, New York, 2005.

Williams, W., Buongiorno, J., and Hu, L.W. Experimental investigation of turbulent convective heat transfer and pressure loss of alumina/water and zirconia/water nanoparticle colloids (nanofluids) in horizontal tubes. *ASME J. Heat Transfer*, 130(4), 042412, 2008.

Wu, Z., Feng, Z., Sundén, B., and Wadsö, L. A comparative study on thermal conductivity and rheology properties of alumina and multi-walled carbon nanotube nanofluids. *Front. Heat Mass Transfer*, 5, 18, 2014.

Wu, Z. and Sundén, B. On further enhancement of single-phase and flow boiling heat transfer in micro/minichannels. *Renew. Sustain. Energ. Rev.*, 40, 11–27, 2014.

Wu, Z. and Sundén, B. Convective heat-transfer performance of aggregate-laden nanofluids. *Int. J. Heat Mass Transfer*, 93, 1107–1115, 2016.

Wu, Z., Wang, L., and Sundén, B. Pressure drop and convective heat transfer of water and nanofluids in a double-pipe helical heat exchanger. *Appl. Therm. Eng.*, 60(1), 266–274, 2013.

Wu, Z., Wang, L., Sundén, B., and Wadsö, L. Aqueous carbon nanotube nanofluids and their thermal performance in a helical heat exchanger. *Appl. Therm. Eng.*, 96, 364–371, 2016.

Xuan, Y., and Li, Q. Investigation on convective heat transfer and flow features of nanofluids. *ASME J. Heat Transfer*, 125(1), 151–155, 2003.

Yu, W., France, D.M., Routbort, J.L., and Choi, S.U.S. Review and comparison of nanofluid thermal conductivity and heat-transfer enhancements. *Heat Transfer Eng.*, 29(5), 432–460, 2008.

4.3 CONVECTION HEAT TRANSFER IN CONDUITS WITH NANOFLUIDS

Clement Kleinstreuer and Zelin Xu

Introduction

For convection heat transfer in conduits, according to Newton's law of cooling ($Q = kA\Delta T$), the heat transfer rate is proportional to the thermal conductivity of the fluid k, the heat transfer area A, and the temperature difference ΔT, being the "driving force." This implies that, next to larger heat transfer areas, fluids with enhanced k values would effectively increase the heat transfer rate (see also Book Section 3.2.3). This can be achieved by dispersing nanometer-size solid particles in liquids, called *nanofluids*, which elevates the effective thermal conductivity of the mixture. Advancements in nanotechnology allow the fabrication of such spheres, ellipsoids, tubes, or fibers with diameters less than 100 nm. This new class of heat transfer fluids is a suspension of nanoparticles (NPs) with typically high thermal conductivity in the base fluid at low volume fractions, say, less than 6%. In nanofluids, Brownian agitation of nanoparticles overcomes particle sedimentation, resulting in mixture properties that are determined by characteristics of both the NPs and the base fluid. The NPs that have been investigated include, but are not limited to, oxides such as Al_2O_3, CuO, TiO_2, Fe_3O_4, and ZnO, metallic solids such as Cu and Ag, nonmetallic solids such as diamond, SiC, as well as carbon nanotubes. Typical base fluids include water, oil, and ethylene glycol (EG). The thermal conductivities of different nanoparticle materials are shown in Table 4.3.1, while a few examples of nanofluids as well as dispersed NPs are shown in Figure 4.3.1.

The nanofluid-flow concept can be further expanded to develop additional applications; for example, to include inhaled toxic/therapeutic nanosize aerosols or nanodroplet–vapor–air mixtures, as well as nanodrugs infused into the blood stream to combat cancer.

The term "nanofluid" was first used by Choi (1995) to refer to the fluid–nanoparticle mixture. Later experimental results showed that adding a small amount of nanoparticles (usually less than 6%) to the base fluids measurably increased the thermal conductivity of the particle suspensions in batch processes (Kleinstreuer and Feng, 2011). For heat transfer in conduits, adding NPs to the base fluid also increased the convection heat transfer coefficient and the Nusselt number (Kleinstreuer, 2016; see also Section 4.2). While the initial applications of nanofluids have been driven primarily by the potential to significantly improve convective heat transfer performances, presently much broader applications are being implemented. Specifically, its unique properties are being explored in drug delivery (Peer et al., 2007; Kleinstreuer and Childress, 2015), mass transport (Fang et al., 2009; Veilleux and Coulombe, 2011), boiling phenomena (Kim et al., 2007; Kwark et al., 2010), absorption and conversion of radiation (Otanicar et al., 2010; Xu and Kleinstreuer, 2014a), optics (Liu et al., 2006; Taylor et al., 2011), surfaces and catalysts (Taylor et al., 2013), and lubrication (Hu et al., 2015).

Preparation of Nanofluids

The preparation of nanofluids is more involved than just mixing the nanoparticles into the base fluid. Chemically inert NPs usually agglomerate due to the van der Waals force. Metal NPs oxidize easily in air due to their large surface-to-volume ratio and high surface reactivity. Even after the particles are dispersed in the carrier fluid, it is difficult to keep the mixture stable and uniform for a long time. Therefore, special procedures are necessary in the preparation of stable nanofluids.

Synthesis

There are primarily two methods to prepare nanofluids: two-step preparation and one-step preparation. In the two-step method, nanoparticles are first produced as dry powders via chemical or physical methods, and then mixed with a base fluid. Aggregation may occur because of the strong van der Waals force acting among adjacent nanoparticles. Hence, ultrasonic vibration or high-shear mixing

TABLE 4.3.1
Thermal Conductivities of Common Liquids and Nanoparticles

Coolants		Metals		Metal Oxides		Nonmetallic	
Material	**k (W/m·K)**	**Material**	**k (W/m·K)**	**Material**	**k (W/m·K)**	**Material**	**k (W/m·K)**
Water	0.61	Cu	401	CuO	77	SiO_2	1.4
EG	0.25	Al	237	ZnO	89	Si	148
Mineral Oil	0.15	Zn	120	Al_2O_3	36	SiC	120
		Ni	93	TiO_2	11	CNT	3000
		Ag	429	Fe_3O_4	7		
		Fe	60				

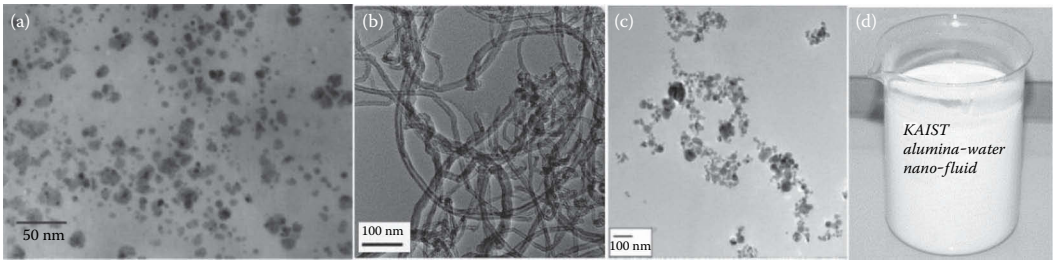


FIGURE 4.3.1 Nanofluids: (a) TEM image of Cu nanoparticles produced by direct evaporation into ethylene glycol (Eastman et al., 2001); (b) TEM image of carbon nanotubes (Ding et al., 2006); (c) TEM image of CuO nanoparticles prepared by rapid drying of a diluted sample (Lee et al., 2006); and (d) Al_2O_3–water nanofluid (Bang and Chang, 2005).

over extended periods of time is usually employed to stir nanopowders and form well-dispersed suspensions. This is the most economical method and is more suitable for preparing nanofluids with oxide NPs than those with metallic NPs.

The more advanced one-step technique simultaneously produces and disperses nanoparticles in fluids. There are several specific methods in this category. For example, in the direct evaporation and condensation method, metals are vaporized and cooled into solids to obtain nanofluids. This technique has better control of the NP morphology and can greatly reduce aggregation. However, it is not suitable for large-scale syntheses. Also, there may be residuals remaining due to incomplete reaction.

Stabilization

Nanofluid stability is essential for practical applications, avoiding particle agglomeration with severe effects. For example, the thermophysical properties of nanofluids will change and any enhanced heat transfer deteriorates as a result. The aggregates may settle and hence deposit on the conduit wall, adding interface thermal resistance and possibly clogging microchannels or filters. Furthermore, due to strong van der Waals interactions between nanoparticles, it may be difficult to maintain a stable suspension in practical applications. Specifically, the attractive potential is proportional to the particle diameter d_p, the material (or Hamaker) constant A_H, and is inversely proportional to the distance of separation, h:

$$\psi_{LD} = -\frac{A_H}{12}\left[\frac{1}{x(x+2)} + \frac{1}{(x+1)^2} + 2\ln\frac{x(x+2)}{(x+1)^2}\right] \tag{4.3.1}$$

where $x = h/d_p$. For small particle separations ($h \ll d_p$), this reduces to

$$\psi_{LD} = -\frac{A_H d_p}{24h} \tag{4.3.2}$$

The average (minimum) separation distance between two spheres in a well-dispersed suspension can be obtained via

$$h = \frac{d_p}{2}\left[\left(\frac{0.64}{\varphi}\right)^{\frac{1}{3}} - 1\right] \tag{4.3.3}$$

where φ is the NP volume fraction. According to Equation 4.3.3, the average particle distance scales with the particle diameter and hence becomes very small for NPs.

There are two main methods of stabilizing nanofluids, that is, steric stabilization, where adsorbed polymeric molecules separate nanoparticles away from each other, and electrostatic or charge stabilization, where electrical double layers are formed at particles surfaces to generate a repulsive force.

Adding surfactants to the nanofluids also enhances the stability of NPs in aqueous suspensions. The reason is that the hydrophobic surfaces of nanoparticles/nanotubes are modified to become hydrophilic and *vice versa* for nonaqueous liquids. This causes steric repulsion between NPs. However, surfactant functionality may deteriorate under high temperatures when the bonds between the polymer molecules and the NPs weaken. The stability can also be improved by altering the pH value of the mixture. The repulsive forces overcome the van der Waals force and keep the particles from aggregating. Figure 4.3.2 illustrates the two stabilization mechanisms.

The potential of the electrostatic repulsive force can be expressed as

$$\psi_{EDL} = 32\pi d_p \kappa_B T \kappa^{-2} \rho_{el} \tanh^2\left(\frac{\sigma e \phi_0}{4\kappa_B T}\right)\exp(-\kappa h) \tag{4.3.4}$$

where κ_B is the Boltzmann constant, σ is the valence of the ions, κ is the inverse of the Debye length, ρ_{el} is the concentration of electrons, e is the electric charge, and ϕ_0 is the electric double layer (EDL) potential at the surface of the nanoparticles. The sum of the potentials due to the van der Waals attraction and the electrostatic repulsion determines the stability of the nanofluid. Higher potentials lead to bigger potential barriers for aggregation. The interaction potentials at different pH-values are shown in Figure 4.3.3a.

Because pH is a measure of the electrostatic charge on the particle surface, it has a significant influence on the thermal property and convection heat transfer performance of nanofluids (see Figure 4.3.3b). Generally, pH must be far from the IEP (isoelectric point, the pH at which the particle carries no net electrical charge) to maintain good nanofluid stability.

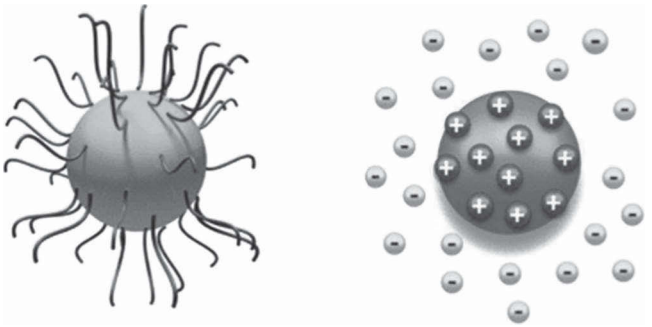


FIGURE 4.3.2 Stabilization methods. (From Yu, W. and Xie, H., *J. Nanomater.*, 1, 2012.)

Thermophysical Properties of Nanofluids

The basic particle–fluid mixture properties are greatly a function of the particle volume fraction. If particles are assumed to be spherical, monodisperse, and forming a homogeneous dilute suspension, the nanofluid density and heat capacity in the case of low NP-volume fractions can be expressed as

$$\rho_{nf} = \varphi\rho_p + (1-\varphi)\rho_{bf} \tag{4.3.5a}$$

$$(\rho c_p)_{nf} = \varphi(\rho c_p)_p + (1-\varphi)(\rho c_p)_{bf} \tag{4.3.5b}$$

where the subscripts bf and p indicate base fluid and particle, respectively.

For the effective thermal conductivity k_{nf} and effective viscosity μ_{nf} of nanofluids, classical mixture theories fail to predict accurate values.

Thermal Conductivity of Nanofluids

Certain nanofluids have shown significantly enhanced thermal conductivities over their base fluids. It has been shown that the effective thermal conductivity of the mixture k_{nf} increases with the NP volume fraction and temperature and with the decrease of particle size (see Figure 4.3.4). In addition, particle shape, pH, and aggregation also have significant effects on k_{nf}. As can be observed from Figure 4.3.4, the percentages of enhancement in thermal conductivity cover a wide range.

Originally, an expression for k_{nf} of a two-phase mixture with relatively large spheres has been derived by Maxwell in 1881. It assumes that the increase in mixture thermal conductivity is due to the high thermal conductivity of the particles. In this case, the effective thermal conductivity depends only on the particle volume fraction. Assuming a monodisperse mixture with spherical particles, the effective thermal conductivity is given as

$$k_{eff} = k_{bf} \cdot \left(1 + \frac{3\left(\frac{k_a}{k_{bf}} - 1\right)\varphi}{\left(\frac{k_a}{k_{bf}} + 2\right) - \left(\frac{k_a}{k_{bf}} - 1\right)\varphi} \right) \tag{4.3.6}$$

where k_{bf} and k_a are the thermal conductivities of the base fluid and the nanoparticle material, respectively. The Maxwell model predicts the k_{eff} of mixtures with particle sizes greater than 1 μm. However, it significantly underestimates the effective thermal conductivity of nanofluids.

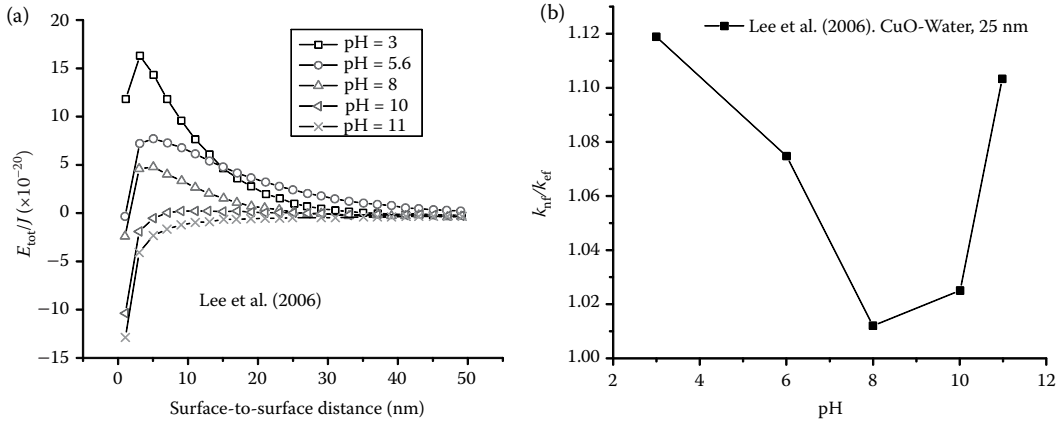


FIGURE 4.3.3 Electrostatic stabilization: (a) The interaction potentials at various pH values as a function of interparticle distance and (b) effect of pH on the effective thermal conductivity of CuO–water nanofluid.

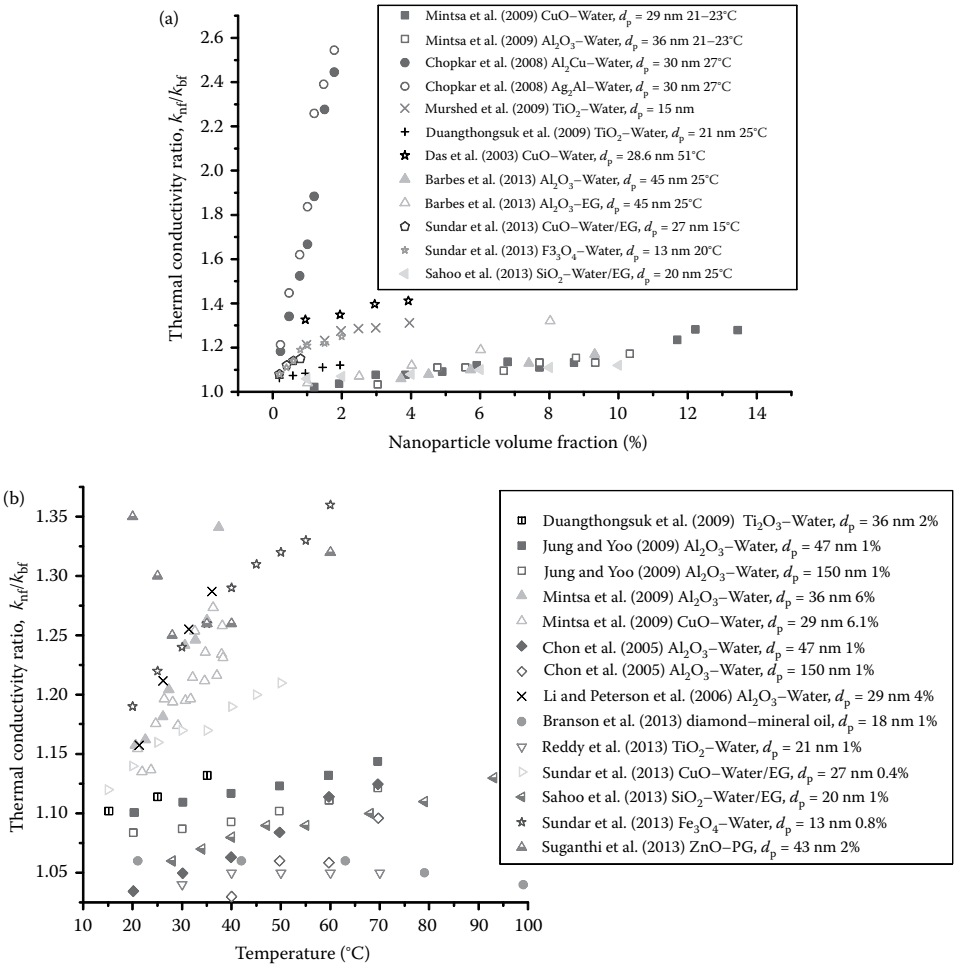


FIGURE 4.3.4 Nanofluid thermal conductivity as a function of (a) volume fraction ϕ and (b) temperature.

To explain this discrepancy, a number of models have been proposed. For example, Brownian motion and particle aggregation are arguably the most fundamental effects that characterize nanofluids. Nanoparticles undergo vigorous Brownian motion due to the random bombardment of liquid molecules, carrying with them relatively large volumes of surrounding liquid. This microscale interaction may occur between hot and cold regions of the fluid, resulting in micromixing effects and causing an enhanced effective thermal conductivity. The randomly moving particles may collide with each other at high frequency, forming locally small cluster-like structures. Such NP clusters can form thermal pathways and augment heat transfer, resulting in enhanced k_{nf}-values (see Figure 4.3.5).

For example, Kleinstreuer and Feng (2012) proposed a model based on Brownian motion-induced microconvection. It divides the thermal conductivity k_{nf} of the nanofluid into two parts: a static part, k_{static}, similar to Maxwell's model, and a micromixing part, k_{mm}, due to Brownian motion of nanoparticles. However, this model does not consider NP clustering, which may be also important. So, Xu and Kleinstreuer (2014a) enhanced the F-K model to include the local clustering effect as well as the interfacial thermal resistance effect. The modified model has the same basic form $k_{nf} = k_{static} + k_{mm}$; however, with a different k_{mm} expression:

$$k_{mm} = 19631 \cdot C_c \varphi \frac{\kappa_B \tau_a}{m_a} \left(\rho c_p\right)_{nf} \left(\overline{T} \cdot \ln \overline{T} - \overline{T}\right) \tag{4.3.7}$$

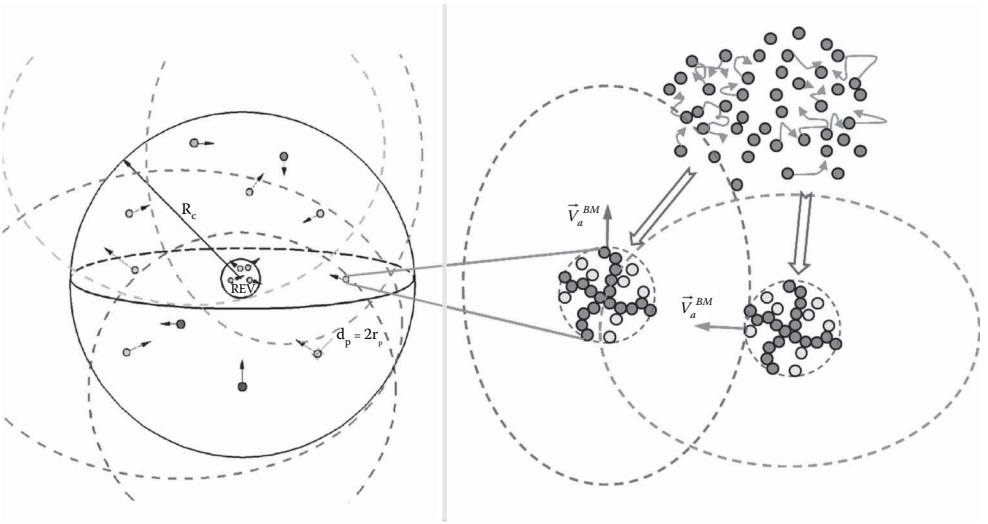


FIGURE 4.3.5 Mechanisms causing enhanced thermal conductivity of nanofluids: nanoparticle interaction and clustering due to Brownian motion. (From Xu, Z., An improved thermal conductivity model for nanofluids with applications to concentration photovoltaic-thermal systems. MS Thesis, North Carolina State University, 2014.)

Now, the NP thermal conductivity in the original Maxwell model is replaced by the effective thermal conductivity of particle aggregate unit k_a:

$$k_a = k_{de} \cdot \frac{3+\varphi_b\left[2\beta_{11}\left(1-L_{11}\right)+\beta_{33}\left(1-L_{33}\right)\right]}{3-\varphi_b\left(2\beta_{11}L_{11}+\beta_{33}L_{33}\right)} \tag{4.3.8}$$

where

$$\begin{aligned} k_{de} &= (3\varphi_{de}-1)k_{peff}+(2-3\varphi_{de})k_{bf} \\ &+\sqrt{[(3\varphi_{de}-1)k_{peff}]^2+[(2-3\varphi_{de})k_{bf}]^2+2[2+9\varphi_{de}(1-\varphi_{de})]k_{peff}k_{bf}} \end{aligned} \tag{4.3.9}$$

The effective thermal conductivity of the NPs including the thermal contact resistance, k_{peff}, is given as

$$k_{peff} = \frac{k_p}{1+R_K k_p/d_p} \tag{4.3.10}$$

In these equations, k_{de} is the effective thermal conductivity of the mixture of base fluid and dead-end particles, φ_{de} is the volume fraction of dead-end particles in an aggregate unit, φ_b is the volume fraction of backbone particles in an aggregate unit, L_{ii} and β_{ii} are parameters appearing in the derivation of the effective thermal conductivity of arbitrary isotropic particulate composites with interfacial thermal resistance. The readers are referred to Xu (2014) Section 2.5 for a complete explanation of the meanings of these parameters. Moreover, R_K is the Kapitza resistance, ρ is the density, c_p is the specific heat capacity, and $\overline{T}$ is the time-averaged temperature. The subscripts p and a denote nanoparticle and aggregate unit, respectively. C_c is a correction factor that differs among different NP–base fluid pairings, but the value is around unity. For example, for Al_2O_3–water nanofluids, $C_c = 1.1$. The characteristic time interval τ_a is expressed as

$$\tau_a = \frac{m_a}{3\pi\mu_{bf}R_g} \tag{4.3.11}$$

where R_g is the average radius of the aggregate units:

$$R_g = 2.5 d_p \sqrt{\frac{\varphi}{0.01}} \cdot \sqrt{\frac{T_0}{T}} \tag{4.3.12}$$

Here, d_p is the nanoparticle diameter, and $T_0 = 273\,\mathrm{K}$ is the reference temperature.

In light of experimental evidence, this model is suitable for several types of metal oxide nanoparticles in water with volume fractions up to 10% and mixture temperatures below 350 K (see Figure 4.3.6).

A few different mechanisms have been proposed to explain the anomalous enhancement of the mixture thermal conductivity. For example, the base-fluid molecules form layered structures at the particle surface, acting as thermal bridge between the liquid and the solid particles, and thus enhance the k_{nf} (Yu and Choi, 2003). However, there are strong debates about whether they are really important to the enhanced thermal conductivity.

Viscosity of Nanofluids

Apart from thermal conductivity, the viscosity of nanofluids, μ_{nf}, is also important in the evaluation of heat transfer performance, influencing both the necessary power supply as well as the convective heat transfer coefficient. Specifically, measurements of the rheological properties of nanofluids show that particle loading, temperature, particle shape and size, surfactants, and pH have direct impacts on the viscosity of nanofluids. For example, viscosity increases with higher particle volume fractions, while it decreases with elevated temperatures (see Figure 4.3.7). Still, at low particle volume concentrations, most nanofluids behave as Newtonian-fluid mixtures (Mahbubul et al., 2012; Kleinstreuer, 2014). Due to the aggregation effect of NPs, it is difficult to determine how the particle size would affect the effective viscosity. Surfactants and pH affect μ_{nf} mainly through enhancing the dispersion of nanoparticles.

Einstein (1906) developed a formula for the effective viscosity of fluids containing spherical particles at low volume fractions (<2%). It shows a linear increase in viscosity with particle volume fraction:

$$\frac{\mu_{eff}}{\mu_{bf}} = 1 + 2.5\varphi \tag{4.3.13}$$

where μ_{efl} and μ_b are the viscosities of the mixture and the base fluid, respectively. Brinkman (1952) extended this model to consider moderate particle concentrations up to 4%:

$$\frac{\mu_{eff}}{\mu_{bf}} = (1-\varphi)^{2.5} \tag{4.3.14}$$

Comparison with experimentally measured data showed that the Einstein model and the Brinkman model both tend to underestimate the viscosity of nanofluids. The reasons are that these two models only account for the effect of particle loading on the mixture viscosity, but neglect other important influencing factors such as temperature, particle size, and pH. In addition, newly proposed theoretical formulas are not able to accurately predict the effective viscosity for a wide range of nanofluids (Mahbubul et al., 2012). Theoretical models for the viscosity of the mixture usually rely on the homogeneous-dispersion assumption, which makes it nearly impossible to derive a general correlation that accurately predicts μ_{nf}, given that the aggregation of nanoparticles almost always lead to a heterogeneous mixture. Still, due to the substantial difference among the available experimental data of μ_{nf}, specific empirical correlations are only applicable to one or a few types of nanofluids within specific ranges of pertinent parameters. For example, Corcione (2011) obtained an empirical correlation for μ_{nf} with NP ranges in diameter from 25 to 200 nm, volume fraction 0.01%–7.1%, and temperature 293–333 K. The correlation reads:

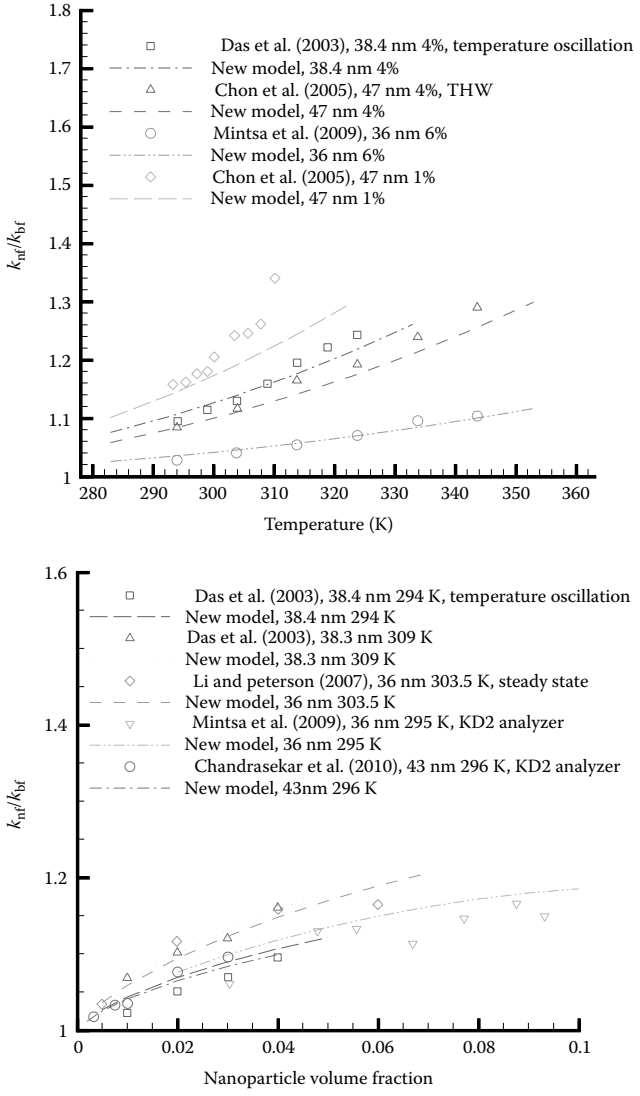


FIGURE 4.3.6 Model prediction of nanofluid thermal conductivity as a function of temperature and particle loading, and comparison with measurements. (Xu, Z. and Kleinstreuer, C., *Energy Convers. Manage.*, 87: 504–512, 2014a.)

$$\mu_{nf} = \frac{\mu_{bf}}{1 - 34.87\left(d_p/d_{bf}\right)^{-0.3}\varphi^{1.03}} \tag{4.3.15}$$

where d_{bf} is the equivalent diameter of a base fluid molecule, given by

$$d_{bf} = \left(\frac{6M}{N\pi\rho_{bf0}}\right)^{1/3} \tag{4.3.16}$$

in which M is the molecular weight of the base fluid, N is the Avogadro number, and ρ_{bf0} is the mass density of the base fluid calculated at temperature $T_0 = 293$ K.

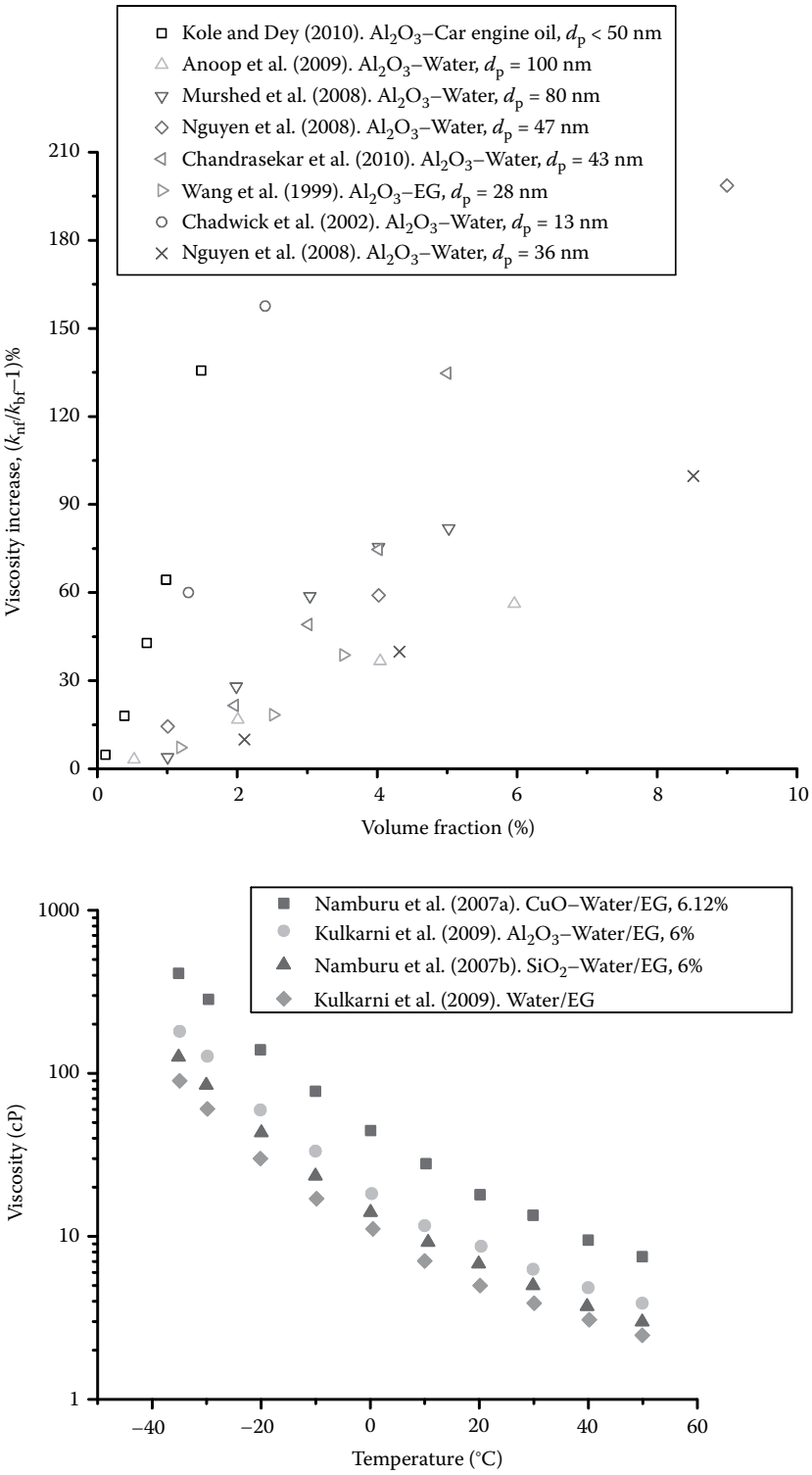


FIGURE 4.3.7 Nanofluid viscosity as a function of particle concentration and temperature.

As mentioned, the implication of elevated viscosities is that higher pressure drops are needed to power nanofluid flows. This drawback may eliminate the gain of enhanced heat transfer in some cases.

Heat Transfer in Nanofluid Flow

For thermal nanofluid flow in conduits, convection heat transfer performance is the most important issue. Heat transfer performance for Fourier-type processes is characterized by the convective heat transfer coefficient or the Nusselt number, that is,

$$h = \frac{q_w}{T_w - T_f} \tag{4.3.17}$$

while

$$\mathrm{Nu}_x = \frac{h \cdot x}{k} \tag{4.3.18}$$

where q_w is the local wall heat flux, and Nu_x is the Nusselt number at location x.

Experimental Evidence

Laminar Flow

The Nusselt number increases with the Reynolds number and the Prandtl number, hence the thermal Peclet number (i.e., $\mathrm{Re} \cdot \mathrm{Pr}$) as well as NP volume fraction (see figures in Book Section 7.1). The enhancement is especially significant in the entrance region, where the nonuniform local particle concentration may be the reason. The heat transfer enhancement reported in many studies ranges from 5% to 45% for spherical nanoparticles; however, for nanofluids with high aspect ratios, such as carbon nanotubes, the enhancement can be much greater (Ding et al., 2006).

It should be noted that the enhancement of the heat transfer coefficient is often found to be larger than that of the thermal conductivity for the same NP-volume fraction. Especially in laminar flows, particle migration effects, for example, thermophoresis, Brownian motion, shear stresses, and/or viscosity gradients are the reasons for the enhanced heat transfer performance. Again, focusing on the entrance region, the thermal conductivity in the wall-shear layer is the main parameter that affects the convection heat transfer rate. Thus, locally enhanced thermal conductivities of nanofluids due to particle migration can substantially increase the convection heat transfer rates. Also, such developing flow fields feature nonuniform thermal conductivity and viscosity profiles as well as reduced boundary layer thicknesses.

Turbulent Flow

For turbulent convective heat transfer, elevated, unchanged, and reduced h_{nf} values, in comparison to h_{bf} values, have all been reported. So, whether the dispersion of nanoparticles in the base fluid enhances the convective heat transfer is still under debate, although the number of studies that observed increased convective heat transfer coefficients with Reynolds number, implied random NP dispersion, and particle volume fraction are in the majority (Kleinstreuer and Feng, 2011; Kleinstreuer and Xu, 2015). Another widely recognized character is the NP size effect on the convective heat transfer coefficient. Larger particles seem to give better h_{nf} results for the same particle volume fraction, possibly due to greater particle slip caused by higher inertia effects. Also, thermophoresis as well as Brownian diffusion of nanoparticles may lead to increased particle–fluid slip, which enhances the convective heat transfer. In the boundary layer, the particle concentration is smaller than in the bulk due to radial flow, which reduces the viscosity and the thickness of the boundary layer, hence promoting heat transfer. Still, some argue that at high Reynolds numbers the

heat transfer is dominated by convection, implying that NPs provide only a small contribution to the overall heat transfer.

The Nusselt number excludes the effect of increased pumping cost due to the dispersion of NPs in the base fluid. Hence, to measure the performance of thermal nanofluid flow, the thermal performance factor may be used:

$$\eta = \frac{\mathrm{Nu}_{\mathrm{nf}}/\mathrm{Nu}_{\mathrm{bf}}}{\left(f_{\mathrm{nf}}/f_{\mathrm{bf}}\right)^{\frac{1}{3}}} \tag{4.3.19}$$

Convection Heat Transfer Correlations

As mentioned, nanofluids have potentially shown superior heat transfer capabilities than the base fluid in both laminar and turbulent regimes. The heat transfer coefficient appears to increase more than the thermal conductivity. The friction factor correlations developed for pure fluids may still be used for nanofluids in both laminar and turbulent flow regimes, given that the corresponding nanofluid properties are used. However, it is under debate whether the conventional correlations for Nusselt number can accurately predict the nanofluid heat transfer characteristics using the measured nanofluid properties (Sarkar, 2011; Sundar and Singh, 2013). New correlations have been proposed to fit the experimental data for Nusselt numbers of nanofluids. Unfortunately, in both laminar and turbulent flow regimes, strong disagreements have been observed between proposed correlations. For forced convection in conduits, the heat transfer coefficient may be an alternative parameter to evaluate nanofluids. For example, Yu et al. (2012) reviewed experimental data sets for turbulent heat transfer of nanofluids in horizontal tubes. They concluded that if properly compared, the actual heat transfer enhancement of nanofluids over their base fluids can be predicted quite accurately using the Dittus–Boelter equation, that is,

$$\mathrm{Nu} = 0.023\mathrm{Re}^{4/5}\mathrm{Pr}^{n} \tag{4.3.20}$$

where $n = 0.4$ for heating (wall hotter than the bulk fluid) and $n = 0.33$ for cooling. It should be noted that while using Equation 4.3.20, the nanofluid properties have to be used to calculate the Reynolds number and the Prandtl number.

Numerical Studies

Numerical simulations of nanofluids allow to depict fluid flow and heat transfer characteristics under different scenarios. The numerical methods used in the literature can be categorized into two groups, that is, the macroscale continuum approach and the meso/microscale approach (see selectively Book Sections 5.1 through 5.9). The continuum approach can be further divided into a single-phase and two-phase approach. The meso/microscale models for nanofluid flow simulations require much smaller length scales, and hence are computationally much more expensive. Examples include the lattice Boltzmann method (LBM), molecular dynamics (MD) simulations, direct numerical simulation (DNS), and dissipative particle dynamics (DPD) (as discussed in Karimipour et al., 2015; Sun et al., 2011; Yang et al., 2014; Abu-Nada, 2015).

Continuum Mechanics Analyses

Single-Phase Approach

In the single-phase approach, it is assumed that the suspended nanoparticles are in thermal equilibrium with the liquid phase, and that the relative velocity between the two phases is negligible. The reasoning is that NPs are so small, say, $d_{\mathrm{p}} < 100$ nm, that they closely follow the fluid streamlines, making the nanofluid to behave like a homogeneous mixture (Kleinstreuer, 2014; among others). Thus, the set of governing equations for pure fluids can be used with the effective thermophysical properties of nanofluids replacing the fluid properties.

Continuity equation:

$$\frac{\partial \rho_{\mathrm{nf}}}{\partial t}+\frac{\partial}{\partial x_i}(\rho_{\mathrm{nf}}u_i)=0 \tag{4.3.21}$$

Momentum equation:
(Laminar flow)

$$\frac{\partial(\rho_{\mathrm{nf}}u_i)}{\partial t}+\frac{\partial}{\partial x_j}(\rho_{\mathrm{nf}}u_iu_j)=-\frac{\partial p}{\partial x_i}+\frac{\partial}{\partial x_j}\left[\mu_{\mathrm{nf}}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)\right] \tag{4.3.22}$$

(Turbulent flow)

$$\frac{\partial(\rho_{\mathrm{nf}}u_i)}{\partial t}+\frac{\partial}{\partial x_j}(\rho_{\mathrm{nf}}u_iu_j)=-\frac{\partial p}{\partial x_i}+\frac{\partial}{\partial x_j}\left[\mu_{\mathrm{nf}}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)-\rho_{\mathrm{nf}}\overline{u_iu_j}\right] \tag{4.3.23}$$

Energy equation:
(Laminar flow)

$$\frac{\partial(\rho_{\mathrm{nf}}h)}{\partial t}+\frac{\partial}{\partial x_i}\left[(\rho c_{\mathrm{p}})_{\mathrm{nf}}u_iT\right]=\frac{\partial}{\partial x_i}\left(k_{\mathrm{nf}}\frac{\partial T}{\partial x_i}\right)+\mu_{\mathrm{nf}}\Phi \tag{4.3.24}$$

(Turbulent flow)

$$\frac{\partial(\rho_{\mathrm{nf}}h_{\mathrm{tot}})}{\partial t}+\frac{\partial}{\partial x_i}(\rho_{\mathrm{nf}}u_ih_{\mathrm{tot}})=\frac{\partial}{\partial x_i}\left(k_{\mathrm{nf}}\frac{\partial T}{\partial x_i}-\rho_{\mathrm{nf}}\overline{u_ih}\right)+\mu_{\mathrm{nf}}\Phi-\frac{\partial u_j}{\partial x_i}(\rho_{\mathrm{nf}}\overline{u_iu_j}) \tag{4.3.25}$$

where $h = u + pv$ is the enthalpy, and Φ is the energy dissipation term, that is,

$$\Phi=\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)\frac{\partial u_i}{\partial x_j} \tag{4.3.26}$$

Clearly, the single-phase approach is easier to implement, requires less computational time, and achieves comparison with selected experimental data sets. However, the results depend strongly on the selected thermophysical property models, especially for the thermal conductivity and viscosity. As mentioned, particle migration can significantly affect the heat transfer results by creating nonuniform concentration fields. They occur most prominently in entrance regions and boundary layers, especially when the Peclet number, $\mathrm{Pe} = \mathrm{Re} \cdot \mathrm{Pr}$, is large.

To improve the single-phase model, some modifications have been applied to include the slip between particles and the base fluid by adding a virtual term in the thermal conductivity expression (Mokmeli and Saffar-Awal, 2010). Also, solving the NP-mass transfer equation together with the momentum and energy equations to account for nonuniform concentration distributions can also increase the accuracy (Kamyar et al., 2012).

Two-Phase Approach

Though ultrafine particles may easily be uniformly suspended, several factors like Brownian motion, thermophoresis, and aggregation can significantly change the relative particle motion, leading to velocity differences between the two phases. Hence, two-phase models, which solve one set of governing equations for each of the phases, may be more suitable (see Book Section 2.8). Most of the

numerical studies used the so-called Eulerian–Eulerian approach to investigate the fluid flow and heat transfer characteristics of nanofluids, where both phases are considered as "fluids" with volume fractions summing up to unity. Specifically, the governing equations for phase α are as follows.

Continuity equation:

$$\frac{\partial(\varphi_\alpha \rho_\alpha)}{\partial t} + \nabla(\varphi_\alpha \rho_\alpha u_\alpha) = 0 \tag{4.3.27}$$

Momentum equation:

$$\frac{\partial}{\partial t}(\varphi_\alpha \rho_\alpha u_\alpha) + \nabla(\varphi_\alpha \rho_\alpha u_\alpha u_\alpha) = -\varphi_\alpha \nabla p + \nabla\left[\varphi_\alpha \mu_\alpha \left(\nabla u_\alpha + (\nabla u_\alpha)^T\right)\right] + F_\alpha \tag{4.3.28}$$

Energy equation:

$$\frac{\partial(\varphi_\alpha \rho_\alpha h)}{\partial t} + \nabla\left[\varphi_\alpha \rho_\alpha u_\alpha h_\alpha\right] = \nabla\left(k_\alpha \nabla T_\alpha\right) + Q_\alpha \tag{4.3.29}$$

where F_α is the surface force due to the existence of another phase and Q_α is the interface heat transfer term. Aside from the inclusion of volume fraction φ, the differences of Equations 4.3.27 through 4.3.29 from Equations 4.3.21 through 4.3.26 are due to interfacial momentum and energy change. Specifically, the interfacial forces that may apply to nanofluid flow are drag, lift, wall lubrication, interphase turbulent dispersion, Brownian diffusion, Magnus, and thermophoresis forces. The rate of interface heat transfer between the nanoparticles (α) and the fluid (β) reads:

$$Q_{\alpha,\beta} = h_{\alpha,\beta} A_{\alpha,\beta}(T_\alpha - T_\beta) \tag{4.3.30}$$

where $A_{\alpha,\beta}$ is the interfacial area per unit volume.

Alternatively, the Eulerian–Lagrangian framework can be employed, which tracks the particles in the fluid phase. Still, a mixture model that introduces a term due to drift of NPs into the momentum equation of the base fluid is feasible (Behroyan et al., 2016). However, the comparisons between different two-phase models suggest no obvious winner in terms of prediction accuracy. Hence the Eulerian two-phase model is favored due to ease of use and low computational cost. It should be noted that although the two-phase models are capable of capturing the nonuniform concentration fields, most studies assumed that the turbulence of the fluid phase is not directly affected by the presence of the nanoparticle phase.

Meso- and Microscale Analyses

High-resolution methods of present research interest are the lattice Boltzmann method, molecular dynamics simulation, and direct numerical simulation. Here, only a brief summary is provided as additional discussions may be found in Book Section 2.8.

Lattice Boltzmann method. In addition to the classical numerical methods, an alternative approach to analyze thermal nanofluid flow is the lattice Boltzmann method (LBM), which fills the gap between microscopic and macroscopic phenomena (Kamyar et al., 2012). In LBM, the discrete Boltzmann equation is solved to simulate nanofluid flow with a suitable particle collision model. This computational fluid dynamics-method is advantageous for systems with complex geometric boundaries and for massive parallel processing. Despite the high computational cost, LBM has shown satisfactory results in thermal nanofluid flow applications. Due to the limited studies available, more research may be needed using this method to see the applicability as well as accuracy.

Molecular dynamics simulation. Molecular Dynamics simulation solves Newton's equation of motion for a system of particles interacting with a prescribed potential. It is able to directly calculate the movement of particles at the atomic level, providing insights from the microscopic scale through statistical processing (Sun et al., 2011). Therefore, MD simulation can be used to directly calculate

the thermal conductivity of nanofluids. Assuming that the micromixing effect is the main reason for the enhanced thermal conductivity of nanofluids, MD simulations show that k_{nf} strongly depends on the shear rate, γ. For large shear rates, the rotation of the particles, being subjected to strong velocity gradients, will be higher. Also, k_{nf} increases with the particle Peclet number, where $\mathrm{Pe} = \gamma d^2/\alpha$, and $\alpha = k/\rho c_p$ is the thermal diffusivity (Sun et al., 2011). The thermal conductivity ratios obtained from such simulations significantly exceed unity. As can be expected, the increase in thermal conductivity ratio with shear rate is less for higher particle loadings due to slower particle rotation.

Direct numerical simulation. Direct numerical simulation can resolve the flow field down to the smallest temporal and spatial scales, even for turbulent flows. Hence, it can directly simulate the forces and interactions between individual NPs as well as the coupling between the particles and the fluid, provided that the mesh spacing is finer than the nanoparticles. This eliminates many of the simplifications made for the theoretical closure models and provides a better understanding to the mechanisms of heat transfer enhancement in nanofluids. For example, the Brownian force, thermophoresis force, and van der Waals force experienced by nanoparticles (and the resulting NP aggregates), as well as the two-way temperature coupling between the continuous phase and the discrete phase can be directly simulated (Kondaraju et al., 2010). Thus, the effective thermal conductivity of the nanofluid can be directly calculated.

Natural Convection Heat Transfer with Nanofluids

As discussed, in *forced* convection applications, nanofluids have generally shown superior heat transfer performance when compared to their base fluids. In *natural* convection, however, the thermal performance of nanofluids is limited. In fact, experimental evidence shows that adding NPs to the base fluid deteriorates natural convection heat transfer either systematically or when a critical particle concentration is used (Haddad et al., 2012). For example, the Nusselt number of nanofluid natural convection decreases with increasing particle concentration for a given Rayleigh number Ra, where $\mathrm{Ra} = g\beta_{nf}\Delta T L^3/\nu_{nf}\alpha_{nf}$, with β being the thermal expansion coefficient, ΔT the temperature difference between the heat transfer surface and the fluid, L a characteristic length, ν the kinematic viscosity, and α the thermal diffusivity. This effect was attributed to a number of factors, including sedimentation of NPs, particle–fluid slip, and modifications of the dispersion properties. Clearly, according to its definition, the Rayleigh number is reduced by an increase in the mixture kinematic viscosity. Owning to this effect, the stress–strain relation of the nanofluid will have a substantial influence on the natural convection heat transfer of nanofluids. This can be seen from the results of numerical studies concerning nanofluid natural convection heat transfer. In computer simulations of single phase natural convection, nanofluids have been shown to be more stable (i.e., delayed onset of convection), exhibiting higher heat transfer rates at least for low Ra numbers than the base fluids. The Nu number increases with NP volume fraction up to a critical Ra number. Hence, using an accurate μ_{nf} model is critical when simulating single-phase nanofluid natural convection. On the other hand, some studies adopted a two-phase mixture model for numerical evaluation of nanofluid natural convection. For example, the model developed by Buongiorno (2006) considers particle–fluid slip caused by Brownian diffusion and thermophoresis. The critical Rayleigh number of nanofluids (i.e., when flow instability occurs) is shown to be much lower than that of the base fluids.

Entropy Generation in Thermal Nanofluid Flow

The use of thermal nanofluid flow for enhanced heat transfer has been the major focus for most studies. However, for efficient removal of high heat fluxes and proper optimization of thermal systems, not only the heat transfer has to be maximized but the increase in entropy generation has to be minimized as well. Entropy generation is the measure of process irreversibility caused by both heat transfer and friction effects. It can be employed as a criterion to assess the performance of thermal devices. For convection heat transfer, the local entropy-generation rate (S_{gen} in W/K) can be expressed as (Bejan, 1996; Kleinstreuer, 2014; among others):

$$S_{\text{gen}} = S_{\text{gen}}(\text{thermal}) + S_{\text{gen}}(\text{friction}) \tag{4.3.31}$$

Clearly, Equation 4.3.31 encapsulates the irreversibility due to heat transfer and frictional effects, where

$$S_{\text{gen}}(\text{thermal}) = \frac{k}{T^2}\left[\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2\right] \tag{4.3.32a}$$

and

$$\begin{aligned} S_{\text{gen}}(\text{frictional}) &= \frac{\mu\Phi}{T} = \frac{\mu}{T}\left(\overline{\Phi} + \Phi'\right) \\ &= \frac{\mu}{T}\left[\left(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i}\right)\frac{\partial \overline{u_i}}{\partial x_j} + \overline{\frac{\partial u_i}{\partial x_j}\frac{\partial u_i}{\partial x_j}}\right] \end{aligned} \tag{4.3.32b}$$

Bejan (1996) pointed out that for turbulent flow the dissipation due to mean flow is roughly the same as that due to turbulent fluctuation. In fact, the frictional entropy generation rate accounts for a negligible part of S_{gen}, unless the Reynolds number becomes very large. It should be noted that when using the previous equations for nanofluids, the corresponding nanofluid properties should be used.

Adding nanoparticles to the base fluid reduces entropy generation in the laminar regime of nanofluid flow in conduits, with an optimal volume fraction corresponding to minimized entropy generation (Li and Kleinstreuer, 2010). Also, a nanofluid decreases the total generated entropy of the system without significantly increasing the required pumping power (Feng and Kleinstreuer, 2010). Although adding NPs to the base fluid increases both k and μ in Equation 4.3.32, it also reduces the temperature and velocity gradients, which leads to a reduction in S_{gen} (Xu and Kleinstreuer, 2014b). Thus, from a perspective of the second law of thermodynamics, adding NPs to the base fluid can be very beneficial. It simply decreases the entropy generation rate, especially in the low-to-medium Reynolds number regime. When the Reynolds number is very high, the part of frictional entropy generation contributes more to the total entropy generation than the thermal part, pushing up the entropy generation rate due to the augmented nanofluid viscosity (Mahian et al., 2013). Clearly, the thermophysical properties of nanofluids play an important role in determining the entropy generation rate.

Discrepancies among Enhanced Heat Transfer Studies Using Nanofluids

Flow of nanofluids is a relatively new concept, and experimental/theoretical research efforts are still ongoing. As a result, there have been strong debates on almost all aspects of nanofluid-flow studies when applied to enhanced heat transfer (Prabhat et al., 2012). Perhaps the most notable one is the International Nanofluid Property Benchmark Exercise (Buongiorno et al., 2009), in which the thermal conductivities of identical samples of stable nanofluids were measured by more than 30 organizations worldwide, using a variety of measurement techniques. They found that when normalized to the measured base fluid thermal conductivity, the thermal conductivity of the nanofluid can be well predicted by the effective medium theory (Maxwell, 1881), suggesting no anomalous enhancement of thermal conductivity. In addition, controversies can be found over nanofluid viscosity, heat transfer coefficient, Nusselt number, and pressure drop. During the first few years, experimental studies often showed enhanced values for the thermal conductivity and heat transfer coefficient (Kleinstreuer and Xu, 2015; Mohamad, 2015). However, in recent years, a lot of studies reported that there is negligible increase of these parameters, occasionally even a decrease when adding nanoparticles to liquids. The controversy over nanofluid viscosity is especially noticeable. For example, some studies found positive relationships between particle diameter and viscosity, and some found negative ones. Even theoretical analyses focusing on the same mechanisms resulted

in contradictory conclusions. Though Brownian motion and aggregation are considered the most important characteristics of nanofluids, *how* they influence the thermophysical properties and hence the heat transfer performance of nanofluids is still under debate.

Nanofluid Flow Applications

Nanofluids have abundant potential applications in a variety of thermal processes occurring in conduits. It would be impractical to exhaust the list here; therefore, only a few important applications are discussed in this section. Clearly, nanofluids were first investigated to achieve enhanced heat transfer for cooling purposes. Hence, the heat transfer performance of a nanofluid in a double-tube heat exchanger will be presented as an example (see also Book Section 4.2). In contrast, magnetic nanofluids have great potential in certain unique thermal applications due to its distinctive properties. Novel applications of nanofluids pertaining to internal flow thermal processes occur in biomedical engineering, as demonstrated with discussions of two state-of-the-art research directions.

Nanofluid Flow in a Heat Exchanger

Thermal nanofluid flow in a double-tube heat exchanger is presented to evaluate the heat transfer performance of an alumina–water nanofluid. The configuration is depicted in Figure 4.3.8 as part of a novel solar collector system (Kleinstreuer and Xu, 2015). Hot water flows in the annular, cold nanofluid flowing in the inner tube exchanges heat with water through a copper wall in between. The convectional dynamic viscosity model by Brinkman (1952) and the empirical model by Corcione (2011) were compared. The effective thermal conductivity model developed by Xu and Kleinstreuer (2014a) was employed, which relied on the single-phase approach for simulations.

Figure 4.3.9 compares the pressure gradients at different Reynolds numbers for water and nanofluids with different NP volume fractions. The conventional Brinkman viscosity model underestimates the pressure drop, especially when the NP loading is high. On the other hand, the nanofluid yields a higher heat transfer coefficient than pure water at the same Reynolds number. Meanwhile, the heat transfer coefficient predicted using the Corcione model is higher than that using the Brinkman model.

Magnetic Nanofluids

Magnetic nanofluids (MNFs) or ferrofluids are nanofluids that consist of a nonmagnetic base fluid and magnetic nanoparticles. The NPs are made from ferromagnetic materials such as iron, nickel, and cobalt, or their oxides such as spinel-type ferrites, and magnetite (Fe_3O_4). In order to prevent the coagulation, particles are coated with surfactants, either long chained molecules, or an electrostatic layer. In the absence of a magnetic field, the fluid behaves as a normal single component fluid. When an external magnetic field is applied, the fluid is magnetized, and the apparent fluid properties such as viscosity or thermal conductivity may be changed (see Figure 4.3.10). In fact, the properties of magnetic nanofluids may be controlled and tuned based on application requirements

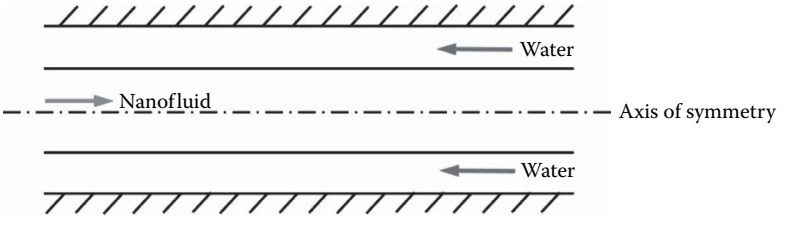


FIGURE 4.3.8 Schematics of a counterflow, double-tube heat exchanger.

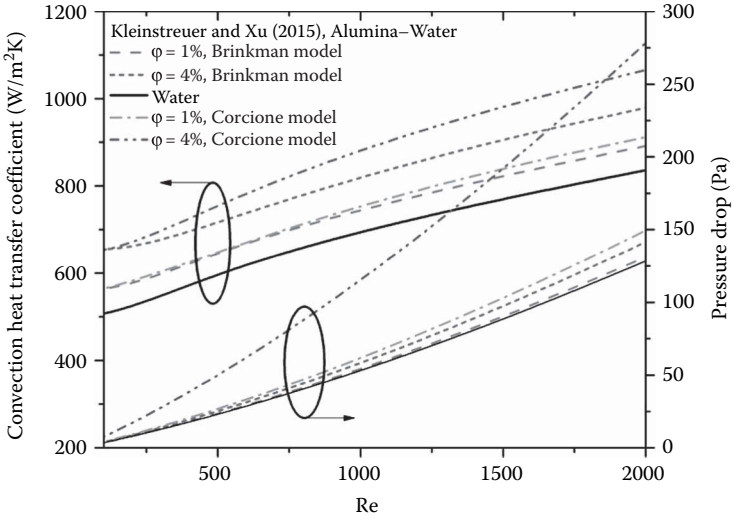


FIGURE 4.3.9 Pressure drop and convection heat transfer coefficient as a function of Re number.

(Shima et al., 2009). The magnetic field can be applied in a number of different ways depending on application requirements (see Figure 4.3.11).

Magnetic nanofluid behaves like a paramagnetic material. In the absence of magnetic field, the magnetic moments of particles are randomly orientated due to the Brownian motion, and the fluid has no net magnetization. When an external magnetic field is applied, the magnetic moments align along the magnetic field direction, generating net magnetization to the fluid. In the presence of a magnetic field, the governing equations of the fluid flow and heat transfer need to be modified to include the Lorentz force and the corresponding energy. The continuity equation remains unchanged, while the momentum equation becomes

$$\rho\frac{D\vec{v}}{Dt}=-\nabla p+\nabla\cdot\vec{\vec{\tau}}+\rho\vec{g}+\vec{J}\times\vec{B} \tag{4.3.33}$$

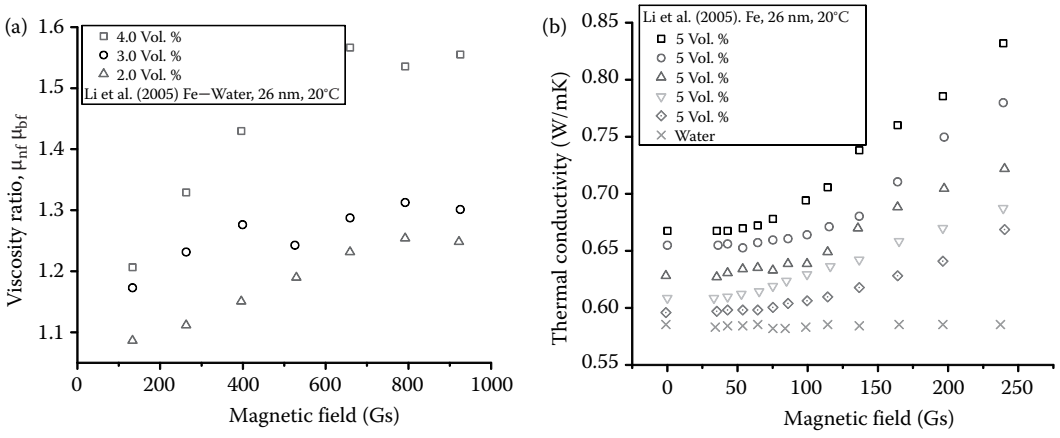


FIGURE 4.3.10 Properties of MNFs: (a) Relative viscosity of Fe–water magnetic fluid parallel to the magnetic field; and (b) thermal conductivity of Fe–water magnetic fluid in the magnetic field parallel to temperature gradient.

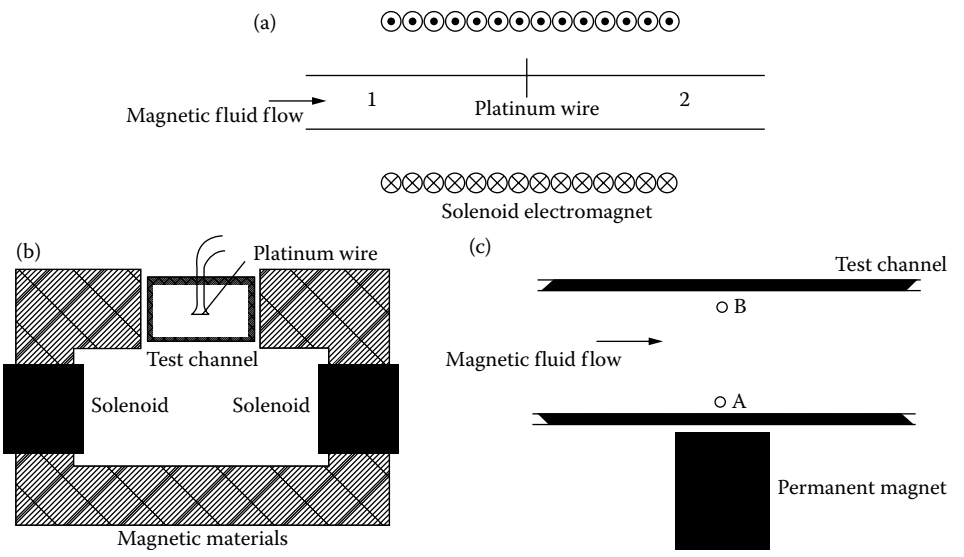


FIGURE 4.3.11 Schematic diagram of magnetic system (*not in scale*): (a) Solenoid electromagnet (parallel field); (b) electromagnet (perpendicular field); and (c) permanent magnet (perpendicular field). (From Li, Q. and Xuan, Y., *Exp. Therm. Fluid Sci.*, 33 (4): 591–596, 2009.)

where $\vec{v}$ is the velocity field, p is the pressure, D/Dt is the total derivative, $\vec{g}$ is the gravitational acceleration, $\vec{\vec{\tau}}$ is the deviatoric stress tensor, $\vec{J}$ is the electric current density, and $\vec{B}$ is the magnetic flux density. Owing to Ampere's law, the Lorentz force term $\vec{J}\times\vec{B}$ can be written in the form

$$\vec{J}\times\vec{B}=\frac{(B\cdot\nabla)B}{\mu_0}-\nabla\left(\frac{B^2}{2\mu_0}\right) \tag{4.3.34}$$

The energy equation has the form

$$\left[\rho c_V-\mu_0\vec{H}\cdot\left(\frac{\partial M}{\partial T}\right)\right]\left(\frac{\partial T}{\partial t}+\vec{v}\cdot\nabla T\right)+\mu_0 T\left(\frac{\partial M}{\partial T}\right)\cdot\frac{\partial\vec{H}}{\partial T}=k_{\mathrm{nf}}\nabla^2 T+\mu_{\mathrm{nf}}\Phi \tag{4.3.35}$$

where $\mu_0=4\pi\times10^{-7}\,\mathrm{H}\cdot\mathrm{m}^{-1}$ is the magnetic constant, $\vec{H}$ is the magnetic field intensity, M is the magnetization of the fluid, and $\mu_{\mathrm{nf}}\Phi$ is the viscous dissipation term. As such, the fluid flow and heat transfer of magnetic nanofluids can be manipulated by applying different magnetic fields.

The thermophysical properties of MNFs are different with and without the presence of a magnetic field. In addition, the strength and orientation of the applied magnetic field affect the properties (see Figure 4.3.12). For example, applying a magnetic field parallel to the temperature gradient will significantly enhance the thermal conductivity along that direction (Bahiraei and Hangi, 2015). This is probably due to the particle alignment and chain-like structure formation in that direction. Moreover, the thermal conductivity increases with the magnetic field strength.

MNF-convection heat transfer in conduits can be controlled by applying magnetic fields to all or a specific section of the flow field, which provides great flexibility. This influence is due both to the modified thermophysical properties of the MNFs and to the motion and migration of nanoparticles under the magnetic field. Using alternating and nonuniform magnetic fields to tune the heat transfer characteristics of MNFs is feasible. For example, increasing the frequency of the alternating magnetic field can lead to greater heat transfer enhancement, as shown in Figure 4.3.12b. Applying a magnetic field such that the field gradient is along the main flow will enhance the convection heat transfer, while a field gradient apposite to the main stream suppresses the heat transfer (see Figure 4.3.12a).

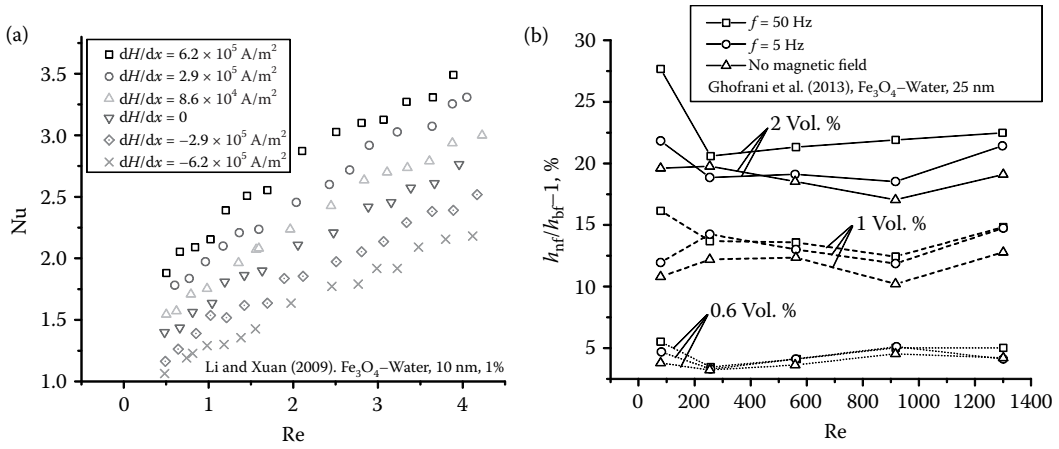


FIGURE 4.3.12 Heat transfer performance of MNFs (a) Nusselt numbers as a function of the Re number with the magnetic field gradient along or opposite to the main flow direction of the magnetic fluid, and (b) variation of the enhancement of heat transfer coefficient with Reynolds number as a function of the frequency of the alternating magnetic field and nanoparticle volume concentration.

Meanwhile, using a magnetic field with negative gradient along the flow stream requires a higher pumping power, while a magnetic field with positive gradient acts as a pump (Aminfar et al., 2011).

Aside from thermal engineering, magnetic nanofluids have also broad applications in biomedical engineering (Kleinstreuer et al., 2008, 2014; Sharifi et al., 2012; Lin et al., 2013; Miaskowski and Sawicki, 2013) and electronics (Li et al., 2008). An interesting application of MNFs is due to the unique characteristics of a specific type of MNF, that is, temperature-sensitive magnetic fluids whose magnetization is dependent on temperature. This property gives rise to thermomagnetic convection, where the MNF continuously circulates in the conduits without additional energy consumption. This is particularly useful in cases where pressure-driven flows are not feasible or inadequate, as the driving force here is the magnetic force induced by a temperature difference. Hence, higher heat load will result in greater velocity and higher heat transfer coefficient (Mukhopadhyay et al., 2005). An important biomedical application is external magnetic field-guided delivery of nanodrugs in either the pulmonary tract (Dames et al., 2007; Azarmi et al., 2008; Pourmhran, et al. 2015) or the cardiovascular system (Arruebo et al., 2007; Veiseh et al., 2010). Another example is cancer treatment taking advantage of magnetic nanofluid hyperthermia (Lin et al., 2013).

It should be noted that in the applications using MNFs, the long-term stability of the suspension is of uttermost importance, because formation of aggregates can lead to unexpected instabilities and system failure.

Nanodroplet–Vapor–Air Mixture Dynamics in Models of Human Lung Airways

As discussed, nanofluids have broad applications in biomedical engineering. For instance, *direct* nanodrug delivery has been extensively studied in recent years, believing to be a major future treatment option for many diseases. A specific application is the convection heat transfer with phase change of inhaled nanodroplets in lung airways (see Figure 4.3.13). Examples include administered drug-aerosols using inhalers, fume- or mist-nanodroplet inhalation from fuel handling, welding operation, spray processes as well as smoke-aerosol inhalation from conventional cigarettes and electronic cigarettes. Their transport and conversion phenomena involve complex physical and thermodynamic interactions between air, vapors, and nanodroplets with their therapeutic (or toxic) compounds and the respiratory tract walls. This is an area of interdisciplinary applications of modern thermal internal nanofluid-flow engineering. A basic mathematical model can be outlined as follows.

Any given air–vapor mixture is described as a single continuous phase, that is, the conservation laws for mass, momentum, and energy, describing the air–vapor mixture transport and the advection-diffusion equation describing vapor-species transport. Droplet–vapor interactions, that is, liquid–vapor mass change due to evaporation or condensation, are implemented by: (1) introducing source terms into the energy equation and vapor-species transport equations (see Equations 4.3.36 through 4.3.39); and (2) employing the local vapor-mass fraction in the droplet mass conservation equation (see Equation 4.3.41). Specifically, the energy equation of the air–vapor mixture reads:

$$\frac{\partial\left(\rho c_p T\right)_{a-v}}{\partial t}+\frac{\partial\left(\rho c_p u_j T\right)_{a-v}}{\partial x_j}=$$

$$\frac{\partial}{\partial x_j}\left[\left(k_{a-v}+\frac{\rho_{a-v}c_{a-v,p}\nu_T}{\mathrm{Pr}_T}\right)\frac{\partial T_{a-v}}{\partial x_j}\right]+\Phi+\frac{\partial}{\partial x_j}\left[\sum_{s=1}^{N}h_s\rho_{a-v}\left(\tilde{D}_{a-v}+\frac{\nu_T}{\mathrm{Sc}_T}\right)\frac{\partial Y_{v,s}}{\partial x_j}\right]+S_{v-d}^{(E)} \quad (4.3.36)$$

where the energy source term $S_{v-d}^{(E)}$ is due to the latent heat of evaporation or condensation, which is released or absorbed by the droplets per local mesh cell:

$$S_{v-d}^{(E)}=\sum_{s=1}^{4}S_{v-d,s}^{(E)}=\left\{\sum_{i=1}^{N_{d,\mathrm{cell}}}\left[\left(\sum_{s=1}^{4}L_s\overline{j_s}\right)A_d\right]_i\right\}/V_{\mathrm{cell}} \quad (4.3.37)$$

where $N_{d,\mathrm{cell}}$ is the total droplet number in a specified mesh cell.

The governing equation for advection and diffusion of the sth vapor species is expressed as

$$\frac{\partial(\rho_{a-v}Y_{v,s})}{\partial t}+\frac{\partial}{\partial x_j}\left(\rho_{a-v}u_jY_{v,s}\right)=\frac{\partial}{\partial x_j}\left[\rho_{a-v}\left(\tilde{D}_{a-v,s}+\frac{\nu_T}{\mathrm{Sc}_T}\right)\frac{\partial Y_{v,s}}{\partial x_j}\right]+S_{v-d,s}^{(Y)} \quad (4.3.38)$$

where $\mathrm{Sc}_T=0.9$ is the turbulence Schmidt number for $Y_{v,s}$, ν_T is the turbulent viscosity, and $\tilde{D}_{a-v,s}$ is the molecular diffusivity of the sth vapor species in the air–vapor mixture. The local vaporized/condensed vapor-mass flow rate of the aerosol components are added to its advection-diffusion equation as a source term $S_{v-d,s}^{(Y)}$ (kg m^{-3} s^{-1}), that is,

$$S_{v-d,s}^{(Y)}=\int_{t_{i,\mathrm{start}}}^{t_{i,\mathrm{start}}+\Delta t_f}\left(\sum_{i=1}^{N_{d,\mathrm{cell}}}\left(\overline{j_s}A_d\right)_i\right)dt_d\,/\left(V_{\mathrm{cell}}\Delta t_f\right) \quad (4.3.39)$$

where $\overline{j_s}$ is the average evaporation/condensation mass flux normal to the droplet surface of the sth component (i.e., $\overline{j_s}>0$ for evaporation and $\overline{j_s}<0$ for condensation), and A_d is the droplet surface area. Also, $\mathrm{d}t_\mathrm{d}$ represents the droplet phase time differential and Δt_f is the flow time step.

For the multicomponent droplet size change and transport, the governing equations for discrete droplets are the translational equation as well as the mass and energy conservation laws.

Newton's second law of motion states:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m_d u_{d,i}\right)=F_i^D+F_i^L+F_i^{BM}+F_i^G \quad (4.3.40)$$

where F_i^D, F_i^L, F_i^{BM}, and F_i^G are the drag force, lift force, Brownian motion-induced force, and gravity, respectively.

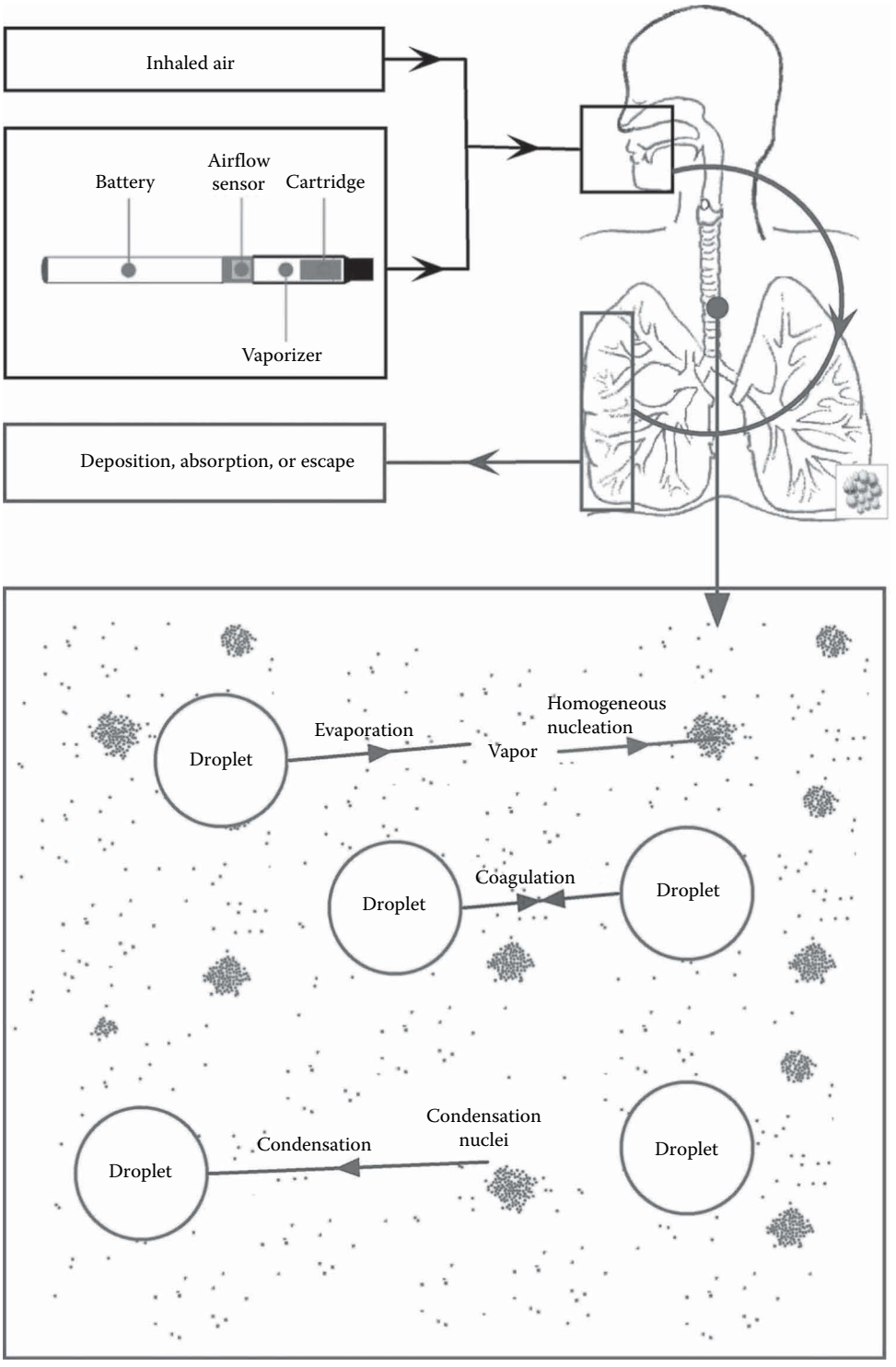


FIGURE 4.3.13 Illustration of inhalation of multicomponent droplet–vapor mixtures and their fate in human lung airways. (From Feng, Y. and Kleinstreuer, C., A novel computational fluid-particle dynamics model for the simulation of multicomponent droplet-vapor transport/deposition in an idealized human upper airway configuration. In *2015 Annual Meeting of the Biomedical Engineering Society*, Tampa, FL, 2015.)

Droplet mass changes due to condensation/evaporation can be described as

$$\frac{\mathrm{d}m_d}{\mathrm{d}t} = -\sum_{s=1}^{m} \int_{\text{surf}} j_s \, dA \approx -\sum_{s=1}^{m} (\overline{j_s} \cdot A_d) \tag{4.3.41}$$

where the average mass flux $\overline{j_s}$ is given by Turns (1996):

$$\overline{j_s} = \rho_{a-v} \cdot C_m \cdot \mathrm{Sh} \cdot \tilde{D}_{a-v,s} \cdot d_d^{-1} \ln\left[\left(1 - Y_{v,s,\,\text{cell}}\right) / \left(1 - Y_{v,s,\,\text{surf}}\right)\right] \tag{4.3.42}$$

Here, Sh is the Sherwood number (see the Nomenclature for definition), while $Y_{v,s,\text{surf}}$ and $Y_{v,s,\text{cell}}$ are the mass fractions of the sth vapor phase at the droplet surface and at the center of the cell, where the droplet currently resides. The correction factor C_m for submicron droplets, considering noncontinuum effects, can be expressed as

$$C_\text{m} = \frac{1 + \mathrm{Kn}}{1 + \left(\dfrac{4}{3\alpha_\text{m}} + 0.377\right)\mathrm{Kn} + \dfrac{4}{3\alpha_\text{m}}\mathrm{Kn}^2} \tag{4.3.43}$$

where Kn is the Knudsen number, $\mathrm{Kn} = 2\lambda_{a-v}/d_d$, in which λ_{a-v} is the mean free path of the air–vapor mixture surrounding the droplet, and α_m is the mass accommodation coefficient, where $\alpha_\text{m} = 1$ was used in the present study (Hinds, 1999). The derivation of Equation 4.3.42 is based on the assumption that the distance between the droplet mass center and the mesh cell center is much larger than the droplet radius. Specifically, $Y_{v,s,\text{cell}}$ is determined by the advection-diffusion equation (see Equation 4.3.38), while $Y_{v,s,\text{surf}}$ can be calculated by

$$Y_{v,s,\,\text{surf}} = \frac{\rho_{v,s,\,\text{surf}}}{\rho_{a-v}} = \frac{\left(P_{v,s,\,\text{surf}} / \Re_s T_d\right)}{\rho_{a-v}} = \frac{\gamma_s \cdot K_s \cdot X_{s,d} \cdot P_{v,s,\text{sat}}(T_d)}{\rho_{a-v}\left(\Re / M_s\right) T_d} \tag{4.3.44}$$

Here, $P_{v,s,\text{surf}}$ is the equivalent density of vapor species s in the air–vapor mixture at the droplet surface; γ_s is the activity coefficient of species s, which is a correction of the evaporation/condensation characteristics of a certain liquid component due to the molecular bonding changes in the pure droplets and in the multicomponent droplets (Tu and Ray, 2005); $X_{s,d}$ is the mole fraction of the sth component in the droplet; $\Re_s$ is the species gas constant; T_d is the droplet temperature; and $P_{v,s,\text{sat}}$ is the temperature-dependent saturation pressure of the pure sth species. Clearly, parameters γ_s and γ_s serve as corrections to the vapor pressure at the surface of multicomponent droplets (Raoult's law), while K_s is the correction factor for the Kelvin effect (Hinds, 1999):

$$K_s = \exp[4\sigma_s M_s / (\Re \rho_s d_d T_d)] \tag{4.3.45}$$

where σ_s is the surface tension of component s at the droplet surface.

The droplet heat transfer equation takes the form:

$$m_d c_{p,d}\left(\frac{\mathrm{d}T_d}{\mathrm{d}t}\right) = C_T\left(k_{a-v} \cdot \frac{\mathrm{Nu}}{d_d}\right)(T_{a-v,\text{cell}} - T_d)A_d - \left(\sum_{s=1}^{N} L_s \overline{j_s}\right)A_d \tag{4.3.46}$$

where L_s is the latent heat of liquid–vapor phase transition of the sth species (L_s is always larger than zero). Also, C_T is the correction factor for submicron droplets:

$$C_T = \frac{1 + \mathrm{Kn}}{1 + \left(\dfrac{4}{3\alpha_T} + 0.377\right)\mathrm{Kn} + \dfrac{4}{3\alpha_T}\mathrm{Kn}^2} \tag{4.3.47}$$

where α_T is the thermal accommodation coefficient with the value of $\alpha_T = 1$ (Hinds, 1999).

The governing equations of the transport and phase change of the multicomponent droplet–vapor mixtures can be solved using computational fluid-particle dynamics (CF-PD) methods (Kleinstreuer, 2014). For example, the transient transport and conversion phenomena of electronic cigarette aerosol under nonisothermal conditions leading to local particle/droplet/vapor depositions and possible species mass transfer into systemic regions was studied by Feng and Kleinstreuer (2015). The transient puff topography is shown in Figure 4.3.14, whereas the deposition patterns of both the droplets and the vapor are shown in Figure 4.3.15.

Future Work

Nanofluids have been studied for more than two decades, especially for enhanced cooling applications (see also Book Section 4.20). However, there are still a lot of questions that remain to be answered. Specifically, the fundamental physics underlying its anomalous characteristics, if any, and the reliability and robustness of its applicability require further attention.

Because contradictory results have been reported, partially due to different preparation methods and measurement procedures adopted, standardized methods for synthesis and characterization of nanofluids are needed to prove reliability and repeatability.

The macroscale characteristics of nanofluids are due to microscale effects that are still elusive to a certain extent. Hence, to understand and predict the behavior of nanofluids, the microscale physics need to be recognized. This requires more experimental studies and high-fidelity simulations that can reveal the interactions between nanoparticles and the base fluid that may give rise to any anomalous behavior of nanofluids in forced convection heat transfer.

A major issue in real systems and engineering applications with nanofluids is instability, which is difficult to model and predict theoretically. The instabilities are often absent when systems are studied using idealized models. The reason is that the unstable behavior of a complex system is often caused by secondary effects, which are neglected in simplified models. Therefore, more experimental measurements and theoretical analyses are needed to determine the most important effects that need to be considered in modeling efforts.

Nanofluid-flow applications with heat transfer in biomedical engineering and toxicology have enormous potential, as indicated in this section. Proof-of-concepts have been so far the focus of

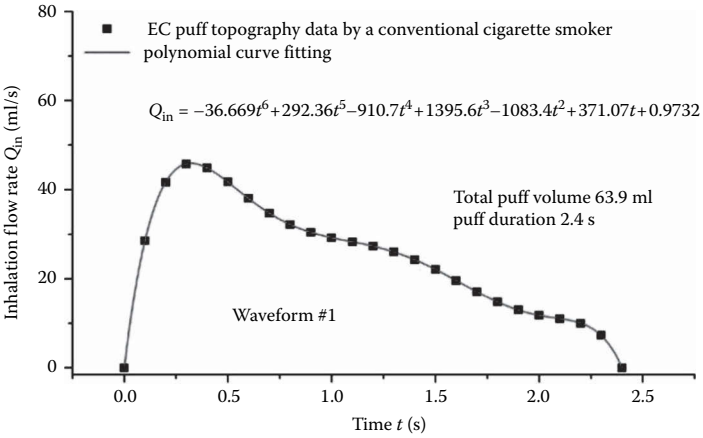


FIGURE 4.3.14 Transient puff topographies. (From Feng, Y. and Kleinstreuer, C., A novel computational fluid-particle dynamics model for the simulation of multicomponent droplet-vapor transport/deposition in an idealized human upper airway configuration. In *2015 Annual Meeting of the Biomedical Engineering Society*, Tampa, FL, 2015.)

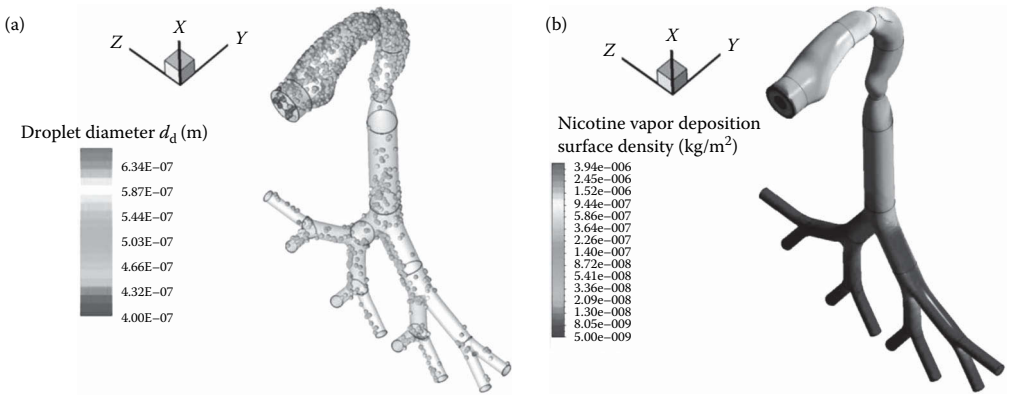


FIGURE 4.3.15 (a) Droplet and (b) nicotine vapor deposition patterns after $t = 2.4$ s. (From Feng, Y. and Kleinstreuer, C., A novel computational fluid-particle dynamics model for the simulation of multicomponent droplet-vapor transport/deposition in an idealized human upper airway configuration. In *2015 Annual Meeting of the Biomedical Engineering Society*, Tampa, FL, 2015.)

most studies. To evaluate real-world performance and applicability of internal thermal nanofluid flow in various scenarios, basic experimental, theoretical, and numerical studies are most desirable.

Conclusions

Nanofluid flow in conduits has shown remarkable convective heat transfer performance mainly due to the small particle size. Brownian motion and nanoparticle aggregation may be the two most important effects that give rise to potentially superior thermal properties of nanofluid flow. Additional factors include nanoparticle loading, particle–liquid slip, temperature, particle size and shape, pH, and favorable aggregation. However, long-term stability of nanofluids is difficult to achieve due to random aggregation effects. Nanofluid flow is still an active field of study, partly because of the contradictory results concerning its most important characteristics. These controversial studies put the effectiveness of nanofluids as coolants in debate and cast a doubt on its engineering applicability. Hence, high-resolution studies are needed to better understand nanofluid flow dynamics and to determine its effectiveness in enhanced cooling or heating.

In contrast, thermal internal nanofluid-flow applications in biomedical engineering and toxicology are, without doubt, one of the important research areas in modern thermal fluid-particle dynamics.

REFERENCES

Abu-Nada, E. 2015. Dissipative particle dynamics simulation of natural convection using variable thermal properties. *International Communications in Heat and Mass Transfer*, 69: 84–93.

Aminfar, H., Mohammadpourfard, M., and Kahnamouei, Y.N. 2011. A 3D numerical simulation of mixed convection of a magnetic nanofluid in the presence of non-uniform magnetic field in a vertical tube using two phase mixture model. *Journal of Magnetism and Magnetic Materials*, 323 (15): 1963–1972.

Anoop, K.B., Kabelac, S., Sundararajan, T., and Das, S.K. 2009. Rheological and flow characteristics of nanofluids: Influence of electroviscous effects and particle agglomeration. *Journal of Applied Physics*, 106 (3): 034909.

Arruebo, M., Fernández-Pacheco, R., Ibarra, M.R., and Santamaría, J. 2007. Magnetic nanoparticles for drug delivery. *Nano Today*, 2 (3): 22–32.

Azarmi, S., Roa, W.H., and Löbenberg, R. 2008. Targeted delivery of nanoparticles for the treatment of lung diseases. *Advanced Drug Delivery Reviews*, 60 (8): 863–875.

Bahiraei, M. and Hangi, M. 2015. Flow and heat transfer characteristics of magnetic nanofluids: A review. *Journal of Magnetism and Magnetic Materials*, 374: 125–138.

Bang, I.C. and Chang, S.H. 2005. Boiling heat transfer performance and phenomena of Al_2O_3-water nano-fluids from a plain surface in a pool. *International Journal of Heat and Mass Transfer*, 48 (12): 2407–2419.

Barbés, B., Páramo, R., Blanco, E., Pastoriza-Gallego, M.J., Piñeiro, M.M., Legido, J.L., and Casanova, C. 2013. Thermal conductivity and specific heat capacity measurements of Al_2O_3 nanofluids. *Journal of Thermal Analysis and Calorimetry*, 111 (2): 1615–1625.

Behroyan, I., Vanaki, S. M., Ganesan, P., and Saidur, R. 2016. A comprehensive comparison of various CFD models for convective heat transfer of Al_2O_3 nanofluid inside a heated tube. *International Communications in Heat and Mass Transfer*, 70: 27–37.

Bejan, A. 1996. *Entropy Generation Minimization: The Method of Thermodynamic Optimization of Finite-Size Systems and Finite-Time Processes*. Boca Raton, FL: CRC Press. Chapter 3.

Branson, B.T., Beauchamp, P.S., Beam, J.C., Lukehart, C.M., and Davidson, J.L. 2013. Nanodiamond nanofluids for enhanced thermal conductivity. *ACS Nano*, 7 (4): 3183–3189.

Brinkman, H.C. 1952. The viscosity of concentrated suspensions and solutions. *The Journal of Chemical Physics* 20: 571–571.

Buongiorno, J. 2006. Convective transport in nanofluids. *Journal of Heat Transfer*, 128 (3): 240–250.

Buongiorno, J., Venerus, D. C., Prabhat, N., McKrell, T., Townsend, J., Christianson, R., Tolmachev, Y.V., Keblinski, P., Hu, L., Alvarado, J.L., Bang, I.C., Bishnoi, S.W., Bonetti, M., Botz, F., Cecere, A., Chang, Y., Chen, G., Chen, H., Chung, S.J., Chyu, M.K., Das, S.K., Di Paola, R., Ding, Y., Dubois, F., Dzido, G., Eapen, J., Escher, W., Funfschilling, D., Galand, Q., Gao, J., Gharagozloo, P.E., Goodson, K.E., Gutierrez, J.G., Hong, H., Horton, M., Hwang, K.S., Iorio, C.S., Jang, S.P., Jarzebski, A.B., Jiang, Y., Jin, L., Kabelac, S., Kamath, A., Kedzierski, M.A., Kieng, L.G., Kim, C. Kim, J.H., Kim, S., Lee, S.H., Leong, K.C., Manna, I., Michel, B., Ni, R., Patel, H.E., Philip, J., Poulikakos, D., Reynaud, C., Savino, R., Singh, P.K., Song, P., Sundararajan, T., Timofeeva, E., Tritcak, T., Turanov, A.N., Van Vaerenbergh, S., Wen, D., Witharana, S., Yang, C., Yeh, W.H., Zhao, X.Z., and Zhou, S.Q. 2009. A benchmark study on the thermal conductivity of nanofluids. Journal of Applied Physics, 106(9): 094312

Chadwick, M.D., Goodwin, J.W., Vincent, B., Lawson, E.J., and Mills, P.D.A. 2002. Rheological behaviour of titanium dioxide (uncoated anatase) in ethylene glycol. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 196 (2): 235–245.

Chandrasekar, M., Suresh, S., and Bose, A.C. 2010. Experimental investigations and theoretical determination of thermal conductivity and viscosity of Al_2O_3/water nanofluid. *Experimental Thermal and Fluid Science*, 34 (2): 210–216.

Choi, S.U.S. 1995. Enhancing thermal conductivity of fluids with nanoparticles. In *The Proceedings of the 1995 ASME International Mechanical Engineering Congress and Exposition*, San Francisco, CA. ASME, FED 231/MD 66, 1995, 99–105.

Chon, C.H., Kihm, K.D., Lee, S.P., and Choi, S.U. 2005. Empirical correlation finding the role of temperature and particle size for nanofluid (Al_2O_3) thermal conductivity enhancement. *Applied Physics Letters*, 87: 153107.

Chopkar, M., Sudarshan, S., Das, P.K., and Manna, I. 2008. Effect of particle size on thermal conductivity of nanofluid. *Metallurgical and Materials Transactions A*, 39 (7): 1535–1542.

Corcione, M. 2011. Empirical correlating equations for predicting the effective thermal conductivity and dynamic viscosity of nanofluids. *Energy Conversion and Management,* 52 (1): 789–793.

Dames, P., Gleich, B., Flemmer, A., Hajek, K., Seidl, N., Wiekhorst, F., Eberbeck, D., Bittmann, I., Bergemann, C., Weyh, T., Trahms, L., Rosenecker, J., and Rudolph, C. 2007. Targeted delivery of magnetic aerosol droplets to the lung. *Nature Nanotechnology*, 2 (8): 495–499.

Das, S.K., Putra, N., Thiesen, P., and Roetzel, W. 2003. Temperature dependence of thermal conductivity enhancement for nanofluids. *Journal of Heat Transfer*, 125 (4): 567–574.

Ding, Y., Alias, H., Wen, D., and Williams, R.A. 2006. Heat transfer of aqueous suspensions of carbon nanotubes (CNT nanofluids). *International Journal of Heat and Mass Transfer*, 49 (1): 240–250.

Duangthongsuk, W. and Wongwises, S. 2009. Measurement of temperature-dependent thermal conductivity and viscosity of TiO_2-water nanofluids. *Experimental Thermal and Fluid Science*, 33 (4): 706–714.

Eastman, J.A., Choi, S.U.S., Li, S., Yu, W., and Thompson, L.J. 2001. Anomalously increased effective thermal conductivities of ethylene glycol-based nanofluids containing copper nanoparticles. *Applied Physics Letters*, 78 (6): 718–720.

Einstein, A. 1906. Eine neue Bestimmung der Moleküldimensionen. *Annalen der Physik*, 324 (2): 289–306.

Fang, X., Yimin, X., and Qiang, L. 2009. Experimental investigation on enhanced mass transfer in nanofluids. *Applied Physics Letters*, 95 (20): 203108.

Feng, Y. and Kleinstreuer, C. 2010. Nanofluid convective heat transfer in a parallel-disk system. *International Journal of Heat and Mass Transfer*, 53: 4619–4628.

Feng, Y. and Kleinstreuer, C. 2015. A novel computational fluid-particle dynamics model for the simulation of multicomponent droplet-vapor transport/deposition in an idealized human upper airway configuration. In *2015 Annual Meeting of the Biomedical Engineering Society*, Tampa, FL.

Feng, Y., Kleinstreuer, C., and Rostami, A. 2015. Evaporation and condensation of multicomponent electronic cigarette droplets and conventional cigarette smoke particles in an idealized G3–G6 triple bifurcating unit. *Journal of Aerosol Science*, 80: 58–74.

Ghofrani, A., Dibaei, M.H., Sima, A.H., and Shafii, M.B. 2013. Experimental investigation on laminar forced convection heat transfer of ferrofluids under an alternating magnetic field. *Experimental Thermal and Fluid Science*, 49: 193–200.

Haddad, Z., Oztop, H.F., Abu-Nada, E., and Mataoui, A. 2012. A review on natural convective heat transfer of nanofluids. *Renewable and Sustainable Energy Reviews*, 16 (7): 5363–5378.

Hinds, W.C. 1999. *Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles*, 2nd edn. New York: John Wiley & Sons.

Hu, C., Bai, M., Lv, J., and Li, X. 2015. Molecular dynamics simulation of mechanism of nanoparticle in improving load-carrying capacity of lubricant film. *Computational Materials Science*, 109: 97–103.

Jung, J. and Yoo, J.Y. 2009. Thermal conductivity enhancement of nanofluids in conjunction with electrical double layer (EDL). *International Journal of Heat and Mass Transfer*, 52 (1): 525–528.

Kamyar, A., Saidur, R., and Hasanuzzaman, M. 2012. Application of computational fluid dynamics (CFD) for nanofluids. *International Journal of Heat and Mass Transfer*, 55 (15): 4104–4115.

Karimipour, A., Nezhad, A.H., D'Orazio, A., Esfe, M.H., Safaei, M.R., and Shirani, E. 2015. Simulation of copper–water nanofluid in a microchannel in slip flow regime using the lattice Boltzmann method. *European Journal of Mechanics-B/Fluids*, 49: 89–99.

Kim, S.J., Bang, I.C., Buongiorno, J., and Hu, L.W. 2007. Surface wettability change during pool boiling of nanofluids and its effect on critical heat flux. *International Journal of Heat and Mass Transfer*, 50 (19): 4105–4116.

Kleinstreuer, C. 2014. *Microfluidic and Nanofluidics: Theory and Selected Applications*. Hoboken, NJ: John Wiley & Sons.

Kleinstreuer, C. 2016. Computational nanofluid flow and heat transfer in microchannels. In *CRC Handbook of Fluid Dynamics*, 2nd edn, ed. R.W. Johnson. Boca Raton, FL: CRC Press.

Kleinstreuer, C. and Childress, E. 2015. Nanodrug delivery for tumor treatment. In *Encyclopedia of Microfluidics and Nanofluidics*, ed. D. Li, 2348–2358. New York: Springer.

Kleinstreuer, C. and Feng, Y. 2011. Experimental and theoretical studies of nanofluid thermal conductivity enhancement: A review. *Nanoscale Research Letters*, 6 (1): 1–13.

Kleinstreuer, C. and Feng, Y. 2012. Thermal nanofluid property model with application to nanofluid flow in a parallel-disk system—Part I: A new thermal conductivity model for nanofluid flow. *Journal of Heat Transfer*, 134 (5): 051002.

Kleinstreuer, C., Feng, Y., and Childress, E.M. 2014. Drug-targeting methodologies with applications: A review. *World Journal of Clinical Cases*, 2 (12): 742–756.

Kleinstreuer, C., Li, J., and Koo, J. 2008. Microfluidics of nano-drug delivery. *International Journal of Heat and Mass Transfer*, 51 (23): 5590–5597.

Kleinstreuer, C. and Xu, Z. 2015. Thermal nanofluid flow in microchannels with applications. In *Heat Transfer Enhancement with Nanofluids*, eds. V. Bianco, O. Manca, S. Nardini, and K. Vafai. Boca Raton, FL: CRC Press.

Kole, M. and Dey, T.K. 2010. Viscosity of alumina nanoparticles dispersed in car engine coolant. *Experimental Thermal and Fluid Science*, 34 (6): 677–683.

Kondaraju, S., Jin, E.K., and Lee, J.S. 2010. Direct numerical simulation of thermal conductivity of nanofluids: The effect of temperature two-way coupling and coagulation of particles. *International Journal of Heat and Mass Transfer*, 53 (5): 862–869.

Kulkarni, D.P., Das, D.K., and Vajjha, R.S. 2009. Application of nanofluids in heating buildings and reducing pollution. *Applied Energy*, 86 (12): 2566–2573.

Kwark, S.M., Kumar, R., Moreno, G., Yoo, J., and You, S.M. 2010. Pool boiling characteristics of low concentration nanofluids. *International Journal of Heat and Mass Transfer*, 53 (5): 972–981.

Lee, D., Kim, J.W., and Kim, B.G. 2006. A new parameter to control heat transport in nanofluids: Surface charge state of the particle in suspension. *The Journal of Physical Chemistry B*, 110 (9): 4323–4328.

Li, J. and Kleinstreuer, C. 2010. Entropy generation analysis for nanofluid flow in microchannels. *Journal of Heat Transfer*, 132: 122401-1-8.

Li, Q., Lian, W., Sun, H., and Xuan, Y. 2008. Investigation on operational characteristics of a miniature automatic cooling device. *International Journal of Heat and Mass Transfer*, 51 (21): 5033–5039.

Li, C.H. and Peterson, G.P. 2006. Experimental investigation of temperature and volume fraction variations on the effective thermal conductivity of nanoparticle suspensions (nanofluids). *Journal of Applied Physics*, 99 (8): 084314.

Li, C.H. and Peterson, G.P. 2007. The effect of particle size on the effective thermal conductivity of Al_2O_3-water nanofluids. *Journal of Applied Physics*, 101 (4): 44312–44312.

Li, Q. and Xuan, Y. 2009. Experimental investigation on heat transfer characteristics of magnetic fluid flow around a fine wire under the influence of an external magnetic field. *Experimental Thermal and Fluid Science*, 33 (4): 591–596.

Li, Q., Xuan, Y., and Wang, J. 2005. Experimental investigations on transport properties of magnetic fluids. *Experimental Thermal and Fluid Science*, 30 (2): 109–116.

Lin, M., Zhang, D., Huang, J., Zhang, J., Xiao, W., Yu, H., Zhang, L., and Ye, J. 2013. The anti-hepatoma effect of nanosized Mn–Zn ferrite magnetic fluid hyperthermia associated with radiation in vitro and in vivo. *Nanotechnology*, 24 (25): 255101.

Liu, G.L., Kim, J., Lu, Y.U., and Lee, L.P. 2006. Optofluidic control using photothermal nanoparticles. *Nature Materials*, 5 (1): 27–32.

Mahbubul, I.M., Saidur, R., and Amalina, M.A. 2012. Latest developments on the viscosity of nanofluids. *International Journal of Heat and Mass Transfer*, 55 (4): 874–885.

Mahian, O., Kianifar, A., Kleinstreuer, C., Moh'd A, A. N., Pop, I., Sahin, A.Z., and Wongwises, S. 2013. A review of entropy generation in nanofluid flow. *International Journal of Heat and Mass Transfer*, 65: 514–532.

Maxwell, J.C. 1881. *A treatise on electricity and magnetism*. Oxford: Clarendon Press.

Miaskowski, A. and Sawicki, B. 2013. Magnetic fluid hyperthermia modeling based on phantom measurements and realistic breast model. *IEEE Transactions on Biomedical Engineering*, 60 (7): 1806–1813.

Mintsa, H.A., Roy, G., Nguyen, C.T., and Doucet, D. 2009. New temperature dependent thermal conductivity data for water-based nanofluids. *International Journal of Thermal Sciences*, 48 (2): 363–371.

Mohamad, A.A. 2015. Myth about nano-fluid heat transfer enhancement. *International Journal of Heat and Mass Transfer*, 86: 397–403.

Mokmeli, A. and Saffar-Avval, M. 2010. Prediction of nanofluid convective heat transfer using the dispersion model. *International Journal of Thermal Sciences*, 49 (3): 471–478.

Mukhopadhyay, A., Ganguly, R., Sen, S., and Puri, I.K. 2005. A scaling analysis to characterize thermomagnetic convection. *International Journal of Heat and Mass Transfer*, 48 (17): 3485–3492.

Murshed, S.M.S., Leong, K.C., and Yang, C. 2008. Investigations of thermal conductivity and viscosity of nanofluids. *International Journal of Thermal Sciences*, 47 (5): 560–568.

Murshed, S.M.S., Leong, K.C., and Yang, C. 2009. A combined model for the effective thermal conductivity of nanofluids. *Applied Thermal Engineering*, 29 (11): 2477–2483.

Namburu, P.K., Kulkarni, D.P., Misra, D., and Das, D.K. 2007a. Viscosity of copper oxide nanoparticles dispersed in ethylene glycol and water mixture. *Experimental Thermal and Fluid Science*, 32 (2): 397–402.

Namburu, P.K., Kulkarni, D.P., Dandekar, A., and Das, D.K. 2007b. Experimental investigation of viscosity and specific heat of silicon dioxide nanofluids. *Micro and Nano Letters, IET*, 2 (3): 67–71.

Nguyen, C.T., Desgranges, F., Galanis, N., Roy, G., Maré, T., Boucher, S., and Angue Mintsa, H. 2008. Viscosity data for Al_2O_3–water nanofluid—Hysteresis: Is heat transfer enhancement using nanofluids reliable? *International Journal of Thermal Sciences*, 47 (2): 103–111.

Otanicar, T.P., Phelan, P.E., Prasher, R.S., Rosengarten, G., and Taylor, R.A. 2010. Nanofluid-based direct absorption solar collector. *Journal of Renewable and Sustainable Energy*, 2 (3): 033102.

Peer, D., Karp, J.M., Hong, S., Farokhzad, O.C., Margalit, R., and Langer, R. 2007. Nanocarriers as an emerging platform for cancer therapy. *Nature Nanotechnology*, 2 (12): 751–760.

Pourmhran, O., Rahimi-Gorji, M., Gorji-Bandpy, M., and Gorji, T.B. 2015. Simulation of magnetic drug targeting through tracheobronchial airway in presence of an external non-uniform magnetic field using Lagrangian magnetic particle tracking. *Journal of Magnetism and Magnetic Materials*, 393: 380–393.

Prabhat, N., Buongiorno, J., and Hu, L.-W. 2012. Convective heat transfer enhancement in nanofluids: Real anomaly or analysis artifact? *Journal of Nanofluids*, 1: 55–62.

Reddy, M.C.S. and Rao, V.V. 2013. Experimental studies on thermal conductivity of blends of ethylene glycol-water-based TiO_2 nanofluids. *International Communications in Heat and Mass Transfer*, 46: 31–36.

Sahoo, B.C., Das, D.K., Vajjha, R.S., and Satti, J.R. 2013. Measurement of the thermal conductivity of silicon dioxide nanofluid and development of correlations. *Journal of Nanotechnology in Engineering and Medicine*, 3 (4): 041006.

Sarkar, J. 2011. A critical review on convective heat transfer correlations of nanofluids. *Renewable and Sustainable Energy Reviews*, 15 (6): 3271–3277.

Sharifi, I., Shokrollahi, H., and Amiri, S. 2012. Ferrite-based magnetic nanofluids used in hyperthermia applications. *Journal of Magnetism and Magnetic Materials*, 324 (6): 903–915.

Shima, P.D., Philip, J., and Raj, B. 2009. Magnetically controllable nanofluid with tunable thermal conductivity and viscosity. *Applied Physics Letters*, 95 (13): 133112.

Solangi, K.H., Kazi, S.N., Luhur, M.R., Badarudin, A., Amiri, A., Sadri, R., Zubir, M.N.M., Gharehkhani, S., and Teng, K.H. 2015. A comprehensive review of thermo-physical properties and convective heat transfer to nanofluids. *Energy*, 89: 1065–1086.

Suganthi, K.S., Anusha, N., and Rajan, K.S. 2013. Low viscous ZnO–propylene glycol nanofluid: A potential coolant candidate. *Journal of Nanoparticle Research*, 15 (10): 1–16.

Sun, C., Lu, W.Q., Liu, J., and Bai, B. 2011. Molecular dynamics simulation of nanofluid's effective thermal conductivity in high-shear-rate Couette flow. *International Journal of Heat and Mass Transfer*, 54 (11): 2560–2567.

Sundar, L.S., Sharma, K.V., Naik, M.T., and Singh, M.K. 2013. Empirical and theoretical correlations on viscosity of nanofluids: A review. *Renewable and Sustainable Energy Reviews*, 25: 670–686.

Sundar, L.S. and Singh, M.K. 2013. Convective heat transfer and friction factor correlations of nanofluid in a tube and with inserts: A review. *Renewable and Sustainable Energy Reviews*, 20: 23–35.

Taylor, R., Coulombe, S., Otanicar, T., Phelan, P., Gunawan, A., Lv, W., Rosengarten, G., Prasher, R., and Tyagi H. 2013. Small particles, big impacts: A review of the diverse applications of nanofluids. *Journal of Applied Physics*, 113 (1): 011301.

Taylor, R.A., Phelan, P.E., Otanicar, T.P., Adrian, R., and Prasher, R. 2011. Nanofluid optical property characterization: Towards efficient direct absorption solar collectors. *Nanoscale Research Letters*, 6 (1): 1–11.

Tu, H.H. and Ray, A.K. 2005. Measurement of activity coefficients from unsteady state evaporation and growth of microdroplets. *Chemical Engineering Communications*, 192: 474–498.

Turns, S.R. 1996. *An Introduction to Combustion*, Vol. 287. New York: McGraw-Hill.

Veilleux, J. and Coulombe, S. 2011. A dispersion model of enhanced mass diffusion in nanofluids. *Chemical Engineering Science*, 66 (11): 2377–2384.

Veiseh, O., Gunn, J.W., and Zhang, M. 2010. Design and fabrication of magnetic nanoparticles for targeted drug delivery and imaging. *Advanced Drug Delivery Reviews*, 62 (3): 284–304.

Wang, X., Xu, X., and Choi, S.S.U. 1999. Thermal conductivity of nanoparticle-fluid mixture. *Journal of Thermophysics and Heat Transfer*, 13 (4): 474–480.

Xu, Z. 2014. *An Improved Thermal Conductivity Model for Nanofluids with Applications to Concentration Photovoltaic-Thermal Systems*. MS Thesis, North Carolina State University.

Xu, Z. and Kleinstreuer, C. 2014a. Concentration photovoltaic–thermal energy co-generation system using nanofluids for cooling and heating. *Energy Conversion and Management*, 87: 504–512.

Xu, Z. and Kleinstreuer, C. 2014b. Computational analysis of nanofluid cooling of high concentration photovoltaic cells. *Journal of Thermal Science and Engineering Applications*, 6 (3): 031009.

Yang, J.C., Li, F.C., Cai, W.H., Zhang, H.N., and Yu, B. 2014. On the mechanism of convective heat transfer enhancement in a turbulent flow of nanofluid investigated by DNS and analyses of POD and FSP. *International Journal of Heat and Mass Transfer*, 78: 277–288.

Yu, W. and Choi, S.U.S. 2003. The role of interfacial layers in the enhanced thermal conductivity of nanofluids: A renovated Maxwell model. *Journal of Nanoparticle Research*, 5 (1–2): 167–171.

Yu, W., France, D.M., Timofeeva, E.V., Singh, D., and Routbort, J.L. 2012. Comparative review of turbulent heat transfer of nanofluids. *International Journal of Heat and Mass Transfer*, 55 (21–22): 5380–5396.

Yu, W. and Xie, H. 2012. A review on nanofluids: Preparation, stability mechanisms, and applications. *Journal of Nanomaterials*, 2012: 1.

4.4 FOULING IN CRUDE OIL AND FOOD RELATED HEAT-TRANSFER EQUIPMENT

D. Ian Wilson and Graham T. Polley

Introduction: The Major Unresolved Problem in Heat Transfer

Fouling is the unwanted accumulation of material on the surface of process equipment to form deposit layers, which impose additional resistances to heat transfer. These fouling layers can also affect the flow of the fluid, increasing surface roughness and/or pressure drop. As a result, the thermal and hydraulic performance of the unit will deteriorate over time.

Fouling occurs because species in the fluid being processed form the layers in response to being heated or cooled. This section addresses fouling and its mitigation in the food processing sector and crude oil refining operations, where fouling is commonly encountered. The chapter *does not* deal with water scaling, particulate fouling from gases, or biofouling: the reader is advised to consult Bott (1995), ESDU (2003, 2008), Bott (2008), and the proceedings of the conferences on fouling and cleaning in heat exchangers at http://www.heatexchanger-fouling.com/ for detailed information on these topics to complement the general points made here.

Fouling is a dynamic phenomenon and requires consideration at each stage of a heat exchanger's life in design, commissioning, and during operation. Fouling requires active management. Many problems with fouling arise because it is treated as a steady-state phenomenon at the design stage, or the heat exchanger is not operated in a manner which will mitigate fouling.

Quantitative Impact of Fouling

The thermal resistance of a deposit layer is usually lower than that of the volume of moving fluid it displaces, so the overall resistance to heat transfer increases. For heat transfer across a tube of inner and outer radii r_i and r_o, respectively, the clean overall heat transfer coefficient is given by

$$1/U_{j,\text{clean}}r_j = 1/r_i h_i + \ln\left(r_o/r_i\right)/k_{\text{wall}} + 1/r_o h_o . \tag{4.4.1}$$

Here j can be either i or o. The presence of fouling layers adds further terms. Where fouling layers are thin, the thermal resistance can be approximated as that of a thin slab with thermal conductivity k_f and thickness l_f. The overall heat transfer coefficient is then

$$\begin{aligned} 1/U_j r_j &= 1/U_{j,\text{clean}}r_j + l_{f,o}/r_o k_{f,o} + l_{f,i}/r_i k_{f,i} \\ &= 1/U_{j,\text{clean}}r_j + R_{f,o}/r_o + R_{f,i}/r_i , \end{aligned} \tag{4.4.2}$$

where R_f is the fouling resistance and subscripts o and i denote the surface on which the fouling layer is formed. If a fouling layer is not thin, the slab approximation will be inaccurate and the film heat transfer coefficient is also likely to change. A rough deposit layer will also change the film heat transfer coefficient, causing it to increase: this can result in an apparently negative fouling resistance if the impact of roughness on $U_{j,\text{clean}}$ is not considered. The pressure drop will, however, increase: this is one of the reasons why pressure drop as well as thermal performance should be monitored on heat exchangers subject to fouling.

In practice, it is not possible to distinguish where in a heat exchanger a deposit is formed on the basis of steady state thermal measurements alone, and a single term is used to quantify the additional resistance due to fouling:

$$1/U_j = 1/U_{j,\text{clean}} + R_{f,j} . \tag{4.4.3}$$

In several texts, $R_{f,j}$ is called the *fouling factor*. It follows that

$$U_j = \frac{U_{j,clean}}{\left(1 + R_{f,j} U_{j,clean}\right)}. \tag{4.4.4}$$

The group $R_{f,j}\ U_{j,clean}$ is dimensionless and is sometimes called the fouling Biot number. Another measure used for the impact of fouling is the cleanliness factor, CF, defined as

$$\begin{aligned} CF &= 100 \times U_j / U_{j,clean} \\ &= 100 / (1 + R_{f,j} U_{j,clean}). \end{aligned} \tag{4.4.5}$$

Both U_j and CF decrease as the fouling resistance increases.

In the rest of this chapter, the j subscript on U is no longer used.

Equation 4.4.4 shows that the impact of fouling depends on the magnitude of the clean heat transfer coefficient. Figure 4.4.1 shows the effect of a linear increase in fouling resistance with time on the overall heat transfer coefficient for two values of U_{clean}. When U_{clean} is large, a small fouling resistance has a significant effect, whereas if U_{clean} is small, the effect of fouling will not be noticed, or measurable in practical terms, until R_f is large. In the latter case, the effect of fouling on pressure drop and/or flow rate is likely to be noticed before that on heat transfer (see Figure 4.4.4).

It is also important to note that the definition of R_f in Equation 4.4.3 includes the inaccuracy in the calculation of the clean heat transfer coefficient and any uncertainty in its estimation. In practice, the thermophysical properties of the fluids are seldom known accurately, or may be specified wrongly. Differences in heat transfer performance can then be incorrectly attributed to fouling.

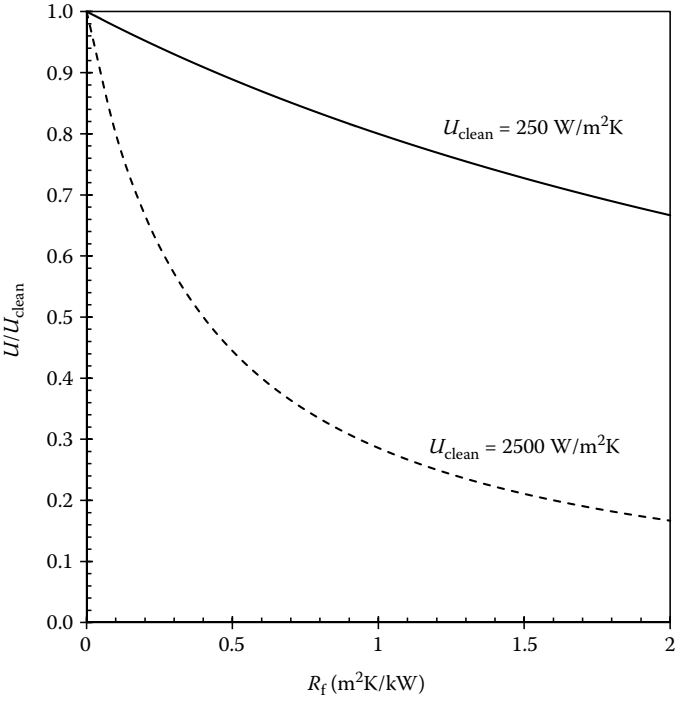


FIGURE 4.4.1 Effect of fouling resistance on overall heat transfer coefficient for two values of U_{clean}.

Robust Design

The impact of fouling on a heat exchanger depends on the unit's design: some designs are more *robust* toward the impact of fouling. It must be stressed that this does not mean that they mitigate fouling by reducing the occurrence of fouling or the rate of fouling layer deposition: it refers to the effect of fouling on the unit's performance. Fouling changes the local heat fluxes and temperature driving forces within a unit and the total heat duty, q, is determined by the contacting arrangement.

This is illustrated in Figure 4.4.2 for the case of a single-pass countercurrent heat exchanger. Recall that the heat exchanger effectiveness (see NTU-effectiveness method, Chapter 4.1) is determined by the contacting arrangement and the number of transfer units, NTU = UA/C, where A is the heat transfer area, and C is the smaller heat capacity rate. If the approach temperatures are constant, q is proportional to the effectiveness ε, given by

$$\varepsilon = \frac{\left(1 - \exp\left(-\mathrm{NTU}\left(1 - C_r\right)\right)\right)}{\left(1 - C_r \exp\left(-\mathrm{NTU}\left(1 - C_r\right)\right)\right)}, \tag{4.4.6}$$

where C_r is the ratio of the heat capacity rates. Figure 4.4.2 shows that ε is more sensitive to NTU when NTU is less than about 3. Fouling reduces U, the reduction in ε (and q) depending on C_r and the initial value of NTU, that is, the surface area of the exchanger and U_{clean}. A robust design would either start with a large value of NTU or be configured so that there is little fouling.

All four variables in Equation 4.4.6, namely U, A, C, and C_r, are determined by the unit design. The *historical* method for designing for fouling service has been to set C and C_r at the values determined by the process design (with neither recycle nor bypass modifying a flow rate) and to calculate U_{clean} based on clean performance (via Equation 4.4.1). The U value used for the design is then calculated via Equation 4.4.2, using estimates of fouling resistance expected to be experienced in practice for $R_{f,i}$ and $R_{f,o}$. These are often referred to as fouling factors. The heat exchanger area is

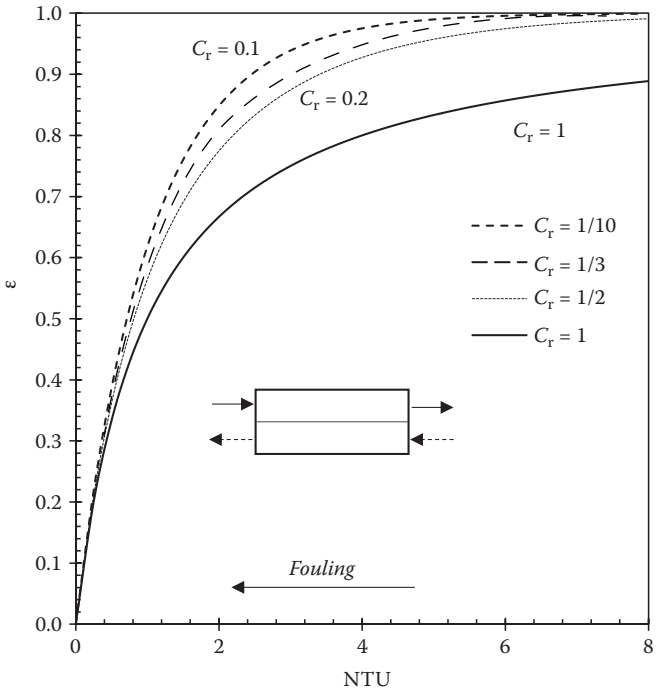


FIGURE 4.4.2 Effect of fouling, via NTU, on the effectiveness of a simple countercurrent heat exchanger for different values of capacity rate ratio. Fouling causes NTU to decrease, reducing ε.

then larger than that required for the clean case: excess area has to be added. NTU is then greater than required, and Figure 4.4.2 shows that the design is then apparently more robust. Initially, however, it will overperform and some control action (typically involving bypassing) will be needed to adjust the duty to its target value.

Tables of "fouling factors" are available in many sources, including the web, design handbooks (e.g., Saunders, 1988; Kakaç and Liu, 2002), as well as those based on local experience. If the historical method is to be used, the latter source is to be preferred.

Equation 4.4.4 shows that injudicious selection of such fouling factors can lead to large amounts of excess area, with associated penalties in terms of cost, pressure drop, and even promotion of fouling resulting in a "self-fulfilling prophecy" (see Bott, 1995; Ishiyama et al., 2008). An alternative rule of thumb used elsewhere is the "no-foul" approach whereby fouling factors are not used, and the excess area added for fouling and uncertainty in the calculation of heat transfer coefficients is limited to a fixed percentage, for example, 20% (Gilmour, 1965; Bennett and Nesta, 2004).

Other contacting arrangements give different relationships between ε and NTU, and therefore different sensitivity to fouling. The contacting arrangement also determines the distribution of surface temperatures within the heat exchanger: for fouling processes that are sensitive to surface temperature, this will determine the rate of deposition and is discussed further in the "Use of Fouling Models in Exchanger Design" section.

Fouling Dynamics

The rate of fouling is determined by the supply of precursors in the fluid, flow rates, bulk and surface temperatures, and the physical and chemical properties of the heat transfer surface. The factors particular to food processing and crude oil are discussed further in "Fouling in Food Processing" and "Crude Oil Fouling" sections, respectively.

The inset in Figure 4.4.3 shows two characteristic fouling behaviors that are often observed in practice: (1) linear fouling, where the fouling resistance increases at a constant or near-constant rate over time, and (2) asymptotic fouling, where the rate of increase decreases over time and may approach an asymptotic value, labelled $R_{f,\infty}$. The fouling resistances are plotted as the product $R_f U_{clean}$, and the effect on heat transfer performance of a simple countercurrent unit, expressed as the effectiveness, is shown for two values of NTU. It can be seen that a linear increase in R_f over time gives a nonlinear change in effectiveness (and thus duty). After extended time, the performance of the linear fouling case is poor, and consideration should be given to taking the unit offline for cleaning.

The asymptotic trend has the same initial behavior as the linear case but approaches a steady value of fouling resistance and effectiveness over time. This is the basis for the design approach based on fouling factors, which assumes that the fouling resistance will reach an asymptotic limit and the unit is designed on this basis. The figure highlights the pitfalls associated with this approach: asymptotic fouling behavior is often observed with cooling water systems but is less common in refinery crude oil and continuous food processing applications.

Both trends show an initial period in which there is no measured change in fouling resistance, called the *induction period*. This is often associated with the modification of the original, clean surface to a form that supports the attachment of fouling layers. It can involve surface oxidation, corrosion, or adsorption of precursors. Once the original surface is coated with a fouling layer, the nature of the surface plays little role in deposition (apart from heat transfer considerations), but it can influence the rate of cleaning. Surface modification to prolong induction periods and thereby mitigate fouling is the subject of ongoing research (e.g., Gomes da Cruz et al., 2015).

Timescales

The length of time over which fouling occurs varies between applications. *Acute* fouling, where the overall heat transfer coefficient changes strongly over a relatively short period, is associated with fouling precursors being present in the bulk fluid in high concentration. *Chronic* fouling is

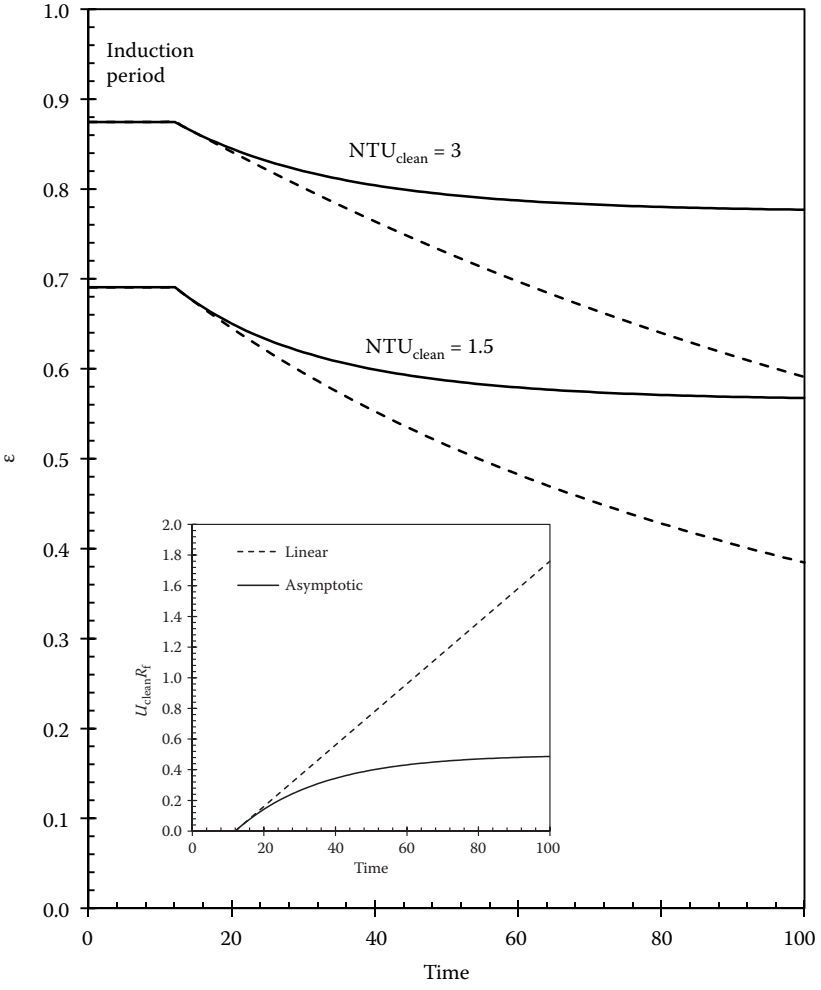


FIGURE 4.4.3 Effect of linear and asymptotic fouling on effectiveness of a simple countercurrent heat exchanger (see Figure 4.4.2). Inset shows evolution of $U_{clean}R_f$, which is dimensionless. Time is in arbitrary units.

associated with fouling precursors being present in low concentration, or being formed in the heated or cooled region near the heat transfer surface. Operations that normally operate subject to chronic fouling can experience acute fouling when process conditions or bulk chemistry are perturbed, such as when large amounts of material from a slop tank are added to the crude oil fed to a refinery preheat train or when incompatibility problems arise in blends of crude oils.

Pressure Drop

Fouling affects pressure drop by changing the roughness of the surface and by narrowing the cross-sectional area of the duct for flow. The effect of fouling on pressure drop can be estimated from the following simple model if the thermal conductivity of the fouling layer is known (or can be calculated). Assume that fraction x of the overall fouling resistance is caused by a layer of uniform thickness l_f and thermal conductivity k_f, on the side (hot or cold) under consideration. The thickness of the layer, assuming the thin slab approximation, is given by

$$l_f = xk_f R_f . \tag{4.4.7}$$

The hydraulic diameter of the duct, and thus the pressure drop, can then be calculated.

This is illustrated here for a shell-and-tube unit with no fouling on the shell side. Only the tube side pressure drop is then affected. For the tube side, $x=1$ and the internal diameter of the fouled tube, d_f, is given by

$$d_f = d_i - 2l_f = d_i - 2\ k_f R_f . \tag{4.4.8}$$

The pressure drop along the tube is proportional to the friction factor, C_f, and mean velocity, with the latter proportional to V/d_f^2. If the volumetric flow rate V is held constant, the ratio of the pressure drop under fouled and clean conditions is given by

$$\frac{\Delta P_{\text{fouled}}}{\Delta P_{\text{clean}}} = \frac{C_{\text{f,fouled}}}{C_{\text{f,clean}}} \times \frac{1}{\left(1 - 2k_f R_f / d_i\right)^4} . \tag{4.4.9}$$

Figure 4.4.4 plots the effect of fouling on pressure drop, given by Equation 4.4.9, against the effect of fouling on the overall heat transfer coefficient, Equation 4.4.4, for different values of the ratio of friction factors, U_{clean} and k_f/d_i. The U_{clean} values and the range in R_f are those used in Figure 4.4.1.

Over time the pressure drop increases and the overall heat transfer coefficient decreases. It can be seen that (1) the pressure drop is very sensitive to changes in roughness, and (2) the relationship between pressure drop and heat transfer coefficient depends on the value of U_{clean}. This is because a thin fouling layer has a significant effect if U_{clean} is large, but the effect on pressure drop is small. If U_{clean} is small, the fouling resistance (and thus the change in duct dimension) must be large in order to affect the heat transfer coefficient.

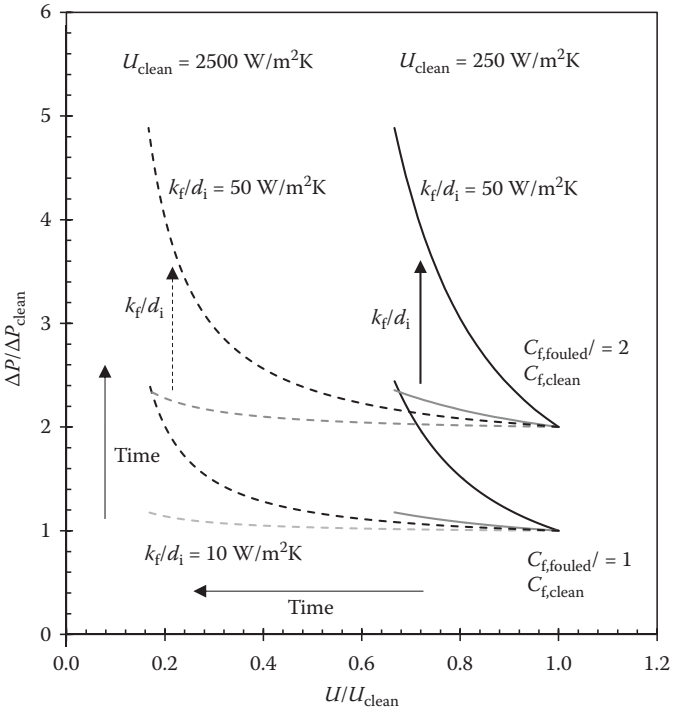


FIGURE 4.4.4 Effect of fouling on pressure drop, expressed as $\Delta P/\Delta P_{\text{clean}}$, and thermal performance, expressed as U/U_{clean}, for different values of U_{clean} and friction coefficient C_f. Line patterns indicate different values of k_f/d_i.

Fouling and Cleaning

A heat exchanger subject to fouling is likely to require cleaning in order to restore its performance or to avoid phenomena promoted by the presence of fouling layers such as microbial growth (in the food sector) and corrosion of the heat transfer surface. It is essential that cleaning is considered at the design stage, as certain configurations of heat exchanger are not compatible with some cleaning methods. A description of cleaning technologies can be found in Müller-Steinhagen and Zettler (2011). The aim of a cleaning operation is to remove any fouling deposit present and to restore the surface to its original condition. Ineffective cleaning will leave residues of foulant, which will promote fouling growth, so that the induction period will be very short or nonexistent (see Figure 4.4.3).

Cleaning also needs to be considered when selecting the materials of construction, seals, and pump sizing so that these can handle the chemical agents, conditions, and flow rates required for effective cleaning. In the food sector, the design must also consider the agents used for disinfection such as live steam.

Where cleaning is performed regularly, there is often an optimal run time between cleaning actions. Mathematical algorithms for scheduling cleaning operations have been developed for individual exchangers (see Cosado, 1990; Sheikh et al., 1996; Ishiyama et al., 2011). These rely on the availability of models describing the change in heat duty over time. The calculations for isolated exchangers can be performed on spreadsheets. Scheduling cleaning for units in networks requires detailed simulation of the network (see Ishiyama et al., 2009) and commercial software tools for this purpose have been developed (e.g., SmartPM from IHS; Studio from Hexxcell).

If the process is continuous and a unit cannot be taken offline for cleaning, it is common practice to split the unit into several smaller units, one of which is cleaned and is then available on standby.

Fouling Mechanisms

Fouling layers are often classified in terms of the mechanism by which they are formed, as this determines the nature of the deposit and the steps controlling the rate of formation. The latter information is important for heat exchanger design and control as it can guide how design and operating parameters can be selected to reduce fouling rates or possibly avoid fouling altogether.

The classification reported by Epstein (1983) is as follows:

Crystallization fouling is the formation and growth of crystalline deposits from species in solution. Crystallization occurs when the concentration of solute exceeds the equilibrium value, which is usually strongly related to temperature. A hot or cold surface can result in the local equilibrium concentration being lower than the bulk value, promoting crystallization at the surface. The equilibrium concentration increases with temperature for normal solubility species, such as fats, which leads to fats forming fouling layers on chilled surfaces. It decreases with temperature for inverse solubility salts such as calcium and magnesium carbonates and sulfates, and also calcium phosphates: the latter is responsible for the fouling of the higher temperature milk heaters in dairy ultra-high temperature (UHT) processing.

Particulate fouling occurs when solid particulates present in the bulk fluid are transported to the heat transfer surface where they attach. Particulates can be conveyed into the exchanger in the process steam or generated in the bulk by a chemical reaction such as crystallization.

Chemical reaction fouling occurs when species present in solution undergo a chemical transformation at the surface to form a fouling layer. The surface may catalyze the transformation. Surface temperature is an important factor, and this parameter will determine which reactions take place and their rate. Chemical reaction, like crystallization, can also occur in the bulk fluid and can lead to the formation of insoluble species which deposit by a particulate fouling mechanism.

Corrosion fouling is where the heat transfer surface undergoes a transformation under the operating conditions to form corrosion products with increased heat transfer resistance. Corrosion products can also spall off, generating particulates which deposit elsewhere in the exchanger or in a unit downstream.

Biofouling is the formation of fouling layers by organisms. Heat transfer surfaces provide environments which favor the attachment of macroorganisms such as shellfish as well as colonization by biofilm-forming microorganisms. Biofouling is rarely seen in crude oil processing but is encountered in food processing operations due to the availability of nutrients in aqueous solution.

Many fouling layers in practice are generated by a mixture of fouling mechanisms. Biofouling can promote corrosion fouling when anaerobic bacteria generate acidic conditions at the exchanger surface; a particulate fouling layer can become harder to remove owing to crystallization, chemical reaction, or biofouling "cementing" particulates in the deposit.

Table 4.4.1 shows where the different fouling mechanisms arise in the food and crude oil processing sectors.

The rate of fouling is related to the concentration of precursors entering the system. This can be subject to natural variations, for example, seasonal changes in foodstuffs and lactation cycles for milk. In processing crude oil, refineries often blend crudes together so the nature and concentration of fouling precursors can vary regularly: in these systems, it is important to avoid combining so-called incompatible crudes to give a mixture in which some components are insoluble. Those components then precipitate in the bulk fluid and the heat exchangers are subject to acute particulate fouling (Wiehe, 2008).

Many deposits undergo *ageing*, where the initially deposited species are transformed over time owing to exposure to the temperature and flow conditions active at the heat transfer surface. This is the equivalent to baking the layer to give a harder deposit. Wax and fat layers can similarly harden over time on cold surfaces. Ageing can create problems for cleaning, in which case cleaning should be performed earlier and thus more often. Ageing also disguises the identity of the fouling precursors responsible for deposit growth. Where fouling does occur, it is recommended that samples of the fouling layer are obtained intact for analysis in order to identify the precursor(s) and mechanism(s) involved.

TABLE 4.4.1
Fouling Mechanisms Encountered in the Food and Crude Oil Processing Sectors

Fouling Mechanism	Food Processing	Crude Oil
Crystallization	Fats—gelation on cold surfaces	Waxes in chiller units
	Calcium phosphate in UHT milk processing (milk stone)	Upstream of a desalter and downstream of a poorly operated desalter
	Heating and cooling water	Heating and cooling water
Particulate	Suspensions	Iron sulfide and rust particles introduced during crude storage and transport
	Protein aggregates formed in milk heated above the protein denaturation temperature	Sediments
		Sand particles
		Asphaltene precipitates arising from blending of incompatible crudes
Chemical reaction	Protein solutions heated above the protein denaturation temperature (e.g., milk and egg)	Degradation of asphaltenes
		Polymerization of unsaturated compounds (both in the presence and absence of oxygen)
	Unsaturated fats exposed to air (autoxidative polymerization, e.g., in fryers)	Condensation reaction products
		Pyrolysis and coking reactions (high temperature)
	Viscous sauces	
Corrosion	Avoid by material selection: can be caused by cleaning solutions	Acidic crudes
Biofouling	Heating and cooling water	Heating and cooling water
	Many dilute aqueous solutions	

Designing for Fouling Service

Fouling can be avoided or controlled to a low level in some circumstances by the addition of antifoulant chemicals. This is often only economically feasible for closed-loop systems and is unlikely to be acceptable in food processing operations. Surface treatments that deter the attachment of fouling deposits are available for some niche applications but have yet to reach widespread use in the food and crude oil sectors.

A second option is to employ an antifouling exchanger or fouling mitigation measures. Table 4.4.2 lists some of the devices available. These incur additional capital cost and can also require additional pressure drop in order to handle a given throughput. Design of such units is specialized and require liaison with the device manufacturer.

The final option is to utilize a conventional heat exchanger and select a combination of surface temperatures, flow velocity, and contacting arrangement which will reduce the rate of fouling. Fouling is still likely to occur so the design specification must include targets for the heat duty and maximum allowable pressure drop across the unit in the fouled state, noting that fouling will increase the pressure drop for a given flow rate.

The time for which the unit is to be available, before cleaning or switching to a standby unit, must also be specified. This sets an acceptable fouling rate. If fouling models are available for estimating the effect of flow velocity, bulk and surface temperature on the fouling rate, the heat exchanger design can be optimized to mitigate fouling. This is considered further for crude oil applications in Section 4.5.

Some general guidelines for designing shell-and-tube heat exchangers for fouling service presented by Gilmour in 1965 are summarized in Table 4.4.3 and still apply to these and other configurations. Many of his recommendations focus on ensuring that flow patterns are close to theoretical. One point which Gilmour made 50 years ago and still applies is that poor exchanger design is often attributed, wrongly, to fouling: according to Equation 4.4.3, R_f includes all errors in estimating U_{clean}.

The choice of exchanger type should include cleaning considerations, including (1) the requirements of the likely cleaning method; (2) the ability to access both sides of the heat transfer surface if fouling occurs on both; and (3) whether blockage caused by fouling layers will render the unit difficult to clean. Plate and frame exchangers, which can be disassembled to expose individual plates for cleaning and inspection, are good in this regard but care must be taken to maintain the integrity of the seals. In contrast, tube inserts may need to be removed and replaced in order to achieve effective cleaning.

Particular care should be taken when selecting shell-and-tube configurations. It is conventional with shell-and-tube designs for the process stream more prone to fouling to be on the tube side, particularly if cleaning required mechanical action such as pigging. Some tube pitch patterns render the shell-side surface difficult to access by water jets, while U-bend tubes are not suitable for a range of cleaning technologies. More details about configurations and cleaning are given in Müller-Steinhagen and Zettler (2011).

Control

The information used to design the exchanger should also determine the control strategy.

When an exchanger is designed with excess area in order to mitigate the effect of fouling, it will overperform when clean. The standard control technique for a unit exchanging heat between two process streams is to bypass a fraction of one of the streams in order to maintain the exit temperatures at the target values. If one of the streams is not prone to fouling, this is the stream that should be manipulated unless other criteria such as boiling will arise (see Figure 4.4.5a). If the fouling stream is manipulated, the initial flow rate (and velocity through the unit) will be lower, reducing the shear stress acting on the surface and the film heat transfer coefficient. The surface temperature will then approach that of the other stream. Both these effects tend to promote fouling, and will result in high initial fouling rates, negating the benefit of the excess area.

TABLE 4.4.2
Fouling Mitigation: Antifouling Heat Exchangers and Devices

Heat Exchanger Type	Origin of Mitigation Action	Comments and Source
Scraped surface	Annular heat exchanger with rotating blades which scrape material off the outer, heat transfer surface	Standard technology in the manufacture of lubricants and ice cream, and for processing of viscous foods Power required for motor
	Solano et al. (2011) report a novel double-pipe device with reciprocating scraper elements	Blades require regular maintenance
Fluidized bed	Scouring particles are fluidized by the passage of the process stream prone to fouling: the particles remove fouling deposits from the heat transfer surface	Specialized units with proven record for streams prone to severe fouling Klaren et al. (2007) Klaren and de Boer (2011) Müller-Steinhagen and Lancefield (2005)
	Shell and Tube	
Taprogge	Abrasive balls are circulated through tubes by fluid flow, captured and reinjected	Widely used in cooling water and power generation applications Not all tubes may be cleaned
	Used intermittently to remove fouling layers	Abradents need replenishment periodically Müller-Steinhagen and Zettler (2011)
Spirelf	Coiled insert for tubes. Fluid flow causes the inset to vibrate, scouring deposit layer	Increases heat transfer and pressure drop Krueger and Pouponnot (2009)
Turbotal	Helical ribbon insert for tubes. Fluid flow causes the insert to rotate, scouring deposit layer	Increases heat transfer and pressure drop Krueger and Pouponnot (2009)
Fixotal	Fixed helical insert for tubes. Promotes turbulence	Does not scour/remove deposit Krueger and Pouponnot (2009)
HITRAN	Coiled wire inserts for tubes, which increase fluid mixing and wall shear stress	Film heat transfer coefficient is increased, accompanied by higher pressure drop Difficult to clean Ritchie et al. (2007) Drögemüller et al. (2013)
Twisted tubes	Tubes are rolled such that they appear to be twisted. They are primarily oval in shape and can be aligned so that they touch intermittently along the length of the bundle. The result is a very rigid structure. Significantly more twisted tubes than plain tubes can be accommodated in a shell of given diameter.	These units cannot be cleaned by standard mechanical means: specialist cleaning is required Most frequently used with crude oil Al-Hadhrami et al. (2010)
Helical baffles	These consist of a series of interlocking plates positioned to provide a screw-type motion along the shell. The spacing of the plates is a function of shell diameter (unlike segmental baffles where the spacing can be changed independently). The larger the shell diameter, the lower the shell side velocity, and the lower the pressure drop. This is accompanied by a lower heat transfer coefficient. Later designs utilized a "double helix" that gave some improvement	In the case of crude oil duties, low heat transfer coefficients usually lead to increased fouling rates. Consequently, the processing of crude on the shell side of such exchangers is not recommended. There is a "niche" application for these units, in the cooling of vacuum and atmospheric residues The viscosity of such residues is generally high (>2 cps). Therefore they benefit from being passed through the shell side of the unit. The crude can be passed through the tubes and inserts used to promote good heat transfer van der Zijden et al. (2013), Brignone et al. (2015)

TABLE 4.4.3
Summary of Gilmour (1965) Recommendations for Designing Shell-and-Tube Heat Exchangers for Fouling Service

Feature	Recommendation	Reason
Flow maldistribution	Site inlets, outlets, and impingement plates in order to ensure uniform flow in tubes (and other channels)	Channels subject to low flow velocity will give poor heat transfer and are also prone to fouling and blockage.
Avoid shell-side bypassing	Use a baffle cut of 20% and never more than 25% Do not use vertical baffle cuts for sensible heat transfer Avoid shell-side bypass streams by insisting on TEMA shell-baffle clearances Use tie rods and spacers to block bypass routes around a shell Use dummy tubes to block central bypass paths Avoid notches for drainage and venting	Bypassing reduces the flow pattern from plug flow to mixed flow, giving poorer heat transfer The associated low velocities promote fouling and deposition in stagnant areas
Materials of construction	Select materials of construction carefully, giving consideration to fluid conditions as well as bimetallic corrosion	Saving money on a one-off capital purchase often proves more expensive over time due to higher operating costs and/or replacement
Viscous streams	Put through the channel that is easier to clean	Some guidelines insist on putting viscous fluids through tubes rather the shell: if the residue is viscous rather than hard, shell-side cleaning can be faster

Alternative control strategies include recycling a fraction of the outlet stream and mixing it with the feed in order to raise the inlet temperature, reducing the maximum heat duty (Figure 4.4.5b). This requires a pumparound and has the advantage of increasing the velocity in the unit, albeit at the cost of a pump. This strategy can also be used to enhance the flow rate when high velocities are required. It is not suitable for many food applications owing to the effect of heating or cooling operations on the product microstructure, or the potential for long residence times to promote microbial growth via the recycle of material.

When one of the process streams is a heating or cooling utility, the inlet temperature of this stream should be manipulated where this is feasible (e.g., by adjusting steam pressure: see Figure 4.4.5c).

Fouling in Food Processing

Fouling is encountered in a wide range of food applications. Many food products are temperature sensitive yet need to be heated for pasteurization or sterilization. Dairy processing units, for example, experience acute fouling primarily due to denaturation of proteins at temperatures below 110°C. At higher temperatures, crystallization of calcium phosphate also contributes, giving a hard protein/mineral deposit sometimes called “milk stone.”

Many foodstuffs are mixtures of species that individually could cause fouling, so deposition can involve several mechanisms. Milk stone is an example, where denatured proteins (chemical reaction fouling) glue together calcium phosphate solids (crystallization fouling). It is advised to determine as part of the design specification whether any temperature conditions should be avoided: this will set a constraint or a guideline for the surface temperatures to be reached in the unit.

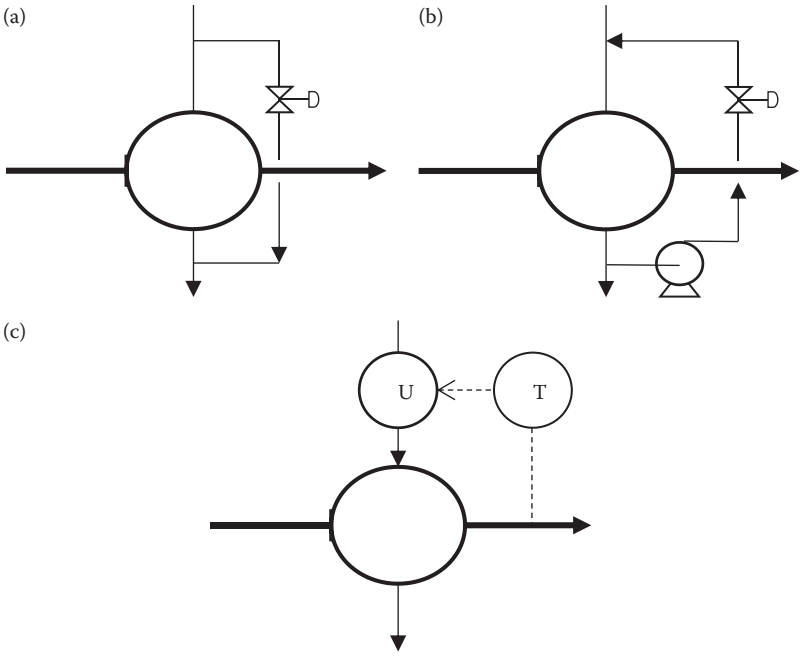


FIGURE 4.4.5 Schematic of control schemes to mitigate fouling. Solid line is stream subject to fouling: (a) bypassing of nonfouling stream, (b) recycle of nonfouling stream (not suitable if extended residence times for this stream in the heat exchanger cause problems), (c) modification of inlet temperature of nonfouling stream using utility (heating or cooling).

Heat exchangers used for food processing must therefore be designed to be cleaned and to be disinfected. The materials of construction must be compatible with the temperatures and pH of the product stream(s), the cleaning solutions (alkali and acids are both used), and the sterilizing/disinfecting agents. Stainless steel is frequently used except where the product contains high levels of salt. Cleaning is required for

1. *Microbiological integrity:* Hygienic operation is critical and units will often be taken offline for cleaning before fouling has reduced process performance in order to minimize microorganism growth. Many food fouling layers promote microorganism attachment and can shield thermophilic bacteria or their spores from cleaning agents.
2. *Product changeover:* Food processing plants often manufacture a range of products, so the units have to be flushed out on a regular basis. This means that the designer must consider a range of fluid properties. It also poses challenges for surface treatments to reduce fouling, as changes in product composition will determine the fouling precursors and mechanisms.
3. *Restoring process performance:* It should be noted that many food products are manufactured in batch mode and the batch run time determines the period over which acceptable process performance must be maintained. Many food fouling layers age to give more tenacious deposits, which are harder to clean. For units running for extended periods, this should be considered in the specification of the cleaning system and protocols.

While the operating period may be set by hygiene or batch manufacturing criteria, it is nevertheless important to minimize fouling in food heat exchangers because cleaning can take a significant length of time, affecting plant productivity. Hygienic operation usually requires cleaning-in-place (CIP), whereby solutions are circulated through a unit for a set length of time or until a monitoring program determines that the end point has been reached. A CIP cycle often includes purging,

TABLE 4.4.4
Heat Exchanger Types Regularly Used in Food Processing Operations

Exchanger Type	Fouling/Cleaning	Examples
Plate and frame	Complex flow patterns: deposition often occurs in regions of slower flow Can be opened for inspection or rigorous cleaning	Widely used in milk processing
Scraped surface	The rotating blade provides a self-cleaning action Start-up and shutdown must be managed to avoid a thick layer building up	Ice cream (water freezes to form ice crystals, which are scraped into the bulk fluid) High viscosity sauces, pastes
Concentric tube (double pipe, triple tube)	Well-defined flow field to minimize dead zones, promoting uniform fouling and uniform access for CIP fluids	Viscous sauces, creams Products containing larger particulates or pieces (this dictates the tube size)
Multiple tube	A small number of tubes with heating utility in a common shell. Careful design of flow manifold required	As above
Evaporators (falling film, rising film)	Concentrations of fouling precursors increase as evaporation occurs, giving regions where fouling starts Uniform distribution of liquid product and CIP solutions is critical to avoid dry patches/hot spots on surface	Concentration units upstream of spray driers
Fryers	Gum formation by oxidation of unsaturated fats/oils	Snack food manufacture, e.g., potato chips, french fries

cleaning, rinsing, disinfection, and clean water stages, which can be reduced in length and resource consumption if there is little fouling.

These considerations result in the range of exchanger types used in heating or cooling food products being quite small, summarized in Table 4.4.4.

In all food processing units, it is important not to create dead zones in the fluid pathway, both to avoid broad residence time distributions (affecting product quality) and inhibiting biofilm formation. These are covered by hygienic design standards. The exchanger specification should include instrumentation for monitoring the performance of the unit during normal operation and a CIP cycle.

Fouling can also occur in sections where there is no driving force for heat transfer such as holding sections, distributors, and headers. Deposition in these cases can be caused by the laydown of species generated by heat treatment upstream, such as protein aggregates, or by particulates created by cooling oils and fats.

Studies of food fouling have tended to focus on the influence of composition and temperature regime on fouling mechanisms, the deposits generated, and on cleaning (e.g., Changani et al., 1997; Visser, 1997; Bansal and Chen, 2006). Most studies of fouling rates, useful for guiding heat exchanger design, are for dairy applications. Most studies have been performed using simple tubular devices, shell and tube, or plate heat exchangers with bulk temperatures below 100°C, where fouling is dominated by the deposition of proteinaceous deposits. One of the models that describes the observed fouling rates is that of Paterson and Fryer (1988), in which the *initial* rate follows

$$\mathrm{d}R_\mathrm{f} / \mathrm{d}t \propto \exp\left(-E_\mathrm{act} / RT_\mathrm{w}\right) / u, \tag{4.4.10}$$

where t is the time, E_act is an apparent activation energy (a positive quantity), T_w is the temperature at the heat transfer surface, and u is the local flow velocity. More detailed models including protein

kinetics have been proposed, for example, Petermeier et al. (2002). Equation 4.4.10 indicates that high surface temperatures and low flow velocities will give high fouling rates so these should be avoided where possible. If the heating utility temperature is kept constant, the temperature at the surface of the deposit in contact with the milk will decrease: in this case, as deposit grows, the fouling rate will change and falling rate or asymptotic fouling behavior may be observed. If the heating utility temperature is adjusted in order to maintain the outlet temperature at its target value, the deposit surface temperature will not change as much and fouling is likely to proceed at a similar rate. The dynamics are, however, complicated by ageing and modification of the flow pattern by deposit buildup.

When an exchanger design features excess area, consideration should thus be given to manipulating the heating (or cooling) utility temperature in order to reduce the initial surface temperature. The excess area will result in the product still reaching the desired outlet temperature. When steam is used as a heating utility, this can be achieved by modifying the steam pressure. As fouling occurs, the utility temperature can be adjusted to maintain the target performance. This control strategy has the added advantage of minimizing ageing of deposit driven by high wall temperatures.

Plate and frame heat exchangers are widely used in dairy applications. These feature heating, holding, and cooling sections within a unit, and the configuration of the exchanger can be manipulated to give temperature distributions that mitigate fouling (see De Jong, 1997). The plate layout (chevron angle, number and spacing of channels, and distributor layout) dictates the flow distribution, pressure drop, and overall heat transfer coefficient (see Picon-Nunez et al., 2010) as well as fouling behavior, and manufacturers may advocate different designs for units in fouling service. Detailed numerical models of plate heat exchangers subject to milk fouling have been developed (e.g., Georgiadis and Macchietto, 2000; Grijspeerdt et al., 2004) but have yet to find widespread use in industrial design.

Crude Oil Fouling

Crude oil is a mixture of many organic components and may contain inorganic species as well as water. The simplest characterization of crudes classifies species by solvency into saturates, aliphatics, resins, and asphaltenes. Crudes vary according to source and transportation history, and the crude oil entering a refinery will often contain a water fraction with dissolved salts as well as suspended solids (sand, clay particles, and minerals such as iron sulfide). No two crudes are the same, and refineries frequently blend batches of crudes to give the required product slate.

Crude oil fouling is therefore complex. The multicomponent nature of a crude or crude blend means that fouling can arise from several mechanisms (see Watkinson and Wilson, 1997). Figure 4.4.6 shows a schematic of a refinery distillation unit preheat train and identifies the regions where different fouling mechanisms occur, which is chiefly determined by the temperature regime.

Two of the phenomena in the figure are related to phase behavior. Most preheat trains feature a desalter to remove water and its dissolved salts because the water will flash off at elevated temperature, giving a vapor phase and an associated high-pressure drop. This deprives the salts of a solvent and they will form crystalline deposits. Poor control of a desalter will promote fouling downstream of the unit. Similarly, if the crude is allowed to vaporize at higher temperatures before it reaches the furnace, this will increase the pressure drop (possibly limiting throughput) and the loss of volatile species from the liquid will change the solubility of some of the species. The solubility of the many species collectively known as asphaltenes is a complex phenomenon (see Wiehe, 2008), and every effort should be made to ensure that these stay in solution before the crude reaches the distillation column.

Crude Oil Fouling Models

The historical method for designing heat exchangers based on fouling factors (see "Quantitative Impact of Fouling" section) is widely used for units in crude oil service and is described in texts

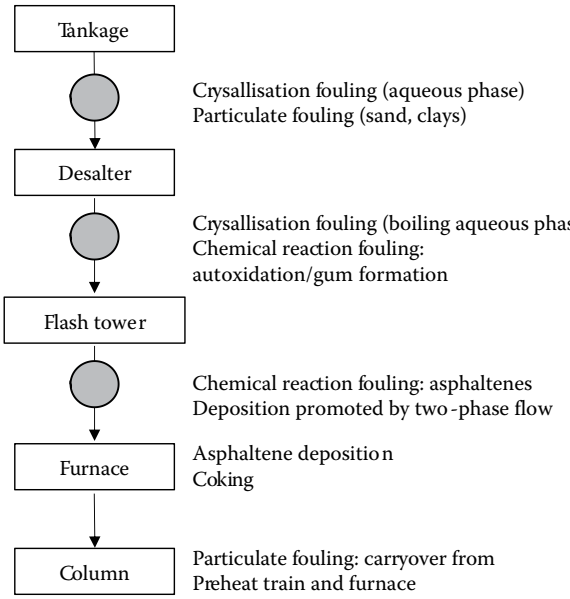


FIGURE 4.4.6 Schematic of crude oil fouling behaviors observed in an oil refinery preheat train.

such as Saunders (1988) and Kakaç and Liu, 2002. Bennett and Nesta (2004) presented an alternative, the "no foul" approach, in which fouling factors are *not* used. They advocated the application of guidelines for exchanger design based on critical velocities and maximum surface temperatures which should preclude significant fouling, and the use of a safety margin (in effect, excess area) to account for uncertainties in predicting heat transfer and a small amount of fouling. Table 4.4.5 is a summary of their "no foul" approach.

Since 1995, considerable attention has been paid to developing quantitative fouling models to provide a means of relating the rate of fouling encountered within a heat exchanger processing crude

TABLE 4.4.5
Summary of Bennett and Nesta (2004) "No Foul" Approach for Shell-and-Tube Units in Liquid Hydrocarbon Service

Parameter	Guideline
Fluid	Liquids with API gravity <45° No heavy particulate matter such as catalyst fines No or low salt content
Fouling factors	Set to zero unless the fluid is outside the scope and is known to foul. A modest safety margin (20% or less excess area, depending on experience) may be used
Tube-side velocity	$u \geq 2$ m/s d_o up to 25.4 mm $u \geq 2.2$ m/s for $d_o \geq 31.75$ mm
Wall temperature	The maximum tube wall temperature should be 300°C
Shell-side design	Single segmental baffles with baffle cut of 20% Spacing shell-side velocity should be at least 0.6 m/s and selected to give a shear stress similar to that for the tube-side guideline if fouling is expected Baffle cut orientation, end spaces, impingement protection, longitudinal baffles: see Bennett and Nesta (2004)
Pressure drop	As required to meet tube-side and shell-side velocity guidelines

oil to its operating condition. This approach is consistent with the Bennet and Nesta guidelines and is discussed here in detail as it is expected to be the subject of future developments in the field.

The purpose of a fouling model is to indicate how the operating conditions in a unit can be specified or changed in order to reduce fouling within the exchanger. It can be used to design a new exchanger, for example, by identifying critical velocities and surface temperatures, and also to predict how an existing exchanger will perform in a revised network.

A variety of models have been proposed for fouling caused by heating of crude oils in refinery preheat train applications. These models have incorporated a variety of considerations, including mass transfer, chemical reaction, physical adhesion rates, and ageing of deposits. Modelling of crude oil fouling was reviewed by Wilson et al. (2017).

To be of practical use, a fouling model must satisfy two requirements. It needs a reasonably sound physical basis and it should have a small number of characterizing parameters. There is a conflict between these requirements, as crude oil fouling is complex and it may not be possible to simplify reality to a small number of parameters. For instance, it is known that the thermal conductivity of asphaltene deposits increase over time as ageing converts the layer to coke. This means that the thermal resistance presented by the deposit will reduce over time. Mathematical models that account for this process have been developed but the model parameters are often obtained by fitting the models to noisy plant data, which are inherently of questionable accuracy as they are subject to instrument error etc. It is difficult to reproduce crude oil fouling conditions in laboratory tests so the methodology described here is based on interrogating plant data sets to identify parameters for simple fouling models, and to apply these to exchanger design and operation. Detailed descriptions of the approach are given in Wilson et al. (2002); Polley et al. (2002a, b), and Yeap et al. (2005).

Three fouling models that fit the abovementioned requirements of reasonable physical basis and small number of parameters are the (1) Ebert–Panchal model (1997) and its variants; (2) the Epstein model (see Epstein, 1994; Yeap et al., 2005); and (3) the asphaltene precipitation model (Polley et al., 2011a; see Wilson et al., 2017). A particular feature of the Epstein model is its capacity to describe fouling controlled by mass transfer or by chemical reaction, such that there can be a maximum in the observed fouling rate as velocity is increased. It does contain more parameters as a result. The Ebert–Panchal model is used here owing to its frequent use in published work at the time of writing. It is recognized that other models have been, and will be, developed over time.

Ebert–Panchal Model

This model has been used widely to relate fouling rate data obtained from plant operating data for exchangers subject to chemical reaction fouling from crude oil streams. It is similar in form to the Paterson and Fryer model, Equation 4.4.10, but includes a second term describing fouling suppression by fluid shear.

$$\mathrm{d}R_\mathrm{f} / \mathrm{d}t = C_1 \exp\left(-E_\mathrm{A} / RT_\mathrm{f}\right) / h - C_2 \tau_\mathrm{w}, \tag{4.4.11}$$

where C_1, C_2, and E_A are fouling model parameters, which have to be identified: h is the relevant film heat transfer coefficient (tube or shell side), τ_w is the local wall shear stress, and T_f is the film temperature, given by

$$T_\mathrm{f} = 0.55\ T_\mathrm{w} + 0.45\ T_\mathrm{bulk}. \tag{4.4.12}$$

The fouling suppression term, $-C_2\ \tau_\mathrm{w}$, is sometimes labelled as a "removal" term. The latter suggests that deposit, once attached, can be removed by fluid shear alone. This has not been confirmed in any systematic study. Equation 4.4.11 does, nevertheless, suggest that an acceptably low fouling rate can be achieved by selecting appropriate combinations of T_f, τ_w, and h. All three quantities are related by the flow velocity (and the design): this is the basis of the "threshold fouling" approach to mitigating fouling described in "Threshold Fouling and the Temperature Field Plot" section. This

extends the critical velocity and critical temperature guidelines of the "no foul" approach in Table 4.4.5 to allow different exchanger designs to be identified and an optimal case selected.

Values of E_A obtained from regression fitting to plant data range from 20 to 80 kJ/mol and particular care needs to be taken in estimating this parameter owing to the sensitivity of the exponential term. Polley et al. (2010) have suggested, following work by Wiehe on the reaction of maltenes to asphaltenes and subsequent deposition, that an E_A value of 44.3 kcal/g mol (185 kJ/mol) should be used for fouling dominated by asphaltene deposition. This is typically at the hot end of a preheat train and in a furnace.

Threshold Fouling and the Temperature Field Plot

The "threshold fouling" concept arises from the form of Equation 4.4.11, which suggests that an acceptable level of fouling (or fouling rate) can be achieved by judicious selection of flow velocity and operating temperatures. The label "fouling threshold" arises from the case where no fouling is observed (i.e., $dR_f/dt = 0$). Rearranging Equation 4.4.11 gives

$$\exp\left(-E_A / RT_f\right) = h / C_1 \times \left(\left[dR_f / dt\right]_{target} + C_2 \tau_w\right), \tag{4.4.13}$$

where $[dR_f/dt]_{target}$ is the acceptable fouling rate. For turbulent flow through a tube, this can be written in terms of the mean velocity, u:

$$\exp\left(-E_A / RT_f\right) \propto u^{0.8} / C_1 \times \left(\left[dR_f / dt\right]_{target} + C_3 u^{1.75}\right) \tag{4.4.14}$$

For a given value of u, the film temperature required can then be calculated if the constants are known. The corresponding wall temperature can then be calculated.

In practice, in a refinery heat exchanger network, the film and surface temperature in a unit are set by the network configuration. Crude oil passing from one exchanger to the next can experience a large change in film temperature owing to the difference in hot stream temperatures. The temperature field plot is a construction that allows the effect of network configuration on fouling behavior to be assessed.

The crude bulk temperature is plotted on the x-axis, which thus represents the location in the network. The hot stream temperature and the film temperature are plotted on the y-axis. Accurate determination of the film temperature requires calculation of the wall temperature, and the rigor which is used to do this depends on the stage of design. For a simple estimate, the film temperature could be taken as the arithmetic average of the hot and cold stream temperatures. Each exchanger is plotted as a line linking the temperature conditions (T_{crude}, T_f) at its inlet and outlet: in Figure 4.4.7, the exchangers are assumed to operate in countercurrent mode. The distance along the x-axis is related to the gain in crude temperature and thus the exchanger duty.

Also plotted are the solutions to Equation 4.4.13 for different values of u, which is set by the number of tubes, passes, and size of tubes in a shell-and-tube exchanger. For each value of T_{crude}, there is a corresponding value of T_f. If the heat exchanger line lies on or below a fouling rate locus, this means that a unit designed with that value of u will be subject to acceptable levels of fouling. The further a heat exchanger line lies above a locus, the larger the fouling rate will be: a higher velocity would be needed to avoid fouling.

Two exchanger lines are plotted on the figure, corresponding to different hot streams for the same crude duty. Match A features lower hot stream temperatures and thus lower T_f values, and requires a lower velocity to avoid fouling. The temperature driving force in this match is smaller so the exchanger would need to be larger, but the pressure drop is likely to be smaller. Critically, the pressure drop is unlikely to change much over time as fouling will be modest. It should be noted that no excess area has been added at this stage. Match B, with higher hot stream temperatures, offers a larger temperature driving force and smaller area but would require higher velocities to avoid fouling. If higher velocities were not allowed, any excess area added to accommodate fouling could offset the difference in areas for the two clean designs.

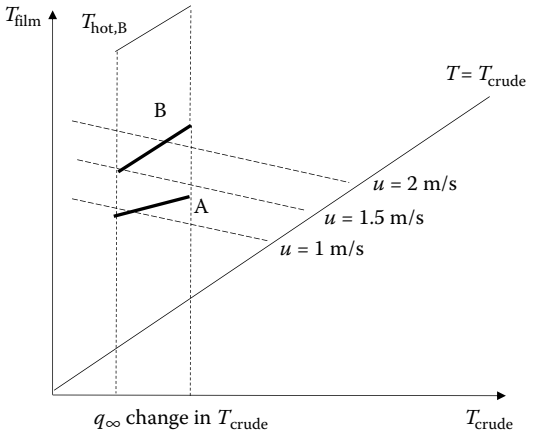


FIGURE 4.4.7 Temperature field plot construction. Dashed lines show T_{film} locus for the acceptable fouling rate for the u value indicated. Solid lines indicate T_{film} values expected in exchangers A and B. Vertical dotted lines show construction for exchanger B, where T_{film} is a weighted average of T_{hot} and T_{crude}. Exchanger A features a cooler hot stream.

The effect of configuration and fouling on pressure drop across the exchanger can also be included in the construction, in the modified temperature field plot (see Wilson et al., 2002).

Figure 4.4.7 provides insights into how fouling should be considered when considering network placement. The temperature field plot can also be used for retrofits and revamps (see Yeap et al., 2005). It can also be seen that as the crude temperature increases, higher velocities are needed (for the use of inserts or alternative designs, which will be described by different fouling model parameters, see "Use of Fouling Models in Exchanger Design" section).

The film temperature, and thus the fouling rate, can differ significantly between the crude inlet and outlet. Since the performance of the unit is determined by an average of the fluxes across the exchanger, an equivalent average fouling rate and fouling resistance is required. Ishiyama et al. (2008) compared different averages for the case where the film temperature increased linearly across the exchanger. Good agreement was obtained for an expression based on the exponential integral which can be evaluated using modern spreadsheets: the geometric average will give a reasonable estimate as long as the fouling rates at each end are non-zero.

Use of Fouling Models in Exchanger Design

Shell-and-tube exchangers are commonly used for crude oil service owing to their ability to handle the process conditions (temperatures and pressures) at the scale required, and the availability of techniques for cleaning fouled tube bundles. This section will therefore focus on the design of shell-and-tube units for refinery preheat train service.

The flow rates in refinery preheat trains are large. It is therefore common to find many trains with exchangers of large diameter (greater than 1 m). This is often the result of arguments based on economy of scale. However, such arguments only apply to the capital cost of the exchanger; the "lifetime" cost of an exchanger is dominated by operational costs, particularly by costs associated with exchanger downtime and reduced productivity. Consideration of fouling needs to be an integral part of exchanger design.

The design process involves the following steps:

a. Specifying the process streams, their properties, inlet temperatures, and any outlet target temperatures. This sets the exchanger duty or range of duties.

b. Specifying the maximum pressure drop for each stream, which may include an absolute pressure to avoid boiling or condensation.
c. Identifying likely exchanger configurations (size, type, and number of tubes; number of tube passes, shell passes, etc.) as well as baffle arrangement.
d. Calculating the heat transfer coefficients and duty for given configurations, including considerations of turndown and process variation.
e. Identifying optimal designs for final selection.

The fouling model approach provides guidance for steps (c) and (d), for use in tandem with computational design tools.

Simple guidelines for step (c) include the following:

Baffle Configuration

Exchanger design software often allows the baffle cut to be specified anywhere within the range 15–45%, and any spacing that will fit within the shell. Little or no guidance is given on how the baffle configuration will affect fouling within the shell. Polley et al. (2011b) and Bouhairie (2012) both used computational fluid dynamics simulations (CFD) to study flow fields within heat exchanger shells and recommended that only geometries providing close to uniform flow be used, confirming Gilmour's advice. This means baffle cuts are in the range 15–25% and a window-to cross flow area ratio in the range 0.8–1.2 (see Tables 4.4.3 and 4.4.5). It is also important that a sufficiently high velocity is maintained within the exchanger shell. Before that advent of computer programs, it was standard practice to ensure that the cross-flow velocity exceeded 0.5 m/s.

Cleaning

The need to clean can preclude some bundle layouts: a square pitch or 45° layout is preferred. It also affects the suitability of U-tube and twisted tube designs as well as other technologies (plate units, plate-and-shell units etc.).

Fouling Maps

"Threshold Fouling and the Temperature Field Plot" section introduced the concept of the "fouling threshold," which is a combination of flow velocity and film temperature which will give little fouling. In many applications, it is not possible to identify a design that will allow operation below the fouling threshold and an acceptable fouling rate, $[\mathrm{d}R_\mathrm{f}/\mathrm{d}t]_\mathrm{target}$ will need to be identified. This provides an indication of how frequently the unit may have to be cleaned and where the installation of tube inserts will be beneficial (see below).

Equation 4.4.13 can then be used to sketch a "fouling map" to compare the designs identified in step (c) that satisfy the pressure drop and heat transfer requirements. The fouling map is a plot of the film temperature which will give the specified fouling rate ($[\mathrm{d}R_\mathrm{f}/\mathrm{d}t]_\mathrm{target}$) for a given tube or shell-side velocity. The latter parameter is set by the flow configuration; the film temperature is determined by the resultant heat transfer calculation.

The calculation method is as follows:

1. From the local flow velocity, u, determine the wall shear stress, τ_w, and the film heat transfer coefficient, h;
2. Solve Equation 4.4.13 to find T_f;
3. For each candidate design, calculate T_f at the inlet and outlet, and other salient points (see below) to see where the design lies on the fouling map. The film temperature is calculated using Equation 4.4.12, with T_w given by

$$T_\mathrm{w} = T_\mathrm{cold} + \left(T_\mathrm{hot} - T_\mathrm{cold}\right) \times \left(1 + r_\mathrm{i} h_\mathrm{i} \ln\left(r_\mathrm{o} / r_\mathrm{i}\right) / k_\mathrm{wall} + r_\mathrm{i} h_\mathrm{i} / r_\mathrm{o} h_\mathrm{o}\right). \tag{4.4.15}$$

Figure 4.4.8 shows an example for an application with fouling on the tube side. For a tube side velocity of 2 m/s and a fouling rate of 10^{-6} m^{2}K/Wh, the critical film temperature is 169°C: for a tube side velocity of 2.5 m/s, the critical film temperature is about 180°C. If the bulk crude temperature is 150°C, the corresponding wall temperatures are 185°C and 205°C, respectively. These values can be compared with T_w calculated using Equation 4.4.15 or generated by the heat exchanger design software using its internal algorithms. The design can then be rejected at this level of fouling, or retained for comparison with others.

The same information can be used in the Poddar plot construction for identifying candidate designs (see Polley et al., 2002b).

It is important to note that the shell configuration determines the wall temperature distribution and thus the location where the largest fouling rate is likely to arise. Figure 4.4.9 summarizes some important differences between standard TEMA (Tubular Manufacturer's Manufacturing Association) configurations. Similarly, for multiple shells, Figure 4.4.10 shows how the flow arrangement can be manipulated to reduce the maximum wall temperature. It is important to note that configurations or arrangements that are favorable in terms of fouling may not be optimal in terms of clean heat transfer (e.g., capital cost), but are more likely to perform reliably over time (e.g., operating cost), when operated correctly.

The fouling model approach can be extended to determine fouling on the shell side of an exchanger as well as tubes fitted with inserts (see Table 4.4.2). In both cases, a component of the pressure drop is associated with form drag in addition to skin friction, and this has to be compensated for in calculating the wall shear stress that appears in the fouling model.

"Self-Cleaning" Inserts

As more heat is recovered in a preheat train, the crude oil bulk temperature reaches levels where the associated wall temperature is too high to be countered by allowed liquid velocities. The fouling rate can be so high that regular cleaning is required to maintain productivity. As a result, bypass exchangers, fluidized bed units, and inserts are regularly employed at the hot end of the preheat train.

The performance of Turbotal inserts has been successfully described by the fouling model approach. These devices feature a rotating element that reduces the rate at which fouling occurs

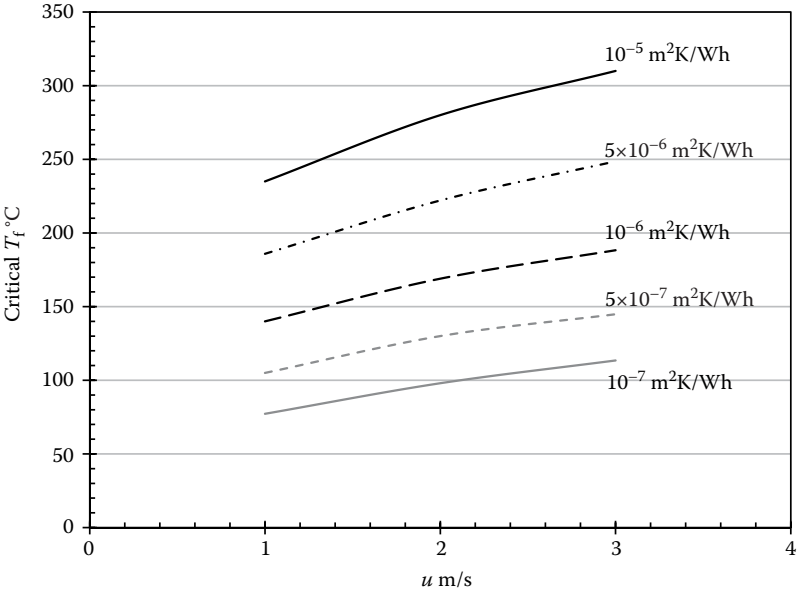


FIGURE 4.4.8 Example of fouling map.

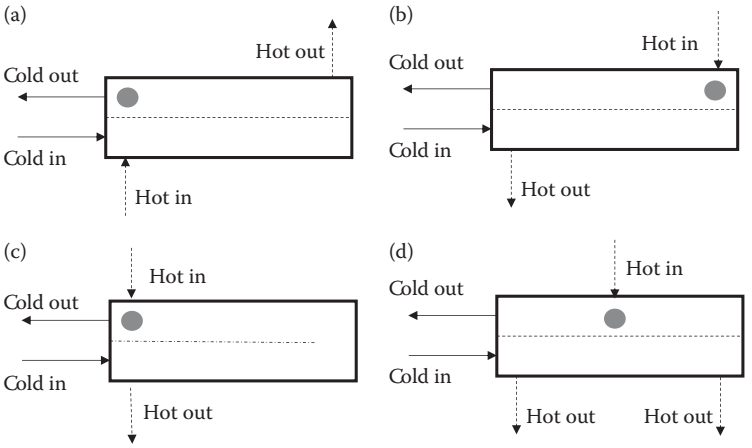


FIGURE 4.4.9 Thermal contacting patterns in single shell, twin tube pass configurations. The designs are equivalent in terms of heat transfer but the T_w distributions differ. The circles indicate regions with high T_w, where the hot stream enters the shell and where the cold stream is warm (After Polley, G.T. et al., *Appl. Therm. Eng.*, 22: 777–788, 2002b.) (a, b) Two tube side pass, one shell side, (c) TEMA F shell, and (d) TEMA J shell (lower shell-side pressure drop than F, less good in terms of fouling propensity).

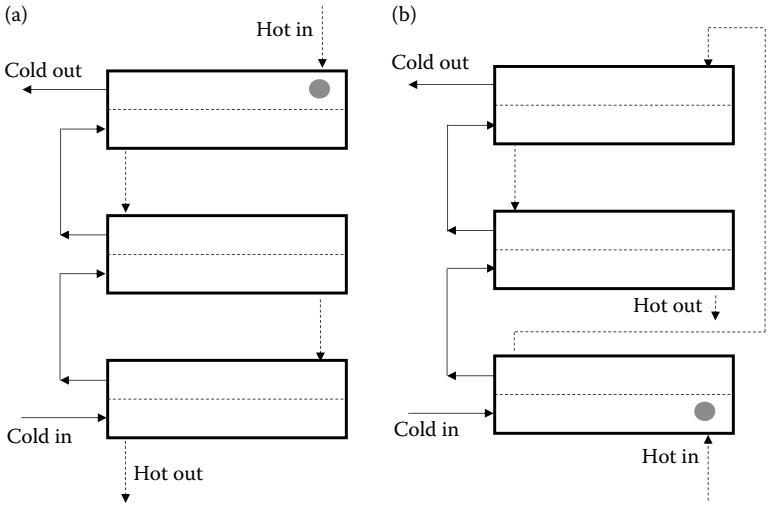


FIGURE 4.4.10 Shells-in-series arrangement: (a) normal pattern and (b) one that reduces fouling. Circles indicate the region of thermal contact where the hot stream enters the shell: T_w differs significantly between the two designs. (After Polley, G.T. et al., *Appl. Therm. Eng.*, 22: 777–788, 2002b.)

compared to a plain tube and subsequently control the thickness of the deposit formed such that the fouling resistance is maintained at a constant value. Aquino et al. (2007) and Polley and Gonzales-Garcia (2009) presented a modified form of the Ebert–Panchal model for the prediction of the initial fouling rates in tubes fitted with inserts:

$$\begin{aligned} dR_f / dt &= h_{i,plain} / h_{i,enhanced} \times C_1 Re^{-0.66} Pr^{-0.33} \exp\left(-E_A / RT_f\right) \\ &\quad - 0.7\ C_2 \tau_w \times \Delta P_{enhanced} / \Delta P_{plain}. \end{aligned} \tag{4.4.16}$$

The modification is a simple one. The deposition term has been multiplied by the ratio of the plain tube heat transfer coefficient to the enhanced coefficient. This modification quantifies the reduced

thickness of the thermal film, and hence reduction in volume of reacting liquid assumed by Ebert and Panchal. The suppression term has been multiplied by the ratio of the pressure drop encountered in the enhanced tube to that in the plain tube. This term is multiplied by 0.7 to reflect that only 70% of the pressure drop encountered during flow through tubes fitted with inserts is associated with wall friction. The remainder is caused by form drag. These modifications allowed the fouling rate to be estimated.

Eventually, a layer of adherent layer fills the gap between the rotating insert and the wall, and the insert scours any further fresh material. Aquino et al. (2007) found that this gave an asymptotic value of R_f of 0.0035 m^{2}K/W. This was in good agreement with an estimate of R_f based on the insert clearance and a deposit thermal conductivity of 0.43 W/mK (see Watkinson, 1988). Figure 4.4.11 shows how their simulation of the performance of a two-shell unit over 9 months matched the R_f values estimated from plant measurements. Readers unfamiliar with fouling data may find the scatter and uncertainty in this plot disconcerting. It is, however, a feature of many calculations based on plant operating data (see Crittenden et al., 1992; Takemoto et al., 1999).

In considering the use of such inserts, the designer would have to consider the impact of the insert on pressure drop as well as fouling: the absolute pressure at the hot end of the preheat train usually needs to be maintained above the bubble point of the crude in order to avoid flashing of the oil. If there was sufficient pressure drop, a unit could be designed on the basis of the asymptotic fouling resistance. The abovementioned guidance on flow distribution as well as that on operation would still apply.

Further Factors to Consider in Shell-and-Tube Exchanger Design

Effect of Fouling on Shell-Side Flow and Heat Transfer

When fouling does occur on the shell side, the deposits are usually not evenly distributed over the metal surfaces. Deposition is often heaviest down the face of the "trailing" baffle (the one forming

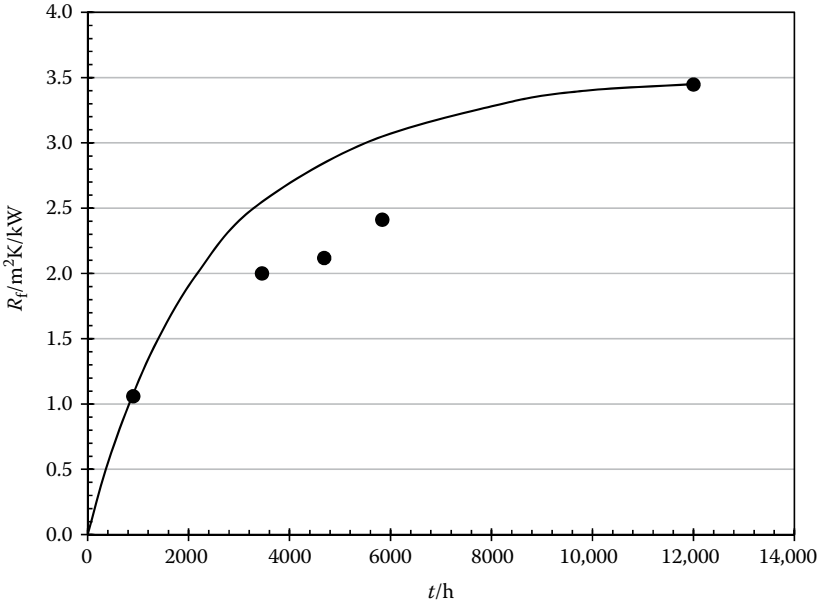


FIGURE 4.4.11 Use of Turbotal insert to mitigate fouling in a crude preheat train exchanger subject to severe fouling. Symbols: R_f estimated from plant measurements: line—predictions of a simulation with fouling predicted using Equation 4.4.16. (Reproduced from Aquino, B. et al., *Proceedings of 7th Engineering Foundation Conference on Heat Exchanger Fouling and Cleaning*, Tomar, Portugal, http://dc.engconfintl.org/heatexchanger2007/51, 2007.)

the window through which flow enters the baffle space), around the perimeter of the baffles, and in the corner between the shell and the baffle diverting the flow into the baffle space. This results in blockage of the clearances between the tubes and the baffle plates and between the baffle plate and the heat exchanger shell.

These blockages result in increased pressure drop but they also lead to an increase in the shell-side heat transfer coefficient. It is this phenomenon that led Gilmour (1965) to state that fouling can be eliminated by putting the fouling stream on the shell side rather than the usual practice of passing it through the tubes. However, this benefit can only be realized if the baffle arrangement provides a good flow field.

Fouling and Cleaning

It is harder for a deposit to grow on a clean surface than on existing deposit: the cleaning methodology influences the speed at which heat exchangers will foul when returned to service.

Chemical cleaning involves the recirculation of solution through the exchanger in order to dissolve the deposit. When strongly aged deposits are present, residual layers are likely to remain and the effectiveness of chemical cleaning will decrease as the number of cleaning cycles increases. For example, gas oil washes are used in some installations but this is unlikely to remove coked material or mineral deposits. Residual layers that will promote fouling are thin (microns thick), as they influence the surface chemistry: it is impossible to determine their presence from flow or heat transfer measurements.

Cleaning methods based on hydraulic forces, termed "mechanical" cleaning, including steam lances and high-pressure water jets, are in widespread use (see Müller-Steinhagen and Zettler, 2011). These techniques do not clean to the microscopic level. More stringent removal is achieved by pigging and by hydro-drilling. Pigging uses high-pressure water to propel a cleaning plug along a tube, driving debris ahead of it before being collected by a basket at the exit. In hydro-drilling, high-pressure water lances fitted with rotating nozzles are passed along the exchanger tube. A very high-speed water jet is then directed onto any deposit. In both cases, the shear forces applied can remove hard deposits and also the outer layers of metal (removing any corrosion deposits that may have initiated fouling in the first instance). Hydro-drilling and pigging are readily applied to straight, empty tubes but cannot be used in every exchanger type. Notably, they cannot be used with twisted tubes or in welded plate units.

The use of acoustic resonance in conjunction with chemical cleaning has proved effective for cleaning badly fouled units in some instances. This requires specialist contractors, and its effectiveness is not guaranteed in every instance.

Fired Heaters and Two-Phase Flow

Preheat trains and their exchangers are often designed to avoid two-phase flow owing to the high-pressure drop associated with the higher volumetric flow rate. Some units do nevertheless operate with flashing of the crude. The Ebert–Panchal fouling model can also be applied to situations in which two-phase flow is occurring within plain tubes: the h value used is that due to two-phase convective heat transfer, and τ_w is taken from the frictional component of the two-phase pressure loss.

Fired heaters are designed to vaporize part of the feed. The wall temperatures are often high enough to promote coking, and it is important to recognize the different requirements of fired heaters designed for coking operations from those preheating distillation column feeds: coke generated in a fired heater can cause significant operating problems for a column.

Fired heater design is complicated as the thermal and hydraulic performance of the unit are intimately linked. The optimization of preheat train fired heaters was studied in detail by Morales-Fuentes et al. (2014). They highlighted one aspect of current fired heater design practice that requires examination. Traditionally, fired heaters consist of two separate sections; a convective section and a radiant section, each with a number of parallel pipes. The crude is fed to the convective section before passing to the hotter radiant section, with the pipes in the convective section feeding a common manifold which feeds the radiant tubes.

In the sections of the convective section where the crude is liquid, the heat transfer coefficient will be low, giving high wall temperatures and high fouling rates. It can lead to excessive fuel consumption as higher wall temperatures will be required for a given duty, and consequently faster fouling. Allowing the crude to vaporize in the convective section will increase heat transfer rates and reduce fouling. However, if the fluid leaving the convective section contains vapor, the two phases are likely to separate in the manifold, and this could give rise to some tubes in the radiant section handling vapor and other liquid. The tubes handling vapor could subsequently fail through excessive wall temperatures. It follows that designers should seek to modify the construction of fired heaters such that a two-phase flow can be generated as early as possible and maintained throughout the unit.

Closing Remarks

Fouling is the major unresolved problem in heat transfer because it is a complex and dynamic phenomenon. This chapter has sought to communicate some simple lessons about fouling, to which local experience and design criteria must be added. The following closing statements by Gilmour in 1965 apply equally well today:

"One could go on and on, citing paradoxical instances in relation to fouling and heat transfer, but it is hoped that by now we have made the point clear that the fouling factor should not be relied upon as the means of combating the fouling problem.

Selling Ideas to Engineers

It has been, of course, impossible for us to condense thirty years of know-how on fouling prevention in this short discourse. It will also be impossible for one to sell these ideas to engineers who have never been in an operating plant, have never evaluated test data, have never issued specification for rebuilding an existing—but inadequately operating—exchanger, and who are concerned only with the inconsistent requisites of lower first cost and highest surface area. It is equally impossible to expect the fabricators to accept these ideas, because their object is to sell lots of surface area, their burden is to guarantee performance on nameless fluids, and their ululation is that they never receive adequate feedback on the performance and maintenance of the heat exchangers they sell. Therefore, they need to use a generous fouling factor to protect their interests.

Foster Dedicated Employees

It follows, therefore, that if a used company desires to reap economic benefits, it behoves that company to foster one or more of their own dedicated employees to become proficient in the design, procurement, testing, and maintenance of heat exchangers for their own requirements. This job entails design, specification, review of vendors' detailed drawings, observance of fabrication and installation, performance testing, and observance of units removed for maintenance."

REFERENCES

Al-Hadhrami, L.M., Ahmad, A., and Al-Qahtani, A. 2010. Performance analysis of heat exchangers of an existing naphtha hydrotreating plant: A case study. *Appl. Therm. Eng.* 30: 1029–1103.

Aquino, B., Derouin, C., and Polley, G.T. 2007. Towards an understanding of how tube inserts mitigate fouling in heat exchangers. *Proceedings of 7th Engineering Foundation Conference on Heat Exchanger Fouling and Cleaning*, Tomar, Portugal. http://dc.engconfintl.org/heatexchanger2007/51/.

Bansal, B.P., and Chen, X.D. 2006. A critical review of milk fouling in heat exchangers. *Compr. Rev. Food Sci. Food Saf.* 5: 27–33.

Bennett, C.A., and Nesta, J. 2004. Reduce fouling in shell-and-tube heat exchangers. *Hydrocarbon Process.* Article 2599305, February 2004 issue. http://www.hydrocarbonprocessing.com/Article/2599305/Reduce-fouling-in-shell-and-tube-heat-exchangers.html.

Bott, T.R. 1995. *Fouling of Heat Exchangers*. Amsterdam: Elsevier.

Bott, T.R. 2008. Biofouling control in cooling water. *Int. J. Chem. Eng.* Article ID 619873. http://dx.doi.org/10.1155/2009/619873.

Bouhairie, S. 2012. Selecting baffles for shell-and-tube heat exchangers. *Chem. Eng. Prog.* 108(2): 27–33.

Brignone, V.M., Perrone, F., Rottoli, M., Pugh, S.J., and Ishiyama, E.M. 2015. EMbaffle® in refinery service on-field study and data validation through SmartPM®. *Proceedings of 11th International Conference on Heat Exchanger Fouling and Cleaning.* http://www.heatexchanger-fouling.com/papers/papers2015/14_Garavaglia_F.pdf.

Changani, S.D., Belmar-Beiny, M.T., and Fryer, P.J. 1997. Engineering and chemical factors associated with fouling and cleaning in milk processing. *Exp. Therm. Fluid Sci.* 14: 392–706.

Cosado, E. 1990. Model optimizes exchanger cleaning. *Hydrocarbon Process.* 69: 71–76.

Crittenden, B.D., Kolaczkowski, D.T., and Downey, I.L. 1992. Fouling of crude-oil preheat exchangers. *Chem. Eng. Res. Des.* 70: 547–557.

De Jong, P. 1997. Impact and control of fouling in milk processing. *Trends Food Sci. Technol.* 8: 401–405.

Drögemüller, P., Osley, W., and Philipp, D. 2013. Increased shear, reduced wall temperatures, use of hiTRAN wire watrix inserts in systems subject to fouling. *Proceedings of 10th International Conference on Heat Exchanger Fouling and Cleaning.* http://www.heatexchanger-fouling.com/papers/papers2013/55_Droegmueller_F.pdf.

Ebert, W., and Panchal, C.B. 1997. Analysis of Exxon crude-oil slip stream coking data. *Fouling Mitigation of Industrial Heat Exchange Equipment*, eds. Panchal, C.B., Bott, T.R., Somerscales, E.F.C., and Toyama, S., pp. 451–460. New York: Begell House.

Epstein, N. 1983. Thinking about heat transfer fouling: A 5×5 matrix. *Heat Transfer Eng.* 4: 43–56.

Epstein, N., 1994. A model of the initial chemical reaction fouling rate for flow within a heated tube, and its verification. *Proceedings of 10th International Heat Transfer Conference, Brighton, IChemE* 4: 225–229.

ESDU. 2003. Fouling in cooling systems using seawater. IHS-ESDU Data Item 03004. London: IHS-ESDU.

ESDU. 2008. Fouling in cooling systems using fresh water. IHS-ESDU Data Item 08002. London: IHS-ESDU. See also Pugh, S.J., Hewitt, G.F., and Müller-Steinhagen, H. 2009. Fouling during the use of "fresh" water as coolant—The development of a "user guide". *Heat Transfer Eng.* 30: 851–858.

Georgiadis, M.C., and Macchietto, S.M. 2000. Dynamic modelling and simulation of plate heat exchangers under milk fouling, *Chem. Eng. Sci.* 55: 1605–1619.

Gilmour, C.H. 1965. No fooling, no fouling. *Chem. Eng. Prog.* 61: 49–54.

Gomes da Cruz, L., Ishiyama, E.M., Boxler, C., Augustin, W.A., Scholl, S., and Wilson, D.I. 2015. Value pricing of surface coatings for mitigating heat exchanger fouling. *Food Bioprod. Process.* 93: 343–363.

Grijspeerdt, K., Mortier, L., De Block, J., and Van Renterghem, R. 2004. Applications of modelling to optimise ultra high temperature milk heat exchangers with respect to fouling. *Food Control* 15: 117–130.

Ishiyama, E.M, Paterson, W.R., and Wilson, D.I. 2008. Thermo-hydraulic channelling in heat exchangers caused by fouling. *Chem. Eng. Sci.* 63: 3400–3410.

Ishiyama, E.M., Paterson, W.R., and Wilson, D.I. 2009. A platform for techno-economic analysis of fouling mitigation options in refinery preheat trains. *Energy Fuels* 23: 1323–1337.

Ishiyama, E.M., Paterson, W.R., and Wilson, D.I. 2011. Optimum cleaning cycles for heat transfer equipment undergoing fouling and ageing. *Chem. Eng. Sci.* 66: 604–612.

Kakaç, S. and Liu, H. 2002. *Heat Exchangers-Selection, Rating and Thermal Design*, 2nd ed. Boca Raton, FL: CRC Press.

Klaren, R.G., and de Boer, E.F. 2011. Revamping existing severely fouling conventional heat exchangers into a self-cleaning (fluidized-bed) configuration: New developments and examples of revamps. *Heat Transfer Eng.* 32: 320–330.

Klaren, R.G., de Boer, E.F., and Sullivan, D.W. 2007. 'Zero fouling' self-cleaning heat exchanger. *Heat Transfer Eng.* 28: 216–221.

Krueger, A., and Pouponnot, F. 2009. Heat exchanger performance enhancement through the use of tube inserts in refineries and chemical plants—Successful applications: Spirelf®, Turbotal® and Fixotal® systems. *Proceedings of 8th International Conference on Heat Exchanger Fouling and Cleaning.* www.heatexchanger-fouling.com/papers/papers2009/56_Krueger_F.pdf.

Morales-Fuentes, A., Picón-Núñez, M., Polley, G.T., and Méndez-Díaz, S. 2014. Analysis of the influence of operating conditions on fouling rates in fired heaters. *Appl. Therm. Eng.* 62: 777–784.

Müller-Steinhagen, H., and Lancefield, D. 2005. Deposit formation in the evaporator of a sulphuric acid recovery plant for TiO_2 pigment production. *Proceedings of 6th International Conference on Heat Exchanger Fouling and Cleaning.* http://dc.engconfintl.org/cgi/viewcontent.cgi?article=1008&context=heatexchanger2005.

Müller-Steinhagen, H., and Zettler, Z.U. 2011. *Heat Exchanger Fouling: Mitigation and Cleaning Technologies*, 2nd ed. Essen: Publico.

Paterson, W.R., and Fryer, P.J. 1988. A reaction engineering approach to the analysis of fouling. *Chem. Eng. Sci*. 43: 1714–1717.

Petermeier, H., Benning, R., Delgado, A., Kulozik, U., J. Hinrichs, J., and Becker, T. (2002) Hybrid model of the fouling process in tubular heat exchangers for the dairy industry. *J. Food Eng.* 55: 9–17.

Picon-Nunez, M., Polley, G.T., and Jantes-Jaramillo, F. 2010. Alternative design approach for plate and frame heat exchangers using parameter plots. *Heat Transfer Eng.* 31: 742–749.

Polley, G.T., and Gonzales-Garcia, G. 2009. Procedure for applying fouling models to predict overall fouling rates in industrial heat exchangers. *Proceedings of 8th International Conference on Heat Exchanger Fouling and Cleaning*, Schladming, Austria. www.heatexchanger-fouling.com/papers/papers2009/51_Polley_Pverall%20fouling%20rate_F.pdf.

Polley, G.T., Tamakloe, E., and Picon-Nunez, M. 2011a. Models for chemical reaction fouling. In AICHE Spring Meeting, (Chicago). www.aiche.org/conferences/aiche-spring-meeting-and-global-congress-on-process-safety/2011/proceeding/paper/94c-models-chemical-reaction-fouling-0.

Polley G.T., Vidal, M.A., and Auribes Ramirez, A. 2011b. Using simple CFD models to identify efficient baffle arrangements for shell-and-tube heat exchangers, AIChE Spring meeting, paper 113c. www.aiche.org/conferences/aiche-spring-meeting-and-global-congress-on-process-safety/2011/proceeding/paper/113c-using-simple-cfd-models-identify-efficient-baffle-arrangements-shell-and-tube-heat-exchangers-0.

Polley, G.T., Tamakloe, E., Wilson, D.I., Coletti, F., and Macchietto, S. 2010. Development of a model for the prediction of fouling in heat exchangers processing crude oil. *Proceedings of AIChE Spring Meeting*, San Antonio, March 21–25.

Polley, G.T., Wilson, D.I., Yeap, B.L., and Pugh, S.J. 2002a. Use of crude oil threshold data in heat exchanger design. *Appl. Therm. Eng.* 22: 763–776.

Polley, G.T., Wilson, D.I., Yeap, B.L., and Pugh, S.J. 2002b. Evaluation of laboratory crude oil fouling data for application to refinery pre-heat trains. *Appl. Therm. Eng.* 22: 777–788.

Ritchie, J.M., Drögemüller, P., and Simmons, M.J.H. 2007. HITRAN wire matrix inserts in fouling applications. *Proceedings of Engineering Foundation Conference on Heat Exchanger Fouling and Cleaning*, Tomar, Portugal. http://dc.engconfintl.org/heatexchanger2007/52/.

Saunders, E.A.D. 1988. *Heat Exchangers: Selection, Design and Construction*. Harlow: Longman.

Sheikh, A.K., Zubair, S.M., and Haq, U.M., 1996. Reliability based maintenance strategies for heat exchangers subject to fouling. *ASME J. Energy Resour. Technol.* 118(4): 306–312.

Solano, J.P., Garcia, A., Vicente, P.G., and Viedma, A. 2011. Performance evaluation of a zero-fouling reciprocating scraped-surface heat exchanger. *Heat Transfer Eng.* 32: 331–338.

Takemoto, T., Crittenden, B.D., and Kolaczkowski, S.T. 1999. Interpretation of fouling data in industrial shell and tube heat exchangers. *Chem. Eng. Res. Des.* 77: 769–778.

van der Zijden, E.J.J., Brignone, V.M., Rottoli, M., and van Lint, C.F.J.M. 2013. EMbaffle® heat exchanger in fouling operation. *Proceedings of 10th International Conference on Heat Exchanger Fouling and Cleaning*. http://www.heatexchanger-fouling.com/papers/papers2013/46_van%20der%20Zijden_F.pdf.

Visser, H. (Ed.). 1997. *Fouling and Cleaning of Heat Treatment Equipment*. Bulletin of the International Dairy Federation, No. 328. Brussels: IDF.

Watkinson, A.P. 1988. Critical review of organic fluid fouling, Argonne National Laboratory report no. ANL/CNSV-TM-208.

Watkinson, A.P., and Wilson, D.I. 1997. Chemical reaction fouling—A review. *Exp. Therm. Fluid Sci.* 14: 361–374.

Wiehe, I.A. 2008. *Process Chemistry of Petroleum Macromolecules*. Boca Raton, FL: CRC Press.

Wilson, D.I., Ishiyama, E.M., and Polley, G.T. 2017. Twenty years of Ebert and Panchal—What next? *Heat Transfer Eng.* 38(7–8), 669–680.

Wilson, D.I., Polley, G.T., and Pugh, S.J. 2002. Mitigation of crude oil preheat train fouling by design. *Heat Transfer Eng.* 23: 24–37.

Yeap, B.L., Wilson, D.I., Polley, G.T., and Pugh, S.J. 2005. Retrofitting crude oil refinery heat exchanger networks to minimise fouling while maximising heat recovery. *Heat Transfer Eng.* 26: 23–34.

4.5 BIOHEAT TRANSFER

John A. Pearce, Kenneth R. Diller, and Jonathan W. Valvano

Abstract

Heat transfer plays a major role in many processes in living systems. This chapter presents a brief review of bioheat transfer from the perspective of those aspects, which distinguish it from processes in inanimate systems. Of particular interest are the effects of blood perfusion on temperature distribution, coupling with biorate processes, measurement of thermal transport properties in tissues, therapeutic processes, human thermoregulation, and effects on living tissues of elevated temperatures and subzero temperatures. An extensive compilation of properties for biomaterials is also presented.

Introduction

Analysis and control of heat transfer processes in living systems involve phenomena that are not encountered in systems that are not alive. For example, there is a continuous flow of blood through tissues and organs via a complex network of branching vessels. Heat exchange processes are influenced by vascular geometry, which consists of arteries and veins with countercurrent flow over large portions of the network of pairs. For some physiological conditions, arterial and venous blood temperatures are different, neither is equal to the local tissue temperature, and these temperature differences may vary as a function of many transient physiological and physical parameters. Mechanisms of regulation for the thermal state in tissue are quite nonlinear and have presented a major challenge to understand and model. Nonetheless, thermoregulatory processes are critical to the maintenance of life and must be accounted for in the design of many types of systems that interface with humans and animals.

Although considerable progress has been made in defining and modeling thermal processes in living systems, and especially so over the past 15 years, many important problems remain. This section provides a brief discussion of the present state of knowledge and understanding of some of the more active areas of bioheat transfer. Greater detail and additional areas of application of bioheat transfer may be found among the references cited.

Coupling of Temperature History to Rate Processes

Although the processes of life are observed at the system (macroscopic) and microscopic levels of resolution, the true essence of life lies in the multitudinous biochemical phenomena that occur continuously throughout all organisms. Over the past 50 years, the science of biochemistry has elucidated many of the governing life processes at the molecular level. Today the major frontier of knowledge in the life sciences is associated with molecular biology. Increasingly, bioengineers are collaborating with molecular biologists to understand and manipulate the molecules and biochemical processes that constitute the basis of life. Temperature is a primary controlling parameter in the regulation of these rate processes. The study of thermal biology has identified that the rates of nearly all physiological functions are altered by 6%–10% per degree Celsius over wide range of thermal states produced in response to environmental conditions.[1] Likewise, temperature is often altered during therapeutic or diagnostic procedures to produce or measure a targeted effect, based on the fact that a change in local temperature will have a large effect on biochemical process rates. Thus, knowledge of how temperature can be monitored and/or controlled in living tissues is of great value in both the assessment of normal physiological function and the treatment of pathological states.

In assessing the effects of temperature alterations on biochemical rate processes, two broad categories of state changes can be considered: temperatures both above and below the normal physiological thermal steady state. An extensive review of both of these thermal domains has been published recently,[2] to which the reader is referred for further details and bibliography.

A summary of some of the most important and recent advances in bioheat transfer will be discussed in this section.

Tissue Thermal Transport Properties

The transport of thermal energy in living tissue is a complex process involving multiple phenomenological mechanisms including conduction, convection, radiation, metabolism, evaporation, and phase change. The equilibrium thermal properties presented here were measured after temperature stability had been achieved.

Thermal probe techniques are used frequently to determine the thermal conductivity and the thermal diffusivity of biomaterials.[3–6] Common to these techniques is the use of a thermistor bead either as a heat source or a temperature sensor. Various thermal diffusion probe techniques[7] have been developed from Chato's first practical use of the thermal probe.[3] Physically, for all of these techniques, heat is introduced to the tissue at a specific location and is dissipated by conduction through the tissue and by convection with blood perfusion.

Thermal probes are constructed by placing a miniature thermistor at the tip of a plastic catheter. The volume of tissue over which the measurement occurs depends on the surface area of the thermistor. Electrical power is delivered simultaneously to a spherical thermistor positioned invasively within the tissue of interest. The tissue is assumed to be homogeneous within the milliliter surrounding the probe. The electrical power and the resulting temperature rise are measured by a microcomputer-based instrument. When the initial tissue temperature is just below the freezing point, the thermistor heat is removed both by conduction and by latent heat. In this situation, the instrument measures effective thermal properties that are the combination of conduction and latent heat. By taking measurements over a range of temperatures, the processes of conduction and latent heat can be separated. When the tissue is perfused by blood, the thermistor heat is removed both by conduction and by heat transfer due to blood flow near the probe. *In vivo*, the instrument measures effective thermal properties that are the combination of conductive and convective heat transfer. Thermal properties are derived from temperature and power measurements using equations that describe heat transfer in the integrated probe/tissue system.

The following five complexities make the determination of thermal properties a technically challenging task. First, tissue heat transfer includes conduction, convection, radiation, metabolism, evaporation, and phase change. It is difficult but necessary to decouple these different heat transfer mechanisms. Second, the mechanical and thermal interactions between the probe and tissue are complex, and must be properly modeled to achieve accurate measurements. When the probe is inserted into living tissue a fluid pool may form around the probe because of the mechanical trauma. Because the probe is most sensitive to the tissue closest to it, the presence of a pool of blood and other fluids will significantly alter the results. Tissue damage due to probe insertion may also occur *in vitro*. Third, the tissue structure is quite heterogeneous within each sample. Thus, the probe (which returns a single measurement value) measures a spatial average of the tissue properties surrounding the active elements. Unfortunately, the spatial distribution is very nonuniform.[8] The probe is most sensitive to the tissue immediately adjacent to it. It is important to control this effective measurement volume. If the effective volume is too small, then the measurement is highly sensitive to the mechanical/thermal contact between the probe and tissue. If the effective volume is too large, then the measurement is sensitive to the boundary conditions at the surface of the tissue sample. Fourth, there are significant sample to sample and species to species variabilities. One must be careful when extrapolating results obtained in one situation to different situations. Fifth, tissue handling is critical. Thermal properties are dependent on temperature and water content.[9–12]

Blood flow, extracellular water, and local metabolism are factors that strongly affect heat transfer in living tissue, but are difficult to determine or control experimentally. Once a tissue dies, if handled improperly there will be significant water fluxes which will affect tissue thermal properties. Tissues should be stored on a saline-moistened surgical gauze pad and wrapped in aluminum foil and kept at 4°C to minimize tissue mass transfer: collagenous tissues are especially hydrophilic and will absorb substantial amounts of water if immersed, significantly altering their physical properties.

Currently, there is no method to quantify simultaneously the major three parameters: the intrinsic tissue thermal conductivity, k_m, the tissue thermal diffusivity, α_m, and perfusion, w. Either the knowledge of k_m is required prior to the perfusion measurement, or even when k_m is measured in the presence of perfusion, the thermal diffusivity cannot be measured.[6,13]

Background

There are many good reviews of techniques to measure thermal properties.[10,14–16] Thermophysical Properties of Matter is a ten-volume set that catalogs thermal properties. Volumes 1 and 3 contain thermal conductivity data, and Volume 10 contains thermal diffusivity data. Extensive reviews of measurement techniques exist as prefaces to each volume of the set. Additional thermal property data can be found in Kreith and Bohn.[17] John Chato has written an excellent chapter in Heat Transfer in Medicine and Biology, edited by Shitzer and Eberhart, which reviews techniques to measure thermal properties of biologic materials.[10] Valvano has documented the temperature dependence of tissue thermal properties.[11,12] Duck has written an excellent review chapter on this subject.[18]

The thermal diffusion probe was conceived by Chato.[3,10] Significant developments were obtained by Balasubramaniam, Bowman, Chen, Holmes, and Valvano.[19–27] Patel and Walsh have applied the self-heated thermistor technique to nondestructive surface measurements.[8,24–26] Unfortunately, surface probes are unreliable due to poor probe/tissue contact and uncertain boundary conditions at the tissue surface.[24,27]

Self-heated thermistors have been used to measure perfusion.[4,9,19,20,22–24] Effective thermal conductivity, k_{eff}, is the total ability of perfused tissue to transfer heat in the steady state; k_{eff} is the combination of conduction (due to intrinsic thermal conductivity, k_m) and convection (due to perfusion). Measurements of k_{eff} are very sensitive to perfusion. The limitation of most techniques is that the intrinsic tissue thermal conductivity must be known in order to accurately measure perfusion. Holmes and Chen use a combination of steady-state and transient heating modes to determine perfusion without requiring a no-flow calibration measurement.[13,19,20] The uncertainty of k_m significantly limits the perfusion accuracy.[9]

Measurement of Thermal Conductivity and Diffusivity

Methods

In the constant temperature heating technique, the instrument first measures the baseline tissue temperature, T_0. Then, an electronic feedback circuit applies a variable voltage, $V(t)$, in order to maintain the average thermistor temperature at a predefined constant, T_h. The applied thermistor power includes a steady state and a transient term:

$$P(t) = A + Bt^{-1/2} \tag{4.5.1}$$

where A and B are system constants that depend on the characteristics of the probe and the thermal interaction with the tissue in which it is embedded.

In order to measure thermal conductivity, thermal diffusivity, and tissue perfusion, the relationship between applied thermistor power, P, and resulting thermistor temperature rise, $\Delta T(t) = T_h - T_0$, must be known. In the constant temperature method, ΔT is constant. The thermistor bead is treated

as a sphere of radius "a" embedded in a homogeneous medium. Since all media are considered to have constant parameters with respect to time and space, the initial temperature will be uniform when no power is supplied to the probe.

$$T_b = T_m = T_0 = T_a + \frac{Q_{met}}{wc_{bl}} \quad \text{at } t = 0 \tag{4.5.2}$$

where w is the rate of perfusion of blood through the local tissue (mL blood/mL tissue $- s$) and c_{bl} is the specific heat of blood kJ/(kg K).

Let $\overline{T}$ be the temperature rise above baseline, $\overline{T} = T - T_0$. Both the thermistor bead temperature rise ($\overline{T}_b$) and the tissue temperature rise ($\overline{T}_m$) are initially zero. The subscripts b and m refer, respectively, to the bead and the tissue (i.e., the surrounding medium).

$$\overline{T}_b = \overline{T}_m = 0 \quad \text{at } t = 0 \tag{4.5.3}$$

Assuming the venous blood temperature equilibrates with the tissue temperature and that the metabolic heat is uniform in time and space, the Pennes's bioheat transfer[28] in spherical coordinates is given by

$$\rho_b c_b \frac{\partial \overline{T}_b}{\partial t} = k_b \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial \overline{T}_b}{\partial r} \right) + \frac{A + Bt^{-1/2}}{\frac{4}{3}\pi a^3} \qquad r < a \tag{4.5.4}$$

$$\rho_m c_m \frac{\partial \overline{T}_m}{\partial i} = k_m \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial \overline{T}_m}{\partial r} \right) - wc_{bl} V_m \qquad r > a \tag{4.5.5}$$

Perfect thermal contact is assumed between the finite-sized spherical thermistor and the infinite homogeneous perfused tissue. At the interface between the bead and the tissue, continuity of thermal flux and temperature leads to the following boundary conditions:

$$\overline{T}_b = \overline{T}_m \quad \text{at } r = a \tag{4.5.6}$$

$$k_b \frac{\partial \overline{T}_b}{\partial r} = k_m \frac{\partial \overline{T}_m}{\partial r} \quad \text{at } r = a \tag{4.5.7}$$

The other necessary boundary conditions are at positions $r \to 0$ and $r \to \infty$. Since no heat is gained or lost at the center of the thermistor,

$$\overline{T}_b = \text{finite}\left(\text{or } k_b \frac{\partial \overline{T}_b}{\partial r} = 0 \right) \quad \text{as } r \to 0 \tag{4.5.8}$$

Because the thermistor power is finite and the tissue is infinite, the tissue temperature rise at infinity goes to zero:

$$\overline{T}_m \to 0 \quad \text{as} \quad r \to \infty \tag{4.5.9}$$

It is this last initial condition that allows the Laplace transform to be used to solve the coupled partial differential equations. The Laplace transform converts the partial differential equations into ordinary differential equations that are independent of time t. The steady-state solution allows for the determination of thermal conductivity and perfusion.[20]

$$\bar{T}_b(r) = \frac{A}{4\partial a k_b}\left\{\frac{k_b}{k_m\left(1+\sqrt{z}\right)} + \frac{1}{2}\left[1-\left(\frac{r}{a}\right)^2\right]\right\} \tag{4.5.10}$$

$$\bar{T}_m(r) = \frac{A}{4\partial r k_m}\left(\frac{e^{(1-r/a)\sqrt{z}}}{1+\sqrt{z}}\right) \tag{4.5.11}$$

where z is a dimensionless Pennes's model perfusion term ($wc_{bl}a^2/k_m$). The measured thermistor response, ΔT, is assumed to be the simple volume average of the thermistor temperature:

$$\Delta T = \frac{\int_0^a \bar{T}_b(r) 4\pi r^2 \, dr}{\frac{4}{3}\pi a^3} \tag{4.5.12}$$

Inserting Equation 4.5.10 into Equation 4.5.12 yields the relationship used to measure thermal conductivity, assuming no perfusion.[4]

$$k_m = \frac{1}{\frac{4\pi a \Delta T}{A} - \frac{0.2}{k_b}} \tag{4.5.13}$$

A similar equation allows the measurement of thermal diffusivity, α, from the transient response, again assuming no perfusion.[22]

$$\alpha_m = \left[\frac{a}{\sqrt{\pi}\, B/A\left(1+0.2\frac{k_m}{k_b}\right)}\right]^2 \tag{4.5.14}$$

Rather than using the actual probe radius (a) and probe thermal conductivity (k_b), the following empirical equations are used to calculate thermal properties:

$$k_m = \frac{1}{\frac{c_1 \Delta T}{A} + c_2} \tag{4.5.15}$$

$$\alpha_m = \left[\frac{c_3}{\sqrt{\pi}\, B/A\left(1+0.2\frac{k_m}{c_4}\right)}\right]^2 \tag{4.5.16}$$

The coefficients c_1, c_2, c_3, and c_4 are determined by operating the probe in two materials of known thermal properties. Typically, agar-gelled water and glycerol are used as thermal standards. This empirical calibration is performed at the same temperatures at which the thermal property measurements will be performed.

It is assumed that the baseline tissue temperature, T_0, is constant during the 30-s transient. Patel has shown that if the temperature drift, dT_0/dt, is larger than 0.1°C/min, then significant errors will occur.[8] The electronic feedback circuit forces T_h to a constant. Thus, if T_0 is constant, then ΔT does not vary during the 30-s transient period.

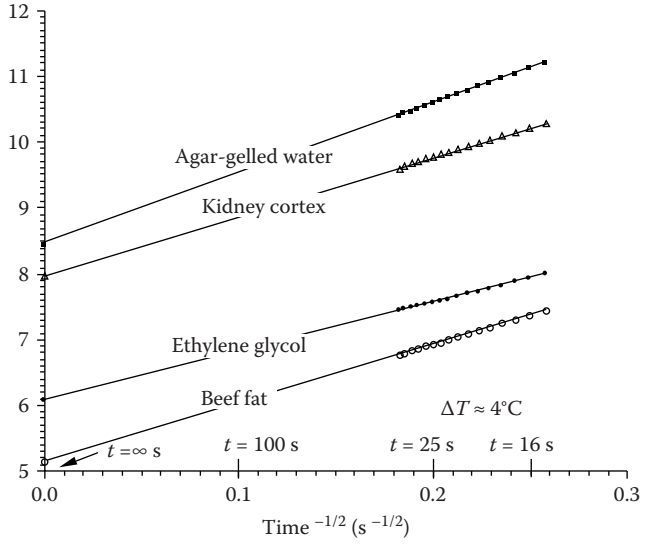


FIGURE 4.5.1 Typical $P/\Delta T$ vs. time$^{-1/2}$ data for the constant temperature heating technique. (From Valvano, J.W., *Optical-Thermal Response of Laser-Irradiated Tissue*, Plenum Press, New York, 1995. With permission.)

The applied power, $P(t)$, varies during the 30-s transient. Linear regression is used to calculate the steady-state and transient terms in Equation 4.5.1. Figure 4.5.1 shows some typical responses. The steady-state response (time equals infinity) is a measure of the thermal conductivity. The transient response (slope) indicates the thermal diffusivity.

The time of heating can vary from 10 to 60 s. Shorter heating times are better for small tissue samples and for situations where there is baseline tissue temperature drift. Another advantage of shorter heating times is the reduction in the total time required to make one measurement. Longer heating times increase the measurement volume and reduce the effect of imperfect thermistor/tissue coupling. Typically, shorter heating times are used *in vivo* because it allows more measurements to be taken over the same time period. On the other hand, longer heating times are used *in vitro* because accuracy is more important than measurement speed.

Probe Design

Thermal probes must be constructed in order to measure thermal properties. The two important factors for the thermal probe are thermal contact and transducer sensitivity. The shape of the probe should be chosen to minimize trauma during insertion. Any boundary layer between the thermistor and the tissue of interest will cause a significant measurement error. The second factor is transducer sensitivity, that is, the slope of the thermistor voltage vs. tissue thermal conductivity. Equation 4.5.13 shows that for a fixed ΔT, k_m, and k_b, the thermistor power (A) increases linearly with probe size (a). Therefore, larger probes are more sensitive to thermal conductivity. Thermometrics P60DA102M and Fenwal 121-102EAJQ01 are glass probe thermistors that make excellent transducers. The glass-coated spherical probes provide a large bead size and a rugged, stable transducer. Thermometrics BR55KA102M and Fenwal 112-102EAJ-B01 bead thermistors also provide excellent results. For large tissue samples multiple thermistors can be wired in parallel, so they act electrically and thermally as one large device. There are two advantages to using multiple thermistors. The effective radius, $a=c_1/4\pi$, is increased from about 0.1 cm for a typical single P60DA102M probe to about 0.5 cm for a configuration of three P60DA102M thermistors. The second advantage is that the three thermistors are close enough to each other that the tissue between the probes will be heated by all three thermistors. This cooperative heating tends to increase the effective measurement volume and

reduce the probe/tissue contact error. Good mechanical/thermal contact is critical. The probes are calibrated after they are constructed, so that the thermistor geometry is incorporated into the coefficients c_1, c_2, c_3, and c_4. The same water bath and probe configuration should be used during the calibration and tissue measurements.

Calibration

Calibration is a critical factor when using an empirical technique. For temperatures below 0°C, ice and ethylene glycol are used as thermal standards. For temperatures between 0°C and 15°C, agar-gelled water and ethylene glycol can be used as thermal standards. For temperatures between 15°C and 75°C, agar-gelled water and glycerol can be used; 1 g of agar per 100 mL of water should be added to reduce convection. The instrument has been used to measure k_m and α_m of various concentrations of agar-gelled water at 37°C. A mixture of water and glycerol can be used to estimate the accuracy of the technique. The mass fraction, m, can be used to determine the true thermal properties of the mixture where k_w and k_g are in mW/(m K):[29]

$$k_m = m\,k_g + (1-m)k_w + 1.4m(m-1)(k_w - k_g - 2) - 0.014m(m-1)(T - 20°C) \tag{4.5.17}$$

$$\alpha_m = m\alpha_g + (1-m)\alpha_w \tag{4.5.18}$$

Estimation of the Convective Heat Transfer Coefficient

Introduction

The heart is an organ in which it is possible to make experimental measurements of convective transport between flowing blood and tissue. The topologies of the inner surfaces of the heart and blood vessels are very complex. Consequently, it is very difficult to model the heat transfer at the wall analytically. Any attempt to estimate the heat transfer coefficient analytically will be inaccurate because of the complex contours of structures like the endocardium. The experimental method presented in this section to estimate the heat transfer coefficient is very appropriate.

The following method can be used to estimate the convective coefficient due to blood flow inside the chambers of the heart. This method can be applied to blood vessels as well. The method to measure the convective coefficient due to blood flow inside the heart is derived from basic heat transfer analysis. Two miniature thermistor (BR11, Thermometrics Inc, Northridge, CA.) probes are inserted at fixed locations near the inner surface of the model. T_1 is the measured temperature nearest the endocardial wall, and T_2 is measured 2 mm inside the myocardial wall. The epicardial surface of the heart is maintained at a steady temperature, T_o. Blood at a different temperature T_f is pumped through the heart. The thermistor probes are used to continuously monitor the spatial temperature gradient in the heart muscle. The basic principle used to estimate the heat transfer coefficient, h, is as follows:

$$h \cdot \left(T_{wall} - T_f\right) = k_s \cdot \left.\frac{dT}{d\hat{n}}\right|_{wall} \tag{4.5.19}$$

where $\hat{n}$ is the direction perpendicular to the surface and k_s is the thermal conductivity of the heart wall. The gradient at the wall of the surface is calculated from the data obtained from the array of thermistors.

Calibration

Ideally, the temperature measured in any experiment should be the temperature of the zero-volume point of interest. But when a measurement is made the temperature sensor has a finite volume and the actual measurement is the volume-averaged measurement of some region around the sensor. Similarly, the gradient measured using two sensors is an estimate of the gradient between the two

sensors. These two errors are responsible for the estimate of the heat transfer coefficient measured using Equation 4.5.19 to be vastly different from the correct value. Hence, the probe assembly is first calibrated by operating the probe in situations where the true convective heat transfer coefficient is known. Two calibration factors, c_1 and c_2, are added to Equation 4.5.19 to adjust for the finite size of the temperature sensors, and for the fact that the measured temperature gradient is not at the wall surface:

$$h = c_1 \exp\left(c_2 \frac{k_s (T_1 - T_2)}{(T_f - T_1)} \right) \tag{4.5.20}$$

where T_1 and T_2 are measured temperatures in the wall. A cylindrical tube made with a material (e.g., silicone) that has thermal properties close to tissue can be used to calibrate the sensor combination. The key to accurate measurements lies in how close the calibration configuration matches the actual experimental conditions.

Verification

The effect of measurement errors on the estimate of the heat transfer coefficient was studied for a tube of inner radius R_i (0.73 cm) and outer radius of R_o (1.33 cm). The temperature distribution inside the tube wall was analytically determined for water flowing inside and the outer wall temperature held at a constant value of 23°C. The heat transfer coefficient was then calculated from the temperature distribution using the relation in Equation 4.5.19. The effect of the measurement errors on the heat transfer coefficient estimate was analyzed by using values for the temperature and the gradient away from the wall of the tube.

The temperature at any point of distance R from the axis of the tube is given by

$$T = T_o - \mathrm{Bi}(T_f - T_w) \ln\left(\frac{R}{R_o} \right), \tag{4.5.21}$$

where the T_w is the inner wall temperature given by

$$T_w = \frac{T_o - \mathrm{Bi} \ln\left(\frac{R_i}{R_o} \right) T_f}{1 - \mathrm{Bi} \ln\left(\frac{R_i}{R_o} \right)} \tag{4.5.22}$$

where T_o is the outer wall temperature, Bi is the Biot number, calculated as (hR_i/k), h is the heat transfer coefficient, and k is the thermal conductivity of the material of the tube.

In Vitro Studies

The first *in vitro* study used two cylindrical tubes. The first tube was used to calibrate the probe and the second to evaluate measurement accuracy. The exact position of the thermistor array for sensing the temperature must be in a similar location relative to the inner wall for both calibration and measurement. Water at 37°C was made to flow through the tube at a rate of L (liters per minute). The flow was measured using a rotameter type flowmeter (accuracy 2%, Omega Engineering Inc, Norwalk, CT.). The temperature gradients were recorded continuously.

The tube was initially maintained at room temperature. For a smooth tube of cylindrical cross section the relation between Nu, Re, and Pr for turbulent flow conditions is given by

$$\mathrm{Nu}_D = \frac{(f/8)(\mathrm{Re}_D - 1000)\mathrm{Pr}}{1 + 12.7(f/8)^{1/2}(\mathrm{Pr}^{2/3} - 1)} \tag{4.5.23}$$

where the friction factor, f, is

$$f = (0.79 \ln\{\mathrm{Re_D}\} - 1.64) - 2 \tag{4.5.24}$$

This correlation is valid for $0.5 < \mathrm{Pr} < 2000$ and $2300 < \mathrm{Re_D} < 5 \times 10^6$. The setup was calibrated with a tube of 1.46 cm diameter and then tested with a tube of 1.27 cm diameter. The flow rates were selected to give a range of h values from 800 to 4000 W/m^2 K. The average accuracy is about 10%.

Temperature-Dependent Thermal Properties

Temperature Dependence of Organ Tissue

When modeling heat transfer in situations where the temperature change exceeds 10°C it is important to consider the temperature dependence of the tissue thermal properties. Valvano measured tissue thermal properties as a function of temperature using the constant ΔT thermistor heating technique.[11,12] The results, shown in Table 4.5.1, were derived from *in vitro* measurements taken at 3°C, 10°C, 17°C, 23°C, 30°C, 37°C, and 45°C.

TABLE 4.5.1
Thermal Properties as a Function of Temperature

Tissue	Species	k_0 mW/(cm °C)	k_1 mW/(cm °C^2)	α_0 (cm^2/s)	α_1 (cm^2/s °C)
Adenocarcinoma of the breast	Human	4.194	0.03911	0.001617	–0.000049
Cerebral cortex	Human	5.043	0.00296	0.001283	0.000050
Colon cancer	Human	5.450	(at 19°C)	0.001349	(at 19°C)
Fat of spleen	Human	3.431	–0.00254	0.001321	–0.000002
Liver	Human	4.692	0.01161	0.001279	0.000036
Liver	Pig	4.981	0.00800	0.001240	0.000053
Liver	Rabbit	4.668	0.02601	0.001370	0.000178
Lung	Human	3.080	0.02395	0.001071	0.000082
Lung	Human	4.071	0.01176	0.001192	0.000031
Lung	Pig	2.339	0.02216	0.000695	0.000080
Myocardium	Dog	4.869	0.01332	0.001296	0.000058
Myocardium	Human	4.925	0.01195	0.001289	0.000050
Myocardium	Pig	4.841	0.01333	0.001270	0.000051
Pancreas	Dog	4.790	0.00849	0.001287	0.000062
Pancreas	Human	4.365	0.02844	0.001391	0.000084
Pancreas	Pig	4.700	0.00194	0.001530	0.000130
Renal cortex	Dog	4.905	0.01280	0.001333	0.000039
Renal cortex	Human	4.989	0.01288	0.001266	0.000055
Renal cortex	Pig	4.967	0.01176	0.001284	0.000039
Renal cortex	Rabbit	4.945	0.01345	0.001311	0.000027
Renal medulla	Dog	5.065	0.01298	0.001305	0.000063
Renal medulla	Human	4.994	0.01102	0.001278	0.000055
Renal pelvis	Dog	4.930	0.01055	0.001334	0.000052
Renal pelvis	Human	4.795	0.01923	0.001329	0.000011
Spleen	Human	4.913	0.01300	0.001270	0.000047
Spleen	Rabbit	4.863	0.01267	0.001257	0.000042

Source: Valvano, J.W., *Optical-Thermal Response of Laser-Irradiated Tissue*, Plenum Press, New York, 1995. With permission.

The animal tissues were measured from freshly sacrificed dogs, rabbits, and pigs. The normal human tissues were obtained by autopsy. The human cancers were freshly excised. The k_0, k_1, α_0, and α_1 values are the linear fit of the thermal properties as a function of temperature.

$$k = k_0 + k_1 T \tag{4.5.25}$$

$$\alpha = \alpha_0 + \alpha_1 T \tag{4.5.26}$$

The average thermal properties of these data are

$$k = 4.574 + 0.01403T \tag{4.5.27}$$

$$\alpha = 0.001284 + 0.000053T \tag{4.5.28}$$

where conductivity is in mW/(cm °C), diffusivity is in cm²/s, and temperature is in °C.

Temperature Dependence of Human Arterial Tissue

Aortic tissue was obtained from a local pathology lab. The thermal probes were placed on the endothelial surface of the aortic wall, and the tissue/probe combination was wrapped in plastic. The tissue surface was kept wet to improve the thermal contact and to prevent drying. The samples were placed in a beaker of saline and the beaker was put into a temperature-controlled water bath. Thermal conductivity and thermal diffusivity were measured 10 times at each temperature 35°C, 55°C, 75°C, and 90°C. The measurement sequence was varied between 35, 55, 75, 90, 95, 75, 55, 35 75, 90, 55, 35 and 55, 35, 90, 75. Measurements were obtained from both normal and diseased tissue. The plaques were categorized by gross visual observation. The calcified plaques were hard and bony. The fibrous plaques were firm but pliable. The fatty plaques were loose and buttery. The results from 54 tissues are presented in Tables 4.5.2 and 4.5.3. The column n refers to the number of tissue samples. The standard deviation is given in the parentheses.

The two-sample *t*-test with $p = 0.05$ was used to determine significant differences. The tissue thermal properties increased with temperature and were significantly less than those of water. The

TABLE 4.5.2
Thermal Conductivity mW/(cm °C) of Human Aorta and Atherosclerotic Plaque

Tissue	*n*	at 35°C	at 55°C	at 75°C	at 90°C
Normal aorta	12	4.76 (0.41)	5.03 (0.60)	5.59 (0.37)	6.12 (0.12)
Fatty plaque	13	4.84 (0.44)	4.97 (0.49)	5.46 (0.54)	5.88 (0.81)
Fibrous plaque	12	4.85 (0.22)	5.07 (0.30)	5.38 (0.38)	5.77 (0.56)
Calcified plaque	17	5.02 (0.59)	5.26 (0.73)	5.81 (0.82)	6.19 (0.85)

Source: Valvano, J.W., *Optical-Thermal Response of Laser-Irradiated Tissue*, Plenum Press, New York, 1995. With permission.

TABLE 4.5.3
Thermal Diffusivity (×1000 cm²/s) of Human Aorta and Atherosclerotic Plaque

Tissue	*n*	at 35°C	at 55°C	at 75°C	at 90°C
Normal aorta	12	1.27 (0.07)	1.33 (0.11)	1.44 (0.10)	1.56 (0.05)
Fatty plaque	13	1.28 (0.05)	1.32 (0.06)	1.41 (0.11)	1.46 (0.15)
Fibrous plaque	12	1.29 (0.03)	1.36 (0.07)	1.41 (0.10)	1.52 (0.20)
Calcified plaque	17	1.32 (0.07)	1.37 (0.12)	1.53 (0.17)	1.66 (0.20)

Source: Valvano, J.W., *Optical-Thermal Response of Laser-Irradiated Tissue*, Plenum Press, New York, 1995. With permission.

measurement order did not affect the measured thermal properties. There was no significant difference between the thermal conductivity of normal aorta, fatty plaque, and fibrous plaque. The thermal conductivity and thermal diffusivity of calcified plaque were slightly higher than normal aorta, fatty plaque, and fibrous plaque.

Temperature Dependence of Canine Arterial Tissue

Carotid and femoral arteries were harvested immediately postmortem. The thermal probes were placed on the endothelial surface of the arterial wall. Thermal conductivity and thermal diffusivity were measured 10 times at each temperature 25°C, 35°C, 45°C, 55°C, 65°C, 75°C, 85°C, and 95°C. Measurements were obtained only from normal tissue. The results from 18 tissue samples are summarized as follows:

Canine Femoral Artery

$$k\text{mW/(cm °C)} = 3.688 + 0.0062014T(°\text{C}) \tag{4.5.29}$$

$$\alpha(\text{cm}^2/\text{s}) = 0.001003 + 0.000001381T(°\text{C}) \tag{4.5.30}$$

Canine Carotid Artery

$$k\text{mW/(cm °C)} = 4.480 + 0.0000164T(°\text{C}) \tag{4.5.31}$$

$$\alpha(\text{cm}^2/\text{s}) = 0.001159 + 0.000003896T(°\text{C}) \tag{4.5.32}$$

The two-sample *t*-test with $p = 0.01$ shows that both thermal conductivity and thermal diffusivity are higher in carotid vs. femoral artery. These results could be explained from the fact that the carotid artery contains more collagen than femoral artery. A tissue with a higher percentage of collagen would have lower thermal properties because collagen is a thermal insulator.

Temperature Dependence of Swine Left Ventricle

Swine myocardial samples were harvested immediately postmortem. The thermal probes were placed on the left ventricular muscle. Thermal conductivity and thermal diffusivity were measured 10 times at each temperature 25°C, 37°C, 50°C, 62°C, and 76°C. Measurements were obtained only from normal tissue. The results are summarized in Tables 4.5.4 and 4.5.5.

Thermal Properties of Frozen Tissue

The thermal properties of frozen tissue are significantly higher than normal tissue. Valvano measured frozen tissue thermal properties using the constant ΔT thermistor heating technique.[30] The results, shown in Table 4.5.6, were derived from *in vitro* measurements taken at –18°C, –5°C, and +0.1°C.

Thermal Properties as a Function of Water and Fat Content

In a global sense, the thermal properties of tissue are determined by the relative concentrations of its constituent parts. Spells found a linear relationship between tissue thermal conductivity and water content:[31]

$$k\text{mW/(cm °C)} = 0.54 + 5.73m_{\text{water}} \text{ for } m_{\text{water}} > 0.2 \tag{4.5.33}$$

where m_{water} is the mass fraction of water in the tissue. Cooper and Trezek found an empirical relationship between thermal conductivity and mass fractions of water, protein, and fat:[32]

$$k\text{mW/(cm °C)} = \rho\sum_n \frac{k_n m_n}{\rho_n} = \rho\left(6.28m_{\text{water}} + 1.17m_{\text{protein}} + 2.31m_{\text{fat}}\right) \tag{4.5.34}$$

TABLE 4.5.4
Thermal Conductivity mW/(cm·K) of Myocardial Tissue

Temperature	25°C	37°C	50°C	62°C	76°C
	5.23	5.14	5.17	4.39	5.24
	5.07	5.12	4.75	3.30	4.29
	5.30	5.21	5.61	5.67	4.83
	5.43	5.54	4.22	4.16	5.89
	4.68	5.35	4.93	5.33	5.23
	5.25	5.08	4.84	5.70	5.39
	5.27	5.48	4.42	5.11	4.75
	5.28	4.57	4.93	4.99	3.25
	5.86	5.76	5.52	5.03	2.69
	4.78	5.10	5.88	5.30	5.28
	4.75	5.35	5.35	4.67	5.60
	4.92	6.02	5.60	5.49	4.68
Mean	5.15	5.31	5.1	4.93	4.76
Standard deviation	0.33	0.37	0.51	0.70	0.95

TABLE 4.5.5
Thermal Diffusivity (cm²/s) of Myocardial Tissue

Temperature	25°C	37°C	50°C	62°C	76°C
	0.00151	0.00170	0.00165	0.00159	0.00167
	0.00154	0.00147	0.00203	0.00235	0.00249
	0.00143	0.00165	0.00151	0.00169	0.00166
	0.00146	0.00143	0.00116	0.00191	0.00229
	0.00159	0.00160	0.00176	0.00167	0.00173
	0.00141	0.00178	0.00179	0.00163	0.00185
	0.00165	0.00149	0.00235	0.00143	0.00185
	0.00132	0.00206	0.00179	0.00170	0.00199
	0.00141	0.00144	0.00147	0.00143	0.00062
	0.00168	0.00179	0.00160	0.00180	0.00167
	0.00154	0.00156	0.00173	0.00161	0.00173
	0.00164	0.00138	0.00171	0.00169	0.00192
Mean	0.00152	0.00161	0.00171	0.00171	0.00179
Standard deviation	0.00012	0.00020	0.00031	0.00025	0.00047

Cooper and Trezek[32] found similar relationships for specific heat and density:

$$c\text{J/(g °C)} = \sum_n c_n m_n = 4.2 m_{\text{water}} + 1.09 m_{\text{protein}} + 2.3 m_{\text{fat}} \tag{4.5.35}$$

$$\rho(\text{g/cm}^3) = \frac{1}{\sum_n \frac{m_n}{\rho_n}} = \frac{1}{m_{\text{water}} + 0.649 m_{\text{protein}} + 1.227 m_{\text{fat}}} \tag{4.5.36}$$

TABLE 4.5.6
Average Thermal Properties

T (°C)	*N*	*M*	Species Tissue	*k*m mW/(cm °C)	αm (1000*cm²/s)
+0.1	45	6	Bovine kidney cortex	4.54 (±0.16)	1.18 (±0.09)
–5	15	4	Bovine kidney cortex	15.35 (±1.09)	4.71 (±0.99)
–18	18	3	Bovine kidney cortex	13.72 (±0.73)	6.84 (±0.83)
+0.1	66	9	Bovine liver	4.17 (±0.13)	1.05 (±0.09)
–5	66	9	Bovine liver	13.96 (±2.49)	4.77 (±0.58)
–18	56	8	Bovine liver	9.89 (±0.44)	5.71 (±0.74)
+0.1	48	6	Bovine muscle	4.25 (±0.37)	1.05 (±0.11)
–5	42	7	Bovine muscle	13.93 (±1.23)	5.37 (±0.97)
–18	60	8	Bovine muscle	10.76 (±1.14)	6.84 (±1.10)
+0.1	21	3	Bovine fat	1.93 (±0.12)	0.59 (±0.13)
–5	32	4	Bovine fat	2.66 (±0.38)	0.98 (±0.19)
–18	24	4	Bovine fat	2.80 (±0.53)	1.54 (±0.57)

Source: Valvano, J.W., *Optical-Thermal Response of Laser-Irradiated Tissue*, Plenum Press, New York, 1995. With permission.

Note: *N* is the number of measurements and *M* is the number of tissues. (±0.xx) is the standard deviation of the average.

Effect of Blood Flow on Temperature

Bioheat transfer processes in living tissues are often influenced by blood perfusion through the vascular network. When there is a significant difference between the temperature of blood and the tissue through which it flows, convective heat transport will occur, altering the temperatures of both the blood and the tissue. Perfusion-based heat transfer interaction is critical to a number of physiological processes such as thermoregulation and inflammation.

The blood/tissue thermal interaction is a function of several parameters including the rate of perfusion and the vascular anatomy, which vary widely among the different tissues, organs of the body, and pathology. Appendix B contains an extensive compilation of perfusion rate data for many tissues and organs and for many species. The literature on mathematical modeling of the influence of blood perfusion on bioheat transfer phenomena has been reviewed recently by Charney,[33] and this reference is highly recommended for readers desiring an in-depth presentation of the topic.

The rate of perfusion of blood through different tissues and organs varies over the time course of a normal day's activities, depending on factors such as physical activity, physiological stimulus, and environmental conditions. Further, many disease processes are characterized by alterations in blood perfusion, and some therapeutic interventions result in either an increase or decrease in blood flow in a target tissue. For these reasons, it is very useful in a clinical context to know what the absolute level of blood perfusion is within a given tissue. There are numerous techniques that have been developed for this purpose over the past several decades. In some of these techniques, the coupling between vascular perfusion and local tissue temperature is applied to advantage to assess the flow through local vessels via inverse solution of equations which model the thermal interaction between perfused blood and the surrounding tissue.

Pennes[28] published the seminal work on developing a quantitative basis for describing the thermal interaction between tissue and perfused blood. His work consisted of a series of experiments to measure temperature distribution as a function of radial position in the forearms of nine human subjects. A butt-junction thermocouple was passed completely through the arm *via* a needle inserted as a temporary guideway, with the two leads exiting on opposite sides of the arm. The subjects were unanesthetized so as to avoid the effects of anesthesia on blood perfusion. Following a period of normalization,

the thermocouple was scanned transversely across the mediolateral axis to measure the temperature as a function of radial position within the interior of the arm. The environment in the experimental suite was kept thermally neutral during the experiments. Pennes's data showed a temperature differential of three to four degrees between the skin and the interior of the arm, which he attributed to the effects of metabolic heat generation and heat transfer with arterial blood perfused through the microvasculature.

Pennes proposed a model to describe the effects of metabolism and blood perfusion on the energy balance within tissue. These two effects were incorporated into the standard thermal diffusion equation, which is written in its simplified form as

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot k \nabla T + (\rho c)_b \omega_b (T_a - T) + q_{met} \tag{4.5.37}$$

Metabolic heat generation, q_{met}, is assumed to be homogeneously distributed throughout the tissue of interest as the rate of energy deposition per unit volume. It is assumed that the blood perfusion effect is homogeneous and isotropic and that thermal equilibration occurs in the microcirculatory capillary bed. In this scenario, blood enters capillaries at the temperature of arterial blood, T_a, where heat exchange occurs to bring the temperature to that of the surrounding tissue, T. There is assumed to be no energy transfer either before or after the blood passes through the capillaries, so that the temperature at which it enters the venous circulation is that of the local tissue. The total energy exchange between blood and tissue is directly proportional to the density, ρ_b, specific heat, c_b, and perfusion rate, ω_b, of blood through the tissue, and is described in terms of the change in sensible energy of the blood. This thermal transport model is analogous to the process of mass transport between blood and tissue, which is confined primarily to the capillary bed.

A major advantage of the Pennes model is that the added term to account for perfusion heat transfer is linear in temperature, which facilitates the solution of Equation 4.5.37. Since the publication of this work, the Pennes model has been adapted by many researchers for the analysis of a variety of bioheat transfer phenomena. These applications vary in physiological complexity from a simple homogeneous volume of tissue to thermal regulation of the entire human body.[34,35] As more scientists have evaluated the Pennes model for application in specific physiological systems, it has become increasingly clear that many of the assumptions foundational to the model are not valid. For example, Chato,[36] Chen and Holmes,[37] and Weinbaum et al.[38–50] have demonstrated very convincingly that thermal equilibration between perfused blood and local tissue occurs in the precapillary arterioles, and that by the time blood flows into vessels 60 μm in diameter and smaller, the equilibration process is complete. Therefore, no significant heat transfer occurs in the capillary bed; the exchange of heat occurs in the larger components of the vascular tree. The vascular morphology varies considerably among the various organs of the body, which contributes to the need for specific models for the thermal effects of blood flow (as compared to the Pennes model that incorporates no information concerning vascular geometry). It would appear as a consequence of these physiological realities that the validity of the Pennes model is questionable.

Many investigators have developed alternative models for the exchange of heat between blood and tissue. These models have accounted for the effects of vessel size (Chato,[36] Chen and Holmes,[37] Mooibroek and Lagendijk[51]), countercurrent heat exchange (Baish,[52] Huang et al.,[53] Keller and Seiler,[54] Mitchell and Meyers,[55] Mooibroek and Lagendijk[51]), as well as a combination of partial countercurrent exchange and bleed-off perfusion (Weinbaum and Jiji[38–50]). All of these models provided a larger degree of rigor in the analysis, but at the compromise of greater complexity and reduced generality. Some of these models have been the subject of considerable debate concerning their validity and range of appropriate application (Baish et al.,[56] Weinbaum and Jiji,[57] Wissler[58,59]). These studies also led to an increased appreciation of the necessity for a more explicit understanding of the local vascular morphology as it governs bioheat transfer, which has given rise to experimental studies to measure and characterize the three-dimensional architecture of the vasculature in tissues and organs of interest.

It is quite interesting that, in the context of the above studies to improve on the widely applied but questionable Pennes model, the 50th anniversary of the publication of Pennes's research was

recognized recently.[28] For this occasion, Wissler[59] returned to Pennes's original data and analysis and reevaluated his work. Given the hindsight of 5 decades of advances in bioheat transfer plus greatly improved computational tools and better constitutive property data, Wissler's analysis pointed out further flaws in Pennes's work which had not been appreciated previously. However, he also showed that much of the criticism that has been directed toward the Pennes model is not justified, in that his improved computations with the model demonstrated a good standard of agreement with the experimental data. Thus, Wissler's conclusion is that "those who base their theoretical calculations on the Pennes model can be somewhat more confident that their starting equations are valid." The quantitative analysis of the effects of blood perfusion on the internal temperature distribution in living tissue remains a topic of active research after a half-century of study.

Limitations of Pennes's Model

Pennes's model was the first major effort in quantifying the heat transfer contribution of perfusion. It was developed for describing the transverse temperature profile in the human forearm. The model is unique in that the perfusion is very simple. The "bioheat equation" was previously shown as Equation 4.5.37.

The limitations of this model arise from the erroneous view of the heat transfer process and its anatomical location. Chen and Holmes's analysis of blood vessel thermal equilibration lengths showed that Pennes's concept is incorrect.[37] The thermal equilibration length is defined as the length at which the difference between the blood and the tissue temperature decreases to $1/e$ of the initial value. They indicated that thermal equilibration occurs predominantly within the terminal arterioles and venules, and that blood is essentially equilibrated prior to entering the capillaries. In considering the contribution of perfusion as a nondirectional term, the directional convective mechanism is neglected. Nor does the model account for specific vascular architecture such as countercurrent arteries and veins. The limitations of Pennes's model have motivated subsequent investigators to develop their own models.

Despite its erroneous concept, the perfusion term of Pennes's model has been widely used, and found to be valid for situations other than the forearm. Its wide usage has been mainly due to its simplicity of implementation, especially in analyses where a closed form analytical solution is sought.[22,60] Investigators have obtained good temperature predictions for the following circumstances:

1. The porcine kidney cortex in the absence of large vessels (diameter > 300 μm)[61]
2. The rat liver[5,22]
3. The capillary bleed-off from large vessels[42]

In the last case, Charny compared Weinbaum and Jiji's countercurrent model and Pennes's model against the experimental results of Pennes. The simulations found that Pennes's model is valid in the initial branchings of the largest microvessels from the countercurrent vessels (diameter > 500 μm) in deep tissue. In this case, the microvessel blood temperature is close to arterial temperature. Arkin et al.[62] provide an explanation of the inconsistencies between the two anatomical sites. They suggest that since blood typically travels down successive generations of the vascular branches before equilibrating with the tissue temperature, Charny's claim actually refers to the collective contribution of numerous smaller thermally significant vessels in a region dominated by the large microvessels. The distinction of being thermally significant is based upon the ratio of thermal equilibration length to actual vessel length (ε):

$$\varepsilon = \frac{L_e}{L} \tag{4.5.38}$$

Along with Xu's observations in porcine kidney,[61] Pennes's model appears to be applicable to regions where the vasculature comprises numerous, small, thermally significant vessels ($\varepsilon \approx 1$).

Continuum Models

Among the continuum formulations of bioheat transfer, the Chen–Holmes model is the most developed. Prior to Chen–Holmes model, continuum formulations by Wulff[63] and Klinger[64] addressed the isotropicity of Pennes's perfusion term. However, they did not challenge the Pennes's concept of the site of heat exchange. Chen and Holmes[37] formulated their model after analyzing the significance of blood vessel thermal equilibration length. Through this analysis, they quantitatively showed that the major heat transfer processes occur in the 50- to 500-μm diameter vessels and refuted Pennes's paradigm. In their model, they proposed that larger vessels be modeled separately from smaller vessels and tissue. Larger vessels were distinguished using the ratio of equilibration length to actual vessel length of about one ($\varepsilon \approx 1$) as the criterion. The smaller vessels and tissue were then modeled as a continuum. In a differential control volume of this continuum, they further separated solid tissue from the blood within the vascular space (Figure 4.5.2).

Subsequently, the heat transfer mechanisms can be divided into the contributions from (1) non-equilibrated blood in the thermally significant vessels, (2) blood that has equilibrated with the surrounding tissue, and (3) nearly equilibrated blood. The perfusion term of Pennes is then replaced by three terms:

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot k \nabla T + (\rho c)_b \omega^* (T_a^* - T) - (\rho c)_b \bar{u} \cdot \nabla T + \nabla \cdot k_p \nabla T + q_m \tag{4.5.39}$$

The second through the fourth term on the right-hand side arise from each of the three categories described above. The second term, $(\rho c)_b \omega^* (T_a^* - T)$, is similar to Pennes except the perfusion and the arterial temperature are specific to the volume being considered. The $(\rho c)_b \bar{u} \nabla T$ term is a directional convective term due to the net flux of equilibrated blood. Finally, the $\nabla \cdot k_p \nabla T$ term is to account for the contribution of the nearly equilibrated blood in a tissue temperature gradient. The nearly equilibrated blood contributes to small temperature fluctuations within the tissue and the effect is modeled as a tensor "perfusion conductivity":

$$k_p = n(\rho c)_b \pi r_b^2 \bar{V} \cos^2 \gamma \sum_{i=1}^{\infty} \frac{L_e}{L_e^2 \beta_1^2 + 1} \tag{4.5.40}$$

which is a function of local average blood flow velocity vector within the vessel ($\bar{V}$) relative angle (γ) between blood vessel direction and the tissue temperature gradient, the number of vessels (n), and vessel radius (r_b). The Fourier integral spectral wave number (β) can be approximated as the inverse of vessel length. The contribution of this conductivity is minimal except when vessels with large equilibration lengths are considered. However, for this situation, Chen and Holmes recommend that these vessels be treated separately. The assumptions that were made for their model include (1) neglecting the mass transfer between vessel and tissue space, and (2) treating the thermal

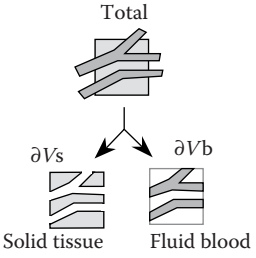


FIGURE 4.5.2 Schematic representation of tissue control volume as used by Chen-Holmes.[37]

conductivity and temperature within the tissue-blood continuum as that of the solid tissue, since the vascular volume is much smaller than that of the solid tissue.

The limitation of this model is that given the detail required, the model is not easy to implement. Also, the perfusion conductivity term is difficult to evaluate, and distinction within the continuum model is not well defined. Furthermore, the model does not explicitly address the effect of closely spaced countercurrent artery–vein pairs. This model has been applied to the porcine kidney and was found to predict temperatures similar to Pennes's model, and thus, given the simplicity of the latter, Xu et al.[61] recommended that Pennes be used. Arkin et al.[62] claim that the Chen–Holmes model can be essentially applied to the same tissue region as that for Pennes.

Vasculature-Based Models

Weinbaum–Jiji–Lemons[38–50]

The modeling of countercurrent vasculature, which was not explicitly addressed by the Chen–Holmes model, developed separately from that of the continuum models.[38–50] Bazett et al.[65] initially presented the countercurrent structure from observations of large arteries and veins in human limbs. The first major quantitative analysis was presented by Mitchell and Myers.[55] It was then followed by the work of Keller and Seiler,[54] which became the predecessor to the Weinbaum–Jiji models. In 1979, Weinbaum and Jiji[43] proposed the initial model of the artery–vein pair as two parallel cylinders of equal diameters with collateral bleed-off in the plane normal to the cylinders. The anatomical configuration is a schematic of an artery and vein pair with branches to the peripheral skin layer (Figure 4.5.3). The contribution of perfusion to heat transfer in tissue was treated as heat transfer in a porous medium. It was considered as a unidirectional convective term normal to the artery–vein pair. Knowledge of vessel density, diameter, and blood velocity was required at the different blood vessel generations.

In 1984, they presented a more thorough model based upon anatomical observations with Lemons.[38,39] This model analyzed three tissue layers of a limb: (1) deep, (2) intermediate, and (3) superficial or cutaneous. For the countercurrent structure of the deep tissue layer, they proposed a system of three coupled equations:

$$(\rho c)_b \pi r_b^2 \overline{V} \cdot \frac{dT_a}{ds} = -q_a \tag{4.5.41}$$

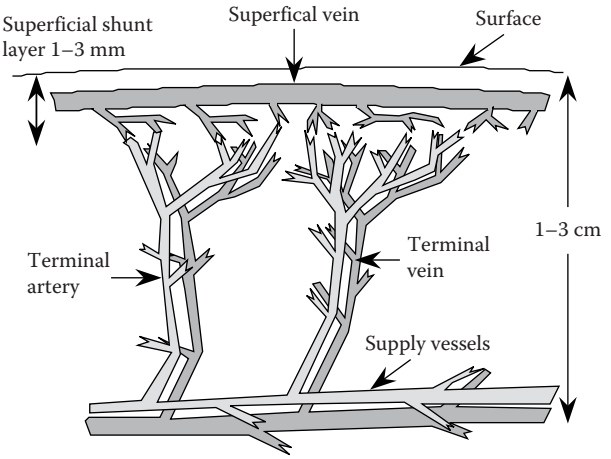


FIGURE 4.5.3 Schematic of artery and vein pair in peripheral skin layer.[43]

$$(\rho c)_b \pi r_b^2 \bar{V} \cdot \frac{dT_v}{ds} = -q_v \tag{4.5.42}$$

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot k \nabla T + n\, g (\rho c)_b \cdot (T_a - T_v) - n \pi r_b^2 (\rho c)_b \bar{V} \cdot \frac{d(T_a - T_v)}{ds} + q_m \tag{4.5.43}$$

The first two equations describe the heat transfer of the thermally significant artery and vein, respectively. The third equation refers to the tissue surrounding the artery–vein pair. For this equation, the middle two right-hand-side terms represent the capillary bleed-off energy exchange, and the net heat exchange between the tissue and artery–vein pair, respectively. The capillary bleed-off term is similar to Pennes's perfusion term except the bleed-off mass flow (g) is used. Their analysis showed that the major heat transfer is due to the imperfect countercurrent heat exchange between artery–vein pairs. They quantified the effect of perfusion bleed-off associated with this vascular structure, and showed that Pennes's perfusion formulation is negligible due to the temperature differential.

Assumptions Include the Following:

1. Neglecting the lymphatic fluid loss, so that the mass flow rate in the artery is equal to that of the vein.
2. Spatially uniform bleed-off perfusion.
3. Heat transfer in the plane normal to the artery–vein pair is greater than that along the vessels (in order to apply the approximation of superposition of a line sink and source in a pure conduction field).
4. A linear relationship for the temperature along the radial direction in the plane normal to the artery and vein.
5. The artery–vein border temperature equals the mean of the artery and vein temperature.
6. The blood exiting the bleed-off capillaries and entering the veins is at the venous blood temperature.

The last assumption has drawn criticism based on studies that indicate the temperature to be closer to tissue.[58,62] Limitations of this model include the difficulty of implementation, and that the artery and vein diameters must be identical. Both of these issues have led to the development of the models described in subsequent sections. Studies using this model have been applied to the peripheral muscle tissue of a limb,[44–46] and the model is accepted as valid for vasculature with diameters <300 μm and $\varepsilon < 0.3$.[62]

Simplified Weinbaum–Jiji[40]

In response to the criticism that their previous model is too difficult and complex to apply, Weinbaum and Jiji simplified the three-equation model to a single equation:[40]

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k_{eff} \frac{\partial T}{\partial x} \right) + q_{met} \tag{4.5.44}$$

In their simplification, they derived an equation based on the temperature of tissue only. The imperfect countercurrent heat exchange is embodied in an effective conductivity tensor term.

$$k_{eff} = k \left\{ 1 + \frac{n\left[(\rho c)_b \pi r_b^2 \bar{V} \cos\gamma \right]^2}{\sigma_\Delta \cdot k^2} \right\} + q_{met} \tag{4.5.45}$$

The k_{eff} term has similar parameters to the tissue and artery–vein pair heat exchange term in Equation 4.5.43, and a shape factor term (σ_Δ). In order to eliminate the blood temperature from their previous formulation, two major assumptions (the closure conditions) were used:

1. The mean tissue temperature $=(T_a+T_v)/2$.
2. Heat from a paired artery is mostly conducted to the corresponding vein:

$$q_a \approx q_v \approx \sigma_\Delta k \left(T_a - T_v\right) \tag{4.5.46}$$

Both of these assumptions were based upon studies in rabbit thigh muscle from their previous formulation. However, to respond to criticism of these assumptions, they performed further mathematical analysis on ε and provided insights into the limits for applying these assumptions.[47] An obvious limitation of this model is that the local temperatures along the countercurrent artery and vein cannot be calculated. Another limitation is that the model is applicable only in situations where $L_e / L \ll 1$. In the example of analyzing the peripheral tissue in the arm, L is equal to the characteristic radius of the arm.[48] Weinbaum and Lemons[48] admit that this assumption breaks down under the following conditions: (1) if blood flow rates significantly increased in the larger vessel pairs of the peripheral tissue layer, and (2) if deeper muscle tissue, where the diameter of the countercurrent pair vessels are less than 300 μm, are included. This model has been tested in the porcine[61] and canine kidney,[66] and continues to be verified by the Weinbaum group[41,50] and other investigators.[67]

Small Artery Model[68,69]

The small artery model was developed by Anderson in studies of the canine kidney cortex.[68,69] The model considers the energy balance in a control volume (i, j, k) which contains either an arterial (Q_a) or venous (Q_v) vessel. For a volume with an artery parallel to the "z" coordinate axis, the equation is

$$Q_a = N(VA)_a(\rho c)_b(1+\lambda-2\lambda\xi)\left(\frac{T_z - T_{z-\Delta z}}{\Delta z}\right) \tag{4.5.47}$$

For a volume with a vein,

$$Q_v = M(VA)_v(\rho c)_b(1+\lambda-2\lambda\xi)\left(\frac{T_{z+\Delta z} - T_z}{\Delta z}\right) \tag{4.5.48}$$

where N and M are the densities of the interlobular arteries and veins in the kidney cortex, respectively; ξ refers to the fraction of the total interlobular artery flow within the control volume; in the kidney cortex, $\xi=1$ at the corticomedullary junction and decreases to $\xi=0$ at the outer capsule. The total flow within this region accounts for bleed-off from the interlobular arteries through the λ term, where $\lambda=1$ represents complete bleed-off.

The discrete representation provides for straightforward numerical implementation when the vessel density within the tissue region of interest is known. Model assumptions include the following:

1. Thermal equilibration length within the volume is much less than vessel length.
2. A linear relation between the effect of bleed-off on arterial flow and location along the length of the vessel.
3. Bleed-off is modeled as a change in the arterial flow.
4. Bleed-off heat transfer is negligible.
5. No major thermally significant vessels (i.e., $\varepsilon \ll 1$) in the region of interest.

In its initial formulation, arbitrarily oriented vessels which would cause more than one vessel to occupy a control volume were not considered. Branching vessels are also not addressed. Even though this is not an inherent limitation of the model, implementation would be more difficult.

Due to its discrete representation, the model cannot solve the inverse problem. The model has been shown to be valid in the canine kidney cortex where there is uniformly oriented countercurrent artery–vein architecture of 70-μm diameter.

Hybrid Models

The lack of an encompassing model which can account for the various tissue structures has led researchers to propose and apply hybrid models. The substantial amount of effort related to the application and investigation of the major models discussed has shown that no one model applies to all the different vascular structures in tissue.[58,62] Wissler[58,59] points to the unlikelihood of a single equation being able to provide a complete description of the heat transfer process in tissue, and thus suggests the use of a combination of equations. The realization of this suggestion is found in Charny's work,[42] in which W-J's and Pennes's models are used to describe peripheral and deep muscle tissue, respectively. When the applicability of each of the major models has been conclusively defined, an algorithm which would enable users to arrive at the appropriate choice of equations given the tissue vasculature of interest would be beneficial.

The applicability of the models discussed requires an understanding of the validity and development of each model. A comparison of the models discussed has been summarized by Charney.[33]

Thermal Measurements of Perfusion

Introduction

Perfusion, the transmission of blood in the microcirculation, is an important factor in surgery, tissue transplants, heart disease, and cancer therapy. Despite its importance, no clinical method of measuring perfusion is currently available for a majority of applications. One technique that shows considerable promise involves the use of self-heated thermistors.[6–13] In this method, a miniature thermistor (0.5–2.5 mm in diameter) is placed invasively in the tissue of interest and heated with a predetermined applied power. Since both tissue conduction and perfusion act to carry heat away from the thermistor, the resulting volumetric-average temperature rise in the thermistor bead, ΔT, is related to both the tissue thermal conductivity and perfusion. By knowing the intrinsic tissue conductivity and the apparent conductivity of the tissue (due to both blood flow and conduction), the perfusion rate can be calculated.

At least two difficulties exist with this technique. The first is that the intrinsic tissue conductivity of perfused tissue must be known in order to calculate the perfusion rate. Although one could stop the blood flow to a tissue and measure its conductivity, this is clearly not desirable, nor is it always practical. In order to overcome this problem, Holmes and Chen have proposed techniques that measure perfusion without interrupting blood flow.[13,20]

Perfusion Resolution

The perfusion resolution, δw, is defined as the smallest change in perfusion that can be detected by the instrument. It can be determined theoretically:

$$\delta w = \delta k \cdot \frac{\partial w}{\partial k} \tag{4.5.49}$$

For the constant temperature heating technique using a P60 thermistor, δk is about 0.02 mW/(cm °C), and the sensitivity is about 100 (mL/100 g min)/(mW/cm °C). Using Equation 4.5.49 gives a perfusion resolution of about 2 mL/100 g min. Due to fluctuations in the baseline tissue temperature, the practical resolution is about 10 mL/100 g min. There are so many experimental and tissue variables that it is extremely important to test these techniques in preparations where the perfusion is known.

Measurement Volume

The measurement volume of a thermistor is a complex function of many factors, including the perfusion rate and vascular anatomy of the tissue of interest. One problem with small thermistors is their small measurement volume. A boundary layer (decoupler) between the thermistor and the tissue causes a significant measurement error. This unwanted boundary layer is often caused by the probe itself during insertion. The larger probes exhibit a smaller error, but are likely to cause a larger decoupler because of the increased trauma during insertion.

Temperature Dependence of Perfusion

Perfusion depends on a wide variety of factors, some local to the tissue (pH, temperature, pO_2), some which are external but directly control local flow (parasympathetic, hormones), and some which indirectly affect local flow (heart rate, blood pressure, skin temperature, needs of other organs). A simple experiment studied the effect of local tissue temperature on muscle perfusion. The constant ΔT method was used to measure perfusion in an anesthetized rat. The muscle temperature was manipulated by placing the hind limb into a water bath. Insulation was carefully placed so as to minimize changes to the body temperature. The rectal and neck temperatures were constant while the muscle in the hind limb was heated. Figure 4.5.4 shows the perfusion as a function of tissue temperature for a typical experiment. The dip in perfusion as a function of temperature, as shown in Figure 4.5.4, consistently occurred, but did not always occur at the same temperature. This dip may be due to an anastomotic shunt attempting to regulate the core body temperature.

Figure 4.5.5 presents the averaged results for 10 rats. Because the dip occurred at different temperatures for the various rats, it does not appear in the average. The large standard deviations are due to perfusion differences from one rat to the next. A linear fit to this averaged data gives the following approximation:

$$w = 1.9720(1 + 0.059T) \tag{4.5.50}$$

where w has units of mL/(100 g min) and T has units of °C.

Yuan et al.[70] measured perfusion and temperatures at various locations within each of the four canine prostates subjected to a transurethral microwave (MW) thermal source. The total number of the perfusion sampling points coupled with temperature is 15. Colored microspheres were used to measure perfusion due to their simplicity compared with radioactively labeled microspheres, and because the microsphere trapping method is regarded as a standard. Temperatures were measured using miniature thermistors. The prostate temperatures were raised to 40°C–45°C by 5-W-step increments of the MW power at hourly intervals to 15 W. Temperatures and perfusion were measured at baseline, and at the beginning and end of each heating interval. Thus, the periods between

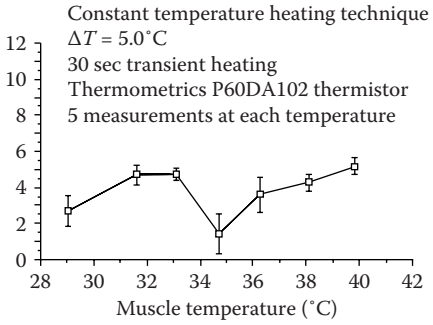


FIGURE 4.5.4 Perfusion vs. muscle temperature during a typical experiment. (From Valvano, J.W., *Optical-Thermal Response of Laser-Irradiated Tissue*, Plenum Press, New York, 1995. With permission.)

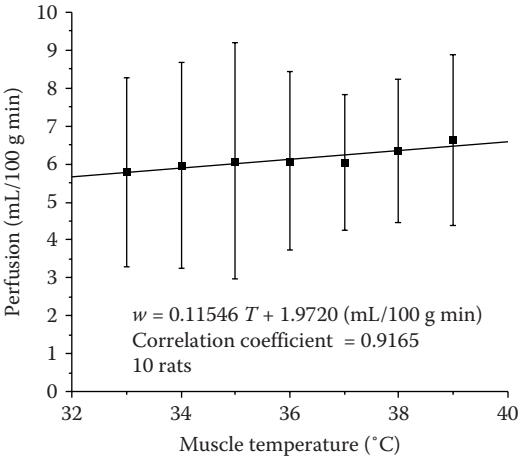


FIGURE 4.5.5 Perfusion vs. muscle temperature averaged over 10 experiments. (From Valvano, J.W., *Optical-Thermal Response of Laser-Irradiated Tissue*, Plenum Press, New York, 1995. With permission.)

perfusion samples were approximately either 5 or 60 min. Under baseline conditions, the temperature fluctuations within the prostate were approximately ±0.3°C. A relative dispersion estimate of 15% was derived from one dog for the fluctuations in baseline perfusion. Thus, changes in absolute perfusion and temperature greater than 15% and 0.3°C, respectively, were considered to be substantial changes.

As heating progressed, a variety of substantial changes were observed, but no uniform pattern emerged. However, the measurements included changes typically expected for hyperthermia:

1. An initial perfusion increase associated with elevating the baseline temperature.
2. A perfusion return towards baseline after this initial increase.
3. A dramatic increase in perfusion at elevated temperatures.

The initial perfusion increases were observed in three dogs when the temperatures exceeded 38°C ± 3°C (mean ± 1 SD, $N = 8$). The perfusion increased 34% from a baseline value of 0.59 ± 0.26 mL/g min over a temperature rise of 1.7°C ± 1.3°C. Half of the measurements in the three dogs subsequently showed a decrease in perfusion ranging from 16% to 25%. In two dogs, dramatic perfusion increases as high as 364% were observed with a corresponding decrease in tissue temperature.

The mean perfusion and temperature measured from all dogs at similar instances during the experimental protocol were calculated. Figure 4.5.6 shows the changes as the experiment progressed from baseline conditions.

The mean behavior indicates no substantial change in perfusion until the tissue temperature exceeded 39.6°C, after which the perfusion increased 17%. This occurred over a 5-min period as the nominal MW power was stepped from 10 to 15 W. The perfusion increased another 35% when the tissue temperature exceeded 41.7°C, and appeared to affect a slight lowering of tissue temperature.

Xu et al.[71] measured perfusion using the pulse-decay self-heated thermistor technique[13,20] in these dog prostates during the same transurethral MW hyperthermia treatments. Interestingly, in the exact same dogs at roughly the same locations, the perfusion response to temperature measured with the thermal technique was roughly linear with temperature as shown in Figure 4.5.7.

The interdependence between perfusion and temperature was observed in these studies: most notably, a decrease in tissue temperature associated with a dramatic increase of perfusion. Such

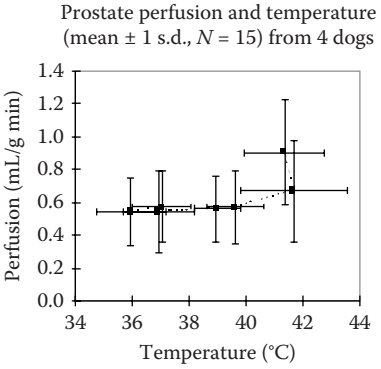


FIGURE 4.5.6 Perfusion vs. prostate temperature.[71]

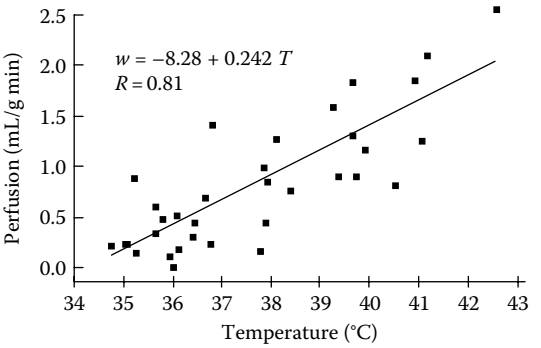


FIGURE 4.5.7 Perfusion vs. prostate temperature.[72]

changes have been modeled previously with Pennes's bioheat equation, using assumed perfusion values and changes for the prostate. The data from this study will provide more realistic estimates of perfusion values and thermoregulation models in hyperthermic canine prostates. Two hypotheses exist that explain the differences between the perfusion responses to hyperthermia as measured by microspheres and the thermal decay probe. The first possibility is that one or both methods have significant measurement errors. A second possibility is that perfusion as measured by spheres trapped in the capillaries is a different parameter than perfusion as measured by enhanced heat transfer within the 70- to 200-μm-diameter vessels.

Human Thermoregulation

Humans possess an elaborate thermal control (thermoregulation) system that ensures maintenance of internal body temperatures near a physiological set point under a large spectrum of environmental conditions and metabolic rate activities. After many years of research, much has been learned of the operation of the human thermoregulatory system and it remains a topic of active investigation. A number of reviews on specific aspects of the physiological basis of thermoregulation are recommended for further background detail.[72–75]

Internal and boundary heat transfer processes are key to thermoregulation. As the operational mechanisms have become better understood, it has been possible to model these functions in an attempt to simulate and predict thermoregulatory behavior and to be able to design systems to interact thermally with the human body (such as an environmental protective garment or a space suit) without compromising the health and safety of the subject.

Physiological Processes of Thermoregulation

The prevailing thermal physiology theory is that the human thermoregulation system operates to maintain the body core temperature at a constant value consistent with that required for regular body function, regardless of the environmental temperature and energy loading. Alternatively, it has been suggested that thermoregulation serves the purpose of maintaining the body's energy balance, with body temperatures being a result, not a cause, of the regulation process.[72,76,77]

Many general observations of the thermal characteristics of human thermoregulation are consistent, even if alternative explanations of the underlying operative mechanisms are proffered. For example, temperature and internal energy storage of the human body vary with time of day, metabolic activity, and individuality of the person.[78] The body incorporates a number of energy production and dissipation mechanisms to preserve thermal homeostasis, many of which are controlled by feedback signals based on specific physiological states. Examples of thermoregulatory processes that are governed by feedback are sweating, shivering, and variable localized blood flow.

According to the theory of thermoregulation by control of the body core temperature, the main control center is located in the hypothalamus of the brain, from which multiple reflex responses operate to maintain the body temperature within a narrow range.[79] The signals that activate the hypothalamic temperature-regulating centers come largely from two sources: the temperature-sensitive cells in the anterior hypothalamus and cutaneous temperature receptors. The cells in the anterior hypothalamus sense the temperature of the body core or, specifically, the temperature of arterial blood which passes through the head.

Webb has presented a theory of thermoregulation to maintain an energy balance within the body, as demonstrated by the existence of temperature sensors at several levels in the skin enabling the sensing of heat flow within and from the body.[77] Evidence also supports neurological sensing of thermal gradients which direct thermoregulation behavior. The hypothesis in support of the theory of energy content regulation based on Webb's experimental observations is:

> Heat (energy) regulation achieves heat (energy) balance over a wide range of heat (energy) loads. Heat flow to or from the body is sensed, and physiological responses defend the body heat (energy) content. Heat (energy) content varies over a range that is related to heat (energy) load. Changes in body heat (energy) content drive deep body temperatures.[77]

The proposed mechanism of energy-driven thermoregulation balances the constantly changing metabolic energy production and the adjustment of heat losses to maintain the body as a system at steady-state. In contrast, the proposed mechanism of temperature-driven thermoregulation affects a coordination of physiological processes to maintain the body core temperature at a set point.

Thermoregulatory Processes

Conservation of energy for the human body must account for internal metabolic energy production plus multiple mechanisms of environmental heat and work exchange.

$$\Delta E = M - (W + Q_{conv} + Q_{cond} + Q_{rad} + Q_{evap} + Q_{resp}) \tag{4.5.51}$$

where ΔE = Rate of energy storage in the body (W)
M = Metabolic energy production (W)
W = External work (W)
Q_{conv} = Surface heat loss by convection (W)
Q_{cond} = Surface heat loss by conduction (W)
Q_{rad} = Surface heat loss by radiation (W)
Q_{evap} = Surface heat loss by evaporation (W)
Q_{resp} = Respiratory heat loss (W)

The human body produces energy, exchanges heat with the environment, and loses heat by evaporation of body fluids. Energy is produced in the body by basal (resting), metabolism defined as the minimal metabolism measured at a temperature of thermal neutrality in a resting homeotherm with normal body temperature several hours after a meal and not immediately after hypothermia, and also at an increased rate due to muscle activity, including physical exercise and shivering, and by food intake. Therefore, the total energy production in the body is determined by the energy needed for basic body processes plus any external work. Since the body operates with less than 100% efficiency, only a fraction of the metabolic rate is applied to work, with the remainder dissipated as heat.[78] The mechanical efficiency, η, associated with metabolic energy utilization, is zero for most activities except when the person is performing external mechanical work such as in walking upstairs, lifting something to a higher level, or cycling on an ergometer.[80] When work transferred from the environment is dissipated as heat in the human body, η is negative. An example of this case is walking downstairs.

Convection, radiation, conduction, and evaporation of sweat at the skin surface all facilitate heat transfer from the body. Heat transfer also occurs via the respiratory tract and lungs, although to a lesser extent in human beings as compared with many mammals and reptiles. Storage of energy takes place whenever there is an imbalance of production and dissipation mechanisms. In many instances, such as astronauts in space suits or military personnel in chemical defense garments, energy storage is forced due to the lack of appropriate heat exchange with the environment.[81]

The human thermoregulatory system is quite complex and behaves mathematically in a highly nonlinear manner. It contains multiple sensors, multiple feedback loops, and multiple outputs.[82] The primary mechanisms by which the body responds to control the storage of energy (positive or negative) include the evaporation of sweat, shivering of the muscles, and vasoconstriction and vasodilation of the blood vessels.

Heat transfer internal to the body is due to the conductance which governs the flow of energy between the core, through the tissue, and to the surface. This transport process is governed significantly by peripheral blood flow, the core-skin temperature gradient, and the conductivity of the various body tissues. Transient blood perfusion provides an effective variable path of peripheral energy transport via convection between blood and tissue and countercurrent heat exchange between the arteries and the veins. Blood flow distribution is controlled according to metabolic needs of the body, local tissue temperature, and the need to maintain the appropriate core temperature. When the core becomes too hot, the blood vessels in the skin dilate to allow increased perfusion to the body surface. The blood is cooled by the environment, and the cooler blood is returned to the core. Increased blood flow to the skin surface also enables greater sweat production, adding to the cooling process. In contrast, when the core becomes too cold, blood flow to the skin is constricted to conserve the body's internal energy. Sweating occurs when the arterial blood is elevated in temperature, causing the hypothalamus to increase nerve impulses to the sweat glands. Shivering, on the other hand, is an involuntary response of the skeletal muscles when passive body cooling exceeds metabolic energy production.[79]

Human thermoregulation mechanisms can be divided into two main categories: autonomic and behavioral.[38] Autonomic thermoregulation is controlled primarily by the hypothalamus, whereas behavioral thermoregulation is controlled consciously by humans and includes active movement and adjustment of clothing. Behavioral thermoregulation is associated with conscious temperature sensation as well as with thermal comfort or discomfort. Some engineered systems are designed to interact with the body to enhance its ability to maintain thermoregulation in harsh or hazardous environments. Examples range from the common HVAC systems encountered in buildings and vehicles to sophisticated garments worn during certain military and space missions.

Wissler Model of Human Thermoregulation

Models that accurately incorporate the transient whole-body behavior during thermoregulation for a wide range of states and environmental challenges may be quite useful in describing and predicting this important human physiological function. Moreover, such a model can be used as a design tool

in the development of systems with which humans must interact for a variety of work, pathological, and recreational circumstances. The development of models of human thermoregulation has proved to be a daunting task that has been addressed by many researchers. The complexity of the coupling among physiological processes involved in thermoregulation and of the control algorithms has dictated that models that incorporate these multiple effects be solved numerically. Thus, the first realistic thermoregulation models appeared in the 1960s and 1970s from the studies of Wissler,[34,83] Stolwijk,[84,85] Mitchell et al.[85,86] Hayward et al.,[87,88] and Kuznetz.[89] An important early application of modeling human thermoregulation was the design and development of active thermal control garments to be worn under the space suit during extravehicular activity.[78,81,90,91] Over the years the Wissler model has been improved and updated on a continuous basis, and it is now applied to a very broad spectrum of human thermal control scenarios.[35,92]

The Wissler human thermoregulation model is a transient-state mathematical formulation used to describe human physiological responses to various levels and combinations of exercise and thermal stress. The model computes thermal, cardiovascular, ventilatory, and metabolic changes that occur during a specified period of time. The values for the independent physiological and environmental variables, such as metabolic rate, environmental conditions, and garment cooling liquid flow rate, are specified as inputs to the model. A finite-difference scheme is applied to numerically compute tissue and blood temperatures throughout the body as a function of time.

The model represents the human body as 15 cylindrical elements for the head, thorax, abdomen, and proximal, medial, and distal portions of each arm and leg. A schematic representation of the multi-element model of the human body is shown in Figure 4.5.8. Each body element is composed of cylindrical shells that are defined by physical and physiological properties for composite layers of viscera, bone, muscle, fat, and skin. In each of the elements metabolic energy may be generated as appropriate due to basal metabolism, exercise, and shivering. The generated energy is either conducted to adjacent body regions, convected to circulating blood, or stored in the element. Each body element contains an arterial and venous pool and a capillary bed. The pools are interconnected serially between adjacent elements. As venous blood is returned to the heart it passes through the lungs for exchange of gases before entering the arterial system.

A number of signal feedback loops are used to account for thermoregulation within the body. Feedback data are the differences between skin and central head temperatures and their respective specified set point values. Error signals specified by these differences drive sweat secretion, shivering, and vasomotion responses. The algorithm equations have the general format of a proportional controller that is activated when a defined threshold condition is satisfied. The control equations do not all have the same level of response to error signals. For example, sweating and vasodilation are more responsive to an increase in central head temperature than to cutaneous temperature, while vasoconstriction is more responsive to cutaneous temperature.

Thermoregulation is also affected via heat exchange during the venous return flow of blood. Venous flow occurs through either deep or superficial veins, thus influencing the rate of countercurrent heat exchange. Modeling of the venous return provides a realistic simulation of the countercurrent heat exchange that takes place in the extremities during exercise in hot environments and immersion in cold water.

Blood flow rates are simulated to satisfy the level of local metabolic need for oxygen in the tissues. Therefore, a lower limit is set for extremity blood perfusion during cold exposure. The model also incorporates material balances for oxygen, carbon dioxide, and lactate as defined for cardiovascular response. In addition to physiological factors, conditions of the environment with which the subject interacts are accounted for. Physical properties of the environmental fluid and characteristics of the flow field in the vicinity of the subject are specified as they determine conductive and convective heat transfer.

The model allows the user to specify layers of clothing described by physical properties of annular shells outside of but interacting with the skin layer. The physical properties of the garments may change with the accumulation of unevaporated sweat, which may saturate the material.

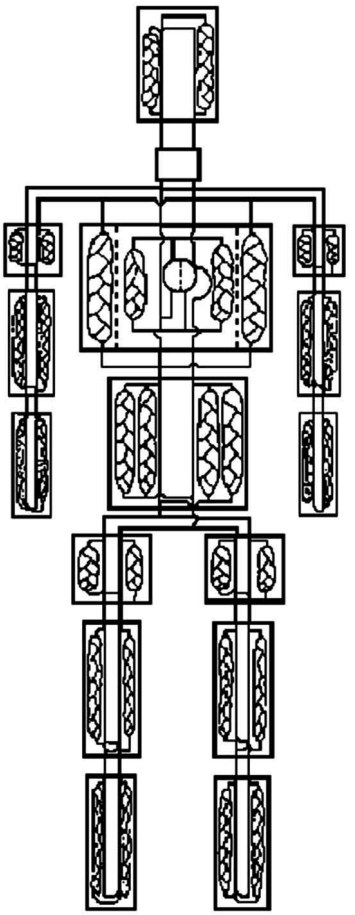


FIGURE 4.5.8 Multielement model representing the human body in the Wissler model. (From Wissler, E.H., *Heat Transfer in Medicine and Biology: Analysis and Applications*, Vol. 1, Plenum Press, New York, 325, 1985. With permission.)

Transient tissue and blood temperatures are calculated during simulated thermal scenarios by Crank-Nicolson-type finite-difference equations. Updated temperatures and other physiological values are computed every 10. The model allows a simulation to be broken down into any number of scenario subintervals, with the following factors being specified in each interval:

- Basal metabolic rate
- Kind and level of exercise being performed
- Environmental conditions (pressure, temperature, dew point, etc.)
- Garment properties
- Use of a fluid-conditioned garment (FCG) (flow rates, inlet temperatures, etc.)

The equations needed to describe a liquid- or air-cooled garment are derived from material and energy balances for the circulating fluid. Sensible heat transfer is treated the same for liquid- and air-cooling situations. However, the evaporation and condensation associated with latent heat transfer must be treated differently for the two fluids. In the case of the liquid-cooled garment, water from sweat and environmental moisture condenses on the cool surface of the garment. On the other hand, air cooling generally facilitates evaporation of sweat by removing water vapor from the garment.

In its early formulation the Wissler model required a large mainframe computer to perform simulations of thermal regulation. However, owing to advances in computational capabilities in recent years, it is now possible to run extensive simulations on a desktop personal computer in an order of magnitude less than real time. The thermoregulation model provides reliable predictions of changes in the human thermal state for many different types of environmental and metabolic stress, and it has been used for the design of new automated personal garments for application in hazardous environments.[92]

Therapeutic Heating

Normal physiologic temperatures range from approximately 30°C to 42°C. The average so-called core (central abdominal) temperature in mammals varies slightly among species between about 36°C and 38°C (37°C for humans, 38°C for most canines). It was noted above that thermal gradients are inherent in metabolically driven open systems at steady state, and measured "core" temperatures will vary slightly depending on the relative metabolic rate, blood flow, and location of the various organs, and from point to point within an organ. Though core temperature is maintained close to a steady-state value, skin surface temperature at rest in a controlled room environment varies between about 30°C and 34°C, depending on climate and season.

Elevation of body, appendage, or organ temperature to between 40°C and 42°C can be used for analgesia (relief of pain) and to provide the most advantageous environment for injury response (wound healing) processes. Note that there is no external intervention that can be claimed to "promote" wound healing—one may only provide the best conditions under which the natural processes will heal wounds. Thermal intervention has been used for many years for these two purposes.

Heat Generation Modalities

Therapeutic heating can be obtained by surface heat transfer means (hot or cold packs), but the effectiveness is limited to surface injury. Nevertheless, contact heating is the method of choice for such injuries as sprains, muscle strain, and postoperative swelling. The standard response to elevated temperatures is to increase the blood flow, and thus perfusion. One applies cold heat sinks to reduce perfusion in the injured area during the first 12–24 h and hot packs or circulating water (up to 45°C) after that to promote perfusion and thus increase the supply of neutrophils and other blood components related to injury response and wound healing.

Deeper injuries require volumetric heating to attain adequate temperatures. The most effective volumetric heating methods to date have been (1) electromagnetic fields at radiofrequencies (RFs) or MW frequencies), and (2) ultrasound (US). Assigned RFs for industrial, scientific, and medical (ISM) use are 6.78, 13.56, 27.12, and 40.68 MHz. Practical ISM frequencies for medical use in the MW range are 915 MHz and 2.45 GHz. Typical ultrasound heating devices operate between about 500 kHz and 10 MHz. In all three cases there is an engineering trade-off among local volumetric heat generation, q''', thermal heat generation field dimensions, and depth of penetration. Higher frequencies have shorter wavelengths and, in general, are absorbed more strongly and thus penetrate less deeply. Tissues are inhomogeneous and, with only a few exceptions, anisotropic. The electrical and acoustic properties of tissues can vary over several orders of magnitude. Consequently, electromagnetic and acoustic boundary conditions often determine the volume generation term in spite of efforts to shape the field by clever applicator design.

For electromagnetic heating in tissues the volume generation term is essentially governed by the electrical conductivity, σ (S/m), the imaginary part of the electric permittivity, ε''—$\varepsilon^* = \varepsilon' - j\varepsilon''$ (F/m)—and the magnitude of the local electric field, $|\mathbf{E}|$:

$$q''' = (\sigma + \omega\varepsilon'')|\mathbf{E}|^2 + \omega\mu''|\mathbf{H}|^2 \tag{4.5.52}$$

where ω is the angular frequency (rad/s). Heating can be due to direct absorption from a magnetic field, $\omega\mu''|\mathbf{H}|^2$ where $\mu^* = \mu' - j\mu''$ is the complex magnetic permeability (H/m). Magnetic field heating is usually negligible except in the case of magnetic "seeds" or magnetic nanoparticles. Yet, RF and MW magnetic fields heat tissue very effectively, too, despite that tissues have the (real) magnetic permeability of free space, μ_0. This is because, by Faraday's law of induction, a time-varying magnetic field will induce an electric field in the tissue, and it is the electric field that actually generates the heat in tissues.

Faraday's Law in integral form:

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{\partial}{\partial t}\left[\int \int \mu \mathbf{H} \cdot d\mathbf{S}\right] \tag{4.5.53}$$

in point form:

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} \tag{4.5.54}$$

A comprehensive discussion of these properties and effects in materials may be found in the work of Green and Diller.[98] Values of relevant electrical and acoustic properties of tissues are given in Tables 4.5.7[93] and 4.5.8,[94] respectively.

Acoustic waves are highly scattered in many tissues, so simple wave-propagation models are inadequate to predict the local acoustic field strength. Nevertheless, a uniform plane-wave description of the heating field often suffices to explain the distribution of the heat generation with acceptable accuracy. The volume heat generation term is determined by the acoustic absorption coefficient, α (m^{-1}), and the wave power density, U_0 (W/m^2) E, approximately according to Beer's law:

$$q'''(z) = \alpha U_0 e^{-\alpha z} \tag{4.5.55}$$

TABLE 4.5.7
Representative Electrical Properties of Tissues

Tissue	Conductivity (S/m)	Real Part, ε_r' ε_r'	Imaginary Part, ε_r''
Fat	0.05–0.09	3.9–7.2	0.67–1.4
Muscle	0.1–1.0	45–48	13–13.9
Bone	0.006–0.05	4.2–5.8 (marrow)	0.7–1.3 (marrow)

Source: Geddes, L.A., and Baker, L.E., *Med. Biol. Eng.*, 5, 271, 1967.

Note: Muscle is highly anisotropic, higher conductivity parallel to fibers. Electric permittivity is as measured at 3 GHz, and $\varepsilon(F/m) = \varepsilon_0\left[\varepsilon_r' - j\varepsilon_r''\right]$ where $j = \sqrt{(-1)}$ and ε_0 = free space permittivity, 8.85×10^{-12} F/m.[94]

TABLE 4.5.8
Representative Acoustic Properties of Tissues[94]

Tissue	Velocity, *a* (m/s)	Absorption at 1 MHz (cm^{-1})
Fat	1450	0.06
Muscle	1585	0.2–0.7
Bone	4080	30

where U_0 is the surface power density at $z=0$ (W/m^2) and α varies according to the acoustic wave frequency, f: $\alpha \approx f^n$, where $1 < n < 2$. Table 4.5.8 units for α are (cm^{-1}). In other references, α may be reported as dB/cm [dB = 10 $\log_{10}(U_{out}/U_{in})$].

Physiologic Effects of Local Heating

The general physiological effects of local heating comprise a cascade of responses.[95,96] As noted in the previous section, systemic response is controlled from the hypothalamus by both neuronal and hormonal signals. Appendage- and organ-level response to local heat may be spinal-cord-mediated; but local heating can trigger local release of bradykinins that induce vascular dilation or constriction, thus affecting temperature rise.

The increase in tissue temperature is accompanied by an increase in cellular metabolism. The arterioles (microscopic arteries, typically about of 40–200 μm inner diameter[95]) dilate in response to heating under smooth muscle control. The downstream capillary pressure increases above the homeostatic level of about 25 torr. Two consequences are an increased capillary flow and capillary pressure. First, at higher capillary pressures the gaps between the endothelial cells (thin epithelial cells comprising the vessel wall) tend to widen and, at higher pressures, edema may form (an increase in the fluid in the extracellular compartment).[97,98] Second, higher capillary flow results in rapid clearing of cellular metabolites and increases convection heat transfer, delivery of tissue oxygen, cellular nutrients, antibodies, and leukocytes (white blood cells) including monocytes which are necessary in the healing response. Neutrophils are white blood cells that initially release proteolytic enzymes to digest dead tissue and, later, phagocytose (eat) the debris.[99] Monocytes are other leukocytes that remove damaged and dead cells by phagocytosis.

Tissue Effects: Elevated Temperatures

Exposure to temperatures above normal physiologic ranges (>42°C) over time can result in measurable irreversible changes in tissue structure or function. Cell death or tissue alterations may be detrimental—for example, skin burns—or beneficial, as in vessel sealing or tumor destruction. Tissues of the central nervous system are the most thermally sensitive, exhibiting irreversible changes for long-term exposures above about 42°C. The specific pathologic end point may be evaluated histologically, histochemically, and/or physiologically. Most assays of thermal alterations are qualitative in nature; however, several end points that are inherently quantitative lend themselves well to rate process descriptions of their thermal kinetics. Examples of quantitative assays[99] include loss of birefringence, vital stain uptake, fluorescent stains, gene expression, apoptosis, necroptosis, traumatic necrosis, and loss of clonogenicity (in cancer cells). Even for qualitative processes, thermal kinetic models often provide useful descriptions and so provide helpful insights into the underlying principles of tissue thermal damage.

Here we use "damage" in a generic sense to denote all types of irreversible alterations, therapeutic or not. Many damage processes can be modeled as first-order rate processes, Arrhenius models, for which two experimentally derived coefficients are sufficient. First-order Arrhenius analysis applies well to thermal alterations in structural proteins (such as collagen, elastin and actin) and to cell death processes at higher temperatures. Other intrinsic programmed cell death processes (with either intrinsic or extrinsic signal mechanisms), such as apoptosis, necroptosis and, possibly, autophagy involve intricate multipath cascades of functional proteins (chiefly enzyme catalysts) that require much higher-order kinetic models to describe adequately. We will illustrate that the net effect of the higher order chemical kinetics is to introduce a temperature-dependent delay into the cell death process, a so-called shoulder region in a cell survival curve. It turns out that many such processes exhibiting shoulder region behavior can be accurately represented by adding a temperature-dependent time delay to a first-order Arrhenius model.

In the Arrhenius models, irreversible thermal damage is exponentially dependent on temperature and linearly dependent on time of exposure. The first-order rate process models often apply well to

the prediction of damage thresholds and less well as the damage becomes complete or severe since several of the fundamental assumptions are violated. In order to be useful in evaluating thermal insult, the kinetic model must be coupled to quantitative pathological analysis. This subsection describes several quantitative markers of thermal damage and experimental methods for estimating relevant kinetic coefficients both in constant temperature and transient thermal history experiments. As expected, transient *in vivo* thermal history data yield a noisy kinetic plot; however, useful estimates of the appropriate rate coefficients can be made.

Theory of Rate Process Descriptions: The Arrhenius Model

The original work on the application of rate process models to thermal damage was reported by Moritz and Henriques in a series of seminal papers entitled "Studies of Thermal Injury" in 1947.[100–103] They applied flowing water at elevated temperatures to pig skin and measured exposure times required to create first-, second-, and third-degree burns. In their work, the damage was quantified using a single parameter, Ω, which ranges on the positive real axis and is calculated from an Arrhenius integral:

$$\Omega(\tau) = \int_0^{\tau} A e^{\left[\frac{-E_a}{RT}\right]} \mathrm{d}t \tag{4.5.56}$$

where A is a frequency factor (s^{-1}), τ the total heating time (s), E_a an activation energy barrier (J/mol), R the universal gas constant (8.314 J/mol/K), and T the absolute temperature (K).

Over the ensuing 50 years, many subsequent investigators have modeled this process[104–107] and made experimental measurements of threshold burn conditions in human,[108–110] animal,[104,111] and cellular and molecular[112–114] systems. Although a considerable body of literature has accrued, there is by no means a consensus on how to accurately predict the occurrence of thermal injury over the wide range of conditions that cause burns.

One difficulty with this model is that a single damage parameter inherently lumps all damage processes into one global measure. In multiple process cases, such as in the case of a skin burn, the lowest temperature process saturates the damage measure, Ω, early during an exposure. In this discussion, we review the underlying assumptions and origin of the terms in Equation 4.5.56 and recast the traditional thermal damage parameter, Ω, into a form suitable for evaluation of multiple-process thermal damage effects.

Kinetic Models of Reaction Product Formation Rates

The basis for rate process models of thermal damage may be obtained from chemical reaction kinetics (see, e.g., the work of Maron and Lando[115]). In a typical reaction process, thermally active reactants surmount an activation barrier to form products, as illustrated in Figure 4.5.9. In the figure,

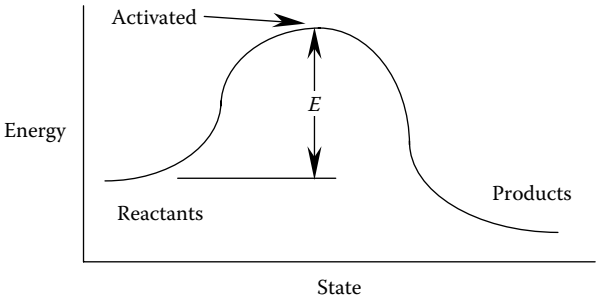


FIGURE 4.5.9 Energy-state diagram of a bimolecular process.

E is the energy barrier in Equation 4.5.56. The collision theory description of ordinary first-order bimolecular reaction kinetics holds that the reactants are activated by collisions; n^* are activated out of n total molecules, and the probability of activation is

$$\frac{n^*}{n} = e^{-\left[\frac{E}{RT}\right]} \tag{4.5.57}$$

In such a process, activated reactants are considered to form an activated "complex" which may either relax to inactivated single reactants or irreversibly progress to form product molecules. The complex has some of the properties of an ordinary molecule and is at least temporarily stable. For reactant molecules *A* and *B*, the sequence of formation is

$$[A]+[B] \underset{k_b}{\overset{k_a}{\rightleftarrows}} [AB^*] \xrightarrow{k} \text{Products} \tag{4.5.58}$$

The overall reaction velocity, k (s^{-1}), determines the rate of formation of product and is related to the equilibrium constant for formation of activated complex, K^*, by

$$k = \frac{RT}{Nh} K^* = \frac{RT}{Nh} e^{-\frac{\Delta G^*}{RT}} \tag{4.5.59}$$

where N is Avogadro's number (6.023×10^{23}), h is Planck's constant (6.627×10^{34} J s), and ΔG^* is the Gibb's free energy of formation of the activated complex. In turn, the free energy of formation is given by

$$\Delta G^* = \Delta H^* - T\Delta S^* \tag{4.5.60}$$

where ΔH^* is the enthalpy (i.e., the total energy) of activation (J/mol) and ΔS^* is the entropy of activation J/(mol K). Employing this relationship results in the Eyring–Polanyi equation.[116,117]

$$k = \frac{RT}{Nh} e^{\left[\frac{\Delta S^*}{R}\right]} e^{-\frac{\Delta H^*}{RT}} \tag{4.5.61}$$

The activation entropy is not calculable except for the simplest reactions in the gas phase, and therefore is usually determined from experimental measurements of the reaction velocity and activation enthalpy. The activation enthalpy, ΔH^*, is determined from the observed activation energy, E_a by

$$\Delta H^* = E_a - iRT \tag{4.5.62}$$

where i is 1 for first-order reactions in solution and gases, 2 for second-order, and 3 for third-order reactions.

Unimolecular Process Descriptions

Thermal damage in tissue is generally a unimolecular process—tissue constituent proteins transition from the native state to the damaged state (Figure 4.5.10). Absolute reaction rate theory can also be used to explain the rate of formation for this process if we assume that a time lag exists between molecular activation and denaturization.[115] During this time lag, the molecules may either denature or relax back to the native state, as illustrated in Figure 4.5.10. Here, ΔH is the enthalpy (internal thermal energy) difference between native state and denatured molecules. The relative barriers are such that, in the thermal damage of tissue, ΔH^* is almost always smaller than ΔH.

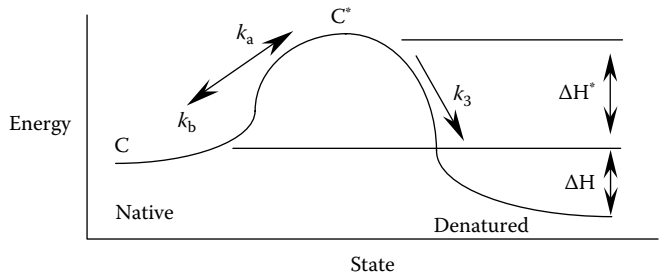


FIGURE 4.5.10 Unimolecular process activation and denaturalization.

So, the activation process may be regarded as reasonably likely, and the probability of denatured tissue relaxing back to native state tissue is near enough to zero that it may be regarded as the impossible event in the absence of an energy-consuming healing process. The rate of damage formation is then proportional to only those molecules that remain activated. For a unimolecular process in the native state C, having an activated state, C^*, with velocity constants k_a, and k_b:

$$[C] \underset{k_b}{\overset{k_a}{\rightleftarrows}} [C] + [C^*] \tag{4.5.63}$$

the activated complex, $[C^*]$ progresses to the damaged state at k_c:

$$[C^*] \xrightarrow{k_c} \text{Damaged state} \tag{4.5.64}$$

so that the rate of disappearance of native state molecules, $[C]$, is given by

$$-\frac{d[C]}{dt} = k_c [C^*] \tag{4.5.65}$$

Generally, $[C^*]$ is neither known nor calculable; however, at sufficiently low concentrations of C^*, the steady-state principle asserts that for short-lived activated states, the rate of formation can be considered equal to the rate of disappearance. The activated state, $[C^*]$, forms at a rate $k_a [C]^2$, relaxes back to inactivated at rate $k_b [C][C^*]$ and denatures at the rate $k_c [C^*]$. Consequently.

$$k_a[C]^2 = k_b[C][C^*] + k_c[C^*] \tag{4.5.66}$$

and so

$$[C^*] = \frac{k_a[C]^2}{k_c + k_b[C]} \tag{4.5.67}$$

We actually need an overall reaction velocity, k, which relates $[C]$ to its rate of disappearance:

$$-\frac{d[C]}{dt} = k[C] \tag{4.5.68}$$

There are two limiting cases for Equation 4.5.67: first, the concentration of remaining undamaged material, $[C]$, may be large enough that deactivation at k_b dominates the k_c pathway, so $[C^*] \cong [C]\, k_a/k_b$ for which the overall formation rate, $k = k_c\, k_a/k_b$ and a first-order process results; and second, if the remaining undamaged material concentration, $[C]$, is small, $k_c \gg k_b\,[C]$ and the process is second order since from Equation 4.5.67 $k = k_a\,[C]$. In liquid-phase systems with appreciable concentrations of native state molecules, the first condition should apply, so the first-order approximation applies.

After a long time of exposure at damaging temperatures such that [C] is very small, $k_c \gg k_b\, C$ and a second-order process results:

$$-\frac{d[C]}{dt} = k[C]^2 \tag{4.5.69}$$

where for simplicity the [C] dependence has been removed from k.

First-Order Solution

Equation 4.5.67, then, is a Bernoulli differential equation with the solution:

$$C(\tau) = C(0)e^{\{-\int k\, dt\}} \tag{4.5.70}$$

Equations 4.5.59 and 4.5.60 may be used to relate k to ΔH^* and ΔS^*. It should be noted at this point that the energy barrier, E_a (Figure 4.5.9) is in fact $\Delta H^* + RT$; however, in practice $\{\Delta H^* \simeq 5\times10^5\} \gg \{RT \cong 3\times10^3\}$, so we may assume that $E_a \simeq \Delta H^*$.

The pre-exponential term in Equation 4.5.59 suggests that it is temperature dependent; however, the linear dependence of A on T is extremely weak compared to the exponential dependence in the final term.

Application of Kinetic Formulations in Thermal Damage and Cell Death Studies

When a quantitative thermal damage phenomenon has been identified, it can be studied as a first-order process. This subsection inspects the more common first-order analysis in some detail and introduces one form of combined process analysis to extend the usefulness of the method. The specific example of Henriques and Moritz original data is also studied. When the damage end point is a qualitative measure, application of these kinetic models is closer to a curve-fitting exercise than to a fundamental study of the underlying phenomena. Therefore, precise identification and strict definition of quantifiable damage end points are required for meaningful analysis.

First-Order Arrhenius Process Analysis

A more useful form of Equation 4.5.56 may be obtained by recasting the result into a volume fraction model. In this formulation, C signifies the remaining concentration of native state (undamaged) tissue constituent. Therefore, the physical significance of the traditional damage measure, Ω, is the logarithm of the ratio of the original concentration of native tissue to the remaining native state tissue:

$$\Omega(\tau) = \ln\left\{\frac{C(0)}{C(\tau)}\right\} = \int_0^\tau k\, dt = \int_0^\tau A e^{-\left[\frac{E_a}{RT}\right]} dt \tag{4.5.71}$$

This form of the damage integral has the advantages that (1) it is easily compared to quantitative pathologic end points such as tissue collagen and muscle birefringence loss, collagen hyaline damage, leakage of fluorescent dyes from dead or dying cells, or some types of cell survival in culture, and (2) multiple thermodynamically independent damage processes may be simultaneously modeled and/or studied in parallel by accumulating separate values for $C(\tau)$. The first-order Arrhenius formulation is effective when the damage process is dominated by irreversible thermal alterations in structural proteins or in cell death processes at higher temperatures—those substantially in excess of the "hyperthermic" range, above about 55°C. Programmed cell death processes often dominate at hyperthermic temperatures—between about 42°C and 50°C; however, whether or not intrinsic processes dominate depends critically on cell type and the particular process studied. Intrinsic cell death processes involve several cascades of functional proteins and require substantially more

sophisticated mathematical formulations to obtain accurate results. Using quantitative damage/death formulations, direct comparisons can be made between computer models of $T(x, y, z, \tau)$ and measured histologic damage.

Confining the discussion to the Arrhenius formulation for the moment, a set of coefficients, A and E_a, is required for each damage process considered in the computer model. For Arrhenius tissue damage processes studied to date, A varies from about 10^{40} to 10^{105} s^{-1}, while E_a usually ranges from about 10^5 to 10^6 J/mol. Each damage process is then allowed to progress in parallel, driven by the calculated thermal field. This formulation assumes that the individual processes are thermodynamically independent, a reasonable description for identifiable thermal damage processes. The concentration of each of the "i" damage markers, $C(\tau)_i$, is accumulated, and a distributed field description of the predicted histologic end point may be generated.

Interestingly, A and E_a are not independent, and comprise a coupled pair. In 2003 two papers appeared that noted this correlation, one by He and Bischof[118] and the other by Neil Wright.[119] The respective correlations are given by Equations 4.5.72 and 4.5.73:

$$\text{He–Bischof: } \ln\{A\} = 3.80\times10^{-4} E_a - 9.236 \tag{4.5.72}$$

$$\text{Wright: } \ln\{A\} = 3.83\times10^{-4} E_a - 10.04 \tag{4.5.73}$$

In these relationships, the units of E_a have been corrected to (J/mol). Over the typical range of E_a, the two relationships are virtually indistinguishable. There has been some controversy regarding the applicability of the correlations. There is support for the correlations in the hypothesis of "Compensation Law" behavior, as presented by Rosenberg et al.[120] Barrie questions the Compensation Law, pointing out that experimental error can result in something that looks like Compensation Law behavior.[121,122] Yelon et al. argue successfully against Barrie's assertion,[123] and the point is finally conceded by Barrie.[124] Wright cited support for the correlation based on a "polymer in a box" construct by Miles and Ghelashvili.[125] Qin et al. argue in favor of the correlation on biochemical grounds.[126] In fact, inspection of the Eyring–Polanyi equation over a fairly narrow temperature range (Equation 4.5.61), as He and Bischof did in 2003, suggests that ΔS^* and ΔH^* should be correlated in some sense. It is also reasonable to expect that a higher energy process would be accompanied by higher entropy.

Functional behavior of the first-order Arrhenius model—The characteristic behavior of the kinetic damage model is that below a threshold temperature the rate of damage accumulation is negligible, but it increases precipitously when this value is exceeded. For purposes of discussion, it is useful to define the critical temperature as the temperature at which the damage accumulation rate, $d\Omega/dt$, is 1.0:

$$T_{crit} = \frac{E_a}{R\ \ln(A)} \tag{4.5.74}$$

A hypothetical damage process with representative coefficients chosen from the approximate center of the He–Bischof line with $E_a = 5\times10^5$ and $A = 3.197\times10^{78}$ has a critical temperature of 59.6°C. Figure 4.5.11 illustrates the damage accumulation rate dependence on temperature for the hypothetical example process.

Constant temperature exposures of the example process will result in a decrease in concentration of native state material depending on the time of exposure. Figure 4.5.12 shows the remaining concentration for this process for constant temperature exposures of $\tau = 0.1$, 1.0, and 10 s; the concentration is seen to gradually decrease with increasing temperature for fixed exposure times, as expected. The strong exponential nature of the process is evident as well. Applying this model framework relies heavily on identifying independent damage processes which can be quantitatively measured. From plots of the form of Figure 4.5.12, estimates of A and E may be made. Each curve will give

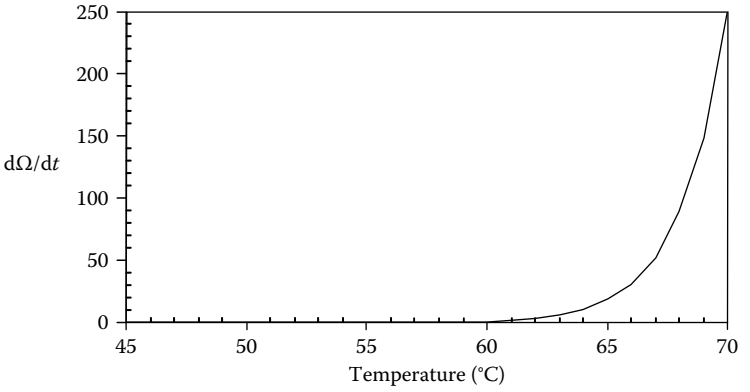


FIGURE 4.5.11 Hypothetical example damage accumulation process has $A = 3.197 \times 10^{78}$ and $E = 5 \times 10^5$ which gives the damage rate, $d\Omega/dt$, shown and a critical temperature of 59.6°C.

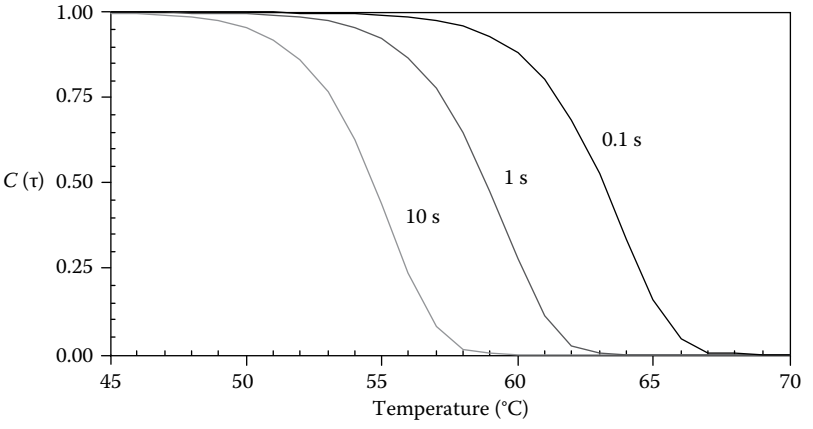


FIGURE 4.5.12 Remaining undamaged concentration vs. temperature at 0.1, 1, and 10 s for the hypothetical example process.

one point on an Arrhenius plot (see above), where $\Omega = 1$. Note the necessity of using exposure times which span several orders of magnitude in order to separate the curves sufficiently to give acceptable accuracy in the determination of the kinetic coefficients.

Though the damage parameter, Ω, cannot be measured directly in histologic section, often a clearer picture of the functional behavior can be obtained from it. The exponential dependence of damage on the inverse of absolute temperature means that the temperature required to obtain comparable damage levels is sensitive to time of exposure. For example, we may define a threshold temperature, T_{TH}, as the temperature at which the damage parameter, Ω, is 1 for a given duration, τ. Assuming a constant temperature exposure, the integral of Equation 4.5.71 reduces to a simple multiplication and the threshold temperature is

$$T_{TH} = \frac{E_a}{R\left[\ln\{\tau\} + \ln\{A\}\right]} \tag{4.5.75}$$

Figure 4.5.13 is a plot of the threshold temperature as a function of duration for the example damage process of Figures 4.5.11 and 4.5.12. Note that threshold temperature is exponentially dependent on

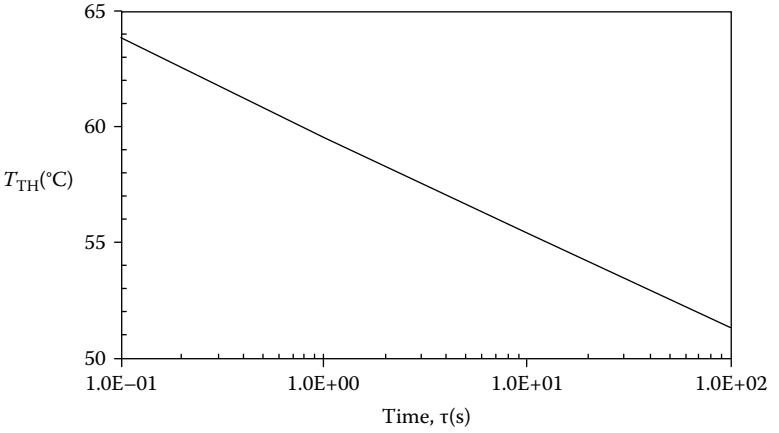


FIGURE 4.5.13 Threshold temperature T_{TH} (where $\Omega = 1$) as a function of duration, τ, for constant temperature exposures for the hypothetical example damage process.

TABLE 4.5.9
Henriques and Moritz Skin Burn Threshold Data ($\Omega = 1$ Means a Second Degree Burn Was Observed)[103]

Exposure Time (s)	Temperature (°C)
25,000	44
11,000	45
5,000	46
2,400	47
1,100	48
570	49
300	50
160	51
90	52
35	54
16	56
5	60
2	65
1	70

duration, as expected, with (in this case) a slope of about −4.58°C per decade. So, while 68.8°C is sufficient to result in $\Omega = 1$ at 0.1 s, 87.1°C would be required for a 1 μs exposure for the hypothetical example process.

Application to the data of Henriques and Moritz[103]—A study of the original skin burn data reported by Henriques and Moritz[100–103] (Table 4.5.9) by Diller and Klutke[127] shows that the fit to a first order process is weak for the higher temperature/shorter exposure time experiments,[127] and that the original values of $A = 3.1 \times 10^{98}$ and $E = 6.27 \times 10^5$ (J/mol)[103] do not fit the data as well as their revised values of $A = 1.3 \times 10^{95}$ and $E = 6.04 \times 10^5$ (for temperatures less than 52°C). In Figure 4.5.14, the skin burn data are also well fit with $A = 8.73 \times 10^{94}$ and $E = 6.03 \times 10^5$ when the last reasonable datum is included (at 52°C).

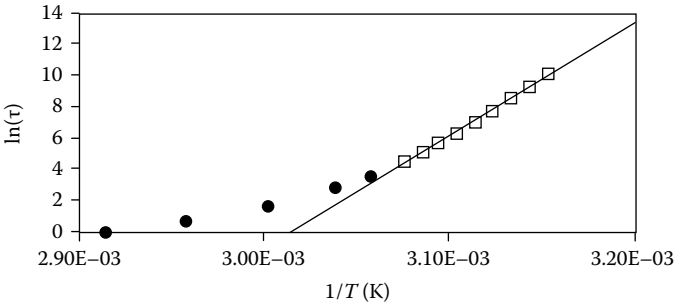


FIGURE 4.5.14 Plot of Henriques and Moritz[107] data on Arrhenius axes with curve fit line for data at 52°C and lower (open diamonds). For the fit $E/R = 72504$ and $\ln\{A\} = 218.61$.

Enzyme Deactivation Model

Xu and Qian[128] presented an enzyme deactivation model for skin burn data that combines some of the first- and second-order aspects into a single calculation. Their method is based on a general simple enzyme-catalyzed reaction sequence from substrate, S, to products, P:

$$E_a + S \underset{k^-}{\overset{k^+}{\rightleftharpoons}} E_aS \xrightarrow{k} E_a + P \tag{4.5.76}$$

and

$$E_a \xrightarrow{k_d} E_i \tag{4.5.77}$$

where here E_a represents activated enzyme, E_i is inactivated denatured enzyme (at velocity k_d), and E_aS is the enzyme–substrate complex. The analysis is completed by assuming that the enzyme inactivation process is slow compared to the main reaction sequence. Nondimensional temperature, z, is used in the analysis, where:

$$z = 1 - \frac{T_0}{T} \tag{4.5.78}$$

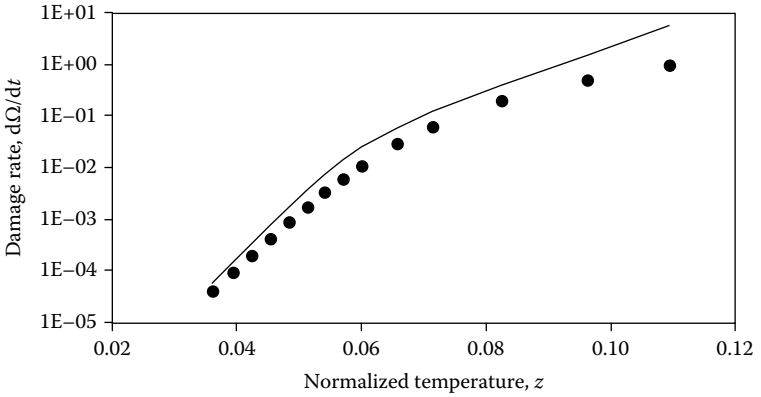


FIGURE 4.5.15 Plot of damage rate for the original Henriques and Moritz data[107] (solid circles) and the enzyme deactivation model of Xu and Qian[128] vs. normalized temperature.

The reference temperature, T_0, as used by Xu and Qian, was representative of the average resting skin temperature: $T_0 = 32.49°C = 305.65$ K.

$$\Omega(\tau) = \ln\left\{\frac{C(0)}{C(\tau)}\right\} = \ln\left\{\frac{E_a(0)}{E_a(\tau)}\right\} = \int_0^{\tau} \frac{Ae^{-\alpha z}}{1 + Be^{-\beta z}}\,dt \tag{4.5.79}$$

This combined formulation fits the data of Henriques and Moritz (Table 4.5.9) well over its entire range if $A = 1.0 \times 10^{-4}$, $\alpha = 100$, $B = 8.0 \times 10^4$, and $\beta = 195$, (see Figure 4.5.15).[128]

Cell Death Processes

Cell death can result in partial or complete death of tissues, organs and, eventually, individuals. The death processes of cells and tissues are divided into two general types: nonprogrammed traumatic cell necrosis and programmed cell death.[136] Nonprogrammed cell necrotic death does not require any premortem genetic activation or expression. It is to be emphasized that any cell can die and undergo postmortem necrosis if presented with a sufficiently severe traumatic injury. On the other hand, programmed cell death processes can only occur in cells that have the specific genes required to initiate and facilitate the mechanisms of cell death.

Nonprogrammed cell and tissue death—Nonprogrammed traumatic necrotic cell death can occur as a result of severe chemical or physical trauma that diffusely and almost instantaneously disrupts all cell function and structure in the targeted volume. As a result, numerous intracellular chemical agents escape into the tissues. These agents include proteases and lipases (digestive enzymes that break down proteins and lipids into small fragments of necrotic debris) and cytokines (small proteins that initiate inflammatory, immune and, later, healing responses to the necrosis). The progression of these postmortem events follows a regular pattern over time (minutes to days) and the resulting necrotic lesions and healing wounds can be described, measured and mapped.

Programmed cell death—Many cells are genetically programmed to die only when a specific chemical or physical stimulus interacts with specific receptors found within or on the surface of the cell. Depending on the receptor type and the cell's genetic programing, these interactions can initiate different cascades of serial catalytic activations of functional proteins, usually enzymes, in the still living cell. However, at some decisive point in a given cascade, a critical activation will occur beyond which the cell will die. Rate analysis and modeling of any cell death process requires the identification of the initiating stimulus/receptor interaction and determination of the critical activation event that leads to the cell death.

Four programmed cell death processes have been identified to date: apoptosis,[99,129] necroptosis,[130,131] autophagy,[132] and pyroptosis.[133] These programmed processes are essential to normal development, maturation and aging, and an individual's response to disease. All consist of signaling, decision-making and execution cascades of functional proteins, usually enzyme catalysts.[99] Catalytic enzymes differ substantially from structural proteins, which derive their function from their molecular formulation and its periodicity.[134] Catalytic enzymes derive their functionality from their geometric conformation: a de-activated enzyme protein is folded in complex ways so that the constituent catalytic segments and binding sites of the long chain protein (which may be separated by many noncatalytic segments) locate closely together in the folded conformation. They may be activated by a slight change in conformation, exposing or assembling the catalytic site.[134]

A comparative description of the processes and their mechanisms was presented by Pearce in 2013.[135] Apoptosis consists of an ordered cleavage of the cell cytoskeleton, DNA, and organelles with bits of nuclear and cytosolic debris completely contained within so-called blebs in the plasma membrane. These "blebs" bud off the cell surface forming "apoptotic bodies" that are soon (on the order of hours) ingested by adjacent phagocytic cells. Apoptosis is identified by the activation of various caspase (i.e., cystic asp.artic protease) enzymes in an intricate and convoluted cascade.[99,129] Additional markers of apoptosis include the microscopic identification

of "apoptotic bodies" or demonstration of endonuclease DNA disruption by gel electrophoresis. The plasma membrane-bound blebs expose "eat me" signals when the normally inactive cytosolic leaflet enzyme "scramblease" is activated and the normally asymmetric membrane proteins distribute uniformly. The symmetrical membrane is a signal for macrophages to consume the bleb. Necroptosis is a sort of ordered cellular explosion marked by receptor interacting proteins (RIP1, RIP3, etc.) in which the cytosolic components are directly released into the extracellular space as random debris.[130,131] In autophagy, malformed or aged proteins and organelles are sequestered into endosomes in the cell cytoplasm. If not too damaged, some of the proteins are repaired. The senescent organelles are digested and compacted into larger, membrane-bound cytoplasmic bodies, the autophagosome, for storage within the cell. Some investigators feel that, in severe cases, the cell may sacrifice itself.[132] However, others (mainly pathologists) think that cell death specifically caused by autophagocytosis has yet to be proven.[136] Pyroptosis is cell death in response to binding of circulating pyrogens–pyrogens, such as interleukins (IL-1, IL-6, etc.) that are produced by host inflammatory and immune cells in response to infections by microorganisms such as pathogenic bacteria, fungi, or viruses.[133]

All of the programmed cell death processes consist of a signaling block, a decision-making block and an executioner block of functional proteins. The best-understood among these processes is apoptosis, so it will be used to illustrate some of the complexity. Apoptosis is of prime interest in the research community because of its implications in embryologic development and maturation, tissue bioengineering, cancer treatment, HIV, and autoimmune diseases. There are tens of thousands of references on apoptosis, the vast majority of which are within the past 15 years. In 2004, Eissing et al.[137] described the decision-making block in apoptosis, including many of the inhibitors and promoters of apoptosis, in an elegant dynamic state-space model. The decision-making block is a bistable switch that introduces a variable delay between signaling and execution depending on the strength of the input signal. Delays between a few minutes and several days may be observed. There are both signaling and executioner caspases among the 14 or so that have been identified. Caspases are numbered in the sequence of their discovery and exist to enzymatically cleave other inactive proenzymes to activated enzymes. Activation of a caspase hetero-tetramer consists of combining the cleaved long and short segments of two pro-caspase molecules after cleaving the *N*-terminus and *C*-terminus.[99]

In Eissing's biochemical model, the initiator (i.e., input signal) caspases (C2, C8, C9, C10, etc.) are represented by activated caspase 8, $C8^*$. There are, of course, many other input signals; for example p52, the "guardian of the genome" and death receptors like FADD (Fas-associated death domain).[129] The executioner caspases (C3, C6, C7, etc.) are represented by activated caspase 3, $C3^*$. Caspase 9 has somewhat of a dual role in that in addition to signaling it participates as the hub of an "apoptosome", a large flower-shaped molecule, with the protein APAF1 forming "petals" tipped with cytochrome *C*.[99,129] There is also intimate coupling between apoptosis and metabolism.[138] Cytochrome *C*, in turn, is a critical participant in the final stages of oxidative phosphorylation on the outer leaflet of the inner membrane of mitochondria, where it is responsible for electron transport between complexes III and IV, immediately upstream of ATP synthase, the final step.[99,129] Presumably, cytochrome *C* plays a similar role at the tips of the apoptosome petals.

The feedback pathways in the decision-making block as modeled by Eissing et al.[137] includes inhibitors of apoptosis proteins (IAPs) and bifunctional apoptosis regulators (BARs), most of which belong to the Bcl2 family.[99,129] This family includes promoters and inhibitors of apoptosis, as well as inhibitors of promoters and inhibitors of inhibitors. The process is an intricate ballet, as it were, in the protein cascade. The feedback system is sketched in Figure 4.5.16. Bid is one of the Bcl2 family, t-Bid is the truncated form, and SMAC is "second mitochondria-derived activator of caspases" (also called DIABLO).

Central to the reaction system describing the feedback pathways are association and dissociation reactions, which are common in all biochemical systems. $C3^*$ is held inactive by binding

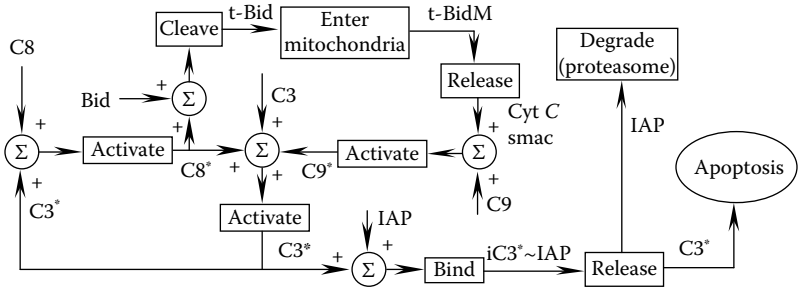


FIGURE 4.5.16 Feedback pathways in the decision block system. (From Pearce, J.A., *Int. J. Hyperthermia*, 29(4), 262, 2013. With permission.)

to IAP, for example, as in *i*C3*~IAP. The four overall governing "law of mass action" reactions are as follows:

$$C8^* + C3 \rightarrow C8^* + C3^* \tag{4.5.80a}$$

$$C8 + C3^* \rightarrow C8^* + C3^* \tag{4.5.80b}$$

$$C3^* + IAP \leftrightarrow iC3^* \sim IAP \tag{4.5.80c}$$

$$C3^* + IAP \rightarrow C3^* \tag{4.5.80d}$$

These four reactions are, in turn, described by a system of eight dynamic state equations. Each of the association and dissociation reactions is a nonlinear component of the state-space system of equations, as presented in the following equations:

$$\frac{d[C8]}{dt} = -k_2[C3^*][C8] - k_9[C8] + k_{-9} \tag{4.5.81a}$$

$$\frac{d[C8^*]}{dt} = -k_2[C3^*][C8] + k_5[C8^*]k_{11}[C8^*][BAR] \tag{4.5.81b}$$

$$\frac{d[C3]}{dt} = -k_1[C8^*][C3] - k_{10}[C3] + k_{-10} \tag{4.5.81c}$$

$$\frac{d[C3^*]}{dt} = k_1[C8^*][C3] - k_3[C3^*][IAP] + k_{-3}[iC3^* \sim IAP] - k_6[C3^*] \tag{4.5.81d}$$

$$\frac{d[IAP]}{dt} = -k_3[C3^*][IAP] + k_{-3}[iC3^* \sim IAP] - k_4[C3^*][IAP] - k_8[IAP] + k_{-8} \tag{4.5.81e}$$

$$\frac{d[C3 \sim IAP]}{dt} = k_3[C3^*][IAP] - k_{-3}[iC3^* \sim IAP] - k_7[iC3^* \sim IAP] \tag{4.5.81f}$$

$$\frac{d[BAR]}{dt} = -k_{11}[C8^*][BAR] + k_{-11}[iC8^* \sim BAR] - k_{12}[BAR] + k_{-12} \tag{4.5.81g}$$

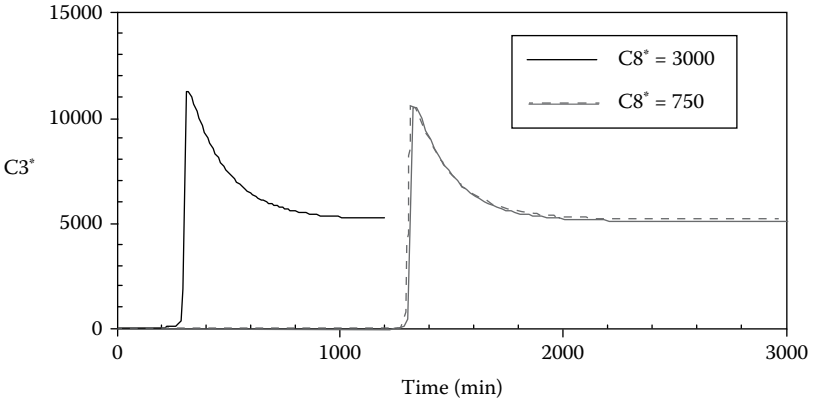


FIGURE 4.5.17 Dynamic response of the biochemical model to two different strengths of input signal, activated caspase 8, C8*, in molecules per cell. The delay in apoptosis varies from 250 to 1300 min (4 to 22 h). (From Pearce, J.A., *Int. J. Hyperthermia*, 29(4), 262, 2013.)

$$\frac{\mathrm{d}\left[C8^{*} \sim \mathrm{BAR}\right]}{\mathrm{d}t} = k_{11}\left[\mathrm{C8}^{*}\right]\left[\mathrm{BAR}\right] - k_{-11}\left[i\mathrm{C8}^{*} \sim \mathrm{BAR}\right] - k_{13}\left[i\mathrm{C8}^{*} \sim \mathrm{BAR}\right] \quad (4.5.81\mathrm{h})$$

where k_i are the reaction velocities, $+i$ for forward reactions, and $-i$ for reverse reactions. The association reactions, negative nonlinear terms, decrease the concentration, while the dissociation reactions, positive nonlinear terms, increase the concentration. The net effect of the nonlinear terms is to cause a variable delay in apoptosis. Figure 4.5.17 illustrates the variable delay in a COPASI model of equations 4.5.81 for inputs of 750 and 3000 molecules/cell of activated caspase 8, C8*. In practice, there may be a delay of as much as a few days in the appearance of apoptosis. However, in health and disease of intact multicellular individuals, biochemical and/or morphological evidence of apoptosis occurs and disappears within 72 h of a single initiating event. But life and death are complicated and numerous waves of programmed cell death may continue as new initiating signals including those for necroptosis and pyroptosis are introduced.

In terms of thermally initiated apoptosis, we might reasonably expect that each of the reaction velocities, k_i, in the system would follow Arrhenius kinetics. If so, we expect the reactions with the higher activation energies would initiate the process since the thermal sensitivity, $\partial k/\partial T$, (ignoring the dependence of A on T) is approximately given by

$$\frac{\partial k}{\partial T} = \left[\frac{E_\mathrm{a}}{RT^2}\right] A e^{\left[\frac{-E_\mathrm{a}}{RT}\right]} \quad (4.5.82)$$

At present, the activation energies of the reaction velocities used by Eissing et al.[137] are not available, however.

The state-space model describes a single cell. In their paper, the authors extend the study by applying a distribution of input signal strengths to simulate a colony of cells. The authors point out that their work was the first attempt to couple an individual cellular response to a colony of cells (Figure 4.5.18).

Modeling Cell Death Processes

The other intrinsic cell death mechanisms are likely to be governed by similarly complex functional protein cascades. Plainly, simpler models are required for practical numerical models. The first-order Arrhenius models are much simpler, but fail to describe these intrinsic processes because they lack sufficient mathematical sophistication (Figure 4.5.19). Arrhenius models work well on structural proteins, and on cell death at higher temperatures, where the intrinsic processes are overwhelmed

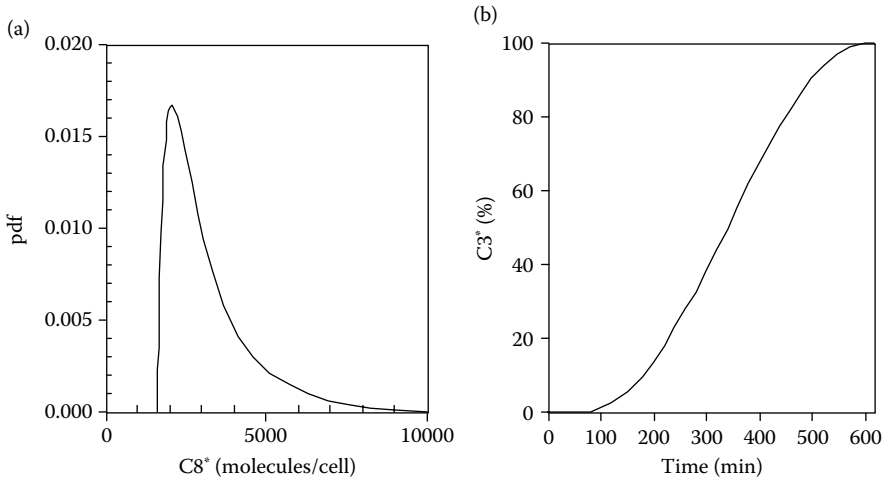


FIGURE 4.5.18. Stochastic response of the biochemical model to distributed input signals, activated caspase 8, C8*, in molecules per cell. (a) The probability density function (pdf) as applied, and (b) the resulting distribution of apoptosis, which may be interpreted as the inverse of a cell survival curve. (From Pearce, J.A., *Int. J. Hyperthermia*, 29(4), 262, 2013. With permission.)

by the applied thermal stress. They fail miserably at hyperthermic temperatures mainly because they initiate immediately and do not have a delay feature. The shoulder region, which typifies many cell survival curves, is bypassed and an unrealistic overestimation of "success" in the form of the irreversible alteration measured results, as in Figure 4.5.19. In the figure, the measured loss of clonogenicity curve at 43°C for Chinese hamster ovary (CHO) cells is compared to the Arrhenius model of equation 4.5.71 using coefficients derived from the constant rate region of the curves. Loss of clonogenicity, strictly speaking, isn't necessarily a cell death process; the cells may simply remain in the G1 phase (i.e., "resting"). However, for cancer cells, such as these are, it amounts to suppressing proliferation and makes a good measure. Since the Arrhenius model initiates immediately without delay, by that measure the "success" is grossly overestimated, a thoroughly unacceptable result.

What is missing, of course, is an appropriate delay. It turns out that by adding a temperature-dependent delay before initiating a first-order decline much more accurate results can be obtained.[140]

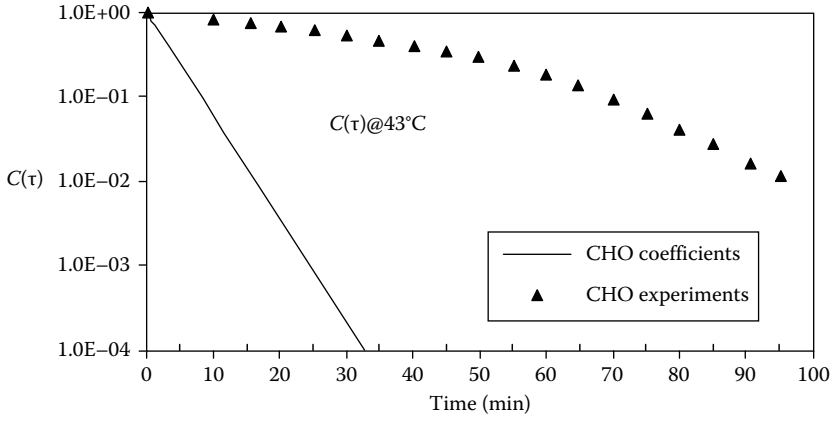


FIGURE 4.5.19 Arrhenius model predictions compared to the cell survival curve at 43°C for the CHO cell loss of clonogenicity from Sapareto and Dewey data.[139] (From Pearce, J.A., *Int. J. Hyperthermia*, 29(4), 262, 2013. With permission.)

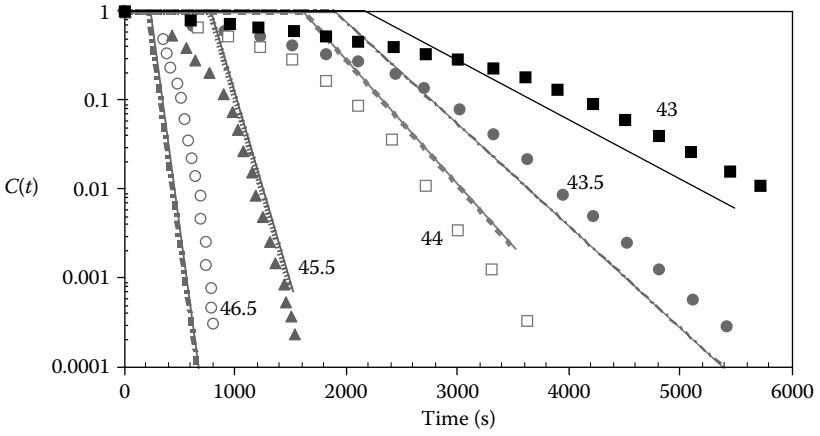


FIGURE 4.5.20 Arrhenius model with temperature-dependent time delay predictions compared to cell survival curves for the CHO loss of clonogenicity from Sapareto and Dewey data.[140] (From Pearce, J.A., *J. Biomech. Eng.*, 137(6), 121006, 2015.)

Looking at the CHO data with an added time delay improves the accuracy substantially, as in Figure 4.5.20. The corrected Arrhenius coefficients for the constant rate region are: $A = 6.355 \times 10^{97}$ and $E_a = 6.102 \times 10^5$ (J/mol). The required time delay from a regression line is given by[140]

$$t_d\,(\text{s}) = 26062 - 555.9T\ (^\circ\text{C}) \quad r^2 = 0.912 \quad T_{max} = 46.9^\circ\text{C} \tag{4.5.83}$$

and the maximum temperature for which a time delay is needed is 46.9°C. The improvement in accuracy is notable, and realistic predictions can now be made with a relatively simple model.

Adding a time delay to the Arrhenius model gives an excellent representation, and thus prediction, of the measurements.

The time delay model also fits other experimental survival curve data with a shoulder region quite well. For example, in a carefully constructed series of experiments on PC3 cells, Feng et al. used propidium iodide (PI) uptake as the cell death measure.[141] PI stains cells that have damaged plasma membranes. The survival curves have a pronounced shoulder region. The required time delay is[140]

$$t_d\ (\text{s}) = 2703 - 49.6T\ (^\circ\text{C}) \quad r^2 = 0.953 \quad T_{max} = 54.5^\circ\text{C} \tag{4.5.84}$$

Elegant and instructive studies of several cell types illustrate that not all death mechanisms require time delays. In 2004, Bhowmick et al. studied the AT1 subline of R3327 benign rat prostate tumor cells with three types of thermal damage assays: calcein fluorescence intensity decay, PI exclusion, and loss of clonogenicity.[142] Of these, only calcein fluorescence decay and loss of clonogenicity and required an added time delay.[140]

AT1 cell calcein fluorescence decay ($E_a = 2.85 \times 10^4$ J/mol and $A = 246.9$ s^{-1}):

$$t_d\,(\text{s}) = 1304 - 18.95T\ (^\circ\text{C}) \quad r^2 = 0.937 \quad T_{max} = 68.8^\circ\text{C} \tag{4.5.85}$$

AT1 cell loss of clonogenicity ($E_a = 5.21 \times 10^5$ J/mol and $A = 5.95 \times 10^{82}$ s^{-1}):

$$t_d\,(\text{s}) = 2436 - 46.1T\ (^\circ\text{C}) \quad r^2 = 0.750 \quad T_{max} = 52.8^\circ\text{C} \tag{4.5.86}$$

The PI uptake studies in AT1 cells were fit very well by the standard Arrhenius model with $E_a = 2.448 \times 10^5$ (J/mol) and $A = 2.99 \times 10^{37}$ (s^{-1}). In another study by the same group, PI uptake was

measured in SN12 human renal carcinoma cells.[143] The standard Arrhenius model also fit the measured data well with no noticeable shoulder region, $E_a = 2.88 \times 10^5$ (J/mol) and $A = 4.46 \times 10^{43}$ (s^{-1}).

These observations present two interesting conundrums. First, the same assay, PI uptake, behaves differently in PC3 (where a time delay is required), in AT1 and in SN12 cells where a standard Arrhenius model fits well. Second, the AT1 cells respond differently to the assays: a time delay is required for calcein fluorescence, but not for PI uptake—both of which assays indicate plasma membrane damage.

In summary, cell survival curves that demonstrate an initial slowly developing shoulder region may be fit by adding a time delay to the standard Arrhenius model.[140] The required time delay is temperature-dependent, and crosses zero typically at some temperature in the range of 46°C–70°C, depending on cell type and particular damage assay. This is in keeping with published observations that the Arrhenius model fits well at higher temperatures. The maximum delay temperature is an indication of the temperature at which intrinsic cell death processes and/or healing responses (such as those involving heat shock proteins) are overwhelmed by the applied thermal stress. Other assays in the same cell line may not require a time delay for adequate accuracy. Of course, this approach discards information in the shoulder region; however, when heating tumors for therapeutic goals the shoulder information is of no interest—except that, of course, the delay time at temperature must be exceeded in order to hope for an effective treatment.

Histologic Markers of Thermal Damage*

Tissue and cell heating produce several pathologic alterations that can be directly attributed to thermal damage mechanisms. Many of these are measurable using quantitative pathologic techniques, and thus are amenable to description by kinetic rate process models. As discussed previously, it is important to recognize that several independent damage processes can be occurring at the same time. Therefore, it is imperative to describe precisely which particular end point will be measured for kinetic analysis. Depending on the tissue, some end points are more easily identified and measured than others. From the pragmatic biological standpoint, cellular and tissue thermal damage end points are divided into two general categories: lethal and nonlethal thermal injury. Details are presented and illustrated by Thomsen and Pearce in several extensive reviews.[144–146]

Nonlethal Low-Temperature Tissue Effects

Most living cells and tissues can tolerate and survive modest temperature elevations for limited time periods depending on the species and the metabolic status of the individual. Recovery after heating is marked by the restoration of normal functions. Most nonlethal thermal injuries of individual cells are secondary either to (1) heat-induced acceleration of metabolism, (2) thermal inactivation of particular enzymes, and/or (3) rupture of cellular membranes. These alterations lead to disruption of several physiologic reactions of which the most critical are respiration and energy-generating metabolic functions.[136,147] The best indicators of reversible, nonlethal thermal injury in the living cell or organism are physiological tests that monitor heat-sensitive biochemical and metabolic changes rather than morphologic alterations.

Immediate effects—Routine light microscopic techniques applied immediately after heating are not useful for the detection of histologic markers of nonlethal thermal injury. However, enzyme histochemistry, a laborious, somewhat capricious methodology, can be used to reveal inactivation of a few heat-labile enzymes.[148] Transmission electron microscopy (TEM) provides sufficient magnification and resolution to show ruptured cellular membranes and distorted organelles; however, it is not quantitative, but subjective, and interrogates only a small tissue region (<1 mm^2). If the ruptures and distortions are infrequently seen, then the damage is not likely lethal. Over time, recovery

*The authors acknowledge the contributions to this subsection by Sharon Thomsen, M.D.; portions excerpted from Pearce, J.A. and Thomsen, S., in Optical-Thermal Response of Laser-Irradiated Tissue, Welch, A.J. and van Germert, M.M.C., eds., Plenum Press, New York, 1995. With permission.

occurs with the synthesis of replacement enzyme proteins and new membranes to patch the holes and reconstitute the normal functions of the organelle and cell.

On the other hand, tissues and organs are composed of a diverse population of cells that have (1) varying levels of heat sensitivity, (2) different types of metabolic functions and repair capabilities, and (3) various contributions to the function of the tissue or organ. Unlike single cells, reversible thermal damage of tissues can and usually does involve some cell death, but it occurs at a level that does not impair the function of the organ or the tissues. Like the situation in single cells, histologic assessment of reversible tissue thermal injury immediately after heating requires enzyme histochemistry and TEM to determine the degree and distribution of thermal damage, because this damage is not detectable with routine light microscopic techniques. The immunofluorescent microscopy based TUNEL (Terminal deoxynucleotidyl transferase dUTP nick end labeling) histochemical reaction reveals endonuclease cleaved ends of DNA but this reaction is fraught with false positive reactions in necrotic lesions.[99]

Delayed effects—Depending on the tissue, intracellular edema, tissue edema, and hyperemia occur within seconds to hours. Delayed effects can only be seen in living, physiologically intact organisms. These are easily identified delayed organismal responses to nonlethal thermal damage in both living single cells and tissues. Intracellular edema is due to abnormal accumulation of fluids secondary to thermally induced metabolic dysfunction, and resolution of the edema upon repair is a marker of healing. Heat-induced tissue edema and transient hyperemia (increased blood flow and blood vessel dilatation) are mediated by the release of vasoactive polypeptides from local inflammatory cells within several seconds of injury. The action of the polypeptides causes blood fluids (primarily plasma) to escape through gaps between the endothelial cells lining the vessel. On the other hand, prolonged hyperemia, like that associated with sunburn, is a delayed response associated with direct endothelial damage. In this case, the histologic indicators of the more severe vascular damage are dilated blood vessels stuffed with red blood cells and, not infrequently, microscopic leakage of the red blood cells through the damaged vessel wall into the adjacent tissue spaces. Recovery, which is secondary to endothelial regeneration and repair, is seen in a few days.[136]

Lethal Low-Temperature Thermal Effects

Nonprogrammed cell and tissue death and postmortem necrosis—Nonprogrammed cell death and subsequent necrosis result when the damage is so severe that the usual repair mechanisms cannot cope and/or the mediators of the repair mechanisms (DNA and RNA transcription enzymes) are thermally destroyed. Again, routine light microscopic techniques cannot be used to identify dead cells immediately after heating, except in the case of severe damage. However, within minutes of lethal thermal damage, TEM reveals the accumulation of chromatin (the DNA-containing chromosomes) at the margins of the nuclei, a recognized TEM marker of cell death, plus more extensive disruption of cellular organelles and limiting cell membranes that are incompatible with survival.[148] Enzyme histochemistry demonstrates enzyme inactivation but, as discussed above, enzyme inactivation by itself cannot be used as a hallmark of cell death. After death, cells and tissues undergo necrosis, which is a natural autolytic process of disintegration that follows a predictable course over time. Tissue necrosis, the "gold standard" marker of death, is easily recognized in light microscopic sections of lesions 24–72 h after the lethal event and, for most tissues, histologic evaluation of the maximal extent of tissue necrosis is seen at about 3–5 days.[149–151] After 5 days, the boundaries of necrosis and other markers of thermal coagulation are obscured by the healing processes and precise measurement of the extent of lethal heating can no longer be done.

In vivo red zone thermal damage—When localized electromagnetic, acoustic, or optical heating results in deep zones of necrosis, as when laser light is delivered through an optical fiber or ultrasound or electromagnetic waves are focused, concentric zones of thermal damage form around the source during heating and immediately thereafter. The boundaries between these zones are often distinct and measurable. Depending on the power density and duration of exposure, the three major zones in the lesion which can be seen by the naked eye are (1) a central ablation hole (at high-power

densities), (2) a whitish zone of coagulation, and (3) when the blood supply is intact, a surrounding peripheral red zone. The red zone is formed as a result of hemostasis, hemorrhage, and hyperemia (increased blood flow). The vascular responses are intermixed; however, hemostasis—due to direct damage to blood cells and vessels—tends to occur at the inner boundary between the surrounding red zone and the more central white coagulation zone. Studies performed in rat liver and goat breast show that the outer boundary of the red zone corresponds to the maximum extent of lethal thermal damage as determined by tissue necrosis evaluated in surviving animals 3 days after treatment.[145,149,150]

Microscopically, additional distinct zones of lethal thermal damage can be detected in the white coagulum. Advancing radially inward from the boundary between the red and white zones, one observes the following quantitative histopathological markers of thermal damage:

1. Lethal cellular injury manifested by cell shrinkage, spindling, and hyperchromasia
2. Collagen hyalinization
3. Collagen birefringence changes.
4. Muscle birefringence changes
5. Water-dominated effects

The changes intensify as the hotter center of the lesion is approached. Additionally, a mass defect created by ablation of the tissue by vaporization and plasma formation is observed in cases of intense heating, but is not considered here.

Thermal Coagulation

As temperatures rise and/or heating times are prolonged, cellular, and tissue structural proteins, which are thermally more stable than the vital, energy-producing enzymes, undergo denaturization and conformational changes—a process defined as thermal coagulation. Coagulation is immediately apparent and always indicates lethal thermal effect. For most tissues, coagulation can be seen with the naked eye as whitening of the tissue associated with increased turgor and opacity. The whitening of the egg white while the egg is being fried is an obvious example of thermal coagulation of protein, in this case, albumin. On the other hand, coagulation of collagen-rich tissues, such as tendon and skin dermis, can be signaled by increasing transparency due, apparently, to decreased optical scattering by the denatured collagen fibrils.

Microscopically, thermal coagulation of cellular and extracellular structural proteins includes an array of morphological alterations that mark a large range of temperatures and exposure times. As tissue temperatures and/or exposure times increase, the coagulative changes become more obvious and light microscopic analysis becomes the investigative tool of choice.

The most useful histologic markers of coagulative thermal damage in tissues are structural alterations of cells and collagens. Thermally coagulated cells and intracellular organelles shrink and undergo characteristic conformational changes. Collagens are a widely dispersed biochemical class of extracellular fibrous proteins that form the supporting scaffolding of nearly all soft tissues and the organic strut system of bone, cartilage, and teeth. These fibrillar proteins swell and form amorphous masses as a unique response to heat.

Cellular shrinkage and conformational changes—Thermally coagulated cells shrink due to denaturation of the elongate cytoplasmic and nuclear proteins that form the three-dimensional scaffolding system of these structures.[144,152] Coagulated surface epithelial cells—such as those that line the gut, bladder, and numerous glands of the body—tend to become elongate (spindle-shaped). The boundary between this change and the normal epithelium can be quite distinct, and thus potentially useful for kinetic analysis.

In contrast, the epithelial cells of the solid organs—such as the liver, pancreas, and kidney—just shrink and do not undergo noticeable conformational change at the light microscopic level. The cellular shrinkage is also due to loss of intracellular water. The coagulated proteins of the desiccated

cells become more densely packed, as reflected by dark cytoplasmic and nuclear staining (hyperchromasia) in routinely prepared histologic sections. Unfortunately, these histologic changes are very subtle and, frequently, cannot be distinguished from technical artifacts due to poor fixation. Therefore, they are not reliable markers of thermal damage, especially in the hands of an inexperienced observer.

Collagen coagulation—Thermal denaturization of extracellular fibrous structural proteins, the collagens, is marked histologically by swelling and an amorphous, glassy transformation of the fibers (hyalinization).[146,149,150,152,153] This transformation is easily seen in those tissues composed of loose networks of collagen fibers, such as the cornea, skin, submucosa of the stomach, intestines, and urinary bladder. However, some collagen fibers, such as those found in arterial media and liver sinusoids, are very thin (reticulin fibers); thus detection of hyalinization is beyond the resolution of the light microscope. And, at the other end of the fiber size spectrum, mild hyalinization and fiber swelling are equally difficult to see in tissues composed of thick, densely packed collagen fibers, such as those found in blood vessel adventitia and dermis. Changes of birefringence of thermally coagulated collagen (described next) are sometimes easier to detect than hyalinization in these circumstances.

Thermally Induced Loss of Birefringence

Birefringence is an optical property of some tissues that rotates incident polarized light. Thermal coagulation is associated with partial and total loss of the native birefringence of muscle tissues and certain fibrillar collagens as revealed by transmission polarizing microscopy (TPM).[144–146,149,153] The birefringence of muscle is due to the very regular arrangement of the fibrillar contractile protein macromolecules, actin and myosin, that form the contractile unit of the muscle. Collagen birefringence is the result of the molecular structure and regimentation of the tropocollagen molecules from which collagen fibrils are built. Tissue birefringence has two components, intrinsic and form birefringence. Intrinsic birefringence is secondary to the α-helical conformation of the polypeptides that form the individual contractile proteins of muscle and tropocollagen molecules and accounts for 12%–30% of the total birefringence of these tissues. The remaining birefringence, form birefringence, is related to the crystalline-like array of the actin and myosin molecules in the sarcomere and the regimented longitudinal array of tropocollagens within the collagen fibrils.[144–146,153–167]

Thermally induced changes of muscle birefringence are associated with dissociation and disruption of the molecules, as revealed by TEM. Partial loss of birefringence in skeletal and cardiac muscle has been identified with the breakdown of the relationships of actin and myosin in the sarcomere (loss of form birefringence). The source of the birefringence of striated muscle is the central portion of the sarcomere, the A or anisotropic band, in which the actin and myosin are arranged in parallel crystalline-like array. The I or isotropic band of the sarcomere includes portions of adjacent sarcomeres that contain actin molecules which are attached to the Z band, an attachment zone for the actin molecules. Total loss of birefringence is associated with increased temperature–time histories and is probably related to the reduction of the individual fibrillar contractile proteins to thermally denatured granular profiles (loss of intrinsic birefringence) as shown in TEM of heated myocardium.[95,168,169]

Thermally induced birefringence changes in collagen involve decreasing intensity of the birefringence image until it is totally lost at higher temperatures. TEM of heated collagens shows a gradual unraveling of the fibers with disappearance of the characteristic periodicity of collagen, reflecting the dissolution of the regimented organization of native collagen as it is heated.[149,167]

The onset of the birefringence changes of collagens varies depending on the experimental conditions, anatomic site, and the age of the organism. Native tissue collagens reside in complex environments of different proteins, glycoproteins, and cells that are greatly influenced by the local and systemic physiology of the living organism. The intimate relationships of the collagens with these tissue constituents are regulated by water and salt concentrations, ionic and covalent bonds, and hormonal and metabolic effects. Collagens extracted from tissues are not in their natural environment,

therefore *in vitro* experiments testing molecular relationships including measurements of birefringence will not predict the native behavior of collagen *in situ* or *in vivo*.[163–165] In addition, as the organism ages, the collagens, especially the Type I collagen found in most tissues, form more molecular cross-links that alter the temperature-time thresholds of fiber dissociation.[164]

TPM can be performed on paraffin sections prepared for routine light microscopy; thus the birefringent images can be correlated directly to the morphologic structure. The intensity of the birefringent images is governed by the thickness of the section, the type of dyes used, and in the case of collagens, the collagen fiber thickness.[150,170]

Collagen Shrinkage

As collagen is heated to denaturing temperatures at which it loses its rope-like native state conformation, it shrinks in length and swells in diameter in an approximately iso-volumic process. An interesting model of the collagen shrinkage process was presented by Chen et al.[171–173] They studied thermally-induced shrinkage in rat *chordae tendineae* between 65°C and 95°C under different levels of applied longitudinal stress between 0 and 650 kPa, and were able to collapse all of their data into a single curve based on a normalized time axis, $\nu = \ln\{t/\tau_2\}$—τ_2 is described in Figure 4.5.22 and Equation 4.5.89. Three nondimensional variables were defined to analyze the data.

The overall shrinkage, ξ (%), is calculated from an interpolation function, $f(\nu)$:

$$\xi = \left(1 - f\left(\nu\right)\right)\left[a_0 + a_1\nu\right] + f\left(\nu\right)\left[b_0 + b_1\nu\right] \tag{4.5.87}$$

where $a_0 = 1.80 \pm 2.25$, $a_1 = 0.983 \pm 0.937$, $b_0 = 42.4 \pm 2.94$, and $b_1 = 3.17 \pm 0.47$ (all in %). The best-fit interpolation function, $f(\nu)$, is given by

$$f\left(\nu\right) = \frac{e^{a(\nu - \nu_m)}}{1 + e^{a(\nu - \nu_m)}} \tag{4.5.88}$$

where $a = 2.48 \pm 0.438$ and $\nu_m = \ln\{\tau_1/\tau_2\} = -0.77 \pm 0.26$. The interpolation function includes all three segments of the shrinkage curve in Figure 4.5.20, excepting the post-heating relaxation curve. And finally, at any temperature, τ_2 is given by

$$\tau_2 = e^{\left[\alpha + \beta P + M/T\right]} \tag{4.5.89}$$

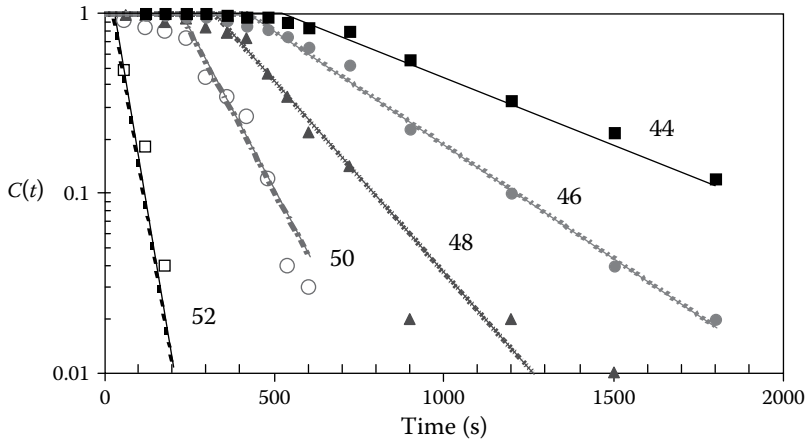


FIGURE 4.5.21 Arrhenius model with temperature-dependent time delay predictions compared to cell survival curves for the PC3 PI uptake assay.[140] (From Pearce, J.A., *J. Biomech. Eng.*, 137(6), 121006, 2015.)

where $\alpha = -152.35$, $\beta = 0.0109$ (kPa^{-1}), P = applied stress (kPa), and $M = 53{,}256$ (K). The exponent in Equation 4.5.89 contains the thermal kinetics of the process: $\alpha = -\ln\{A\}$ and $M = E_a/R$, so the equivalent $A = 1.461 \times 10^{66}$ and $E_a = 4.345 \times 10^5$ (J/mol) for the process (i.e., $T_{crit} = 343$ K = 69.9°C). The signs are reversed compared to Equation 4.5.56 because τ_2 occurs in the denominator of the kinetic equation.

To use the collagen shrinkage model, the equivalent t/τ_2 is accumulated during model execution. That is, at each point in space and time an equivalent value for the increment, $(\Delta t/\tau_2)$ is calculated and accumulated from Equation 4.5.89. Point-specific shrinkage is calculated at selected times by application of Equations 4.5.87 and 4.5.88, respectively.

Water-Dominated Effects

The role of water vaporization becomes dominant as the tissue temperatures approach 100°C. The tissue effects depend on (1) the rate of water vapor formation, (2) the temperature of the water vapor, (3) the rate of diffusion (escape) of the vapor from the tissue, and (4) the tissue mechanical properties.[174–183] All tissues exposed to air will lose water vapor to the atmosphere because of the differences between the partial pressure of water vapor across the tissue/air boundary. Below 100°C water vaporization is a surface phenomenon. As tissues are heated water vapor will be generated and the vapor will diffuse toward the surface to escape. The water vapor is in thermodynamic equilibrium with liquid water and, in the tissues deep to the surface, the relatively low partial pressure gradients favor the liquid over the vapor phase. Meanwhile, at the surface, the tissues dry out as the water vapor escapes leaving the hard, not infrequently brittle, solid tissue components behind.

At or just above 100°C water vapor is generated volumetrically, equilibrium is pushed toward the vapor phase, and, for sufficiently high volume generation rates, q''', more vapor is produced than can escape by simple diffusion. The excess vapor is trapped not only within the cell but also in the extracellular tissue layers forming vacuoles or dissections along planes of mechanical weakness.[183–185] The surface desiccates rapidly because diffusion lengths are short. The temperatures of the tissues that still contain liquid water will either remain close to 100°C or the pressure will rise.

The vapor in the vacuoles will expand quickly compressing the surrounding, rapidly drying tissues that form the vacuolar walls. As the vacuoles expand the walls separating the vapor pockets from each other or the tissue surface become thin. The walls rupture as the force of the increasing pressures of expanding vapor overcomes the mechanical strength of the tissue. The vacuoles coalesce to become larger holes within the tissues or form irregular defects on the surface. When this process takes place over relatively long heating times, the tissue defects are formed as the result of tissue compression, desiccation, and shrinkage, but not to loss of dry mass. On the other hand, rapid heating times result in explosive rupture which causes tissue fragments to be ejected from the surface creating irregular craters formed secondary to dry tissue mass loss. These water-dominated events are the same as those when popcorn is made, hence the tissue changes have been called the "popcorn" effect.[175–179,184,186]

The kinetics of water vaporization will not be treated here; see the discussion of phase change in Chapter 3. Histologically, the zone of water vapor formation manifests as a distinct region populated by vacuoles adjacent to ablation craters formed from volume heat sources. In some experiments, for example, electrosurgical cutting and laser ablation at water-absorbed wavelengths, the vacuolized zone usually has a distinct and measurable border.

Thermal blood vessel sealing—The thermodynamics of tissue water vaporization play an exceedingly important role in RF sealing of large vessels. Hyalinization (gellification) of tissue collagen—which typically occurs between about 60°C and 90°C, depending on heating time—represents the first step in the vessel sealing process. It turns out to be necessary to also thermally denature the vascular elastin, a dominant structural protein in large arteries and in many of the larger veins. This frequently requires temperatures in excess of 140°C, and under ordinary bipolar forceps electrodes explosive boiling destroys tissue structure resulting in a brittle, friable, and unsatisfactory seal. To get a successful seal it is necessary to apply high apposition pressures—at least

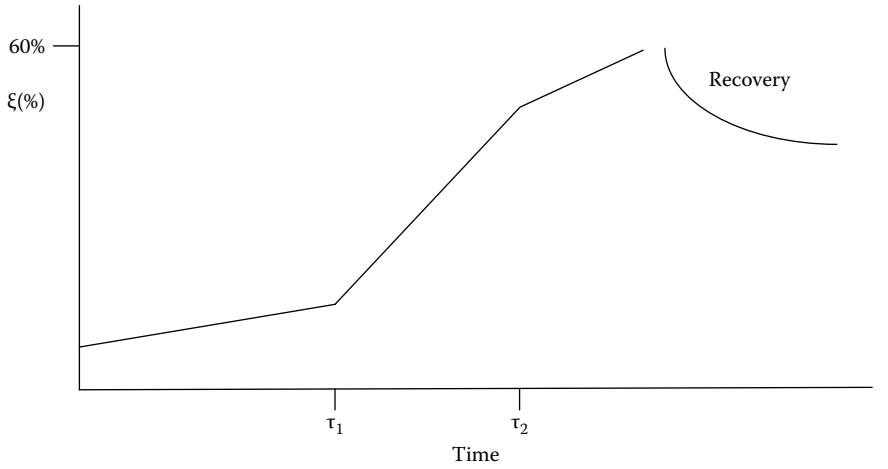


FIGURE 4.5.22 Collagen shrinkage: slow initial shrinkage up to τ_1 followed by rapid shrinkage up to τ_2 and slower shrinkage to a maximum of about 60%. Some relaxation occurs during cooling.

several atmospheres—to raise the equilibrium boiling temperature above the elastin denaturation temperature range. As a result, substantial tissue water is retained, the seal is more pliable, and able to withstand higher burst pressures.

Figure 4.5.23 illustrates the complex anatomy of larger vessels, which is the origin of the vaporization/disruption problem. In Figure 4.5.23a, the cross section of a canine carotid artery at 40× original magnification is shown under Weigert–van Gieson elastin stain. The elastin plates are black, smooth muscle is tan and the adventitial collagen is pink. The internal elastic lamina is at the

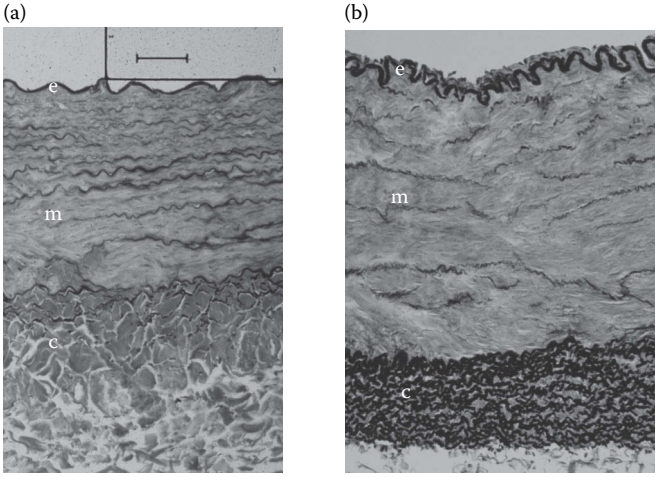


FIGURE 4.5.23 Large vessel anatomical features. (a) Cross section of canine carotid artery. A relatively thick layer of black-staining elastin lamina (e) covers the luminal surface of the artery. The smooth muscle cells (m) form layers between the sheets of elastin in the vessel media. A few black elastin fibers surround the stained collagen fibers (c) in the outer vessel layer, the adventitia. Weigert-van Gieson elastin stain, original magnification 40×, bar = 62 μm. (b) Cross section of porcine femoral artery at the same original magnification with the same stain. By comparison, the femoral arterial structure has the same three layers but the muscle (m) is more abundant in the media and the dense layers of elastin obscure the scant collagen (c) of the adventitia. (Histology provided by Dr. Sharon Thomsen.)

top of the cross-section (the endothelial surface), and the bottom is the adventitia. In Figure 4.5.23b, a porcine femoral artery is shown at the same magnification, and again the internal elastic lamina is at the top (the endothelial surface). Note the different distribution of the elastin layers; this artery has significant elastin in the adventitia and very little collagen by comparison, while the media has a much higher concentration of smooth muscle. In both cases, evolved steam will be trapped in the elastin plates resulting in high internal pressures. Counteracting with high apposition pressures preserves tissue integrity and water retention during *RF* heating.

Damage Rate Process Coefficients

Thermal damage kinetic coefficients are usually determined from constant temperature exposures of relatively long duration. Threshold damage results are selected out of a set of damaged tissue samples for analysis from which estimates of A and E are obtained.

Constant Temperature Exposures

Because of the sensitivity of the damage integral to small changes in temperature, the typical approach in obtaining A and E is to expose the tissue to a constant temperature, identify experiments in which the damage is threshold—i.e., for which $\Omega = 1$, or $C(t) = 36.8\%$ of $C(0)$ — and obtain A from the intercept and E from the slope of an Arrhenius plot of $\ln(t)$ vs. $1/T$ for the threshold experiments. If the temperature is held constant, the integral reduces to a simple multiplication, and $C(\tau) = 0.368\ C(0)$, then Equation 4.5.70 becomes

$$ln\{\tau\} = \frac{E}{R}\frac{1}{T} - \ln\{A\} \tag{4.5.90}$$

The usual experimental method is to expose thin slices of tissue to constant temperature in a water bath, by surface application of heated water,[100–103] or on a heated metallic plate for desired time intervals. Table 4.5.10 lists rate coefficients obtained in various experiments and the relative damage rates ($d\Omega/dt$) are compared for several of them in Figure 4.5.24.

Other experiments in the ensemble for which Ω is not 1 can be included in the estimation of the coefficients by noting that Ω can be included in an equivalent exposure time, τ_{eq}:

$$ln\{\tau\} = \frac{E}{R}\frac{1}{T} - \ln\{A\} + \ln\{\Omega\} \tag{4.5.91a}$$

$$\ln\{\tau_{eq}\} = \ln\{\tau\} - \ln\{\Omega\} = \ln\left\{\frac{\tau}{\Omega}\right\} = \frac{E}{R}\frac{1}{T} - \ln\{A\} \tag{4.5.91b}$$

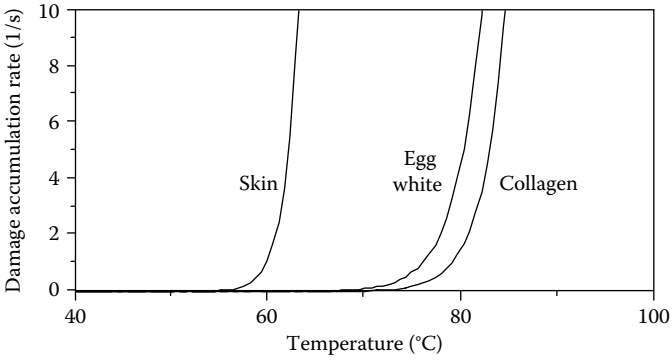


FIGURE 4.5.24 Comparison of selected damage processes from Table 4.5.10: damage rate, $d\Omega/dt$, versus temperature.

TABLE 4.5.10
Experimentally Determined Rate Coefficients

A (s^{-1}) Retina	E (J/mol)	Conditions	
Takata[188]	0	$T \leq 316$ K	
4.3×10^{64}	4.2×10^5	$316 < T \leq 323$	
9.3×10^{104}	6.7×10^5	$T > 323$	
Welch and Polhamus[187]	3.1×10^{99}	6.28×10^5	
Birngruber[189,190]	10^{44}	2.93×10^5	
Skin			
Moritz[103]	3.1×10^{98}	6.27×10^5	Not recommended
Diller and Klutke[127]	8.82×10^{94}	6.03×10^5	$T \leq 53$°C (H&M data)
	1.297×10^{31}	2.04×10^5	$T > 53$°C
Weaver and Stoll[106]	2.2×10^{124}	7.83×10^5	$317 < T < 323$ K
1.8×10^{51}	3.27×10^5	$T \geq 323$	
Egg			
Yang[191]	White 3.8×10^{57}	3.85×10^5	
Yolk 3.05×10^{56}	3.89×10^5		
PI uptake, AT-1 cells			
Bhowmick[142]	2.99×10^{37}	2.448×10^5	
Calcein fluorescence decay, AT-1 cells			
Bhowmick[142]	246.9	2.85×10^4	$300 < T < 328$ K
Time delay: t_d (s) = 1304 – 18.95 T (°C)			
2.7×10^{11}	8.78×10^4	$T > 328$ K	
Loss of clonogenicity, AT-1 cells			
Bhowmick[142]	5.95×10^{82}	5.21×10^5	$300 < T < 328$ K
Time delay: t_d (s) = 2436 – 46.1 T (°C)			
6.11×10^{12}	9.22×10^4	$T > 328$ K	

When using the kinetic models and measured coefficients, it is imperative that an adequate description of the particular end point be given. This is because different end points in the same tissue will have widely varying critical temperatures and damage accumulation rates. For example, in Table 4.5.10 in addition to the variation in tissue, the end points in the various studies differ markedly.

Retina—The measurements of Welch and Polhamus[187] used the diameter of the edge of the visible lesion formed in bovine retina under argon laser irradiation as the end point. The temperatures were not measured directly, but were determined in separate experiments on retinas in which a correlation between temperature and radius was established using microthermocouples (about 5 μm in diameter) advanced from the posterior surface of the eye to a point just below the retina. The correlates were used to estimate the retinal temperature given laser beam power and duration. The critical temperature for these coefficients is 56.0°C. Takata et al.[188] used a similar decision criterion for retinal damage for shorter exposure times. They fit the data with three sets of coefficients because a single first-order model was not sufficient. It would appear that several parallel damage processes were at work in their study, thus a set of coefficients was required. The critical temperature for the high-temperature set of coefficients is 59.9°C. Birngruber et al.[189,190] estimated a frequency factor and activation energy from consideration of the thermodynamics of protein and enzyme denaturization processes. Their estimates have a critical temperature of 74.5°C.

Skin—The end point for $\Omega = 1$ in the original studies of Henriques and Moritz[100–103] corresponded to a continuum of cascaded effects. In their study the skin of pigs was exposed, *in vivo*, to flowing water at a controlled temperature for exposure times varying over several orders of magnitude. They calibrated their coefficients so that $\Omega = 0.53$ corresponded to the onset of erythema (characterized as "first degree" in their paper). Then $\Omega = 1.0$ corresponded to a "second degree", or partial thickness burn, and $\Omega = 104$ to a full thickness, or "third degree" burn. Their coefficients, $A = 3.1 \times 10^{98}$ (s^{-1}) and $E = 6.27 \times 10^5$ (J/mol), have a critical temperature of 59.7°C; however, these coefficients do not fit their data very well and are not recommended despite their having been used by multiple investigators in model studies. In a follow-up study[127] new coefficients were obtained from the original data: $A = 8.82 \times 10^{94}$ (s^{-1}) and $E = 6.03 \times 10^5$ (J/mol) for temperatures below 53°C. These are recommended in place of the original coefficients. Weaver and Stoll[106] used similar criteria to those of Henriques and Moritz and applied two sets of coefficients (the upper values applicable above 50°C, as in the Takata et al. study) to match the experimental data. The critical temperature for their highest temperature coefficient set is 59.4°C.

Egg white and egg yolk—Egg white and yolk were exposed to constant temperature (±0.2°C) in a water bath for varying exposure times.[191] Coagulation was defined as the onset of whiteness (coagulum formation due, apparently, to an increase in scattering in the clear liquid egg white) observed by the naked eye. Water bath temperatures ranged from 60°C to 90°C in 4°C increments. At each temperature, the time to threshold was measured and plotted in accordance with Equation 4.5.22 and rate coefficients determined from linear regression, as described. Approximately 3–5 s were required to obtain whitening in the egg white at 70°C; while in the egg yolk 82°C was required for the same exposure time. The table values have a critical temperature of 76.0°C for egg white and 86.6°C for egg yolk.

The kinetic nature of thermal damage means that the time of exposure is of critical importance when discussing threshold temperatures. For example, it is commonly stated that thermal damage occurs above 45°C—and, in fact, 45°C was identified by Moritz and Henriques as the long-term exposure asymptote for threshold damage. However, at that temperature it would require 9.7 h of exposure to get a burn ($\Omega = 1$). To get a skin burn in 1 s, a temperature of 59°C would be required, and over the 0.1 s typical of the laser activations, 63°C is required. Similarly, using the coefficients for collagen, at 45°C it would require about 73 days of exposure to get $\Omega = 1$ (and that is a huge overprediction since collagen doesn't denature at 45°C); for damage in 1 s, 79°C, and in 0.1 s a constant temperature of 84.5°C is required. So, when comparing the higher temperatures required to obtain collagen denaturization in laser experiments (above about 80°C) to, for example, published values of collagen temperature thresholds (which range from 50°C to 60°C), it is important to note that (1) the times of exposure used in published collagen damage reports are usually in terms of hours and (2) it is entirely likely that collagen in dilute solution has a different set of kinetic coefficients than collagen *in situ* owing to the complex structural interrelationships inherent in the tissue.

Calcein leakage—Bhowmick and Bischof[114] measured calcein leakage rates from type AT-1 prostate tumor cells exposed to elevated temperatures. Calcein is a fluorescent dye that diffuses into cell membranes at varying rates depending on membrane integrity. The coefficients in Table 4.5.10 correspond to a 50% increase in intracellular calcein fluorescence intensity. They also reported cell survival rates (again, for 50% survival). These data are reasonably close to $\Omega = 1$ points (36.8% survival or remaining fluorescence intensity). Many other cell survival studies have been reported in the literature (see, e.g., the works of Mixter et al.[192] and Padanilam[193]).

Transient Thermal History Experiments

Problems with the standard constant temperature approach to determining A and E_a are that

1. With only a very few exceptions, the tissue must be excised to perform the experiments disrupting its blood perfusion and activating autolytic (postmortem degeneration and necrosis) processes.

2. No exposure is truly constant temperature and rise time segments must necessarily be short compared to total exposure time in order to be negligible (restricting one to the study of very slow processes).
3. It is difficult to obtain trustworthy estimates of A since the temperature axis is hyperbolic ($1/T$) and, even if the exposure times vary over several orders of magnitude, small uncertainties in the slope (E_a/R) create very large uncertainties in A.
4. The methods are difficult to use in transient MW, RF or laser damage studies since constant temperature exposures are nearly impossible to obtain.

Working with transient thermal data adds a high level of uncertainty to the damage coefficients since the time of exposure is problematic. Nevertheless, estimates of rate coefficients can be made upon which treatment protocols can be evaluated.

As can be seen from Table 4.5.10, most of the damage processes have activation enthalpies on the order of 10^5 (J/mol). For the hypothetical example damage process, with $E_a = 5.0 \times 10^5$ (J/mol), at 60°C the rate of damage formation would double if the temperature increased to 61.3°C (about 0.4% temperature increase on the absolute scale). If the transient temperature and damage process can be observed as the damage develops, as with fluorescence decay or dye penetration, the reaction velocity is easily determined as a function of temperature and determining the rate coefficients reduces to a curve fitting exercise. If, however, there is only one final observation of the accumulated damage, as in histologic section or a final immunohistochemistry assay, then the most effective approach is to integrate the ensemble of transient temperature records using assumed values for E_a and A from the He-Bischof line evaluated against some sort of "cost" function.

We may rearrange Equation 4.5.71 in which the unknowns are E_a and $\ln\{A\}$:

$$\ln\{A\} = \ln\{\tau\} - \left(\frac{1}{RT}\right)E_a \tag{4.5.92}$$

The value of Ω is calculated for each transient experiment using coefficients from the He–Bischof line and the cost might be some goodness of fit measure such as

$$\text{Cost} = \left| \ln\left\{ \frac{\int_0^{\tau} A\, e^{\left[\frac{-E_a}{RT(t)}\right]} dt}{\Omega(\tau)} \right\} \right| \tag{4.5.93}$$

The He–Bischof line is searched for values that yield minimums (ideally zeroes) in the cost. If the entire E_a and $\ln\{A\}$ plane is searched, the locus of minima is a straight line (a regression line). The solution is where that regression line crosses the He–Bischof line. There should be only one cluster of solution points that fall on the He–Bischof line, so some cluster analysis method may be required.

Summary

First-order kinetic models for tissue damage are useful for predicting trends in damage experiments. Unfortunately, there are only a few damage processes for which the frequency factor and energy have been determined *in situ*—critical temperatures for known processes range from about 50°C to 90°C. Nevertheless, these models can be used to make direct comparisons between numerical predictions of damage and histologic results; something which cannot, as yet, be achieved any other way. A very careful definition of the particular histologic end point is necessary. To avoid the

arbitrariness which characterizes much of the early damage studies, the end point ought to (1) be clearly recognizable and defined, (2) be reproducible under varying experimental conditions, (3) be easily measured, and (4) involve relatively homogeneous and readily available test tissues. For example, thermally induced birefringence image intensity decrease in skeletal and cardiac muscle is easily seen and measured using polarizing microscopy and optical detectors. Beef and pork loin cuts, chicken breasts, and canine or beef hearts can provide the numerous tissue slices required to accumulate adequate data.

In part because of the effects of biologic inhomogeneities and random fluctuations in tissue characteristics, and in part due to the difficulty of resolving small temperature differences, thermal damage data are inherently noisy. Consequently, the prediction of observed damage boundaries in very long-exposure experiments (on the order of minutes to hours) in the presence of substantial spatial thermal gradients may be frustrating. However, in small spot-size exposures of relatively short duration (up to fractions of seconds) the predicted boundaries compare favorably to those observed. This is probably due to the very steep thermal gradients typical of small spot experiments—that is, a rather large error in actual critical temperature may be swamped out by the very steep thermal gradient so that the location of the predicted damage contour may, in fact, agree fairly well with experimental histologic observation. Certainly, even though the results of a particular experiment may eventually prove impossible to duplicate in numerical models, a rather careful analysis of the trends which one would obtain from changing power, contact temperature, spot size, and duration can be studied in detail in the numerical model and on a spatial scale similar to microscopic observation. Also, the numerical model allows dissection of the transient development of thermal damage—something which cannot be achieved in any other way. So, while there are many uncertainties associated with kinetic models of thermal damage, they can be extremely illuminating and helpful in dosimetry planning.

Tissue Effects: Subzero Temperatures

An important area of bioheat transfer application is in the domain of subzero temperatures. As was noted previously, biochemical rate processes are governed by the local system temperature. Lowering the temperature has the effect of reducing reaction rates, and at sufficiently low temperatures a state of suspended animation can be achieved. Owing to the major aqueous component of physiological fluids, temperatures low enough to effect suspended animation normally result in freezing. The freezing of native biomaterials is nearly always lethal to the affected tissue upon thawing. The affected injury can be used beneficially in cryosurgery for the purpose of destroying a target tissue such as cancer. Alternatively, the tissue can be modified prior to freezing by the introduction of a chemical cryoprotective additive (CPA) to afford protection from freeze–thaw injury. Techniques have been developed for the successful long-term cryopreservation of many biomaterials for subsequent applications, including human transplantation. The response of living biomaterials to freezing and thawing is intimately tied to the thermal history during processing, especially at subzero temperatures. Thus, bioheat transfer analysis has played a key role in the design and development of low-temperature bioprocessing protocols.

Cryopreservation

Living tissues may be frozen to deep subzero temperatures to create a state of suspended animation for indefinite periods and recovered with very minimal loss of viability and function. It is necessary to store tissues at below approximately −120°C so the kinetics of chemical reactions and ice nucleation become infinitesimally small. Successful cryopreservation protocols require that subject tissue be modified prior to cooling to subzero (°C) temperatures by addition of a CPA, either to protect against the injurious effects of ice formation or to block the formation of ice so that a glassy state results (vitrification).

Widespread interest has developed in exploiting cryopreservation as a means for reversibly banking a broad spectrum of tissues for transplantation. The seminal paper which first reported this work described the use of glycerol to freeze fowl sperm 67 years ago.[194] Successes were reported in succession for other types of tissues having rather simple cell structures, such as erythrocytes, gametes, and various cells obtained from primary cultures.[195–197] Most of these cryopreservation techniques were derived via largely empirical methods, and starting in the 1970s, it came to be realized that the cryopreservation of more complex systems such as multicellular tissues and whole organs require a more rigorous scientific understanding of the mechanisms of the governing biophysical processes and cellular response to freezing and thawing. Since that time engineers have made significant contributions to the developing science of cryobiology, not the least of which has been to identify some of the key biophysical problems to be solved.[198]

Biophysical Basis of Freezing Processes in Cells

Analysis of cryopreservation by freeze/thaw processes is based on addressing phenomena associated with the solidification of aqueous solutions and their resulting effects on embedded living cells surrounded by semipermeable membranes. When an aqueous solution freezes, water is sequestered into the solid ice phase, resulting in the concentration of solutes in the residual liquid phase solution. If equilibrium is maintained between the liquid and solid phases, the coupling between temperature and solute concentration is described by the phase diagram for the solution. As shown in Figure 4.5.25, as the temperature is depressed below 0°C, the amount of solute in the liquid increases according to a function defined by the liquidus curve until the eutectic state is reached. Thus, cooling a biological tissue until ice is nucleated imposes an osmotic as well as thermal stress.

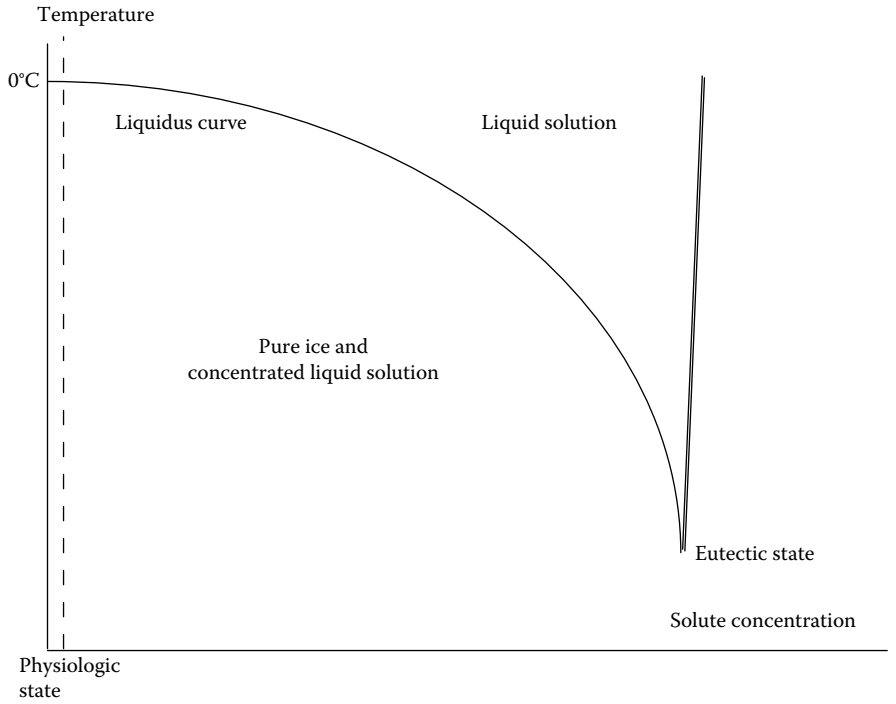


FIGURE 4.5.25 Simplified representation of an aqueous equilibrium phase diagram. Physiological solutions have solute concentrations that are quite small (about 300 mOsm) in comparison with the eutectic state. As freezing progresses, the liquid phase concentration may increase by more than 20-fold, subjecting cells to intense osmotic stress.

A consequence of the osmotic stress is that the individual cells in a tissue will lose intracellular water to the solute-enriched environment. (It is assumed that at subzero temperatures the time scales for transport of other molecular species will be negligible in comparison with that for water.) The rate at which this mass transport occurs across the cell membrane is governed by the membrane permeability to water. This transport process has been modeled as a simple diffusion phenomenon.[199]

$$\frac{\partial \Psi}{\partial T} = -\frac{L_p SBRT}{v} \ln\left(\frac{p_{in}}{p_{ex}}\right) \tag{4.5.94}$$

The Mazur equation predicts the change in the volume, Ψ, of a cell with temperature during freezing at a cooling rate of B in conjunction with the transport of water from within the cell in response to the osmotic differential developed across the membrane as solutes are concentrated during the solidification process, and R is the gas constant. The rate of transport is dependent on the surface area, S, of the cell, the membrane permeability to water, L_p, the molar volume of water, v, and the ratio of intracellular to extracellular vapor pressures of water. Although the latter term is not a rigorous description of the driving potential for the transport process, it is easily replaced by the transmembrane differential in chemical potential or solution concentration.

Dehydration is not the only process by which equilibrium can be achieved between intracellular and extracellular water during freezing. Ice crystals may nucleate within the cell, forming intracellular ice. The occurrence of intracellular ice formation (IIF) is nearly always lethal to cells and is therefore to be avoided during cryopreservation. The balance between equilibration of water across the cell membrane by osmotic dehydration and by IIF is governed primarily by the magnitude of the cooling rate. As depicted in Figure 4.5.26, at rapid cooling rates there is little opportunity for water to escape from the interior of the cell as the extracellular solute concentration increases with progressive freezing. Eventually, the conditions are satisfied for intracellular nucleation of ice

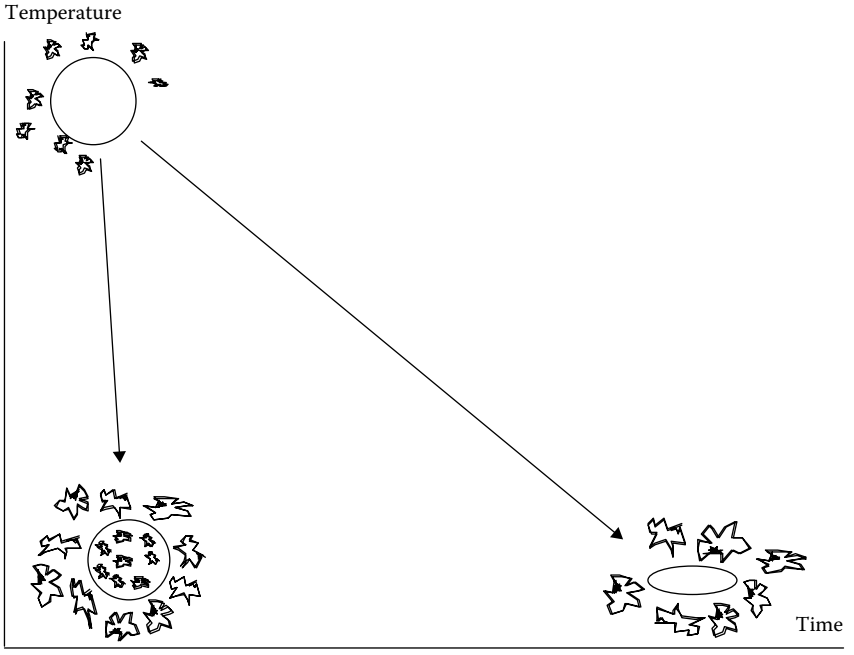


FIGURE 4.5.26 Schematic of the reaction of individual cells to extreme variations in the cooling rate during freezing. At rapid cooling rates, dehydration of the cell is minimal and intracellular ice forms. At slow cooling rates, dehydration is extensive, resulting in loss of intracellular water and morphological distortion.

crystals. At slow cooling rates, the cells are able to dehydrate as the extracellular solute concentration increases, and the intracellular water becomes frozen in the extracellular space. The resulting shrinkage of the cell can give rise to injurious chemical and mechanical stresses. In general, optimal survival from cryopreservation is achieved at intermediate cooling rates, the magnitude of which are dictated by the membrane permeability at subzero temperatures.[200]

Cryoprotective Additives

The sensitivity of cells and tissues to both IIF and osmotic stress and dehydration can be modified substantially by the prefreezing additions of a CPA. As a consequence, virtually all practical cryopreservation procedures are based on the use of a CPA to which the cell membrane is permeable. During both the CPA addition and removal procedures and the freezing and thawing processes, the applied osmotic stress results in coupled transport between the CPA and water. This coupled process is typically described in terms of irreversible thermodynamics using a format defined by Kedem and Katchalsky in which the phenomenological coefficients are related directly to physically measurable properties of a cell membrane.[201,202] These equations are written as

$$J_V = L_p \Delta p - \sum_{i=1}^{n} \sigma_i \Delta \pi_i \tag{4.5.95}$$

$$J_S = C_S (1 - \sigma) J_V + \omega_S \Delta \pi \tag{4.5.96}$$

where L_p is the membrane permeability to water; ω, the permeability to solute (CPA for applications in cryopreservation); and σ, the coupling between the two flows (reflection coefficient). The average concentration of permeable solute between the extracellular and intracellular solutions is given by C_S, and the total volume flux and solute fluxes by J_V and J_S, respectively. Equation 4.5.95 shows a provision for volume flow occurring under the action of multiple solutes.

An alternative, more generalized formulation of this type of phenomenon was also posed by Katchalsky in terms of network thermodynamics.[203] The network thermodynamic model has been applied successfully to the analysis of the cellular response to freezing processes.[204–206] It can readily be applied for inverse solution to quantify the transport coefficients from either freezing data or CPA addition data.[205,207]

Tissues and Organs

The preceding models hold for the analysis of individual cells. However, many important biological systems of interest for cryopreservation consist of organized tissues which have a three-dimensional geometry that exerts a significant effect on the osmotic behavior. For these systems, transport occurs not only across the individual cell membranes, but the interstitial volume is also involved in both the transport and storage of water and CPA. In this case, the model must be expanded to include both parallel and serial transport and storage for all mobile chemical species. Network thermodynamics provides an effective modeling format for describing these processes.

The cryopreservation of whole organs, which is still in the research stage, will demand even more complex coupled thermal and chemical transport analyses. CPAs are added and removed via perfusion through the vascular network, and this hydrodynamic flow is coupled to the osmotic, diffusional, and viscoelastic energy domains of the organ. The network thermodynamic approach to modeling is very well suited to analysis of these coupled multidomain processes and has been applied successfully to the design of CPA perfusion protocols.[208]

Influence of Cell Membrane Permeability

It is well documented that the permeability of cell membranes to water and to CPA is a strong function of the cell species, and that for all species it is a function of temperature.[202] For example,

the water permeability among various species may vary by a factor of 103 or greater. In general, membranes are significantly more permeable to water than to common CPAs, although the opposite case may also occur under unusual conditions.[209] As the temperature is reduced, the permeability is depressed dramatically. For most practical purposes, at temperatures below the range of about –20°C to –30°C the impedance to membrane transport becomes so high that there can be very little molecular exchange between the intracellular and extracellular compartments.

One of the major challenges in applying models for cell membrane transport during freezing has been developing instrumentation to measure the permeability to water and CPAs at subzero temperatures, both in the presence and absence of ice. The approach that has proven to be most successful is to follow visually on a light microscope the transient size of individual cells or tissues when they are subjected to controlled osmotic stress created either by freezing the specimen[210–212] or by changing the chemical environment.[213,214] An effective step change in the concentration of the extracellular solution is produced when ice is nucleated in the extracellular medium and the temperature is subsequently held constant for the duration of the experiment, or when a new bathing medium is perfused rapidly through the specimen chamber on the stage while the cells or tissues are physically immobilized. By both procedures, the concentration of the environmental solution is altered from an initial concentration to a final well-defined state, and the change in cell volume with time can be measured by direct microscopic observation.

The two-dimensional cross-sectional areas of the cells in the micrographs are quantified via digital image analysis and extrapolated to three-dimensional volumes based on an assumption of a consistent geometric morphology throughout the process. These experiments are repeated serially at different temperatures to generate a series of transient volume curves as shown in Figure 4.5.27 for pancreas islets,[215] a multicellular mammalian tissue. A 2 M solution of the CPA dimethyl sulfoxide (Me2SO) was introduced stepwise into the perfusion cryostage to control the osmotic environment of the islets. An initial shrinkage is measured as water leaves the cells more rapidly than the Me2SO can enter. As time progresses the islet reaches a minimum volume and then begins to swell as Me2SO and water enter the cell. The design of protocols for addition and removal of CPAs to tissues is an important and challenging step in developing effective cryopreservation methods. A rational design process is dependent on being able to measure the constitutive transport properties of the tissue of interest and having a model which can be applied to optimize the osmotic process, minimizing stresses (which are potentially lethal) on the cells.

Thus, the utility of the transient volume data is to provide a database from which Equations 4.5.94 through 4.5.96 can be applied to determine the constitutive transport properties for a cell or tissue of interest by inverse solution techniques. Property values are identified at each temperature for which experimental data exist, and the data for each property are used to fit an Arrhenius-type

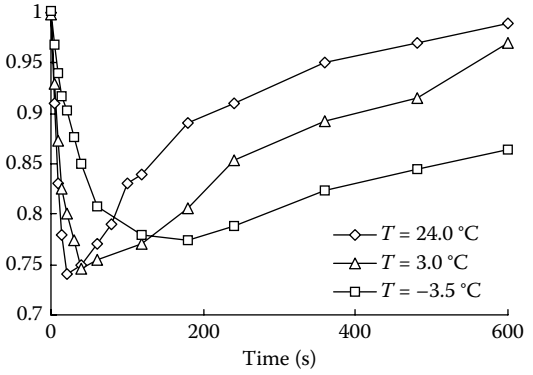


FIGURE 4.5.27 Osmotic response of human pancreas islets to exposure to a 2M solution of dimethyl sulfoxide at the indicated temperatures. (From deFreitas, R.C. et al., *Cryobiology*, 35, 230, 1997. With permission.)

function to characterize the temperature coefficient.[202] Equation 4.5.97 shows a typical formulation used for this purpose.

$$L_p(T) = L_{p,g}(T_g)\exp\left[\frac{\Delta E}{R}\left(\frac{1}{T} - \frac{1}{T_g}\right)\right] \tag{4.5.97}$$

Experiments for many different cell types have shown that the Arrhenius expression provides a model that describes the temperature dependence of the tissue transport properties well.[202]

In order to fully interpret the transient osmotic data described above, it is necessary to determine the fraction of the cell contents which is unable to cross the cell membrane under the types of osmotic stress encountered during cryopreservation. This component of the cytoplasm is often referred to as the osmotically inactive volume. It can be measured by conducting a series of experiments in which cells are exposed to increasing concentrations of impermeable solute and measuring the equilibrium volume at each state. These volumes are then plotted as a function of reciprocal solute concentration and extrapolated to the intercept of the ordinate, which indicates the volume which the cell would assume without injury when exposed to an infinitely large solute concentration. The complementary normalized fraction is a measure of the initial amount of intracellular water which can participate in the response to extracellular osmotic stress. The data are displayed on a Boyle–van't Hoff plot.[216]

A fundamental issue in the analysis of cryopreservation processes that has only recently begun to be addressed via rigorous quantitative modeling is the osmotic behavior of frozen cells during the thawing process. It has now been demonstrated by application of a model based on Equation 4.5.94 that there are combinations of freeze/thaw protocols for which the warming rate may be the governing component of the thermal history in determining the mechanism and extent of injury.[217] The design of optimal cryoprocessing protocols, therefore, may require an analysis of the effect on a tissue of interest of the complete thermal history, including both cooling and warming.

Cryopreservation by Vitrification

Ice nucleation can be avoided by doping the biological specimen with a high concentration of chemicals to induce a vitreous state.[218] The chemicals act to raise the glass transition temperature to a high value which can be reached at cooling rates consistent with standard refrigeration systems. Two major problems involving transport processes remain to be solved before vitrification can be applied to the cryopreservation of whole organs. One is the process by which a 4–5 M solution equilibrates with the entire volume of an organ, including all the constituent cells. Since this high concentration of CPA can be toxic to cells, the addition process is preferably effected at a low temperature, such as 4°C, to reduce the rates of injurious chemical reactions. In addition, the duration of exposure to the CPA prior to further cooling should be minimized to limit the accrued reaction time. However, the trade-off to lowering the temperature is caused by the thermal coefficient of viscosity, which makes it much more difficult to effect the perfusion of CPA solution through the vascular systems of an organ, and the diffusion from the vascular lumen through the interstitial tissue and into the individual cells. Optimal design is dependent upon balancing the coupled hydrodynamic, osmotic, and viscoelastic reaction of the organ to the perfusion process with the biochemical toxic reactions.[218]

Optimization of Heat Transfer during Cryopreservation

The foregoing discussion indicates the importance of the combined osmotic and thermal history at the local cellular level in determining the response to cryopreservation, including the type and extent of injury which may occur. For a given combination of cell type and CPA composition and concentration, the post-thaw viability achieved is governed by the thermal history of the process. The cooling rate determines the frozen state for storage, and the warming rate determines the manifestation of the frozen state on the final level of viability. In general, there are two extremes of

behavior during freezing. For relatively slow cooling rates, the rate at which the extracellular solute concentration increases will have a time scale comparable with that for the osmotic dehydration of the cell at high subzero temperatures. Therefore, the cell will lose a considerable fraction of its water and experience extensive dehydration as indicated in Figure 4.5.25. This large loss of volume is known to damage cells by one or more molecular-level mechanisms which have not been fully identified.[200] Increasing the concentration of CPA will provide an osmotic buffer against this type of dehydration-induced injury, as well as possibly providing other means of protection. Alternatively, when the cooling rate is rapid, the temperature is reduced much more rapidly than the cell can respond osmotically (because the membrane permeability becomes greatly reduced at low temperatures, as will be explained subsequently). Therefore, there is no opportunity for intracellular water to osmotically equilibrate with the extracellular water which is being progressively diminished in concentration. When the transmembrane water disequilibrium (as manifested in terms of liquid phase supercooling) reaches a critical value, the conditions will be satisfied for nucleation of ice in the intracellular volume. It is important to note that the cell membrane acts as a barrier to the growth of a continuous ice phase between the intracellular and extracellular compartments. The crystalline structure of intracellular ice will be considerably smaller than that of extracellular ice owing to the greater extent of supercooling prior to nucleation. Consequently, the propensity for recrystallization will be relatively large in the intracellular volume, and this process has been demonstrated to be nearly universally lethal to cells.

The combination of different injury mechanisms associated with rapid and slow cooling rates during freezing gives rise to a two-factor theory of cryoinjury.[200,219] Between the extremes of rapid and slow cooling there will be a range of intermediate rates at which survival will be maximized. The magnitude and breadth of this range is defined uniquely for each cell type as a function of its membrane permeability and sensitivity to injury. Cryopreservation protocols are targeted to produce a cooling process which lies within the optimal range of this survival signature that will produce the highest possible biological function following thawing. The design of a cryopreservation protocol for a given cell type must take into account the combined effects of the CPA composition and concentration plus the cooling and warming rates.

The above analysis of the freezing process is presented from the perspective of the local thermal history for an individual cell. However, practical cryopreservation procedures are conducted either on tissues and organs of finite macroscopic dimensions or on containers of solutions of cells or microscopic tissue suspensions. The consequence is that although the governing mass transport processes occur across microscopic scale dimensions (a single membrane and the associated boundary layers may be less than 0.1 μm), the governing heat transport occurs on a scale measured in millimeters or centimeters. It is this macroscopic heat transfer that drives all components of the cryopreservation process. Because of the macroscopic dimensions and the restriction that the refrigeration source be applied only at the external boundary of the systems during cooling, there can be significant spatial gradients in cooling rate manifested within the system. Also, in general the cooling rate will not be constant over the entire cooling process. The cooling rate should be defined over the range of temperatures most important to determining the frozen state of a cell. The critical temperature range is defined as that for which transport occurs across the cell membrane and for which IIF is most probable. Analysis of this process involves solution of the classic Stefan moving boundary problem.

For the freezing of biological systems, it is most important to accurately determine the cooling at high subzero temperatures within the mushy zone[220] during which the primary membrane transport processes occur. It is precisely this range of temperatures for which simplifying assumptions concerning the pattern of latent heat release frequently are made in order to render the mathematical solution more tractable. Therefore, it is important to match as well as possible the coupling between temperature and changing solute concentration as defined by the phase diagram. Only when the actual process physics are matched can the cooling rate for the critical temperature range be described accurately. The combination of a nonlinear pattern of latent heat release with temperature

and the often complex system geometry dictates that the governing heat transport equations be solved via numerical methods. The thermal histories calculated may then be correlated with the cooling rates on the survival signature and with the membrane mass transport equations, such as Equation 4.5.94, to predict the biological response of a given tissue to a defined boundary cooling protocol.[221]

Thermal Stress Effects

Another analysis problem to be addressed in cryopreservation is that of thermal stresses in an organ that has been vitrified. Although the constitutive properties are yet to be measured, it is apparent that a biological tissue with dimensions on the order of several centimeters will be subject to fracture at even moderate cooling rates below −120°C after a glass phase has formed.[222] A mechanical fracture of a whole organ will render it useless for subsequent transplantation. It is important to measure the viscoelastic properties of biological materials loaded with high concentrations of CPA over the temperature range between the glass transition and storage temperatures (usually at or near liquid nitrogen temperature at −196°C). Complementary to the property measurements is a need for a thermal stress analysis that can be applied with a versatility of geometric details.

Rubinsky et al.[223] have developed an initial model for how mechanical stresses may develop in the interior of an organ during solidification and the subsequent cooling process. Different organs have widely divergent shapes and sizes, and constitutive property data for frozen and vitrified tissues are practically nonexistent. In addition, an organ is likely to be preserved in a container of fluid with which it will be coupled thermally and mechanically. Factors that can be considered in a process design are manipulation of the boundary thermal protocol, shape of the overall container into which an organ is placed, and coordination of the pattern of distributed internal deposition with cooling at the boundary.

Recent Developments and Innovations

Recent innovative experiments by Fowler and Toner[224] have demonstrated that it is possible at the micro (cellular) scale to melt and recool cells containing intracellular ice at a rate that can produce a vitrified state without the introduction of the CPA to block the initial ice nucleation event. The frozen specimen is irradiated with a laser wavelength that is absorbed preferentially by the cell cytoplasm over the extracellular solution. For a solution of individual cells having characteristic diameters of about 6 μm, warming rates on the order of 1011°C/s can be achieved within the cells, while the extracellular matrix remains frozen. When the heating process is terminated in an approximate step-wise manner after 7 ns, heat transfer from the extracellular matrix to the cells produces an intracellular cooling rate of about 106°C/s, which is adequate to produce a glass state in an aqueous solution. The vitrified cells may then be warmed rapidly to suprafreezing temperatures (>0°C) without nucleation and with no measured injury. If this process is to realize broad clinical or commercial application, there is considerable engineering work to be accomplished in scale-up from these microscopic level processes to applications in larger tissue and organ systems.

Since the initial achievements in the cryopreservation of cells, many applications have been developed in both the biological and medical fields. There are now numerous for-profit concerns that have developed successful techniques for preserving human tissues for subsequent transplantation. Many of these corporations use proprietary processing protocols to preserve specific tissues, and the number of commercial participants in this field has been growing steadily over the past 10 years. In like manner, there are many for-profit and not-for-profit groups which are marketing the cryopreservation of a very broad spectrum of nonhuman tissues. Some of the more prominent applications involve mammalian and amphibian gametes, tissue culture collections, and plant germplasm. A potentially important new area of application will be the cryopreservation of living materials manufactured by the tissue engineering techniques which are being developed.[225] Bioartificial systems such as skin and organs that contain living cells will require a technology for maintaining them in a viable but latent state between the time of manufacture and implementation in a clinical

setting. Cryopreservation provides a tool of suitable potential for meeting this unique need, but the processing protocols will have to be designed to ensure successful storage and recovery for each type of manufactured tissue.

Cryosurgery

All of the foregoing discussion has been devoted to cryopreservation, for which the objective is to maximize the survival of cells from the frozen state to ensure their living function after thawing. The antithesis of this work is cryosurgery, for which the objective is to maximize destruction of a target tissue *in situ*.[226] Cryosurgery has long been practiced as an effective means for killing surface lesions, and recently it has been adopted more widely as a tool for treating internal tumors that are difficult to resect mechanically.[227]

Many of the recent gains in the successful application of cryosurgery are due to the development of new imaging methods that enable the surgeon to follow the growth of the solid-liquid interface, to which the zone of cellular destruction is coupled, in real time during the freezing process. The initial imaging methodology adapted for this purpose was ultrasonography.[228] Subsequently, other imaging modalities that can be coupled to the transient temperature field are being developed, including NMR[229] and spectroscopy.[230] Advances in cryosurgical probes that provide more effective removal of heat from tissues via enhanced thermal transport mechanisms have also contributed to the recent growth in cryosurgery.[231] Finally, over the past decade there has been a steady advance in understanding the mechanisms by which freezing and thawing act to cause tissue death as a function of the thermal parameters of the solidification process and tissue properties. Many of these fundamental advances have resulted from the application of bioengineering analysis of the tissue freezing process,[232–234] which portends the potential for engineering to further contribute to the field of cryosurgery.

A new approach to improving the efficacy of cryosurgery is derived from techniques long applied to enhance cryopreservation processes. Namely, the tissue is modified by addition of chemical agent prior to the initiation of freezing. However, for applications in cryosurgery the desired result is an increased level of cell killing. Antifreeze proteins (AFPs) are proving to be effective for this purpose.[235,236] AFPs are chemical compounds synthesized by many differing types of plants and animals to provide protection against freezing injury at high subzero temperatures.[237,238] It has also been demonstrated that AFPs modify ice crystals to needlelike shapes that can destroy cells during freezing to deeper subzero temperatures.[239]

In summary, clearly there continue to be many opportunities in low-temperature biological heat transfer for contributions with the potential for substantial positive impact on health-care delivery.

ACKNOWLEDGMENTS

This section was prepared with the support of NSF grant CBET 1250659 and NIH grant R01 EB015522 to K.R. Diller, and NIH grant #1 R01 HL56143-01 to JWV, and to the Temple Foundation Professorship in Engineering at the University of Texas at Austin (JP), and the Robert and Prudie Leibrock Professorship in Engineering (KRD). The perfusion measurements were supported in part by a grant to J.W. Valvano from Urologix Inc. The subsections on convective heat transfer and thermal properties of swine myocardium were derived in part from Naresh C. Bhavaraju's PhD. dissertation, and the subsections on thermal models and microsphere perfusion measurements in the canine prostate from David Yuan's PhD dissertation, both at the University of Texas at Austin. The section on human thermoregulation was derived in part from Karen Nyberg's PhD dissertation at the University of Texas at Austin. Dr. Pearce acknowledges the substantial contributions to the thermal damage discussion by the T.L.L. Temple and O'Donnell Foundations, by the American Heart Association, the Texas Advanced Research Program, and Valley Lab Inc. (now Covidien). Collaborations with colleagues in writing prior reviews in bioheat transfer are acknowledged.

In particular, we note the contribution of Dr. Sharon Thomsen to the subsection on histological markers and Dr. Thomas Ryan for a broad review.[240]

The physiological properties in Appendices A and B were compiled by Professor Kenneth R. Holmes (krholmes@uxl.cso.uiuc.edu) and were published in part previously.[241] The tabulation includes values for both the native thermal conductivity of biomaterials (Appendix A) and the blood perfusion rates for specific tissues and organs (Appendix B). Original sources are documented in the dedicated list of references at the end of each Appendix Knowledge of the perfusion behavior of tissues is important in that the flow of blood can have a direct quantitative effect on the temperature distribution within living tissue.

APPENDIX A

Thermal Conductivity Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Tissue	k (W/m K)	$\%H_2O$	Appendix Reference
Kidney			
Whole (rabbit)	0.502	81	6
Whole (rabbit)	0.495		16
Whole (human)	0.543	84	4
Cortex (rabbit)	0.465–0.490 ($n=7$)	76.6–79.8	1
Cortex (dog)	0.491		16
Cortex (human)	0.499		16
Medulla (rabbit)	0.502–0.544 ($n=7$)	82.0–86.0	1
Medulla (dog)	0.507		16
Medulla (human)	0.499		16
Aorta			
Human	0.476 ± 0.041 (SD) ($n=12$)		2
Arterial plaque			
Fatty	0.484 ± 0.044 (SD) ($n=13$)		2
Fibrous	0.485 ± 0.022 (SD) ($n=12$)		2
Calcified	0.502 ± 0.059 (SD) ($n=17$)		2
Artery			
Femoral (dog)	0.391		2
Carotid (dog)	0.448		2
Blood			
Whole	0.505		10
Whole (human)	0.492 ± 0.009 (SD) ($n=10$)		5
Plasma	0.582		10
Plasma (human)	0.570 ± 0.010 (SD) ($n=10$)		5
Spleen			
Human	0.543	80	4
Liver			
Rabbit	0.493	72	6

(*Continued*)

APPENDIX A (*Continued*)

Thermal Conductivity Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Tissue	k (W/m K)	%H_2O	Appendix Reference
Liver			
Rat	0.498–0.528 (n=4)		3
Sheep	0.495±0.017 (SD) (n=9)		5
Dog	0.550±0.010 (SD) (n=7)		5
Dog	0.574		11
Pig	0.498		16
Human	0.469		16
Human	0.564	77	4
Heart			
Rat	0.521–0.531 (n=2)		3
Dog	0.527		11
Dog	0.487		16
Pig	0.484		16
Human	0.493		16
Human	0.585	81	4
Muscle			
Rat	0.505–0.530 (n=7)		3
Cow	0.410–0.426		7
Pig	0.518–0.618		7
Pig	0.460		17
Sheep	0.478±0.019 (SD) (n=10)		5
Fat			
Cow	0.185–0.233		7
Cow	0.222		17
Pig	0.317–0.402		7
Pig	0.159		17
Blubber (Minke whale)	0.200–0.280	3–30	8
Blubber (Harp seal)	0.190±0.010		9
Blubber (Human)	0.201–0.217		17
Brain			
Whole	0.527	78	4
Cerebral cortex	0.564	83	4
Cerebral cortex	0.504		16
White matter	0.502	71	4
Tumor			
Periphery	0.511		11
Core	0.561		11
Colon cancer (human)	0.545		16
Bone	0.410–0.630		15
Skin	0.210–0.410		15

(*Continued*)

APPENDIX A (*Continued*)

Thermal Conductivity Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Tissue	k (W/m K)	%H_2O	Appendix Reference
Crocodile, middle of back	0.432		12
Crocodile, tail, ventral	0.334		12
Epidermis			
Foot pad (cat)	0.116–0.290		13
(Hydrated)	0.295–0.580		13
Elephant (freeze/thaw)	0.475	62.5	14
Rhinoceros (freeze/thaw)	0.452	60.1	14
Giraffe (formalin fixed)	0.442	60.0	14
Epidermis (human)	0.209		17
Epidermis (pig)	0.209		17
Dermis (human)	0.293–0.322		17
Pure water	0.627		17

Note: $c = [\%H_2O + 0.4 * (100 - \%H_2O)] * 41.9$ J/kg K (from Footnote 4).

$\rho = 1.05(103)$ kg/m^3 *(from Footnote 4).*

1. Holmes, K.R., Ryan, W., and Chen, M.M., Thermal conductivity and H_2O content in rabbit kidney cortex and medulla, *J. Therm. Biol.*, 8, 311–313, 1983.
2. Valvano, J.W. and Chitsabesan, B., Thermal conductivity and diffusivity of arterial wall and atherosclerotic plaque. *Lasers Life Sci.*, 1, 219–229, 1987.
3. Valvano, J.W., Allen, J.T., and Bowman, H.F., The simultaneous measurement of thermal conductivity, thermal diffusivity, and perfusion in small volumes of tissue, ASME 81-WA/HT-21, 1981.
4. Cooper, T.E. and Trezek, G.J., Correlation of thermal properties of some human tissue with water content, *Aerosp. Med.*, 42, 24–27, 1971.
5. Balasubramaniam, T.A. and Bowman, H.F., Thermal conductivity and thermal diffusivity of biomaterials. A simultaneous measurement technique, *J. Biomech. Eng.*, 99, 148–154, 1977.
6. Grayson, J., Internal calorimetry in the determination of thermal conductivity and blood flow, *J. Physiol.*, 118, 54–72, 1952.
7. Chato, J.C., A method for the measurement of the thermal properties of biological materials, in *Thermal Problems in Biotechnology*, Chato, J.C., Ed., American Society of Mechanical Engineers, New York, 1968.
8. Kvadsheim, P.H., Folkow, L.P., and Blix, A.S., Thermal conductivity of Minke whale blubber, *J. Therm. Biol.*, 21, 123–128, 1996.
9. Kvadsheim, P.H., Folkow, L.P., and Blix, A.S., A new device for measurement of the thermal conductivity of fur and blubber, *J. Therm. Biol.*, 19, 431–435, 1994.
10. Dumas, A. and Barozzi, G.S., Laminar heat transfer to blood flowing in a circular duct, *Int. J Heat Mass Trans.*, 27, 391–398, 1984.
11. Holmes, K.R. and Chen, M.M., Local thermal conductivity of Para-7 fibrosarcoma in hamster. *1979 Advances in Bioengineering*, American Society of Mechanical Engineers, New York, 1979, 147–149.
12. Drane, C.R., The thermal conductivity of the skin of crocodilians, *Comp. Biochem. Physiol.*, 68A, 107–110, 1981.
13. Holmes, K.R. and Adams, T., Epidermal thermal conductivity and stratum corneum hydration in cat footpad, *Am. J. Physiol.*, 228, 1903–1908, 1975.
14. *Holmes, K.R., Unpublished data.*
15. Chato, J.C., Fundamentals of bioheat transfer, in *Thermal Dosimetry and Treatment Planning*, Gautherie, M., Ed., Springer-Verlag, New York, 1990, 51.
16. Valvano, J.W., Cochran, J.R., and Diller, K.R., Thermal conductivity and diffusivity of biomaterials measured with self-heated thermistors, *Int. J. Thermophys.*, 6, 301–311, 1985.
17. Adapted from Cohen, M.L., Measurement of the thermal properties of human skin. *A review, J. Invest. Dermatol.*, 69, 333–338, 1977, and its references.

APPENDIX B

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Organ	Blood Flow (mL/min/g)	Species (*=Unanesthetized)	Appendix Reference	Technique
Brain	0.36±0.04 (SE)	Dog (n=11)	36	μs (pentobarbital)
Brain	0.78±0.06 (SE)	Rat (n=7)	30	μs
Brain	1.52±0.23 (SE)	Fetal sheep* (n=12)	32	μs
Brain	0.82±0.06 (SE)	Neonatal sheep* (n=13)	37	μs
Brain	0.64±0.03 (SE)	Sheep* (n=7)	37	μs
Brain	1.14±0.14 (SD)	Rat* (12 months old) (n=5)	33	μs
Brain	0.65±0.07 (SD)	Rat (n=10)	33	μs (pentobarbital)
Brain	1.31±0.22 (SD)	Rat* (24 months old) (n=5)	33	μs
Brain	0.43±0.14 (SD)	Rat (n=7)	33	μs (decerebrate)
Mean Cerebral	0.373±0.062	Dog	2	
Mean Cerebral	0.491±0.078	Baboon	2	
Cerebrum	0.48±0.03	Cat	18	
Cerebrum	0.53±0.06	Cat	18	
Cerebrum	0.32±0.06	Monkey	18	
Cerebrum	0.39±0.03	Monkey	18	
Cerebrum	0.79±0.12	Pig*	19	
Cerebrum	1.17±0.11 (SE)	Rat* (n=8)	31	μs
Cerebrum	0.61±0.06 (SE)	Rat (n=7)	31	μs (pentobarbital)
Cerebrum	0.72±0.05 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Cerebrum	0.50±0.04 (SE)	Rat (n=7)	31	μs (decerebrate)
Cerebrum	0.60±0.06 (SE)	Cat (n=8)	34	μs (α-chloralose)
Cerebrum	1.29±0.03 (SE)	Cat (n=8) (hypercapnic)	34	μs (α-chloralose)
Cerebral cortex	0.446±0.061	Dog	2	
Cerebral cortex	0.447±0.036	Dog	3	
Cerebral cortex	0.64±0.05	Cat	18	
Cerebral cortex	0.75±0.11	Cat	18	
Cerebral cortex	0.40±0.11	Monkey	18	
Cerebral cortex	0.50±0.11	Monkey	18	
Frontal cortex	1.06±0.05	Rat	4	
Parietal cortex	1.07±0.05	Rat	4	
Occipital cortex	0.98±0.08	Rat	4	
Cerebral white	0.240±0.037	Dog	2	
Cerebral white	0.305±0.022	Dog	3	
Cerebral white	0.20±0.05	Monkey	18	
Cerebral white	0.21±0.06	Monkey	18	
Cerebral white	0.29±0.03	Cat	18	
Cerebral white	0.20±0.03 (SE)	Cat (n=8)	34	μs (α-chloralose)
Cerebral white	0.35±0.09 (SE)	Cat (n=8) (hypercapnic)	34	μs (α-chloralose)
White matter	0.164±0.022	Baboon	2	
Corpus callosum	0.249±0.035	Dog	3	
Corpus callosum	0.62±0.03	Rat	4	

(*Continued*)

APPENDIX B (*Continued*)

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Organ	Blood Flow (mL/min/g)	Species (*=Unanesthetized)	Appendix Reference	Technique
Gray matter	0.552±0.077	Baboon	2	
Cortical gray	0.87±0.11 (SE)	Cat (n=8)	34	μs (α-chloralose)
Cortical gray	1.94±0.37 (SE)	Cat (n=8) (hypercapnic)	34	μs (α-chloralose)
Caudate nucleus	0.88±0.03	Rat	4	
Caudate nucleus	0.65±0.06	Cat	18	
Caudate nucleus	0.87±0.11	Cat	18	
Caudate nucleus	0.39±0.05	Monkey	18	
Caudate nucleus	0.47±0.05	Monkey	18	
Caudate nucleus	0.85±0.10 (SE)	Cat (n=8)	34	μs (α-chloralose)
Caudate nucleus	1.31±0.19 (SE)	Cat (n=8) (hypercapnic)	34	μs (α-chloralose)
Choroid plexus	4.18±0.60 (SE)	Neonatal sheep* (n=13)	37	μs
Choroid plexus	7.53±1.11 (SE)	Sheep* (n=7)	37	μs
Hypothalamus	1.03±0.06	Rat	4	
Pituitary	1.08±0.24 (SE)	Neonatal sheep* (n=13)	37	μs
Pituitary	1.13±0.16 (SE)	Sheep* (n=7)	37	μs
Thalamus	0.92±0.07	Rat	4	
Thalamus-midbrain	0.62±0.07 (SE)	Cat (n=8)	34	μs (α-chloralose)
Thalamus-midbrain	1.64±0.32 (SE)	Cat (n=8) (hypercapnic)	34	μs (α-chloralose)
Hippocampus	0.68±0.06	Rat	4	
Cerebellum	0.63±0.04	Cat	18	
Cerebellum	0.70±0.07	Cat	18	
Cerebellum	0.60±0.08 (SE)	Cat (n=8)	34	μs (α-chloralose)
Cerebellum	1.63±0.31 (SE)	Cat (n=8) (hypercapnic)	34	μs (α-chloralose)
Cerebellum	0.40±0.08	Monkey	18	
Cerebellum	0.50±0.02	Monkey	18	
Cerebellum	0.78±0.06	Pig*	19	
Cerebellum	1.31±0.19 (SE)	Rat* (n=8)	31	μs
Cerebellum	1.02±0.12 (SE)	Rat (n=7)	31	μs (pentobarbital)
Cerebellum	1.04±0.08 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Cerebellum	0.78±0.07 (SE)	Rat (n=7)	31	μs (decerebrate)
Pons	0.50±0.03 (SE)	Cat (n=8)	34	μs (α-chloralose)
Pons	1.18±0.25 (SE)	Cat (n=8) (hypercapnic)	34	μs (α-chloralose)
Medulla	0.37±0.06 (SE)	Cat (n=8)	34	μs (α-chloralose)
Medulla	1.07±0.20 (SE)	Cat (n=8) (hypercapnic)	34	μs (α-chloralose)
Medulla	0.84±0.04	Rat	4	
Brain stem	0.52±0.03	Cat	18	
Brain stem	0.58±0.03	Cat	18	
Brain stem	0.33±0.06	Monkey	18	
Brain stem	0.36+0.03	Monkey	18	
Brain stem	0.46±0.06	Pig*	19	
Brain stem	0.55±0.05 (SE)	Cat (n=8)	34	μs (α-chloralose)
Brain stem	1.44±0.28 (SE)	Cat (n=8) (hypercapnic)	34	μs (α-chloralose)
Spinal cord	0.22±0.02	Cat	18	

(*Continued*)

APPENDIX B (*Continued*)

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Organ	Blood Flow (mL/min/g)	Species (* = Unanesthetized)	Appendix Reference	Technique
Spinal cord	0.31±0.03	Cat	18	
Spinal cord	0.20±0.03	Monkey	18	
Spinal cord	0.21±0.03	Monkey	18	
Nerve, sciatic	0.27±0.03 (SE)	Rat (n=14)	29	[^{14}C]iodoantipyrine
Kidney	3.018±0.412 (SE)	Dog (n=11)	36	μs (pentobarbital)
Kidney	3.99±0.88 (SD)	Rat* (12 months old) (n=5)	33	μs
Kidney	3.64±0.92 (SD)	Rat (n=10)	33	μs (pentobarbital)
Kidney	4.82±1.39 (SD)	Rat* (24 months old) (n=5)	33	μs
Kidney	2.31±0.82 (SD)	Rat (n=7)	33	μs (decerebrate)
Kidney	1.54±0.12 (SE)	Fetal sheep* (n=12)	32	μs
Kidney	3.00±0.14 (SE)	Neonatal sheep* (n=13)	37	μs
Kidney	4.19±0.32 (SE)	Sheep* (n=7)	37	μs
Kidney	4.32±0.30	Dog	3	
Kidney	3.2±0.3	Rabbit	17	
Kidney	2.80±0.05	Pig*	19	
Kidney	4.7±0.3	Rat	13	
Kidney	5.55 (4.87–8.76)	Rat (n=14)	26	
Kidney	6.23±0.49	Rat (n=9)	27	
Kidney	4.38±0.51 (SE)	Rat (n=7)	30	μs
Kidney	4.38±0.44 (SE)	Rat (n=7)	30	μs
Kidney	7.63±0.46 (SE)	Rat* (n=8)	31	μs
Kidney	6.16±0.39 (SE)	Rat (n=7)	31	μs (pentobarbital)
Kidney	5.96±0.42 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Kidney	5.46±0.53 (SE)	Rat (n=7)	31	μs (decerebrate)
Whole cortex	4.0–5.0	Human*	6	
Whole cortex	3.63±0.54 (range=3.0–4.4)	Pig	8	
Whole cortex	3.71±0.284	Pig (n=4)	25	
Whole cortex	4.69±0.347	Pig (n=6)	25	
Whole cortex	5.06±0.251	Pig (n=6)	25	
Whole cortex	3.25±0.20 (SE)	Sheep* (n=6)	40	μs
Outer cortex	3.81±0.23 (SE)	Sheep* (n=6)	40	μs
Outer cortex	6.32±0.33	Dog	3	
Outer cortex	7.56±0.70	Dog	11	
Outer cortex	4.62	Dog	7	
Outer cortex	8.84±0.014	Rabbit	5	
Outer cortex	8.7±0.6	Rat	13	
Middle cortex	9.53±0.013	Rabbit	5	
Inner cortex	2.57±0.22 (SE)	Sheep* (n=6)	40	μs
Inner cortex	4.38±1.19	Dog	11	

(*Continued*)

APPENDIX B (*Continued*)

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Organ	Blood Flow (mL/min/g)	Species (* = Unanesthetized)	Appendix Reference	Technique
Inner cortex	3.35±0.42	Dog	3	
Inner cortex	3.96	Dog	7	
Inner cortex	3.24±0.004	Rabbit	5	
Inner cortex	4.7±0.4	Rat	13	
Outer 1/4 cortex	4.24±0.40	Dog	12	
Next 1/4	6.24±0.61	Dog	12	
Next 1/4	3.55±0.45	Dog	12	
Inner 1/4	1.73±0.13	Dog	12	
Medulla	0.35±0.09 (SE)	Sheep* (n=6)	40	μs
Medulla	0.77±0.064	Dog	3	
Outer medulla	1.2	Human*	6	
Outer medulla	1.30	Dog	7	
Outer medulla	2.10±0.28	Dog	11	
Inner medulla	0.25	Human*	6	
Adrenal gland	7.62±2.55 (SE)	Fetal sheep* (n=12)	32	μs
Adrenal gland	1.83±0.12 (SE)	Neonatal sheep (n=13)	37	μs
Adrenal gland	1.99±0.13 (SE)	Sheep* (n=7)	37	μs
Adrenal gland	1.341±0.297 (SE)	Dog (n=11)	36	μs (pentobarbital)
Muscle				
Skeletal	0.027 (average)	Human*	16	
Skeletal	0.12±0.02 (SE)	Neonatal sheep* (n=13)	37	μs
Skeletal	0.04±0.01 (SE)	Sheep* (n=7)	37	μs
Skeletal	0.18±0.07 (SD)	Rat* (12 months old) (n=5)	33	μs
Skeletal	0.03±0.01 (SD)	Rat (n=10)	33	μs (pentobarbital)
Skeletal	0.22±0.12 (SD)	Rat* (24 months old) (n=5)	33	μs
Skeletal	0.05±0.02 (SD)	Rat (n=7)	33	μs (decerebrate)
Skeletal	0.39±0.20 (SD)	Rat (n=12)	41	l33Xe (ether)
Resting thigh				
1.5 cm depth	0.018±0.011	Human*	14	
3.0 cm depth	0.026±0.013	Human*	14	
Resting thigh	0.020–0.022	Human*	15	
Vastus medialis and triceps	0.14±0.03	Pig*	19	
Gracilis	0.049±0.064	Dog	2	
Masseter	0.09±0.01 (SE)	Rat (n=7)	30	μs
Cremaster	0.16±0.03 (SE)	Rat (n=7)	30	μs
Cremaster	0.07±0.018 (SE)	Rat* (n=8)	31	μs
Cremaster	0.06±0.008 (SE)	Rat (n=7)	31	μs (pentobarbital)
Cremaster	0.05±0.013 (SE)	Rat (n=7)	31	μs (chloralose-urethan)

(*Continued*)

APPENDIX B (*Continued*)

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Organ	Blood Flow (mL/min/g)	Species (*=Unanesthetized)	Appendix Reference	Technique
Cremaster	0.04±0.012 (SE)	Rat (n=7)	31	μs (decerebrate)
Diaphragm	0.93±0.19 (SE)	Rat* (n=8)	31	μs
Diaphragm	0.22±0.02 (SE)	Rat (n=7)	31	μs (pentobarbital)
Diaphragm	0.24±0.04 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Diaphragm	0.35±0.04 (SE)	Rat (n=7)	31	μs (decerebrate)
Diaphragm	0.144±0.017 (SE)	Dog (n=11)	36	μs (pentobarbital)
Rectus abdominus	0.18±0.029 (SE)	Rat* (n=8)	31	μs
Rectus abdominus	0.04±0.008 (SE)	Rat (n=7)	31	μs (pentobarbital)
Rectus abdominus	0.09±0.025 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Rectus abdominus	0.09±0.027 (SE)	Rat (n=7)	31	μs (decerebrate)
Gastrocnemius	0.11±0.022 (SE)	Rat* (n=8)	31	μs (decerebrate)
Gastrocnemius	0.04±0.003 (SE)	Rat (n=7)	31	μs (decerebrate)
Gastrocnemius	0.15±0.084 (SE)	Rat (n=7)	31	μs (decerebrate)
Gastrocnemius	0.06±0.019 (SE)	Rat (n=7)	31	μs (decerebrate)
Tibialis anterior	0.36±0.068 (SE)	Rat* (n=8)	31	μs (decerebrate)
Tibialis anterior	0.03±0.004 (SE)	Rat (n=7)	31	μs (decerebrate)
Tibialis anterior	0.17±0.11 (SE)	Rat (n=7)	31	μs (decerebrate)
Tibialis anterior	0.08±0.03 (SE)	Rat (n=7)	31	μs (decerebrate)
Psoas	0.26±0.034 (SE)	Rat* (n=8)	31	μs (decerebrate)
Psoas	0.04±0.004 (SE)	Rat (n=7)	31	μs (decerebrate)
Psoas	0.09±0.037 (SE)	Rat (n=7)	31	μs (decerebrate)
Psoas	0.07±0.022 (SE)	Rat (n=7)	31	μs (decerebrate)
Heart	3.72±0.45 (SE)	Fetal sheep* (n=12)	32	μs
Heart	2.16±0.23 (SE)	Neonatal sheep* (n=13)	37	μs
Heart	1.13±0.09 (SE)	Sheep* (n=7)	37	μs
Heart	5.45±0.59 (SE)	Rat (n=7)	30	μs
Heart	5.31±0.59 (SE)	Rat* (n=8)	31	μs
Heart	2.83±0.60 (SE)	Rat (n=7)	31	μs (pentobarbital)
Heart	2.23±0.37 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Heart	3.19±0.55 (SE)	Rat (n=7)	31	μs (decerebrate)
Heart	6.61±2.20 (SD)	Rat* (12 months old) (n=5)	33	μs
Heart	5.45±1.62 (SD)	Rat (n=10)	33	μs (pentobarbital)
Heart	8.66±2.74 (SD)	Rat* (24 months old) (n=5)	33	μs
Heart	3.67±1.03 (SD)	Rat (n=7)	33	μs (decerebrate)
Right ventricle	0.92±0.12	Pig*	19	
Left ventricle	1.45±0.09	Pig*	19	
Left ventricle	1.58±0.13	Dog	20	
Left ventricle	1.09±0.35	Dog	21	
Left ventricle	0.61±0.41	Dog	22	
Left ventricle	0.88±0.14	Dog	23	

(*Continued*)

APPENDIX B (*Continued*)

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Organ	Blood Flow (mL/min/g)	Species (*=Unanesthetized)	Appendix Reference	Technique
Epicardium	0.89±0.14	Dog	24	
Epicardium	0.86±0.14	Dog	24	
Endocardium	1.14±0.18	Dog	24	
Endocardium	1.08±0.18	Dog	24	
GI				
Intestine	0.47±0.05 (SE)	Fetal sheep* (n=12)	32	μs
Intestine	0.389±0.04	Dog	3	
Tract	2.05±0.25	Rat* (n=8)	31	μs
Tract	1.31±0.11	Rat (n=7)	31	μs (pentobarbital)
Tract	1.26±0.09	Rat (n=7)	31	μs (chloralose-urethan)
Tract	1.19±0.14	Rat (n=7)	31	μs (decerebrate)
Stomach	1.27±0.21 (SE)	Rat* (n=8)	31	μs
Stomach	0.46±0.04 (SE)	Rat (n=7)	31	μs (pentobarbital)
Stomach	0.53±0.05 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Stomach	0.52±0.07 (SE)	Rat (n=7)	31	μs (decerebrate)
Stomach	1.31±0.30 (SD)	Rat* (12 months old) (n=5)	33	μs
Stomach	0.55±0.21 (SD)	Rat (n=10)	33	μs (pentobarbital)
Stomach	1.50±0.45 (SD)	Rat* (24 months old) (n=5)	33	μs
Stomach	0.42±0.21 (SD)	Rat (n=7)	33	μs (decerebrate)
Stomach	0.53±0.26 (SD)	Dog (n=12)	39	Rb86(pentobarbital)
Duodenum	1.11±0.46 (SD)	Dog (n=12)	39	Rb86(pentobarbital)
Duodenum	0.901±0.132 (SE)	Dog (n=11)	36	μs (pentobarbital)
Duodenum	3.63±0.70 (SD)	Rat* (12 months old) (n=5)	33	μs
Duodenum	2.33±0.89 (SD)	Rat (n=10)	33	μs (pentobarbital)
Duodenum	3.31±0.64 (SD)	Rat* (24 months old) (n=5)	33	μs
Duodenum	2.00±0.76 (SD)	Rat (n=7)	33	μs (decerebrate)
Jejunum	0.642±0.086 (SE)	Dog (n=11)	36	μs (pentobarbital)
Jejunum	0.263±0.034 (SE)	Cat (n=12	35	μs (pentobarbital)
Ileum	0.498±0.083 (SE)	Dog (n=11)	36	μs (pentobarbital)
Small bowel	1.00±0.45 (SD)	Dog (n=12)	39	Rb86(pentobarbital)
Small bowel	2.09±0.17 (SE)	Neonatal sheep* (n=13)	37	μs
Small bowel	0.81±0.06 (SE)	Sheep* (n=7)	37	μs
Small bowel	2.81±0.46 (SE)	Rat* (n=8)	31	μs
Small bowel	1.73±0.18 (SE)	Rat (n=7)	31	μs (pentobarbital)
Small bowel	1.67±0.13 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Small bowel	1.52±0.20 (SE)	Rat (n=7)	31	μs (decerebrate)
Cecum	1.89±0.19 (SE)	Rat* (n=8)	31	μs
Cecum	1.52±0.08 (SE)	Rat (n=7)	31	μs (pentobarbital)

(*Continued*)

APPENDIX B (*Continued*)

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Organ	Blood Flow (mL/min/g)	Species (*=Unanesthetized)	Appendix Reference	Technique
Cecum	1.51±0.15 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Cecum	1.55±0.22 (SE)	Rat (n=7)	31	μs (decerebrate)
Colon	1.24±0.57 (SD)	Dog (n=12)	39	Rb86 (pentobarbital)
Colon	0.711±0.065 (SE)	Dog (n=11)	36	μs (pentobarbital)
Large bowel	0.92±0.06 (SE)	Neonatal sheep* (n=13)	37	μs
Large bowel	0.55±0.09 (SE)	Sheep* (n=7)	37	μs
Large bowel	1.33±0.29 (SE)	Rat* (n=8)	31	μs
Large bowel	0.66±0.09 (SE)	Rat (n=7)	31	μs (pentobarbital)
Large bowel	0.59±0.07 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Large bowel	0.70±0.13 (SE)	Rat (n=7)	31	μs (decerebrate)
Pancreas	0.008–0.016	Dog	1	
Pancreas	0.545±0.070 (SE)	Dog (n=11)	36	μs (pentobarbital)
Pancreas	0.652±0.26 (SD)	Dog (n=12)	39	Rb86 (pentobarbital)
Liver				
Hep art. + portal	2.65±0.530	Pig (n=4)	25	
Hep art.+portal	2.53±0.184	Pig (n=6)	25	
Hep art.+portal	3.17±0.228	Pig (n=6)	25	
Hep art.+portal	1.40±0.15 (SE)	Rat* (n=8)	31	μs
Hep art.+portal	1.19±0.07 (SE)	Rat (n=7)	31	μs (pentobarbital)
Hep art.+portal	1.22±0.09 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Hep art.+portal	1.25±0.16 (SE)	Rat (n=7)	31	μs (decerebrate)
Via portal vein	1.25	Rat	25	
Via portal vein	1.19±0.12 (SE)	Rat* (n=8)	31	μs
Via portal vein	0.98±0.05 (SE)	Rat (n=7)	31	μs (pentobarbital)
Via portal vein	0.98±0.10 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Via portal vein	0.93±0.12 (SE)	Rat (n=7)	31	μs (decerebrate)
Via hepatic art.	0.24±0.04	Pig	3	
Via hepatic art.	0.22±0.04	Pig*	19	
Via hepatic art.	0.15±0.03	Rat (n=9)	27	
Via hepatic art.	0.22±0.05 (SE)	Rat* (n=8)	31	μs
Via hepatic art.	0.21±0.05 (SE)	Rat (n=7)	31	μs (pentobarbital)
Via hepatic art.	0.23±0.02 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Via hepatic art.	0.31±0.07 (SE)	Rat (n=7)	31	μs (decerebrate)
Via hepatic art.	0.10±0.05 (SE)	Fetal sheep* (n=12)	32	μs
Via hepatic art.	0.10±0.03 (SE)	Neonatal sheep* (n=13)	37	μs
Via hepatic art.	0.14±0.05 (SE)	Sheep* (n=7)	37	μs
Via hepatic art.	0.02±0.02 (SD)	Rat* (12 months old) (n=5)	33	μs
Via hepatic art.	0.07±0.04 (SD)	Rat (n=10)	33	μs (pentobarbital)
Via hepatic art.	0.03±0.01 (SD)	Rat* (24 months old) (n=5)	33	μs
Via hepatic art.	0.11±0.07 (SD)	Rat (n=7)	33	μs (decerebrate)

(*Continued*)

APPENDIX B (*Continued*)

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Organ	Blood Flow (mL/min/g)	Species (*=Unanesthetized)	Appendix Reference	Technique
Via hepatic art.	0.178±0.050 (SE)	Dog (n=11)	36	μs (pentobarbital)
Via hepatic art.	0.31±0.30 (SD)	Rabbit (n=5)	42	μs
Spleen	1.46±0.14	Dog	3	
Spleen	1.381±0.199 (SE)	Dog (n=11)	36	μs (pentobarbital)
Spleen	3.60±0.16	Pig*	19	
Spleen	1.64±0.14 (SE)	Rat* (n=8)	31	μs
Spleen	1.52±0.22 (SE)	Rat (n=7)	31	μs (pentobarbital)
Spleen	2.35±0.49 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Spleen	2.01±0.22 (SE)	Rat (n=7)	31	μs (decerebrate)
Spleen	5.43±1.13 (SE)	Fetal sheep* (n=12)	32	μs
Spleen	4.03±0.73 (SE)	Neonatal sheep* (n=13)	37	μs
Spleen	1.88±0.27 (SE)	Sheep* (n=7)	37	μs
Spleen	1.51±0.54 (SD)	Rat* (12 months old) (n=5)	33	μs
Spleen	0.75±0.37 (SD)	Rat (n=10)	33	μs (pentobarbital)
Spleen	1.13±0.38 (SD)	Rat* (24 months old) (n=5)	33	μs
Spleen	0.40±0.15 (SD)	Rat (n=7)	33	μs (decerebrate)
Prostate	0.17 (.11–.21)	Rat (n=14)	26	
Prostate	0.18±0.02	Rat (n=9)	27	
Prostate	0.31–0.79	Dog (n=11)	38	(Pentobarb, halothane, or Nembutal™)
Bone				
Femur	0.096±0.039	Rat	9	
Tibia	0.083±0.032	Rat	9	
Femur	0.14±0.01	Pig*	19	
Skin	0.20±0.02 (SE)	Neonatal sheep* (n=13)	37	μs
Skin	0.09±0.01 (SE)	Sheep* (n=7)	37	μs
Skin	0.18±0.03 (SE)	Rat* (n=8)	31	μs
Skin	0.12±0.03 (SE)	Rat (n=7)	31	μs (pentobarbital)
Skin	0.15±0.02 (SE)	Rat (n=7)	31	μs (chloralose-urethan)
Skin	0.17±0.02 (SE)	Rat (n=7)	31	μs (decerebrate)
Skin				
In the cold	~0.02	Human forearm*	10	
Thermoneutral	0.04–0.05	Human forearm*	10	
Hyperthermic	Near or >0.20	Human forearm*	10	
Thermoneutral	0.9±0.09	Pig*	19	
Subcutaneous tissue	0.48±0.15	Rat (n=29)	41	l33Xe (ether)
Tumor	0.25±0.11	Rat (n=12)	41	l33Xe (ether)
	1.79±0.70	Rabbit (n=5)	42	μs

(*Continued*)

APPENDIX B (*Continued*)

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

Organ	Blood Flow (mL/min/g)	Species (* = Unanesthetized)	Appendix Reference	Technique
Fat	0.21 ± 0.06	Pig*	19	
Testes	0.45 ± 0.06	Rat (n = 9)	27	
Lungs	1.14 ± 0.14 (SE)	Fetal sheep* (n = 12)	32	μs

SE, standard error; SD, standard deviation.

1. Rappaport, A.M., Kawamura, T., Davidson, J.K., Lin, B.J., Ohira, S., Zeigler, M., Coddling, J.A., Henderson, M.J., and Haist, R.E., Effects of hormones and blood flow on insulin output of isolated pancreas in situ, *Am. J. Physiol.,* 221, 343, 1971.
2. Marcus, M.L., Bischof, C.J., and Heistad, D.D., Comparison of microsphere and xenon-133 clearance method in measuring skeletal muscle and cerebral blood flow, *Circ. Res*., 48, 748–761, 1981.
3. Fan, F.-C., Schuessler, G.B., Chen, R.Y.Z., and Chien, S., Determinations of blood flow and shunting of 9-and 15-μm spheres in regional beds. *Am. J. Physiol*., 237, H25–H33, 1979.
4. Mraovitch, S., Iadecola, C., and Reis, D. J., Vasoconstriction unassociated with metabolism in cerebral cortex elicited by electrical stimulation of the parabrachial nucleus in rat, *J. Cereb. Blood Flow Metab.*, 3 (Suppl. 1), S196–S197, 1983.
5. Bhattacharya, J. and Beilin, L.J., Left ventricular cannulation for microsphere estimation of rabbit renal blood flow, *Am. J. Physiol.*, 238, H736–H739, 1980.
6. Thurau, K. and Levine, D.Z., The renal circulation, in *The Kidney: Morphology, Biochemistry, Physiology*, Rouiller, C. and Muller, A.F., Eds., Academic Press, New York, 1971, 1–70.
7. Passmore, J.C., Neiberger, R.E., and Eden, S.W., Measurement of intrarenal anatomic distribution of krypton-85 in endotoxic shock in dogs, *Am. J. Physiol.*, 232, H54–H58, 1977.
8. Gyrd-Hansen, N., Renal clearances in pigs, *Acta Vet. Scand.*, 9, 183–198, 1968.
9. Schoutens, A., Bergmann, P., and Verhas, M., Bone blood flow measured by 85Sr microspheres and bone seeker clearances in the rat, *Am. J. Physiol.*, 236, H1–H6, 1979.
10. Johnson, J.M., Brengelmann, G.L., Hales, J.R.S., Vanhoutte, P.M., and Wenger, C.B., Regulation of the cutaneous circulation, *Fed. Proc.*, 45, 2841–2850, 1986.
11. Neiberger, R.E. and Passmore, J.C., Effects of dopamine on canine intrarenal blood flow distribution during hemorrhage, *Kidney Int.*, 15, 219–226, 1979.
12. Youichi, A., Okahara, T., Kishimoto, T., Yamamoto, K., and Ueda, J., Relationship between intrarenal distribution of blood flow and renin secretion, *Am. J. Physiol.*, 225, 319–323, 1973.
13. Mimran, A. and Casellas, D., Microsphere size and determination of intrarenal blood flow distribution in the rat, *Pflugers Arch.*, 382, 233–240, 1979.
14. Sekins, K.M., Dundore, D., Emery, A.F., Lehmann, J.F., McGrath, P.W., and Nelp, W.B., Muscle blood flow changes in response to 915 MHz diathermy with surface cooling as measured by Xe133 clearance, *Arch Phys. Med. Rehabil.*, 61, 105–113, 1980.
15. Lassen, N.A., Lindbjerg, J., and Munck, O., Measurement of bloodflow through skeletal muscle by intramuscular injection of xenon-133, *Lancet*, 1, 686–689, 1964.
16. Keele, C.A. and Neil, E., Eds., *Samson Wright's Applied Physiology*, 12th ed., Oxford Press, London, 1971, 62.
17. Ott, C.E. and Vari, R.C., Renal autoregulation of blood flow and filtration rate in the rabbit, *Am. J. Physiol.*, 237, F479–F482, 1979.
18. Haws, C.W. and Heistad, D.D., Effects of nimodipine on cerebral vasoconstrictor responses, *Am. J. Physiol.*, 247, H170–H176, 1984.
19. Tranquilli, W.J., Parks, C.M., Thurmon, J.C., Benson, G.J., Koritz, G.D., Manohar, M., and Theodorakis, M.C., Organ blood flow and distribution of cardiac output in nonanesthetized swine, *Am. J. Vet. Res.*, 43, 895–897, 1982.
20. Hernandez, E.J., Hoffman, J.K., Fabian, M., Siegel, J.H., and Eberhart, R.C., Thermal quantification of regional myocardial perfusion and heat generation, *Am. J. Physiol.*, 236, H345–H355, 1979.

(*Continued*)

APPENDIX B (*Continued*)

Blood Perfusion Data for Specific Tissues and Organs for Humans and Other Mammalian Species

21. Drake, A.J. and Noble, M.I.M., Myocardial blood flow measured by carbonized microspheres before and after cardiac denervation, *Bibl. Anat.*, 15, 53–56, 1976.
22. Baer, R.W., Payne, B.D., Verrier, E.D., Vlahakes, G.J., Molodowitch, D., Uhlig, P.N., and Hoffman, J.I.E., Increased number of myocardial blood flow measurements with radionuclide-labeled microspheres, *Am. J. Physiol.*, 246, H418–H434, 1984.
23. Crystal, G.J., Boatwright, R.B., Downey, H.F., and Bashour, F.A., Shunting of microspheres across the canine coronary circulation, *Am. J. Physiol.*, 236, H7–H12, 1979.
24. Capurro, N.L., Goldstein, R.E., Aamodt, R., Smith, H.J., and Epstein, S.E., Loss of microspheres from ischemic canine cardiac tissue, *Circ. Res.*, 44, 223–227, 1979.
25. Lovell, R.A., The toxicity of microcystin-LR in swine and mice, PhD thesis, University of Illinois, 1989.
26. Damber, J.-E., Bergh, A., Daehlin, L., Petrow, V., and Landstrom, M., Effects of 6-methylene progesterone on growth, morphology, and blood flow of the Dunning R3327 prostatic adenocarcinoma, *The Prostate*, 20, 187–197, 1992.
27. Jonsson, O., Widmark, A., Grankvist, K., Damber, J.-E., and Henriksson, R., Effects of clonidine-induced hypertension on blood flows in prostatic adenocarcinoma (Dunning R3327) and normal tissue, *The Prostate*, 20, 225–232, 1992.
28. Gores, G.J., Kost, L.J., and LaRusso, N.F., The isolated perfused rat liver: conceptual and practical considerations, *Hepatology*, 6, 511–517, 1986.
29. Rundquist, I., Smith, Q.R., Michel, M.E., Ask, P., Öberg, P.A., and Rapoport, S.I., Sciatic nerve blood flow measured by laser Doppler flowmetry and [^{14}C] iodoantipyrine, *Am. J. Physiol.*, 248, H311–H317, 1985.
30. Proctor, K.G. and Busija, D.W., Relationships among arteriolar, regional, and whole organ blood flow in cremaster muscle, *Am. J. Physiol.*, 249, H34–H41, 1985.
31. Seyde, W.C., McGowan, L., Lund, N., Duling, B., and Longnecker, D.E., Effects of anesthetics on regional hemodynamics in normovolemic and hemorrhaged rats, *Am. J. Physiol.*, 249, H164–H173, 1985.
32. Crandell, S.S., Fisher, D.J., and Morriss, F.H., Jr., Effects of ovine maternal hyperglycemia on fetal regional blood flows and metabolism, *Am. J. Physiol.*, 249, E454–E460, 1985.
33. Tuma, R.F., Irion, G.L., Vasthare, U.S., and Heinel, L.A., Age-related changes in regional blood flow in the rat, *Am. J. Physiol.*, 249, H485–H491, 1985.
34. Baumbach, G.I. and Heistad, D.D., Heterogeneity of brain blood flow and permeability during acute hypertension, *Am. J. Physiol.*, 249, H629–H637, 1985.
35. Kvietys, P.R., Shepherd, A.P., and Granger, D.N., Laser-Doppler, H2 clearance, and microsphere estimates of mucosal blood flow, *Am. J. Physiol.*, 249, G221–G227, 1985.
36. Kapin, M.A. and Ferguson, J.L., Hemodynamic and regional circulatory alterations in dog during anaphylactic challenge, *Am. J. Physiol.*, 249, H430–H437, 1985.
37. Koehler, R.C., Traystman, R.J., and Jones, M.D., Jr., Regional blood flow and O2 transport during hypoxic and CO hypoxia in neonatal and adult sheep, *Am. J. Physiol.*, 248, H118–H124, 1985.
38. Andersson, L., Dahn, I., Nelson, K.E., and Norgren, A., Method for measuring prostate blood flow with Xenon133 in the dog, *Invest. Urol.*, 5, 140–148, 1967.
39. Goodhead, B., Acute pancreatitis and pancreatic blood flow, *Surg. Gynecol. Obstet.*, 129, 331–340, 1969.
40. Reddy, V.K., Zamora, C.S., Frandle, K.A., and Samson, M.D., Regional renal blood flow in ewes, *Am. J. Vet. Res.*, 42, 1802–1803, 1981.
41. Mattsson, J., Alpsten, M., Appelgren, L., and Peterson, H.-I., Influence of Noradrenalin on local tumor blood flow, *Eur. J. Cancer*, 16, 99–102, 1980.
42. Burton, M.A., Kelleher, D.K., Gray, B.N., and Morgan, C.K., Effect of temperature on liver tumour blood flow, *Eur. J. Cancer*, 26, 999, 1990.

REFERENCES

1. Johnston, K. A., and Bennett, A. F., 1996, *Animals and Temperature: Phenotypic and Evolutionary Adaptation*, Cambridge University Press, Cambridge.
2. Diller, K. R., 1992, Modeling of bioheat transfer processes at high and low temperatures, *Advances in Heat Transfer*, 22, pp. 157–357.
3. Chato, J. C., 1968, A method for the measurement of thermal properties of biologic materials, *Symposium on Thermal Problems in Biotechnology, American Society of Mechanical Engineers*, pp. 16–25.
4. Balasubramaniam, T. A., and Bowman, H. F., 1977, Thermal conductivity and thermal diffusivity of biomaterials: a simultaneous measurement technique, *Journal of Biomechanical Engineering*, 99, pp. 148–154.
5. Valvano, J. W. et al., 1984, An isolated rat liver model for the evaluation of thermal techniques to measure perfusion, *Journal of Biomechanical Engineering*, 106, pp. 187–191.
6. Arkin, H. et al., 1986, Thermal pulse decay method for simultaneous measurement of local thermal conductivity and blood perfusion, *Journal of Biomechanical Engineering*, 108, pp. 208–214.
7. Kress, R., 1987, A comparative analysis of thermal blood perfusion measurement techniques, *Journal of Biomechanical Engineering*, 109, pp. 218–225.
8. Patel, P. A. et al., 1987, A finite element analysis of a surface thermal probe, *Thermodynamics, Heat, and Mass Transfer in Biotechnology, American Society of Mechanical Engineers*, BED-5, pp. 95–102.
9. Bowman, H. F., 1985, Estimation of tissue blood flow, *Heat Transfer in Medicine and Biology: Analysis and Applications*, A. Shitzer, and R. C. Eberhart, eds., Plenum Press, New York, (pp. 193–230).
10. Chato, J. C., 1985, Measurement of thermal properties of biological materials, *Heat Transfer in Medicine and Biology: Analysis and Applications*, A. Shitzer, and R. C. Eberhart, eds., Plenum Press, New York, (pp. 167–192).
11. Valvano, J. W. et al., 1985, Thermal conductivity and diffusivity of biomaterials measured with self-heated thermistors, *International Journal of Thermophysics*, 6, pp. 301–311.
12. Valvano, J. W., and Chitsabesan, B., 1987, Thermal conductivity and diffusivity of arterial wall and atherosclerotic plaque, *Lasers Life Sciences*, 1, pp. 219–229.
13. Holmes, K. R., and Chen, M. M., 1980, In vivo tissue thermal conductivity and local blood perfusion measured with heat pulse-decay method, *Advances in Bioengineering*, V. C. Mow and M. K. Wells, eds., pp. 113–115.
14. Touloukian, Y. S. et al., 1970, *Thermophysical Properties of Matter: Thermal Conductivity*, IFI/Plenum, New York.
15. Touloukian, Y. S. et al., 1970, *Thermophysical Properties of Matter: Thermal Conductivity*, IFI/Plenum, New York, p. Preface.
16. Touloukian, Y. S. et al., 1970, *Thermophysical Properties of Matter: Thermal Diffusivity*, IFI/Plenum, New York, pp. 15a–42a.
17. Kreith, F., and Bohn, M. S., 1993, *Principles of Mass Transfer*, West Publishing, St. Paul, MN.
18. Duck, F. A., 1991, *Physical Properties of Tissue: A Comprehensive Reference Book*, Academic Press, London.
19. Chen, M. M. et al., 1981, Pulse-decay method for measuring the thermal conductivity of living tissue, *Journal of Biomechanical Engineering*, 103, pp. 253–260.
20. Holmes, K. R., and Chen, M. M., 1983, Local tissue heating, microbead pulse decay technique for heat transfer parameter evaluation, *Measurement of Blood Flow and Local Tissue Energy Production by Thermal Methods*, Muller-Schenburg, ed., Thieme-Stratton, New York.
21. Valvano, J. W. et al., 1983, A finite element analysis of self-heated noninvasive thermistors, *Advances in Bioengineering*, American Society of Mechanical Engineers, New York, (pp. 149–150).
22. Valvano, J. W. et al., 1984, The simultaneous measurement of thermal conductivity, thermal diffusivity and perfusion in small volume of tissue, *Journal of Biomechanical Engineering*, 106, pp. 192–197.
23. Valvano, J. W., Badeau, A. F., and Pearce, J. A., 1987, Simultaneous measurement of intrinsic and effective thermal conductivity, *American Society of Mechanical Engineers*, New York, pp. 31–36.
24. Patel, P. A. et al., 1987, A self-heated thermistor technique to measure effective thermal properties from the tissue surface, *Journal of Biomechanical Engineering*, 109, pp. 330–335.
25. Patel, P. A., Valvano, J. W., and Hayes, L. J., 1987, Perfusion measurement by a surface thermal probe, *IEEE Engineering in Medicine and Biology*, Vol. 1, pp. 28–29. Boston.
26. Walsh, J. T., 1984, *A Noninvasive Thermal Method for the Quantification of Tissue Perfusion*, Massachusetts Institute of Technology, Cambridge, MA.
27. Patera, P. A. et al., 1979, Prediction of tissue perfusion from measurement of the phase shift between heat flux and temperature, Paper #79-WA/HT-71, American Society of Mechanical Engineers, New York.

28. Pennes, H. H., 1948, Analysis of tissue and arterial blood temperatures in the resting forearm, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 1, pp. 93–122.
29. Rastorguev, Y. L., and Ganiev, Y. A., 1966, Thermal conductivity of aqueous solutions or organic materials, *Russian Journal of Physical Chemistry*, 40, pp. 869–871.
30. Valvano, J. W., 1988, Low temperature tissue thermal properties, *Low Temperature Biotechnology: Emerging Applications and Engineering Contributions*, J. J. McGrath and K.R. Diller, eds., American Society of Mechanical Engineers, New York, pp. 331–346.
31. Spells, K. E., 1960, The thermal conductivities of some biological fluids, *Physics in Medicine and Biology*, 5, pp. 139–153.
32. Cooper, T. E., and Trezck, G. J., 1971, Correlation of thermal properties of some human tissues with water content, *Aerospace Medicine*, 42, pp. 24–27.
33. Charney, C. K., 1992, Mathematical models of bioheat transfer, *Advances in Heat Transfer*, 22, pp. 19–155.
34. Wissler, E. H., 1961, Steady-state temperature distribution in man, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 16, pp. 734–740.
35. Wissler, E. H., 1985, Mathematical simulation of human thermal behavior using whole body models, *Heat Transfer in Medicine and Biology: Analysis and Applications*, Plenum Press, New York, pp. 325–373.
36. Chato, J. C., 1980, Heat transfer to blood vessels, *Journal of Biomechanical Engineering*, 102, pp. 110–118.
37. Chen, M. M., and Holmes, K. R., 1980, Microvascular contributions in tissue heat transfer, *Annals of the New York Academy of Sciences*, 335, pp. 137–150.
38. Weinbaum, S., Jiji, L., and Lemons, D. E., 1984, Theory and experiment for the effect of vascular temperature on surface tissue heat transfer I. Anatomical foundation and model conceptualization, *Journal of Biomechanical Engineering*, 106, pp. 246–251.
39. Weinbaum, S., Jiji, L., and Lemons, D. E., 1984, Theory and experiment for the effect of vascular temperature on surface tissue heat transfer II. Model formulation and solution, *Journal of Biomechanical Engineering*, 106, pp. 331–341.
40. Weinbaum, S., and Jiji, L., 1985, A new simplified bioheat equation for the effect of blood flow on average tissue temperature, *Journal of Biomechanical Engineering*, 107, pp. 131–139.
41. Zhu, L., Lemons, D. E., and Weinbaum, S., 1994, A new approach for predicting the enhancement in the effective conductivity of perfused tissue due to hyperthermia, *American Society of Mechanical Engineers*, 288, pp. 37–43.
42. Charny, C. K., Weinbaum, S., and Levin, R. L., 1990, An evaluation of the Weinbaum-Jiji bioheat equation for normal and hyperthermic conditions, *Journal of Biomechanical Engineering*, 112, pp. 80–87.
43. Weinbaum, S., and Jiji, L. M., 1979, A two phase theory for the influence of circulation on the heat transfer in surface tissue, *Advances in Bioengineering*, M. K. Wells, ed., American Society of Mechanical Engineers, New York, pp. 179–182.
44. Dagan, Z., Weinbaum, S., and Jiji, L. M., 1986, Parametric study of the three layer microcirculatory nodel for surface tissue energy exchange, *Journal of Biomechanical Engineering*, 108, pp. 89–96.
45. Song, W. J., Weinbaum, S., and Jiji, L. M., 1988, A combined macro and microvascular model for whole limb heat transfer, *Journal of Biomechanical Engineering*, 110, pp. 259–267.
46. Song, W. J., Weinbaum, S., and Jiji, L. M., 1987, A theoretical model for peripheral heat transfer using the bioheat equation of Weinbaum and Jiji, *Journal of Biomechanical Engineering*, 109, pp. 72–78.
47. Weinbaum, S., and Jiji, L. M., 1989, The matching of thermal fields surrounding countercurrent microvessels and the closure approximation in the Weinbaum-Jiji equation, *Journal of Biomechanical Engineering*, 111, pp. 271–275.
48. Weinbaum, S., and Lemons, D. E., 1992, Heat transfer in living tissue: the search for a blood-tissue energy equation and the local thermal microvascular control mechanism, *BioMedical Engineering Society Bulletin*, 16, pp. 38–43.
49. Zhu, M., Weinbaum, S., and Lemons, D. E., 1988, On the generalization of the Weinbaum-Jiji equation to microvessels of unequal size: the relation between the near field and local average tissue temperatures, *Journal of Biomechanical Engineering*, 110, pp. 74–81.
50. Lemons, D. E., Weinbaum, S., and Jiji, L. M., 1987, Experimental studies on the role of the micro and macro vascular system in tissue heat transfer, *American Journal of Physiology*, 253, p. R128.
51. Mooibroek, J., and Lagendijk, J. J. W., 1991, A fast and simple algorithm for the calculation of convective heat transfer by large vessels in 3-dimensional inhomogeneous tissue, *IEEE Transactions on Biomedical Engineering*, 38, pp. 490–501.

52. Baish, J. W., 1990, Heat transport by countercurrent blood vessels in the presence of an arbitrary temperature gradient, *Journal of Biomechanical Engineering*, 112, pp. 207–211.
53. Huang, H. W., Chen, Z. P., and Roemer, R. B., 1996, A countercurrent vascular network model of heat transfer in tissues, *Journal of Biomechanical Engineering*, 118, pp. 120–129.
54. Keller, K. H., and Seiler, L., 1971, An analysis of peripheral heat transfer in man, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 30, p. 779.
55. Mitchell, J. W., and Myers, G. E., 1968, An analytical model of the countercurrent heat exchange phenomena, *Biophysical Journal*, 8, pp. 897–911.
56. Baish, J. W., Ayyaswamy, P. S., and Foster, K. R., 1986, Heat transport mechanisms in vascular tissues: a model comparison, *Journal of Biomechanical Engineering*, 108, pp. 324–331.
57. Weinbaum, S., and Jiji, L. M., 1987, Discussion of papers by Wissler and Baish et al. concerning the Weinbaum-Jiji bioheat equation, *Journal of Biomechanical Engineering*, 109, pp. 234–237.
58. Wissler, E. H., 1987, Comments on Weinbaum and Jiji's discussion of their proposed bioheat equation, *Journal of Biomechanical Engineering*, 109, pp. 355–356.
59. Wissler, E. H., 1998, Pennes' 1948 paper revisited, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 85, pp. 35–41.
60. Huang, H. W., Chan, C. L., and Roemer, R. B., 1994, Analytical solutions of Pennes bioheat transfer equation with a blood vessel, *Journal of Biomechanical Engineering*, 116, pp. 208–212.
61. Xu, L. X., Chen, M. M., Holmes, K. R., and Arkin, H., 1991, The evaluation of the Pennes, the Chen-Holmes, the Weinbaum-Jiji bioheat transfer models in the pig kidney cortex, *ASME HTD, American Society of Mechanical Engineers*, 189, pp. 15–21.
62. Arkin, H., Xu, L. X., and Holmes, K. R., 1994, Recent developments in modeling heat transfer in blood perfused tissues, *IEEE Transactions on Biomedical Engineering*, 41, pp. 97–107.
63. Wulff, W., 1974, The energy conservation equation for living tissue, *IEEE Transactions on Biomedical Engineering*, 21, pp. 494–495.
64. Klinger, H. G., 1974, Heat transfer in perfused biological tissue I. General theory, *Bulletin of Mathematical Biology*, 36, pp. 403–415.
65. Bazett, H. C., and McGlone, B., 1927, Temperature gradients in the tissue in man, *American Journal of Physiology*, 82, p. 415.
66. Valvano, J. W., Nho, S., and Anderson, G. T., 1994, Analysis of the Weinbaum-Jiji model of blood flow in the canine kidney cortex for self-heated thermistors, *Journal of Biomechanical Engineering*, 16, pp. 201–207.
67. Brinck, H., and Werner, J., 1994, Estimation of the thermal effect of blood flow in a branching countercurrent network using a three-dimensional vascular model, *Journal of Biomechanical Engineering*, 116, pp. 324–330.
68. Anderson, G. T., and Valvano, J. W., 1989, An interlobular artery and vein based model for self-heated thermistor measurements of perfusion in the canine kidney cortex, *Bioheat Transfer – Applications in Hyperthermia, Emerging Horizons in Instrumentation and Modeling*, American Society of Mechanical Engineers, New York, Vol. HTD-126, pp. 29–35.
69. Anderson, G. T., and Valvano, J. W., 1994, A small artery heat transfer model for self-heated thermistor measurements of perfusion in the canine kidney cortex, *Journal of Biomechanical Engineering*, 116, pp. 71–78.
70. Yuan, D. Y., Zhu, L., Holmes, K. R., and Valvano, J. W., 1998, Perfusion and temperature measurements in hyperthermic canine prostates, *17th Southern Biomedical Engineering Conference*, San Antonio, TX, February 6–8, 1998, p. 85.
71. Xu, L. X., Zhu, L., and Holmes, K. R., 1998, Thermoregulation in canine prostate during transurethral microwave hyperthermia. II. Blood flow response, *International Journal of Hyperthermia*, 14, pp. 65–73.
72. Bligh, J., 1985, Regulation of body temperature in man and other mammals, *Heat Transfer in Medicine and Biology: Analysis and Applications*, A. Shitzer, and Robert C. Eberhart, eds., Plenum Press, New York, pp. 15–52.
73. Rowell, L. B., and Wyss, C. R., 1985, Temperature regulation in exercising and heat-stressed man, *Heat Transfer in Medicine and Biology: Analysis and Applications*, A. Shitzer, and Robert C. Eberhart, eds., Plenum Press, New York, pp. 53–78.
74. Lipton, J. M., 1985, Thermoregulation in pathological states, *Heat Transfer in Medicine and Biology: Analysis and Applications*, A. Shitzer, and Robert C. Eberhart, eds., Plenum Press, New York, pp. 79–106.
75. Heller, H. C., and Glotzbach, S. F., 1985, Thermoregulation and sleep, *Heat Transfer in Medicine and Biology: Analysis and Applications*, A. Shitzer, and Robert C. Eberhart, eds., Plenum Press, New York, pp. 107–134.

76. Chappuis, P., Pittet, P., and Jequier, E., 1976, Heat storage regulation in exercise during thermal transients, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 40, pp. 384–392.
77. Webb, P., 1995, The physiology of heat regulation, *American Journal of Physiology*, 268, pp. R838–R850.
78. Nunneley, S. A., 1970, Water cooled garments: a review, *Space Life Sciences*, 2, pp. 335–360.
79. Ganong, W. F., 1993, *Review of Medical Physiology*, Appleton & Lange, Norwalk, CT.
80. Fanger, P. O., 1970, *Thermal Comfort: Analysis and Applications in Environmental Engineering*, McGraw-Hill, New York.
81. Fulcher, C. W. G., 1970, Control of a liquid cooling garment for extravehicular astronauts by cutaneous and external auditory meatus temperatures, PhD Dissertation, University of Houston.
82. Hensen, J. L. M., 1990, Literature review on thermal comfort in transient conditions, *Building and Environment*, 25, pp. 309–316.
83. Wissler, E. H., 1970, Comparison of results obtained from two mathematical models. A simple 14-node model and a complex 250-node model, *Journal of Physiology (Paris)*, 63, pp. 455–458.
84. Stolwijk, J. A. J., and Hardy, J. D., 1966, Temperature regulation in man. A theoretical study, *Pflugers Archiv (European Journal of Physiology)*, 291, pp. 129–162.
85. Nadel, E. R., Mitchell, J. W., Saltin, B., and Stolwijk, J. A. J., 1971, Peripheral modifications to the central drive for sweating, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 31, pp. 828–833.
86. Mitchell, J. W., Galvez, T. L., Hengle, J., Myers, G. E., and Siebecker, K. L., 1970, Thermal response of human legs during cooling, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 29, pp. 859–865.
87. Hayward, J. S., Eckerson, J. D., and Collis, M. L., 1977, Thermoregulatory heat production in man: prediction equation based on skin and core temperatures, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 42, pp. 377–384.
88. Hayward, J. S., 1975, Thermal balance and survival time prediction of man in cold water, *Canadian Journal of Physiology and Pharmacology*, 53, pp. 21–32.
89. Kuznetz, L. H., 1979, A two-dimensional transient mathematical model of human thermoregulation, *American Journal of Physiology*, 6, pp. R266–R277.
90. Shitzer, A., and Chato, J. C., 1985, Thermal interaction with garments, *Heat Transfer in Medicine and Biology: Analysis and Applications*, Plenum Press, New York, pp. 375–394.
91. Shitzer, A., Chato, J. C., and Hertig, B. A., 1973, Thermal protective garment using independent regional control of coolant temperature, *Aerospace Medicine*, 44, pp. 49–59.
92. Nyberg, K. L., Diller, K. R., and Wissler, E. H., 1997, Modeling of human regulation for liquid cooling garment applications, *Advances in Heat and Mass Transfer, American Society of Mechanical Engineers*, Vol. HTD-355 and BED-37, pp. 119–126.
93. Geddes, L. A., and Baker, L. E., 1967, The specific resistance of biological material. A compendium of data for the biomedical engineer and physiologist, *Medical and Biological Engineering*, 5, pp. 271–293.
94. Schwan, H. P., 1972, Biophysics of diathermy, *Therapeutic Heat and Cold*, E. Licht, ed., Waverly Press, Baltimore, MD.
95. Fawcett, D. W., 1994, *Textbook of Histology*, Chapman & Hall, New York.
96. Hall, J. E., 2016, *Guyton and Hall Textbook of Physiology*, Elsevier-Saunders, Philadelphia, PA.
97. Stillwell, G. K., 1972, General principles of thermal therapy, *Therapeutic Heat and Cold*, E. Licht, ed., Waverly Press, Baltimore MD.
98. Green, D. M., and Diller, K. R., 1978, Measurement of burn-induced leakage of macromolecules in living tissue, *Journal of Biomechanical Engineering*, 100, pp. 153–158.
99. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P., 2008, *Molecular Biology of the Cell*, Garland Science, New York.
100. Moritz, A. R., 1947, Studies of thermal injury III. The pathology and pathogenesis of cutaneous burns: an experimental study, *American Journal of Pathology*, 23, pp. 915–934.
101. Moritz, A. R., and Henriques, F. C., 1947, Studies in thermal injury II: the relative importance of time and surface temperature in the causation of cutaneous burns, *American Journal of Pathology*, 23, pp. 695–720.
102. Henriques, F. C., 1947, Studies of thermal Injury, V. The predictability and significance of thermally induced rate processes leading to irreversible epidermal injury, *Archives of Pathology*, 43, pp. 489–502.
103. Henriques, F. C., and Moritz, A. R., 1947, Studies of thermal injury in the conduction of heat to and through skin and the temperatures attained therein: a theoretical and experimental investigation, *American Journal of Pathology*, 23, pp. 531–549.

104. Büttner, K., 1951, Effects of extreme heat and cold on human skin. I. Analysis of temperature changes caused by different kinds of heat application, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 3, pp. 691–702.
105. Stoll, A. M., 1960, A computer solution for determination of thermal tissue damage integrals from experimental data, *IRE Transactions on Medical Electronics*, 7, pp. 355–358.
106. Weaver, J. A., and Stoll, A. M., 1967, Mathematical model of skin exposed to thermal radiation, *Aerospace Medicine*, 40, pp. 24–30.
107. Takata, A. N., 1974, Development of criterion for skin burns, *Aerospace Medicine*, 45, pp. 634–637.
108. Stoll, A. M., and Green, L. C., 1959, Relationship between pain and tissue damage due to thermal radiation, (republished for 50th anniversary issue of *Journal of Applied Physiology*, 85, 5–34, 1998), 14, pp. 373–382.
109. Stoll, A. M., and Chianta, M. A., 1968, Burn production and prevention in convective and radiant heat transfer, *Aerospace Medicine* (39), pp. 1232–1238.
110. Lawrence, J. C. a. B., J.P., 1976, Thermal conditions which cause skin burns, *Engineering in Medicine*, 5, pp. 61–63.
111. Ross, D. C., and Diller, K. R., 1976, An experimental investigation of burn injury in living tissue, *Journal of Heat Transfer*, 98, pp. 292–296.
112. Moussa, N. A., McGrath, J. J., Cravalho, E. G., and Asimacopoulos, P. J., 1979, Kinetics of thermal injury in cells, *Journal of Biomechanical Engineering*, 101, pp. 213–217.
113. Moussa, N. A., Tell, E. N., and Cravalho, E. G., 1979, Time progression of hemolysis of erythrocyte populations exposed to supraphysiological temperatures, *Journal of Biomechanical Engineering*, 101, pp. 213–217.
114. Bhowmick, S., Swanlund, D. J., and Bischof, J. C. 2000, Supraphysiological thermal injury in dunning AT-1 prostate tumor cells, J*ournal of Biomechanical Engineering*, 122(1), pp. 51–59.
115. Maron, S. H., and Lando, J. B., 1974, *Fundamentals of Physical Chemistry*, Macmillan, New York.
116. Eyring, H., Gershinowitz, H., and Sun, C. E., 1935, The absolute rate of homogeneous atomic reactions, *Journal of Chemical Physics*, 3(12), pp. 786–796.
117. Eyring, H., and Polanyi, M., 1931, Über Einfache Gasreaktionen (On simple gas reactions), *Zeitschrift fur Physikalische Chemie B*, 12, pp. 279–311.
118. He, X., and Bischof, J. C., 2003, Quantification of temperature and injury response in thermal therapy and cryosurgery, *Critical Reviews in Biomedical Engineering*, 31(5 and 6), pp. 355–421.
119. Wright, N. T., 2003, On a relationship between the Arrhenius parameters from thermal damage studies, *Journal of Biomechanical Engineering*, 125(2), pp. 300–304.
120. Rosenberg, B., Kemeny, G., Switzer, R. C., and Hamilton, T. C., 1971, Quantitative evidence for protein denaturation as the cause of thermal death, *Nature*, 232, pp. 471–473.
121. Barrie, P. J., 2012, The mathematical origins of the kinetic compensation effect: 2. The effect of systematic errors, *Physical Chemistry and Chemical Physics*, 14(1), pp. 327–336.
122. Barrie, P. J., 2012, The mathematical origins of the kinetic compensation effect: 1. The effect of random experimental errors, *Physical Chemistry and Chemical Physics*, 14(1), pp. 318–326.
123. Yelon, A., Sacher, E., and Linert, W., 2012, Comment on 'The mathematical origins of the kinetic compensation effect' Parts 1 and 2 by P. J. Barrie, *Physical Chemistry and Chemical Physics*, 14(22), pp. 8232–8234.
124. Barrie, P. J., 2012, Reply to 'Comment on the mathematical origins of the kinetic compensation effect' Parts 1 and 2' by A. Yelon, E. Sacher and W. Linert, *Physical Chemistry and Chemical Physics*, 14(22), pp. 8235–8236.
125. Miles, C. A., and Ghelashvili, M., 1999, Polymer-in-a-box mechanism for the thermal stabilization of collagen molecules in fibers, *Biophysics Journal*, 76, pp. 3243–3252.
126. Qin, Z. P., Balasubramanian, S. K., Wolkers, W. F., Pearce, J. A., and Bischof, J. C., 2014, Correlated parameter fit of arrhenius model for thermal denaturation of proteins and cells, *Annals of Biomedical Engineering*, 42(12), pp. 2392–2404.
127. Diller, K. R., and Klutke, G. A., 1993, Accuracy analysis of the Henriques model for predicting burn injury, *Advances in Bioheat and Mass Transfer*, Roemer, R. B. ed., American Society of Mechanical Engineers, New York, Vol. HTD-268, pp. 117–123.
128. Xu, Y., and Qian, R., 1995, Analysis of thermal injury process based on enzyme deactivation mechanisms, *Journal of Biomechanical Engineering*, 117, pp. 462–465.
129. Weinberg, W. A., 2007, *The Biology of Cancer,* Garland Science, Taylor & Francis Group, New York.

130. Vanden Berghe, T., Vanlangenakker, N., Parthoens, E., Deckers, W., Devos, M., Festjens, N., Guerin, C. J., Brunk, U. T., Declercq, W., and Vandenabeele, P., 2010, Necroptosis, necrosis and secondary necrosis converge on similar cellular disintegration features, *Cell Death and Differentiation*, 17(6), pp. 922–930.
131. Vandenabeele, P., Galluzzi, L., Vanden Berghe, T., and Kroemer, G., 2010, Molecular mechanisms of necroptosis: an ordered cellular explosion, *Nature Reviews. Molecular Cell Biology*, 11(10), pp. 700–714.
132. Zhang, Y., and Calderwood, S. K., 2011, Autophagy, protein aggregation and hyperthermia: a mini-review, *International Journal of Hyperthermia*, 27(5), pp. 409–414.
133. Bergsbaken, T., Fink, S. L., den Hartigh, A. B., Loomis, W. P., and Cookson, B. T., 2011, Coordinated host responses during pyroptosis: caspase-1-dependent lysosome exocytosis and inflammatory cytokine maturation, *Journal of Immunology*, 187(5), pp. 2748–2754.
134. Mathews, C. K., van Holde, K. E., Appling, D. R., and Anthony-Cahill, S. J., 2013, *Biochemistry*, Pearson, Toronto, ON.
135. Pearce, J. A., 2013, Comparative analysis of mathematical models of cell death and thermal damage processes, *International Journal of Hyperthermia*, 29(4), pp. 262–280.
136. Kumar, V., Abbas, A. K., and aster, J. C., 2014, *Robins and Cotran: Pathologic Basis of Disease*, Elsevier-Saunders, Philadelphia, PA.
137. Eissing, T., Conzelmann, H., Gilles, E. D., Allgower, F., Bullinger, E., and Scheurich, P., 2004, Bistability analyses of a caspase activation model for receptor-induced apoptosis, *Journal of Biological Chemistry*, 279(35), pp. 36892–36897.
138. Anderson, J. L., and Kornbluth, S., 2013, The tangled circuitry of metabolism and apoptosis, *Molecular Cell*, 49(3), pp. 399–410.
139. Sapareto, S. A., Hopwood, L. E., and Dewey, W. C., 1978, Combined effects of x irradiation and hyperthermia on CHO cells for various temperatures and orders of application, *Radiation Research*, 73(2), pp. 221–233.
140. Pearce, J. A., 2015, Improving accuracy in arrhenius models of cell death: adding a temperature-dependent time delay, *Journal of Biomechanical Engineering*, 137(6), p. 121006.
141. Feng, Y., Oden, J. T., and Rylander, M. N., 2008, A two-state cell damage model under hyperthermic conditions: Theory and in vitro experiments, *Journal of Biomechanical Engineering*, 130(4), p. 041016.
142. Bhowmick, S., Swanlund, D. J., and Bischof, J. C., 2004, In vitro thermal therapy of AT-1 Dunning prostate tumours, *International Journal of Hyperthermia*, 20(1), pp. 73–92.
143. He, X., and Bischof, J. C., 2005, The kinetics of thermal injury in human renal carcinoma cells, *Annals of Biomedical Engineering*, 33(4), pp. 502–510.
144. Thomsen, S., 2000, Qualitative and quantitative pathology of clinically relevant thermal lesions, *Critical Reviews of Optical Science and Technology*, Ryan, T. P. ed., Society of Photo-optical Instrumentation Engineers, Bellingham, WA, pp. 425–458.
145. Thomsen, S., 2003, Practical pathology for engineers: how to do the job right the first time, *Thermal Treatment of Tissue: Energy Delivdery and Assessment II*, Society of Photo-optical Instrumentation Engineers, San Jose, CA, pp. 1–26.
146. Thomsen, S., and Pearce, J. A., 2011, *Thermal Damage and Rate Processes, Optical-Thermal Response of Laser Irradiated Tissue*. 2nd ed., A. J. Welch, and M. J. D. vanGemert, eds., Springer, New York, pp. 487–549.
147. Pearse, A. G. E., 1980, *Histochemistry: Theoretical and Applied*, Churchill Livingstone, New York.
148. Ghadially, F. N., 1988, *Ultrastructural Pathology of the Cell and Matrix*, Butterworth Publishers, Boston, MA.
149. Thomsen, S., 1995, Identification of lethal thermal injury at the time of photothermal treatment, *Laser-Induced Interstitial Thermotherapy*, Muller, G. and Roggan, A., eds., Society of Photo-Optical Instrumentation Engineers, Bellingham, WA, pp. 459–467.
150. Thomsen, S., 1999, Mapping thermal injury in biologic tissues using quantitative pathologic techniques, *Photonics West, BiOS*, Society of Photo-Optical Instrumentation Engineers, San Jose, CA, pp. 82–97.
151. Ingber, D. E., 1998, The architecture of life, *Scientific American*, 278, pp. 48–57.
152. Schober, R., Ullrich, F., Sander, T., Dunselen, H., and Hessel, S., 1986, Laser-induced alteration of collagen substructures allows microsurgical tissue welding, *Science*, 232, pp. 1421–1422.
153. Thomsen, S., Pearce, J. A., and Cheong, W.-F., 1989, Changes in birefringence as markers of thermal damage in tissues, *IEEE Transactions on Biomedical Engineering*, 36, pp. 1174–1179.
154. Fawcett, D. W., 1994, *A Textbook of Histology*, W.B. Saunders, Philadelphia, PA.
155. Schmidt, S. J., 1937, Die doppelbrechung von karyoplasma, zytoplasma und metaplasma, *Protoplasma-Monographien*, Heilbrunn, L. V. ed., Verlag von Gebruder Borntraeger, Berlin, pp. 154–267.

156. Fisher, E., 1944, The birefringence of striated and smooth muscles, *Journal of Cellular and Comparative Physiology*, 23, pp. 110–130.
157. Ramachandran, G. N., and Ramakrishnan, C., 1976, Molecular structure, *Biochemistry of Collagen*, A. H. Ramachandran G.N. and Reddi, ed., Plenum Press, New York, pp. 45–84.
158. Miller, E. J., and Gay, S., 1987, Collagens in structural and contractile proteins: Part D, extracellular matrix, *Methods in Enzymology*, 144, pp. 3–171.
159. Cohen, K. I., Diegelmann, R.F., and Linblad, W.J., ed., 1992, Structural and regulatory components of wound healing, *Wound Healing: Biochemical and Clinical Aspects*, W.B. Saunders, Philadelphia, PA, pp. 129–151.
160. Peckham, M., and Irving, M., 1989, Myosin crossbridge orientation in demembranated muscle fibres studied by birefringence and x-ray diffraction measurements, *Journal of Molecular Biology*, 210, pp. 113–126.
161. Harris, P., and Heath, D., 1986, Structure and function of vascular smooth muscle, *The Human Pulmonary Circulation: Its Form and Function in Health and Disease*, Churchill Livingstone, New York, pp. 161–182.
162. Canham, P. B., Finlay, H. M., Whittaker, P., and Starkey, J., 1986, The tunica musclularis of human brain arteries: three-dimensional measurements of alignment of the smooth muscle mechanical axis by polarized light and the universal stage, *Neurological Research*, 8, pp. 66–74.
163. Hulmes, D. J. S., Miller, A., Parry, D. A. D., Piez, K. A., and Woodhead-Galloway, J., 1985, Crystalline regions in collagen fibrils, *Journal of Molecular Biology*, 184, pp. 473–477.
164. Flory, P., and Garrett, R. R., 1958, Phase transition in collagen and gelatin systems, *Journal of the American Chemistry Society*, 80, pp. 4836–4845.
165. Wood, G. C., 1963, Spectral changes accompanying the thermal denaturation of collagen, *Biochemical and Biophysical Research Communications*, 13, pp. 95–99.
166. Deak, G., and Romhanyi, G., 1967, The thermal shrinkage process of collagen fibers as revealed by polarization optical analysis of topo-optical staining reactions, *Acta Morphologica Academiae Scientiarum Hungaricae*, 15, pp. 195–200.
167. Lim, J. J., 1976, Transition temperature and enthalpy change dependence on stabilizing and destabilizing ions in the helix-coil transition in native tendon collagen, *Biopolymers*, 15, pp. 2371–2381.
168. Bosman, S., Pickering, J. W., vanMarle, J., and vanGemert, M. J. C., 1995, Ultrastructural alterations in heated canine myocardium, *Lasers in Surgery & Medicine*, 17, pp. 39–48.
169. Thomsen, S., Jacques, S. L., and Flock, S., 1990, Microscopic correlates of macroscopic optical property changes during thermal coagulation of myocardium, *Laser-Tissue Interaction*, Society of Photo-optical Instrumentatin Engineers, Los Angeles, CA, pp. 2–11.
170. Junqueira, L. C. U., Bignolas, G., and Brentani, R. R., 1979, Picrosirius staining plus polarization microscopy, a specific method for collagen detection in tissue sections, *Histochemical Journal* 11, pp. 447–455.
171. Chen, S. S., Wright, N. T., and Humphrey, J. D., 1997, Heat-induced changes in the mechanics of a collagenous tissue: Isothermal free shrinkage, *Journal of Biomechanical Engineering*, 119(4), pp. 372–378.
172. Chen, S. S., Wright, N. T., and Humphrey, J. D., 1998, Heat-induced changes in the mechanics of a collagenous tissue: isothermal, isotonic shrinkage, *Journal of Biomechanical Engineering*, 120, pp. 382–388.
173. Chen, S. S., Wright, N. T., and Humphrey, J. D., 1998, Phenomenological evolution equations for heat-induced shrinkage of a collagenous tissue, *IEEE Transactions on Biomedical Engineering*, 45, pp. 1234–1240.
174. McKenzie, A. L., 1986, A three-zone model of soft-tissue damage by a CO_2 laser, *Physics in Medicien & Biology*, 31, pp. 967–983.
175. Partovi, F., Izatt, J. A., Cothren, R. M., Kittrell, C., Thomas, J. E., Strikwerda, S., Kramer, J. R., and Feld, M. S., 1987, A model for thermal ablation of biological tissue using laser radiation, *Lasers in Surgery & Medicine*, 7, pp. 141–154.
176. Rastegar, S., Motamedi, M., Welch, A. J., and Hayes, L. J., 1989, A theoretical study of the effect of optical properties in laser ablation of tissue, *IEEE Transactions on Biomedical Engineering*, 36, pp. 1180–1187.
177. LeCarpentier, G. L., Motamedi, M., Rastegar, S., and Welch, A. J., 1989, Simultaneous analysis of thermal and mechanical events during cw laser ablation of biological media, *Photonics West BiOS, Society of Photo-optical Instrumentation Engineers*, Los Angeles, CA, pp. 107–113.
178. Verdaasdonk, R. M., Borst, C., and vanGemert, M. J. C., 1990, Explosive onset of continuous wave laser tissue ablation, *Physics in Medicine & Biology*, 35, pp. 1129–1144.

179. Van Leeuwen, T. G., van der Veen, J. J., Verdaasdonk, R. M., and Borst, C., 1991, Non-contact tissue ablation by holmium:YAG laser pulsed in blood, *Lasers in Surgery & Medicine*, 11, pp. 26–34.
180. Gijsbers, G. H. M., Selten, F. M., and vanGemert, M. J. C., 1991, CW laser ablation velocities as a function of absorption in and experimental one-dimensional tissue model, *Lasers in Surgery & Medicine*, 11, pp. 287–296.
181. Zweig, A. D., 1991, Infrared tissue ablation: consequences of liquefaction, *Laser-Tissue Interaction, Society of Photo-optical Instrumentation Engineers*, Los Angeles, CA, pp. 2–8.
182. Frenz, M. C., Greber, M., Romano Forrer, V. M., and Weber, H. P., 1991, Damage induced by pulsed IR laser radiation at transitions between different tissues, *Laser-Tissue Interaction, Society of Photo-optical Instrumentation Engineers*, Los Angeles, CA, pp. 9–15.
183. Pearce, J. A., and Thomsen, S., 1992, Kinetic models of tissue fusion processes, *Laser-Tissue Interaction 3, Society of Photo-optical Instrumentation Engineers*, Los Angeles, CA, pp. 251–260.
184. Zweig, A. D., Meierhofer, B., Muller, O. M., Mischler, C., Romano, V., Frenz, M., and Weber, H. P., 1990, Lateral damage along pulsed laser incisions, *Lasers in Surgery & Medicine*, 10, pp. 262–274.
185. Van Leeuwen, T. G., Van Erven, L., Meertens, J. H., Motamedi, M., Post, M. J., and Borst, C., 1992, Origin of wall dissections induced by pulsed excimer and mid-infrared laser ablation in the pig, *Journal of the American College of Cardiology*, 19(7), pp. 1610–1618.
186. Dabby, F. W., and Paek, U., 1972, High-intensity laser-induced vaporization and explosion of solid material, *IEEE Journal of Quantum Electronics*, 8, pp. 106–111.
187. Welch, A. J., and Polhamus, G. D., 1984, Measurement and prediction of thermal injury in the retina of Rhesus monkey, *IEEE Transactions on Biomedical Engineering*, 31, pp. 633–644.
188. Takata, A. N., 1974, Thermal model of laser-induced eye damage, Final Report. Contract no. F41609-74-C-0005, IIT Research Institute, Chicago, IL.
189. Birngruber, R., 1980, Thermal modeling in biological tissue, *Lasers in Biology and Medicine*, F. Hillenkamp, Pratesi, R., and Sacchi, C.A., ed., Plenum Press, New York, pp. 77–97.
190. Birngruber, R., Hillenkamp, F., and Gabel, V.-P., 1985, Theoretical investigations of laser thermal retinal injury, *Health Physics*, 48(6), pp. 781–796.
191. Yang, Y., Welch, A. J., and Rylander, H. G. I., 1991, Rate process parameters of albumen, *Lasers in Surgery & Medicine*, 11, pp. 188–190.
192. Mixter, G., Delhry, G. P., Derksen, W. L., and Monahan, T., 1963, The Influence of time on the death of HeLa cells at elevated temperature, *Temperature: Its Measurement and Control in Science and Industry*, J. D. Hardy, ed., Reinhold, New York.
193. Padanilam, J., Bischof, J. C., Lee, R. C., Cravalho, E. G., Tompkins, R. G., Yarmush, M. L., and Toner, M., 1994, Effectiveness of Poloxamer 188 in arresting calcein leakage from thermally damaged isolated skeletal muscle cells, *Annals of the* N*ew York Academy of Sciences*, 720, pp. 111–123.
194. Polge, C., Smith, A. U., and Parkes, A. S., 1949, Revival of spermatozoa after vitrification and dehydration at low temperatures, *Nature*, 164, p. 666.
195. Lovelock, J. E., 1953, The mechanism of the protective action of glycerol against haemolysis by freezing and thawing, *Biochimica et Biophysica Acta*, 11, pp. 28–36.
196. Strumia, M. M., Clawell, L. S., and Strumia, P. V., 1960, The preservation of blood for transfusion, *Journal of Laboratory and Clinical Medicine*, 56, pp. 576–593.
197. Whittingham, D. G., Leibo, S. P., and Mazur, P., 1972, Survival of mouse embryos frozen to –196°C and –296°C, *Science*, 178, pp. 411–414.
198. McGrath, J. J., and Diller, K. R., 1988, *Low Temperature Biotechnology: Emerging Applications and Engineering Contributions*, American Society of Mechanical Engineers, New York, Vol. HTD-98 and BED-10.
199. Mazur, P., 1963, Kinetics of water loss from cells at subzero temperatures and the likelihood of intracellular freezing, *Journal of General Physiology*, 47, pp. 347–369.
200. Mazur, P., 1990, Equilibrium, quasi-equilibrium and nonequilibrium freezing of mammalian embryos, *Cell Biophysics*, 17, pp. 53–92.
201. Kedem, O., and Katchalsky, A., 1958, Thermodynamic analysis of permeability of biological membranes to nonelectrolytes, *Biochimica et Biophysica Acta*, 27, pp. 229–246.
202. McGrath, J. J., 1988, Membrane transport properties, *Low Temperature Biotechnology: Emerging Applications and Engineering Contributions*, J. J. McGrath, and K. R. Diller, eds., American Society of Mechanical Engineers, New York, pp. 273–330.
203. Oster, G. F., Perelson, A. S., and Katchalsky, A., 1973, Network thermodynamic modelling of biophysical systems, *Quantum Reviews in Biophysics*, 6, pp. 1–134.

204. Diller, K. R., 1988, Energy based network modeling for cryobiology, *Low Temperature Biotechnology: Emerging Applications and Engineering Contributions*, J. J. McGrath, and K. R. Diller, eds., American Society of Mechanical Engineers, New York, pp. 189–202.
205. Diller, K. R., Beaman, J. J., Montoya, J. P., and Breedfeld, P. C., 1988, Network thermodynamic modeling with bond graphs for membrane transport during cell freezing, *Journal of Heat Transfer*, 110, pp. 938–945.
206. deFreitas, R. C., Diller, K. R., Lachenbruch, C. A., and Merchant, F. A., 1998, Network thermodynamic model of coupled transport in a multicellular tissue — the islet of Langerhans, *Annals of the New York Academy of Sciences*, 858, pp. 191–204.
207. Walsh, J. R., Diller, K. R., and Brand, J. J., 1997, Osmomechanical behavior of plant cells across regions of elastic and osmotic stress, *Advances in Heat and Mass Transfer in Biotechnology*, pp. 15–23.
208. Lachenbruch, C. A., Pegg, D. E., and Diller, K. R., 1998, Sensitivity of kidney perfusion protocol design to physical and physiological parameters, *Annals of the New York Academy of Sciences*, 858, pp. 298–309.
209. Walsh, J. R., Diller, K. R., and Brand, J. J., 2004, Measurement and simulation of water and methanol transport in algal cells, *Journal of Biomechanical Engineering*, 126(2), pp. 167–179.
210. Diller, K. R., and Cravalho, E. G., 1970, A cryomicroscope for the study of freezing and thawing processes in biological cells, *Cryobiology*, 7, pp. 191–199.
211. Diller, K. R., 1982, Quantitative low temperature optical microscopy of biological systems, *Journal of Microscopy*, 126, pp. 9–28.
212. Diller, K. R., 1988, Cryomicroscopy, *Low Temperature Biotechnology: Emerging Applications and Engineering Contributions*, J. J. McGrath, and K. R. Diller, eds., pp. 347–362.
213. McGrath, J. J., 1985, A microscope diffusion chamber for the determination of the equilibrium and non-equilibrium response of individual cells, *Journal of Microscopy*, 139, pp. 249–263.
214. Walcerz, D. B., and Diller, K. R., 1991, Quantitative light microscopy of combined perfusion and freezing processes, *Journal of Microscopy*, 161, pp. 297–311.
215. deFreitas, R. C., Diller, K. R., Lakey, J. R. T., and Rajotte, R. V., 1997, Osmotic behavior and transport properties of human islets in a dimethyl sulfoxide solution, *Cryobiology*, 35, pp. 230–239.
216. Levin, R. L., 1988, Osmotic behavior of cells during freezing and thawing, *Low Temperature Biotechnology: Emerging Applications and Engineering Contributions*, J. J. McGrath, and K. R. Diller, eds., American Society of Mechanical Engineers, New York, pp. 177–188.
217. Kasharin, A. V., and Karlsson, J. O. M., 1998, Analysis of mass transport during warming of cryopreserved cells, *Annals of the New York Academy of Sciences*, 858, pp. 163–174.
218. Fahy, G. M., 1988, Vitrification, *Temperature Biotechnology: Emerging Applications and Engineering Contributions*, J. J. McGrath, and K. R. Diller, eds., American Society of Mechanical Engineers, New York, pp. 113–146.
219. Steponkus, P. L., 1984, The role of the plasma membrane in freezing injury and cold acclimation, *Annual Review of Plant Physiology*, 35, pp. 543–584.
220. Kurz, W., and Fisher, D. J., 1992, *Fundamentals of Solidification*, Trans Tech Publications, Aedermannsdorf, Switzerland.
221. Hayes, L. J., Diller, K. R., Chang, H. J., and Lee, H. S., 1988, Prediction of local cooling rates and cell survival during the freezing of cylindrical specimens, *Cryobiology*, 25, pp. 67–82.
222. Fahy, G. M., 1990, Physical problems with the vitrification of large biological systems, *Cryobiology*, 27, pp. 492–510.
223. Rubinsky, B., Cravalho, E. G., and Mikic, B., 1980, Thermal stresses in frozen organs, *Cryobiology*, 17, pp. 66–73.
224. Fowler, A. J., and Toner, M., 1998, Prevention of hemolysis in rapidly frozen erythrocytes by using a laser pulse, *Annals of the* New *York Academy of Sciences*, 858, pp. 245–252.
225. Nerem, R. M., and Sambanis, A., 1995, Tissue engineering: from biology to biological substitutes, *Tussue Engineering*, 1, pp. 3–13.
226. Rubinsky, B., 1995, Biophysical and bioengineering aspects of cryosurgery, *Cryobiology and Cryotechnology*, 41, pp. 67–81.
227. Onik, G., and Rubinsky, B., 1988, Cryosurgery: new developments in understanding and technique, *Low Temperature Biotechnology: Emerging Applications and Engineering Contributions*, J. J. McGrath, and K. R. Diller, eds., American Society of Mechanical Engineers, New York, pp. 57–80.
228. Onik, G., Gilbert, J., Hoddick, W., Filly, R., Callen, P., Rubinsky, B., and Christianson, M., 1984, Ultrasonic characteristics of frozen liver, *Cryobiology*, 21, pp. 331–335.

229. Rubinsky, B., Gilbert, J. C., Onik, G., Roos, H. S., Wong, S. T. S., and Brennan, K. M., 1993, Monitoring cryosurgery in the brain and in the prostate with proton NMR, *Cryobiology*, 30, pp. 191–199.
230. Otten, D. M., Rubinsky, B., Cheong, W.-F., and Benaron, D. A., 1998, Ice front propagation monitoring in tissue by the use of visible-light spectroscopy, *Applied Optics*, 37, pp. 6006–6010.
231. Baust, J. G., and Chang, Z., 1995, Underlying mechanisms of damage and new concept in cryosurgical instrumentation, *Cryosurgery: Mechanism and Applications*, IIR Press, Paris, pp. 21–36.
232. Rubinsky, B., and Pegg, D. E., 1989, A mathematical model for the freezing process in biological tissue, *Proceedings of the Royal Society of London*, B234, pp. 343–358.
233. Bischof, J. C., and Rubinsky, B., 1993, Microscale heat and mass transfer of vascular and intracellular freezing in the liver, *Journal of Heat Transfer*, 115, pp. 1029–1035.
234. Pazhayannur, P. V., and Bischof, J. C., 1997, Measurement and simulation of water transport during freezing in mammalian liver tissue, *Journal of Biomechanical Engineering*, 119, pp. 269–277.
235. Koushafar, H., and Rubinsky, B., 1997, Effect of antifreeze proteins on frozen primary prostatic adenocarcinoma cells, *Urology*, 49, pp. 421–425.
236. Koushafar, H., Pham, L., Lee, C., and Rubinsky, B., 1997, Chemical adjuvant cryosurgery with antifreeze proteins, *Journal of Surgical Oncology*, 66, pp. 114–121.
237. DeVries, A. L., 1988, The role of antifreeze glycopeptides and peptides in the freezing avoidance of antarctic fishes, *Comparative Biochemistry and Physiology*, 90, pp. 611–621.
238. Ananthanarayanan, V. S., 1989, Antifreeze proteins: structural diversity and mechanism of action, *Life Chemistry Reports*, 7, pp. 1–32.
239. Ishiguro, H., and Rubinsky, B., 1998, Influence of fish antifreeze proteins on the freezing of cell suspensions with cryoprotectant penetrating cells, *International Journal of Heat and Mass Transfer*, 41, pp. 1907–1915.
240. Diller, K. R., and Ryan, T. P., 1998, Heat transfer in living systems: Current opportunities, *Journal of Heat Transfer*, 120, pp. 810–829.
241. Holmes, K.R., 1998, Thermal conductivity of selected tissues, in *Biotransport: Heat and Mass Transfer in Living Systems*, Diller, K.R., ed., New York Academy of Science, New York.

4.6 THERMAL INSULATION

David W. Yarbrough

Introduction

In Fourier's law of steady-state heat conduction in one dimension, $Q = -k \cdot A \cdot (dT/dx)$, k, the material thermal conductivity, relates the flow of heat to the temperature gradient in the material. In many cases, the thermal conductivity is treated as a constant, but for more general cases, it can be a function of temperature, material density, or direction of heat flow. Thermal insulation design is an area of engineering that deals with the minimization of heat flux and control of temperature through proper selection and use of the material thermal conductivity of insulation surrounding or adjacent to a region or body.

Thermal insulation is used either to minimize the energy loss from a body or area such as a building, or to control the surface temperature, for example, of the outer shell of a furnace for safety reasons. It can also be used to control an object's rate of temperature change, for example, biological material, while it is being frozen or thawed.

Thermal insulation is important for all temperature ranges, from near absolute zero in cryogenic applications up to the highest melting point materials which can be used as insulation, ~3000 K. Thermal insulation beyond the highest melting point is no longer a material consideration, since solid materials can no longer exist in that condition and still function as insulators, but rather become purely a heat transfer phenomena involving radiation and convection. This is the case in gas plasmas where temperatures exceed 10,000 K.

The thermal conductivity of materials that could conceivably be used as insulation is shown as a function of temperature in Figure 4.6.1. At very low temperatures, the thermal conductivity approaches zero for all materials. At a few degrees above absolute zero, the thermal conductivity of materials such as pure copper or aluminum exhibits a spike of several thousand W/m·K in the thermal conductivity and then quickly decreases as temperature increases. As the temperature is further increased, the conductivity of metals ranges up to 50–100 W/(m·K). The thermal conductivity of materials can range from 10^{-6} to 1000 W/(m·K). This wide range of thermal conductivity provides the means to control temperatures or heat flows for a wide variety of applications.

Heat Transfer in Thermal Insulation

Efficient use of thermal insulation requires an understanding of the heat transfer that occurs in insulations. Heat transfer occurs by three mechanisms (taken to be independent): conduction through the solid material and interstitial gas in the insulation, gas-phase convection, and thermal radiation. Each will be discussed separately and then combined to give an overall apparent thermal conductivity. Additional discussion can be found in Refs. [1–3].

Solid Conduction

Solid conduction occurs as heat is transferred through the solid portions of the insulation material, from the hot to the cold boundaries. It is characterized by the thermal conductivity of the continuous solid material of the insulation. In a fibrous insulation, such as mineral fiber insulation, this is the conductivity of the material which forms the fibers, the type of contact between the fibers, and the number of fibers per unit volume, i.e., the insulation density. The solid conductivity of the insulation can be modeled by the following equation:

$$k_s = A \cdot k_m \tag{4.6.1}$$

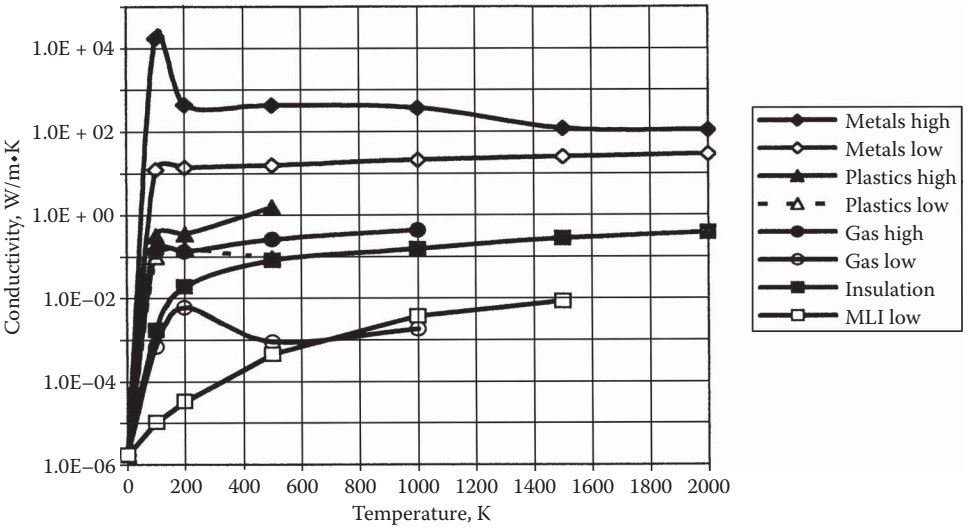


FIGURE 4.6.1 Material thermal conductivity ranges.

where k_s = the thermal conductivity of the insulation material

k_m = the thermal conductivity of the solid-phase material

A = empirical constant which can be related to the load on the insulation, the nature of the contacts or structure of the solid phase, and temperature.

The reader is referred to Refs. [4–6] for models that describe this interaction.

Gas Conduction and Convection

Gas-phase convection is normally suppressed in fibrous or cellular industrial insulation materials because of the insulation density, the closed-cell properties of some solid insulations, or the gas being at a low density in evacuated insulation. In some insulation systems, such as double- or triple-pane windows or low-density open-cell or fibrous insulation, convection can be of importance and will be included in the manufacturers' data sheets showing thermal resistance values. Gas conduction is of importance in most insulations.

Gas conduction at higher pressure ranges will follow a similar equation as solid conduction:

$$k_g = Bk'_g \tag{4.6.2}$$

where k_g = the thermal conductivity of the gas

k'_g = the thermal conductivity of the gas at the mean temperature and pressure of interest

B = an empirical constant.

This shows the insulation conductivity can be reduced by replacing the air in the insulation with a low-conductivity heavy gas such as CO_2, argon, krypton, or high molecular weight hydrocarbons. The thermal conductivity of the insulation can be increased if the air in the insulation is replaced by a high-conductivity gas such as helium or hydrogen. The thermal conductivity ratio of krypton to air is 0.21 while helium to air is 5.7 at STP (273 K and 1 atmosphere).

Trapping a fuel such as hydrogen or a reactant such as oxygen in the insulation can be very dangerous and precautions must be taken to purge or vent the insulation space as a safety measure. In a cryogenic system, there is also the possibility of condensing liquid air in the insulation. The liquid air is enriched in oxygen and will be hazardous if the insulation is combustible.

Equation 4.6.2 is valid as long as the mean free path of the gas is much shorter than the average spacing between solid particles in fibrous or powder insulation or the pore size in a cellular

insulation. In such cases, the gas molecules collide with each other more often than with the insulation surfaces and the gas behaves as if it were at the mean temperature of the insulation. As the gas pressure is reduced, the mean free path of the gas increases and the gas thermal conductivity decreases. At low pressure, k_g is proportional to the gas pressure:

$$k_g = C \cdot k'_g \cdot L \cdot P \tag{4.6.3}$$

where C = a constant which is dependent on the gas at a reference pressure and temperature

L = average distance a gas molecule moves before striking a surface in the porous insulation; this is smaller than the mean free path of the gas.

P = the absolute gas pressure

k'_g = the gas conductivity at the mean insulation temperature and the gas reference pressure, usually taken as 1 atmosphere.

As the gas pressure is lowered, the gas conductivity at some point becomes negligible compared to the radiation and solid conduction components of heat transport. If the insulation interparticle spacing is made smaller than the mean free path of the gas at atmospheric pressure, then the effective gas conductivity in the insulation will be less than the conductivity of the gas at the mean temperature and pressure of the insulation. This is the mechanism by which low-density aerogel insulation or fine particles can attain thermal conductivities less than air at ambient pressure and the mean insulation temperature. More detailed discussions can be found in Refs. [7–9].

Radiation Heat Transfer

The basics of thermal radiation are discussed in Chapter 3 along with the radiation mechanism for insulation containing multiple low-emittance radiation barriers. In powders, cellular or fibrous insulation systems the radiation resistance is provided by reflection, scattering, and absorption.[10,11] The details depend on the optical properties of the material, the size of the particle, and the wavelength of the radiation, which is related to the temperature of the material. In some circumstances, the radiation conductivity for a particulate insulation can be expressed as

$$k_r = \sigma \cdot T^3 / D \tag{4.6.4}$$

where T = absolute temperature (K)

σ = Stefan–Boltzmann constant 5.67×10^{-8} W/m^2·K^4

D = insulation extinction coefficient per unit thickness.

Multilayer insulations consist of multiple low-emissivity metal foils separated by low-conductivity solid spacers. The radiation conductivity for multilayer insulation can be expressed as

$$k_r = \sigma \cdot \varepsilon \cdot T^3 / N \tag{4.6.5}$$

where ε is the emittance of the radiation barrier material, and N is the number of radiation barriers per unit thickness. This shows that in multifoil systems, the radiation heat transfer can be significantly reduced by increasing the number of barrier layers. The emittance for metals at ambient temperature that are used as radiation barriers includes aluminum ($\varepsilon = 0.03$), gold ($\varepsilon = 0.01$–0.02), molybdenum ($\varepsilon = 0.1$), and tantalum ($\varepsilon = 0.05$).

Combined Thermal Conductivity

The combined point thermal conductivity of an insulation can be expressed by

$$k = k_s + k_g + k_r \tag{4.6.6}$$

or

$$k = A \cdot k_{\mathrm{m}} + B \cdot k'_{\mathrm{g}} + \sigma \cdot T^3 / D \tag{4.6.7}$$

The average apparent conductivity of the insulation between two temperatures becomes

$$k_{\mathrm{a}} = A \cdot k_{\mathrm{m}} + B \cdot k'_{\mathrm{g}} + \sigma \cdot \left(T_2^4 - T_1^4\right) / \left(4 \cdot D \cdot \left(T_2 - T_1\right)\right) \tag{4.6.8}$$

Application of this equation to specific types of insulation systems is discussed in the following sections. Most manufacturers report thermal conductivity as the average between two temperatures, typically ambient and a higher or lower temperature. Equation 4.6.9 is the relationship between the point thermal conductivity, $k(t)$, and the average, ka:

$$k_{\mathrm{a}} = \left(\frac{1}{T_2 - T_1}\right) \cdot \int_{T_1}^{T_2} k(t)\mathrm{d}t \tag{4.6.9}$$

Insulation Systems (Nonvacuum and Vacuum)

Powder Systems

Powder insulation consists of particulates such as perlite, vermiculite, beaded Styrofoam, microcellular material, silica, and some types of aerogels. The advantage of this type of insulation is that when poured or blown into the space to be insulated it will flow around obstructions such as piping. Particulates can also be removed by gravity flow, fluidized transfer, or evacuation. This is an advantage in complex process equipment that must be accessible for repair or maintenance. The disadvantage of powder insulation is that a container has to be placed around the region that is to be insulated in order to keep the powder insulation in place and in below ambient temperature systems to eliminate water vapor migration into the powder where it can condense or freeze. If condensation or freezing occurs, the thermal resistance of the insulation will be degraded.

The thermal performance of the insulation is limited by the conduction of the gas filling the pores of the powder. The solid conductivity component is dependent on the multiple point contacts between particles and by selecting a material with low intrinsic solid conductivity. In some powders, the particles are porous rather than solid. Examples are perlite, vermiculite, or agglomerates of smaller particles, such as in fumed silica. Radiation in powder insulation can be reduced through the selection of the proper particle size or the addition of opacifiers such as carbon black or reflective metal flakes of aluminum or copper.[12,13]

Some of the most common powder insulation material properties and thermal conductivities between various temperature ranges are given in Table 4.6.1. Thermal conductivity curves for perlite, vermiculite, and aerogel at various densities are shown in Figure 4.6.2. As shown in Figure 4.6.2, the conductivities of gas-filled perlite and vermiculite powders at the same temperature do not vary greatly from one another since both are dominated by the gas in the insulation pores. The aerogel powder insulation has a pore size approaching the mean free path of the gas, resulting in greatly reduced thermal conductivity.[14,15]

To reduce the conductivity, the gas can be removed from the insulation. This will reduce the apparent thermal conductivity of the insulation by an order of magnitude or more. The level of vacuum which has to be achieved to reduce the gas conductivity to a negligible value is of practical importance. Therefore, the thermal conductivity of various powder insulations is presented as a function of the gas pressure. The evacuated thermal conductivity curves for perlite, aerogel, and microcel with 10% aluminum flakes are shown in Figure 4.6.3. The differences in the powder and the effect of particle size can be seen in these figures.

Handling considerations of powder insulation deserve special mention due to the small particle size of many of the insulations, which can result in airborne dust or respirable hazard while loading

TABLE 4.6.1
Powder Insulation

Materials	Maximum Operating Temperature (K)	Density (kg/m³)	Hot Side Temp (K)	Cold Side Temp (K)	Apparent Thermal Conductivity (W/m·K)
Perlite	1255	112	327	294	0.0418
Vermiculite	1033	144	327	294	0.0677
Silica aerogel	977	80	327	294	0.0303
Silica aerogel with 15 wt% carbon	977	116	327	294	0.0135
Cork	366	208	327	294	0.049
Diatomaceous silica	1144	240	327	294	0.0577

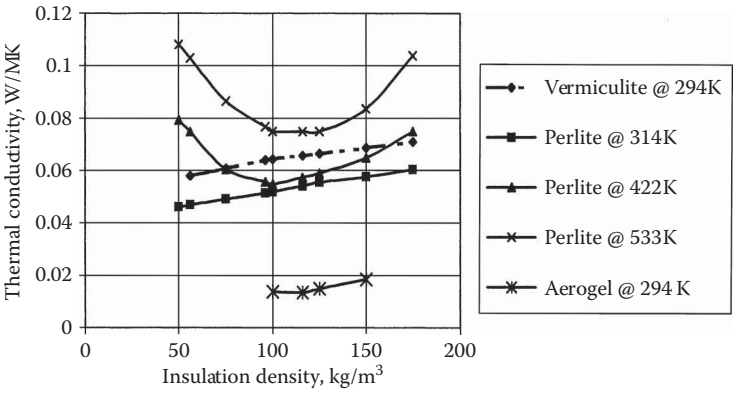


FIGURE 4.6.2 Powder insulation thermal conductivity.

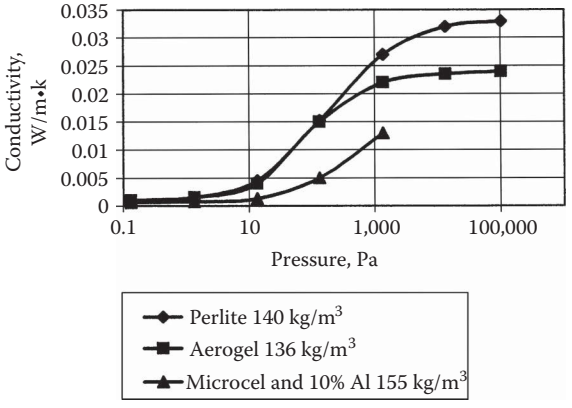


FIGURE 4.6.3 Evacuated powder thermal conductivity from 77 to 300 K.

the insulation powder. If the insulation powder is combustible, the airborne dust can pose an explosion hazard.

Nonvacuum perlite powder insulation is used in cryogenic process cold boxes. Evacuated perlite powder insulation is used in mobile carriers for cryogenic liquid oxygen and nitrogen. Evacuated carbon powder insulation is used in some metal thermos bottles. Powders opacified with metal

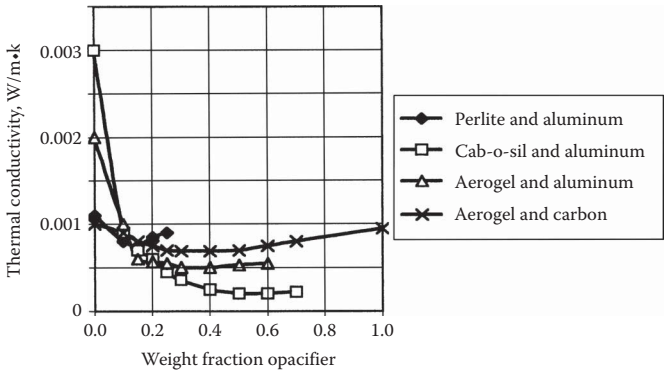


FIGURE 4.6.4 Evacuated opacified powder insulation thermal conductivity.

flakes to reduce radiation heat transfer were used in early small portable liquid oxygen and nitrogen containers, but were later replaced with higher performing evacuated multifoil insulation. The effect of the addition of opacifiers to perlite, Cab-O-Sil (fumed silica), and aerogel is shown in Figure 4.6.4.

Evacuated Multifoil Insulation

Evacuated multifoil insulation has the lowest thermal conductivity of any insulation. This is achieved by removing the gas from the insulation space, using very low solid-conductivity lightweight fiberglass or plastic spacers to separate the radiation shields, and reducing the radiation heat transfer with many low-emittance metal foils or metallized films. This type of insulation has been used successfully from liquid helium temperatures to 1400 K with the correct selection of supporting materials.[16–19]

The selection of material is critical for the insulation to achieve a low-thermal conductivity and to maintain its performance over the life of the system. Problems that can occur are out-gassing or off-gassing of the materials, which can degrade the vacuum or condense on the cooler radiation barrier surfaces and increase the foil emittance. Reactions between the radiation barriers and the spacer material can also increase the solid conduction and reduce the foil emittance. In high-temperature systems, the insulation assembly can be altered to use reduced thermal conductivity materials as the operating temperature decreases. Multifoil insulation systems, which have been used successfully in the past, are listed in Table 4.6.2 along with the operating temperature range.

Commercial Fibrous and Solid Insulation Systems for Moderate and High Temperatures

Introduction

This subsection will deal with systems operating at temperatures above 300 K. Processes operating at moderate or high temperatures are insulated to conserve energy, reduce risk of injury, and control temperature. A variety of products are available for a given temperature range and the thermal properties of a given type of insulation at a given temperature can vary significantly among vendors. It is important, therefore, to verify the temperature limits and the measured thermal conductivity for the specific product under consideration.

Health and Safety

Occupational Safety and Hazard Administration guidelines limit the operating temperatures of accessible surfaces. In some regions, state or local codes may also apply. Safety considerations for systems operating above ambient temperature are often taken to be more crucial than energy savings. The skin damage resulting from hot surface contact depends on a combination of factors such as surface temperature, contact time, and properties of the surface material such as thermal conductivity, specific heat, and density.[20,21] A suggested design for personnel protection is to limit accessible surface temperatures to about 52°C for highly conductive (metal) surfaces and 65°C for nonconductive surfaces.[22]

TABLE 4.6.2
Multifoil Vacuum Insulation

Materials							
Radiation Barrier	Spacer	Maximum Operating Density (K)	Layer Density (layers/cm)	Density (kg/m³)	Hot Side Temp (K)	Cold Side Temp (K)	Apparent Thermal Conductivity (W/m·K)
Aluminum foil	Glass paper	750	24	87	294	77	0.000036
Aluminum foil	Glass mat	750	4	19	294	77	0.00017
Aluminized Mylar	None	370	55	50	294	77	0.000042
Copper foil	Quartz paper	1000	31	339	1000	294	0.00095
Nickel foil	Quartz cloth	1200	42	384	1200	294	0.0021
Molybdenum foil	Zirconia cloth	1600	50	459	1500	294	0.0026

Note: Vacuum level = 0.1 μm.

In some cases, insulating materials contain binders or additives which may volatilize near the upper temperature limit of their range of operation. This upper limit temperature may be approached for the insulation immediately adjacent to a hot surface. The out-gassing is usually limited to the first few hours of heating but must be taken into consideration, especially for insulations used in enclosed areas. The manufacturer's literature should always be examined for warnings about initial off-gassing.

Heat Loss Calculations

Process piping and flat furnace insulations are the most common types of insulation to be considered. Virtually all exposed surfaces operating at temperatures above about 60°C are candidates for thermal insulation. Economic and safety considerations become more important as the temperature increases.

The calculated heat loss from high-temperature surfaces requires consideration of the entire system, inside material to outside air. In many process flow situations, steam lines, for example, the inside heat transfer coefficient is large so the thermal resistance attributable to the inside film coefficient can be neglected. Similarly, the thermal resistance of metal walls and pipes is generally neglected in comparison with the thermal resistance of even a minimal thickness of insulation. The thermal conductivity or thermal resistance of insulation materials to be applied to an outside surface, the external heat transfer coefficient for conduction–convection, and the loss to the surroundings due to radiation are of primary importance.

Discussions of the calculation of heat loss from pipes or flat surfaces at temperatures above ambient are discussed elsewhere in this handbook and in texts on heat transfer. A "standard practice" for calculating heat gain or heat loss complete with supporting computer code has been developed by the American Society for Testing and Materials (ASTM).[23] This Standard Practice published by the ASTM forms the basis for a program entitled "3E Plus, Insulation Thickness Computer Program" and is obtainable from the North American Insulation Manufacturers Association.[24]

The temperature change across moderate or high-temperature insulation can be very large. Since the thermal conductivity of insulations used in such systems increases with temperature, it is necessary to use either mean values for the insulation properties over the range of temperatures to be encountered or use temperature-dependent expressions in Fourier's law for calculating steady-state heat loss from hot systems.

Heat loss calculations require reliable thermal properties for the insulations that are used, exterior surface emittances, and knowledge of the exterior conditions such as air temperature and wind speed. The following discussion contains selected thermal data for a number of materials. Products that are selected for use should be validated by thermal test data obtained in accordance with ASTM C 177,[25] ASTM C 335,[26] ASTM C 518,[27] or C 1033.[28]

The heat transfer through thermal insulations used to reduce heat loss is often described by expressions like Equation 4.5.10 with T being on an absolute temperature scale:

$$k_a(T) = a + b \cdot T + c \cdot T^3 \tag{4.6.10}$$

where the T^3 term is associated with radiative transport, and the linear part of the equation is related to conduction. The radiative term should be written in terms of absolute temperature and the coefficients a, b, and c are positive. Equations 4.6.11–4.6.13 are also used to correlate thermal conductivity with temperature:

$$k_a(T) = a + b \cdot T + c \cdot T^2 \tag{4.6.11}$$

$$k_a(T) = \exp(a + b \cdot T) \tag{4.6.12}$$

$$k_a(T) = a + b \cdot T \tag{4.6.13}$$

Care must be exercised in interpreting laboratory k_a data, especially when large temperature differences are anticipated.[29] Measured k_a are commonly reported as the average of the cold and hot side temperatures. This value is then taken to be a point value for the apparent thermal conductivity. The measurement, however, was actually for the mean value of k_a on the temperature interval of the test. In the case of large measurement intervals, data are best described by an integral method.[30]

Thermal Insulations

A variety of thermal insulating materials are available for use on surfaces operating above 300 K. The maximum allowable temperature for an insulation type is the first consideration in selecting an insulation for a particular application. Table 4.6.3 contains a list of commonly used insulations, an

TABLE 4.6.3
Commonly Used Industrial Insulations

Number and Type of Insulation		Maximum Temperature (K)	ASTM Standard Specification[a]
1	Mineral fiber pipe (Type 1)	727	C 547
2	Mineral fiber pipe (Type 2)	923	C 547
3	Mineral fiber pipe (Type 3)	923	C 547
4	Mineral fiber board (Type 1A, 1B)	505	C 612
5	Mineral fiber board (Type 2)	727	C 612
6	Mineral fiber board (Type 3)	811	C 612
7	Mineral fiber board (Type 4)	922	C 612
8	Mineral fiber board (Type 5)	1255	C 612
9	Calcium silicate (Type 1)	922	C 533
10	Calcium silicate (Type 2)	1200	C 533
11	Perlite block and pipe	922	C 610
12	Cellular glass	700	C 552
13	Polystyrene	348	C 578
14	Polyisocyanurate board	366	C 1289
15	Loose-fill perlite	1033	C 549
16	Loose-fill vermiculite	1033	C 516

[a] These standard specifications can be found in Vol. 04.06 of the *2015 Annual Book of ASTM* Standards from the American Society for Testing and Materials, West Conshohocken, PA.

TABLE 4.6.4
Thermal Conductivity Data Taken from ASTM Standard Specifications

	Coefficients for Equation 4.6.10			k_a (W/m·K)			
Insulation[a]	***a*[b]**	***b***	***c***	**300 K**	**500 K**	**700 K**	**900 K**
1	–0.1342 E–1	0.1405 E–3	0.1666 E–9	0.033	0.078	0.142	–
2	0.1763 E–1	0.3736 E–4	0.2579 E–9	0.036	0.069	0.132	0.239
3	0.1763 E–1	0.3736 E–4	0.2579 E–9	0.036	0.069	0.132	0.239
4 (I-A)	0.1131 E–1	0.4574 E–4	0.4704 E–9	0.038	0.093	–	–
4 (I-B)	–0.2721 E–2	0.1182 E–3	0.1711 E–9	0.038	0.078	–	–
5	0.1224 E–1	0.4380 E–4	0.4296 E–9	0.037	0.088	0.190	–
6	0.2223 E–1	0.4020 E–5	0.5123 E–9	0.037	0.088	0.201	–
7	0.9717 E–2	0.6254 E–4	0.2300 E–9	0.035	0.070	0.132	0.234
8	–0.2303 E–1	0.1888 E–3	0.7477 E–10	0.036	0.081	0.135	0.201
9	0.1910 E–1	0.1230 E–3	0.1488 E–10	0.056	0.082	0.110	0.141
10	0.4241 E–1	0.1015 E–3	–0.2550 E–10	0.072	0.090	0.105	0.115
11	0.4082 E–1	0.9171 E–4	0.9363 E–10	0.072	0.098	0.137	0.192
12	0.1793 E–1	0.7351 E–4	0.3796 E–9	0.050	0.102	0.200	–
13	18 kg/m^3			0.038	–	–	–
	26 kg/m^3			0.029	–	–	–
	48 kg/m^3			0.029	–	–	–
14	Type 1 Class 1	Foil faced		0.022	–	–	–
15	Expanded	105 kg/m^3		0.053[31]	–	–	–
16	Expanded	122 kg/m^3		0.068[31]	–	–	–

Note: Maximum k_a values.

[a] Refer to Table 4.5.1.

[b] E–1 means $\times 10^{-1}$.

indication of the maximum application temperature, and a reference to a standard specification. The table contains materials used to insulate pipes, furnaces, and process vessels.

The standard specifications for insulations 1–12 in Table 4.6.3 include sufficient data to determine coefficients for Equation 4.6.10. The coefficients are listed in Table 4.6.4 along with apparent thermal conductivities at selected temperatures.

Table 4.6.5 contains k_a values for calcium silicate and perlite pipe insulation. These data are representative of commercially available products. Table 4.6.6 contains k_a values for several fibrous-type insulations that are representative of commercially available products. Since the data are from various manufacturers, it is not possible to interpolate in density.

Insulation Application

Thermal Insulation Products and Installation

The thermal data contained in the previous paragraphs are representative of the various types of insulation that are available for pipes and flat surfaces. The technical data for a specific product should be consulted before design calculations are completed. The performance of thermal insulations of all types requires installation in accordance with manufacturer's instructions and regular maintenance to remove and replace damaged sections. In general, the thermal resistance of porous or cellular insulation decreases if water is sorbed into the insulation. Thermal insulation is covered with mastics or water impermeable sheathing to prevent water uptake.

TABLE 4.6.5
Thermal Conductivity Data for Calcium Silicate and Perlite Pipe Insulations

Insulation Type (Density kg/m³)	Coefficients for Equation 4.6.10			k_a (W/m·K)		
	a	*b*	*c*	300 K	500 K	700 K
			Calcium Silicate			
(below 130)	–0.5770 E–2[a]	0.1356 E–3	0.6982 E–10	0.0368	0.0708	0.113
(130–190)	0.4841 E–1	–0.3478 E–4	0.2031 E–9	0.0435	0.0564	0.0937
(190–260)	0.3998 E–1	0.3480 E–4	0.1165 E–9	0.0536	0.0719	0.104
(260–400)	0.6243 E–1	–0.1467 E–4	0.1838 E–9	0.0630	0.0781	0.115
(400–560)	0.5573 E–1	0.7523 E–4	0.2692 E–10	0.0790	0.0967	0.118
			Perlite			
(below 160)	0.9902 E–2	0.1543 E–3	0.1233 E–10	0.0565	0.0886	0.122
(above 160)	0.8481 E–1	–0.6834 E–4	0.2992 E–9	0.0724	0.0880	0.140

[a] E–2 means $\times 10^{-2}$.

Vacuum Maintenance

In order to achieve the design thermal conductivity of a vacuum-insulated system over the life of the system, the vacuum level of the insulation must be maintained at the design value. This is achieved by checking the joints in the enclosure to eliminate air leaks into the vacuum space, calculating the gas permeation through the wall of the enclosure, off-gassing the insulation for a sufficient period of time before sealing the system, and through the addition of active chemical getters and adsorbents.

In cryogenic liquid oxygen, nitrogen, and argon systems, oxygen, nitrogen, and water can be adsorbed using molecular sieves placed on the cold wall of the system. Hydrogen, which will outgas from the metal walls of the enclosure and from the metal radiation foils, will not be adsorbed in significant quantities until the cold side attains a temperature that is below liquid hydrogen. Therefore, an active metal getter has to be placed in the vacuum space. These materials combine chemically

TABLE 4.6.6
Thermal Conductivity Data for Fibrous Pipe and Board Insulation

Insulation Type (density)	Coefficients for Equation 4.6.10			k_a (W/m·K)		
	a	*b*	*c*	300 K	400 K	500 K
			Mineral Fiber Board			
(64 kg/m³)	–0.1074 E–1[a]	0.1279 E–3	0.2621 E–9	0.0347	0.0572	0.0860
(96 kg/m³)	0.2173 E–1	0.1827 E–4	0.2796 E–9	0.0348	0.0469	0.0658
(160 kg/m³)	0.1185 E–1	0.6506 E–4	0.1375 E–9	0.0351	0.0467	0.0616
			Mineral Fiber Pipe			
(128 kg/m³)	0.5427 E–2	0.7864 E–4	0.1472 E–9	0.0330	0.0463	0.0631
Fiberglass pipe						
(64 kg/m³)	0.7111 E–2	0.5306 E–4	0.2805 E–9	0.0306	0.0463	0.0687
(224 kg/m³)	0.4276 E–1	0.3082 E–4	0.1446 E–9	0.0559	0.0643	0.0762
			Ceramic Fiber Pipe			
(218 kg/m³)	0.8725 E–1	–0.1180 E–3	0.3667 E–9	0.0618	0.0635	0.0741

[a] E–1 means $\times 10^{-1}$.

TABLE 4.6.7
Vacuum Maintenance Materials

Material	Temperature Range (K) Low	High	Gettered Gas	Capacity (torr·cm³/g)	Activation Temperature (K)
Palladium oxide	273	600	H_2	158,000	NA
SAES St 707 Zr 70%, V 24.6%, Fe 5.4%	273	973	H_2 at 298 K	1,100	720–620
			CO at 298 K	40	
			N_2 at 298 K at 750 K activation	9	
SAES St 172 St 707 + zirconium powder	273	973	H_2 at 298 K	420	673–1173
			CO at 298 K at 750 K activation	14	
SAES flash barium getter $BaAl_4$ + Nickel			H_2 at 298 K	110,000	
			CO at 298 K	14,000	
			N_2 at 298 K	26,000	
			O_2 at 298 K	50,000	

with the reactive gas to remove it from the vacuum space. SAES Getters Inc. produces commercial getters for this purpose.

In ambient and elevated temperature systems, adsorption will not work. The only methods available for vacuum maintenance are chemical getters or very long evacuations combined with high-temperature out-gassing of the insulation, or the use of insulation with very small pore sizes that can tolerate high vacuum levels. Vacuum maintenance materials and temperature levels are given in Table 4.6.7. References 32 and 33 are sources of additional information.

ACKNOWLEDGMENT

Insulation product data were generously provided by the Pabco division of the Fibreboard Corporation, the Rock Wool Manufacturing Co., and the North American Insulation Manufacturers Association.

REFERENCES

1. Kagan, M.G., *Thermal Insulation in Cryogenic Engineering*, IPST Press, Jerusalem, 1969.
2. Flynn, T.M., *Cryogenic Engineering*, Marcel Dekker, New York, 1997.
3. Barron, R., *Cryogenic Systems*, McGraw-Hill, New York, 1966
4. Strong, H.M., Bundy, F.P., and Bovenkerk, H.P., Flat panel vacuum thermal insulation, *J. Appl. Phys.*, 31(1), 1960.
5. Nayak, A.L. and Tien, C.L., Lattice-vacancy analysis for packed-sphere conductance, *Prog. Astronaut. Aeronaut.*, 56, 113–125, 1977.
6. Hammond, M.B., An analytical model for determining the thermal conductivity of closed-cell foam insulation, *Adv. Cryog. Eng.*, 15, 332–341, 1969.
7. Dushman, S., *Scientific Foundations of Vacuum Technique*, John Wiley & Sons, New York, 1966.
8. Kistler, S.S., The relation between heat conductivity and structure in silica aerogel, *J. Phys. Chem.*, 39, 79–85, 1935.
9. Yarbrough, D.W., Literature review: Heat transfer through two-phase insulation systems consisting of powders in a continuous gas phase, Oak Ridge National Laboratory, Report ORNL/M-2426, December 1992, Oak Ridge, TN.
10. Larkin, B.K. and Churchill, S.W., Heat transfer by radiation through porous insulation, *AIChE J.*, 5(4), 467–474, 1959.

11. Verschoor, J.D. and Greebler, P., Heat transfer by gas conduction and radiation in fibrous insulation, *Trans. ASME*, 74(6), 961–967, 1952.
12. Kropshot, R.H. and Burgess, R.W., Perlite for cryogenic insulation, *Adv. Cryog. Eng.*, 8, 425–436, 1963.
13. Hunter, B.J. and Kropshot, R.H., Metal powder additives in evacuated powder insulation, *Adv. Cryog. Eng.*, 5, 146–156, 1960.
14. Smith, D.M., Aerogel-based thermal insulation, *J. Non-Crystl. Solids*, 225, 254–259, 1998.
15. Deissler, R.G. and Boegly, Y.S., An investigation of effective thermal conductivities of powders in various gases, *Trans. ASME*, 80(7), 1417, 1958.
16. Matsh, L.C., Advances in multilayer insulations, *Adv. Cryog. Eng.*, 7, 413–418, 1962.
17. Kropshot, R.H., Multiple-layer insulations for cryogenic application, *Cryogenics*, 1(3), 171–177, 1961.
18. NCR-2 Sales Literature, King-Seely Thermos Co., Winchester, MA.
19. Grunert, W.E. et al., Research and development in a thermal insulation study, Final Report ALO 3632-43 September 1967 to September 1970, U.S. Atomic Energy Commission, AT(29-2)-2832, Washington, D.C., 1970.
20. ASTM C 1055, Standard guide for heated system surface conditions that produce contact burn injuries, *2015 Annual Book of ASTM Standards*, Vol. 04.06, American Society for Testing and Materials, West Conshohocken, PA, 2015, 551–558.
21. ASTM C 1057, Standard practice for determination of skin contact temperature from heated surfaces using a mathematical model and thermesthesiometer, *2015 Annual Book of ASTM Standards*, Vol. 04.06, American Society for Testing and Materials, West Conshohocken, PA, 2015, 559–564.
22. Anon., Industrial insulation for systems operating above ambient temperature, ORNL/M-4678, Oak Ridge National Laboratory, Oak Ridge, TN, 1995.
23. ASTM C 680, Standard practice for determination of heat gain or loss and the surface temperatures of insulated pipe and equipment systems by the use of a computer program, *2015 Annual Book of ASTM Standards*, Vol. 04.06, American Society for Testing and Materials, West Conshohocken, PA, 2015, 311–339.
24. 3E Plus, Insulation Thickness Computer Program, North American Insulation Manufacturers Association. http://3e-plus.software.informer.com/4.0/.
25. ASTM C 177, Standard test method for steady-state heat flux measurements and thermal transmission properties by means of the guarded-hot-plate apparatus, *2015 Annual Book of ASTM Standards*, Vol. 04.06, American Society for Testing and Materials, West Conshohocken, PA, 2015, 23–45.
26. ASTM C 335/C335M, Standard test method for steady-state heat transfer properties of horizontal pipe insulation, *2015 Annual Book of ASTM Standards*, Vol. 04.06, American Society for Testing and Materials, West Conshohocken, PA, 2015, 79–92.
27. ASTM C 518, Standard test method for steady-state heat flux measurements and thermal transmission properties by means of the heat flow meter apparatus, *2015 Annual Book of ASTM Standards*, Vol. 04.06, American Society for Testing and Materials, West Conshohocken, PA, 2015, 156–171.
28. ASTM C 1033, Standard test method for steady-state heat transfer properties of pipe insulation installed vertically, *1998 Annual Book of ASTM Standards*, Vol. 04.06, American Society for Testing and Materials, West Conshohocken, PA, 1998, 505–512 (withdrawn 2003).
29. Langlais, C., Thermal gradients effect on thermal property measurements, *J. Therm. Insul.*, 11, 189–195, 1998.
30. Hust, J.G. and Langford, A.B., Comments of the measurements of thermal conductivity and presentation of a thermal conductivity integral method, *Int. J. Thermophys.*, 3(1), 67–77, 1982.
31. Rosenow, W.M., Harnett, J.P., and Cho, Y.I., *Handbook of Heat Transfer*, 3rd ed., McGraw-Hill, New York, 1998.
32. SAES Getters, St 707 Non-evaporable getters that activate at low temperatures, SAES Getters U.S., www.saesgetters.com.
33. Boffito, C. et al., Gettering in cryogenic applications, *J. Vac. Sci. Technol., A*, 5(6), 1987.

4.7 ENERGY AUDIT FOR BUILDINGS

Moncef Krarti

Introduction

Since the oil embargo of 1973, significant improvements have been made in the efficiency of new buildings. However, the vast majority of the existing stock of buildings are more than a decade old and do not meet current energy-efficient construction standards (EIA, 1991). Therefore, energy retrofits of existing buildings will be required for decades to come if the overall energy efficiency of the building stock is to meet the standards.

Investing to improve the energy efficiency of buildings provides an immediate and relatively predictable positive cash flow resulting from lower energy bills. In addition to the conventional financing options available to owners and building operators (such as loans and leases), other methods are available to finance energy retrofits for buildings. One of these methods that is becoming increasingly common is performance contracting, in which payment for a retrofit project is contingent upon its successful outcome. Typically, an energy service company (ESCO) assumes all the risks for a retrofit project by performing the engineering analysis and obtaining the initial capital to purchase and install equipment needed for energy efficiency improvements. Energy auditing is an important step used by ESCOs to ensure the success of their performance-contracting projects.

Currently, several large industrial and commercial buildings have established internal energy management programs, based on energy audits, to reduce waste in energy use or to comply with the specifications of some regulations and standards. Other building owners and operators take advantage of available financial incentives typically offered by utilities or state agencies to perform energy audits and implement energy conservation measures (ECMs).

In the 1970s, building energy retrofits consisted of simple measures such as shutting off lights, turning down heating temperatures, turning up air-conditioning temperatures, and reducing the hot water temperatures. Today, building energy management infers complete control of the energy systems and consumption within a facility. Therefore, the energy auditor should be aware of key energy issues such as the subtleties of electrical utility rate structures and of the latest building energy efficiency technologies and their applications.

This section describes a general but systematic procedure for energy auditing that is suitable for both commercial buildings and industrial facilities. Some of the commonly recommended ECMs are briefly discussed. A case study for an office building is presented to illustrate the various tasks involved in an energy audit. Finally, an overview is provided to outline the existing methods for measurement and verification of energy savings incurred by the implementation of ECMs.

Types of Energy Audits

The term "energy audit" is widely used and may have different meanings depending on the ESCOs. Energy auditing of buildings can range from a short walk-through of the facility to a detailed analysis with hourly computer simulation. Generally, four types of energy audits can be distinguished as briefly described as follows.

Walk-Through Audit

This audit consists of a short on-site visit of the facility to identify areas where simple and inexpensive actions can provide immediate energy and/or operating cost savings. Some engineers refer to these types of actions as operating and maintenance (O&M) measures. Examples of O&M measures include lowering heating set point temperatures, replacing broken windows, insulating exposed hot water or steam pipes, and adjusting boiler fuel–air ratio.

Utility Cost Analysis

The main purpose of this type of audit is to carefully analyze the operating costs of the facility. Typically, the utility data over several years are evaluated to identify the patterns of energy use, peak demand, weather effects, and the potential for energy savings. To perform this analysis, it is recommended that the energy auditor conducts a walk-through survey to get acquainted with the facility and its energy systems.

It is important that the energy auditor understands clearly the utility rate structure that applies to the facility for several reasons, including the following:

- Checking the utility charges and ensuring that no mistakes were made in calculating the monthly bills. Indeed, the utility rate structures for commercial and industrial facilities can be quite complex, with ratchet charges and power factor penalties.
- Determining the most dominant charges in the utility bills. For instance, peak demand charges can be a significant portion of the utility bill, especially when ratchet rates are applied. Peak usage shaving measures can then be recommended to reduce these demand charges.
- Identifying if the facility can benefit from using other utility rate structures or purchasing cheaper fuel to reduce its operating costs. This analysis can provide a significant reduction in the utility bills, especially with the electrical industry deregulation and the advent of real time pricing (RTP) rate structures.

Moreover, the energy auditor can determine whether or not the facility is prime for energy retrofit projects by analyzing the utility data. Indeed, the energy use of the facility can be normalized and compared to indices (for instance, the energy use per unit of floor area for commercial buildings—or per unit of a product for industrial facilities).

Standard Energy Audit

The standard audit provides a comprehensive analysis of the energy systems of the facility. In addition to the activities described for the walk-through audit and the utility cost analysis described earlier, the standard energy audit includes the development of a baseline for the energy use of the facility and the evaluation of the energy savings and the cost-effectiveness of appropriately selected ECMs. The step-by-step approach of the standard energy audit is similar to that of the detailed energy audit, which is described in the following subsection.

Typically, simplified tools are used in the standard energy audit to develop baseline energy models and to predict the energy savings of ECMs. Among these tools are the degree-day methods and linear regression models (Fels, 1988). In addition, a simple payback analysis is generally performed to determine the cost-effectiveness of ECMs.

Detailed Energy Audit

This is the most comprehensive but also time-consuming energy audit type. Specifically, the detailed energy audit includes the use of instruments to measure energy use for the whole building and/or for some energy systems within the building (for instance, by end uses: lighting, office equipment, fans, chiller, etc.). In addition, sophisticated computer simulation programs are typically employed for detailed energy audits to evaluate and recommend energy retrofits for the facility.

The techniques available to perform measurements for an energy audit are diverse. During an on-site visit, handheld and clamp-on instruments can be used to determine the variance of some building parameters such as the indoor air temperature, the luminance level, and the electrical energy use. When long-term measurements are needed, sensors are typically used and are connected to a data acquisition system so measured data can be stored and be remotely accessible. Recently, nonintrusive load monitoring (NILM) techniques have been proposed (Shaw et al., 1998). The NILM technique can determine the real-time energy use of the significant electrical loads in

a facility by using only a single set of sensors at the facility service entrance. The minimal effort associated with using the NILM technique compared to the traditional multimetering approach (which requires a separate set of sensors to monitor energy consumption for each end use) makes the NILM a very attractive and inexpensive load-monitoring technique for ESCOs and facility owners.

The computer simulation programs used in the detailed energy audit typically provide the energy use distribution by load type (i.e., energy use for lighting, fans, chillers, boilers, etc.). They are often based on dynamic thermal performance of the building energy systems and usually require a high level of engineering expertise and training. These simulation programs range from those based on the bin method (Knebel, 1983) to those that provide hourly building thermal and electrical loads such as DOE-2 (LBL, 1980).

In the detailed energy audit, more rigorous economical evaluation of the ECMs is generally performed. Specifically, the cost-effectiveness of energy retrofits may be determined based on the life-cycle cost (LCC) analysis rather than the simple payback period analysis. LCC analysis takes into account a number of economic parameters such as interest, inflation, and tax rates.

General Procedure for a Detailed Energy Audit

To perform an energy audit, several tasks are typically carried out depending on the type of the audit and the size and function of the building. Some of the tasks may have to be repeated, reduced in scope, or even eliminated based on the findings of other tasks. Therefore, the execution of an energy audit is often not a linear process and is rather iterative. However, a general procedure can be outlined for most buildings.

Step 1: Building and Utility Data Analysis

The main purpose of this step is to evaluate the characteristics of the energy systems and the patterns of energy use for the building. The building characteristics can be collected from the architectural/mechanical/electrical drawings and/or from discussions with building operators. The energy use patterns can be obtained from a compilation of utility bills over several years. Analysis of the historical variation of the utility bills allows the energy auditor to determine any seasonal and weather effects on the building energy usage. Some of the tasks that can be performed in this step are presented as follows, with the key goals expected from each task noted in italics:

- Collect at least 3 years of records of utility data [*to identify a historical energy use pattern*].
- Identify the fuel types used (electricity, natural gas, oil, etc.) [*to determine the fuel type that accounts for the largest energy use*].
- Determine the patterns of fuel use by fuel type [to identify the peak demand for energy use by fuel type].
- Understand utility rate structure (energy and demand rates) [to evaluate if the building is penalized for peak demand and if cheaper fuel can be purchased].
- Analyze the effect of weather on fuel consumption.
- Perform utility energy use analysis by building type and size (building signature can be determined including energy use per unit area [*to compare against typical indices*]).

Step 2: Walk-Through Survey

This step should identify potential energy savings measures. The results of this step are important since they determine if the building warrants any further energy auditing work. Some of the tasks involved in this step are as follows:

- Identify the customer's concerns and needs.
- Check the current O&M procedures.

- Determine the existing operating conditions of major energy use equipment (lighting, heating, ventilating, and air conditioning (HVAC) systems, motors, etc.).
- Estimate the occupancy, equipment, and lighting (energy use density and hours of operation).

Step 3: Baseline for Building Energy Use

The main purpose of this step is to develop a base-case model that represents the existing energy use and operating conditions for the building. This model will be used as a reference to estimate the energy savings due to appropriately selected ECMs. The major tasks to be performed during this step are as follows:

- Obtain and review architectural, mechanical, electrical, and control drawings.
- Inspect, test, and evaluate building equipment for efficiency, performance, and reliability.
- Obtain all occupancy and operating schedules for equipment (including lighting and HVAC systems).
- Develop a baseline model for building energy use.
- Calibrate the baseline model using the utility data and/or metered data.

Step 4: Evaluation of Energy-Saving Measures

In this step, a list of cost-effective ECMs is determined using both energy savings and economic analysis. To achieve this goal, the following tasks are recommended:

- Prepare a comprehensive list of ECMs (using the information collected in the walk-through survey).
- Determine the energy savings due to the various ECMs pertinent to the building by using the baseline energy use simulation model developed in Step 3.
- Estimate the initial costs required to implement the ECMs.
- Evaluate the cost-effectiveness of each ECM using an economical analysis method (simple payback or LCC analysis).

Tables 4.7.1 and 4.7.2 provide summaries of the energy audit procedure recommended, respectively, for commercial buildings and for industrial facilities. Energy audits for thermal and electrical systems are separated since they are typically subject to different utility rates.

Common Energy Conservation Measures

In this subsection, some ECMs commonly recommended for commercial and industrial facilities are briefly discussed. It should be noted that the list of ECMs presented later does not pretend to be exhaustive nor comprehensive. It is provided merely to indicate some of the options that the energy auditor can consider when performing an energy analysis of a commercial or an industrial facility. However, it is strongly advised that the energy auditor keeps abreast of any new technologies that can improve the facility energy efficiency. Moreover, the energy auditor should recommend the ECMs only after he performs an economical analysis for each ECM.

Building Envelope

For some buildings, the envelope (i.e., walls, roofs, floors, windows, and doors) can have an important impact on the energy used to condition the facility. The energy auditor should determine the actual characteristics of the building envelope. During the survey, a sheet for the building envelope should be established to include information such as materials of construction (for instance, the level of insulation in walls, floors, and roofs) and the area and number of various assemblies of

TABLE 4.7.1
Energy Audit Summary for Residential and Commercial Buildings

Thermal Systems	Electric Systems
Utility Analysis	
Thermal energy use profile (building signature)	Electrical energy use profile (building signature)
Thermal energy use per unit area (or per student for schools or per bed for hospitals)	Electrical energy use per unit area (or per student for schools or per bed for hospitals)
Thermal energy use distribution (heating, domestic hot water (DHW), process, etc.)	Electrical energy use distribution (cooling, lighting, equipment, fans, etc.)
Fuel types used	Weather effect on electrical energy use
Weather effect on thermal energy use	Utility rate structure (energy charges, demand charges, power factor penalty, etc.)
Utility rate structure	
On-Site Survey	
Construction materials (thermal resistance type and thickness)	HVAC system type
HVAC system type	Lighting type and density
DHW system	Equipment type and density
Hot water/steam use for heating	Energy use for heating
Hot water/steam for cooling	Energy use for cooling
Hot water/steam for DHW	Energy use for lighting
Hot water/steam for specific applications (hospitals, swimming pools, etc.)	Energy use for equipment
	Energy use for air handling
	Energy use for water distribution
Energy Use Baseline	
Review architectural, mechanical, and control drawings	Review architectural, mechanical, electrical, and control drawings
Develop a base-case model (using any baselining method ranging from very simple to more detailed tools)	Develop a base-case model (using any baselining method ranging from very simple to more detailed tools)
Calibrate the base-case model (using utility data or metered data)	Calibrate the base-case model (using utility data or metered data)
Energy Conservation Measures	
Heat recovery system (heat exchangers)	Energy-efficient lighting
Efficient heating system (boilers)	Energy-efficient equipment (computers)
Temperature setback	Energy-efficient motors
EMCS	HVAC system retrofit
HVAC system retrofit	EMCS
DHW use reduction	Temperature setup
Cogeneration	Energy-efficient cooling system (chiller)
	Peak demand shaving
	TES system
	Cogeneration
	Power factor improvement
	Reduction of harmonics

the envelope (for instance, the type and the number of panes for the windows should be noted). In addition, comments on the repair needs and recent replacement should be noted during the survey.

Some of the commonly recommended ECMs to improve the thermal performance of building envelope are as follows:

1. *Addition of Thermal Insulation*. For building surfaces without any thermal insulation, this measure can be cost-effective.

TABLE 4.7.2
Energy Audit Summary for Industrial Facilities

Thermal Systems	Electric Systems
Utility Analysis	
Thermal energy use profile (building signature)	Electrical energy use profile (building signature)
Thermal energy use per unit of a product	Electrical energy use per unit of a product
Thermal energy use distribution (heating, process, etc.)	Electrical energy use distribution (cooling, lighting, equipment, process, etc.)
Fuel types used	
Analysis of the thermal energy input for specific processes used in the production line (such as drying)	Analysis of the electrical energy input for specific processes used in the production line (such as drying)
Utility rate structure	Utility rate structure (energy charges, demand charges, power factor penalty, etc.)
On-Site Survey	
List of equipment that use thermal energy	List of equipment that use electrical energy
Perform heat balance of the thermal energy	Perform heat balance of the electrical energy
Monitor thermal energy use of all or part of the equipment	Monitor electrical energy use of all or part of the equipment
Determine the by-products of thermal energy use (such as emissions and solid waste)	Determine the by-products of electrical energy use (such pollutants)
Energy Use Baseline	
Review mechanical drawings and production flow charts	Review electrical drawings and production flow charts
Develop a base-case model (using any baselining method)	Develop a base-case model (using any baselining method)
Calibrate the base-case model (using utility data or metered data)	Calibrate the base-case model (using utility data or metered data)
Energy Conservation Measures	
Heat recovery system	Energy-efficient motors
Efficient heating and drying system	Variable speed drives
EMCS	Air compressors
HVAC system retrofit	Energy-efficient lighting
Hot water and steam use reduction	HVAC system retrofit
Cogeneration (possibly with solid waste from the production line)	EMCS
	Cogeneration (possibly with solid waste from the production line)
	Peak demand shaving
	Power factor improvement
	Reduction of harmonics

2. *Replacement of Windows.* When windows represent a significant portion of the exposed building surfaces, using more energy-efficient windows (high-thermal resistance frame and glazing, low-emissivity glazing, airtight, etc.) can be beneficial in both reducing the energy use and improving the indoor comfort level.
3. *Reduction of Air Leakage.* When the infiltration load is significant, leakage areas of the building envelope can be reduced by simple and inexpensive weather stripping techniques.

The energy audit of the envelope is especially important for residential buildings. Indeed, the energy use from residential buildings is dominated by weather since heat gain and/or loss from direct conduction of heat or from air infiltration/exfiltration through building surfaces accounts for

a major portion (50%–80%) of the energy consumption. For commercial buildings, improvements to the building envelope are often not cost-effective due to the fact that modifications to the building envelope (replacing windows, adding thermal insulation in walls) typically are very expensive. However, it is recommended to systematically audit the envelope components not only to determine the potential for energy savings but also to ensure the integrity of its overall condition. For instance, thermal bridges, if present, can lead to heat transfer increase and to moisture condensation. The moisture condensation is often more damaging and costly than the increase in heat transfer since it can affect the structural integrity of the building envelope.

Electrical Systems

For most commercial buildings and a large number of industrial facilities, electrical energy cost constitutes the dominant part of the utility bill. Lighting, office equipment, and motors are the electrical systems that consume the major part of energy usage in commercial and industrial buildings.

1. *Lighting.* Lighting for a typical office building represents, on average, 40% of the total electrical energy use. There are a variety of simple and inexpensive measures to improve the efficiency of lighting systems. These measures include the use of energy-efficient lighting lamps and ballasts, the addition of reflective devices, delamping (when the luminance levels are above the recommended levels by the standards), and the use of daylighting controls. Most lighting measures are especially cost-effective for office buildings for which payback periods are less than 1 year.
2. *Office Equipment.* Office equipment constitutes the fastest growing part of the electrical loads, especially in commercial buildings. Office equipment includes computers, fax machines, printers, and copiers. Today, there are several manufacturers that provide energy-efficient office equipment (such as those that comply with U.S. EPA Energy Star specifications). For instance, energy-efficient computers automatically switch to a low-power "sleep" mode or off mode when not in use.
3. *Motors.* The energy cost to operate electric motors is a significant part of the operating budget of any commercial and industrial building. Measures to reduce the energy cost of using motors include reducing operating time (turning off unnecessary equipment), optimizing motor systems, using controls to match motor output with demand, using variable speed drives for air and water distribution, and installing energy-efficient motors. Table 4.7.3 provides typical efficiencies for several motor sizes.

In addition to the reduction in the total facility electrical energy use, retrofits of the electrical systems decrease the cooling loads and, therefore, further reduce the electrical energy use in the building. These cooling energy reductions, as well as possible increases in thermal energy use (for space heating), should be accounted for when evaluating the cost-effectiveness of improvements in lighting and office equipment.

The integration of daylighting controls in office buildings is used as an illustrative example of proven energy efficiency measures used to reduce lighting energy consumption. Indeed, several studies have indicated that daylighting can offer a cost-effective alternative to electrical lighting for commercial and institutional buildings. Through sensors and controllers, daylighting can reduce and even eliminate the use of electrical lighting required to provide sufficient illuminance levels inside office spaces. A simplified calculation method has been developed by Krarti et al. (2005) to estimate the reduction in the total lighting energy use due to daylighting with dimming controls for office buildings. The method has been shown to apply for office buildings in the United States as well as in other countries (Al-Mohimen et al., 2005). The simplified calculation method can be used as a predesign tool to assess the potential of daylighting in saving electricity use associated with artificial lighting for office buildings.

TABLE 4.7.3
Typical Efficiencies of Motors

Motor Size (HP)	Standard Efficiency (%)	High Efficiency (%)
1	72	81
2	76	84
3	77	89
5	80	89
7.5	82	89
10	84	89
15	86	90
20	87	90
30	88	91
40	89	92
50	90	93

To determine the percent savings, f_d, in annual use of artificial lighting due to implementing daylighting using daylighting controls in office buildings, Krarti et al. (2005) found that the following equation can be used:

$$f_d = b\left[1 - \exp\left(-a\tau_w A_w / A_p\right)\right]\frac{A_p}{A_f}, \tag{4.7.1}$$

where

- A_w/A_p: window to perimeter floor area. This parameter provides a good indicator of the window size relative to the daylit floor area.
- A_p/A_f: perimeter to total floor area. This parameter indicates the extent of the daylit area relative to the total building floor area. Thus, when $A_p/A_f = 1$, the whole building can benefit from daylighting.
- a and b: coefficients that depend only on the building location and are given by Table 4.7.4 for various sites throughout the world.

HVAC Systems

The energy use due to HVAC systems can represent 40% of the total energy consumed by a typical commercial building. The energy auditor should obtain the characteristics of major HVAC equipment to determine the condition of the equipment, its operating schedule, its quality of maintenance, and its control procedures. A large number of measures can be considered to improve the energy performance of both primary and secondary HVAC systems. Some of these measures are listed as follows:

1. *Setting up/back thermostat temperatures.* When appropriate, setback of heating temperatures can be recommended during unoccupied periods. Similarly, setup of cooling temperatures can be considered.
2. *Retrofit of constant air volume (CAV) systems.* For commercial buildings, variable air volume (VAV) systems should be considered when the existing HVAC systems rely on constant-volume fans to condition part or the entire building.
3. *Installation of heat recovery systems.* Heat can be recovered from some HVAC equipment. For instance, heat exchangers can be installed to recover heat from air handling unit (AHU) exhaust air streams and from boiler stacks.

TABLE 4.7.4
Coefficients *a* and *b* of Equation 4.7.1 for Various Locations throughout the World

Location	*a*	*B*	Location	*a*	*b*
Atlanta	19.63	74.34	Casper	19.24	72.66
Chicago	18.39	71.66	Portland	17.79	70.93
Denver	19.36	72.86	Montreal	18.79	69.83
Phoenix	22.31	74.75	Quebec	19.07	70.61
New York City	18.73	66.96	Vancouver	16.93	68.69
Washington DC	18.69	70.75	Regina	20.00	70.54
Boston	18.69	67.14	Toronto	19.30	70.48
Miami	25.13	74.82	Winnipeg	19.56	70.85
San Francisco	20.58	73.95	Shanghai	19.40	67.29
Seattle	16.60	69.23	Kuala Lumpur	20.15	72.37
Los Angels	21.96	74.15	Singapore	23.27	73.68
Madison	18.79	70.03	Cairo	26.98	74.23
Houston	21.64	74.68	Alexandria	36.88	74.74
Fort Worth	19.70	72.91	Tunis	25.17	74.08
Bangor	17.86	70.73	Sao Paulo	29.36	71.19
Dodge City	18.77	72.62	Mexico City	28.62	73.63
Nashville	20.02	70.35	Melbourne	19.96	67.72
Oklahoma City	20.20	74.43	Roma	16.03	72.44
Columbus	18.60	72.28	Frankfurt	15.22	69.69
Bismarck	17.91	71.50	Kuwait	21.98	65.31
Minneapolis	18.16	71.98	Riyadh	21.17	72.69
Omaha	18.94	72.30			

4. *Retrofit of central heating plants*. The efficiency of a boiler can be drastically improved by adjusting the fuel–air ratio for proper combustion. In addition, installation of new energy-efficient boilers can be economically justified when old boilers are to be replaced.
5. *Retrofit of central cooling plants:* Currently, there are several chillers that are energy efficient and easy to control and operate and are suitable for retrofit projects.

It should be noted that there is a strong interaction between various components of a heating and cooling system. Therefore, a whole-system analysis approach should be followed when retrofitting a building HVAC system. Optimizing the energy use of a central cooling plant (which may include chillers, pumps, and cooling towers) is one example of using a whole-system approach to reduce the energy use for heating and cooling buildings.

Compressed Air Systems

Compressed air has become an indispensable tool for most manufacturing facilities. Its uses range from air-powered hand tools and actuators to sophisticated pneumatic robotics. Unfortunately, staggering amounts of compressed air are wasted in a large number of facilities. It is estimated that only about 20%–25% of input electrical energy is delivered as useful compressed air energy. Leaks are reported to account for 10%–50% of the waste while misapplication accounts for 5%–40% of the loss of compressed air (Howe and Scales, 1998).

To improve the efficiency of compressed air systems, the auditor can consider several issues including whether compressed air is the right tool for the job (for instance, electric motors are more energy efficient than air-driven rotary devices), how the compressed air is applied (for instance, lower pressures can be used to supply pneumatic tools), how it is delivered and controlled (for

instance, the compressed air needs to be turned off when the process is not running), and how the compressed air system is managed (for each machine or process, the cost of compressed air needs to be known to identify energy and cost savings opportunities).

Energy Management Controls

Because of the steady decrease in the cost of computer technology, automated control of a wide range of energy systems within commercial and industrial buildings is becoming increasingly popular and cost effective. An energy management and control system (EMCS) can be designed to control and reduce the building energy consumption within a facility by continuously monitoring the energy use of various equipment and making appropriate adjustments. For instance, an EMCS can automatically monitor and adjust indoor ambient temperatures, set fan speeds, open and close AHU dampers, and control lighting systems.

If an EMCS is already installed in the building, it is important to recommend a system tune-up to ensure that the controls are operating properly. For instance, the sensors should be calibrated regularly in accordance with manufacturers' specifications. Poorly calibrated sensors may cause an increase in heating and cooling loads and may reduce occupant comfort.

As an example of the impact of improved controls on energy use, the performance of demand ventilation control options for car garages are discussed. Automobile parking garages can be partially open or fully enclosed. Partially open garages are typically above grade with open sides and do not need generally mechanical ventilation. However, fully enclosed parking garages are usually underground and require mechanical ventilation. Indeed, in absence of ventilation, enclosed parking facilities present several indoor air quality problems. The most serious is the emission of high levels of carbon monoxide (CO) by cars within the parking garages. Other concerns related to enclosed garages are the presence of oil and gasoline fumes, and other contaminants such as oxides of nitrogen (NO_x) and smoke haze from diesel engines.

To determine the adequate ventilation rate for garages, two factors are typically considered: the number of cars in operation and the emission quantities. The number of cars in operation depends on the type of the facility served by the parking garage and may vary from 3% (in shopping areas) up to 20% (in sports stadium) of the total vehicle capacity (ASHRAE, 2015). The emission of carbon monoxide depends on individual cars including such factors as the age of the car, the engine power, and the level of car maintenance.

For enclosed parking facilities, ASHRAE standard 62-1989 specifies fixed ventilation rate of below 7.62 L/s m^2 (1.5 cfm/ft^2) of gross floor area (ASHRAE, 1989). Therefore, a ventilation flow of about 11.25 air changes per hour is required for garages with 2.5-m ceiling height. However, some of the model code authorities specify an air change rate of 4–6 air changes per hour. Some of the model code authorities allow ventilation rate to vary and be reduced to save fan energy if CO demand-controlled ventilation is implemented, that is, a continuous monitoring of CO concentrations is conducted, with the monitoring system being interlocked with the mechanical exhaust equipment. The acceptable level of contaminant concentrations varies significantly from code to code. A consensus on acceptable contaminant levels for enclosed parking garages is needed. Unfortunately, ASHRAE standard 62-1989 does not address the issue of ventilation control through contaminant monitoring for enclosed garages. Thus, ASHRAE commissioned a research project 945-RP (Krarti et al. 1999) to evaluate current ventilation standards and recommend rates appropriate to current vehicle emissions/usage. Based on this project, a general methodology has been developed to determine the ventilation requirements for parking garages.

Figure 4.7.1 indicates also the fan energy savings achieved by the On–Off and VAV systems (relative to the fan energy use by the CV system). As illustrated in Figure 4.7.1, significant fan energy savings can be obtained when demand CO ventilation control strategy is used to operate the ventilation system while maintaining acceptable CO levels within the enclosed parking facility. These energy savings depend on the pattern of car movement within the parking facility. Figure 4.7.2 indicates three types of car movement profiles considered in the analysis considered by Krarti et al. (1999).

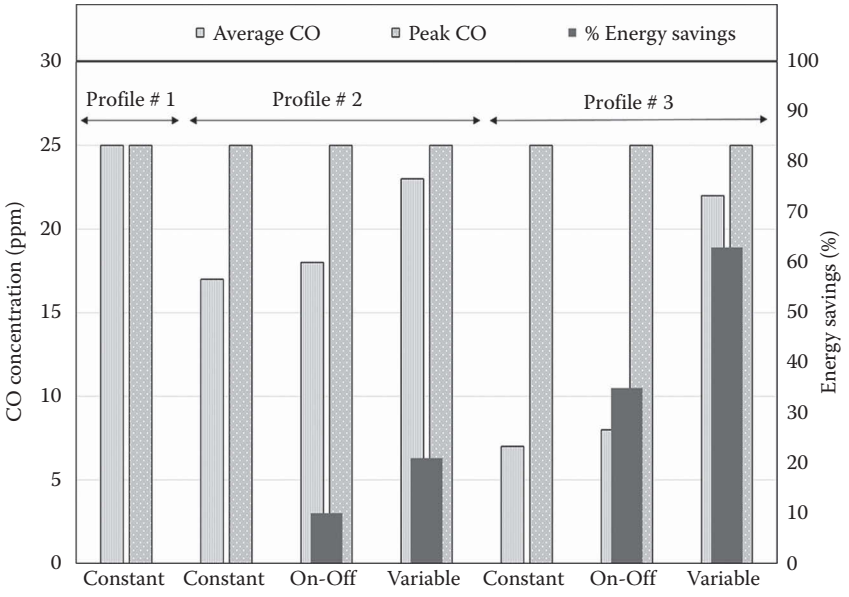


FIGURE 4.7.1 Typical energy savings and maximum CO level obtained for demand CO ventilation controls.

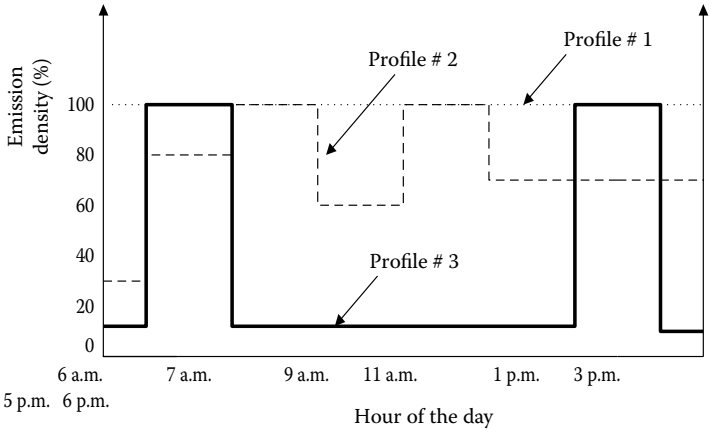


FIGURE 4.7.2 Car movement profiles used in the analysis conducted by Krarti et al. (1999).

Indoor Water Management

Water and energy savings can be achieved in buildings by using water-saving equipment instead of the conventional fixtures for toilets, faucets, showerheads, dishwashers, and clothes washers. Savings can also be achieved by eliminating leaks in pipes and fixtures.

Table 4.7.5 provides the typical water usage of conventional and water-efficient fixtures. In addition, Table 4.7.5 indicates the hot water consumption by each fixture as a fraction of the total water usage. With water-efficient fixtures, a savings of 50% of water use can be achieved.

New Technologies

The energy auditor may consider the potential of implementing and integrating new technologies within the facility. It is, therefore, important that the energy auditor understands these new

TABLE 4.7.5
Characteristics of Water-Using Fixtures

End Use	Conventional Fixtures	Water-Efficient Fixtures	Usage Pattern	Hot Water Usage (%)
Toilets	3.5 gal/flush	1.6 gal/flush	4 flushes/person/day	0
Showers	5.0 gal/min	2.5 gal/min	5 min/shower	60
Faucets	4.0 gal/min	2.0 gal/min	2.5 min/person/day	50
Dishwashers	14.0 gal/load	8.5 gal/load	0.17 loads/person/day	100
Clothes washers	55.0 gal/load	42.0 gal/load	0.3 loads/person/day	25
Leaks	10% total use	2% total use	N/A	50

technologies and knows how to apply them. Among the new technologies that can be considered for commercial and industrial buildings include:

Building Envelope Technologies

Recently, several materials and systems have been proposed to improve the energy efficiency of the building envelope, especially windows, including the following:

- Spectrally selective glasses that can optimize solar gains and shading effects.
- Chromogenic glazing that change their properties automatically depending on temperature and/or light-level conditions (similar to sunglasses that become dark in sunlight).
- Building integrated photovoltaic panels that can generate electricity while absorbing solar radiation and reducing heat gain through the building envelope (typically roofs).

Light Pipe Technologies

While the use of daylighting is straightforward for perimeter zones that are near windows, it is not usually feasible for interior spaces, particularly those without skylights. Recent but still emerging technologies allow one to "pipe" light from roof or wall-mounted collectors to interior spaces that are not close to windows or skylights.

HVAC Systems and Controls

Several strategies can be considered for energy retrofits, including the following:

- Heat recovery technologies such as rotary heat wheels and heat pipes can recover 50%–80% of the energy used to heat or cool ventilation air supplied to the building.
- Desiccant-based cooling systems are now available and can be used in buildings with large dehumidification loads during long periods (such as hospitals, swimming pools, and supermarket fresh produce areas).
- Geothermal heat pumps can provide an opportunity to take advantage of the heat stored underground to condition building spaces.
- Thermal energy storage (TES) systems offer a mean of using less-expensive off-peak power to produce cooling or heating to condition the building during on-peak periods; several optimal control strategies have been developed in recent years to maximize the cost savings of using TES systems.

Cogeneration

This is not really a new technology. However, recent improvements in its combined thermal and electrical efficiency have made cogeneration cost-effective in several applications including institutional buildings such as hospitals and universities.

Case Study

To illustrate the energy audit process described previously, a case study is presented. The activities performed for each step of the energy audit are briefly described. For more details, refer to Kim et al. (1998). The building analyzed in this case study is a medium-size office building located in Seoul, Korea.

STEP 1: BUILDING AND UTILITY DATA ANALYSIS

The first step in the building energy audit process is to collect all available information about the energy systems and the energy use pattern of the building. This information was collected before the field survey, in particular, from the architectural, mechanical, and electrical drawings and utility bills.

Building Characteristics

The building is a 26-story office building with a 2-story penthouse and 4-story basement. It is located in downtown Seoul, Korea. The structure of the building consists of modular concrete and steel frame. The building area is 3920 m^2 and the site area is 6555 m^2. Single-glazed windows are installed throughout the building. Figure 4.7.3 shows a typical floor plan of the building. Table 4.7.6 describes the various construction materials used throughout the building.

Energy Use

Figure 4.7.4 summarizes the monthly electrical energy use of the building for 1993. The monthly average dry-bulb outdoor air temperature recorded for 1993 is also presented in Figure 4.7.4. It is clear that the electrical energy use increases during the summer months (June through October) when the outdoor temperatures are high. During the other months, the electrical energy use is almost constant and can be attributed mostly to lighting and office equipment loads. Preliminary analysis of the metered building energy indicates that natural gas consumption is inconsistent

TABLE 4.7.6
Building Construction Materials

Component	Materials
Exterior wall	5 cm tile
	16 cm concrete
	2.5 cm foam insulation
	0.6 cm finishing material
Roof	5 cm lightweight concrete
	15 cm concrete
	2.5 cm foam insulation
Interior wall	2 cm finishing cement mortar
	19 cm concrete block
Glazing	1.2 cm thick single-pane glazing
Underground wall	25 cm concrete
	Asphalt shingle
	Air space
	10 cm brick
	2 cm finishing cement mortar
Underground floor	15 cm concrete
	Asphalt shingle
	12 cm concrete
	2 cm finishing cement mortar

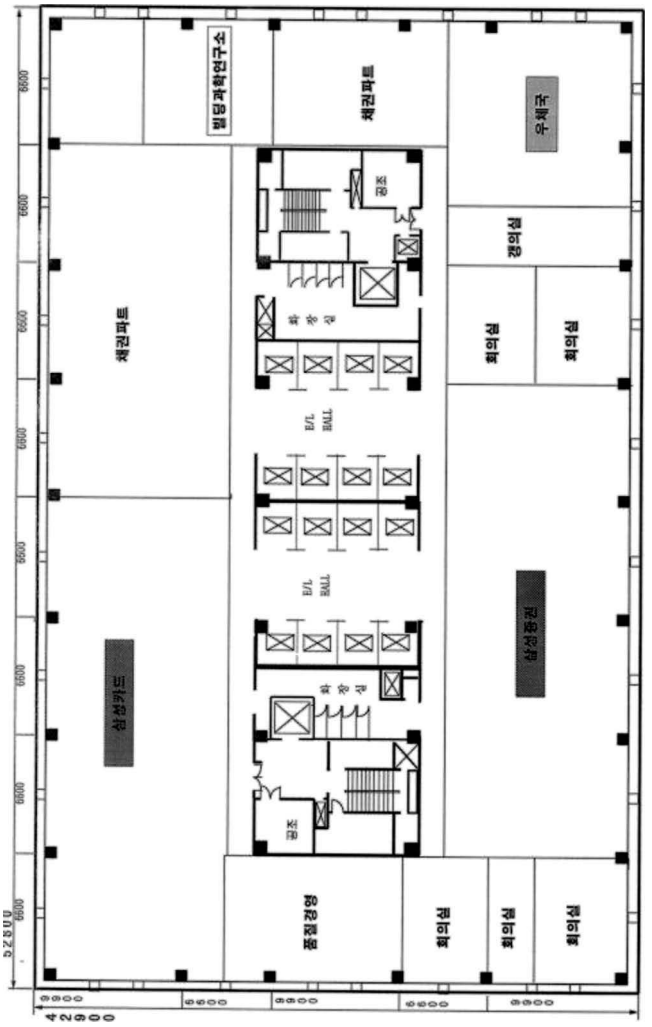


FIGURE 4.7.3 Typical floor plan of the audited office building.

from month to month. For example, gas consumption during January is six times higher than during December, even though the weather conditions are similar for both months. Therefore, the recorded data for natural gas were considered unreliable and only metered electrical energy data were analyzed.

STEP 2: ON-SITE SURVEY

A one-day field survey was conducted with the assistance of the building operator in the summer of 1996. During the survey, several useful and revealing information and engineering data were collected. For instance:

- It was found that the building had been retrofitted with energy-efficient lighting systems. The measurement of luminance levels throughout the working areas indicated adequate lighting. To determine an estimate of the energy use for lighting, the number and type of lighting fixtures were recorded.
- It was observed that the cooling and heating temperature set points were 25.5°C and 24.5°C, respectively. However, indoor air temperature and relative humidity measurements during the field survey revealed that during the afternoon the thermal conditions

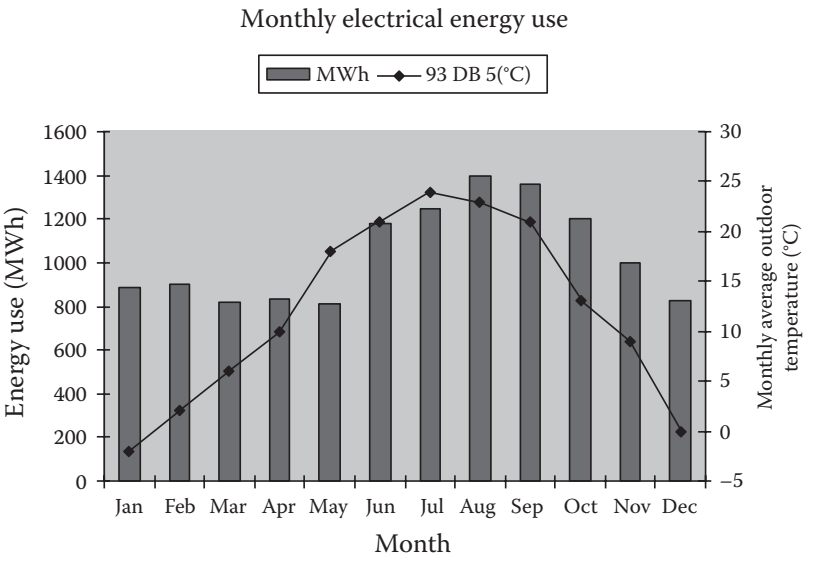


FIGURE 4.7.4 Monthly actual electrical energy consumption.

were uncomfortable in several office spaces, with average air dry-bulb temperature of 28°C and relative humidity of 65%. A discussion with the building operator indicated that the chillers were no longer able to meet the cooling loads following the addition of several computers in the building during the last few years. As a solution, an ice storage tank was then added to reduce the peak cooling load.

- It was discovered during the survey that the building is heated and cooled simultaneously by two systems: CAV and fan coil unit (FCU) systems. The CAV system is complemented by the FCU system as necessary. Two air-handling units serve the entire building, and about 58 FCUs are on each floor.
- The heating and cooling plant consists of three boilers, six chillers, three cooling towers, and one ice storage tank. The capacities of the boilers and the chillers are
  Boilers: 13 MBtu/h (2 units) and 3.5 MBtu/h (1 unit)
  Chillers: 215 tons (5 units) and 240 tons (1 unit)
  The TES system consists of a brine ice-on-coil tank. The hours of charging and discharging are 10 and 13, respectively. The TES system is currently controlled using simple and nonpredictive storage-priority controls.

The internal heat gain sources are shown in Table 4.7.7. Operating schedules were based on the discussion with the building operators and on observations during the field survey.

TABLE 4.7.7
Internal Heat Gain Level for the Office Building

Internal Heat Gain	Design Load
Occupancy	17 m^2/person
	Latent heat gain: 45 W
	Sensible heat gain: 70 W
Lighting	14 W/m^2
Equipment	16 W/m^2
Ventilation	14.7 CFM/person

STEP 3: ENERGY USE BASELINE MODEL

To model the building using DOE-2, each floor was divided into two perimeter and two core zones. Figure 4.7.5 shows the zone configuration used to model the building floors. The main reason for this zoning configuration is the lack of flexibility in the DOE-2 SYSTEMS (Energy Information Agency, 1980) program. Although the actual building is conditioned by the combination of CAV and FCU systems, the SYSTEM module of DOE-2.1E cannot model two different types of HVAC systems serving one zone. Therefore, a simplification has been made to simulate the actual HVAC system of the building. This simplification consists of the following: The perimeter zone is conditioned by the FCUs, while the core zone is conditioned by CAV. Since all the FCUs are located at the perimeter, this simplification is consistent with the actual HVAC system operation.

Figure 4.7.6 shows the monthly electrical energy consumption predicted by the DOE-2 base model and the actual energy use recorded in 1993 for the building. It shows that DOE-2 predicts the actual energy use pattern of the building fairly well, except for the months of September and October. The difference between the annual metered energy use in the building and the annual predicted electricity use by the DOE-2 base-case model is about 762 MWh. DOE-2 predicts that the building consumes 6% more electricity than the actual metered annual energy use. To perform the DOE-2 base-case model, a TRY-type weather file of Seoul was created using the

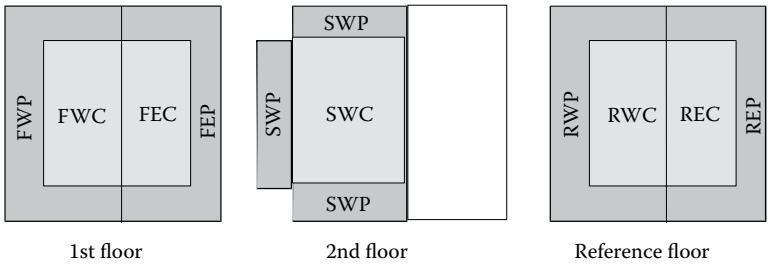


FIGURE 4.7.5 Building zoning configuration for DOE-2 computer simulation.

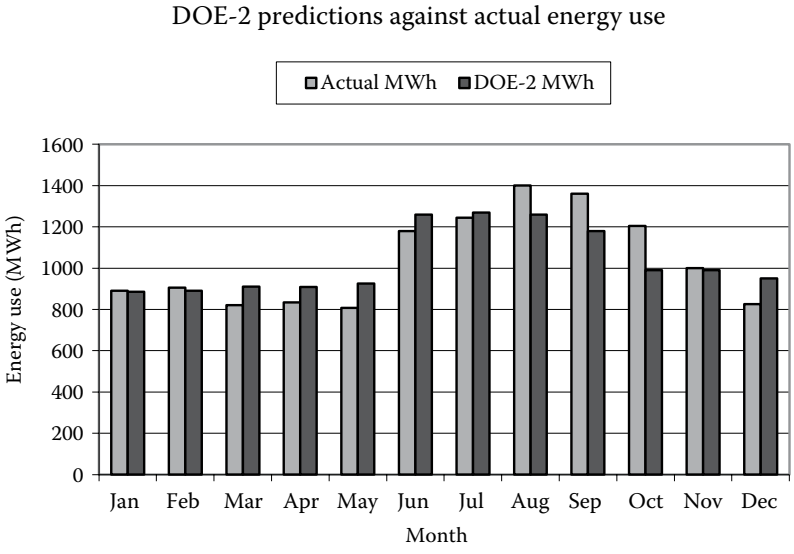


FIGURE 4.7.6 Comparison of the DOE-2 prediction and actual building electrical energy use.

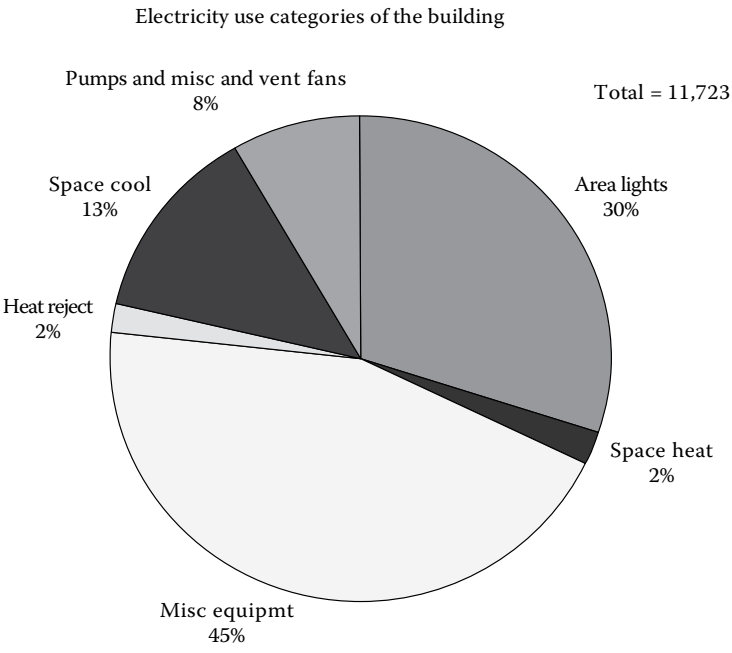


FIGURE 4.7.7 Electricity use distribution.

raw weather data collected for 1993. Using the DOE-2 base-case model, a number of energy conservation opportunities (ECOs) can now be evaluated.

Figure 4.7.7 shows the distribution by end uses of the building energy consumption. The electrical energy consumption of the building is dominated by lighting and equipment. The electricity utilized for lighting and office equipment represents about 75% of the total building electric consumption. As mentioned earlier, a recent lighting retrofit was installed in the building, comprising electronic ballasts and energy-efficient fluorescent fixtures. Therefore, it was decided not to consider a lighting retrofit as an ECO for this study. The electricity consumption for cooling is about 13.1%. The ECOs selected for this building aim mostly at reducing the cooling loads.

STEP 4: EVALUATION OF ENERGY CONSERVATION OPPORTUNITIES

Based on the evaluation of the energy use pattern of the building, several ECOs for the building were analyzed. Among the ECOs considered in the study, six of them successfully reduced energy consumption:

1. *ECO #1: CAV to VAV conversion*. The present AHU fans are all constant-speed fans. They supply conditioned air through a constant-volume air supply system to the selected zones. The system is designed to supply enough air to heat or cool the building under design conditions. Under nondesign conditions, more air than needed is supplied. Changing the system to a VAV system would reduce the amount of air supplied by the AHUs and result in less energy to condition the various zones. For this ECO, a constant-volume reheat fan system assigned to the core zones in the building was changed to a VAV system. In particular, VAV boxes, controlled by the space thermostat, vary the amount of conditioned air supply to the building zones to control the indoor temperature. Both labor and equipment costs were included in the estimation of the payback period for this measure.
2. *ECO #2: Optimal ice storage control*. The current TES system is operated using a nonpredictive storage-priority control. To improve the benefits of the TES system, a

near-optimal controller is suggested. This ECO is analyzed to determine if the cost of electrical energy consumption for the building is reduced when a near-optimal control strategy is used. To determine the savings of this option, the simulation environment developed by Henze et al. (1997) is used. This simulation environment is based on a dynamic programming technique and determines the best operating controls for the TES system given the cooling and noncooling load profiles and electrical rate structure. No DOE-2 simulation is performed for this ECO. The results of the dynamic programming simulation indicate that an energy reduction of 5% can be achieved using a near-optimal control in lieu of the storage-priority control. To implement this near-optimal control, a predictor is required to determine future building cooling/noncooling loads. An example of such a predictor could be based on neural networks (Kreider et al., 1995). The labor cost and the initial cost of adding some sensors and a computer were considered in determining the payback period for this measure.

3. *ECO #3: Glazing retrofit*. For this building, low-e glazing systems were considered to reduce the internal heat gain due to solar radiation. Thus, the cooling load is reduced. In addition, the increased U-value of the glazing reduces the heating load. For this ECO, the existing single-pane windows with the glass conductance of 6.17 W/m^2 K and the shading coefficient of 0.69 were changed to double-pane windows of 1.33 W/m^2 K and 0.15 shading coefficient.
4. *ECO #4: Indoor temperature setback/setup*. In this ECO, the impact of the indoor temperature setting on energy use was analyzed using the DOE-2 simulation program; the heating temperature was changed from 24.5°C to 22.5°C and the cooling temperature was changed from 25.5°C to 27.5°C. There is only a labor cost associated with this measure.
5. *ECO #5: Motor replacement*. Increasing the efficiency of the motors for fans and pumps can reduce the total electric energy consumption in the building. In this ECO, the existing efficiencies for the motors were assumed to range from 0.85 (10 HP) to 0.90 (50 HP). Energy-efficient motors have efficiencies that range from 0.91 (10 HP) to 0.95 (50 HP). Only the differential cost was considered in the economic analysis.
6. *ECO #6: Daylighting control plan*. A continuous dimming control would regulate the light level so that the luminance level inside the zones remains constant. The electricity consumption of the building can be significantly reduced, while the gas consumption can be slightly increased because of the reduced space heat gain from the lighting system. For this ECO, a daylighting system with continuous dimming control was considered for perimeter office zones. The cost of adding the sensors was considered.

TABLE 4.7.8
Economic Analysis of the ECOs

	Electricity Cost (MWon)	LNG Cost (MWon)	Total Cost (MWon)	Capital Cost (MWon)	Saving (%)	Savings (MWon)	Payback Period (Years)
Base Case	984.4	139.1	1123.5	—	—	—	—
ECO #1	940.8	49.8	990.5	465.5	11.8	133.0	3.5
ECO #2	979.1	139.1	1118.2	42.4	0.6	5.3	8.0
ECO #3	977.9	126.9	1104.8	280.5	1.7	18.7	15.0
ECO #4	983.7	106.4	1090.1	16.7	3.0	33.4	0.5
ECO #5	972.6	138.7	1111.4	60.5	1.1	12.1	5.0
ECO #6	911.6	144.8	1056.4	268.4	6.0	67.1	4.0

The impact of the selected ECOs on the electricity use in the building as predicted by DOE-2 simulations is shown in Table 4.7.8. Based on these results, converting CAV to VAV and implementing a daylighting control system with dimming control reduced the total electricity consumption of the building 5.2% and 7.3%, respectively. These savings are significant considering that the electricity consumption for the cooling plant alone is about 13.1% of the total electricity consumption of the building.

The economic analysis was performed using the utility rate of Seoul, Korea. Table 4.7.8 presents the energy cost savings of ECOs in Korean currency (1000 Won = $1). In addition to the electricity cost, the natural gas cost was also included in the economic analysis. The natural gas is used only for heating the building. The economic analysis shows that the VAV conversion reduces the total energy consumption by more than 10%, and the daylighting control saves about 6% of the total energy costs.

Recommendations

The results for economic analysis, the VAV conversion (ECO #1), adjustment of temperature set point (ECO #4), and the daylighting control (ECO #6) are the recommended energy-saving opportunities to be implemented for the audited office building. In particular, the VAV conversion reduces the total energy consumption more than 10%, and the daylighting control saves about 6% of the total energy consumption.

Verification Methods of Energy Savings

Energy conservation retrofits are deemed cost-effective based on predictions of the amount of energy and money a retrofit will save. However, several studies have found that large discrepancies exist between actual and predicted energy savings. Due to the significant increase in the activities of ESCOs, the need became evident for standardized methods for measurement and verification of energy savings. This interest has led to the development of the *North American Energy Measurement and Verification Protocol* published in 1996 and later expanded and revised under the *International Performance Measurement and Verification Protocol*.

In principle, the measurement of the retrofit energy savings can be obtained by simply comparing the energy use during pre- and post-retrofit periods. Unfortunately, the change in energy use between the pre- and post-retrofit periods is not only due to the retrofit itself but also to other factors such as changes in weather conditions, levels of occupancy, and HVAC operating procedures. It is important to account for all these changes to accurately determine the retrofit energy savings.

Several methods have been proposed to measure and verify savings of implemented ECMs in commercial and industrial buildings. Some of these techniques are briefly described as follows.

Regression Models

The early regression models used to measure savings adapted the variable-base degree-day method. Among these early regression models, the PRInceton Scorekeeping Method (PRISM) that uses measured monthly energy consumption data and daily average temperatures to calibrate a linear regression model and determine the best values for nonweather-dependent consumption, the temperature at which the energy consumption began to increase due to heating or cooling (the change point or base temperature), and the rate at which the energy consumption increased. Several studies have indicated that the simple linear regression model is suitable for estimating energy savings for residential buildings. However, subsequent work has shown that the PRISM model does not provide accurate estimates for energy savings for most commercial buildings (Ruch and Claridge, 1992).

Single-variable (temperature) regression models require the use of at least four-parameter segmented linear or change-point regressions to be suitable for commercial buildings.

Katipamula et al. (1994) proposed multiple linear regression models to include as independent variables internal gain, solar radiation, wind, and humidity ratio, in addition to the outdoor temperature. For the buildings considered in their analysis, Katipamula et al. found that wind and solar radiations have small effects on the energy consumption. They also found that internal gains have generally modest impact on energy consumption. Katipamula et al. (1998) discuss in more detail the advantages and the limitations of multivariate regression modeling.

Time Variant Models

Several techniques have been proposed to include the effect of time variation of several independent variables on estimating the energy savings due to retrofits of building energy systems. Among these techniques are the artificial neural networks (Krarti et al., 1998), Fourier series (Dhar et al., 1998), and NILM (Shaw et al., 1998). These techniques are very involved and require a high level of expertise and training.

Summary

An energy audit of commercial and industrial buildings encompasses a wide variety of tasks and requires expertise in a number of areas to determine the best ECMs suitable for an existing facility. This section provided a description of a general but systematic approach to perform energy audits. If followed carefully, the approach helps facilitate the process of analyzing a seemingly endless array of alternatives and complex interrelationships between the building and its energy system components.

REFERENCES

Al-Mohimen, M., Hanna, G., and Krarti, M., 2005. Analysis of daylighting benefits for office buildings in Egypt. *ASME Sol. Energy Eng.*, 127, 366–370.

ASHRAE, 1989. *Ventilation for Acceptable Indoor Air Quality*, *Standard 62-1989.* American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., Atlanta, GA.

ASHRAE, 2015. *Handbook of HVAC Applications*. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., Atlanta, GA.

Dhar, A., Reddy, T.A., Claridge, D.E., 1998. Modeling hourly energy use in commercial buildings with Fourier series functional forms. *ASME Sol. Energy Eng. J.*, 120(3), 217.

Energy Information Agency, 1980. DOE-2 Users Guides, Ver. 2.1, Report No. LBL-8689, Lawrence Livermore National Laboratory, Berkeley, CA.

Energy Information Agency, Energy facts, 1991. DOE/EIA-0469(91), Department of Energy, Washington, DC.

Fels, J., 1998. Special issue devoted to measuring energy savings: the scorekeeping approach. *Energy Build.*, 12(2), 113–127.

Henze, G.P., Krarti, M., and Brandemuehl, M.J., 1997. A simulation environment for the analysis of ice storage control. *Int. J. HVAC Res.*, 3, 128–148.

Howe, B. and Scales, B., 1998. Beyond leaks: demand-side strategies for improving compressed air efficiency. *Energy Eng.*, 95, 31–39.

Katipamula, S., Reddy, T.A., and Claridge, D.E., 1994. Development and application of regression models to predict cooling energy consumption in large commercial buildings. *1994 ASME-JSEME-JSES International Solar Energy Conference*, San Francisco, CA, March 27–30, 307–322.

Katipamula, S., Reddy, T.A., and Claridge, D.E., 1998. Multivariate regression modeling. *ASME J. Sol. Energy Eng.*, 120, 177–184.

Kim, K., Yoon, H., Lee, E., Choi, S., and Krarti, M., 1998. Building energy performance simulations to evaluate energy conservation measures. *ASME Solar Energy Engineering Conference Proceedings*, Albuquerque, NM, June 14–17, 45–55.

Knebel, D.E., 1983. *Simplified Energy Analysis Using the Modified Bin Method*. American Society of Heating, Refrigeration, and Air-Conditioning Engineers, Atlanta, GA.

Krarti, M., Kreider, J.F., Cohen, D., and Curtiss, P., 1998. Estimation of energy savings for building retrofits using neural networks. *ASME J. Sol. Energy Eng.*, 120, 211–216.

Krarti, M., Ayari, A. Grot, D., 1999. Ventilation requirements for enclosed vehicular parking garages, Final Report for ASHRAE RP-945, American Society of Heating, Refrigerating, and Air Conditioning Engineering, Atlanta, GA.

Krarti, M., Erickson, P., Hillman, T., 2005. A simplified method to estimate energy savings of artificial lighting use from daylighting. *Build. Environ.*, 40, 747–754.

Kreider, J.F., Claridge, D.Z., Curtis, P., Dodier, R., Haberl, J.S., and Krarti, M., 1995. Building energy use prediction and system identification using recurrent neural networks. *J. Sol. Energy Eng.*, 117, 161–166.

LBL, 1980. DOE-2 User Guide, Version 2.1, Report No. LBL-8689, Rev. 2, Lawrence Livermore National Laboratory, Berkeley, CA.

Ruch, D. and Claridge, D.E., 1992. A four-parameter change-point model for predicting energy consumption in commercial buildings. *ASME J. Sol. Energy Eng.*, 104, 177–184.

Shaw, S.R., Abler, C.B., Lepard, R.F., Luo, D., Leeb, S.B., and Norford, L.K., 1998. Instrumentation for high performance non-intrusive electrical load monitoring. *ASME J. Sol. Energy Eng.*, 120, 177–184.

4.8 ADVANCED ENERGY-EFFICIENT BUILDING ENVELOPE SYSTEMS

Moncef Krarti and John Zhai

Introduction

Building heating and cooling energy use accounts for up to 40% of the total energy consumption in developed countries (Pérez-Lombard et al., 2008). Heat losses and gains through the building envelope greatly contribute to thermal heating and cooling loads and subsequently to the overall energy performance of buildings (Tuhus-Dubrow and Krarti, 2010; Sadineni et al., 2011). Several advances in building technology have been made that continue to transform building energy performance and promote new and innovative construction techniques and challenges traditional practices. In fact, some of the existing systems that are often required by energy efficiency codes and standards may not provide the optimal energy performance and indoor environmental quality (IEQ). For instance, leaky buildings traditionally perform very poorly in terms of energy consumption, but in general, their IEQ—as a result of the incoming outside air—is fairly good. For the sake of energy efficiency, the trend has been tighter, more effectively sealed buildings, which in turn has led to more IEQ, mold, and sick building syndrome problems. As the push for improved energy performance points designers and builders toward tighter construction, the very principle that reduces the building's energy consumption—reduced infiltration—is a net loser for IEQ.

In this chapter, some advanced technologies of building envelope are presented to challenge the existing paradigms for high-performance buildings. Specifically, four building envelope systems are presented to enhance the energy efficiency and optimize the IEQ of buildings. These advanced building envelope systems include:

- *Dynamic Insulation Materials* (DIMs)*:* These materials are assemblies that can change their thermal resistance through, for instance, a controlled exchange of liquid or gaseous media, and are applied to the exterior walls and roofs of buildings.
- *Variable Reflectivity Coating roofs*: These systems are associated with coatings that can change their reflective properties over time depending on desired controlled strategies. It is well known that static cool roofs can save energy used during cooling periods but may increase energy consumption during swing or heating seasons due to lower heat gains. The variable reflectivity coatings allow the reduction in heating penalty associated with conventional cool roofs.
- *One-Layer Breathing Walls*: These systems provide air ventilation through the walls while recovering thermal energy. Essentially using heat exchanging mechanism, filtrated fresh air is introduced through the walls from outside to inside and is heated or cooled by the thermal energy wasted by conduction and convection with the walls.
- *Multilayer Living Walls:* These walls utilize basic biomimetic principles to adapt to the change in climatic conditions. In particular, these climate-adaptive intelligent walls have embedded systems of air, water, and phase change material (PCM) and can be controlled to maintain acceptable indoor environment under changing outdoor conditions.

In the following sections, an overview of the basic operating concepts of these advanced envelope systems is first provided. Then, their energy performance is summarized based on the reported results and based on either experimental or modeling analyses.

Dynamic Insulation Materials

Background

High thermal resistance materials and assemblies are effective in reducing heat transfer through the building envelope when there is a significant indoor–outdoor temperature difference. However,

building envelopes with high RSI value (*R*-value in SI unit) can be beneficial for one season of the year but are not necessarily beneficial for the other season. For example, in cold climates, high thermal insulation can actually increase building thermal loads—especially during mild periods when it may be beneficial to let heat that is trapped indoors to escape outdoors.

Some studies have theorized on DIMs and systems in which the thermal conductivities of wall assemblies are controllable and tunable, either due to applied electrical voltage to conducting polymers or other electrochromic materials (Jelle et al., 2010, Jelle, 2011; Kalnæs and Jelle, 2014). While some prototypes of electrochromic materials are available, these technologies have only been applied to smart windows rather than opaque building envelope systems. The operating principle of one of the DIM technologies is depicted in Figure 4.8.1. In lieu of traditional, static insulation assemblies (e.g., extruded polysterene (XPS), vacuum insulation panels (VIPs), gas insulation panels (GIPs)), a DIM is a rigid, cellular panel that is placed within the external wall cavities of a building. The DIM panel is designed with either an open cavity or an interconnected system of conduits. Once installed, the bulk thermal diffusivity of the assembly is tailored by the introduction of inert gases with variable conductivities (e.g., carbon dioxide, nitrogen, helium), thereby affecting the rate of heat transfer through the envelope. It is envisioned that the insulation media can be regulated via control strategies based on indoor–outdoor temperature differences and exchanged daily, monthly, or seasonally, depending on the climate and necessary heating and cooling demands. While many have speculated on the potential benefits of DIMs, no studies have yet quantified the potential energy savings of such materials.

Energy Performance

In this section, the potential energy savings of DIMs for buildings is briefly presented based on reported work by Park et al. (2015). Using a dual-control strategy, the DIM thermal resistance is changed from low to high values to accommodate the variations in outdoor conditions and reduce thermal loads for buildings. Total energy savings associated with DIMs is estimated and compared with benchmarks achievable by high-performance static insulation in representative US climates.

A comparative analysis is presented to evaluate the effect of DIMs on the annual energy use for a single-zone residential building located in three US climates: Golden, CO; Madison, WI; Phoenix, AZ. Specifically, four control scenarios of RSI-value of the exterior walls are considered:

- CASE 1: Building envelope with RSI-1.0
- CASE 2: Building envelope with RSI-2.5
- CASE 3: Two-step RSI-value control strategy (switch between RSI-1.0 and RSI-2.5)
- CASE 4: Two-step RSI-value control strategy (switch between RSI-0.5 and RSI-2.5)

The two-step control strategies are achieved by modulating the RSI values of the walls based on heating or cooling requirements and changes in outdoor temperatures. The main goal of the

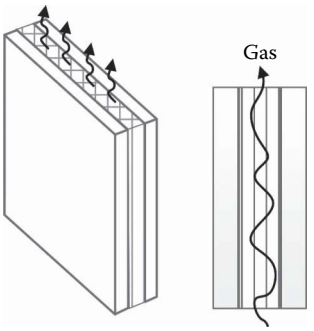


FIGURE 4.8.1 Operating principle of DIMs.

two-step RSI-value control strategy is to maximize the heat gains through the building envelope, and minimize the heat losses through the walls during the heating season. The same control strategy minimizes heat gains and maximizes heat losses during the cooling season.

The heating and cooling set point temperature for a single zone residential building is 19°C and 25°C for 24 h/day during winter and summer season, respectively. Table 4.8.1 and Figure 4.8.2 compare the annual heating and cooling energy end uses obtained for the four DIM control cases outlined above for a residential building located in three US locations. Solar heat gains transmitted through the window assist in heating the zone air, and thus reduce the heating thermal load for the building. However, the addition of heat gains from transmitted solar radiation typically increases building cooling thermal load in most climates. As shown in Figure 4.8.2, low RSI-value walls are more effective to release the heat trapped indoors during cooler nighttime periods than the high RSI-value walls. The two-step control strategies take advantage of the DIM low RSI-value during nighttime, and minimize heat gains through the exterior walls during daytime by increasing their DIM RSI value. Specifically, the two-step-RSI-1/RSI-2.5 control strategy reduces the cooling

TABLE 4.8.1
Heating and Cooling Energy Use Savings Compared to CASE 1 (RSI-1.0 Exterior Walls) When the Building Has Windows and Internal Gains in Three US Locations: Golden, CO; Madison, WI; and Phoenix, AZ

	Heating Energy Use Savings (GJ)			% Savings		
	RSI-2.5	Two-Step RSI-1/RSI-2.5	Two-Step RSI-0.5/RSI-2.5	RSI-2.5	Two-Step RSI-1/RSI-2.5	Two-Step RSI-0.5/RSI-2.5
Golden, CO	–5.55	–5.65	–5.77	–29.0	–29.5	–30.1
Madison, WI	–10.14	–10.14	–10.15	–30.6	–30.6	–30.6
Phoenix, AZ	n/a	n/a	n/a	n/a	n/a	n/a
	Cooling Energy Use Savings (GJ)			**% Savings**		
	RSI-2.5	Two-Step RSI-1/RSI-2.5	Two-Step RSI-0.5/RSI-2.5	RSI-2.5	Two-Step RSI-1/RSI-2.5	Two-Step RSI-0.5/RSI-2.5
Golden, CO	–0.24	–0.67	–1.05	–4.3	–11.8	–18.6
Madison, WI	0.40	–0.18	–0.69	8.5	–3.7	–14.5
Phoenix, AZ	–3.86	–4.70	–5.34	–13.2	–16.0	–18.2

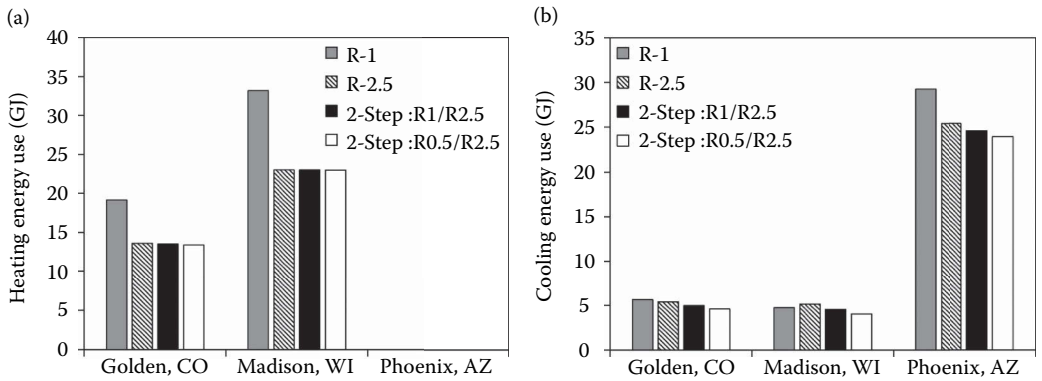


FIGURE 4.8.2 (a) Annual heating energy use and (b) annual cooling energy use for four DIM control strategy when the building has windows and internal gains in three US locations: Golden, CO; Madison, WI; and Phoenix, AZ.

energy use by up to 8% compared to RSI-2.5 wall setting in Golden, CO. While, the two-step-RSI-0.5/ RSI-2.5 control strategy reduces the cooling energy use by up to 15% compared to RSI-2.5 wall setting in Golden, CO as shown in Table 4.8.1. Cooling energy use can be reduced by 11% and 21%, respectively, with two-step-RSI-1/RSI-2.5 and two-step-RSI-0.5/RSI-2.5 control strategy in Madison, WI. Two-step DIM can achieve <6% cooling energy savings in Phoenix, AZ. Two-step DIM has little effect on the heating energy savings. This is because DIM wall maintains the high RSI value during the cold winter.

The number of hours when each DIM exterior wall is set at the low-thermal resistance over a year is shown in Figure 4.8.3 for all orientations and three US climates. DIM walls are mostly activated during cooling season in heating dominated climates (i.e., Golden, CO and Madison, WI) and during the heating season in cooling climate (i.e., Phoenix, AZ). In particular, Figure 4.8.3a and b indicates that the thermal resistance for the northern oriented walls is the most actively controlled during the heating and cooling operation. The solar radiation incident on north-facing walls is smaller than that on other walls, thus the outside surface temperature of the northern wall is generally cooler. Therefore, heat trapped inside the thermal zone can be easily released through the northern oriented walls.

Figure 4.8.4 summarizes the performance of two-step DIM control strategies compared to static RSI-2.5 wall setting during heating and cooling seasons for the residential building with and without windows and internal gains located in the three US locations. The two-step DIM control strategies

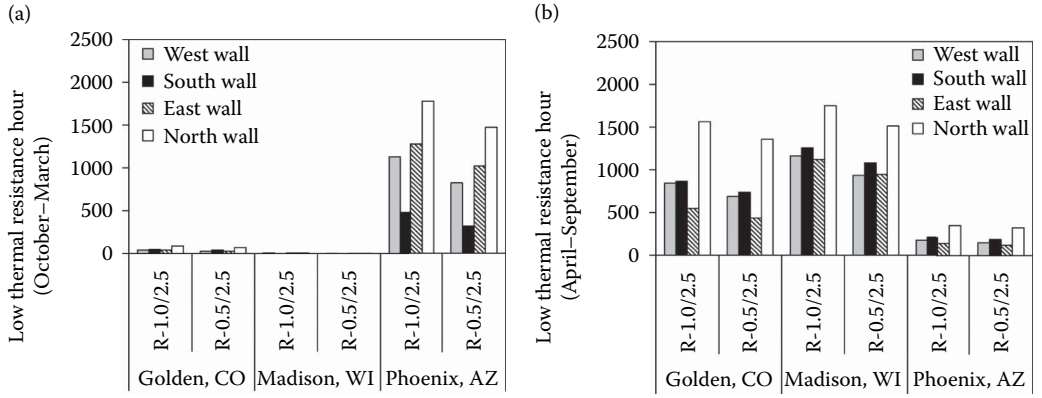


FIGURE 4.8.3 Number of hours for maintaining low RSI value for the exterior walls in the building with windows and internal load in three US locations: Golden, CO; Madison, WI; and Phoenix, AZ.

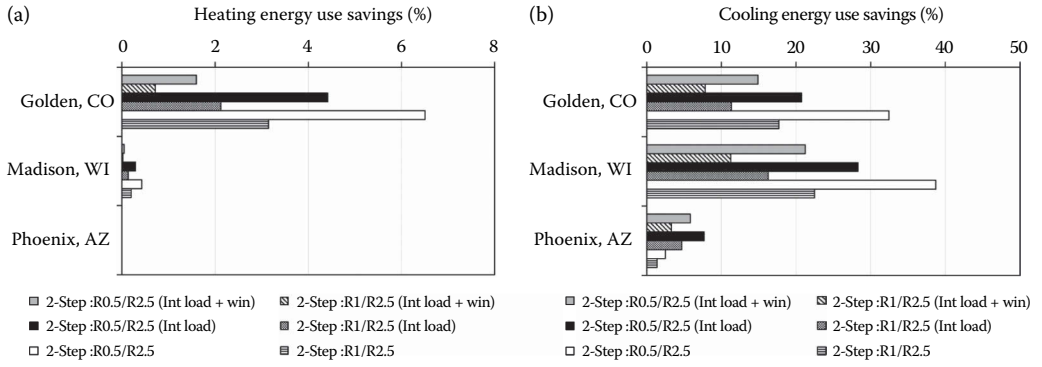


FIGURE 4.8.4 Percent annual energy use savings in (a) heating and (b) cooling for two-step DIM control strategies relative to RSI-2.5 wall setting associated to a residential building in three US climates.

are effective in cooling season regardless of locations. The two-step DIM control wall can achieve more cooling energy savings in heating dominant climates (i.e., Golden, CO and Madison, WI) during the cooling season. As the results of Figure 4.8.4 indicate, the two-step DIM control strategies are effective when the internal loads are minimal.

Variable Reflectivity Coating Roofs

Background

Cool roofs have been utilized as one of passive cooling techniques to reduce cooling energy demand for buildings. Using highly reflective coatings, cool roofs reduce solar heat gains during the summer and can keep exterior roofs cooler compared to roofs with standard reflectivity properties (Akbari and Levinson, 2008).

The benefits of cool roofs to reduce cooling energy use for air conditioning of buildings during the summer and improve thermal comfort for nonair conditioning buildings have been evaluated and tested for several building types and regions around the world (Akbari et al., 2001; Levinson et al., 2005; Akbari and Levinson, 2008; Gao et al., 2014). However, cool roofs have been found to be not necessarily beneficial during winter seasons when buildings are required to be heated. In fact, several studies have noted that cool roofs actually increase building heating loads by reducing the amount of solar heat gains.

In this section, the performance of variable reflectivity roofs, referred to throughout this section as variable reflectivity coating (VRC) roofs, to the static conventional cool roofs in order to assess the potential benefits of developing optically active coatings suitable for opaque building envelope systems. The analysis is based on the work of Park and Krarti (2016).

Thermal Analysis

A series of simulation are conducted to investigate the effectiveness of VRCs applied on the exterior opaque building envelope surfaces. Various combinations of reflectivity settings for building envelope systems (i.e., walls and roof) are considered as shown in Table 4.8.2 based on the potential achievable optical properties for thermochromic materials (Kamalisarvestani et al., 2013). Cases 1 and 2 are considered as the reference cases: Case 1 represents the configuration when the office building has conventional finish coatings for all its walls and roof while Case 2 is associated with configuration when a static (i.e., with constant reflectivity value) cool roof is utilized for the building.

For the VRC material, the solar reflectivity is changed from a high value (i.e., 0.55 in this study) during the cooling period to a low value (i.e., 0.30 in this study) during the heating period.

TABLE 4.8.2
Options for the Reflectivity Settings Applied to the Office Building Exterior Walls and Roof

No. Case	Reflectivity of Exterior Walls	Reflectivity of Exterior Roof
1	0.30	0.30
2	0.30	0.55
3	0.30	VRC
4	VRC	0.30
5	0.55	0.55
6	0.55	VRC
7	VRC	0.55
8	VRC	VRC

TABLE 4.8.3
Insulation Level of Roof and Wall Constructions Used for the Analysis

Insulation Level	Construction Type	*R*-Value (K-m²/W)
High	Roof	5.25
	Wall	2.02
Low	Roof	0.15
	Wall	0.36

Depending on the climate, the heating and cooling periods are adjusted for the office building. It should be noted that the solar reflectivity of the building envelope affects the level of solar radiation absorbed by the building exterior surfaces and thus their temperature. Any variation of outside surface temperature has a great effect on the building thermal load during both heating and cooling operation, especially for poorly insulated envelope systems. In the analysis carried out in this study and to investigate the impact of *R*-value of the building envelope on the effectiveness of the VRC application for both walls and roof, two insulation levels are considered for the analysis as outlined in Table 4.8.3.

First, the analysis is performed when the office building is located in Golden, CO to assess the impact of the application of VRC on both exterior walls and roof. In particular, the analysis considers the variations of building envelope inner and outer surface temperatures as well as of the building heating and cooling thermal loads, and annual energy costs. The results are then utilized to assess the performance and the cost-effectiveness of the application of VRC materials to building envelope surfaces for the office buildings in four US climates.

Impact of Heating Thermal Loads

In this section, the impact of VRC on heating thermal loads for the office building is determined for four US sites representing different climatic conditions including Golden, CO; Chicago, IL; Phoenix, AZ; and Tampa, FL. Figure 4.8.5 summarizes the annual heating loads for four US sites, two thermal insulation levels, and several reflectivity cases listed in Table 4.8.2. In particular, Figure 4.8.5 indicates that the heating load increases for all sites and insulation levels due to the application of a high solar reflectivity coating or VRC on the opaque building envelope. The results outlined in Figure 4.8.5 show, however, that the heating load is more sensitively affected by high reflectivity coatings when the building envelope is poorly insulated. Indeed, a high reflectivity coating roof installed on a well-insulated roof increases the heating load by 1.5 MJ/m^2 in both Chicago, IL and Golden, CO. While on poorly insulated building envelope, the same high reflectivity coating increases the heating load by 26.8 and 31.5 MJ/m^2 in Chicago, IL and Golden, CO, respectively.

Based on the results, the application of high reflectivity coating has more impact in hot climates (i.e., Phoenix, AZ and Tampa, FL) even though the absolute increase in heating load per unit area is not significant. Indeed, the increase in annual heating load in Phoenix, AZ ranges from 0.1 to 6.9 MJ/m^2 depending on the insulation level of the building envelope. This range represents 23%–62% increase in heating load caused by high reflectivity coating compared to maintaining low reflectivity surfaces. Similar results are observed for the heating load in Tampa, FL.

Impact of Cooling Thermal Loads

Figure 4.8.6 shows the variations in the annual cooling loads for four US sites (i.e., Golden, CO; Chicago, IL; Phoenix, AZ; and Tampa, FL), two thermal insulation levels (i.e., *R*-values listed in Table 4.8.3), and eight reflectivity options (i.e., cases listed in Table 4.8.2). The results outlined in Figure 4.8.6 clearly indicate that high reflectivity coating or VRC can greatly reduce the building

cooling load in all four US climates regardless of the thermal insulation level. Similar to the results obtained for the heating loads, the application of high solar reflectivity coatings on the envelope outer surfaces strongly affects the cooling load when the building is poorly insulated. It should be noted, however, that significant cooling load savings can be achieved by employing cool roofs or VRCs for the building envelope surfaces in all locations. The cooling load reduction per unit area is estimated to be 0.4 and 7.4 kWh/m^2 when the office building is located in Golden, CO with, respectively, high and low insulation level. The cooling load per unit area is reduced by 1.6 kWh/m^2 (well-insulated case) and 18.9 kWh/m^2 (poorly insulated case) when the building is in Phoenix, AZ. The percent savings of the cooling load is smaller in Phoenix, AZ and Tampa, FL, even though the reduction per unit area is two times larger than those of Chicago, IL and Golden, CO.

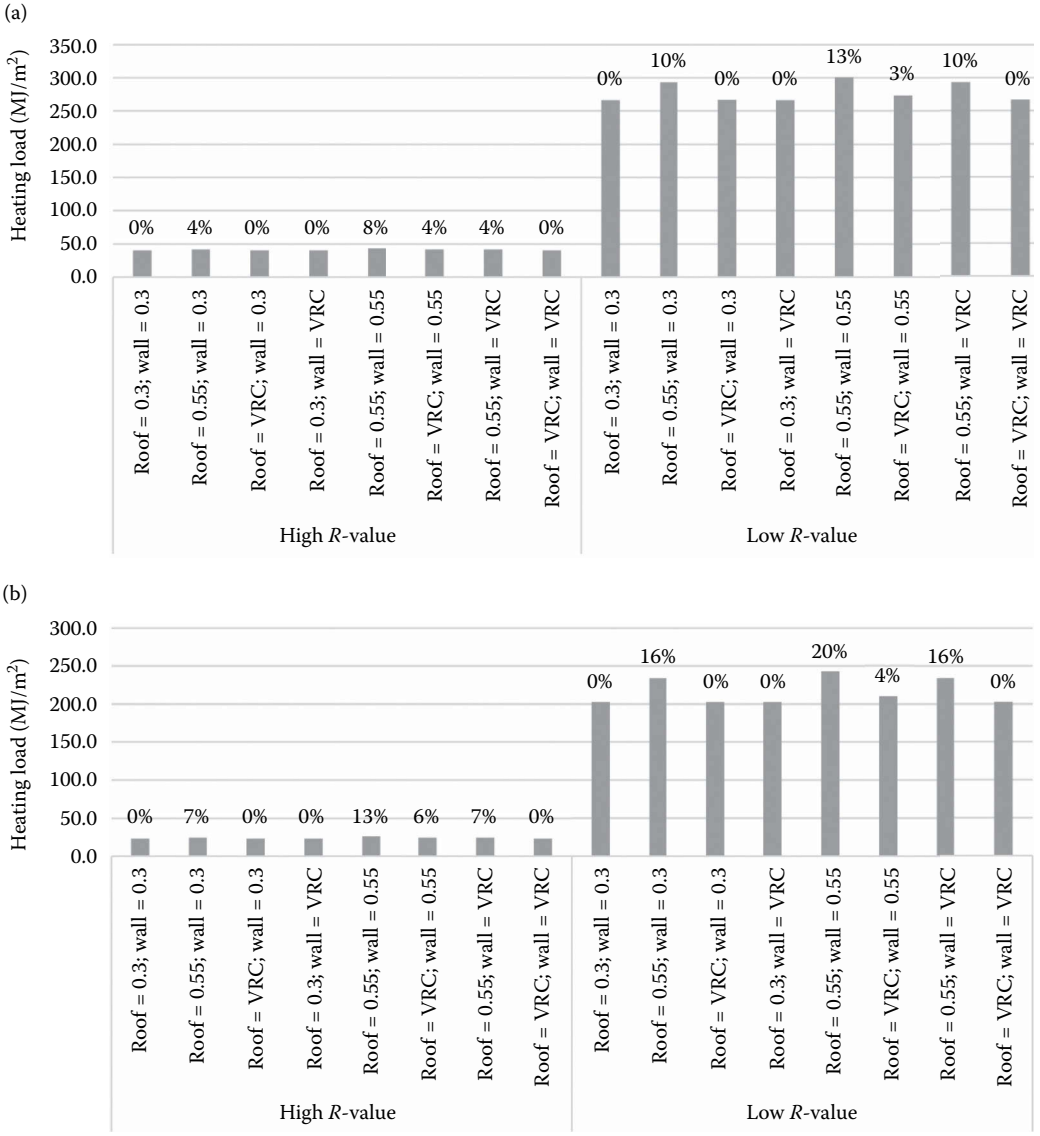


FIGURE 4.8.5 Heating load per unit area of the building with various reflectivity combinations of exterior walls and roof in (a) Chicago, IL; (b) Golden, CO; (c) Phoenix, AZ; and (d) Tampa, FL.

(*Continued*)

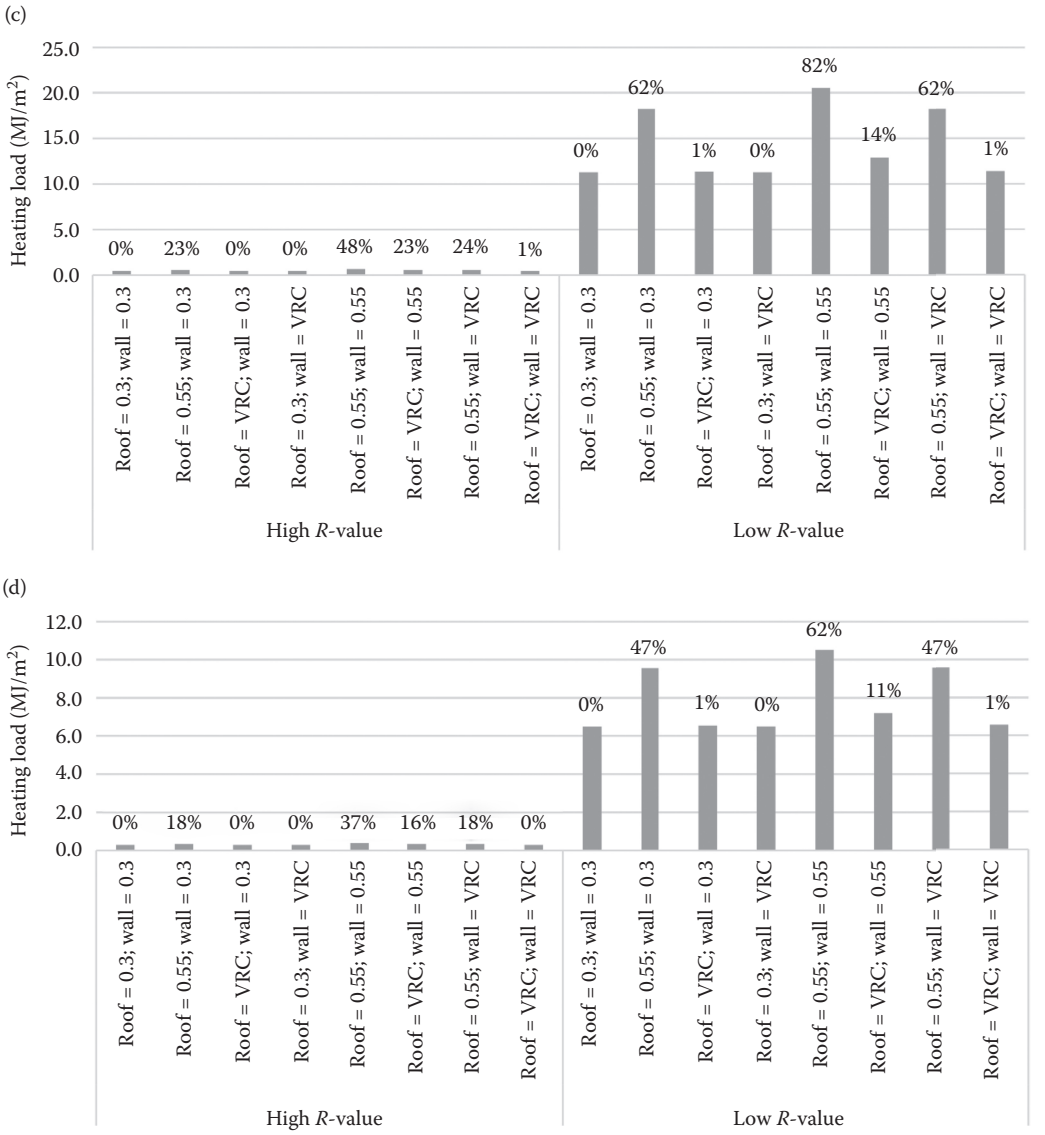


FIGURE 4.8.5 (CONTINUED) Heating load per unit area of the building with various reflectivity combinations of exterior walls and roof in (a) Chicago, IL; (b) Golden, CO; (c) Phoenix, AZ; and (d) Tampa, FL.

Impact on Annual Energy Costs

Since the seasonal reflectivity control strategies, associated with VRCs considered in this study, affect both heating and cooling thermal loads, annual energy cost analysis combining the costs for both heating and cooling is required to determine the potential benefits and cost-effectiveness of various coating options. For this study, average utility rates of \$7.1/GJ and \$0.1/kWh are considered for, respectively, natural gas (used for heating) and electricity (used for cooling). In this section, a sensitivity analysis is carried out to assess the cost benefits of applying VRCs for the office building located in four US climates. Figure 4.8.7 presents the variations in the annual energy costs per floor area cooling loads for all the insulation and reflectivity cases considered in this study. In particular, the results of Figure 4.8.7 show that the application of VRC to poorly insulated building is effective in reducing the annual heating and cooling energy costs compared to the other coating options

including the static cool roof. Indeed, the dynamic cool roof, through the application of VRC on the roof alone, can save up to 7% in energy costs when compared to the static cool roof option for poorly insulated buildings located in Golden, CO. It should be noted that the simple use of static cool roofs is highly effective in reducing energy costs, especially in hot climates. Indeed, static cool roofs reduce the office building annual energy costs by 25% and 28% for the poorly insulated building located in Phoenix, AZ and Tampa, FL, respectively. Lower but still noticeable energy cost reduction of 14% and 17% can be achieved from static cool roofs in Chicago, IL and Golden, CO, respectively.

When VRCs are considered for both roof and exterior walls, up to 28% of energy cost savings can be achieved when the building is located in Golden, CO, representing 11% additional cost reduction compared to the static cool-roof system. In hot climates, such as that of Tampa, FL, the

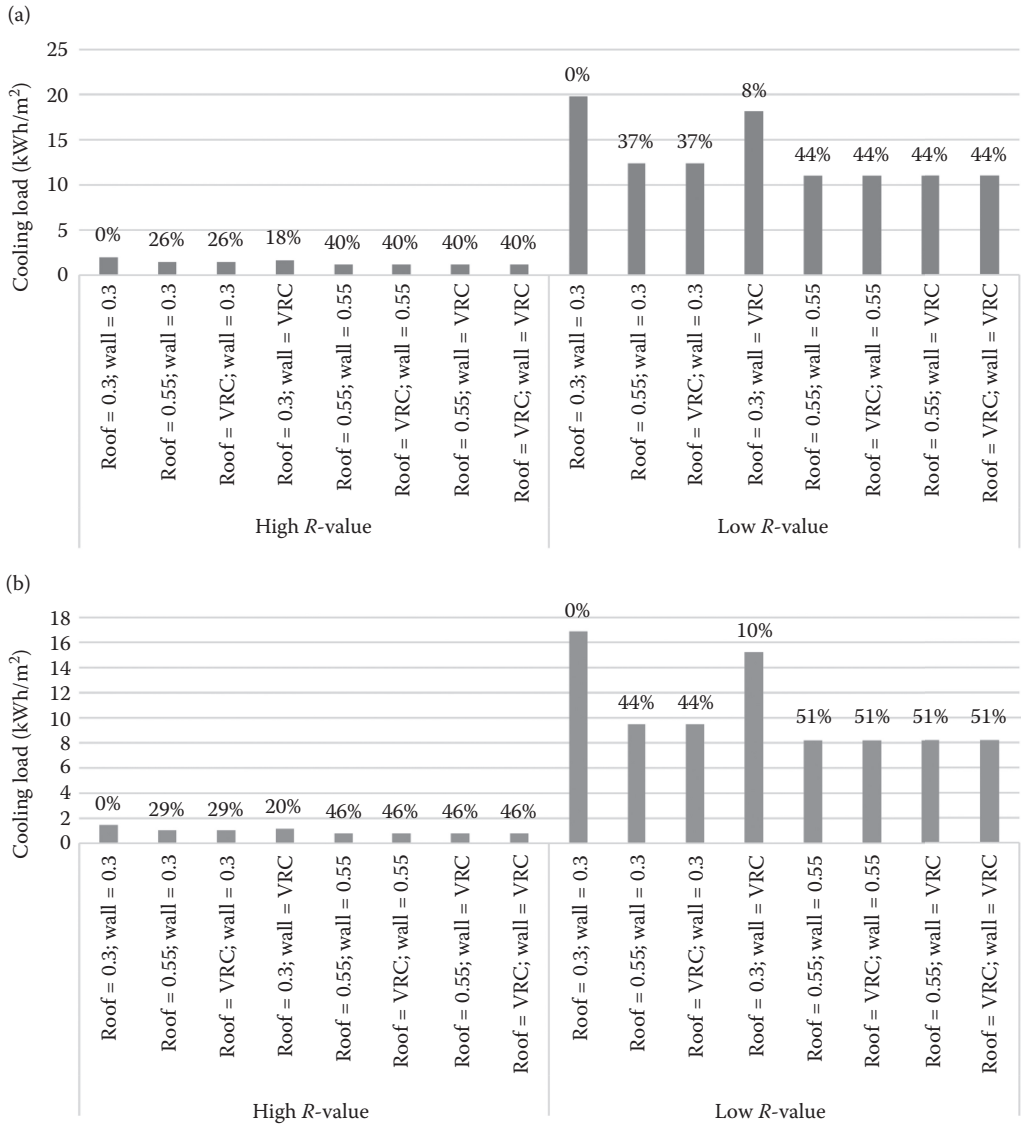


FIGURE 4.8.6 Cooling load per unit area of the building with various reflectivity combinations of exterior walls and roof in (a) Chicago, IL; (b) Golden, CO; (c) Phoenix, AZ; and (d) Tampa, FL.

(*Continued*)

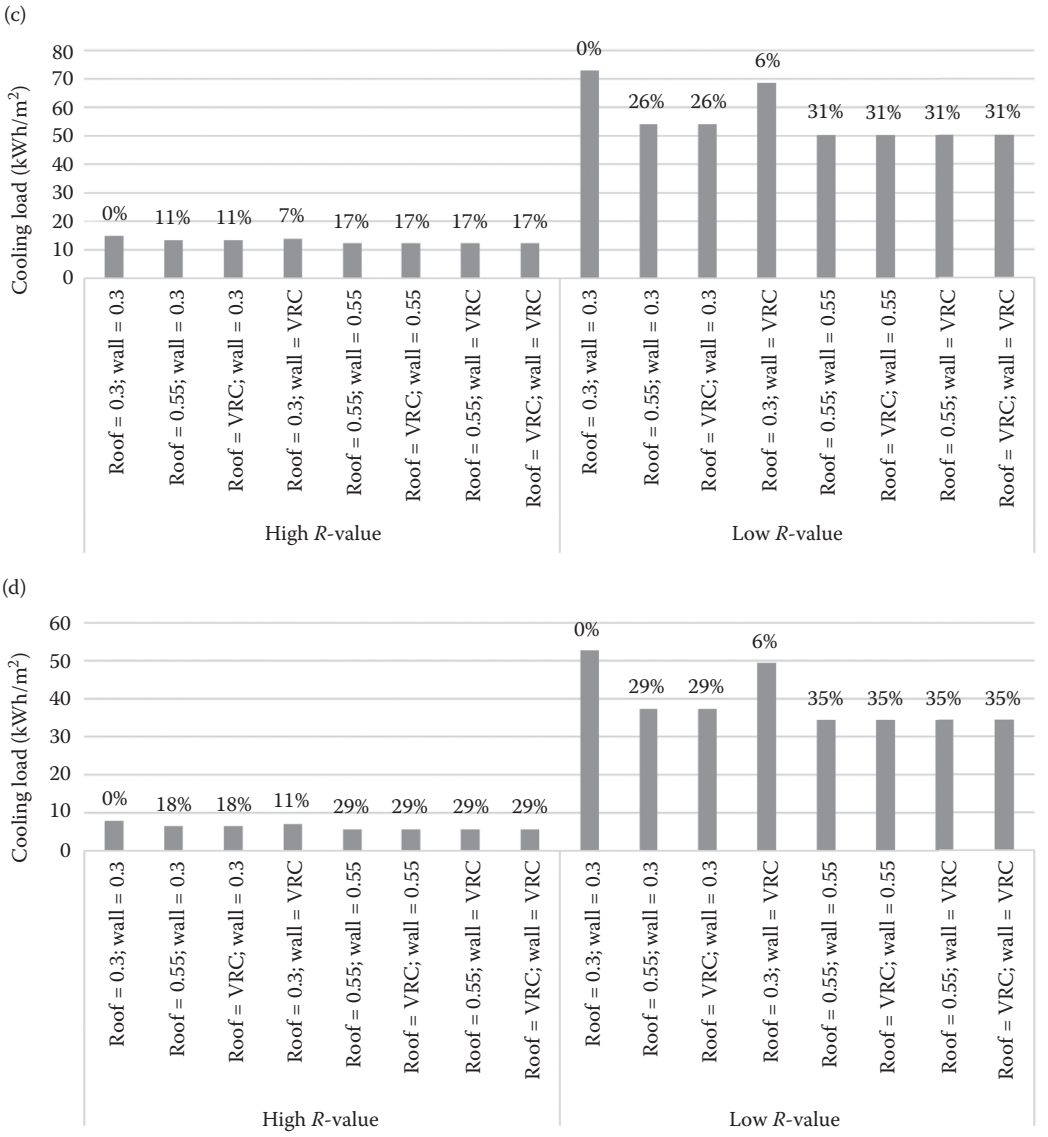


FIGURE 4.8.6 (CONTINUED) Cooling load per unit area of the building with various reflectivity combinations of exterior walls and roof in (a) Chicago, IL; (b) Golden, CO; (c) Phoenix, AZ; and (d) Tampa, FL.

annual energy cost savings can increase to 34% when VRCs are applied to both roof and walls from 29% when VRC is used for only the roof. However, as expected, the effectiveness of dynamic cool roofs compared to the static cool roofs is higher for cold climates and poorly insulated buildings. Indeed, VRCs can reduce the penalties associated with static cool roofs in heating dominated climates. The cost-effectiveness of VRCs depending highly on the climate, the energy prices, and the building thermal loads will be discussed in the following section.

Cost–Benefit Analysis of VRCs

Figures 4.8.8 and 4.8.9 show the variations in annual heating and cooling energy cost per unit floor area as a function of the cool roof and VRCs reflectivity for, respectively, well and poorly insulated office building located in four US sites (i.e., Golden, CO; Chicago, IL; Phoenix, AZ; and Tampa, FL). For VRC applied to opaque building envelope surfaces, the solar reflectivity is changed

from a high value as noted in the *x*-axis during the cooling season to a low value (i.e., 0.30 in this study) during the heating season. While cool roofs have the same reflectivity value shown in *x*-axis throughout the year. The results of Figures 4.8.8 and 4.8.9 clearly indicate that total annual energy cost decreases as the roof reflectivity increases for all four US locations regardless of the thermal insulation level. As noted earlier, dynamic cool roofs (i.e., with VRC applied only on the roofs) reduce the heating energy penalties associated with static cool roofs in heating dominated climates, especially for poorly insulated buildings. However, the application of VRC on both the roof and the exterior walls can considerably reduce the annual energy cost in hot as well as cold climates. In fact, VRC applied to all opaque envelope surfaces results in higher savings relative to either static or dynamic cool roofs in hot climates. For instance and when a well-insulated office building is located in Tampa, VRCs on both roof and walls reduce energy cost by 15% relative to the cool roof options when the coating high reflectivity is 0.55 and by 32% when the coating high reflectivity is 0.9.

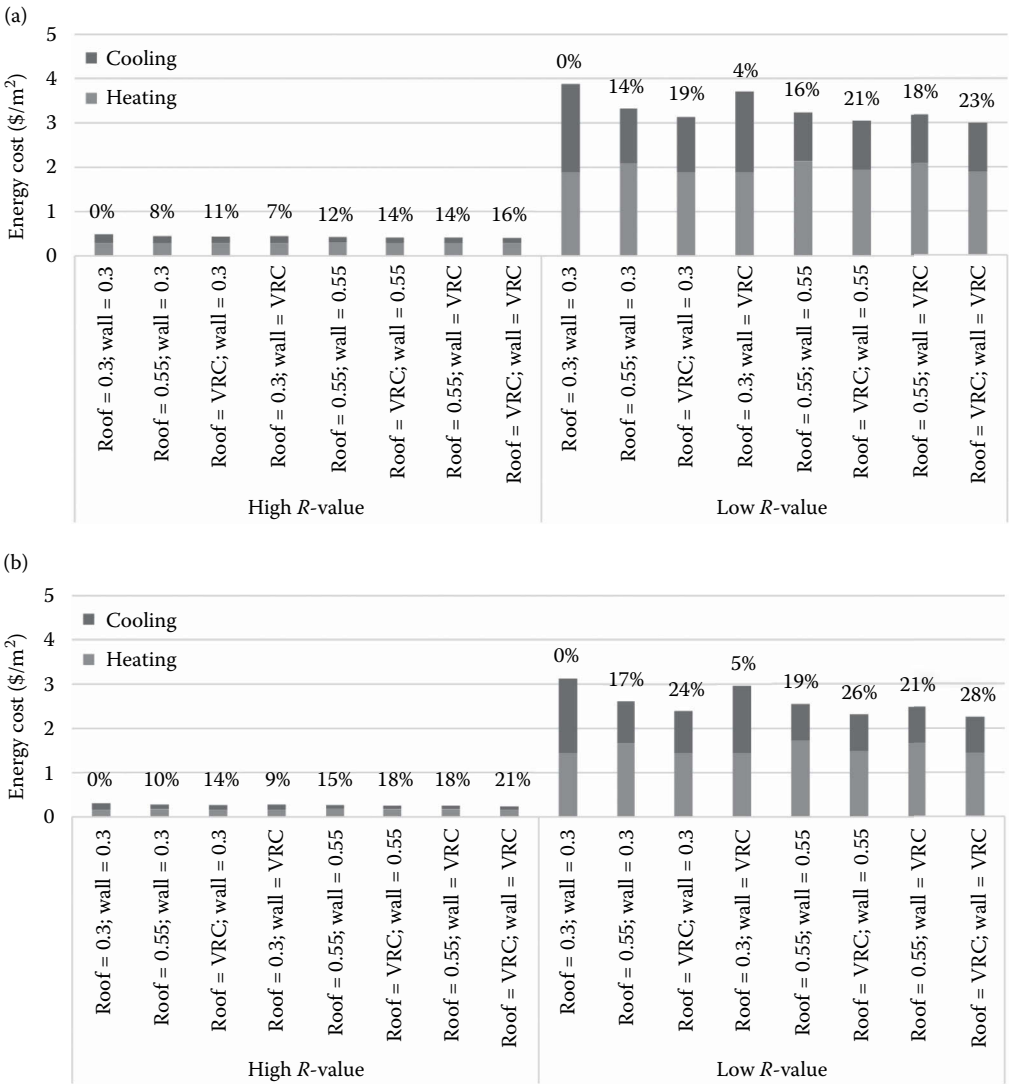


FIGURE 4.8.7 Energy cost of the building with various reflectivity combinations of exterior envelopes in (a) Chicago, IL; (b) Golden, CO; (c) Phoenix, AZ; and (d) Tampa, FL.

(*Continued*)

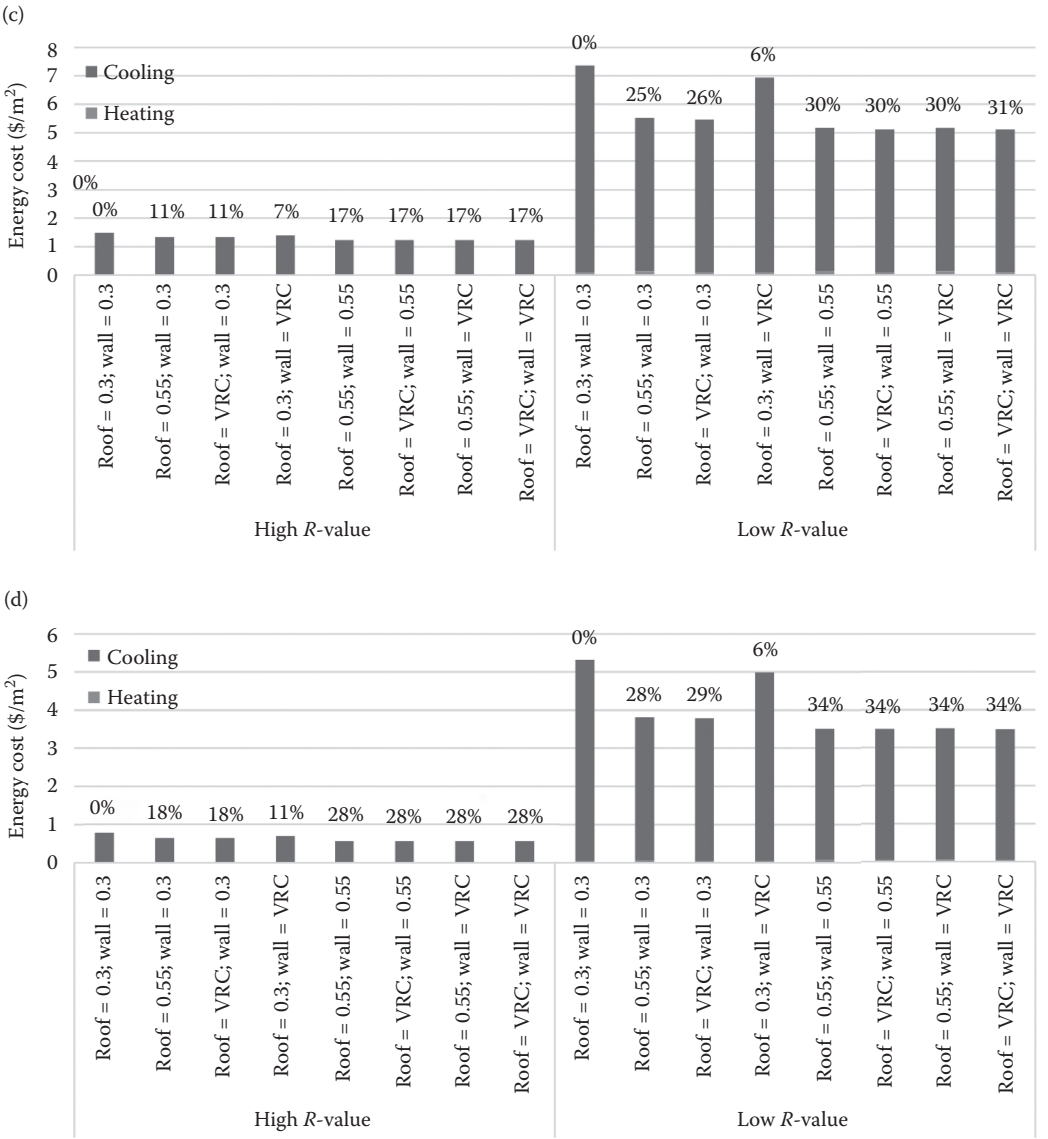


FIGURE 4.8.7 (CONTINUED) Energy cost of the building with various reflectivity combinations of exterior envelopes in (a) Chicago, IL; (b) Golden, CO; (c) Phoenix, AZ; and (d) Tampa, FL.

In fact, the ideal VRC systems with a reflectivity of 0.9 during the summer season and reflectivity of 0.3 during the winter season minimize energy costs for all climates and insulation levels considered in the analysis.

Single-Layer Breathing Wall

Introduction

A promising new technology introduces a method of avoiding the efficiency/air quality compromise, yielding better energy efficiency and improved IEQ. The technology, referred to as a "breathing wall," draws a steady stream of filtered air through the walls and into the building at all times, providing exceptionally clean ventilation air to the occupants. A schematic breathing wall diagram can be seen in Figure 4.8.10. Whereas higher ventilation rates traditionally produce higher energy

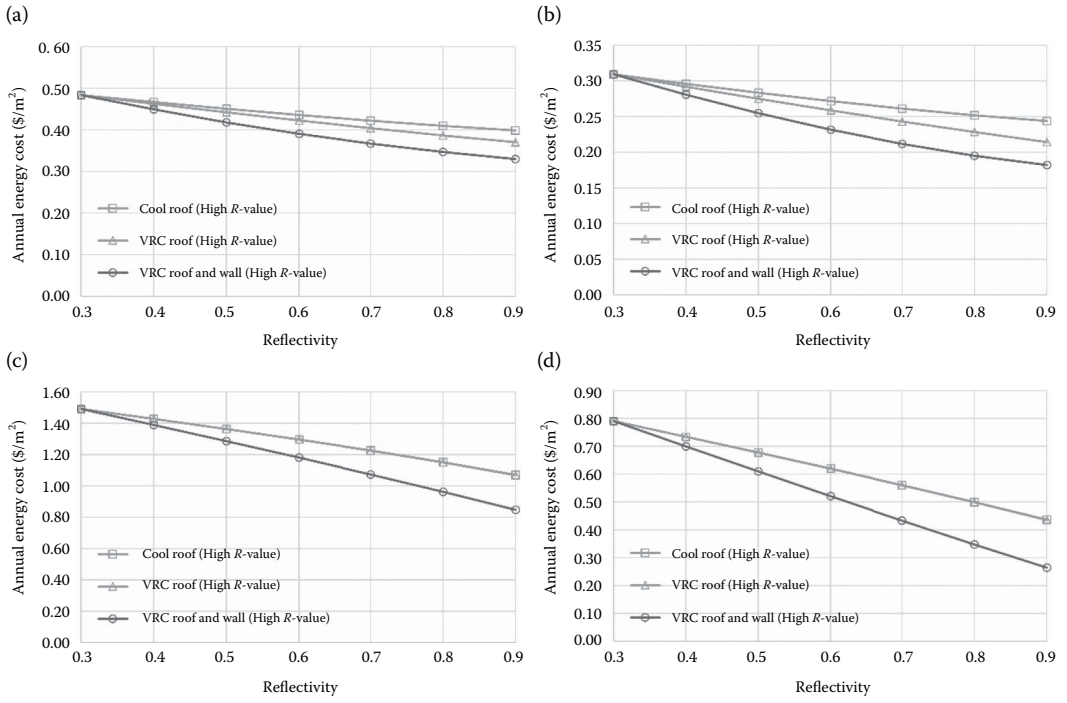


FIGURE 4.8.8 Annual energy cost savings as a function of reflectivity value of cool roof and VRCs for well-insulated building located in (a) Chicago, IL; (b) Golden, CO; (c) Phoenix, AZ; and (d) Tampa, FL.

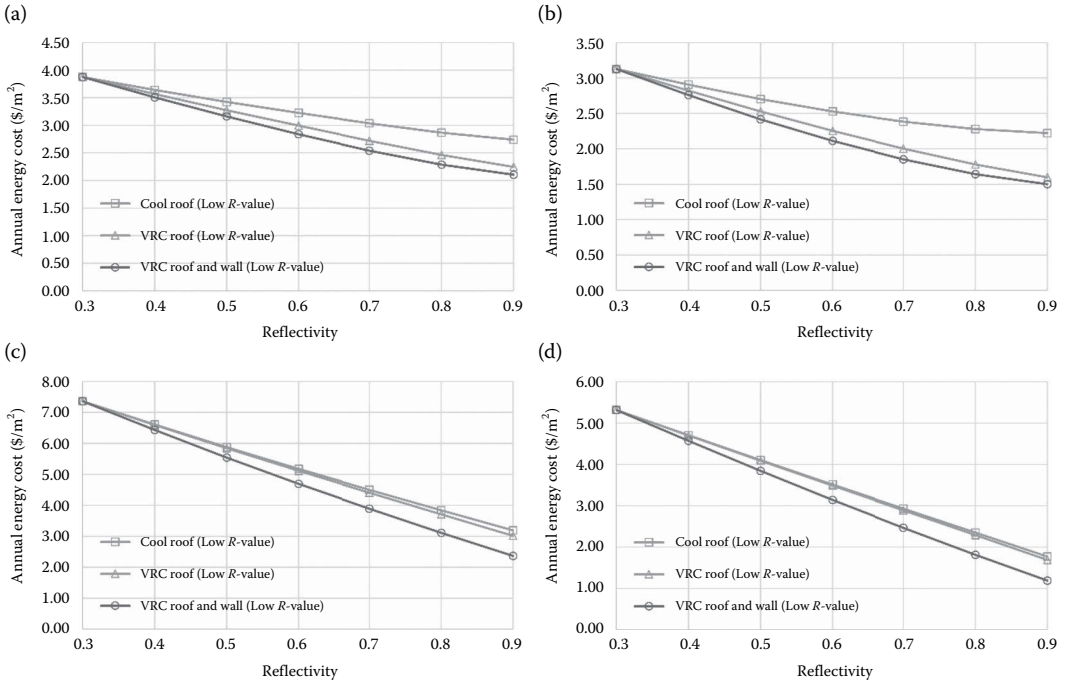


FIGURE 4.8.9 Annual heating and cooling energy cost as a function of reflectivity value of cool roof and VRCs for poorly insulated building located in (a) Chicago, IL; (b) Golden, CO; (c) Phoenix, AZ; and (d) Tampa, FL.

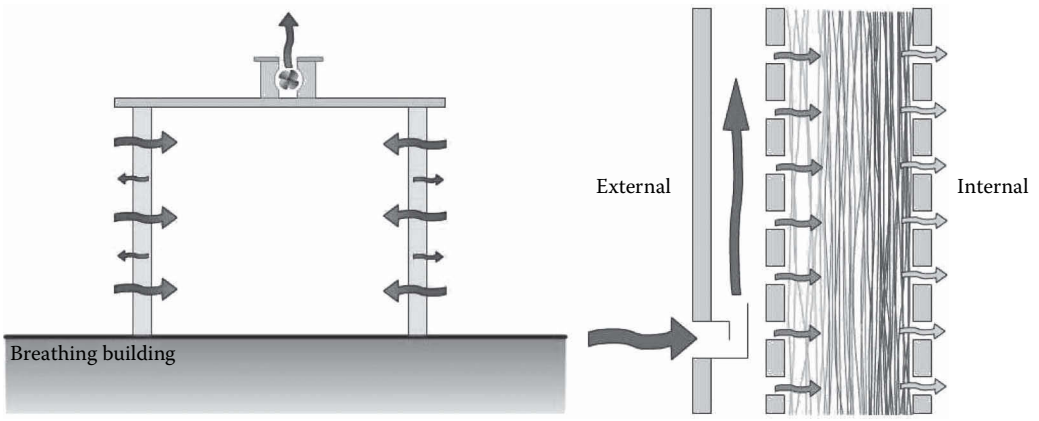


FIGURE 4.8.10 Breathing wall configurations, as proposed by Imbabi and Peacock (2003).

loads in buildings, the "dynamic insulation" used in breathing walls actually works to reduce that load, effectively creating efficient, superinsulated walls.

The projected energy savings and air quality implications associated with breathing wall technology are astounding. Previous studies of the energy and air filtration efficiencies of breathing walls estimated that such technologies can reduce year-round heating and cooling loads between 10% and 40%, while providing a steady stream of fresh ventilation air, filtered to high energy particulate air (HEPA) standards, 365 days per year (Imbabi and Peacock, 2003). Additionally, breathing walls may be able to clean up polluted cities too, as the filtered air exhausted by "breathing wall buildings" will effectively contain lower concentrations of pollutants and particulate matter than the outdoor air. The study also suggests that the filter mechanism of the walls will last throughout the lifetime of the building, providing the energy savings and air filtration for 60+ years, without requiring replacement.

While a few preliminary reports have projected promising energy and IEQ benefits of breathing walls, much research remains to be done. One major issue at hand pertains to understanding the heat transfer mechanism between the breathing wall media and incoming air, particularly under varied ambient conditions. Some of the previous research provides a simple estimation of the wall's insulative properties (U-value), involving only the conventional U-value and the flow rate of incoming air (Taylor et al., 1996). Using this method, the insulation of the wall at certain airflow rates is effectively infinite. While this method will suffice to make broad estimations of the energy savings associated with breathing walls, it may not prove accurate under changing environmental conditions. It is important to understand the behavior of breathing walls under a wide range of conditions, including temperature, humidity, pressure, and airflow rate.

Background

Among the first to conduct experimental studies on the effects of infiltration heat recovery within the building envelope were Bhattacharyya and Claridge (1995). Their one-dimensional model effectively predicted the heat transfer between moving infiltration air and the exterior building envelope. The so-called "heat exchanger" model was then revised and simplified by Barhoun and Guarracino (2004), who studied the particular case of a vertical air cavity within a wall. Separate studies by Buchanan and Sherman (2000), as well as Abadie et al. (2002), added a simulation of the airflow within the interior to a model of the heat recovery within the exterior wall element. Further research related to the movement of air throughout the building envelope has investigated the thermal and heat transfer properties of hollow bricks and double-skin façades. Additionally, Chen and Liu (2004) studied the airflow through porous media for applications within the solar-collector industry.

Several European studies, including Bailly (1985) and Kohonen and Ojanen (1987), have shown the potential energy saving impacts of dynamic insulation in buildings. Full-scale tests were conducted by Baker (2003), Dimoudi et al. (2004), and Morrison and Karagozis (1992) to study the energy performance of a dynamically insulated room. Bailly found that dynamic insulation can potentially save up to 14% of the heating load during the wintertime.

Krarti (1994) developed an analytical model to study the impacts of combined conduction and infiltration/exfiltration upon the thermal performance of dynamic insulation, and demonstrated that the heating load in applications of coupled infiltration and conduction was less than the simple summation of isolated conduction and infiltration. He demonstrates that there is a significant thermal coupling between the leakage and insulation layers that modifies the heat transmission in building thermal envelopes. The correlations developed by Krarti use the Peclet number to indicate the velocity of air moving through the wall. However, the savings is limited due to the fact that the amount of air forced through the dynamic wall must be equal to the air being exhausted. Analysis of the wall temperature profiles shows that the temperature profile within the wall largely depends upon the direction and rate of airflow through the wall. When no air is passing through the wall, the profile is linear, while it takes on a convex shape when air is exfiltrating, and a concave shape when air is infiltrating. The temperature profile throughout the wall is important, as it affects the heat loss due to conduction through the wall. Krarti concludes that the wall as a heat exchanger is capable of reducing the daily heating load due to conduction and infiltration by up to 22%.

Taylor et al. (1996) likewise developed a one-dimensional steady-state model of multilayer breathing wall heat transfer, and proposed to measure the performance through a dynamic *U*-value. One of Taylor's key findings is that very different types of construction will have the same thermodynamic performance with the same airflow, provided the thermal resistance is equal. Another significant finding of the analytical research conducted by Taylor et al. (1996) was that pro-flux flow (heat and airflow in the same direction) and contra-flux flow (heat and airflow in opposite directions) exhibit asymmetrical behavior. Imbabi and Peacock (2003) conducted physical and analytical experiments on dynamically insulated walls. Their research suggests that a layer of wet-blown cellulose with a static *U*-value of 0.168 W/m^2 K, when subjected to an infiltration airflow rate of 1 m/h can achieve a dynamic *U*-value of 0.058 W/m^2 K, which falls further to 1.7E−8 W/m^2 K at an airflow rate of 10 m/h.

Qiu and Haghighat (2007) published research regarding the temperature profile at various depths within a ventilated wall, as well as the concept of treating the wall as a combination of ventilated and nonventilated areas. They developed a numerical simulation in order to study the thermal performance of the diffuse building envelope, and conducted physical testing in order to corroborate the numerical results. Their computer simulations indicate that at low airflow rates, the temperature of the air rises very quickly to approximate the temperature of the solid matrix, allowing for a single medium analytical treatment. This research relies heavily upon the Peclet number (closely related to airflow rate) as a determining factor for breathing wall thermal performance. A key finding of the research is that the breathing wall can be broken up into regions affected by the heat transfer of infiltration, and regions that remain unaffected by infiltration, and are therefore dominated by pure conduction heat transfer. Additionally, it was determined that the ratio of these two areas is one of the key indicators of breathing wall performance.

Existing Products and Technologies

Among the earliest applications of dynamic insulation within buildings were the ceilings of agricultural buildings in Norway in the 1960s (Græe, 1974). Improvements in dynamic insulation technology led to the application of breathing walls within traditional barns, and then in the 1980s, inside a small number of traditional residential dwellings. Among the more contemporary installations of breathing wall technologies is the first major building in the United Kingdom to incorporate such a strategy, the McLaren Community Leisure Centre, in Callander, Scotland.

There are several products currently on the market that incorporate dynamic insulation or some of the properties of breathing walls. A start-up company in the United Kingdom and associated with the research team of Imbabi and Peacock markets "innovative modular breathing wall systems to provide clean, preheated, filtered ventilation air to the building." This project has received substantial funding from the Scottish Executive SMART Program, in order to facilitate development. The specific product sold by EBP is called an Energy*flo* Cell, and it can be easily integrated into a building structure to incorporate the benefits of dynamic insulation. Additionally, Gaia Research Group included dynamically insulated squash courts, bowling hall, sports hall, and swimming pool rooms in the McLaren Community Leisure Centre, which was finished in 1998 in Callander, Scotland. This building was the first major building in the United Kingdom to use dynamic insulation, and is undergoing a detailed monitoring program. Several other companies market "breathing wall" technologies as well, but many of these deal with breathability as it relates to building moisture and mold, rather than the energy-saving performance of dynamic insulation.

Measured and Predicted Thermal Performance

A small-scale breathing wall test facility is developed and used to characterize the performance of a prototype breathing wall (Zhai and Slowinski, 2013). Various experimental parameters, including temperature, pressure, airflow rate, and auxiliary heating load were measured and analyzed.

Test Chambers and Breathing Wall

In order to simulate a wide variety of exterior and interior environmental conditions, separate test chambers of 1.1176 m × 1.1176 m × 1.1176 m each, simulating interior and exterior environmental conditions, are built with R-13 (h-ft^2-°F/Btu) polystyrene panels and can be heated to varying degrees by a system of dimming incandescent lamps. Allowing the temperatures to reach a steady state, this system facilitates breathing wall simulations at winter conditions with a temperature differential of up to 38°C. The test chambers also possess the capability to vary pressure differentials and therefore airflow rates by means of an exhaust fan and a variable outlet damper, in order to develop a pressure drop–flow rate curve for each type of breathing wall construction. Pressures are varied from 2 to 12 Pa and the corresponding air volume flow rates are measured. A small fan is located within the interior space, to encourage even mixing of the interior air. The final configuration can be seen in Figure 4.8.11.

The breathing wall is constructed with outside dimensions of 1.1176 m by 1.1176 m, and features interior and exterior cladding, made of 6.35 mm thick plywood. The exterior façade features an inlet grille located in the center of the wall, 0.767 m up from the bottom. The interior façade features a similar exhaust grille located 0.2 m from the bottom. An air gap is created between the cladding element and the porous breathing wall material, such that air is allowed to freely flow away from the inlet vent and penetrate the porous material in a relatively even fashion. The prototype breathing wall was constructed using commercially available unfaced fiberglass batt insulation, which was spread out across a fiberglass window screen and secured within the wood plane of the wall.

Temperature is measured by thermocouples at 44 different locations, both inside and outside of the test chamber (shown in Figure 4.8.11). The selected locations allow for the determination of average interior and exterior temperatures, average interior wall temperatures, exhaust air temperature, laboratory ambient temperatures, breathing wall vent temperatures, as well as simple wall temperature profiles at four poles throughout the breathing wall media. Thermocouples are embedded at various depths (25%, 50%, and 75%) throughout the breathing wall, to form four different "poles" (see "filter media poles" in Figure 4.8.11) so as to obtain simple temperature profiles throughout the wall in desired locations. Additional thermocouples are located in both the interior and exterior air gaps of the poles. One thermocouple is located within the exterior vent and another within the interior vent, in order to measure both the inlet and exhaust breathing wall air temperatures. Agilent VEE Pro 2.4 is used to monitor and record all of the data from the thermocouples, resistance temperature detector, and digital manometer. Experiments continue as long as necessary

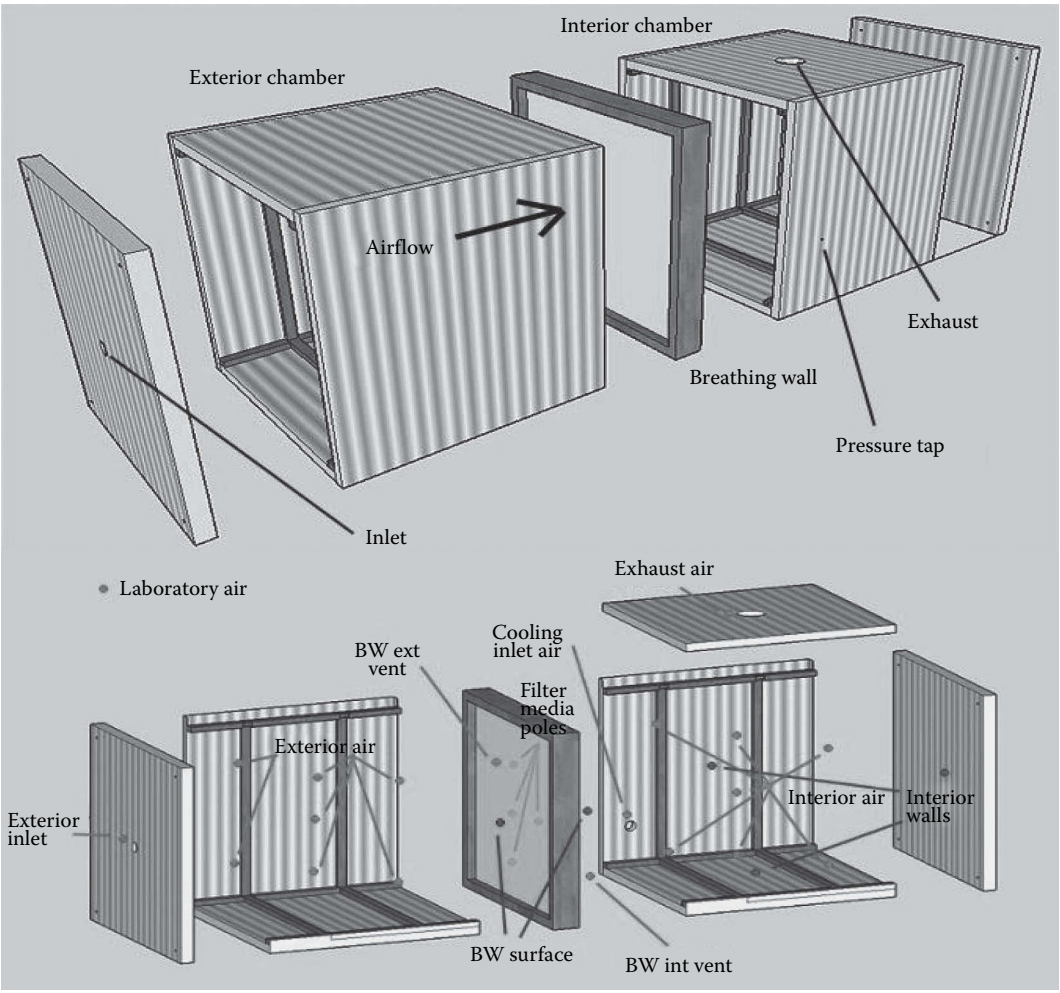


FIGURE 4.8.11 Experimental test chamber setup and thermocouple locations.

to achieve average temperature changes <0.5°C per hour. In some cases, steady-state conditions can be achieved within 2 h, while in others, it may take more than 24 h. The measurement accuracy meets or exceeds all relevant ASHRAE/ASTM standards.

Measurement Results and Analysis

In order to properly document the energy savings of the ventilated breathing wall assembly over a static wall construction, a base case is defined, from which the static wall heat loads can be calculated. The base case consists of a conditioned stream of ventilation air entering through a simple vent in one of the test chamber walls, rather than passing through the breathing wall. Winter conditions are simulated by heating the interior test chamber up to temperatures as high as 60°C, to provide contrast with the "outdoor" environmental conditions found in the exterior chamber, at 22°C. In each of the 13 base case runs, temperature, pressure, and airflow measurements were taken in the same fashion as in the dynamic breathing wall cases. Twenty experiments were performed with varying "outdoor" conditions simulated in the exterior chamber and with ventilation air flowing through the breathing wall. The 20 dynamic runs consist of permutations of four different airflow rates across the breathing wall and five different temperature differentials, with a temperature range similar to that of the base case runs.

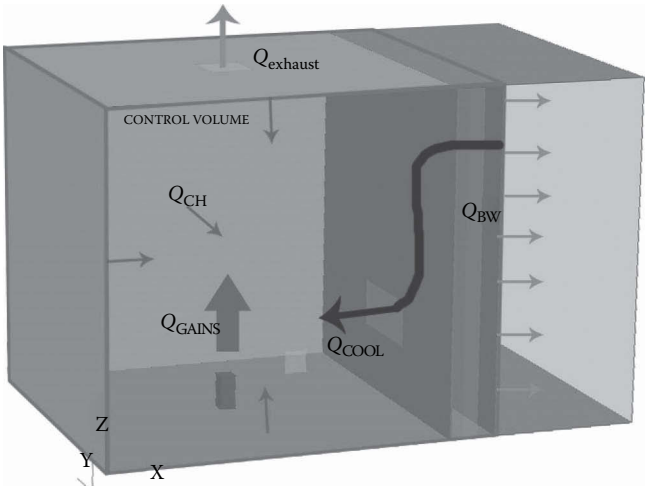


FIGURE 4.8.12 Energy balance model for the test chamber.

It is important to quantify the thermal performance and potential energy saving of breathing walls, such that the benefits of breathing wall technology can easily and effectively be incorporated into whole building energy simulations. Efficiency values that allow for interpretation of the breathing wall's performance can be developed from the energy balance equation. The control volume is defined as the interior surfaces of the interior chamber walls, extending out to the exterior surface of the breathing wall. The energy balance equation accounts for HVAC energy supply within the interior space, as well as the air being exhausted by the exhaust fan and heat loss through chamber walls. A conceptual model of the experimental model energy balance can be seen in Figure 4.8.12.

Equations in the experimental model energy balance include the exhaust energy flow, Q_{exhaust}, as seen in Equation 4.8.1, and the enthalpy being introduced into space as ventilation air through the breathing wall, Q_{cool}, seen in Equation 4.8.2. Note that in the dynamic cases, $T_{\text{cooling-inlet}}$ represents the temperature at the breathing wall exterior vent, while in the base cases, it represents the temperature at the supply inlet on one of the chamber walls:

$$Q_{\text{exhaust}} = \dot{m} \times C_{\text{p}} \times T_{\text{exhaust}} \tag{4.8.1}$$

$$Q_{\text{cool}} = \dot{m} \times C_{\text{p}} \times T_{\text{cooling-inlet}} \tag{4.8.2}$$

Equation 4.8.3 shows the heat flow through the test chamber walls. The UA value of the test chamber is derived from the tests in Section 3.1:

$$Q_{\text{ch}} = UA_{\text{ch}} \times \left(T_{\text{interior-walls}} - T_{\text{lab}}\right) \tag{4.8.3}$$

Q_{gains} is the combination of the heat input from the bulb and from the mixing fan. The final energy balance equation for the test chamber provides an analytical solution for Q_{BW}, shown in Equation 4.8.4:

$$Q_{\text{BW}} = Q_{\text{cool}} + Q_{\text{gains}} - Q_{\text{ch}} - Q_{\text{exhaust}} \tag{4.8.4}$$

Krarti (1994) defined the efficiency of breathing walls (referred to as η_{l}) as a function of the total energy needed by a case with breathing walls, Q_{dynamic}, and that used in a similar scenario with dedicated vents, Q_{static}, as in Equation 4.8.5:

$$\eta_{\text{l}} = 1 - \frac{Q_{\text{dynamic}}}{Q_{\text{static}}} \tag{4.8.5}$$

Taylor et al. (1996) defined efficiency (referred to as η_2) as a function of breathing wall conduction heat exchange (excluding the heat being carried by airflow) and static breathing wall heat exchange, as seen in Equation 4.8.6:

$$\eta_2 = 1 - \frac{Q_{\text{BW, dynamic}}}{Q_{\text{BW, static}}} \tag{4.8.6}$$

In order to account for any increase or decrease in conduction heat transfer through the breathing wall, this study defined the term U_{dynamic} in Equation 4.8.7 simply as the heat flow through the non-vent portion of the breathing wall, divided by the product of wall area and temperature differential across the wall surfaces:

$$U_{\text{dynamic}} = \frac{Q_{\text{BW, dynamic}}}{A_{\text{BW}} \times \left(T_{\text{inside-surface}} - T_{\text{outside-surface}}\right)} \tag{4.8.7}$$

The third definition of efficiency, referred to as η_3, is defined as in Equation 4.8.8 (Zhai and Slowinski, 2013):

$$\eta_3 = 1 - \frac{U_{\text{dynamic}}}{U_{\text{static}}} \tag{4.8.8}$$

Many building energy simulation programs, such as DOE-2 or EnergyPlus, calculate the amount of energy load dedicated to conduction losses through building walls and the amount of heating outside air for ventilation purposes. If U_{dynamic} for a dynamic wall can be determined and used to replace the corresponding U_{static} in the input file of a building energy simulation program, the potential energy savings on an hourly basis using the dynamic wall can then be predicted with these energy simulation programs.

Multiple regression analyses were conducted on the experimental data points, and correlations were developed, respectively, for Q_{static}, Q_{dynamic}, $Q_{\text{BW,static}}$, $Q_{\text{BW,dynamic}}$, U_{static}, and U_{dynamic}. The developed correlation between interior–exterior temperature differential, airflow rate, and dynamic case Q_{dynamic} can be seen in Equation 4.8.9. The calculated R^2 value for this correlation is 93.15%, indicating a fairly close match. In particular, the correlation predicts values with lower values of Q_{dynamic} very accurately, while the data points with higher Q_{dynamic} values appear more scattered:

$$Q_{\text{dynamic,exp}} = \left(6.176 \times \ln\left(1.249 \times \dot{V}\right) - 5.556\right) \times \Delta T_{(\text{avg, inside-avg, outside})} \tag{4.8.9}$$

where $\dot{V}$ is the average "breathing" air velocity over the entire wall surface in CFM/m^2. The applicable mass flow and temperature differential ranges are 5.9–16 CFM and 9°C–38°C, respectively. In contrast to the dynamic case values, the Q_{static} base case values conform much more closely to the determined correlation, shown in Equation 4.8.10. This correlation has an R^2 value of 99.53%, indicating a very close match. Consequently, there is no trend of overprediction or underprediction from the correlation:

$$Q_{\text{static,exp}} = 1.108 \times \dot{V} \times \Delta T_{(\text{avg, inside-avg, outside})} \tag{4.8.10}$$

Similarly, $Q_{\text{BW, static}}$, $Q_{\text{BW, dynamic}}$, U_{static}, and U_{dynamic} are obtained below:

$$Q_{\text{BW, dynamic, exp}} = 1.781 \times \dot{V} + 3.658 \times \Delta T_{(\text{avg, inside-avg, outside})} - 15.483 \tag{4.8.11}$$

$$Q_{\text{BW, static, exp}} = (0.241 \times \dot{V} + 1.866) \times (\Delta T_{(\text{avg, inside-avg, outside})} \times 0.821 + 2.739) \tag{4.8.12}$$

$$U_{\text{dynamic,exp}} = 0.103 \times \dot{V} + 2.165 \tag{4.8.13}$$

$$U_{\text{static,exp}} = 0.211 \times \dot{V} - 0.097 \times \Delta T_{(\text{avg, inside-avg, outside})} + 2.883 \tag{4.8.14}$$

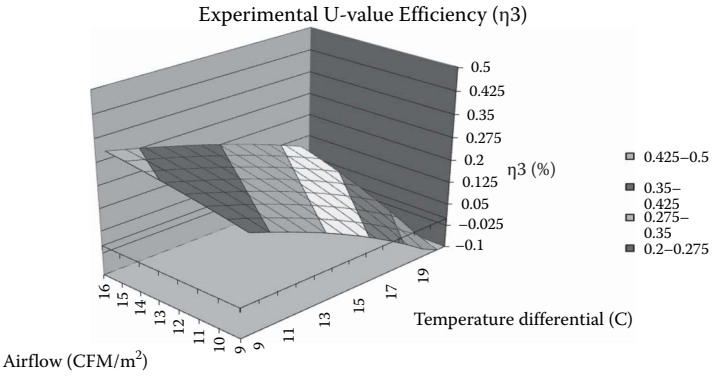


FIGURE 4.8.13 Experimental efficiency (η_3) for varying temperature gradients and airflow rates.

Now that trends in the base and dynamic case indicators have been obtained, they can be compared to examine the measures of efficiency. Previous research has suggested that efficiency measures, at least for one-dimensional cases and for breathing walls with no internal air cavities, depend only upon the airflow rate and not upon the temperature differential. The developed correlations certainly suggest that flow rate plays more of a role, as the coefficients of flow rate in the equations are generally larger. However, the experimental results suggest that temperature differential does have some impact upon efficiency trends. Figure 4.8.13 shows the η_3 contour (defined in terms of U_{dynamic}). The experimental results show a clear positive correlation between airflow rate and efficiency, as well as a clear negative correlation between temperature differential and efficiency. Efficiencies range from –10% to 30%.

Predicted Results with Analytical Formula and Computational Fluid Dynamics Simulation

The study compared the calculated efficiency of the tested wall to the analytical solutions obtained by Krarti (1994) and Taylor et al. (1996). Both analytical solutions were developed for ideal and uniform flow through one-dimensional porous walls. The three-dimensional experiment has small inlet and outlet vents on the walls as well as interior wall cavities, which will introduce different heat transfer behaviors. Figure 4.8.14a demonstrates that the efficiency η_1 does not vary with temperature differential. Taylor's efficiency η_2 appears to vary with temperature differential as noted in Figure 4.8.14b. Experimental efficiencies have similar trends to the analytical solutions, although the efficiencies are much lower, with negative efficiencies below certain airflow rates.

A computational fluid dynamics (CFD) model was created to explore the flow details in and out of the breathing wall that could not be revealed in the physical experiment. A Reynolds Average Navier–Stokes (RANS) CFD simulation tool (with the KE Renormalization Group (KERNG) turbulence model) was used to simulate various tested chamber conditions with a 500 *K* nonuniform grid (identified via a rigorous grid-independence study). Figure 4.8.15 illustrates that free convection circulation loops are formed at both the interior and exterior air gaps within the breathing wall. These circulation loops may play a part in increasing convection heat transfer within the wall and decreasing the efficiency as temperature differential across the wall increases. CFD provides similar efficiency results and trends as the experiments as qualitatively illustrated in Figure 4.8.16b.

Multilayer Living Walls

Background

Using biomimetic principles, multilayer living wall concept can be proposed and developed. Figure 4.8.16 illustrates the vision of such a climate-adaptive intelligent wall structure with embedded systems of air (respiration in body), water (blood in body), and PCM (fat in body). It shows a

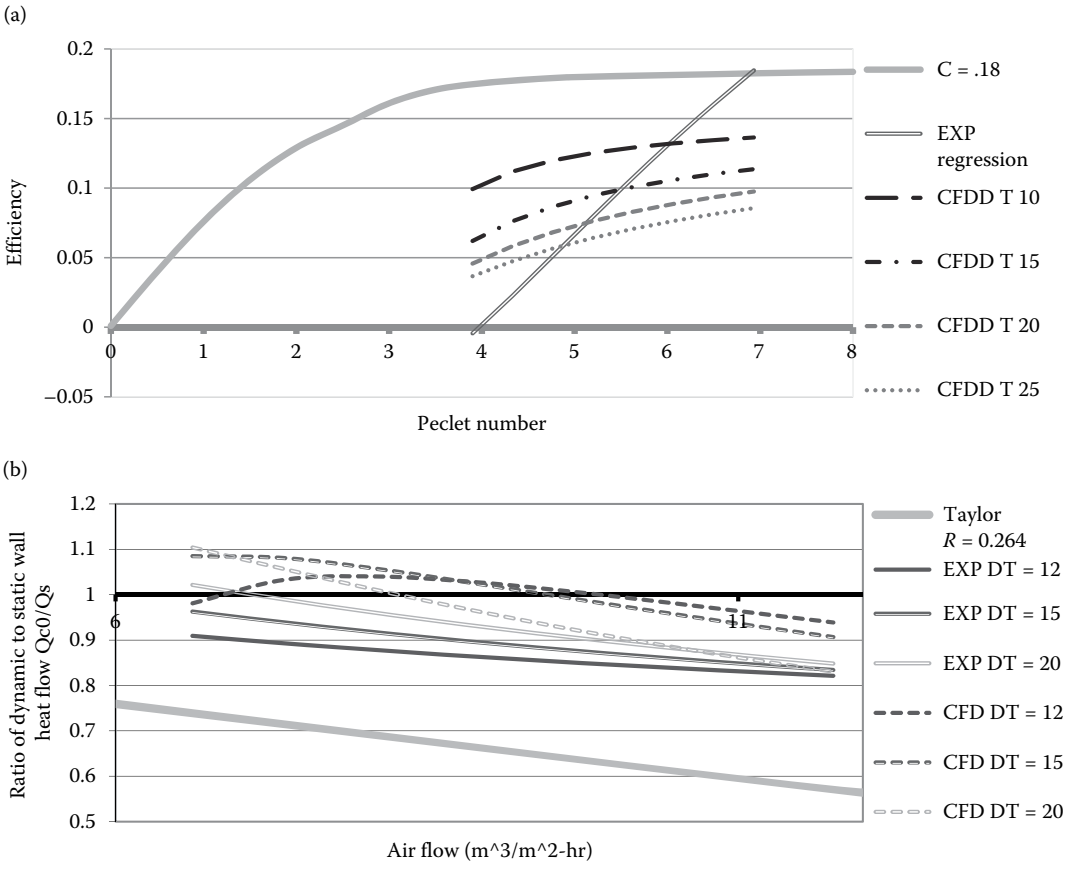


FIGURE 4.8.14 Comparison of analytical efficiency of Krarti and Taylor and CFD results with tested efficiency.

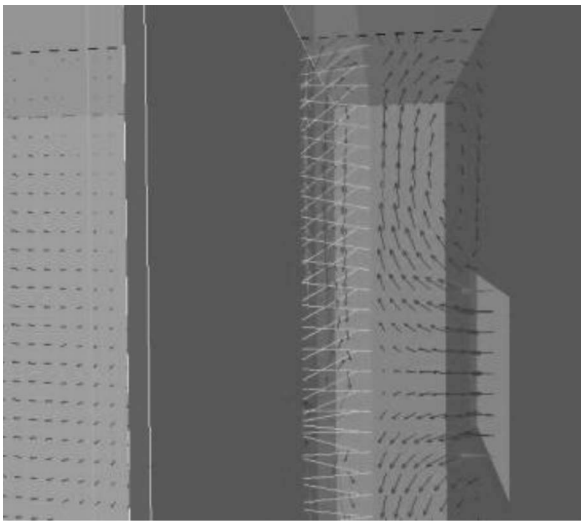


FIGURE 4.8.15 Airflow vectors within the breathing wall, indicating circulation loops. Color profiles are based upon air velocity.

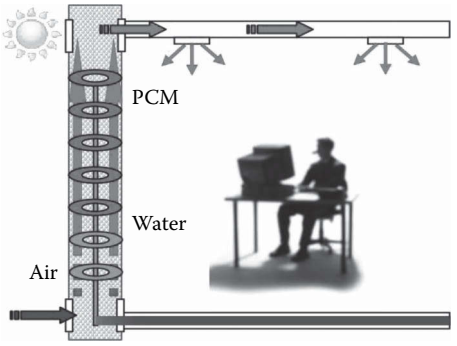


FIGURE 4.8.16 Illustration of a living wall system design; solar energy is harvested via a slab water radiation system and a buoyancy-driven overhead air distribution system.

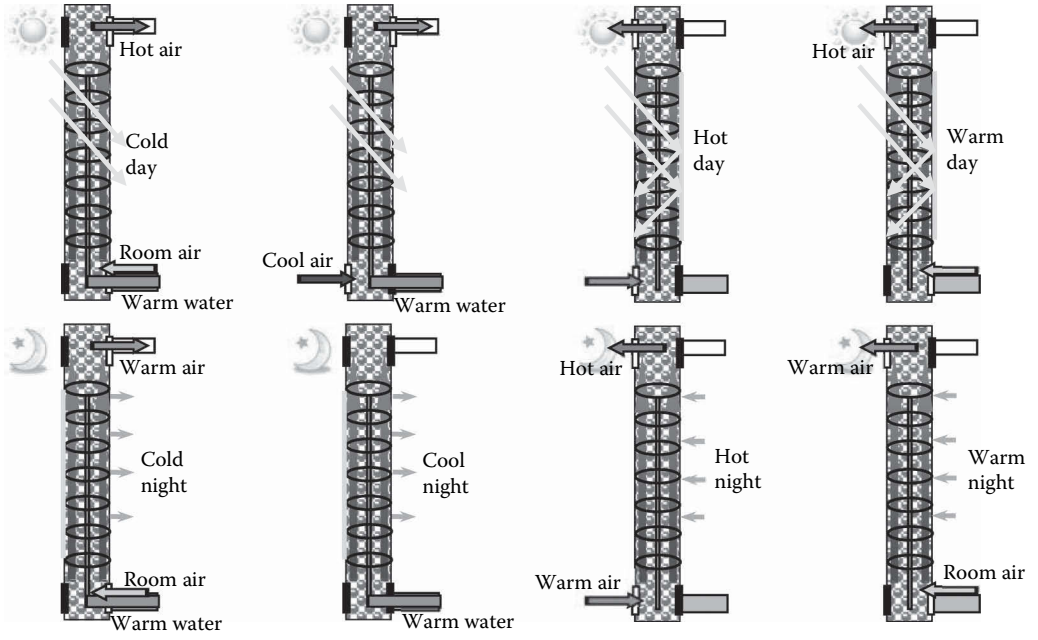


FIGURE 4.8.17 Illustration of different environmental and system operating conditions.

potential design of the living wall system for transferring and use of solar heat via a slab water radiation system and a buoyancy-driven overhead air distribution system. Proper design and control of the heat charge/discharge of PCM and circulation of gas and liquids in microvascular fluid systems can effectively manage the heat gain and loss through the structure, and maintain a pleasant indoor environment with comfortable temperature/relative humidity and adequate fresh air under dynamic climate conditions. This naturally controlled system will not only significantly reduce the total energy cost required to heat, cool, vent, and light the spaces, but also create attractive architectural features. Figure 4.8.17 illustrates different environmental and system operating conditions.

Thermal Models for Advanced Façade Systems

Modeling heat transfer processes of the air, liquid and PCM systems and exploring the thermal coupling fundamentals between various heat sources and sinks are the most important but challenging work. This requires establishing more sophisticated fluid dynamics and heat transfer models that can

account for the coupled multiphysics and multiphase fluid flow and heat transfer processes with connection to both indoor and outdoor environments. A generic program for studying advanced façade systems (named "AdvFacSy") has developed and validated for this purpose using Simulink environment. The program includes (1) weather data reader; (2) solar radiation model; and (3) models for different wall layers (e.g., solid, cavity, glass, PCM, etc.) that can be used to form different wall configurations (AL Saadi and Zhai, 2015).

Using "AdvFacSy," a variety of wall designs can be easily generated and assessed under different climatic conditions. In addition to envelope designs, the program can be used to evaluate different operational strategies such as to charge and discharge energy stored in PCMs using the air in the cavity as a heat transfer medium. The air is driven by thermal buoyancy or mechanical fan. The air can be induced from the outdoor environment or indoor via air vents located at top and bottom of the façade system. The controlling mechanism facilitates the charging and discharging of PCMs in the wall unit. The recovered heat can then be directly used to meet the heat demand, or used to preheat fresh air before treated centrally in air conditioning system, or transported to other heat storage medium for later usage. The program has been well validated against experiments for both conventional multilayer walls and ventilated cavity walls (EL Mankibi et al., 2015).

Simulated Thermal Performance of Stand-Alone Advanced Multilayer Wall

Simulation Parameters The AdvFacSy tool can be used to study the performance of a stand-alone exterior envelope assuming a predefined indoor air temperature representing interior boundary conditions. As a demonstration, the heating season from January to the end of April and again from October to the end of December is modeled. During this period, the indoor air temperature is unchanged at 22°C. The cooling season is from May to the end of September where the cooling set point is constant at 24°C. The exterior boundary conditions are provided using the weather file in EnergyPlus format (EPW) for Golden, CO. The inside and outside convective heat transfer coefficients are important factors that may influence the thermal performance of the PCM-enhanced walls. For this study, the internal convective heat transfer coefficient is assumed to be 4.43 W/m^2 K based on the experimental results from Liu and Awbi for PCM wallboards under natural convection (Liu and Awbi, 2009). The outside convective heat transfer coefficient is based on ASHRAE winter design conditions (ANSI/ASHRAE, 2011).

The main objective of this analysis is to study the general behaviors and identify the overall trends of thermal performance of the PCM-enhanced ventilated walls under typical meteorological weather data. Four performance indicators have been selected to investigate the impact of PCM: peak heating load, annual heating load, peak cooling load, and annual cooling load. The peak heating load is the maximum instantaneous heat flux from indoor environment to the wall surface. The peak heating load is used for sizing heating systems. The annual heating load is the yearly summation of the instantaneous heat flux from the indoor to the wall's surface. The peak cooling load, on the other hand, is the maximum instantaneous heat flux from interior wall's surface to the indoor. The peak cooling load is used for sizing cooling systems. The annual cooling load is the yearly summation of the instantaneous heat flux from the wall's surface to the indoor environment.

For all the cases, the PCMs are assumed to be concentrated as a separate layer with an overall thickness of 12.5 mm. Such systems have been proved to offer higher potential for heating and cooling application when compared to the randomly distributed mixed system (Kim and Darkwa, 2003). A lightweight wall system for residential building is selected. The reference wall has been developed for Golden, Colorado climate based on the Building America House Simulation Protocols (Hendron and Engebrecht, 2010; Engebrecht-Metzger et al., 2012). This analysis can be used for new design or for retrofitting existing buildings.

A ventilated cavity wall design is proposed and tested as schematically shown in Figure 4.8.18, where the PCM is located to the interior side of the wall (the most efficient location as proved previously). The ventilated cavity wall design consists of several components: the main multilayer wall, a cavity, and a glazing, whereas the cavity and glazing characteristics are of top design interest. The

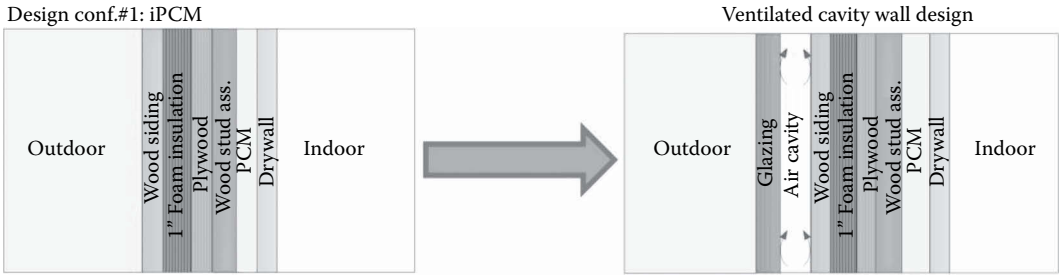


FIGURE 4.8.18 Multilayer wall configuration and the ventilated cavity wall design for harvesting solar for heating purpose.

cavity depth of 0.15 m and vent area of 1.05 m^2 are adopted. A typical double glazing is used and the air is naturally induced from the indoor environment in this design analysis. The rationale behind this design selection has been thoroughly investigated by Al-Saadi (2014). Upon the same design of the ventilated cavity, a base model without the PCM layer is simulated as a comparison. A series of simulations are performed for various PCM characteristics including the latent heat between 50 and 300 kJ/kg at an increment of 50, the melting temperature of 20°C–25°C at an increment of 1°C, and the melting range of 0.1°C, 1°C, 2°C, 4°C, and 6°C.

Simulation Results and Analysis Figure 4.8.19 shows the simulated results of all orientations across all the PCM parameters. The dotted line below all curves represents the contribution of the cavity when no PCM is used. It is apparent that the performance of PCM has been improved by introducing the cavity. There is a clear distinction in PCM's performance as the melting temperature varies. When no cavity is used, the maximum savings in annual heating load was found ~3%. When a cavity is introduced in the south wall, the maximum savings in annual heating loads reaches above 95% as shown in Figure 4.8.19a for the south wall case. Cavity contributes 65% of savings in annual heating loads and the rest (i.e., of more than 30%) is due to PCM layer.

The PCM contributions in west and east are lower than the south wall as shown in Figure 4.8.19c and g. The savings in annual heating loads due to PCM is around 20%, if cavity contribution due to conduction is subtracted. The cavity in North wall is not effective in improving the performance of PCM as shown in Figure 4.8.19e. When PCM is used, the savings in annual heating loads is less than the case when no PCM is used. This is likely due to the storage of PCM which is not utilized internally to reduce heating demand. In all the cases, the optimal melting temperature is close to the heating set point of 22°C with a narrow melting range of 0.1°C. At melting temperature away from the heating set point, the PCMs with wide melting range give more savings.

The savings in peak heating load has improved when the cavity design is compared to the non-cavity design. For all orientations, the savings in peak heating loads are similar as shown in Figure 4.8.19b, d, f and h. For south multilayer PCM wall, a maximum savings in peak load achieved with the noncavity case is 20%. For the cavity design, the maximum savings in peak heating load is around 50%. When cavity contributions are ignored (i.e., 15.40%), 34.5% savings in peak heating load is due to the PCM layer only. This is a 14.5% improvement over the noncavity design. The savings in other orientations are also similar to the south wall. In all cases, the optimal melting temperature is below 20°C with a narrow melting range of 0.1°C.

A clear distinction between PCMs with different melting range is observed. Those with narrow melting range achieve their maximum savings in annual heating loads at melting temperature close to the heating set point, following a triangular shape. This group then tends to give low savings as the melting temperature moves away from the set point. The savings of the wide melting range is flat and therefore are less sensitive to the heating set point. This group can be considered conservative. When the cavity contributions are subtracted, the cavity enhances the performance of interior PCM from 3% with no cavity to 30% for south and 20% for east and west. The cavity on north did

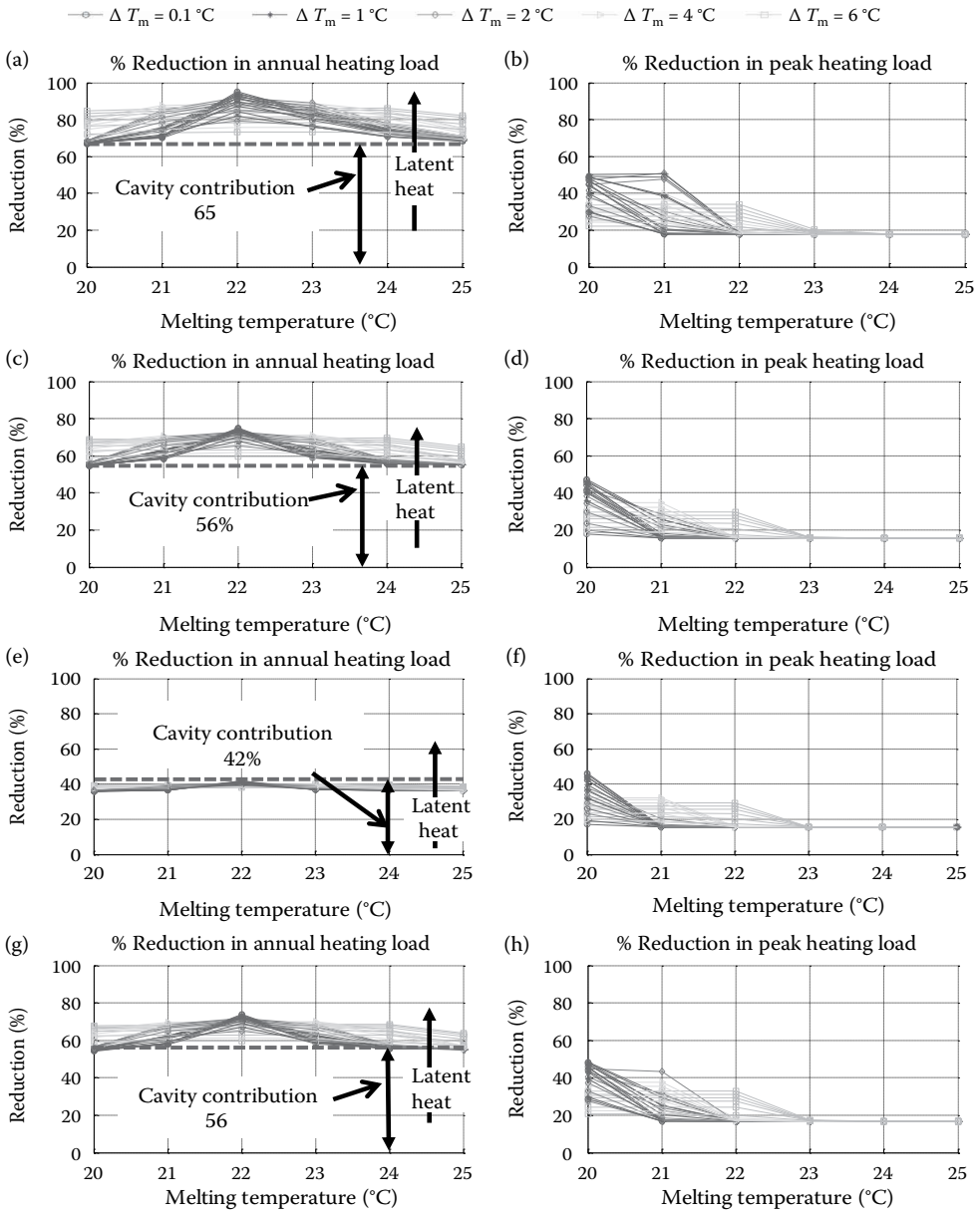


FIGURE 4.8.19 Percentage reduction in heating loads across all PCM's thermal parameters for ventilated cavity PCM-enhanced wall under Golden CO (a) % reduction in annual heating load for south wall, (b) % reduction in peak heating load for south wall, (c) % reduction in annual heating load for west wall, (d) % reduction in peak heating load for west wall, (e) % reduction in annual heating load for north wall, (f) % reduction in peak heating load for north wall, (g) % reduction in annual heating load for east wall, and (h) % reduction in peak heating load for east wall.

not improve the performance of PCM. In all cases, the optimal melting temperature is close to the heating set point of 22°C with a narrow melting range of 0.1°C.

Simulated Thermal Performance of Whole Buildings with Advanced Multilayer Walls

The developed and validated AdvFacSy model was integrated into a full-scale building simulation model, the HYBCELL 1.2 (El Mankibi et al., 2006). HYBCELL 1.2 was developed under

MATLAB/Simulink environment by coupling a thermal model and a pressure airflow model. The building thermal model can be divided into three parts: the envelope model based on the finite volume method, the short-wavelength radiation model based on an accurate description of the sunspot inside the simulated building, and the long-wavelength radiation model using the radiosity method. The airflow model is a pressure model that takes into account the effect of the wind and the thermal buoyancy of air to calculate the airflow through the openings of the building. Indoor air temperature is calculated considering several thermal evolution phenomena, such as heat transfer through the walls, air infiltration and ventilation, internal heat gains, and auxiliary heating or cooling. Schedule and occupation patterns can also be integrated into the model.

The tested building is an 80 m^2 single zone dwelling having a rectangular shape (10 m long, 8 m wide, and 3 m high). The studied multilayer wall represents the south facade. The properties of the walls (refer to Table 4.8.4) and the windows (illustrated by Figure 4.8.20) were defined to represent several configurations (i.e., high and low inertia, large and small glazing area). Table 4.8.4 presents material layer thicknesses adopted in the case of a building envelop with high inertia. In order to simulation building envelop with low inertia, the internal layer (Layer 1) of vertical walls and Layer 2 of roof and floor has been divided by 10 (0.02 m instead of 0.2 m). The building configurations

TABLE 4.8.4
Properties for Building Materials in Multilayer Walls

Wall	Thickness (m)		Density (kg/m^3)	Specific Heat (J/kg K)	Conductivity (W/m K)
	Inertia	Less Inertia			
Vertical walls (except south)	0.2	0.02	2500	800	1.5
	0.2	0.2	30	700	0.04
	0.1	0.1	1500	800	0.2
Floor and roof	0.01	0.01	1500	700	1
	0.2	0.02	2500	700	1.5
South wall	Studied multilayer wall				

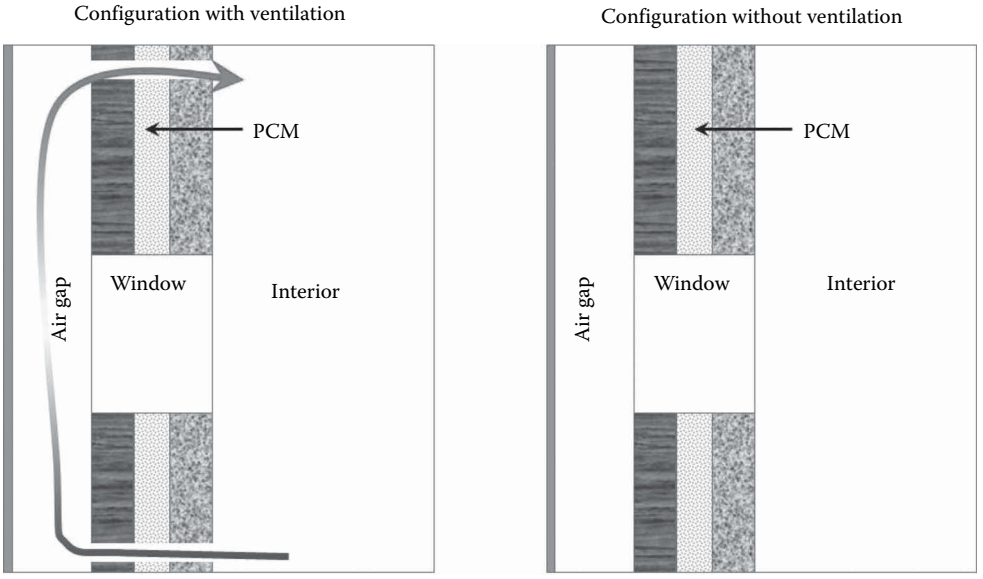


FIGURE 4.8.20 Multilayer wall configurations coupled to a full-scale building.

with large glazing have two windows in each façade except north façade (Figure 4.8.20) with a percentage of 26.39% of glazed area, and the building configuration with less glazed area (13.19%) has only one window per façade except north façade.

The tested multilayer wall is composed of three layers (external sensible storage layer, intermediate latent heat storage layer—PCM, and internal sensible storage layer). The goal is to investigate and optimize the thickness, density, and conductivity of the three layers and the melting temperature range and latent heat of PCM layer. To perform the optimization, NSGA-II (Deb, 2015), a nondominated sorting genetic algorithm (NSGA) tool coupled to Simulink model has been used. The ranking criteria (to minimize) are: heating energy consumption (kWh) and thermal discomfort risk (°C-hours) related to the number of hours when indoor air temperature is above 25°C. The energy consumption is evaluated calculating the injected heat (convective heating) at each step time in order to keep indoor air temperature above the set point. A proportional–integral–derivative (PID) controller maintains the indoor air temperature above 18°C and a shading system is activated to keep indoor air temperature under 26°C. Thermal discomfort risk is calculated using the following equation:

$$\text{Discomfort risk} = \int_{0}^{\text{Simulation time}} \left(\frac{\max\left(0; \left(T_{\text{in}}(t) - 26\right)\right) \cdot \mathrm{d}t}{3600} \right) \tag{4.8.15}$$

where $T_{\text{in}}(t)$ is the indoor air temperature at step time t(°C), dt is the step time (s). Discomfort risk is related to the inertia and greenhouse effect of the multilayer wall cavity. Thus, there is a risk to overheat the building even if the heating system is switched off. If this risk exists in winter conditions it will be higher in summer conditions.

The number of NSGA generations considered is 20 generations with 50 as the size of population. The simulated period is the first week of January. Figures 4.8.21 and 4.8.22 show the examples of

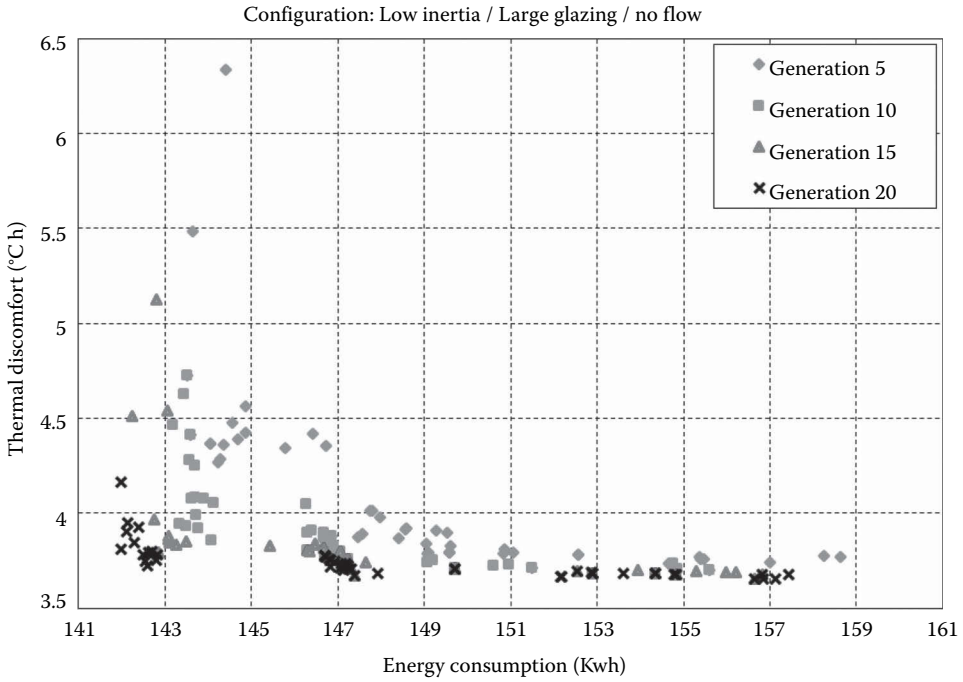


FIGURE 4.8.21 The 5th, 10th, 15th, and 20th generations from GA-based optimization results for low inertia building configuration.

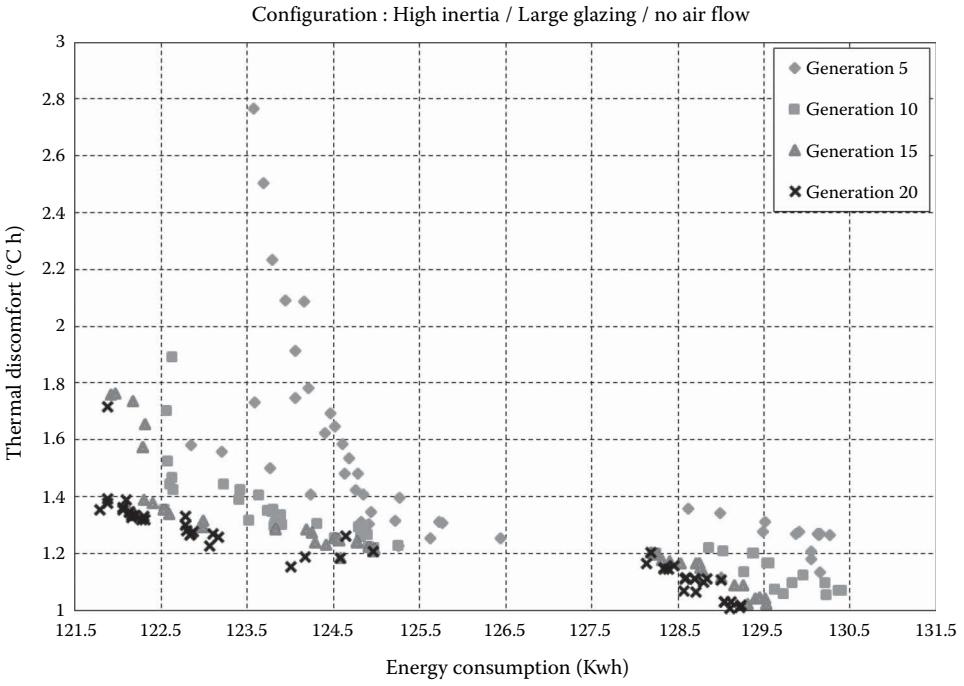


FIGURE 4.8.22 The 5th, 10th, 15th, and 20th generations from GA-based optimization results for high inertia building configuration.

optimization results for the building with high inertia configurations. Each configuration optimization results are shown at the 5th, 10th, 15th, and 20th generations.

Table 4.8.5 summarizes the optimal multilayer wall materials and thicknesses for each tested configuration. Table 4.8.6 shows the energy and thermal discomfort reduction related to the selected optimal multilayer wall configurations shown in Table 4.8.5. The optimal solutions (multilayer wall configurations) are compared to the reference case results. The reference case is a building that has the same characteristics except the fact that the south façade is constituted by a classical wall like all other vertical walls. Thus, there are four reference case studies (with/without inertia and with/without large glazing).

The results show that the performance of multilayer system improves the performance of the reference case study by 28% (building with less inertia configuration) and up to 38% (building with high inertia configuration) in terms of energy consumption. The multilayer configuration without cavity ventilation is thermally more efficient but induces high risk of thermal discomfort. Thus, control strategies of shading and outdoor air circulation should be provided to avoid this risk.

It was found that the PCM layer thickness varies between 1 and 4 cm depending on the whole building inertia. The external and internal layers have to be conductive with high thermal inertia. The thickness depends on the ventilation configuration and the whole building inertia. The internal layer is found to be more sensitive to these two parameters. The same conclusion applies to the glazing ratio and ventilation rate.

Summary and Conclusions

In this chapter, innovative and advanced building envelope systems have been described and their energy performance documented. These advanced wall systems can improve the building energy and thermal performance when designed properly. However and in order to optimize their design and operation, several factors should be considered. Model-based optimal controls may have to be

TABLE 4.8.5
Identified Best Configurations (Inputs)

Building Configuration			Living Wall Layers Thickness (cm)			Sensible Storage		Sensible Storage		Phase Change Layer		
Inertia	Glazing	Cavity Air Flow	External	PCM	Internal	External Layer	Internal Layer	External Layer	Internal Layer	Melting Temperature	Melting Range	Latent Heat
Low	Small	No	6	1	10	1.47	1.46	2500	2500	20.84	3.15	27607
Low	Small	Yes	6	3	12	1.07	1.40	985	2167	21.00	1.00	89705
Low	Large	No	7	4	3	1.50	1.43	2406	2398	19.46	2.31	37695
Low	Large	Yes	11	2	6	1.50	0.96	2419	2033	18.73	3.09	51661
High	Small	No	1	4	2	0.92	1.05	1288	1935	21.01	1.00	89766
High	Small	Yes	18	16	2	0.07	1.14	30	2092	21.92	1.00	70223
High	Large	No	9	2	1	1.45	1.05	2162	1648	20.03	1.94	56357
High	Large	Yes	1	3	13	0.81	1.26	1576	1761	20.52	1.51	88046

TABLE 4.8.6
Predicted Best Configurations (Outputs)

Building Configuration			Building with Living Wall Performances		Reference without Living Wall	
Inertia	**Glazing**	**Cavity Air Flow**	**Energy (kWh)**	**Discomfort (°C-h)**	**Energy (kWh)**	**Discomfort (°C-h)**
Low	Small	No	99.09	5.01	138.50	0
Low	Small	Yes	98.47	5.36		
Low	Large	No	141.97	3.81	203.80	3.16
Low	Large	Yes	142.92	4.50		
High	Small	No	87.21	3.87	134.00	0
High	Small	Yes	83.21	4.81		
High	Large	No	121.78	1.35	176.30	0
High	Large	Yes	121.87	4.40		

developed for these systems. A brief summary of the main conclusions and findings outlined for each technology is provided below:

- DIMs have the potential to reduce heating and cooling loads in all three climates compared to statically high insulated walls. In particular, it is found that the two-step DIMs result in higher cooling energy savings in heating dominant locations (i.e., Golden, CO and Madison, WI). Without any internal loads, the maximum cooling energy savings compared to high insulated wall are 32%–39%, respectively, with two-step-RSI-0.5/ RSI-2.5 in heating dominant locations. Moreover, it is found that the effectiveness of DIMs in reducing cooling energy use increase with lower internal heat gains and minimal window size on the wall.
- *Variable Reflectivity Coatings or VRCs* have been shown to offer an effective alternative to static cool roofs to reduce building cooling loads compared without increasing the heating loads. Specifically, it is found that the penalty of increased heating energy consumption associated with static cool roofs is avoided when VRCs are utilized. In particular, the dynamic cool roofs can save up to 15% and 10% in heating energy use compared to static cool roofs for an office building located in, respectively, Chicago, IL, and Golden, CO.
- *Breathing Walls* have a potential to save energy under varying environmental conditions. Both experimental and simulation studies prove that breathing walls can save energy during winter conditions above a certain airflow rate, and their efficiencies, defined by three different metrics, tend to increase with airflow rates. Additionally, in most of the cases, efficiencies tend to show a slight negative correlation with temperature differential, indicating that convection might play a greater role in breathing walls with cavities when the temperature differential is greater. Specifically, the potential energy savings from breathing walls have been shown to range between −10% and +30%, depending upon the airflow rate and temperature differential between indoors and outdoors.
- *Multilayer Living Walls* can offer adaptive envelope systems to changing outdoor conditions. Preliminary analysis results indicate that thermal performance of living walls is enhanced with a PCM as part of the interior layer. For the multilayer PCM-enhanced wall design, a narrow melting range coupled with optimal melting temperature close to the heating and cooling set points achieve maximum savings in annual and peak heating and cooling loads. The full-scale building model results show that when optimized, the living wall system can allow 27%–38% of reduced heating energy consumption while avoiding thermal discomfort.

REFERENCES

Abadie, M.O., Finlayson, E.U., Gadgil, A.J. Infiltration heat recovery in building walls: computational fluid dynamics investigations results. Lawrence Berkeley Laboratory Report. LBNL-51324, August 2002.

Akbari, H., Levinson, R. Evolution of cool-roof standards in the US. *Advances in Building Energy Research* 2–1 (2008), 1–32.

Akbari, H., Pomerantz, M., Taha, H. Cool surfaces and shade trees to reduce energy use and improve air quality in urban areas. *Solar Energy* 70(3) (2001), 295–310.

Al-Saadi, S.N. Modeling and simulation of PCM-enhanced facade systems. Ph.D., University of Colorado at Boulder, 2014.

AL Saadi, S.N., Zhai, Z. Systematic evaluation of mathematical methods and numerical schemes for modeling PCM-enhanced building enclosure. *Energy and Buildings* 92 (2015), 374–388.

ANSI/ASHRAE Standard 140–2011. *Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs*. Atlanta, GA: ASHRAE, 2011.

Bailly, N.R. Dynamic insulation systems and energy conservation in buildings. *ASHRAE Transactions* 91(1) (1985), 447–466.

Baker, P.H. The thermal performance of a prototype dynamically insulated wall. *Building Services Engineering Research and Technology* 24(1) (2003), 25–34.

Barhoun, H., Guarracino, G. Evaluating the energy impact of air infiltration through walls with a coupled heat and mass transfer method. *Proceedings of the CIB Conference*, Toronto, ON, May 1–7, 2004.

Bhattacharyya, S., Claridge, D.E. Energy impact of air leakage through insulated walls. *Journal of Solar Energy Engineering* 117(3) (1995), 167–172.

Buchanan, C.R., Sherman, M.H. A mathematical model for infiltration heat recovery. Lawrence Berkeley Laboratory Report. LBL-44294, 2000.

Chen, W., Liu, W. Numerical analysis of heat transfer in a composite wall solar-collector system with a porous absorber. *Applied Energy* 78(2) (2004), 137–149.

Deb, K. Kanpur Genetic Algorithms Laboratory. NSGA-II source code, http://www.iitk.ac.in/kangal/codes .shtml, Accessed June 06, 2015.

Dimoudi, A., Androutsopoloulus, A., Lykoulis, S. Experimental work on a linked, dynamic and ventilated wall component. *Energy and Buildings* 36(5) (2004), 443–453.

El Mankibi, M., Cron, F., Michel, P., Inard, C. Prediction of hybrid ventilation performance using two simulation tools. *Solar Energy* 80(8) (2006), 908–926.

EL Mankibi, M., Zhai, Z., Al-Saadi, S.N., Zoubir, A. Numerical modeling of thermal behaviors of active multi-layer living wall. *Energy and Buildings* 106 (2015), 96–110.

Engebrecht-Metzger, C., Wilson, E., Horowitz, S. Addendum to the Building America house simulation protocols, 2012.

Gao, Y., Xu, J., Yang, S., Tang, X., Zhou, Q., Ge, J., Xu, T., Lenvinson, R. Cool roofs in China: Policy review, building simulations, and proof-of-concept experiments. *Energy Policy* 74 (2014), 190–214.

Græe, T. *Breathing Bulding Construction*. Stillwater, OK: ASEA Stillwater, 1974.

Hendron, R., Engebrecht, C. Building America house simulation protocols, (Revised), 2010.

Imbabi, M.S., Peacock, A.D. Smart breathing walls for integrated ventilation, heat exchange, energy efficiency and air filtration. *ASHRAE/CIBSE Conference*, September 24–26, 2003.

Jelle, B.P. Traditional, state-of-the-art and future thermal building insulation materials and solutions–Properties, requirements and possibilities. *Energy and Buildings* 43(10) (2011), 2549–2563.

Jelle, B.P., Gustavsen, A., Baetens, R. The path to the high performance thermal building insulation materials and solutions of tomorrow. *Journal of Building Physics* 34(2) (2010), 99–123,.

Kalnæs, S.E., Jelle, B.P. Vacuum insulation panel products: a state-of-the-art review and future research pathways. *Applied Energy* 116 (2014), 355–375.

Kamalisarvestani, M., Saidur, R., Mekhilef, S., Javadi, F.S. Performance, materials and coating technologies of thermochromic thin films on smart windows. *Renewable and Sustainable Energy Reviews* 26 (2013), 353–364.

Kim J.S., Darkwa, K. Simulation of an integrated PCM–wallboard system. *International Journal of Energy Research* 27(3) (2003), 215–223.

Kohonen, R., Ojanen, T. Non-steady state coupled diffusion and convection heat and mass transfer in porous media. *5th International Conference on Numerical Methods in Thermal Problems*, Montreal, QC, July 6–10, 1987.

Krarti, M. Effect of airflow on heat transfer in walls. *Journal of Solar Energy Engineering* 116(1) (1994), 35–42.

Levinson, R., Akbari, H., Konopacki, S., Bretz, S. Inclusion of cool roofs in nonresidential Title 24 prescriptive requirements. *Energy Policy* 33(2) (2005), 151–170.

Liu, H., Awbi, H.B. Performance of phase change material boards under natural convection. *Building and Environment* 44(9) (2009), 1788–1793.

Morrison, I.D., Karagiozis, A.N. Energy impact of dynamic wall ventilation. *Proceedings of the 18th Annual Energy Society of Canada*. Edmonton, AB, 1992.

Park, B., Krarti, M. Energy performance analysis of variable reflectivity envelope systems for commercial buildings. *Energy and Buildings*, 124 (2016), 88–98.

Park, B., Srubar, W.V., Krarti, M. Energy performance analysis of variable thermal resistance envelopes in residential buildings. *Energy and Buildings* 103 (2015), 317–325.

Pérez-Lombard, L., Ortiz, J., Pout, C. A review on buildings energy consumption information. *Energy and Buildings* 40(3) (2008), 394–398.

Qiu, K., Haghighat, F. Modeling the combined conduction-air infiltration through diffusive building envelope. *Energy and Buildings* 39 (2007), 1140–1150.

Sadineni, S.B., Madala, S., Boehm, R.F. Passive building energy savings: A review of building envelope components. *Renewable and Sustainable Energy Reviews* 15(8) (2011), 3617–3631.

Taylor, B.J., Cawthorne, D.A., Imbabi, M.S. Analytical investigation of the steady-state behavior of dynamic and diffusive building envelopes. *Building and Environment* 31(6) (1996), 519–525.

Tuhus-Dubrow, D., Krarti, M. Genetic-algorithm based approach to optimize building envelope design for residential buildings. *Building and Environment* 45(7) (2010), 1574–1581.

Zhai, Z., Slowinski, R. Investigation on thermal performance of ventilated wall in winter conditions. *International Symposium on Green Manufacturing and Applications (ISGMA 2013)*, June 25–29, Honolulu, HI, 2013.

4.9 USE OF PHASE CHANGE MATERIALS IN BUILDINGS

Jan Kośny and David W. Yarbrough

Introduction and Scope

Thermal storage applications in buildings play an important role in energy conservation. It is greatly assisted by the incorporation of latent heat storage (LHS) in various building components and mechanical systems. It has been observed that building components, which can store excess heat available due to the building's periodic power operation, can also reduce space-conditioning energy consumption. Research has demonstrated that phase change materials (PCMs) improve building energy performance by reducing peak-hour cooling loads and by shifting the peak-demand time. PCMs have been tested as a thermal mass component in buildings for at least 40 years. A key goal in low-energy building research is to find ways to manage differences in time between energy sources and consumption (i.e., building loads). This chapter provides general information about PCM use, discusses the performance of the PCM-enhanced building envelopes, and presents the active space-conditioning systems using PCMs. It reviews the most common PCMs suitable for building applications, and discusses PCM encapsulation and packaging. In addition, this chapter examines the performance of a number of PCM-enhanced building elements and whole-building effects. Laboratory and field experiments used for evaluations of PCM-enhanced building components as well as mechanical systems using PCMs are described. A wide range of commercially available PCM products are identified. Finally, references are given for determining dynamic thermal characteristics of buildings utilizing PCM technologies and numerical tools which can be used to study PCM-enhanced building components.

Properties of Commonly Used PCMs

PCMs are substances with a high heat of fusion which are melting and freezing at a certain temperature or temperature range. They are capable of storing and releasing large amounts of energy. In PCMs, heat is absorbed or released when the material changes from solid to liquid and vice versa; thus, PCMs are classified as LHS materials. The material's ability to store the latent heat is a consequence of different types of phase transitions such as solid–solid, solid–liquid, solid–gas, or liquid–gas phase changes. Of these, however, the solid–liquid transitions are the most commonly used for LHS purposes. A great amount of research data is available today for many different types of PCMs, which are currently available as commercial products for a wide range of phase change temperatures—Farid et al. (2004). For example, Zalba et al. (2003) and Lane (1978, 1980, 1983, 1989) identified over 200 organic substances, inorganic substances, organic eutectics, inorganic eutectics, fatty acids, and noneutectic mixtures with potential for use as a PCM and which are suitable for temperature control applications.

The best-known phase-changing substances, which are commonly used today in technical applications are: paraffins (organic), fatty acids (organic), and salt hydrates (inorganic)—as listed in IEA Task 32, (2005). Different blends of useful eutectic PCMs have also been identified. Figure 4.9.1 depicts phase change enthalpies (Δh, kJ/L) as a function of phase change temperatures (°C) for 12 most common types of PCMs. Phase-changing substances used in commercial applications need to have a possibly high energy of phase transition and, at the same time, they should exhibit melting points that are characteristic for the specific usage. It is expected that PCMs used for building applications should be able to store and release significant amounts of energy on a daily basis for a minimum of 25 years (the common estimate for lifespan of the building components). In addition, this amount of latent heat should be significant compared to the sensible heat gains/losses characteristic for typical building materials.

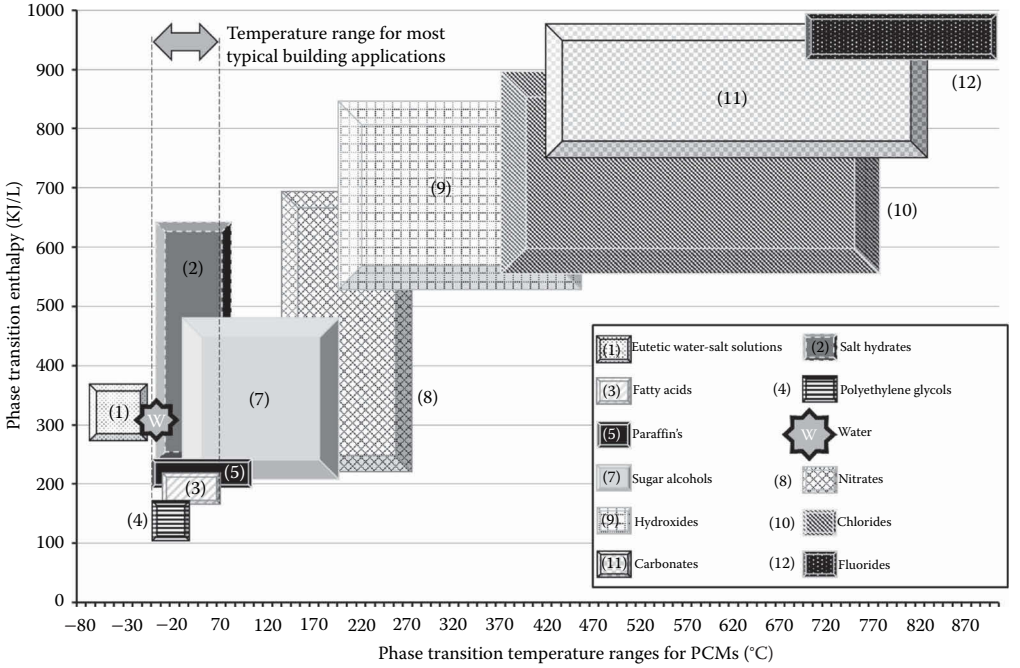


FIGURE 4.9.1 Phase change enthalpy values and transition temperatures for various PCMs.

The majority of PCMs which can be utilized for building applications have Δh values between 100 and 250 kJ/kg. These PCMs are usually 2–10 times lighter than structural materials. They can also store up to 15 times more heat per volume unit, when compared to construction materials such as masonry, concrete, or stone, for a 10 K temperature change. PCM-enhanced building components are often referred to as a lightweight thermal mass. As discussed in Lingayat and Suple (2013), Mehling and Cabeza (2008), and Kośny (2015), numerous differences exist between organic and inorganic PCMs. This includes operating temperature range, phase transition enthalpy, chemical stability, long-term durability, and flammability. These characteristics are especially important for their application in building structures, since the PCM has to be incorporated into the architectural components, finish materials, or thermal insulation.

The category of organic PCMs includes paraffins, fatty acids, fatty acid esters, and alcohols. The key advantages of organic PCMs are their capability to melt and freeze repeatedly without loss of heat storage capability and a fact that they can crystallize with little or no supercooling. Organic PCMs are also noncorrosive. Unfortunately, some organic PCMs exhibit a relatively low fire resistance and readily burn. To enable commercial applications, organic PCMs are often microencapsulated. Sometimes, they are directly blended with carrier materials, or enclosed in plastic containers or foils, arranged in arrays of small pouches. Sometimes, cellular polymers are utilized as carriers for organic PCM. Such process is called shape stabilization—see Inaba and Tu (1997) and Sari (2004). In shape-stabilized PCM products, PCM conforms to the shape of the solid structure during phase transition (Xu et al., 2005). Table 4.9.1 lists the basic physical properties of paraffins having carbon numbers between 14 and 33. They have melting points between 6°C and 71°C, which is within the expected range for building applications.

Salt hydrates, salt mixtures, and metallic materials are frequently utilized for heat storage applications. Of these, hydrated salts are the most commonly used inorganic PCMs because of their low cost. Compared to organic PCMs, they have similar values of Δh (mass basis), but higher values per volume due to their higher density—Kośny (2015). They are also nonflammable and only slightly

TABLE 4.9.1
Thermal Properties of *n*-Paraffins with Carbon Numbers between 14 and 33

Name	Chemical Symbol	Melting Point[a]	Heat of Fusion[b]	Mol. Weight[a]
Units		°C	kJ/kg	g/mol
n-Tetradecane	$C_{14}H_{30}$	6	228	198.39
n-Pentadecane	$C_{15}H_{32}$	10	205	212.42
n-Hexadecane	$C_{16}H_{34}$	18	237	226.44
n-Heptadecane	$C_{17}H_{36}$	22	213	240.47
n-Octadecane	$C_{18}H_{38}$	28	244	254.5
n-Nonadecane	$C_{19}H_{40}$	32	222	268.53
n-Eicosane	$C_{20}H_{42}$	37	246	282.55
n-Henicosane	$C_{21}H_{44}$	40.5	200	296.58
n-Docosane	$C_{22}H_{46}$	44.5	249	310.61
n-Tricosane	$C_{23}H_{48}$	47.5	232	324.63
n-Tetracosane	$C_{24}H_{50}$	52	255	338.66
n-Pentacosane	$C_{25}H_{52}$	54	238	352.69
n-Hexacosane	$C_{26}H_{54}$	56.5	256	366.71
n-Heptacosane	$C_{27}H_{56}$	59	236	380.74
n-Octacosane	$C_{28}H_{58}$	64.5	253	394.77
n-Nonacosane	$C_{29}H_{60}$	65	240	408.79
n-Triacontane	$C_{30}H_{62}$	66	251	422.82
n-Hentriacontane	$C_{31}H_{64}$	67	242	436.85
n-Dotriacontane	$C_{32}H_{66}$	69	170	450.87
n-Triatriacontane	$C_{33}H_{68}$	71	268	464.9

[a] Becker (1997).
[b] Sharma et al. (2009).

toxic (Sharma et al., 2009). Historically, inorganic PCMs were the first LHS materials used in buildings (Telkes, 1952, 1978). These materials were also often installed in solar walls containing massive metal containers. Table 4.9.2 lists the basic phase change properties of selected inorganic PCMs of known phase transition characteristics. In general, for inorganic salt hydrates, densities of solid and liquid phases vary from 1400 to 2000 kg/m^3 (for compounds with melting point suitable for building envelope applications). Hydrated salts are readily available at a low cost. They have a phase change Δh of around 200 kJ/kg and a reasonably high thermal conductivity of about 0.5 W/(m·K). Their thermal conductivity is often at least twice as high as that of organic materials.

The phase transition process for some organic PCMs that occurs at the melting and freezing temperatures involves hydration and dehydration. If water released during the melting process cannot dissolve the crystalline structure of the PCM and there is a significant difference in densities between the solid and liquid phases, then segregation occurs and the heavy component settles on the bottom of the container. Similarly, during the solidification process, poor nucleation of the hydrated salts can result in supercooling and a reduction in the overall performance (Hawes et al., 1993). If an inorganic PCM is not properly formulated, then the situation may become irreversible and the long-term system performance is affected (Farid et al., 2004; IEA, 2005). Building applications are impacted by material compatibility issues that can result in corrosion (IEA, 2005; Mehling and Cabeza, 2008). There are a significant number of other inorganic compounds with desired LHS properties that need to be investigated (IEA, 2005). Since salt hydrates are often not stable at all humidities, they need to be protected from the environment and, for example, loss of water. Plastic pouches or containers are used to protect PCMs from the environment. In many cases, salt hydrates are compatible with plastics.

TABLE 4.9.2
Solid-Liquid Phase Change Properties for Selected Inorganic PCMs

Name	Chemical Formula	Melting Point	Heat of Fusion	References
Units		**°C**	**kJ/kg**	
Dipotassium phosphate hexahydrate	$K_2HPO_4{\cdot}6H_2O$	14	109	Sharma et al. (2009)
Potassium fluoride tetrahydrate	$KF{\cdot}4H_2O$	18.5	231	Abhat (1983), Heckenkamp and Baumann (1997), Zalba et al. (2003), and Khudhair and Farid (2004)
Ferric bromide hexahydrate	$FeBr_3{\cdot}6H_2O$	21	105	Sharma et al. (2009)
Manganese nitrate hexahydrate	$Mn\ (NO_3)_2{\cdot}6H_2O$	25.8	125.9	Nagano et al. (2000)
Calcium chloride hexahydrate	$CaCl_2{\cdot}6H_2O$	29–30	171, 190.8	Abhat (1983), Heckenkamp and Baumann (1997), and Khudhair and Farid (2004)
Lithium nitrate trihydrate	$Li\ NO_3{\cdot}3H_2O$	30	296	Heckenkamp and Baumann (1997)
Calcium bromide hexahydrate	$CaBr_2{\cdot}6H_2O$	30.2	115.5	Lane (1980) and Dincer and Rosen (2002)
Sodium sulfate decahydrate	$Na_2SO_4.10H_2O$	32.4	254	Telkes (1975) and Dincer and Rosen (2002)
Yavapaiite dodecahydrate	$KFe(SO_4)_2{\cdot}12H_2O$	33	173	Sharma et al. (2009)
Lithium bromide dihydrate	$LiBr_2{\cdot}2H_2O$	34	124	Sharma et al. (2009)
Sodium carbonate decahydrate	$Na_2CO_3{\cdot}10H_2O$	34.2	146.9	Telkes (1975) and Abhat (1983)
Sodium orthophosphate dodecahydrate	$Na_2HPO_4{\cdot}12H_2O$	35.5–44	265	Belton and Ajami (1973) and Telkes (1975)
Zinc nitrate hexahydrate	$Zn\ (NO_3)_2{\cdot}6H_2O$	36.2	246.5	Lane (1980) and Dincer and Rosen (2002)
Ferric chloride hexahydrate	$FeCl_3{\cdot}6H_2O$	37	223	Sharma et al. (2009)
Manganese nitrate tetrahydrate	$Mn(NO_3)_2{\cdot}4H_2O$	37.1	115	Sharma et al. (2009)
Cobalt sulfate heptahydrate	$CoSO_4{\cdot}7H_2O$	40.7	170	Sharma et al. (2009)
Potassium fluoride dihydrate	KF_2H_2O	42	162	Sharma et al. (2009)
Magnesium iodide octahydrate	$MgI_2{\cdot}8H_2O$	42	133	Sharma et al. (2009)
Calcium iodide hexahydrate	$CaI_2{\cdot}6H_2O$	42	162	Sharma et al. (2009)
Potassium orthophosphate heptahydrate	$K_2HPO_4{\cdot}7H_2O$	45	145	Sharma et al. (2009)
Zinc nitrate tetrahydrate	$Zn(NO_3)_2{\cdot}4H_2O$	45	110	Sharma et al. (2009)
Magnesium nitrate tetrahydrate	$Mg(NO_3){\cdot}4H_2O$	47	142	Sharma et al. (2009)
Calcium nitrate tetrahydrate	$Ca(NO_3){\cdot}4H_2O$	47	153	Sharma et al. (2009)
Ferric nitrate nonhydrate	$Fe(NO_3)_3{\cdot}9H_2O$	47	155	Sharma et al. (2009)
Sodium silicate tetrahydrate	$Na_2SiO_3{\cdot}4H_2O$	48	168	Sharma et al. (2009)
Potassium orthophosphate trihydrate	$K_2HPO_4{\cdot}3H_2O$	48	99	Sharma et al. (2009)
Sodium thiosulfate pentahydrate	$Na_2S_2O_3{\cdot}5H_2O$	48–55	187, 209	Telkes (1975) and Heckenkamp and Baumann (1997)

(*Continued*)

TABLE 4.9.2 (*Continued*)
Solid-Liquid Phase Change Properties for Selected Inorganic PCMs

Name	Chemical Formula	Melting Point	Heat of Fusion	References
Units				
Magnesium sulfate heptahydrate	$MgSO_4 \cdot 7H_2O$	48.5	202	Sharma et al. (2009)
Calcium nitrate trihydrate	$Ca(NO_3)_2 \cdot 3H_2O$	51	104	Sharma et al. (2009)
Zinc nitrate trihydrate	$Zn(NO_3)_2 \cdot 2H_2O$	55	68	Sharma et al. (2009)
Ferric chloride dihydrate	$FeCl_3 \cdot 2H_2O$	56	90	Sharma et al. (2009)
Nickel nitrate hexahydrate	$Ni(NO_3)_2 \cdot 6H_2O$	57	169	Sharma et al. (2009)
Sodium acetate trihydrate	$Na(CH_3COO) \cdot 3H_2O$	58	226, 264	Heckenkamp and Baumann (1997)
Magnesium chloride tetrahydrate	$MnCl_2 \cdot 4H_2O$	58	151, 178	Sharma et al. (2009)
Ferric nitrate hexahydrate	$Fe(NO_3)_2 \cdot 6H_2O$	60.5	126	Sharma et al. (2009)
Sodium aluminum sulfate decahydrate	$NaAl(SO_4)_2 \cdot 10H_2O$	61	181	Sharma et al. (2009)
Sodium phosphate dodecahydrate	$Na_3PO_4 \cdot 12H_2O$	65	190	Sharma et al. (2009)
Tetrasodium pyrophosphate decahydrate	$Na_2P_2O_7 \cdot 10H_2O$	70	184	Heckenkamp and Baumann (1997)
Lithium acetate dihydrate	$LiCH_3COO \cdot 2H_2O$	70	150	Sharma et al. (2009)

Utilization of PCMs in the Building Envelope and Interior Fabric

PCMs can be employed in building components for many different purposes, including reduction of space-conditioning energy consumption, thermal peak load shaving and shifting, or improvement of interior thermal comfort. In addition, PCMs can generate a shift of building thermal loads (Kośny and Kossecka, 2013; Kośny et al., 2014). This time-shift can be controlled by the form of PCM application, the PCM location in the case of condensed PCM applications, and the total thermal resistance (*R*-value) of the building component insulation. An approximate temperature range for PCMs used in buildings is from −5°C (building envelope surface temperature during winter) to 70°C (peak roof surface temperature during summer). For more traditional internal building applications, like wall, floor, or ceiling finish, PCM's phase transition temperature should be within the human comfort range of 20°C–30°C. A large variety of research studies demonstrated that an application of thermal mass in well-insulated structures could generate heating and cooling energy savings between 5% and 30% in residential buildings, e.g., see Kośny (2015). Reduction of the space-conditioning energy consumption can be achieved either by improvements in building thermal shell components (i.e., through increasing the thickness of the insulation), or by adding PCM to the walls, roof, attic floor insulation, floors and ceilings, and fenestration components. In early applications, PCMs were often incorporated into conventional construction materials including PCMs mixed with structural concrete, gypsum, or stucco. Manufacturing techniques have been developed for impregnating concrete, gypsum, or ceramic masonry with PCM. It is important to mention that a great variety of PCMs had been tried, including inorganic salt hydrates, organic fatty acids and eutectic mixtures, fatty alcohols, neopentyl glycol, and paraffinic hydrocarbons, e.g., see Feldman et al. (1989), Hawes et al. (1993), Demirbas (2006), and Khudhair and Farid (2007). However, today, after several decades of lab testing and full-scale trials, the most popular PCMs are microencapsulated paraffinic hydrocarbons (Kośny, 2015). Usually, an incorporation of PCMs into the building fabric components involves an application of relatively complex technological processes. It also requires detailed thermal analysis. Please note that a great majority of PCMs used in building envelope applications need to go through a complete phase transition during 24 h time

periods in order to be fully effective. It is crucial that the temperature in the location where PCM is installed cycles above and below the PCM functional temperature range.

In the 1980s, a wide application of organic paraffinic PCMs in gypsum boards took place based on pioneering work of Feldman and Shapiro (1989), Salyer and Sircar (1989), and Tomlinson et al. (1992). In Europe, similar work was initiated by the development of encapsulated paraffinic PCM products. Blends of microencapsulated PCMs with building structural and board materials became at that time a common practice in low energy buildings—see Banu et al. (1998), Hawlader et al. (2003), Zalba et al. (2003), Cabeza et al. (2007), and Kośny (2015). In addition, a variety of manufacturing techniques have been developed for concrete blocks and other ceramic building products containing PCM as discussed by Khudhair and Farid (2004) and Zalba et al. (2003). In the early 2000s, PCM-enhanced building thermal insulations based on incorporating microencapsulated PCM into different types of fiber insulations and plastic foams was introduced (Kośny et al., 2006, 2007a,b). Figure 4.9.2 shows electron microscope images of gypsum and thermal insulation containing microencapsulated paraffinic PCM. PCM-enhanced gypsum boards are still widely considered as the most popular building material using latent heat thermal storage.

Several generations of novel roof designs showing large reductions in peak loads and overall loads have been developed. Results of about 95% reduction of peak-hour cooling loads and over 30% overall cooling load reduction compared with traditional roofs design of the same insulation *R*-value have been reported (Kośny et al., 2010). Figure 4.9.3 shows examples of a standing seam metal roofing system containing arrays of PCM pouches, above sheathing ventilation, reflective insulation, and building integrated PV laminate (Kośny et al., 2012a). Wall applications such as

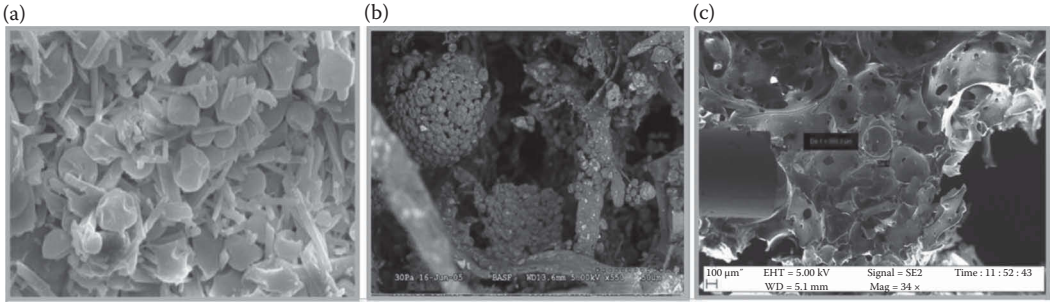


FIGURE 4.9.2 Electron microscope images of PCM-enhanced building materials: (a) PCM-enhanced gypsum composite; (b) cellulose-PCM blend—visible clusters of PCM pellets; and (c) polyurethane foam-PCM blend.

FIGURE 4.9.3 Standing seam metal roof system containing arrays of PCM pouches, above deck ventilation, reflective insulation, and roof-integrated PV laminate.

gypsum or composite boards containing PCM have been tested for the last 40 years, with most of the published research data showing that PCMs can notably enhance energy performance. North American studies (Tomlinson et al., 1992; Feustel, 1997; Kissock et al., 1998; Stovall and Tomlinson, 1995; Zhang et al., 2005; Kośny et al., 2006, 2010; Murugananthama et al., 2010) demonstrated that these applications could generate heating and cooling energy savings of up to 30% in well insulated residential buildings in the southern United States. Similar performance advantages can be found in many reports available worldwide, e.g., see Khudhair and Farid (2004), Zalba et al. (2003), and Kośny (2015).

The major goal of using PCM in fenestration systems is to reduce the building heat gain by absorbing the solar radiation before it reaches the indoor space. An addition of PCMs to the glazing is usually made with transparent plastic containers. PCM-enhanced windows typically consist of a thin layer of translucent PCM combined with a thermally insulating window located on the exterior side. In addition to the exterior thermal envelope applications, PCMs can be incorporated in building components such as internal walls, floors, and ceilings. In these applications, PCMs are usually used to mitigate the interior space temperature fluctuations (Kośny, 2015).

Distributed and Localized PCM Applications in Buildings

Current developments in the area of building technology suggest that both commercial and residential buildings will soon be routinely constructed with low heating and cooling loads. The use of novel building materials containing active thermal components such as PCM will be an important step toward achieving zero-energy buildings. The design of a localized PCM application includes specification of the PCM location in the building enclosure and specification of the thermal resistances bounding the PCM. The location specified must cycle across the phase change temperature for a large part of the year. The amount of PCM used is determined from estimates of the total heat flow across the enclosure during a cycle of operation. The selection of the PCM locations, PCM transition temperature range, and the amount of used PCM are essential for effective and durable use of the PCM-enhanced technologies, considering a relatively long lifespan of building envelopes.

Different configurations of PCM-enhanced gypsum or composite boards, as well as arrays of PCM containers, have been extensively tested as thermal mass components in buildings worldwide. These experiments demonstrated that PCMs can notably enhance energy performance of walls, roofs, and attics. Recent North American studies, performed by Feustel (1995), Athienitis et al. (1997), Kissock et al. (1998), Athienitis et al. (2005), Zhang et al. (2005), Kośny et al. (2006), Murugananthama et al. (2010), and Kośny et al. (2010), have proven that the application of localized PCMs could generate heating and cooling energy savings of up to 30% in well-insulated residential buildings in the southern United States. Childs and Stovall (2012) used numerical analysis to study the thermal performance of different configurations of layered wall cavity insulation using microencapsulated PCM and demonstrated savings.

Paraffinic PCMs are most widely used today as building LHS materials. Though, due to the relatively high cost of paraffin-based PCMs, their flammability, and their origin from petrochemicals, other organic PCMs are being explored today. This includes fatty acids, coming from agricultural and food industry waste. In buildings, interior surfaces such as walls, ceilings, or floors have been traditionally considered as the best locations for the PCM. In case where PCM-enhanced gypsum board or plasters are installed on the interior surfaces, they are expected to stabilize the temperature inside the building. As depicted in Figure 4.9.4, PCM concentrated in the gypsum boards interacts mostly with the building interior space. The combined latent and sensible heat storage capacity of the PCM-enhanced gypsum is used to reduce interior space temperature swings and absorb solar gains coming through the glazing. In this scenario, heat stored in PCM is often discharged by ventilation during the night or by programmed conditioning systems.

Similar to previously described concentrated PCM applications, PCMs can be also mixed with a variety of carrier materials such as concretes, gypsum, foam, or fibrous insulations to

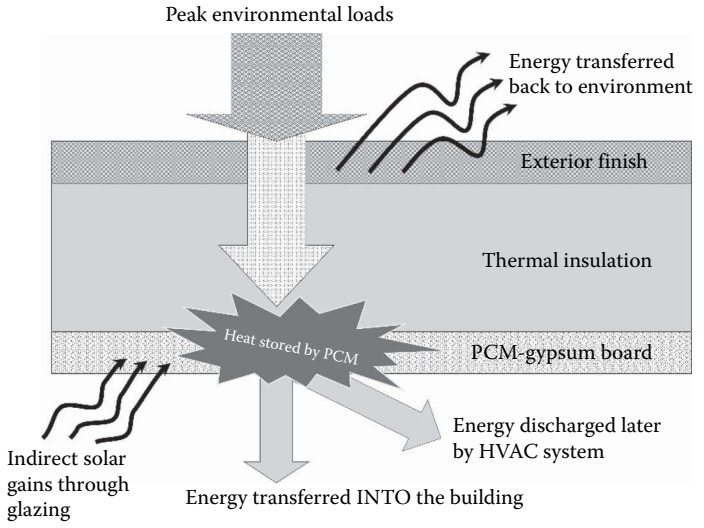


FIGURE 4.9.4 Concentrated PCM application as a part of the interior surface of the building envelope.

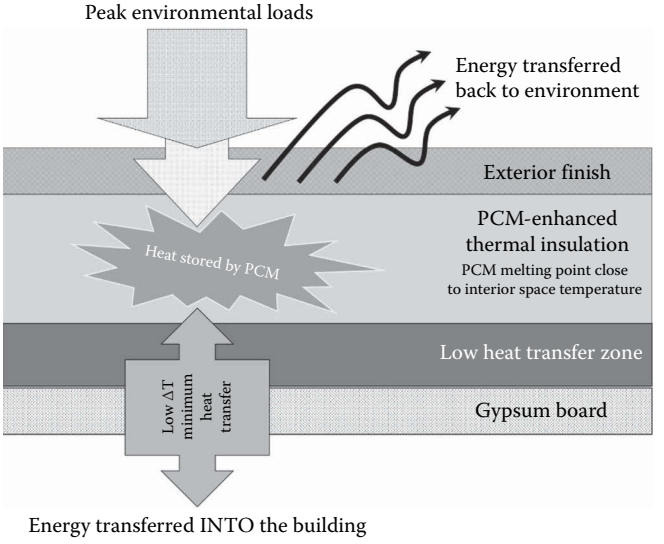


FIGURE 4.9.5 Schematic of a distributed PCM application.

allow distributed applications. As shown in Figure 4.9.5, an application placing PCM inside the wall cavity takes advantage of the large temperature fluctuations that take place on the exterior building envelope surfaces. Many of the North American experimental studies of dispersed PCM applications have been focused either on concrete masonry technologies (earlier works in 1980s), or on lightweight wood framed walls and attics, e.g., see Petrie et al. (1997), Kośny et al. (2006, 2007a,b), Tabares-Velasco (2012), and Kośny et al. (2012b). PCM-enhanced cellulose insulation has been used. Advantages of cellulose originate the fact that natural cellulose fibers can efficiently mix with microencapsulated PCM, while glass or composite fibers have difficulties in structural support of the weight of PCM micropellets. In early North American experiments, paraffinic PCM was utilized—Kośny et al. (2008) and it was later replaced by less flammable bio-based PCM fabricated from waste fatty oils.

Utilization of PCMs in Building Space-Conditioning and Water Heating Systems

A variety of technologies that couple heat pumps to LHS systems have been investigated. Several novel designs of energy storage exchangers utilizing PCMs for storing heat were introduced to ensure the reliable operation of a heat pump under various weather conditions and enhance the system performance at low ambient temperatures. Agyenim et al. (2010) used LHS for residential heating by using a PCM tank (paraffin) that absorbs and releases heat through internal water tubes. Similarly, Stritih (2003) designed two paraffin containers to store latent heat for residential applications. PCM is used in air ducts for LHS. Dolado et al. (2011) and Zukowski (2007) studied the heat transfer between moving air and different forms of PCM containers. These studies led to the conclusion that the air must be blown at high velocity which results in significant fan energy consumption, higher noise levels, and poorer exergy and energy efficiencies.

Domestic hot water tanks generally store heat in the sensible form. PCM heat storage systems have become popular recently due to their capacity of storing energy in high-density PCMs to achieve reduction of temperature stratification within the tank—see Yamaha et al. (2001), Najafiana et al. (2015), and Cabeza et al. (2006). Some advantages were observed, such as the capability of the PCM to reheat the amount of cold water surrounding the PCM after a partial unload. As expected, the PCM-water tank stores up to 8.1% more energy than the water tank, Cabeza et al. (2006).

Laboratory Evaluations of PCM Thermal Performance

It took several decades of intense research and technology development, for PCMs to become one of the most promising technical means of improving the energy efficiency and interior climate stability in residential and commercial buildings. The main benefit of using PCMs is the availability of significant amounts of latent heat during the phase transition processes. In most PCM building applications, the LHS is based on heat absorption or release during a change of phase from solid to liquid and vice versa. The heat storage capacity of materials containing PCMs for a temperature change from T_1 to T_2 that includes the phase change is determined as follows:

$$Q = \int_{T_1}^{T_2} m \cdot c_p \, dT + m \cdot f_m \cdot \Delta h \tag{4.9.1}$$

where f_m is the fraction of the material ongoing phase change, m is the mass, and c_p is the specific heat.

Today, most of commercially used PCM products, when used in small temperature intervals around their phase change temperature, show superior heat storage density (the amount of heat stored in a given material or system per unit volume or mass) compared to sensible heat storage capacity of conventional building materials, such for example, wood, concrete, or brick. Since there are several types of PCMs used in buildings, they also represent a variety of phase change thermal characteristics. That is why during the design process, the enthalpy curve $h(T)$ has to be known to quantify heat storage capacity. Figure 4.9.6 depicts the enthalpy curves for melting and freezing of an organic PCM. Keep in mind that an ideal PCM can be mathematically characterized by four scalar values, namely, the specific heat, c_p, of solid and liquid phases, the phase change enthalpy Δh, and the melting temperature T_m. Still, real phase change processes in real PCMs can be better represented by the function $h(T)$. The precise determination of this function is the goal of variety of available calorimetric measurement techniques. As shown in Figure 4.9.6, enthalpy curves are not always identical for melting and freezing processes. It is often due to the temperature hysteresis phenomenon, which is of course material specific, but it also strongly depends on the test method and equipment—see IEA (2011). In addition, the supercooling effect can be observed in inorganic PCMs. It is associated with impaired PCM nucleation during the freezing process. In many PCM heat storage applications, supercooling can be a source of thermal performance and durability problems.

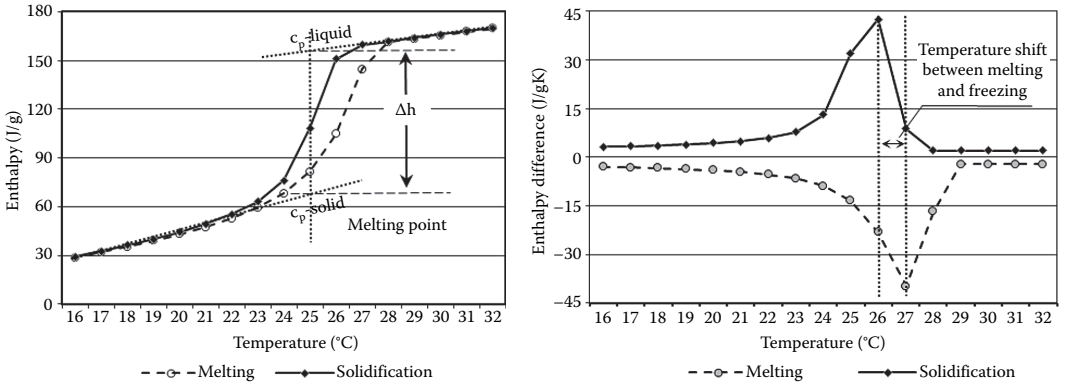


FIGURE 4.9.6 Examples of enthalpy curves for melting and freezing processes showing temperature hysteresis (left) and supercooling effect (right).

In laboratory work, PCM-enhanced building products are usually limited to analysis at the material and systems scale. There are many publications available, which are focused on theoretical and experimental analysis of phase change processes including hysteresis or supercooling phenomena—see Demirel and Paksoy (1993), Martín et al. (2003), Wunderlich (2005), Bony and Citherlet (2007), Mehling and Cabeza (2008), and Günther et al. (2007, 2009). However, the system scale hot-box testing and small building/room experiments can be utilized as well for characterization of building technologies containing PCM—see Kośny (2015).

For a long time, the only readily available material PCM evaluations used differential thermal analysis (DTA) or differential scanning calorimetry (DSC). These techniques were considered as standard tools for the measurement of latent heat of PCMs. These methods, however, require the use of small and relatively uniform material samples (Wunderlich, 2005). Unfortunately, the above requirement is rather unrealistic for most PCM-enhanced building components. Small specimens are not generally representative of PCM-based blends or composites, since these composites are usually heterogeneous. In addition, PCMs used in buildings, due to the cost constrains, are seldom high-quality, pure substances. Typically, PCMs used in buildings contain a wide range of different impurities. That is why a careful sampling is required to obtain the representative test-generated characteristics for building products containing PCMs. This limits the usefulness of DSC and DTA testing procedures to characterize PCMs for use in building applications. Considering the above limitations, Paksoy (1996) introduced the twin water bath method, which determines the thermal storage capacity of PCMs or PCM-based composites. The water bath method is relatively inexpensive, simple and fast, making it widely used today. The method enables evaluation of phase change temperature and the thermal storage capacity of PCM for a wide range of temperatures. In addition, Zhang (1999) and Zhang et al. (2007) proposed the T-history method as an alternative to DSC or DTA testing. It is an inexpensive and easy way to measure the phase change enthalpy of PCM products using test samples significantly larger from what is required for DSC testing. The T-history method can be used for measuring the thermal properties of PCM composites and PCM blends.

Since, most often, PCMs used in building envelopes are difficult to analyze complex heterogeneous composites or structures with multidimensional heat transfer effects, Kośny and Yarbrough introduced a new test method utilizing a ASTM C518 (2013) heat flow meter apparatus—see Kośny et al. (2007b). This method determines the amount of phase change energy available in heterogeneous PCM-enhanced materials to provide thermal storage—Shukla et al. (2013). These developments led to the introduction of a new ASTM standard-ASTM C1784 (2015), "*Standard Test Method for Using a Heat Flow Meter Apparatus for Measuring Thermal Storage Properties of Phase Change Materials and Products.*"

Numerical Tools for PCM Thermal and Energy Performance Analysis

It is expected that the advanced buildings of tomorrow will need to take advantage of the LHS technologies customized for thermal loads mitigation, load shifting from peak to off-peak hours, and for internal thermal comfort improvements. Numerous numerical tools have been developed worldwide to simulate the thermal behavior of PCM-enhanced building components or energy consumption in whole buildings utilizing these technologies. In Crawley et al. (2005), a comprehensive overview was given of various computational capabilities and features which can be found in 20 commonly used building simulation programs. Examples of numerical applications of the programs are listed in Table 4.9.3. It is important to mention here that among a great variety of whole-building computer programs available today, TRNSYS is one of the oldest tools that has been used for decades for analysis of PCM-enhanced building technologies. TRNSYS is the energy simulation software package, developed by the Solar Energy Laboratory at the University of Wisconsin, Madison, USA, that has been extensively used worldwide in the fields of renewable energy engineering, building energy simulations, passive and active solar designs, for over three decades—Pedersen (2007). TRNSYS was adopted by Tomlinson and Heberle (1990) for simulations of the gypsum boards impregnated with paraffin. This work was later continued by Stovall and Tomlinson (1995).

ESP-r is used to simulate PCMs that demonstrate hysteresis. To do this, the program uses different specific heat equations for melting and solidification. ESP-r is a whole-building performance open-source energy tool developed in 1974 by the Energy Systems Research Unit, University of Strathclyde, UK.

The simulation software package "EnergyPlus" was developed by the US Department of Energy. It enables modeling of building energy use with consideration of a variety of HVAC systems. Extensive design options make EnergyPlus a very useful and ideal program for the study of building energy use. EnergyPlus is a collection of modules that are used to calculate heating and cooling requirements. The most recent version of EnergyPlus can be downloaded from the NREL homepage: https://github.com/NREL/EnergyPlusRelease/releases.

CoDyBa is used to analyze transient thermal, and hygrothermal performance of building components subjected to specified climatic conditions (Noel et al., 2001). The program is for the study of heating and cooling strategies including air conditioning, ventilation options, and insulation assemblies. CoDyBa can be used to study the thermal, hygrothermal, and energy performance of building components and buildings of many types and sizes. In addition, it can be used for peak load estimates, building shell optimization, and the design of space-conditioning systems.

TABLE 4.9.3
References for Applications of Heat Transfer Tools and Whole-Building Energy Programs

Computer Program	References
TRNSYS[a]	Stritih and Medved (1994), Stovall and Tomlinson (1995), Koschenz and Lehmann (2000), Ibáñez et al. (2005), Ahmad et al. (2006), Kuznik et al. (2007), and Poulad et al. (2011)
ESP-r	Heim and Clarke (2004), Schossig et al. (2005), Gelissier (2008), Ponechal (2009), Heim (2010), and Fallahi et al. (2012)
EnergyPlus	Almeida et al. (2010), Kośny et al. (2010), Tardieu et al. (2011), Campbell and Sailor (2011), and Tabares-Velasco et al. (2012)
RADCOOL[b]	Feustel and Stetiu (1997)
CoDyBa[c]	Virgone et al. (2009)
BSim[d]	Rode et al. (2003)
WUFI[e]	WUFI (a family of hygrothermal tools with whole-building simulation capabilities)—Shrestha (2012)

[a] http://www.trnsys.com/#1.

[b] https://digital.library.unt.edu/ark:/67531/metadc624249/m2/1/high_res_d/105892.pdf

[c] http://www.jnlog.com/codyba1_en.htm.

[d] http://sbi.dk/en/bsim/about-bsim.

[e] http://www.verlag.fraunhofer.de/wufi/?local=en.

The integrated whole-building numerical tool called BSim was developed by the Danish Building Research Institute, of Aalborg University, Denmark. This is a one-dimensional computer model for analysis of heat and moisture transport. BSim is a successor of several earlier-validated numerical components (Rode et al., 2003). Most recent versions of BSim and program technical literature can be downloaded from the Statens Byggeforskningsinstitut, Aalborg University, Copenhagen, Denmark, http://sbi.dk/en/bsim.

WUFI is a set of numerical routines for modeling of heat and moisture transfer within building structures. WUFI, an abbreviation for "Wärme und Feuchte instationär," was developed by Fraunhofer IBP, Holzkirchen, Germany (Kunzel, 1995). Transient heat and moisture transport simulations can be performed in WUFI using user-defined time steps and user-specified time periods. Finite volume methods are employed with building materials containing PCM simulated in WUFI 5.0 and WUFI Plus using temperature-dependent enthalpy.

Examples of PCM Building Applications

The great diversity of building products containing PCM can help in the near future to achieve peak-hour load and HVAC energy consumption reductions. Notable energy savings can be accomplished in residential and commercial buildings, by means of a variety of applications utilizing external envelopes, internal building fabric technologies, and space-conditioning technologies having the capability for passive and active utilization of sensible and latent heat accumulation. Building components with phase transition capability are allowing large amounts of heat to be stored or released due to the latent heat of fusion or solidification. Residential and commercial air conditioning loads during critical peak periods of the day place a burden on the electricity system. Different forms of PCM (boards, containers, and PCM slurry-based systems) could notably reduce the growing peak electrical demand in many climates worldwide. PCM slurry systems have the advantage of low energy consumption and the ability to absorb large amounts of energy cost-effectively. The feasibility of PCM slurry-based mechanical systems used to absorb peak loads in air conditioning units has been analyzed in many projects—see Yamagishi (1999), Schossig et al. (2003), Egolf et al. (2004), Gschwander et al. (2005), IEA Task 32 (2006), Okamoto (2006), Heinz and Streicher (2006), Wang et al. (2007), Thaicham (2008), Maaraoui et al. (2012), and Niu et al. (2013). The major performance issue that usually needs to be validated is the system ability to develop and maintain high volume fraction PCM slurries and the demonstration that pumping the PCM slurry through the heat exchanger can absorb and reject peak air conditioning loads in a cost-effective manner. Further research should be undertaken to find a PCM material that does not exhibit supercooling or to develop a method to induce phase change at the design temperature, e.g., see Goel et al. (1994), Schossig et al. (2003), Bony et al. (2005), Sabbah et al. (2008), and Valan et al. (2013), etc.

The heat capacity of lightweight-construction buildings can be significantly increased by including LHS materials into the building envelope surfaces as well as within the interior building fabric. The use of PCMs in construction materials in a passive way is relatively not new. There were several moderately successful attempts in the 1970s and 1980s to utilize different types of organic and inorganic PCMs to reduce peak loads and heating and cooling energy consumption, as demonstrated by Balcomb et al. (1983), Hawes et al. (1990), and Salyer et al. (1995). These investigations focused on impregnating concrete, gypsum, or ceramic masonry with salt hydrates or paraffinic hydrocarbons. In the 1980s and 1990s, several researchers have investigated methods for impregnating gypsum wallboard and other architectural materials with PCMs including Salyer et al. (1995), Shapiro et al. (1987), Babich et al. (1994), and Banu et al. (1998). One of the simplest methods for addition of PCM to the concrete or gypsum mixtures is a usage of microencapsulated PCM. As an example, Table 4.9.4 shows the physical properties of a commercially available dry plaster mixture containing microencapsulated PCM—Kośny (2015). Different manufacturing techniques, followed with thermal performance of gypsum wallboards and concrete blocks containing PCM, have been discussed by Khudhair and Farid (2004), Zalba et al. (2003), and Hauer et al. (2005). Example material physical characteristics are

presented in Table 4.9.5 for one of the produced Europe gypsum boards containing microencapsulated PCM—Kośny (2015). In addition, an innovative concept of the PCM-enhanced building insulation was introduced by Oak Ridge National Laboratory, USA—Kośny et al. (2006, 2007a).

PCM-enhanced composite boards or panels are usually manufactured with the use of waste wood, cellulose fibers, or fiber-reinforced mineral composites. Composite wood-based products, also called man-made wood, include a range of derivative wood products which are manufactured by binding the wood strands, particles, fibers, or veneers together with most often adhesives to form composite materials. Plywood and oriented strand board are probably the best-known wood-based composites used for construction purposes today. There is also a group of board products utilizing polymer-based shape-stabilized PCM—see Table 4.9.6. Sometimes, these board products are faced with aluminum foil (Kośny, 2015).

TABLE 4.9.4
Physical Properties of Dry Gypsum Plaster Containing PCM

Physical Property	Value
Compressive strength	≥1.5 N/mm²
Thermal conductivity	>0.3 W/m·K
Water vapor diffusion resistance factor μ	10
Shift range of the PCM (23/26)	23°C–26°C
Enthalpy of fusion of the PCM: encapsulated	About 100 J/g
Enthalpy of fusion (*T*s = 23°C–26°C): plaster composite	18 kJ/kg
Thermal conductivity λ (10°C) =	0.19 W/m·K
Density	~0.90 g/cm³

TABLE 4.9.5
Physical Properties of Gypsum Board Containing Microencapsulated PCM

Physical Property	Value
Specific heat	1.2 kJ/kg·K
Thermal conductivity at 20°C	0.20 W/m·K
Thermal conductivity at 35°C	0.19 W/m·K
Shift range of the PCM	23°C or 26°C
Enthalpy of fusion of the PCM	About 110 J/g
Latent heat	330 kJ/m²

TABLE 4.9.6
Dimensions and Characteristics of Composite Board Product with Shape-Stabilized PCM

Board dimensions	1000 mm × 1198 mm × 5.2 mm
Board area weight	24.5 kg/m²
PCM % weight content	60%
PCM melting point	21.7°C
Latent heat (0°C–30°C)	70 kJ/kg
Total heat storage capacity (0°C–30°C)	140 kJ/kg
Solid-phase thermal conductivity	0.18 W/m·K
Liquid-phase thermal conductivity	0.14 W/m·K

REFERENCES

Abhat, A. Low temperature latent heat thermal energy storage: Heat storage materials. *Solar Energy* 30 (1983), 313–332.

Agyenim, F., Hewitt, N., Eames, P., and Smyth, M. A review of materials, heat transfer and phase change problem formulation for latent heat thermal energy storage systems (LHTESS). *Renewable and Sustainable Energy Reviews* 14 (2010), 614–628.

Ahmad, M., Bontemps, A., Salle H., and Quenard, D. Thermal testing and numerical simulation of a prototype cell using light wallboards coupling vacuum insulation panels and phase change materials. *Energy and Buildings* 38 (2006), 673–681.

Almeida, F., Zhang, D., Fung, A.S., and Leong, W.H. Investigation of multilayered phase-change material modeling in ESP-R. *1st International High Performance Buildings Conference*, Purdue University, West Lafayette, IN (2010).

ASTM C1784. Standard test method for using a heat flow meter apparatus for measuring thermal storage properties of phase change materials and products. *2015 Annual Book of ASTM Standards*, Vol. 04.06 (2015), 1120–1136.

ASTM C518. Standard test method for steady-state thermal transmission properties by means of the heat flow meter apparatus, ASTM International, West Conshohocken, PA, 2017.

Athienitis, A.K., Liu, C., Hawes, D., Banu, D., and Feldman, D. Investigation of the thermal performance of a passive solar test-room with wall latent heat storage. *Building and Environment* 32 (1997), 405–410.

Athienitis, A.K., Zhang, J., and Feldman, D. A study of double facades with phase-change storage and photovoltaics. *Proceedings of International Conference "Passive and Low Energy Cooling for the Built Environment*", May 2005, Santorini, Greece.

Babich, M.W., Benrashid, R., and Mounts, R.D. DSC studies of energy storage materials. Part 3. Thermal and flammability studies. *Thermochimica Acta* 243 (1994), 193–200.

Balcomb, J.D., Kosiewicz, C.E., Lazarus, G.S., McFarland, R.D., and Wray, W.O. *Passive Solar Design Handbook vol. 3*, Los Alamos National Laboratory, Los Alamos, NM, (1983). ISBN 0-89553-106-2.

Banu, D., Feldman, D., Haghighat, F., Paris, J., and Hawes, D. Energy-storing wallboard: Flammability tests. *Journal of Materials in Civil Engineering* 10 (1998), 98–105.

Becker, J.R. *Crude Oil Waxes, Emulsions and Asphaltenes*. Penn Well Publishing Company, Tulsa, OK (1997).

Belton, G., and Ajami, F. Thermochemistry of salt hydrates. National Science Foundation Report No. NSF/RANN/SE/GI27976/TR/73/4 (1973).

Bony, J. and Citherlet, S. Numerical model and experimental validation of heat storage with phase change materials. *Energy and Buildings* 39 (2007), 1065–1072, 0378–7788.

Bony, J., Ibanez, M., Puschnig, P., Citherlet, S., Cabeza, L., and Heinz, A. Three different approaches to simulate PCM bulk elements in a solar storage tank. *Phase Change Material and Slurry: Scientific Conference and Business Forum*, June 15–17, 2005, Yverdon les Bains, Schweiz, S, 99–107.

Cabeza, L.F., Ibáñez, M., Solé, C., Roca, J., and Nogués, M. Experimentation with a water tank including a PCM module. *Solar Energy Materials and Solar Cells* 90 (2006), 1273–1282.

Cabeza, L.F., Castellón, C., Nogués, M., Medrano, M., Leppers, R., and Zubillaga, O. Use of microencapsulated PCM in concrete walls for energy savings. *Energy and Buildings* 39 (2007), 113–119.

Campbell, K.R., and Sailor, D.J. Phase change materials as thermal storage for high performance homes. *Proceedings of the 2011 ASME International Mechanical Engineering Congress & Exposition (IMECE)*, November 11–17, 2011, Denver, CO (2011).

Childs, K., and Stovall, T. Potential energy savings due to phase change material in a building wall assembly: An examination of two climates. Report-ORNL/TM-2012/6 March 2012, Oak Ridge National Laboratory, Oak Ridge, TN (2012).

Crawley, D.B., Hand, J.W., Kummert, M., and Griffith, B.T. Contrasting the capabilities of building energy performance simulation programs. A Joint Report, US DOE, University of Strathclyde UK, University of Wisconsin, MD and National Renewable Energy Renewable Laboratory, Golden, CO (2005).

Demirbas, F.M. Thermal energy storage and phase change materials: An overview. *Energy Sources, Part B*, 1 (2006), 85–95.

Demirel, Y., and Paksoy, H.Ö. Thermal analyses of heat storage materials. *Thermo-Chimica Acta* 213 (1993), 211–221.

Dincer, I., and Rosen, M.A. *Thermal Energy Storage: Systems and Applications*, Wiley, Chichester, UK (2002).

Dolado, P., Lazaro, A., Marin, J.M., and Zalba, B. Characterization of melting and solidification in a real scale PCM-air heat exchanger: Numerical model and experimental validation. *Energy Conversion and Management* 52 (2011), 1890–1907.

Egolf, P.W., Sari, O., Brulhart, J., Gendre, F., Ata-Caesar, D., and Vuarnoz, V. (2004). Physical behaviour of phase change material slurries. Deliverable D6, Workpackage 3, EU project ENK6-CT-2001–00507 (PAMELA).

Fallahi, A., Shukla, N., and Kośny, J. Numerical thermal performance analysis of PCMs integrated with residential attics. *SimBuild2012 Conference*, Madison, WI, August 1–3 (2012).

Farid, M.M., Khudhair, A.M., Ali, K., Razack, S., and Al-Hallaj, F. A review on phase change energy storage: Materials and applications. *Energy Conversion and Management* 45 (2004), 1597–1615.

Feldman, D., Shapiro M. M. Fatty acids and their mixtures as phase-change materials for thermal energy storage. *Solar Energy Materials* 18 (1989), 201–216.

Feldman, D., Shapiro, M., Banu, D., and Fuks, C.J. Fatty acids and their mixtures as phase change materials for thermal energy storage. *Solar Energy Mater* 18 (1989), 201–216.

Feustel, H.E. *Simplified Numerical Description of Latent Storage Characteristics for Phase Change Wallboard. Indoor Environmental Program, Energy and Environment Division*. Lawrence Berkeley National Laboratory. University of California, Berkeley, CA (1997/1995).

Feustel, H. E., and Stetiu, C. Thermal performance of phase change wallboard for residential cooling application. Indoor Environment Program Energy and Environment Division Lawrence Berkeley National Laboratory, University of California Berkeley, CA 94720. Report-LBL-38320 UC-1600 (1997).

Gelissier, A. SPMCMP56 subroutine in ESP-r source standard code (2008). http://www.esru.strath.ac.uk/Programs/ESP-r.htm.

Goel, M., Roy, S.K., and Sengupta, S. Laminar forced convection heat transfer in microencapsulated phase change material suspensions. *Heat and Mass Transfer* 37 (1994), 593–604.

Gschwander, S.P., Schossig, P., and Henning, H.M. Micro-encapsulated paraffin in phase-change slurries. *Solar Energy Materials and Solar Cells* 89 (2005), 307–315.

Günther, E., Mehling, H., and Hiebler, S. Modeling of sub-cooling and solidification of phase change materials. *Modelling and Simulation in Materials Science and Engineering* 15 (2007), 879–892.

Günther, E., Hiebler, S., Mehling, H., and Redlich, R. Enthalpy of phase change materials as a function of temperature: Required accuracy and suitable measurement methods. *International Journal of Thermophysics* 30 (2009), 1257–1269.

Hadorn J. C. "IEA SOLAR HEATING AND COOLING PROGRAMME TASK 32: ADVANCED STORAGE CONCEPTS FOR SOLAR AND LOW ENERGY BUILDINGS," ECOSTOCK 2006, *The Tenth International Conference on Thermal Energy Storage*, Stockton College, NJ, May 31–June 2, (2006).

Hauer, A., Schossig, P., Yamaha, M., Cabeza, L.F., Martin, V, and Setterwall, F. International Energy Agency implementing agreement on energy conservation through energy storage. Annex 17 Final Report (2005).

Hawes D.W., Banu D., Feldman, D. Latent heat storage in concrete. *Solar Energy Mater* 21 (1990), 61–80.

Hawes, D.W., Feldman, D., and Banu, D. Latent heat storage in building materials. *Energy and Buildings* 20 (1993), 77–86.

Hawlader, M.N.A., Uddin, M.S., and Khin, M.M. Microencapsulated PCM thermal-energy storage systems. *Applied Energy* 74 (2003), 195–202.

Heckenkamp, J., and Baumann, H. Latentwärmespeicher. *Sonderdruck aus Nachrichten* 11 (1997), 1075–1081, Germany.

Heim, D. Isothermal storage of solar energy in building construction. *Renewable Energy* 35 (2010), 788–796.

Heim, D., and Clarke, J.A. Numerical modelling and thermal simulation of PCM-gypsum composites with ESP-r. *Energy and Buildings* 36 (2004), 795–805.

Heinz, A., and Streicher, W. Application of phase change materials and PCM slurries for thermal energy storage. *ECOSTOCK – IEA International Conference on Thermal Energy Storage*. Stockton College of New Jersey, May 31–June 2006, Galloway, NJ (2006).

Ibáñez, M., Lazaro, A., Zalba, B., and Cabeza, L.F. An approach to the simulation of PCMs in building applications using TRNSYS. *Applied Thermal Engineering* 25 (2005), 1796–1807.

IEA. Inventory of phase change materials (PCM). A report of IEA Solar Heating and Cooling Programme – Task 32. Advanced storage concepts for solar and low energy buildings. Report C2 from Subtask C. International Energy Agency Solar Heating and Cooling Program (2005). Document can be found online at: http://www.iea-shc.org/data/sites/1/publications/task32-Inventory_of_PCM.pdf.

IEA Task 32. Storage based on phase change materials (PCM)-Selection of concepts. Report C1 from Subtask C. International Energy Agency Solar Heating and Cooling Program (February 2005). Document can be found online at: http://archive.iea-shc.org/publications/downloads/task32-Storage_Based_on_PCM.pdf.

IEA. Development of a test-standard for PCM and TCM characterization Part 1: Characterization of phase change materials technical report. IEA-Solar Heating and Cooling/Energy Conservation through Energy Storage Programme – Task 42/Annex 24: Compact Thermal Energy Storage: Material Development for System Integration (2011). https://www.iea-shc.org/data/sites/1/publications/Task4224-A2-1-Characterization-of-Phase-Change-Materials.pdf.

Inaba, H. and Tu, P. Evaluation of thermophysical characteristics on shape stabilized paraffin as a solid-liquid phase change material. *Heat and Mass Transfer* 32 (1997), 307–312.

Khudhair, A.M., and Farid, M.M. A review on energy conservation in building applications with thermal storage by latent heat using phase change materials. *Energy Conversion* and *Management* 45 (2004), 263–275.

Khudhair, A.M., and Farid, M.M. Use of phase change materials for thermal comfort and electrical energy peak load shifting: Experimental investigations. In: Goswami, D.Y. and Zhao Y., (Eds.), *Solar World Congress*, September 18–21, 2007, Beijing, China.

Kissock, J.K., Hannig, J.M., and Thomas, I. Testing and simulation of phase change wallboard for thermal storage in buildings. *Proceedings of 1998 International Solar Energy Conference*, Albuquerque, June 14–17. J.M. Morehouse and R. E. Hogan, (Eds.). American Society of Mechanical Engineers, New York (1998).

Koschenz, M., and Lehmann B. Handbuch Thermoaktive Bauteilsysteme. TABS, EMPA Abteilung Energiesysteme/Haustechnik, CH-8600, Düsseldorf, Germany.

Koschenz, M., Lehmann, B. Thermoaktive Bauteilsysteme tabs, EMPA, Duebendorf Switzerland, (2006). ISBN 3-905594-19-6.

Kośny, J, Yarbrough, D.W., Wilkes, K. E., Leuthold, D., and Mohiuddin, S. A. PCM-enhanced cellulose insulation-thermal mass in lightweight natural fibers. *ECOSTOCK – IEA International Conference on Thermal Energy Storage*, Stockton College of New Jersey, May 31–June, 2006, Galloway, NJ (2006).

Kośny, J., Yarbrough, D.W., Miller, W., Childs, P., and Syed, A.M. Thermal performance of PCM-enhanced building envelope systems. *Proceedings-Thermal Performance of the Exterior Envelopes of Buildings*, December 2007a, Clearwater Beach, FL.

Kośny, J., Yarbrough, D.W., and Mohiuddin S.A. Performance of thermal insulation containing microencapsulated phase change material. *Proceedings of the International Thermal Conductivity Conference ITCC29 and ITES17*, June 24–27, 2007b, Birmingham, AL.

Kośny, J., Yarbrough, D.W., and Miller, W. Use of PCM-enhanced insulations in the building envelope. *Journal of Building Enclosure Design,* Summer/Fall (2008), 55–59.

Kośny, J., Kossecka, E., and Yarbrough, D.W. Use of a heat flow meter to determine active PCM content in an insulation. *Proceedings of the 2009 International Thermal Conductivity Conference (ITCC) and the International Thermal Expansion Symposium (ITES)*, August 29–September 2, 2009, Pittsburgh, PA (2009).

Kośny, J., Miller, W., and Zaltash, A. Dynamic thermally-disconnected building envelopes—A new paradigm for walls and roofs in low energy buildings. *Proceedings of DOE, ASHRAE, ORNL Conference—Thermal Envelopes XI: Thermal Performance of the Exterior Envelopes of Buildings*, December 2010, Clearwater, FL (2010).

Kośny, J., Biswas, K., Miller, W.A., and Kriner, S. Field thermal performance of naturally ventilated solar roof with PCM heat sink. *Solar Energy* 86 (2012a), 2504–2514.

Kośny, J., Kossecka, E., Brzezinski, A., Tleoubaev, A., and Yarbrough, D.W. Dynamic thermal performance analysis of fiber insulations containing bio-based phase change materials (PCMs). *Energy and Buildings* 52 (2012b), 122–131.

Kośny, J. and Kossecka, E. Understanding a potential for application of phase-change materials (PCMs) in building envelopes. *ASHRAE Transactions* 119:1 (2013), DA-13-001.

Kośny, J., Fallahi, A., Shukla, N., Kossecka, E., and Ahbari, R. Thermal load mitigation and passive cooling in residential attics containing PCM-enhanced insulation. *Solar Energy* 108 (2014), 164–177.

Kośny, J. *PCM-Enhanced Building Components: An Application of Phase Change Materials in Building Envelopes and Internal Structures* (2015). Book, 01/2015; Springer International Publishing AG; Cham, Switzerland, ISBN: 978-3-319-14285-2.

Kunzel, H. Simultaneous heat and moisture transport in building components. Ph.D. Thesis. Wiley IRB Verlag, Stuttgart (1995).

Kuznik, F., Virgone, J., and Noel, J. Optimization of a phase change material wallboard for building use. *Applied Thermal Engineering* 28 (2007), 1291–1298.

Lane, G.A. Macro–encapsulation of PCM, Report No. oro/5117-8, Dow Chemical Company, Midland, MI (1978).

Lane, G.A. Low temperature heat storage with phase change materials. *International Journal of Ambient Energy* 1 (1980), 155–168.

Lane, G.A. *Solar Heat Storage: Latent Heat Materials*, vol. 1, CRC Press, Boca Raton, FL (1983).

Lane, G.A. Chapter 1: Phase change thermal storage materials. In: C. Guyer (ed.), *Handbook of Thermal Design*, McGraw Hill, New York (1989).

Lingayat, A.B., and Suple, Y.S. Review on phase change material as thermal energy storage medium: Materials, application. *International Journal of Engineering Research and Applications* 3(4) (2013), 916–921.

Maaraoui, S., Clodic, D., and Dalicieux, P. Heat pump with a condenser including solid-liquid phase change material. *International Refrigeration and Air Conditioning Conference,* Purdue University, West Lafayette, IN, (2012), Paper 1194. http://docs.lib.purdue.edu/iracc/1194.

Martín, J.M., Zalba, B., Cabeza, L.F., and Mehling, H. Determination of enthalpy–temperature curves of phase change materials with the temperature-history method: Improvement to temperature dependent properties. *Measurement Science and Technology* 14 (2003), 184–489.

Mehling, H., and Cabeza, L.F. *Heat and Cold Storage with PCM.* Springer-Verlag, Berlin, Heidelberg, Aug. 15, 2008.

Murugananthama, K., Phelan, P., Horwath, P., Ludlam, D., and McDonald, T. Experimental investigation of a bio-based phase-change material to improve building energy performance. *Proceedings of 4th ASME International Conference on Energy Sustainability, ES2010*, May 17–22, 2010, Phoenix, AZ (2010), http://proceedings.asmedigitalcollection.asme.org/volume.aspx?volumeid=15077.

Nagano, K., Mochida, T., Iwata, K., Hiroyoshi, H., and Domanski, R. Thermal performance of Mn(NO3)2·6H2O as a new PCM for cooling system, 5th Workshop of the IEA ECES IA Annex 10, Tsu (Japan), 2000.

Najafiana, A, Haghighata, F., and Moreau, A. Influence of GSHP system design parameters on the geothermal application capacity and electricity consumption at city-scale for Westminster, London. *Energy and Buildings* 106 (2015), 3–12.

Niu, F., Ni, L., Yao, Y., Yu, Y., and Li, H. Performance and thermal charging/discharging features of a phase change material assisted heat pump system in heating mode. *Applied Thermal Engineering* 58 (2013), 536–541.

Noel, J., Roux, J., and Schneider, P.S. CoDyBa, a design tool for buildings performance simulation. *Proceedings of the Seventh International IBPSA Conference*, August 13–15, 2001, Rio de Janeiro, Brazil (2001).

Okamoto, S. A heat pump system with a latent heat storage utilizing seawater installed in an aquarium. *Energy and Buildings* 38 (2006), 121–128.

Paksoy, H.O. Determining thermal properties of heat storage materials using the twin bath method. *Energy Conversion and Management* 37(3) (1996), 261–268.

Pedersen, C.O. Advanced zone simulation in EnergyPlus: Incorporation of variable properties and phase change materials (PCM) capability. *Proceedings of the 10th International IBPSA Conference*, Beijing, China, 1341–1345 (2007).

Petrie, T.W., Childs, K.W., Childs, P.W., Christian, J.E., and Shramo, D.J. Thermal behavior of mixtures of perlite and phase change material in a simulated climate, pp. 180–194 in *Insulation Materials: Testing and Applications: Third Volume*, ASTM STP 1320, R.S. Graves and R.R. Zarr (Eds.), American Society for Testing and Materials, W. Conshohocken, PA (1997).

Ponechal, R. An experimental study and simulations of phase change materials in an office thermal environment. *Slovak Journal of Civil Engineering* 2009 (3), 24–29.

Poulad, M.E., Fung, A.S., and Naylor, D. Effects of convective heat transfer coefficient on the ability of PCM to reduce energy demand. *Proceedings of Building Simulation 2011: 12th Conference of International Building Performance Simulation Association*, November 14–16, Sydney (2011).

Rode, C., Salonvaara, M., Ojanen, T., Simonson, C., and Grau, K. Integrated hygrothermal analysis of ecological buildings. *Proceedings of the Second International Building Physics Conference*, September 14–18, 2003, Leuven, Belgium (2003).

Sabbah, R., Farid, M.M., and Al-Hallaj, S. Micro-channel heat sink with slurry of water with micro-encapsulated phase change material: 3D-numerical study. *Applied Thermal Engineering* 29 (2008), 445–454.

Salyer, I.O., and Sircar, A. Development of PCM wallboard for heating and cooling of residential buildings, pp. 97–123, *Thermal Energy Storage Research Activities Review*, U.S. Department of Energy, March 15–17, 1989, New Orleans, LA.

Salyer, I.O., Sircar, A.K., and Kumar, A. Advanced phase change materials technology: Evaluation in light-weight solite hollow-core building blocks. *In: Proceedings of the 30th Intersociety Energy Conversion Engineering Conference*, Orlando, FL, July 30–August 4 (1995), 217–224.

Sari, A. Form-stable paraffin/high density polyethylene composites as solid–liquid phase change material for thermal energy storage: Preparation and thermal properties. *Energy Conversion and Management* 45 (2004), 2033–2042.

Schossig, P., Henning, H.-M., Raicu, A., and Haussmann, T. Mikroverkapselte Phasenwechselmaterialien in Baustoffen, 13. Symposium Thermische Solarenergie OTTI, Technologie-Kolleg, Staffelstein, S. 489–494 (2003).

Schossig, P., Henning H.M., Gschwander, S., and Haussmann, T. Micro-encapsulated phase-change materials integrated into construction materials. *Solar Energy Materials and Solar Cells* 89: 2–3 (2005), 297–306.

Shapiro, M., Feldman, D., Hawes, D., and Banu, D. PCM thermal storage in drywall using organic phase-change materials. *Passive Solar Journal* 4(4) (1987), 417–438.

Sharma, A., Tyagi, V.V., Chen, C.R., and Buddhi, D. Review on thermal energy storage with phase change materials and applications. *Renewable and Sustainable Energy Reviews* 13 (2009), 318–345.

Shrestha, M. PCM Application-effect on energy use and IA temperature. Norwegian University of Science and Technology, Faculty of Architecture and Fine Arts. MSc. Thesis-Sustainable Architecture, September 01, 2012.

Shukla, N., Cao, P., and Kośny, J. Lab-scale dynamic thermal testing of PCM-enhanced building materials. *Proceedings of the ASTM Symposium on Next-Generation Thermal Insulation Challenges and Opportunities*, October 23–24, 2013, Jacksonville, FL.

Stritih, U. Heat transfer enhancement in latent heat thermal storage system for buildings. *Energy and Buildings* 35 (2003), 1097–1104.

Stritih, U., and Medved, S. Use of phase change materials in the wall with TIM. *Strojniski Vestnik/Journal of Mechanical Engineering* 40 (1994), 155–160.

Stovall, T.K. and Tomlinson, J.J. What are the potential benefits of including latent storage in common wallboard? *Journal of Solar Energy Engineering* 117 (1995), 318–325.

Tabares-Velasco, P.C., Christensen, C., and Bianchi, M. Verification and validation of EnergyPlus phase change material model for opaque wall assemblies. *Building and Environment* 54 (2012), 186–196.

Tardieu, A., Behzadi, S., Chen, J.J., and Farid, M. Computer simulation and experimental measurements for an experimental PCM-impregnated office building. *Proceedings of the Building Simulation 2011: 12th Conference of International Building Performance Simulation Association,* vol. 1, 56–63, Sydney, November 14–16, (2011).

Telkes, M. Nucleation of super saturated inorganic salt solution. Industrial and Engineering Chemistry, American Chemical Society, ACS Publications, vol 44 (1952), 1308.

Telkes, M. Thermal storage for solar heating and cooling. *Proceedings of the Workshop, Solar Energy Storage Subsystems for the Heating and Cooling of Buildings*, Charlottesville, VA, April 16–18, (1975).

Telkes, M. Trombe wall with phase change storage material. *Proceedings of the 2nd National Passive Solar Conference*, Philadelphia, PA, March 16–18, (1978).

Thaicham, P. Heat transport in a three-stage thermal energy system using phase change material slurries. *ScienceAsia* 34 (2008), 223–228.

Tomlinson, J., and Heberle, D. Analysis of wallboard containing a phase change material. *Proceedings of the 25th Intersociety Energy Conversion Engineering Conference, Vol. 4*, August 12–17, 1990, Reno, NV Published by IEEE.

Tomlinson, J., Jotshi, C., and Goswami, D. Solar thermal energy storage in phase change materials. *Proceedings of Solar 1992: The American Solar Energy Society Annual Conference*, June 15–18, Cocoa Beach, FL (1992).

Valan, A.A., Sasmito, A.P., and Mujumdar, A.S. Numerical performance study of paraffin wax dispersed with alumina in a concentric pipe latent heat storage system. *Thermal Science* 17 (2013), 419–430.

Virgone, J., Noël, J., and Reisdorf, R. Numerical study of the influence of the thickness and melting point on the effectiveness of phase change materials: Application to the renovation of a low inertia school. *Proceedings of the 11th International IBPSA Conference*, July 27–30, 2009, Glasgow (2009).

Wang, X., Jianlei, N., Li, Y., Chen, B., Zeng, R., Song, Q., and Zhang, Y. Flow and heat transfer behaviours of phase change material slurries in a horizontal circular tube. *International Journal of Heat and Mass Transfer* 50 (2007), 2480–2491.

Wunderlich, B. *Thermal Analysis of Polymeric Materials* (2005). Springer, Berlin, Heidelberg, New York.

Xu, X., Zhang, Y.P., Lin, K.P., Di, H.F., and Yang, R. Modeling and simulation on the thermal performance of shape-stabilized phase change material floor used in passive solar buildings. *Energy and Buildings* 37 (2005), 1084–1091.

Yamagishi, Y. Characteristics of micro-encapsulated PCM slurry as a heat-transfer fluid. *AlChE Journal* 45 (1999), 696–707.

Yamaha, M., Shuku, K., and Misaki, S. A study on thermal characteristics of thermal storage tank using phase change material installed in an air distribution system. *Journal of Architecture and Planning Transaction of AIJ*, 549 (2001), 51–57.

Zalba, B., Marín, J.M., Cabeza, L.F., and Mehling, H. Review on thermal energy storage with phase change: Materials, heat transfer analysis and applications. *Applied Thermal Engineering* 23 (2003), 251–283.

Zhang, Y. A simple method, the t-history method, of determining the heat of fusion, specific heat and thermal conductivity of phase-change materials. *Measurement Science and Technology* 10: 3 (1999), 201–205.

Zhang, M., Medina, M.A., and King, J. Development of a thermally enhanced frame wall with phase-change materials for on-peak air conditioning demand reduction and energy savings in residential buildings. *International Journal of Energy Research* 29:9 (2005), 795–809.

Zhang, Y., Zhou, G., Lin, K., Zhang, Q., and Di, H. Application of latent heat thermal energy storage in buildings: State-of-the-art and outlook. *Building and Environment* 42:6 (2007), 2197–2209.

Zukowski, M. Experimental study of short term thermal energy storage unit based on enclosed phase change material in polyethylene film bag. *Energy Conversion and Management* 48 (2007), 166–173.

4.10 THERMAL BRIDGES IN BUILDING STRUCTURES

Jan Kośny and David W. Yarbrough

Introduction and Scope

Theoretically, the exterior building envelopes should perform a variety of different roles including keeping the exterior climate outside by limiting the transfer of heat, moisture, and air, which helps with maintaining comfortable interior conditions. Additional functions provide a visual and daylight connection to the outdoors, limit noise transmission, support of structural loads, and aesthetically pleasing appearance on both sides. Although building envelopes are usually effective in meeting these requirements, there are situations where they do not perform adequately. Shortcomings in thermal performance are often exhibited by excessive heat transfer (in and out) and unwelcome air or moisture transport across the building envelope that can lead to durability problems and increased energy consumption, as well as to poor thermal comfort within the occupied space. Sometimes, the presence of a thermal bridge results in discoloration on the interior, exterior, or both surfaces of the building envelope. While some cases of poor building envelope performance are associated with inadequate thermal insulation levels, other cases often occur because of construction imperfections; such as discontinuities in the thermal insulation or air barrier systems, compressed insulation, and air-leakage sites.

During the last half-century, numerous wall technologies have been introduced. Some of them are complex three-dimensional networks of structural components and thermal insulation. Moreover, building structural systems are advancing and every year a variety of new materials are introduced. Consequently, buildings are often becoming unforgiving of design errors or assembly imperfections. At the same time, the overall thermal efficiency of a building is a function of the thermal performance of the exterior envelope elements (walls, roofs, and windows), the local heat losses that can occur around the envelope elements where they are penetrated by the heat conducting building components. These areas of intense local heat flow, commonly known as thermal shorts or thermal bridges, can have a significant impact on the overall thermal performance of the building envelope.

With the introduction of advanced materials, a single change in a building envelope configuration may compromise local thermal performance and durability of the exterior envelopes. That is why more and more frequently, only optimized combinations of structural subsystems and insulation materials are specified. In buildings, thermal bridges are responsible for a significant number of thermal performance and durability problems. Thermal bridges, thermal shorts, or cold bridges (different names are used by building professionals in different parts of the world) are understood as unwanted paths for heat flow that bypass the envelope's thermal insulation. Most often this happens when a heat conducting material, for example, a metallic component is placed in the direction of heat flow and parallel to the thermal insulation. In other words, thermal bridging in buildings occurs when a highly conductive (or poorly insulating) material allows an easy pathway for heat flow across the envelope. The most common forms are probably within the eyesight of every reader of this chapter: for example, concrete or metal elements crossing the exterior building skin, structural framing members, perimeters of building openings, and architectural connections. Thermal bridging in buildings contributes to a multitude of problems, including, but not limited to, added energy use during heating and cooling seasons, interior surface condensation problems, mold growth, and material deterioration.

This chapter discusses several types of thermal bridges which can be spotted at various parts of the building skin and architectural interfaces. In order to minimize the resulting energy and durability problems, the proposed engineering solutions need to focus on design that minimizes thermal bridges and the resulting heat losses or gains, and, in many cases, to limit the risk of condensation. Sometimes, these measures occur in parallel with air-sealing efforts. For proper analysis of building thermal bridging, the heat conduction mechanism between adjacent components of the

exterior thermal envelope has to be examined. Materials with high thermal conductivity such as metals or concrete transfer significantly higher quantities of heat than materials with lower thermal conductivity such as wood or thermal insulations. Practical reduction of thermal bridges should be simple, inexpensive, and easy to apply. Several well-known general guidelines for mitigation of thermal bridges include separating highly thermally conductive materials with insulating materials, selecting less thermally conductive materials at the onset of the design, and reducing surface area in contact between highly conductive materials. This chapter will discuss analytical and numerical methods used in the analysis and elimination of thermal bridges. In addition, major building standards regulating this subject are listed.

Definition of Thermal Bridges and Common Consequences of Thermal Bridges

In general, a thermal bridge is defined as a location within the building envelope or architectural component, which exhibits considerably more heat flow than surrounding areas. Thermal bridging refers to the loss of energy by conduction through elements that "bridge" the insulation of a wall or roof enclosure of a conditioned space. While it is a common knowledge that most structural materials contribute to thermal bridging, it is good to remember that, in reality, thermal shorts occur much more often than just when a conductive element passes through or bypasses a thermal insulation barrier. Considering that in building components with temperature gradients, heat flow has multidimensional character and taking into account the complexity of today's building envelopes, these thermal pathways often require three-dimensional numerical analysis.

From the envelope designer's perspective, a thermal bridge is an element or location with missing thermal insulation, insufficient insulation, or reduced insulation performance, relative to adjacent areas of the thermal envelope. The intensity of a thermal bridge depends on differences in the apparent thermal conductivity of neighboring materials. In this light, even thermal insulation materials can generate a thermal bridge effect. As shown in Figure 4.10.1, high-resistance cellular plastic insulation can generate a thermal bridge effect. The infrared image of the building insulated with a use of the vacuum insulation panels (VIPs) packed in the foam casing shows significantly different temperatures in places of VIP connections (Kośny et al., 2014). Bear in mind that foam has about 10%

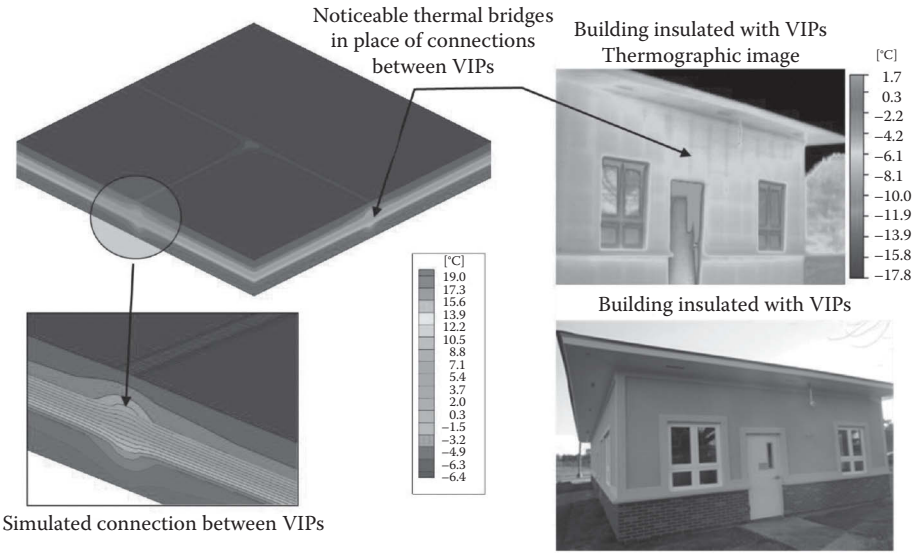


FIGURE 4.10.1 Thermal bridge effect generated by plastic foam surrounding a VIP (left), infrared imaging of a house insulated with VIPs—evidence of thermal bridges in places of VIP connections (right).

of the thermal resistivity than that of the core of a VIP. So, if VIPs are packed with the plastic foam, it is a thermal bridge. The consequences of poorly selected connections between envelope components can produce severe thermal bridging. Taking into account that interface details can have an impact of as much as 50% of the overall wall area for some wall systems, the whole-wall *R*-value (thermal resistance) can be as much as 40% less than what is measured for the clear-wall section (Kośny and Desjarlais, 1994; Christian and Kosny, 1995). Local heat losses through wall interface details may be twice that estimated by design calculations that focus only on the clear wall. Poor interface details also may lead to condensation of water that causes stains and dust markings on the interior finish, thus revealing thermal shorts in an unsightly manner. Moist surface areas can accelerate the propagation of molds and mildews, which can lead to degraded indoor air quality (Hens, 2012). In addition, thermal bridging affects the dynamic thermal performance of a building (Kośny and Kossecka, 2000).

It is good to remember that the exterior skin of a building provides a separation between the interior space and the exterior environment. In a perfect scenario, this separation should be as continuous as possible. This obligation applies to the building thermal insulation as well as to the air and moisture barriers. In the case of discontinuities in the thermal insulation layer, thermal bridges provide unwanted paths of least resistance for heat to move across the building envelope. The following are unwanted effects of thermal bridging:

1. Increased energy consumption: Due to the thermal outflow at the thermally bridged areas, heat is lost from every room (in heating dominated climates) or unwanted heat gains (in cooling dominated climates) resulting in a significant increase in whole-building energy consumption.
2. Mold formation: In cold climates, interior wall temperatures can drop well below the dew point. This leads to condensation, deterioration of plaster and paintwork and conditions for harmful mold formation. Finally, if there is sustained exposure to condensation, the building can experience serious fungi grow that leads to health hazards.
3. Compromised living space thermal comfort: Cold surface temperatures cause uncomfortable living space for occupants.

It is good to remember at this point, that in most of the buildings, even high-performance buildings, the existence of limited thermal bridging is unavoidable, mostly for structural reasons. However, good design and construction can reduce the number of thermal bridges and notably mitigate their effects.

Types of Thermal Bridges and Most Common Locations

The most common flaws in building envelopes are discontinuities in the insulation layer or air/moisture barriers. These problems are caused by improper design, incorrect assembly, or occur over the building's lifetime when the design does not provide adequate consideration of changes due to wind pressures and vibrations. Common imperfections in the thermal insulation barrier include both discontinuities in the insulation layer, installation of insulation in air cavities allowing air circulation around or across the insulation, compression of insulation, and other arrangements which decrease effectiveness. Envelope performance is then compromised by the increased heat-transfer rate. These assemblies usually show increased potential for moisture-related durability issues. Keep in mind that, local condensation and mold formation can be often caused by structural thermal bridges (Trechsel and Bomberg, 2009; Hens, 2012). For example, due to the compromised performance of thermal insulation, during the winter, interior temperatures of the adjacent interior space can drop well below the dew point. This may lead to water condensation on the interior surface and sometimes to deterioration of the interior finish. Considering common electrical penetrations in the exterior walls, it is very likely that water condensation can take place inside the wall cavities

causing a decrease in thermal performance of thermal insulation. Wet insulation, for example, conducts heat by as much as 10 times the design value for dry insulation, e.g., see Langlais et al. (1983), McFadden (1988), and Kumaran (2006). If there is sustained exposure to water condensation inside the building envelopes as well as on the internal surfaces, the building may become a subject to serious material deterioration.

Thermal Bridges Generated by Building Geometry and Architectural Details

Thermal bridges most commonly occur at locations such as corners, overhangs, wall/floor, wall/roof, wall/internal interfaces, and window or door perimeters. Component connections are often high-strength, high-conductivity elements that serve to hold or connect building structural or architectural components. Examples include balconies, wall connections with terraces, perimeters of wall openings, window, and curtain-wall mullions. Geometric thermal bridges do not form literal thermal shorts in the way as, for example, construction thermal bridges do. For instance, more intense heat transfer can occur in the areas where nominal insulation thickness and insulation continuity are maintained. Usually, geometric thermal bridges can be found in the areas where the area of the external building surface is significantly different from the corresponding internal surface area. Identifiable thermal bridge effects occur at corners which accentuate two-dimensional heat-flow paths. As depicted in Figure 4.10.2, some corner constructions may also lead to discontinuities in the envelope insulation and/or the air barrier. Geometric thermal bridges are often unavoidable. However, the energy impact of geometric thermal bridging increases with the complexity of the building form. Therefore, from the thermal and energy standpoint, as well from a long-term durability perspective, it is good to keep the building form as simple as possible.

Thermal Bridges Characteristic of Building Materials and Construction Subsystems

Thermal bridging from heat flow through nonuniform construction elements occurs in two or three dimensions. These thermal bridges are localized and repetitive along the length of the structure. Examples of this type of thermal bridge are provided by metal connectors in concrete sandwich panels as shown in Figure 4.10.3 and small connectors in curtain walls. In high-performance buildings with thick foam insulation installed to provide high *R*-value walls or roofs, the foam sheathing is often connected to the structural substrate using long and relatively thick metal connectors. In addition to being expensive, these connectors often seriously compromise the overall system thermal and hygrothermal performance (Kośny et al., 2013). In longer term, these thermal shorts often

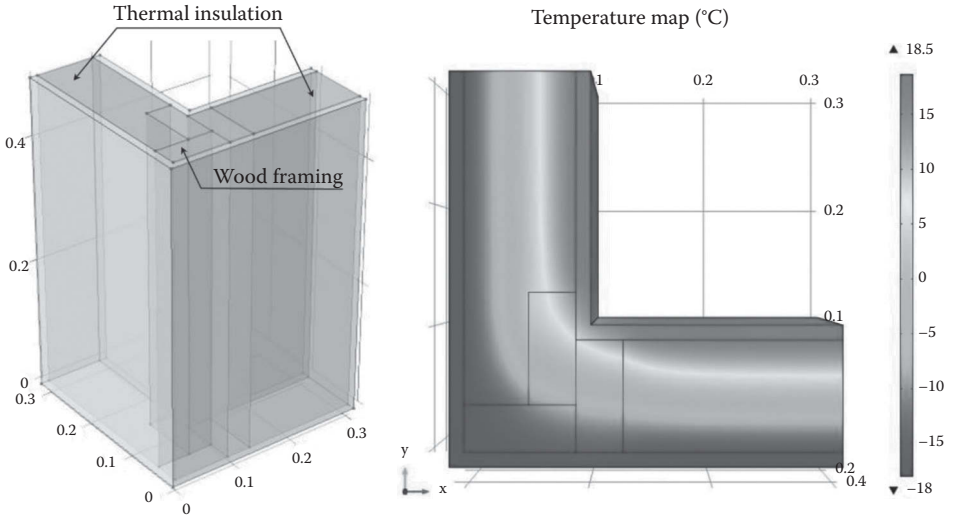


FIGURE 4.10.2 Thermal insulation discontinuity in the corner design using wood studs.

cause durability problems, which are commonly reported for these locations, and which are typically caused by local moisture condensation on thermally bridged cold metal surfaces.

As depicted in Figure 4.10.4, concrete masonry units (CMUs) have different insert designs but high thermal conductivity concrete bridges remain due to two or three concrete webs joining the external and internal concrete skins. Concrete block either empty or filled with thermal insulation will exhibit thermal bridging. The thermal bridge effect can be reduced with interlocking inserts that create a continuous layer of insulation.

Structural Thermal Bridges due to Construction

Structural thermal bridges due to construction are regularly spotted in buildings. This type of thermal bridge is not only the easiest type to recognize, but also easy to comprehend and visualize.

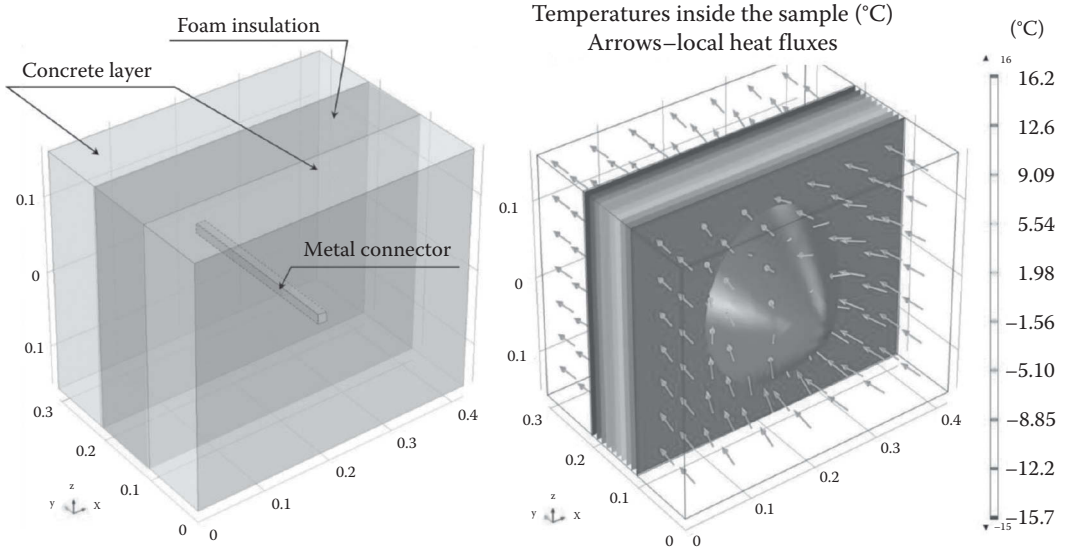
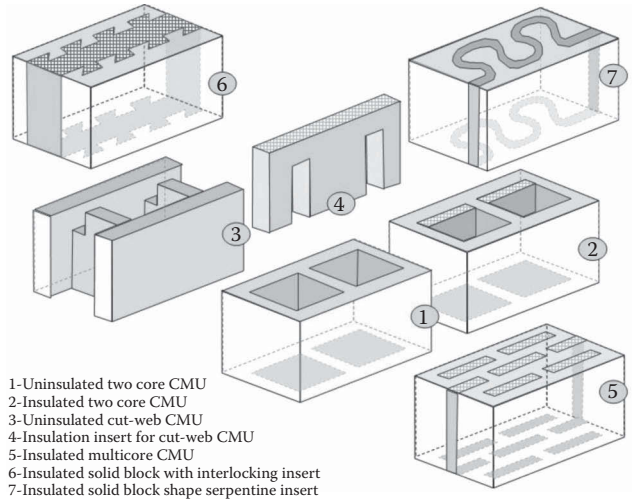
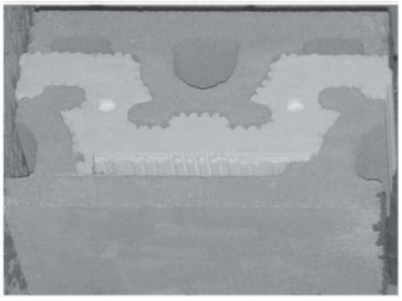


FIGURE 4.10.3 Thermal bridge from metal connector used in concrete-foam-concrete panel.





Example of a concrete block with interlocking foam insert

FIGURE 4.10.4 CMUs and common types of thermal insulation inserts.

Construction thermal bridges occur wherever there is a material or component that conducts heat significantly better than thermal insulation and this material passes through the thermal insulation layer. Structural thermal bridging is usually associated with components like steel or concrete combined with discontinuities in the thermal insulation.

Some examples of heat flow paths that result from the design on the building structure are follows:

- Reinforced concrete or steel frame beams and columns
- Roof rafters or floor joists that pass through the thermal envelope
- Wood, concrete, or metal framing members crossing the thermal insulation zone
- Cantilevered structures passing through or anchored in the thermal envelope
- Lintels and headers that interrupt envelope thermal insulation
- Structural connectors between metal-foam-metal sandwich panels.

Figure 4.10.5 shows the wall–floor intersections for the cases of the light-gage steel framed assembly and the structural penetration of the insulation system created by the concrete floor slab. They are both very common thermal bridge configurations, occurring in many commercial and residential buildings. In this case, both the concrete floor slab and the metal I-beam penetrate the vertical-wall insulation, significantly increasing the heat transmission. Significant disturbances in temperature fields can be observed in the case of the light-gage steel wall–floor intersections. The floor–wall connections are often the site of significant thermal bridging when the metal floor joists or concrete floor slabs pass through the wall insulation layer. The easiest fix for this problem is the use of a continuous layer of external thermal insulation. In the case of building retrofits, a pillow of spray foam is often used where the floor joists meet the wall bond beam (floor–wall intersections). It is not a perfect solution, but it helps in cutting thermal shorts, as well as reducing air movement through this detail.

The wall–roof connection is listed in engineering literature as a common location for thermal shorts and air leakage paths due to discontinuities between the wall and roof thermal and air or moisture barriers. The wall air and moisture barriers do not consistently extend to the roof deck

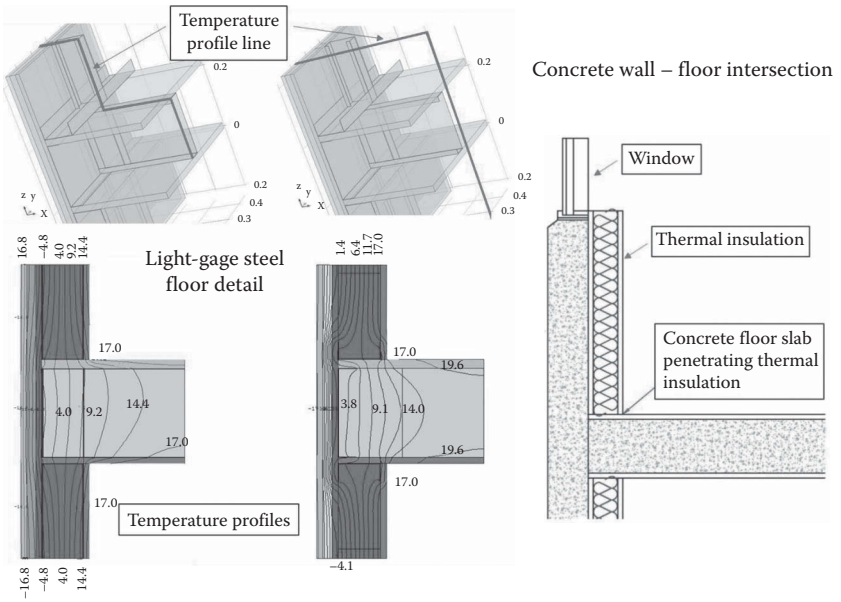


FIGURE 4.10.5 Thermal bridge in wall–floor intersections and the penetration of insulation by concrete construction thermal bridges should be recognized and avoided as much as possibly or minimized by careful design. Keep in mind that any construction thermal bridges will notably compromise local thermal and hygrothermal performance and contribute to the building energy use.

level, and the roof membrane is not always sealed to the impermeable materials in the wall. Instead, the membrane is often turned up at the roof edge, leaving a discontinuity in the envelope air barrier. Depending on the building envelope technology, more or less complex detailing and installation of continues thermal and air/moisture barriers are necessary. As described above, an application of sprayed foam helps in many cases.

Window-wall perimeters are associated with a number of probable thermal defects including highly conducting attachment methods, existing paths for air and moisture leakage, discontinuities in thermal insulation and insulation compression or displacement around the window frames, and placement of the window thermal break in the way that air can easily bypass. Another common mistake is assembly design in which the width of the window frame exposed to the outdoors is greater from the width exposed to the indoors. In cold climates, the greater area for heat loss to the exterior air causes the inner frame to be cold, increasing the potential for local condensation.

The mullions of curtain walls have long been considered as key sources of thermal shorts in commercial glazing systems. This comes from the fact that most curtain wall frames are made of aluminum which is about four times more conductive than steel. Mullions usually connect the exterior surfaces of the building enclosure to the interior, creating a significant thermal bridge. A variety of thermal breaks (added thermal insulation crossing the heat flow paths), which are typically 1–2 cm thick inserts made of less conductive plastic foams or rubbers, have become a characteristic component in modern curtain walls. These thermal breaks are typically located between the face plate and the structural part of the mullion, the rail, in line with the glazing pocket. Sometimes, thermal insulation can be added along the backside of the spandrel panel.* This type of assembly keeps interior and exterior sides of the mullion rail well thermally separated (National Institute of Building Sciences (NIBS), 2012).

Combined Thermal Bridges

In many cases, geometric thermal bridges also include an element of construction thermal bridging. For example, an external wall corner while being a geometric thermal bridge may also have structure creating construction thermal bridging. Similarly, ground floor and external wall junctions often involve a degree of construction thermal bridging. Very good examples of combined thermal bridges are corners with conducting studs or wall intersections with internal partitions. Figure 4.10.6 shows a construction of the corner using light gage steel studs as structural members. In this case, increased heat flow through this architectural detail is caused by the heat transfer through the metal studs, as well as by the corner shape effect (different heat exchange areas on the interior and exterior surfaces).

Durability Consequences of Thermal Bridging

As stated in earlier sections, thermal bridging may lead to condensation in building enclosures. Almost all building materials are susceptible to durability problems associated with increased water content (Sedlbauer, 2001; Kumaran, 2006; Hens, 2012): (i) wet thermal insulation may deteriorate and poorly perform (wet insulation is more conductive), (ii) uncontrolled high moisture content in concretes, mortars, gypsum, or brick will cause material deterioration, (iii) wet bio-based building materials makes great food for fungus and as a result the materials will deteriorate, and (iv) steel components will corrode. In cold climates, one of the most common locations for moisture-related problems is at wall-window perimeters. The positioning of skylights, doors, windows, window walls, curtain walls, and other fenestration within the wall or roof structures may provide a short circuit most often caused by direct connections of the window frame with highly conducting structural components. One of the key recommendations in regard to the window placement is to keep the window frame within the wall thermal insulation plane, e.g., Kośny et al. (2014). Since this recommendation is not always possible due to design or structural reasons, the other option is an application of local thermal insulation inserts, or use of windows with highly insulated frames.

* http://aanb.org/wp-content/uploads/2016/08/Glass-and-Aluminum-Curtain-Wall-Systems.pdf

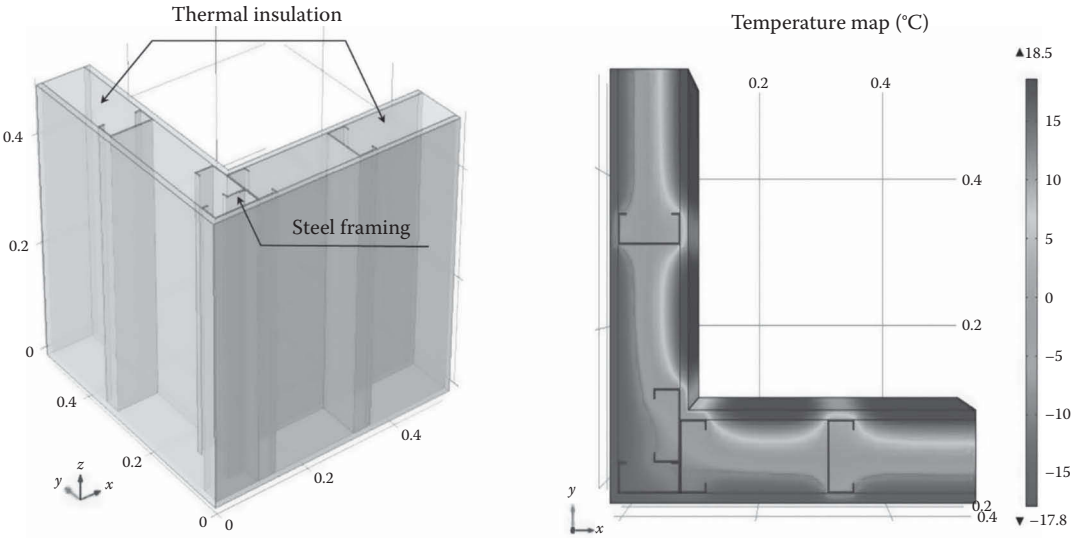


FIGURE 4.10.6 Thermal bridge in corner with light gage steel framing.

Bear in mind that in commercial buildings, a lot of façade shading and decoration details use numerous conductive components; including lintels, anchorage, and hangers that bypass the thermal barrier and produce thermal shorts, which often result in condensation on interior surfaces. Structural-steel framing that extends from the inside to the outside not only makes waterproofing and air barrier details difficult, but also forms a large thermal bridge. Finally, a great number of commercial buildings are constructed today using light gage or heavy gage steel structural profiles. This fact requires special attention during the design of the exterior envelope, because any design mistake coupled with highly conductive metal elements may (most often it will) cause serious thermal shorts and resulting durability problems.

As stated above, thermal bridges may compromise the overall building energy performance, reduce the local thermal and hygrothermal performance of the exterior envelope, and introduce unwanted building durability problems and health risks. Table 4.10.1 summarizes the most typical design problems associated with thermal bridging.

Engineering Methods for Thermal Bridge Analysis

Heat transfer through envelope components is complex and dynamic. The steady-state thermal performance of building envelope technologies and architectural details can be established using engineering calculations, thermal simulations, or through different types of thermal tests. Typically this takes the form of either *R*-value (thermal resistance) or *U*-factor (thermal transmittance) characterizations. Keep in mind that the direction and magnitude of heat flow in building envelope components are affected by solar gains from the sun, outdoor temperature, indoor temperature, and exposed surface area.

R-Value and U-Factor

Thermal resistivity is a material property that is defined as the reciprocal of the apparent thermal conductivity, k_a. Thermal resistance or *R*-value is defined as the thickness of a specimen divided by the apparent thermal conductivity. In either case, these terms are a measure of how effectively the material resists heat flow. Thermal resistance is identified by the term *R*-value with units, $m^2 \cdot K/W$, in SI units or, $ft^{2.\circ}F \cdot h/Btu$, in imperial units. *R*-values are used in the construction industry to describe insulation effectiveness and characterize steady-state heat flow across assemblies. The insulation *R*-value does not describe the overall performance of an assembly. For overall assembly

TABLE 4.10.1
List of Most Typical Design Problems Associated with Thermal Bridging

Typical Consequences	Description of Common Problems Associated with Thermal Bridges
Heat losses	Thermal bridges are the weakest points in the thermal envelope and so, in cold climates, they can contribute considerable heat losses. This reduces the overall building energy efficiency and increases heating costs
Unwanted heat gains	In cooling-dominated and mixed climates, thermal bridges allow unwanted solar heat gains. Roofs and attics are typical building envelope areas which may create intense solar gains if incorrectly designed
Cold internal surfaces	In winter, an intense heat transfer in a thermal bridge area may cause the internal surface temperature to drop, creating a cold spot. The surface relative humidity will thereby increase, yielding a risk of condensation and potential for mold growth. In cold climates, metal window frames in commercial buildings are notorious sources of condensation problems caused by highly conductive materials and insufficient thermal break (Curcija et al., 2013)
Discoloration marks on the interior and exterior surfaces	Depends on climatic conditions, intense thermal bridging (e.g., in case of metal framing) may cause internal or external cold spots, which can yield local condensation. On some surfaces, this may lead to the development of unaesthetic discolorations (e.g., so-called ghosting on the internal gypsum board surfaces in case of steel framed buildings[a])
Risks to thermal comfort and health issues	Cold internal surfaces are uncomfortable to be near and can cause thermal discomfort and perception of draughts. In addition, local cold surfaces within building envelopes usually cause moisture condensation, which in longer term can yield fungus grow, affecting the indoor air quality. It has been well documented in the literature[b, c] that both moisture and mold growth can lead to building occupants health problems (EPA, 2013).[d]

[a] http://www.steelframing.org/PDF/FinalDesignGuideSept82008.pdf.
[b] http://www.cdc.gov/niosh/topics/indoorenv/mold.html.
[c] https://iaqscience.lbl.gov/dampness-risks-house.
[d] http://ec.europa.eu/health/opinions/en/indoor-air-pollution/l-3/8-building-dampness.htm.

performance, the overall thermal resistance of the assembly which is determined from the resistances of the components of the assembly must be determined. The assembly *R*-value of a multilayer structure with one-dimensional heat flow is the sum of the thermal resistances of the layers:

$$R = \sum_{i=1}^{n} \frac{t_i}{k_{ai}} = \sum_{i=1}^{n} R_i \tag{4.10.1}$$

where k_{ai} is the apparent thermal conductivity of the material (layer) i and t_i is the thickness of layer i. R_i is the thermal resistance of the Ith layer.

In a similar way, *U*-factor (thermal transmittance) is used to calculate heat flow across an assembly. It can be calculated as the reciprocal value of the sum of the thermal resistance and the air-film resistance at the exterior and interior surfaces (h_e and h_i). The quantities $1/h_i$ and $1/h_e$ are often referred to as air-film resistances:

$$U = \frac{1}{\frac{1}{h_e} + R + \frac{1}{h_i}} = \frac{1}{R_e + R + R_i} \tag{4.10.2}$$

The *U*-factor in Equation 4.10.3 allows calculation of the rate of steady-state heat flow across an assembly. Most thermal insulations and a great majority of other building materials exhibit heat flow equally in both directions, i.e., these materials are isotropic in this regard. However, some

construction elements, for example, wood, some cellular plastics, reflective insulations or panels using highly conductive materials have different thermal resistances and heat transfer rates in different directions thereby exhibiting anisotropy:

$$Q = U \cdot A \cdot T \tag{4.10.3}$$

Q is the heat flow in Btu/h or W, A is the cross-sectional area (constant in this case), and ΔT is the temperature difference between the exterior air and the interior air.

There are a number of methods for determining the thermal performance of building envelopes, and the appropriate method depends on the types of construction and thermal insulation. The ASHRAE Handbook of Fundamentals (ASHRAE, 2013) provides several methods to determine clear-wall *R*-values and *U*-values. These methods are for assemblies with one- or two-dimensional heat transfer.

Parallel Path Method

A commonly used approach is *the parallel path method*. This method assumes that heat flows independently through regions that are in parallel and there is no heat exchange between regions as shown in Figure 4.10.7.

Considering that some people are more comfortable working with the concept of thermal transmission (U) than thermal resistance (R), the equations are given below for both *R*-values and *U*-factors. The parallel-path calculations yield relatively accurate results for one-dimensional heat transfer, where there is little or no heat exchange between the "parallel" paths in the actual assembly. The parallel-path calculation is useful in cases of simple sandwich constructions or wood-frame assemblies:

$$Q = \left(A_1 / R_{T1} + A_2 / R_{T2} + A_3 / R_{T3} + \ldots\right)\Delta T \tag{4.10.4}$$

$$Q = \left(U_1 \cdot A_1 + U_2 \cdot A_2 + U_3 \cdot A_3 + \ldots\right)\Delta T$$

where

Q is the heat flow through a defined area with multiple adjacent assemblies.

R_{Ti} is the thermal resistance for assembly "*i*", generally obtained by summing the resistance of each layer of material in the assembly including inner and outer air films

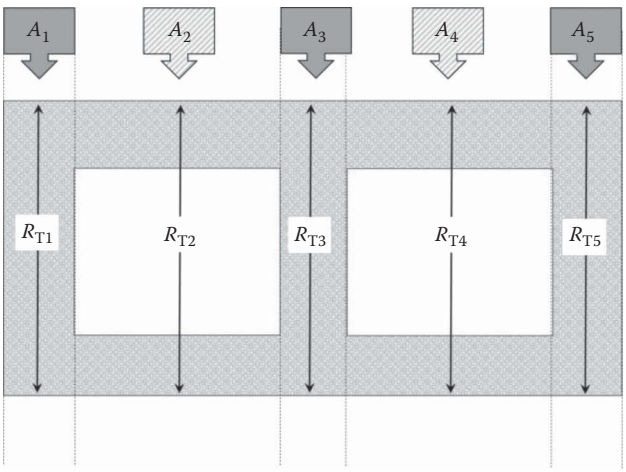


FIGURE 4.10.7 Calculation schematic for the parallel path method.

U_i is the heat transmission coefficient in BTU/(h·ft²·°F) for assembly "*i*" including the effect of the interior and exterior surface films
A_i is the area of assembly "*i*", in ft²
ΔT is the difference between the indoor and outdoor air temperatures.

Effective *R*-value and *U*-factor can be computed using the following equation for a region like that in Figure 4.10.7:

$$R_{\text{eff}} = 1/U = \left(A_1 / R_{T1} + A_2 / R_{T2} + A_3 / R_{T3} + \cdots\right) / A_{\text{total}} \tag{4.10.5}$$

Isothermal Planes (Series–Parallel Path) Method

In buildings with highly conductive structural members or other components made of concrete, steel, aluminum, or glass, the assumption of parallel heat flow is not likely to provide reliable results. In these cases, the *isothermal planes method* provides an alternate procedure for calculating the *U*-value. The name comes from an assumption of uniform temperature (isothermal) for the highly conductive planes separating less conductive layers. This method is used, for example, to analyze heat flow through wood-framed assemblies or masonry walls. The isothermal planes method divides the construction assembly into a series of layers. Hollow masonry units are a very good example of when this calculation method can be applied. The solid webs connecting the face shells are quite conductive compared to the air spaces in the hollow cores, and the face shells conduct heat laterally. As shown in Figure 4.10.8, for a two-core masonry unit, the layer containing the webs and cores is treated with a parallel-path calculation to arrive at an average *R*-Value for the layer. This is then added to the *R*-values of the two face shells in series. In the isothermal planes method, the air-to-air *R*-value is calculated using the following equation:

$$R_{\text{air-to-air}} = \frac{1}{U} = R_e + R_{fe} + \left[\frac{a_w}{R_w} + \frac{a_c}{R_c}\right] + R_{fi} + R_i \tag{4.10.6}$$

where
R_e and R_i are exterior and interior air-film resistances
R_{fe} and R_{fi} are exterior and interior face resistances
a_w and a_c are fractions of the total heat flow area for the webs and air-cores
R_w and R_c are *R*-values (calculated for the space between faces) for all webs and air-filled cores.

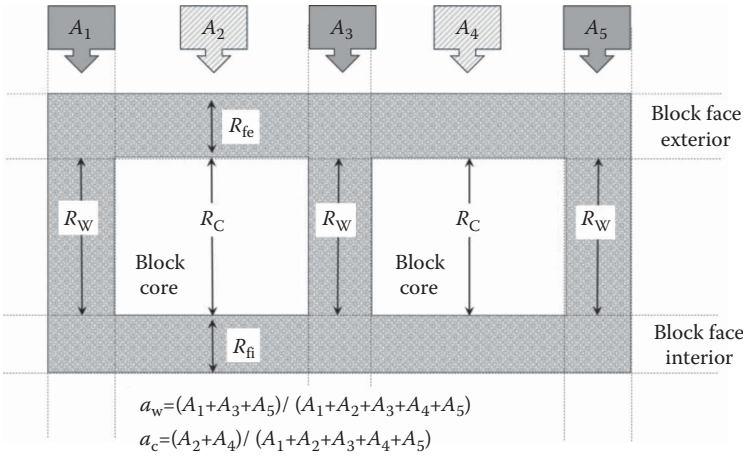


FIGURE 4.10.8 Calculation schematic for the isothermal plane method.

Modified Zone Method

Heat flow through construction assemblies with metal framing or sheathing is more complex and requires special consideration. For walls using light gage steel framing, the ASHRAE Handbook of Fundamentals recommends the use of the ASHRAE Zone Method or modified zone method (ASHRAE, 2013). The modified zone method was designed to improve the accuracy of clear-wall *R*-value calculations for light gage steel-framed walls with insulated cavities containing steel C-shape studs with solid webs (Barbour et al., 1994; Kośny et al., 1995). The modified zone method is similar to the parallel path method and the ASHRAE Zone Method. All three methods are based on parallel-path calculations—the difference being how heat flow through the metal stud (*w* in the figure below) is estimated. In the parallel path method, the "zone of influence *w*" is assumed to be equal to the length of the stud flange—"*L*". In the ASHRAE Zone Method, *w* is assumed to be equal to the length of stud flange plus twice the total thickness of all finish material layers on the thicker side (can be exterior or interior)—as depicted in Figure 4.10.9. In the modified zone method, *w* is determined by the following equation:

$$w = L + z_f \times d_t \tag{4.10.7}$$

where

- w—the zone containing the metal stud
- L—metal stud flange width
- z_f—zone factor determined based on simulations of different steel-stud wall configurations. It depends on the ratio between thermal resistivity of finish material and cavity insulation, depth of stud, and thickness of finish material layers
- d_t—total thickness of all wall finish material layers on the thick side.

The accuracy of the modified zone method was verified by finite-difference modeling of over 200 metal frame walls with insulated cavities (Kośny and Desjarlais, 1994). The modified zone method results are within ±2% of the results of the three-dimensional finite-difference evaluations. Furthermore, for six steel-framed wall assemblies tested at the Oak Ridge National Laboratory and 15 similar wall assemblies tested by Holometrix (Barbour et al., 1994), the results of the modified zone method calculations differed by less than 6.5% from test results—ASHRAE Research Project 785-TRP (1996).

Linear Thermal Transmittance Method

The *linear thermal transmittance method* is a superposition of the local heat flows to calculate the overall heat flow in the building envelope assembly. The heat flows for the full assembly are determined. The methodology for determining the linear thermal bridging coefficient is specified in

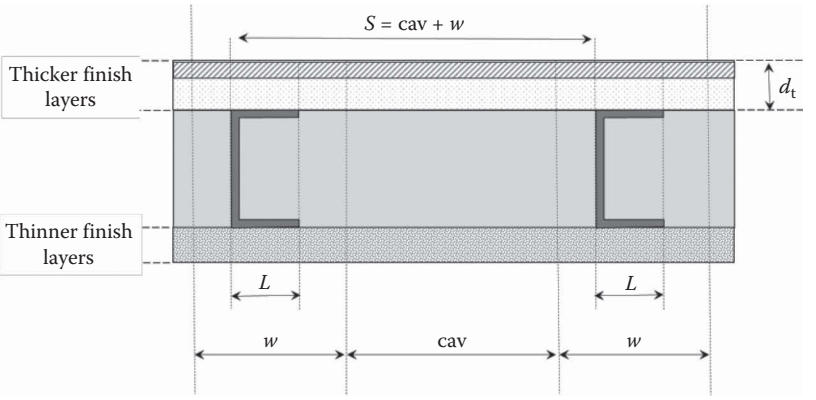


FIGURE 4.10.9 Calculation schematic for the modified zone method.

several ISO standards, including (see ISO 10211, 2007 and ISO 14683, 2007), which deal with methods for determining the heat loss through linear thermal bridges at junctions of building elements. These ISO documents contain specifications for developing thermal models of thermal bridges, which can be utilized for determination of heat losses and surface temperatures—ISO 10211 (2007). The following types of thermal bridges are typically considered:

- Linear thermal bridges: where there is a thermal bridge with a specific length, for example, a lintel.
- Point thermal bridges: where there is a thermal bridge at specific points only, for example, masonry wall ties.
- Repeating thermal bridges: where there is a thermal bridge that repeats at regular intervals within an element of the thermal envelope, for example, timber studs in an insulated wall.
- Nonrepeating thermal bridges: where there is a one-off thermal bridge, for example, a structural column in an insulated wall.

In the linear thermal transmittance method, each construction detail containing a thermal bridge is represented by a heat flow. The thermal effect of the bridge is characterized by its linear thermal transmittance, which is expressed by Ψ [W/m^2·K]. The Ψ-value represents the additional heat flow through the area of the linear thermal bridge, which is in addition to the heat flow for the adjacent, thermally undisturbed area. Two types of thermal bridges are specified in this method: linear and point thermal bridges. A linear thermal bridge is one with a uniform cross section along one of the three orthogonal axes—ISO 14683 (2007), ISO 10211 (2007). It is commonly found in intersections and structural junctions. Examples of point thermal bridges can be visualized by a metal fastener penetrating an envelope assembly, shelf angles, slab edges, balconies, corner framing, parapets, and window interfaces.

A linear thermal bridge is defined mathematically as follows:

$$\psi = L^{2D} - \sum_{j-1}^{N_j} U_j * L_j \tag{4.10.8}$$

where

L^{2D} is the thermal coupling coefficient obtained from 2D simulations—see: ISO 10211 (2007)
U_j is the thermal transmittance of the 1-D component, "*j*," separating the two environments being considered
L_j is the length over which the ψ-value applies.

The linear thermal bridge coefficient can have a negative or positive value. If ψ-value for a specific building component has a negative value, this indicates that the thermal resistance of building envelope has been improved due to the introduction of that component. A positive ψ-value indicates that the presence of the component of interest has reduced the thermal resistance of the building envelope.

The overall thermal resistance $R_{overall}$ [m^2·K/W] of any type of building envelope can be computed with the use of the following equation (assuming that there is only one type of clear field envelope assembly within the area considered for calculations):

$$R_{overall} = \frac{A}{\frac{A}{R} + \sum_{j=1}^{J} L_j \psi_j + \sum_{k=1}^{K} n_k \chi_k} \tag{4.10.9}$$

where

A is the total area of considered assembly [m^2]
R is the thermal resistance of the assembly without thermal bridge effects [m^2·K/W]
L_j is the length of linear thermal bridge [m]
Ψ_j is the linear thermal transmittance of linear thermal bridge [W/(m·K)]

n_k is the number of the point thermal bridge components
χ_k is the point thermal transmittance of the point thermal bridge "k" [W/K].

For building envelope assemblies containing multiple clear-field configurations "i", Equation 4.10.8 has the following form:

$$R_{\text{overall}} = \frac{A}{\sum_{i=1}^{I} \frac{A_i}{R_i} + \sum_{j=1}^{J} L_j \psi_j + \sum_{k=1}^{K} n_k \chi_k} \tag{4.10.10}$$

where
A_i is the area of considered part of the assembly without thermal bridge effects [m^2]
R_i is the thermal resistance of considered part of the assembly without thermal bridge effects [m^2·K/W].

The same calculation methodology is used in ISO standard 6946 (2007) for calculation of the transmission heat transfer coefficients H [W/K], where the contribution due to the thermal bridges is included according to the following equation:

$$H = \sum_{i=1}^{I} \frac{A_i}{R_i} + \sum_{j=1}^{J} L_j \psi_j + \sum_{k=1}^{K} n_k \chi_k = \sum_{i=1}^{I} A_i U_i + \sum_{j=1}^{J} L_j \psi_j + \sum_{k=1}^{K} n_k \chi_k \tag{4.10.11}$$

where
H is the object's total heat transfer coefficient [W/K]
U_i is the thermal transmittance for each considered part of the assembly without thermal bridge effects [W/m^2·K].

Figure 4.10.10 shows basic types of thermal bridges and describes the heat-transfer calculations using the linear transmission method.

Numerical Tools Used in Thermal Performance Analysis

Thermal bridging and construction details in the building envelope impact the thermal performance. These features have both steady-state and dynamic repercussions. Capability for reduction of building loads and successful whole-building integration are dependent, in large scale, on correct predictions

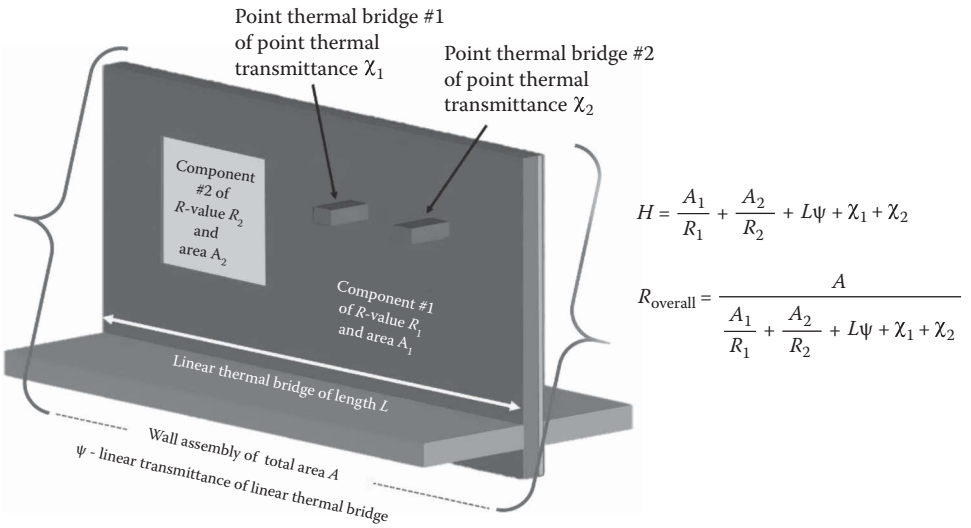


FIGURE 4.10.10 Types of thermal bridges and the linear transmission method.

of building thermal characteristics with building envelopes being one of the most important factors. In other words, building integration and load reductions are only possible when combined with accurate estimations of building envelope-generated thermal loads. Thermal and whole-building energy modeling tools are often useful in determining the local thermal performance as well as the overall energy efficiency of buildings. The energy efficiency for the building enclosure is typically based on individual *R*-values of various building envelope components. The effects of thermal bridges are approximated in these calculations. However, the currently available whole-building energy modeling tools do not always accurately account for the real *R*-values of various interface configurations, such as window-to-wall, wall to partition, wall to foundation, and wall to interfaces (Kośny and Desjarlais, 1994; Kossecka and Kośny, 1998). It is even more difficult to consider thermal bridges formed by steel framing. In these cases, two- and three-dimensional heat transfer programs need to be used to determine local *R*-values; the overall energy analysis can then be modified to consider these effects and arrive at an accurate picture of the overall energy efficiency of the building enclosure.

A great variety of numerical tools for this type of simulations are available and used by practitioners for two-dimensional thermal analysis—Comsol,[*] Heat3[†] AnTherm,[‡] BISCO (for 2D) and TRISCO (for 3D),[§] PSI Therm,[¶] ANSYS Fluent,[**] and PHYSIBEL Voltra.[††] THERM is a free software developed by the Lawrence Berkeley National Laboratory (LBNL), which is commonly used in North America for this purpose.[‡‡] Reports from the European 2010 project funded by the Communities Intelligent Energy Europe Programme[§§]; "Assessment and Improvement of the EPBD Impact (ASIEPI)" lists 29 numerical tools and 26, mostly national, thermal bridge atlases which can be useful in thermal bridge analysis.

Building Performance Standards Dealing with Building Thermal Bridging

In North America, there are three key energy performance standards focused on buildings:

- ASHRAE 90.1 Energy Standard for Buildings Except Low-Rise Residential Buildings-ASHRAE 90.1 (2013)
- ASHRAE 189.1 Standard for the Design of High-Performance Green Buildings Except Low-Rise Residential Buildings—ASHRAE 189.1 (2014)
- National Energy Code of Canada for Buildings 2011 (NECB, 2011)

In Europe, thermal bridge calculations are based on two key ISO standards-ISO 10211 (2007), ISO 6946 (2007) "Thermal bridges in building construction—Heat flows and surface temperatures—Detailed calculations," and "Building Components or Building Elements" which serves for:

- The calculation of minimum (lowest) surface temperatures in order to assess the risk of surface condensation and
- The calculation of heat flows in order to predict overall heat loss from a building (for the constant, steady-state flow case; i.e., time-independent temperature distribution) and
- Determination of linear and point thermal transmittance and surface temperature coefficients (of thermal bridges).

[*] https://www.comsol.com/heat-transfer-module?gclid=CJOBzYmq38wCFUFehgod_MgAHA.
[†] http://www.buildingphysics.com/index-filer/Page691.htm.
[‡] http://www.antherm.at/antherm/EN/.
[§] http://www.physibel.be/index.htm.
[¶] http://www.psitherm.uk/.
[**] http://www.ansys.com/Products/Fluids/ANSYS-Fluent.
[††] http://www.physibel.be/v0n2vo.htm.
[‡‡] https://windows.lbl.gov/software/therm/therm.html.
[§§] https://ec.europa.eu/energy/intelligent/projects/sites/iee-projects/files/projects/documents/asiepi_access_the_results._en.pdf.

The above specifications include the geometrical boundaries and subdivisions of the model, the thermal boundary conditions, and the thermal values and relationships to be used. The second important standard ISO 6946 (2007) "Building Components or Building Elements-Calculation of Thermal Transmittance" is also available. According to this standard, calculation of heat transfer coefficients of parallel plane surface building components shall be performed based on one-dimensional calculations. For all other more complex heat flow cases with multidimensional heat transfer, the standard ISO 10211 (2007) "Thermal Bridges in Building Construction" requires the implementation of numerical methods.

Following is a list of ISO building standards which can be helpful in the analysis of thermal bridges and building design:

- ISO 7345 "Thermal insulation-Physical quantities and definitions"
- ISO 10456 "Building materials and products-Hygrothermal properties-Tabulated design values and procedures for determining declared and design thermal values"
- ISO 13370 "Thermal performance of buildings-Heat transfer via the ground-Calculation methods"
- ISO 13786 "Thermal performance of building components-Dynamic thermal characteristics-Calculation methods"
- ISO 13788 "Hygrothermal performance of building components and building elements-Internal surface temperature to avoid critical surface humidity interstitial condensation-Calculation methods"
- ISO 13789 "Thermal performance of buildings-Transmission and ventilation heat transfer coefficient-Calculation method"
- ISO 14683 "Thermal bridges in building construction-Linear thermal transmittance-Simplified methods and default values."

Methods of Thermal Bridge Mitigation

As stated earlier, in addition to increased building space-conditioning energy consumption, the heat flow through thermally bridged areas often leads to durability problems and material deterioration, which are frequently associated with the local condensation (which is often a direct product of a thermal bridge). Therefore, thermal bridges need to be correctly characterized and mitigated during the building design stage. The choice of construction method dictates how insulation continuity is approached. Components that form the principal insulation layers should be clearly identified on drawings, and their details developed to ensure continuity of these layers between elements of the construction.

Numerous design strategies are available to help ensure a building design contains as few as possible thermal bridges. In terms of thermal bridging, several general "rules of thumb" have been developed. According to Ziegel (2010), they are as follows:

- The *Avoidance Rule* which states that discontinuities in the insulating layer should be avoided
- The *Breakthrough Rule* which says that any breaks in the insulating layer must have as high an insulating value as possible, and
- The *Geometry Rule* states that building geometry should be as simple as possible and corners in building intersections and architectural details should have a rounded shape, if possible

From the practical standpoint, thermal bridging can be avoided by placing insulation in series with a conductive material, instead of installing it in parallel. For example, you can place thermal insulation outside a framed wall or roof assembly instead of placing the insulation between the structural members. Thermal bridging can also be avoided by examining the elevation area for the lowest local

R-value points and improving them through mitigation of thermal shorts. For instance, replacing metal facade connectors with plastic or composite elements.

From the practical standpoint, consider that simpler building geometry and simpler configurations of architectural details and intersection are more likely to be designed and built correctly. That is why the number of different types of construction technologies within the thermal envelope should be minimized, because thermal imperfections are most likely to occur where one type of construction meets another. Also, some construction details include areas that interrupt the insulation layer. These can be often found in areas of intersections of vertical and horizontal structural components. Although it is difficult to completely eliminate these thermal shorts, their effects can be minimized wrapping the conducting element with thermal insulation (even a thin one) and possibly by overlapping the insulating layers in the vicinity of conducting elements, even though these insulations will not necessarily be in direct contact. Local applications of high-performance insulations, such as aerogel and vacuum panels, are strongly recommended in places with limited spaces for conventional foam or fiber insulations.

In addition, all of the following building details need particular attention during the design process, because they are all essential to limiting heat losses through thermal bridging:

- External wall/roof intersection,
- Connections with partition walls that pass through the insulation plane of the ground floor,
- Perimeters around window and door openings (headers, sills, jambs, and lintels),
- External wall/ground floor junctions,
- Wall/floor intersections,
- Internal and external corners in external walls,
- External wall/balcony intersections,
- Chimneys placed in external walls,
- Service boxes and piping/power line passages located in external walls.

REFERENCES

ASHRAE (2013). *2013 ASHRAE Handbook of Fundamentals*. Chapter 27: Heat, air, and moisture control in building assemblies: Examples, ASHRAE, Atlanta, GA.

ASHRAE 90.1 (2013). Energy standard for buildings except low-rise residential buildings. Retrieved May 27, 2016, from ASHRAE.org: https://www.ashrae.org/resources--publications/bookstore/standard-90-1.

ASHRAE 189.1 (2014). Standard for the design of high-performance green buildings. Retrieved May 27, 2016, from ASHRAE.org: https://www.ashrae.org/resources--publications/bookstore/standard-189-1.

Barbour, E., Goodrow, J., Kośny, J., and Christian, J. E. (November 21, 1994). Thermal performance of steel-framed walls, prepared for The American Iron and Steel Institute by NAHB Research Center.

Christian, J. E., and Kosny, J. (1995). Towards a national opaque wall rating label. *ASHRAE, BETEC, U.S. DOE VI Thermal Envelope Conference*, Clearwater, FL, December 1995.

Curcija C., Goudey, H., Mitchell R., and Dickerhoff, E. (2013). Highly insulating window panel attachment Retrofit. Report prepared for the General Services Administration by Lawrence Berkeley National Laboratory-Windows and Envelope Materials Group.

EPA. (2013). Moisture control guidance for building design, construction and maintenance. U.S. Environmental Protection Agency.

Hens, H. S. (2012). *Building Physics—Heat, Air and Moisture: Fundamentals and Engineering Methods with Examples and Exercises.* 2nd edition, Wiley, New York.

ISO 6946. (2007). Building components and building elements—Thermal resistance and thermal transmittance—Calculation method. Retrieved May 27, 2016, from iso.org: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40968.

ISO 7345. (1987). Thermal insulation: Physical quantities and definitions. Retrieved May 27, 2016, from iso.org: http://www.iso.org/iso/catalogue_detail.htm?csnumber=14024.

ISO 10211. (2007). Thermal bridges in building construction—Heat flows and surface temperatures—Detailed calculations. Retrieved May 27, 2016, from iso.org: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=40967.

ISO 10456. (2007). Building materials and products—Hygrothermal properties—Tabulated design values and procedures for determining declared and design thermal values. Retrieved May 27, 2016, from iso.org: http://www.iso.org/iso/catalogue_detail.htm?csnumber=40966.

ISO 13370. (2007). Thermal performance of buildings—Heat transfer via the ground—Calculation methods. Retrieved May 27, 2016, from iso.org: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40965.

ISO 13786. (2007). Thermal performance of building components—Dynamic thermal characteristics—Calculation methods. Retrieved May 27, 2016, from iso.org: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40892.

ISO 13788. (2012). Hygrothermal performance of building components and building elements—Internal surface temperature to avoid critical surface humidity interstitial condensation—Calculation methods. Retrieved May 27, 2016, from iso.org: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=51615.

ISO 13789. (2007). Thermal performance of buildings—Transmission and ventilation heat transfer coefficient—Calculation method. Retrieved May 27, 2016, from iso.org: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40894.

ISO 14683. (2007). Thermal bridges in building construction—Linear thermal transmittance—Simplified methods and default values. Retrieved May 27, 2016, from iso.org: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40964.

Kośny, J., and Desjarlais, A. O. (1994). Influence of architectural details on the overall thermal performance of residential wall systems. *Journal Thermal Insulation and Building Envelopes*, 18(1), 53–69.

Kośny, J., Desjarlais, A. O., and Christian, J. E. (1995). Thermal performance of "energy efficient" metal stud wall systems. *ASHRAE, BETEC, U.S.DOE VI Thermal Envelope Conference*, Clearwater, FL, December 1995.

Kośny, J. and Kossecka, E. (2000). Computer modeling of complex wall assemblies—Some accuracy problems. *The International Building Physics Conference*, Eindhoven, The Netherlands, September 18–21.

Kośny, J., Fallahi, A., and Shukla, N. (2013). Cold climate building enclosure solutions. U.S. DOE Building America Program Report, NREL/SR-5500-55875, http://www.nrel.gov/docs/fy13osti/55875.pdf.

Kośny, J., Misiopecki, C., Fallahi, A., Shukla, N., DuPont, W. C., and Carbary, L. D. (2014). Thermal design of window-wall interface in wall energy retrofits using high-performance vacuum insulation. *ASHRAE Transactions*, 120(SE-14-008), 90.

Kossecka, E., and Kośny, J. (1998). Effect of insulation and mass distribution in exterior walls on the dynamic thermal performance of a whole buildings. *DOE, ASHRAE, ORNL Conference-Thermal Envelopes VII*, Clearwater, FL, December 1998.

Kumaran, M. K. (2006). A thermal and moisture property database for common building and insulation materials, NRCC-45692, NRC, Institute of Research in Construction, Canada.

Langlais, C., Hyrien, M., and Klarsfeld, S. (1983). Influence of moisture on heat transfer through fibrousinsulating materials. ASTM STP 789 project report, ASTM International, West Conshohocken, PA.

McFadden, T. (1988). Thermal performance degradation of wet insulations in cold regions. *Journal Cold Regions Engineering*, 2, 25–34.

NECB. (2011). National Energy Code of Canada for Buildings 2011. Retrieved May 27, 2016, from nrc-cnrc.gc.ca: http://www.nrccnrc.gc.ca/eng/publications/codes_centre/2011_national_energy_code_buildings.html.

NIBS. (2012). Building envelope design guide—Curtain walls. National Institute of Building Sciences (NIBS), Washington DC, June 25, 2012.

Sedlbauer, K. (2001). *Prediction of Mold Fungus Formation on the Surface of and Inside Building Components*. Stuttgart: Fraunhofer Institute for Building Physics.

Trechsel, H. R., and Bomberg, M. (2009). *Moisture Control in Buildings: The Key Factor in Mold Prevention*. West Conshohocken, PA: ASTM International.

Ziegel. (2010). Thermal bridges. Retrieved April 30, 2014, from www.ziegel.at: http://www.ziegel.at/gbc-ziegelhandbuch/eng/ressourcen/energie/prim/waermebruecken.htm.

4.11 COMPRESSORS

Christian K. Bach, Ian H. Bell, Craig R. Bradshaw, Eckhard A. Groll, Abhinav Krishna, Orkan Kurtulus, Margaret M. Mathison, Bryce Shaffer, Bin Yang, Xinye Zhang, and Davide Ziviani

Introduction

Compressors for air and other gases are used for many purposes in industrial, commercial, and residential applications. For example, gas compressors in industrial applications are used for refrigeration, air conditioning, heating, pipeline conveying, natural gas gathering, catalytic cracking, polymerization, and other chemical processes. While the gas compressor may be similar in appearance to an air compressor commonly seen at a construction site, they are specially designed to meet a chemical process requirement of pressure, temperature, or flow, and are capable of handling the specific gases or gas mixtures to be used. Rollins (1989) provides a detailed compilation of the available knowledge about air and gas compressors, compressed air systems, pneumatic tools, and the uses of compressed air [1].

This section contains a review, including descriptions of certain design features and characteristics, of the most widely used and commercially available compressors. Two major applications are considered: (1) compressors that are commonly used in vapor compression cycles for refrigeration, air conditioning, and heat pumping applications and (2) compressors that are commonly used to compress air. The most complete set of technical papers about compressors can be found in the Proceedings of the Purdue Compressor Conferences. References [2–15] contain 10,558 pages (1377 papers).

In refrigeration, air conditioning, and heat pumping systems, the compressor is one of the four major components that are used to implement the widely used vapor compression cycle. The other three components are the condenser, evaporator, and expansion device. The compressor compresses the refrigerant and circulates it in a continuous cycle through the closed system. ASHRAE provides a detailed compilation of the available knowledge about refrigerant compressors [16]. The compressors for these applications have been chosen for this section because they are representative of most of the compressors made in large quantities.

Compressed air is essential to almost all manufacturing and construction enterprises. The compressor manufacturing industry is so large that air compressors have been selected for special consideration in this section.

The capacity of gas and air compressors is specified according to their delivered flow rate, whereas the capacity of refrigerating compressors is specified according to their cooling capacity.

There are two basic types of compressors: positive displacement compressors and dynamic compressors. Positive displacement compressors increase the pressure of the refrigerant vapor, air, or other gas by reducing the volume of the compression chamber by means of work usually applied to the compressor's shaft mechanism. Dynamic compressors increase the pressure by a continuous transfer of angular momentum from the rotating parts of the compressor to the fluid, followed by the conversion of this momentum into a pressure rise. Both types are described in this section.

Positive Displacement Compressors and Their Application to Refrigeration, Air Conditioning, and Heat Pumping

Positive displacement compressors depend on a volume change for the compression process, in contrast to dynamic compressors, which depend on acceleration, deceleration, and conversion of kinetic energy into pressure. They are designed for a wide range of flow rates and suction and discharge pressures, and can be used in many applications. Typically, they are used for refrigeration, air conditioning, and heat pumping applications, and for compressing air and other gases for virtually all uses except aircraft propulsion.

When used for cooling and heating, compressors, whether classified as positive displacement compressors or dynamic compressors, are used in vapor compression cycles for applications of air conditioning, refrigeration, and heat pumping. As the key component in vapor compression systems, the compressor raises the pressure of the working fluid so that it can be condensed to liquid at the heat sink temperature, expanded or throttled to reduce the pressure, and evaporated into vapor to absorb heat from the refrigerated source. In these systems, refrigeration compressors are a component in a closed-loop system, which is tightly sealed (often hermetically) to prevent refrigerant loss. These systems are often sealed for life without service or lubricant changes, operating anywhere from 300 to 8000 h per year with a life expectancy of 5 to 15 years, i.e., operating from 10,000 to 120,000 h with no maintenance.

When used in air conditioning, refrigeration, and heat pumps, they must be designed for a variety of refrigerants and compatible lubricants. Commonly used refrigerants include hydrofluorocarbons such as R-410A, R-407C, R-404A, and R-134a, ammonia, carbon dioxide, and various hydrocarbons.

Positive displacement compressors are classified according to their construction, e.g., whether or not they are enclosed in a hermetic housing. They are also classified according to the kinematic arrangement of their parts to accomplish the change in volume.

The particular types of refrigerant compressors to be described in this section are identified according to how they fit into the range on the application chart. This chart is shown in Figure 4.11.1, which shows the application across the top line, from the smallest for household uses to the largest for possible use in freezing plants and for air conditioning tall buildings [17]. Typically, compressors are specified by the horsepower or the kilowatts needed to drive them, or their refrigerating capability measured in tons of refrigeration, Btu/h, or kW.

Modifications to Conventional Vapor Compression Cycles

Vapor compression systems employing pure refrigerants or mixtures of refrigerants have reached a plateau in system optimization, motivating the search for novel approaches for designing systems with higher levels of efficiency and capacity. Two of such approaches that have a direct impact on the design of compressors are described next.

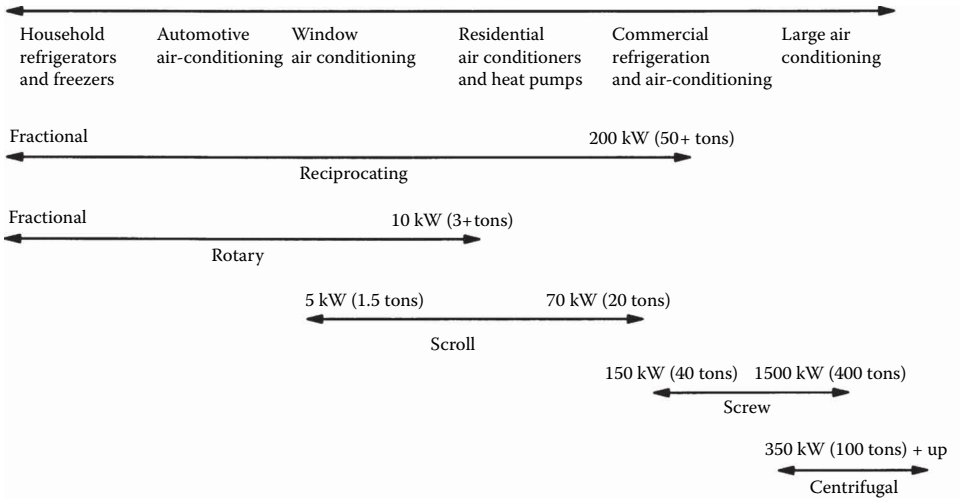


FIGURE 4.11.1 Range of applications of refrigerating compressors. (Data from International Institute of Refrigeration, Paris.)

Liquid Flooded Compression

One means to achieve higher levels of system efficiency is the addition of a significant amount of oil to the refrigerant before compressing it in the compressor, and the addition of an internal heat exchanger after the condenser, as can be seen in Figure 4.11.2 [18]. The ideal flooding liquid has a high density, a high specific heat, extremely low volatility, and good compatibility with the compressor materials. Water meets the first two criteria, but oils are more commonly used as they can better maintain lubrication, while still having reasonable capacity to cool the working fluid as it is being compressed.

This modification to the standard four-component vapor compression system has the following impacts on system design and performance:

- Significant reduction in discharge temperature of the compressor due to cooling by the injected oil. This is particularly relevant for high heat of compression fluids such as the refrigerant blend R404A or refrigerant R32.
- Addition of a large amount of oil allows for improved compressor efficiency, as has been demonstrated experimentally for scroll compressors [19,20]. Experimental results show a monotonic increase in efficiency with increased oil injection rate.
- While the flooding oil takes up a fraction of the suction swept volume of the compressor, it is not usually more than a few percent by volume due to its high density, and this consumption of the suction volume is partially offset by a decrease in internal leakage.
- The addition of the internal heat exchanger increases system efficiency by increasing the latent heat of cooling in the evaporator.

The addition of a large amount of oil flow into the compressor in vapor compression systems offers the dual promise of better compressor performance and an improvement in system efficiency, particularly in large-temperature-lift applications.

Vapor Injected Compression

Another means to achieve higher levels of system efficiency is the use of vapor injected compression with economizing. In a conventional compressor, the compression process takes place in a single process. During this process, the refrigerant heats up which increases the required compression work. A vapor injected compressor uses the injection of saturated vapor to reduce the temperature of the refrigerant during the compression process. Figure 4.11.3 shows conventional and vapor

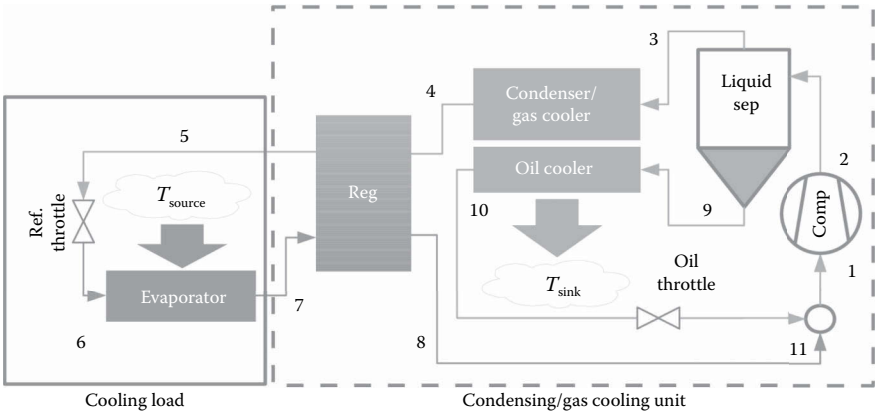


FIGURE 4.11.2 Schematic of the liquid flooded compression (Reg: Internal heat exchanger). (Adapted from Bell, I. H., *Theoretical and experimental analysis of liquid flooded compression in scroll compressors*, Ph.D. Dissertation, Purdue University, 2011. With permission.)

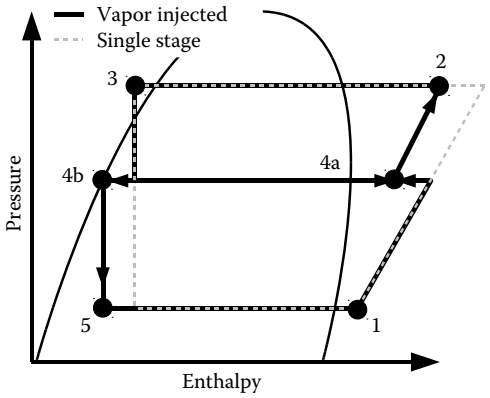


FIGURE 4.11.3 Vapor injected and conventional process in p–h diagram.

injected cycles in a log(p)–h diagram. The refrigerant that is injected into the compression process is taken from the expansion process, either using a vapor separator as shown in Figure 4.11.4 or an economizer. This vapor injection has three effects: (1) it reduces the discharge temperature, which increases the range of the operating envelope; (2) it decreases the inlet quality to the evaporator, reducing the evaporator pressure drop, especially at low evaporation pressures; and (3) it increases the total mass flow rate of the compressor, since vapor injection is typically done after the suction pocket of the compressor closes. This leads to a larger capacity at low suction pressures, which is of advantage for heat pump applications. This increase of capacity can be reduced by bypassing the injected vapor (e.g., point 4a) to the suction side of the compressor. However, this will also reduce the heating COP of the system.

Vapor injection is employed primarily with scroll compressors, which offer the unique opportunity for locating the injection port within a specific angular location of the scroll. Stationary and moving scroll form compression pockets in which volume decreases (and pressure increases) as they move along the compression path. As the compression pockets move past a specific angular position they swipe a certain pressure range, allowing the implementation of vapor injection without the need for implementing actively controlled injection valves. Besides from scroll compressors, vapor injected compression has also been employed in screw compressors and spool compressors, which offer similar characteristics.

Vapor injected compression can be applied to similar applications as cascade cycles, e.g., when large pressure ratios are needed for low-temperature cooling or cold climate heat pump applications. Vapor injected compression can—in principle—be more cost-effective since it does not require two compressors with the associated oil management system. However, vapor injected compressors have the disadvantage of poor performance at low compression ratios due to reexpansion losses and do

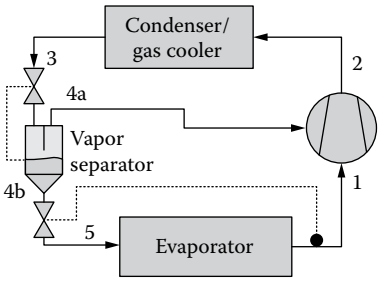


FIGURE 4.11.4 Single-port vapor-injected compression cycle using vapor separator.

not have the option to cost-effectively change the overall compression ratio by bypassing compression stages (as one can do with cascade cycles).

Due to better cycle coefficient of performance for high-pressure ratios than single-stage systems at a larger capacity for low ambient temperatures—at a lower cost than conventional cascade type systems—it is expected that vapor injected compression will eventually become adapted by the industry for low-temperature heat pump applications.

Compressor Type by Construction

Hermetic Compressors

Hermetic compressors, used for many refrigeration applications, have both the electric motor and the compressor inside a welded steel housing, providing a true hermetic seal. The motor shaft is normally connected directly to the compressor shaft. The housing is a steel shell that is welded at the joints to ensure the hermetic seal. If the gas in the housing is suction gas, then the compressor is called a low-side compressor because of the low suction pressure. If the gas in the housing is discharge gas, then the compressor is called a high-side compressor. A high-side design can often be very advantageous by serving as a source of high-pressure oil in the sump to lubricate the compressor components. When oil is injected into or present during the compression process for cooling or sealing, the high-side design provides a cavity where the oil can be separated from the discharge gas before it is delivered from the compressor to the condenser. Advantages of the low-side compressor are improved motor and oil cooling and lighter housing designs because of lower pressure. Hermetic shells are limited to small sizes because of the high cost of producing the larger ones made by a deep-drawing process. Therefore, hermetic compressors are used for smaller capacities. Figure 4.11.5 shows a hermetic reciprocating compressor with part of the shell cut away.

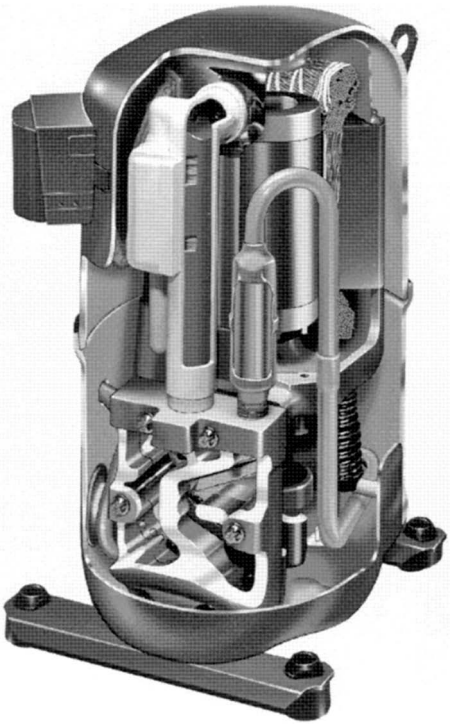


FIGURE 4.11.5 Cutaway drawing of a hermetic reciprocating compressor. (Courtesy of the Copeland Corporation, Sydney, OH.)

Semi-Hermetic Compressors

Semi-hermetic compressors are typically used for medium cooling capacities. They are similar to hermetic compressors except that the outer housing is normally bolted together so that the compressor is accessible for service. An advantage of this design is the ease of construction of multicylinder, large compressors without the need for the expensive deep-drawing manufacturing process. The externally bolted joints are sealed with gaskets that are a source of refrigerant leakage over a long period of time. Figure 4.11.6 shows a cross-sectional drawing of a typical semi-hermetic reciprocating compressor with part of the housing cut away. The electric motor drive is cooled by suction gas as well as by conduction through the motor housing. In many applications, fans are used to force air over the motor housing to aid in cooling. Figure 4.11.7 shows that the semi-hermetic reciprocating compressor provides excellent efficiency as compared to other positive displacement compressors in the medium capacity range. Semi-hermetic reciprocating compressors are normally rugged and durable machines; however, they tend to have relatively high vibration and noise levels.

Open Drive Compressors

Open drive compressors have their motors external to their compressor housings. The shaft connecting the motor and compressor passes through the compressor housing, requiring a shaft seal. The open drive compressor is more sensitive to leakage than the semi-hermetic compressor due to the use of the shaft seal. Open drive compressors are usually used for larger capacities, and often for automotive applications where the power to drive them is taken from the engine.

The semi-hermetic and open drive types make it easier to have multiple cylinders, which is a popular option for larger capacities and higher pressure ratios. The open drive compressors offer more flexibility, especially when repairs are needed. Installing the larger sizes in tandem is a low-cost way to obtain multiple capacities for improved part load efficiencies.

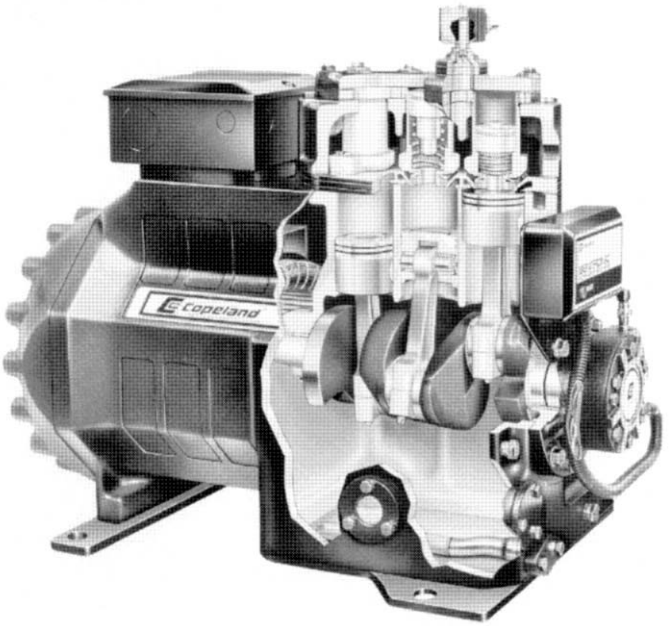


FIGURE 4.11.6 Cutaway drawing of a semi-hermetic reciprocating compressor. (Courtesy of the Copeland Corporation, Sydney, OH.)

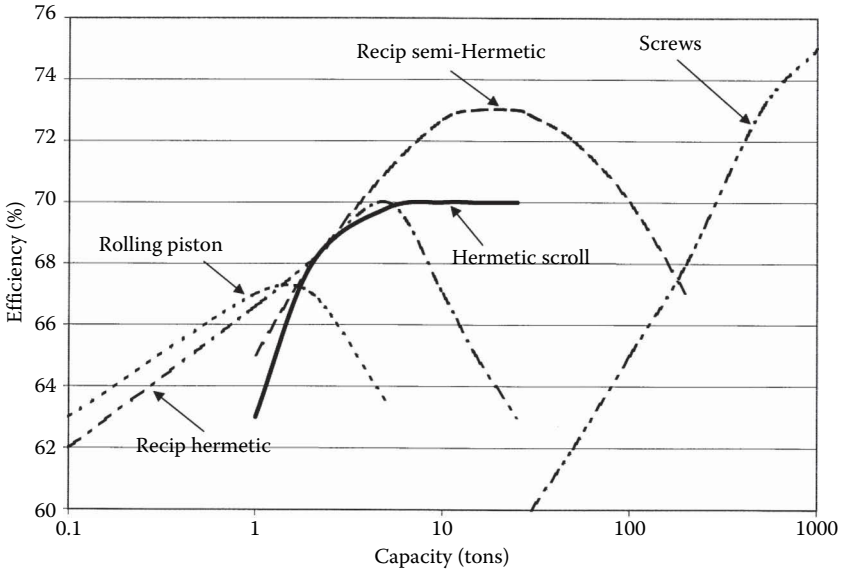


FIGURE 4.11.7 Performance and capacity range for positive displacement refrigeration compressors. (Courtesy of the Copeland Corporation, Sydney, OH.)

Compressor Type by Kinematic Linkages

Reciprocating Compressors

The reciprocating compressor is the most widely used positive displacement compressor, although the other types described here are being produced in ever-increasing numbers.

Most reciprocating compressors are single-acting machines that use a piston inside a cylinder. The piston is usually driven by a crankshaft through a connecting rod and wrist pin to compress the gas. However, other driving mechanisms are used sometimes, e.g., scotch-yoke mechanisms, electromagnetic drives, or swash-plate, wobble-plate, and Sanderson rocker arm motion mechanisms for axial multipiston arrangements. Figure 4.11.8 shows a cross-sectional drawing of a typical reciprocating compressor. The motor is often cooled by suction gas, although conduction through the shell also plays a part in the cooling. In many larger applications, fans are used to force air over the motor housing to aid cooling. Discharge and suction mufflers are used to reduce gas pulsations and sound. Pulsations are mainly caused by the reciprocating action of the piston and the intermittent opening and closing of the suction valve. Gas pulsations are very important in reciprocating compressors because these have a negative effect on performance and noise characteristics. Sometimes mufflers are an internal component in the compressor. A well-designed muffler can reduce pulsations and improve reciprocating compressor performance. Reciprocating compressors are normally mounted on a spring suspension system external to the housing and/or sometimes inside a hermetic shell to help reduce vibration and noise.

The reciprocating compressor has single or multiple cylinders. Every piston–cylinder setup has suction and discharge valves. The valves operate by differential pressure where gas only runs through them in one direction. Reciprocating compressors operate by drawing refrigerant through a suction valve into the cylinder during the downward stroke of the piston. At bottom dead center when the cylinder has reached its maximum volume and the piston starts its upward travel, the suction valve closes. The gas is then compressed until it reaches the desired discharge pressure at which time a discharge valve opens, allowing the compressed gas to flow into the compressor's discharge plenum and then into the system.

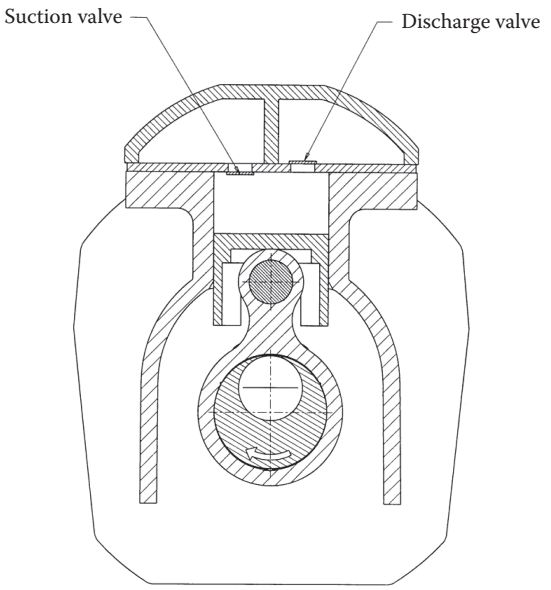


FIGURE 4.11.8 Cross-sectional drawing of a typical reciprocating compressor. (Courtesy of the Copeland Corporation, Sydney, OH.)

The volume inside the cylinder at top dead center is called the clearance volume. The clearance volume can never be reduced to zero, and is typically 1%–5% of the displaced volume due to piston movement. As the piston begins its downward travel, this gas must first be expanded back to suction conditions before the suction valve opens and new suction gas can enter the cylinder; thereby reducing the volumetric efficiency and capacity of the compressor. Single-stage reciprocating compressors are used in applications with pressure ratios ranging from 1.5:1 to 15:1, and two-stage units for applications with pressure ratios of 10:1 and above. At pressure ratios about 15:1, the capacity of high clearance volume compressors is reduced and can approach zero. A certain superheat is usually maintained at the inlet of the reciprocating compressor to avoid liquid entering into the in-cylinder volume which may result in the liquid slugging phenomenon. However, a further superheat in the suction plenum is unwanted as it reduces the compressor volumetric efficiency since the actual density of the gas flowing into the in-cylinder volume is decreased [21]. Figure 4.11.9 shows a typical *P*–*V* diagram of a reciprocating compressor, which indicates the relation between the gas pressure inside the cylinder and piston displacement. Process 1–2 indicates the suction process. Process 2–3 is the compression process during which the gas is compressed to a high pressure when the discharge valve starts to open. Process 3–4 is the process when the in-cylinder gas flows into the discharge plenum. There is some pressure fluctuation during this process due to the vibration of the discharge valve and gas pulsation in the discharge plenum. Process 4–1 is the reexpansion process when the piston starts to move downward and the gas begins expansion until the gas pressure inside the cylinder is low enough for the suction valve to open.

Both the heat transfer inside the suction plenum and the heat transfer between the in-cylinder gas and cylinder wall impact the compressor performance. The cylinder wall temperature changes little compared to its time-averaged value in a single cycle [22]. The heat transfer rate can be determined using developed correlations [23,24]. The heat transfer rate depends on the piston mean velocity and in-cylinder gas instantaneous temperature and pressure. The heat transfer between the gas and its surroundings when it enters into the cylinder and out of the cylinder during the discharge process also needs to be considered [25].

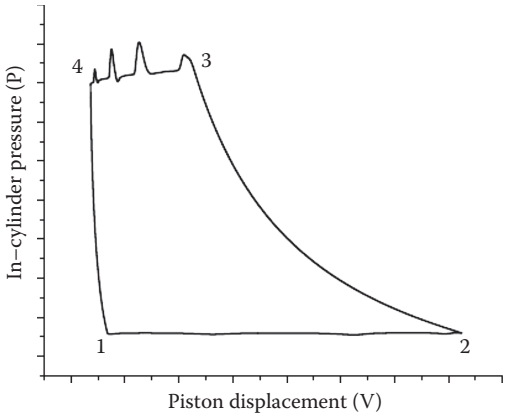


FIGURE 4.11.9 Typical *P–V* diagram of a reciprocating compressor.

Reciprocating compressors with piston ring seals can wear significantly but still provide outstanding performance. Therefore, even as wear increases, they often continue to operate adequately after long service. They are usually considered easy to repair compared to other types, and there is a large cadre of service personnel available. The piston rings in the reciprocating compressors can be lubricated by oil or nonlubricated using self-lubricating piston ring materials instead. The lubricating performance depends on not only the profile and dimensions of the piston ring, but also the properties of the lubricating oil. There is a minimum oil film thickness along the piston ring height between ring profile and cylinder wall which reaches the maximum magnitude when the piston moves in the middle of the downward/upward stroke. The friction between piston ring and cylinder wall gets worse when the piston approaches both the top dead center and bottom dead center due to the slow piston movement and in-cylinder gas pressure [26]. The series of crevice among the piston rings (usually the compressor has several piston rings) creates a path for the gas leaking from or into the in-cylinder volume. The gas pressures in these crevices are different and uniformly distributed. The first piston ring which is closest to the in-cylinder volume usually bears up to 75% of total pressure difference [27].

Capacity modulation can be accomplished in several ways in reciprocating compressors. On multicylinder compressors, some suction valves are often blocked. In some cases the suction process is throttled by passing discharge gas back to the suction side of the system. This type of capacity modulation method has the advantage of being infinitely controllable. However, from the energy utilization point of view, the power used to compress the bypassed gas from suction pressure to discharge pressure brings no beneficial results. In addition, the compressor always works with full load no matter what the required capacity is, i.e., the compressor is always running with high power consumption. There is an alternative method called stepless capacity control and was proposed by Hoerbiger compressor controls [28]. This method uses a pneumatically actuated finger-type unloader to keep the in-cylinder gas from compression in a portion of the stroke so that the discharged flow can be reduced to the partial flow. The actuation of the unloader is controlled by a panel by monitoring the flow requirements. Sometimes the speed of the compressor is varied by mechanical or electrical means, although designers must be careful with valve design to ensure that they work adequately over the range of speeds used. If a variable speed AC motor is used to change the compressor speed, the torsional response of the drive shaft needs to be considered. The capacity modulation of a reciprocating compressor can be related to its volumetric efficiency change which is not only depending on the pressure ratio, but also on the clearance volume. By controlling the auxiliary clearance volume, the volumetric efficiency is changed. This method is a more effective control method at higher pressure ratios. The auxiliary clearance volume can be fixed or variable. Varying the piston stroke by using a movable pin, an adjustable rod, or a linear electric motor can

also have a significant impact on the compressor capacity. According to the report by Danny et al. [29], using a linear electric motor to change the stroke is a very promising approach. In some reciprocating compressors such as swash plate compressors, the piston stroke can be varied by changing the inclined angle of the swash plate. Similarly, in the compressor using the novel Sanderson-Rocker Arm Motion (S-RAM) mechanism, the stroke is also variable by altering the inclined angle between the wobble plate and the machine driving shaft [30].

Hermetic reciprocating compressors for refrigeration and air conditioning applications are produced in sizes ranging from a fraction, up through 25 tons of refrigeration, and are typically used in residential, commercial, and industrial air conditioning, heat pump and refrigeration applications with capacities ranging from ¼ to over 200 tons of refrigeration. Reciprocating compressors may have up to 10 cylinders in radial, in-line, or multiple bank arrangements.

Computer-aided design has greatly improved performance and efficiency of reciprocating compressors, primarily due to new and innovative valve, port and flow passage designs [31]. Computer-aided design studies are also used to improve internal heat transfer, so that further improvements in performance and efficiency can be expected. However, it appears that other designs have potentially inherently superior characteristics for better reliability, volumetric efficiency, pressure oscillations, etc.

Rotary Compressors

Rotary compressors are produced in two general types: the rolling piston compressor and the rotating vane compressor. The most common, shown in Figure 4.11.10, is known as the rolling piston or fixed vane rotary compressor. It is commonly used in small refrigeration and air conditioning applications. The second type, the rotating vane compressor, is shown in Figure 4.11.11. It is most commonly used in automotive applications.

The *rolling piston compressor* is sometimes called a fixed (or stationary) vane rotary compressor because the vane does not rotate with the rotor, but rather slides in a fixed slot in the stationary part of the compressor. The rolling piston name is given to the cylindrical part of the compressor that is mounted on an eccentric shaft since it appears to roll on the surface of the cylinder wall. Further, its position is the controlling element for the compression volume that is formed by the rolling piston, the fixed vane, and the cylinder wall, thereby having a similar function as the piston in a reciprocating compressor. During the suction part of the cycle, suction gas is drawn through a suction port into the increasing volume just described. During the suction process, compression and discharge processes are taking place in the decreasing volume on the opposite side of the piston and vane. These compressors tend to be more reliable than reciprocating compressors primarily because no suction valve is needed and breakage of the suction valve has been a common failure mode.

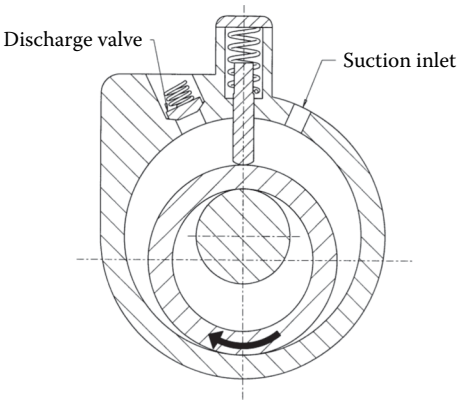


FIGURE 4.11.10 Cross-sectional drawing of a typical rolling piston (fixed vane) rotary compressor. (Courtesy of the Copeland Corporation, Sydney, OH.)

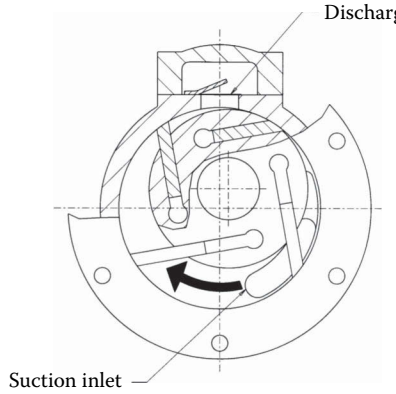


FIGURE 4.11.11 Cross-sectional drawing of a typical rotating vane rotary compressor. (Courtesy of the Copeland Corporation, Sydney, OH.)

Discharge flow is controlled by use of a pressure-operated discharge valve. Another advantage of rolling piston compressors is their small size for a given capacity. However, since rotary compressors are intolerant of contamination, they are usually limited to uses for which the systems are not expected to be opened for servicing. Furthermore, any wear in the compressor, whether it be the bearings, the shaft, the roller, or the vane, increases the clearances, and therefore, has a significant effect on the compressor performance.

To simplify the construction and lubrication system for rolling piston compressors, high side designs are common. However, the high motor and oil temperatures that accompany high side designs limit the pressure ratio (or condensing pressures) that can be achieved without extra cooling.

In air conditioning and refrigeration uses, rolling piston compressors are produced with capacities varying from 600 to 60,000 Btu/h (0.18–17.6 kW). Although there are mechanical means to vary the capacity of rolling piston compressors, variable-speed motors are normally used to achieve capacity modulation.

Pressure ratios range from 2:1 to 10:1. Performance data are shown in Figure 4.11.7. The rolling piston compressor can achieve comparatively high efficiencies, especially when fabricated for units with less than 24,000 Btu/h (7 kW) cooling capacity. The eccentric movement of the piston in the rolling piston is less than that of a piston in a reciprocating compressor of the same displacement. Thus, there is comparatively less vibration and noise compared to reciprocating compressors. However, the rolling piston mechanism is commonly counterweighted to mitigate the imbalance caused by the eccentric movement of the piston. The compressor and motor are normally rigidly mounted in the compressor shell and, therefore, it has relatively high torsional vibration because of the single-cylinder torque pulse.

To reduce the torque pulsations, especially at low speeds, compressors are produced sometimes with two out-of-phase cylinder assemblies. In twin rolling piston compressors, the two cylinder assemblies provide parallel flow paths, which can be useful for increasing the flow rate through the compressor. In two-stage rolling piston compressors, the gas flows through the two cylinder assemblies in series. Breaking the compression process into two stages reduces the pressure ratio acting across each chamber, which can be advantageous for reducing leakage in applications with large temperature lifts. Furthermore, the two-stage configuration allows for the implementation of modifications such as intercooling and economizing. Cooling the working fluid between stages can both increase the compressor efficiency and decrease the discharge temperature, thereby decreasing wear on the compressor.

The *rotating vane rotary compressor*, shown in Figure 4.11.11, has a rotor with slots that contain several vanes. The vanes slide in and out of the rotor and are centrifugally pressed against the

cylinder wall as the assembly rotates. The shaft center is eccentric from the center of the cylinder; therefore, the gas volumes in the pockets between the rotor, vanes, and cylinder vary as the unit rotates. Suction gas enters through a suction port. Discharge gas flow is commonly controlled with a discharge valve. The rotating vane compressor is subject to wear at several sealing surfaces, but is somewhat self-compensating for the wear. It can be more tolerant to contamination than the rolling piston compressor. This type of compressor has cooling capacities ranging from 12,000 to 120,000 Btu/h (3.5–35 kW). It is used primarily in automotive air conditioning applications.

Scroll Compressors

The scroll compressor is the newest of the various types of compressors to be used for the mass market. However, it is rapidly gaining market share. Figure 4.11.12 is a cross section of a typical hermetic scroll compressor. The steel shell is an integral part of the compressor–motor assembly. An internal spring suspension is not required because of the low vibration levels produced by the scroll mechanism. In the dominant conception of the scroll compressor, suction gas is normally used to cool the electric motor. Cooling is also accomplished with direct conduction from the motor to the shell and then to the atmosphere via radiation and natural (or forced) convection.

The compression mechanism consists of two matching spiral scroll members opposite each other as shown in Figure 4.11.13. These scrolls are typically identical, but can be asymmetric. One of the scrolls is typically fixed and the other, driven by a crankshaft, moves in an orbit within the fixed scroll without rotation. Figure 4.11.14 illustrates the process of compression in gas pockets formed by contact of the two scrolls at several radial points. The refrigerant gas enters the lateral openings at the outer diameter between the two scrolls. The lateral openings are then sealed as the orbiting scroll moves, forming two trapped gas pockets indicated by the shaded area at the end of the suction stroke. The gas is then compressed due to the decreasing volume of the pockets as the scroll continues its orbiting motion, moving the compressed gas toward the center of the scroll elements.

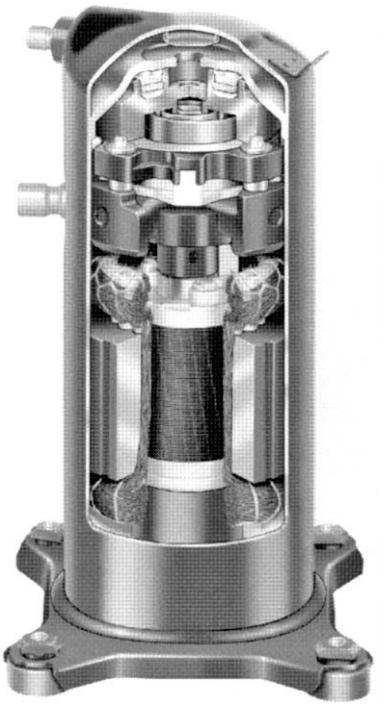


FIGURE 4.11.12 Cutaway drawing of a hermetic scroll compressor. (Courtesy of the Copeland Corporation, Sydney, OH.)

As the gas reaches the center, a discharge port is opened (or uncovered by the motion of the orbiting scroll) which allows the compressed gas to move to the discharge chamber of the compressor and into the system.

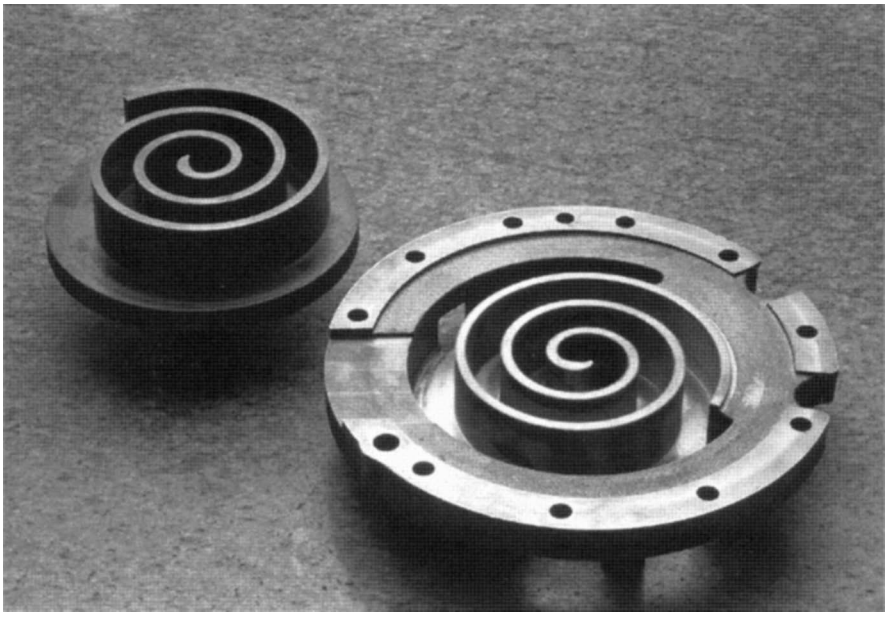


FIGURE 4.11.13 A pair of matching scroll members. (Courtesy of the Copeland Corporation, Sydney, OH.)

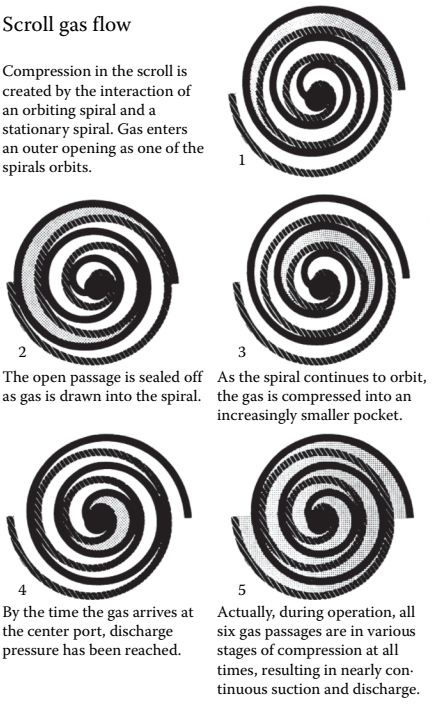


FIGURE 4.11.14 Gas flow in a scroll compressor. (Courtesy of the Copeland Corporation, Sydney, OH.)

The development of advanced machining techniques made high-volume manufacturing of scroll compressors possible. An Oldham coupling (or other analogous methods like the use of idler bearings) is required to keep the orbiting scroll in an orbiting, nonrotating, trajectory. The scroll compressor has a substantial reduction in part count and operational complexity as compared with other positive displacement compressors. The complexity of scroll compressors is typically in the geometry and the addition of compliant mechanisms, not the part count. Operationally, there are many inherent technical advantages with scroll compressors.

The discharge and suction processes in the scroll compressor are completely separated; therefore, there is no clearance volume effect on the volumetric efficiency and capacity of the compressor. As a result, the volumetric efficiencies of scroll compressors are almost independent of the pressure ratio for a given suction condition (i.e., suction density), due to the lack of a clearance volume effect. Thus, virtually all the refrigerant mass that enters the suction cavity of a scroll compressor eventually exits through the discharge port except a small amount lost due to leakage from higher-pressure chambers and heat transfer.

Heat transfer between the suction and discharge processes is also greatly reduced, therefore reducing the heat transferred to the suction gas and improving overall efficiency. The scroll compressor has a constant suction volume flow rate regardless of its operating pressure ratio and the volumetric efficiency remains high, above 90 in most applications. By placing a port in the scroll member, access can be obtained to a point in the compression process. This port can be used to inject vapor or liquid in the middle of compression, making it possible to compound the refrigeration cycle to take advantage of an economizer cycle to achieve high system efficiency and capacity, or to inject oil to cool the working fluid.

There are no valves or valve losses (aside from high-compression ratio applications where discharge valves are commonly employed), and the absence of valves improves reliability and reduces complexity. An almost constant steady flow and relatively low heat transfer gives the scroll compressor the capability of high efficiency. Being dynamically balanced and without valves allows the scroll to operate efficiently at speeds much higher (up to 10,000 rpm) than the rotary or reciprocating compressors, and over a broader range. These features, plus the nearly steady torque and smaller pressure oscillations, result in very low vibration and noise for scroll machines.

Capacity modulation is achieved by varying the speed of the compressor or by delaying the beginning of the compression process. This is done by opening a port that allows gas to flow back to the suction side during the initial compression process. Alternatively, in some scroll compressors, the scroll wraps are allowed to separate for part of the compression process, allowing for smooth capacity control.

By the use of a compliant mechanism, scroll compressors are very tolerant to ingesting liquid and dirt without failure or significant loss in performance. The wear of bearings and other parts does not affect the sealing and has almost no effect on the performance of the compressor. Compliance is accomplished by holding the two scrolls together by means of gas pressure and centrifugal forces, not the mechanical structure.

Scroll compressors are used in a wide range of applications from refrigeration, air conditioning, and heat pumping to air supply for fuel cells. Deployment in vapor compression cycles accounts for a large majority of the applications in air conditioners and commercial refrigeration equipment. Current production sizes range in cooling capacities from 1 to 25 tons of refrigeration (3.5–88 kW). Like the rotary compressors, scroll compressors are smaller in size than reciprocating compressors for the same capacity. They can be designed for the high-pressure ratios needed for large temperature lift heat pump applications.

The increasing market share of the scroll compressors is due to their high reliability, low noise, and several performance advantages with respect to reciprocating compressors.

1. In refrigeration applications, a smaller-displacement compressor can be used to provide adequate refrigerant capacity at very low evaporation temperatures or very high-pressure ratios. This allows the compressor to operate efficiently over a much broader operating range with the same motor.

2. In heat pump applications, the heating capacity at low ambient heating conditions and high-pressure ratios is higher, resulting in a higher heating seasonal performance factor, a measure of the performance of heat pumps.
3. In air conditioning applications, the system is normally sized to ensure adequate cooling at relatively high ambient temperatures. When operating at normal or lower ambient conditions, the capacity of the scroll compressor does not increase rapidly, resulting in less excess capacity. Therefore, there is a reduced load on the coils, which increases overall system efficiency, i.e., which increases the seasonal energy efficiency ratio (SEER). The SEER is a measure of the performance of an air conditioner.
4. Scroll compressors do not usually employ valves, improving their reliability as compared with reciprocating compressors.

Screw Compressors

Screw compressors are generally obtained from the gearing theory of different meshing pairs. They are available in two basic configurations, namely: the Lysholm (or twin screw) type with a pair of meshing rotors, and the single screw type with a central rotor meshing with two gate rotors.

The *twin screw compressor* type is the most common. It utilizes two intermeshing helical rotors, one a male and one a female to achieve the compression process. The male rotor, which is generally the driving rotor, and the female rotor are enclosed in a casing. As their meshing progresses during the rotation, a volume is isolated between the rotors and the casing. Such volume represents the working chamber which is reduced progressively during the compression process. Most of the twin screw compressors present a 4/6 configuration with four lobes in the male rotor and 6 lobes in the female rotor. Both rotors have the same outer diameter. By modifying the number of lobes in the male and female rotors different compression ratios can be achieved as well as the possibility to be used in dry or oil-flooded applications [32].

The suction and discharge processes occur without the use of valves. In particular, when the working chamber at the front ends of the rotors passes across the suction port, the filling process begins until the cut-off point at which the suction volume has a maximum value. At the end of the compression process, the working chamber reaches the rear ends of the rotors and it is exposed to the discharge port. The compression process of a twin screw compressor is shown in Figure 4.11.15 [33]. The port design helps reduce suction and discharge flow losses, and provides a built-in volume ratio that gives peak performance at the corresponding pressure ratio.

The compression process generates axial and bending loads on the rotors as well as contact forces between the rotor lobes which need to be absorbed by the shaft bearings. The bearing loads of twin screw compressors are relatively high due to higher compression ratios achievable especially with oil-flooded twin screw compressors, and the small space between the shafts limits the diameter of the bearing that can be utilized. Normally, antifriction rolling element bearings are used on the twin screws, which result in a finite life under the high loads and often require replacement during the life of the compressor. The overall noise level of the twin screw compressor tends to be high and is at a higher frequency than the other positive displacement compressors. With only two major moving parts, the twin screw has high reliability.

In twin screw compressors, there are two main sources of internal leakages, which are the contact line between the meshing rotors and the clearance between the rotors and the casing. The rotor profile design and the manufacturing of the rotors ensure the existence of a sealing line throughout the compression process, which reduces the amount of pressurize gas that flows backward to low-pressure zones (typically the suction pocket). With respect to the manufacturing process, achievable clearances between the rotors are below 15 μm [32]. The clearances between rotors and casing are more difficult to quantify under running conditions due to the combined effect of thermal expansion and deformation of the rotors. The difficulty encountered in providing a good seal between the rotors and housing is usually met with the use of a large amount of oil (typically called “oil

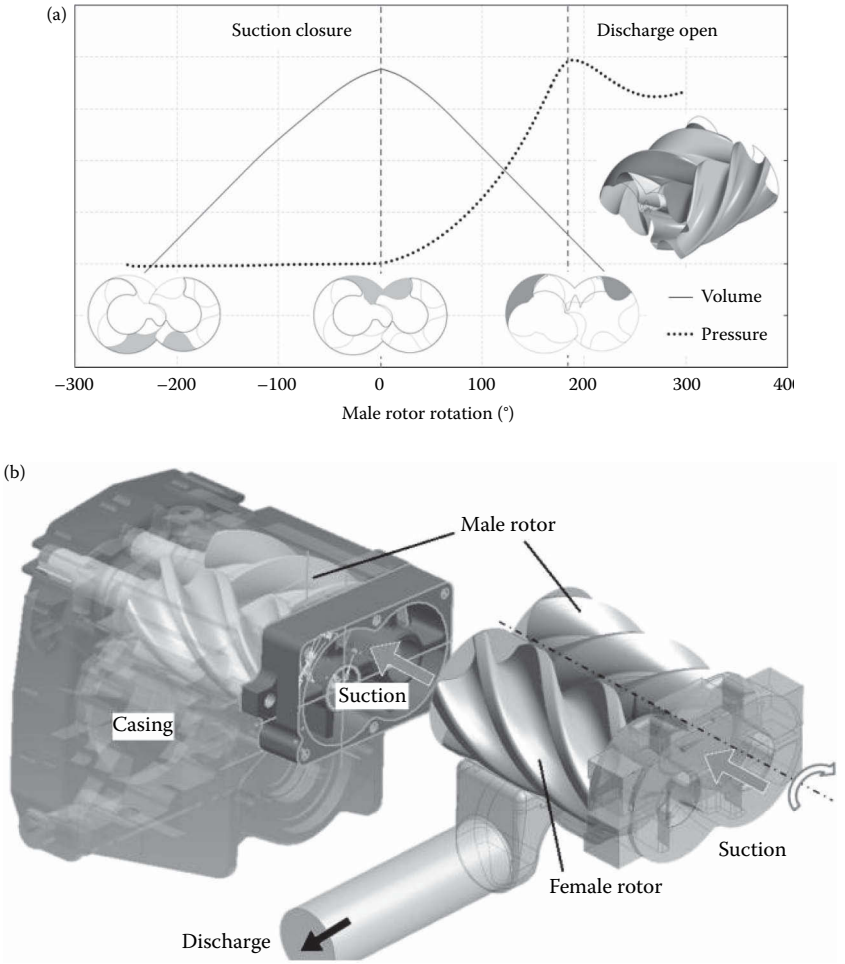


FIGURE 4.11.15 (a) Twin-screw compressor assembly and (b) working process of a twin-screw compressor. (Adapted from Rollins, J. P., Ed., *Compressed Air and Gas Handbook*, 5th ed., Compressed Air and Gas Institute, Cleveland, OH, 1989. With permission.)

flooding") during the compression process. The oil also serves as a coolant during the compression process. Depending on the rotor design and application, the quantity of oil can be substantial and must be separated from the discharge gas flow in most applications, usually by the use of oil separators (semi-hermetic designs). Thus, oil circulation can have a large effect on efficiency and the superheat of the compressor in air conditioning and refrigeration applications. At lower capacities, it is difficult to maintain good sealing. The result is poorer performance at lower capacities and that has traditionally limited applying screw compressors in air conditioning and refrigeration applications below 50 tons of refrigeration (175 kW).

During operation, the capacity of the twin screw compressor is varied to match load requirements by means of different control methods: suction throttling valve, variable speed control, and slide valves. The suction process can be modulated with a gate valve that extends the suction port, delaying the beginning of the compression process, and thus reducing the amount of inlet gas available for compression. Some screw compressors are also equipped with slide valves that shorten the effective length of the rotors by moving parallel to the rotors. Hence, the internal volume ratio is adjusted to the load requirements. The capacity of the screw compressor can be modulated in the range from 100% to 10%.

Besides the traditional twin screw configuration, an oil-flooded *tri-rotor screw compressor* with variable speed control concept has also been commercialized for the HVAC industry in the range of 50–500 tons (175–1750 kW), as shown in Figure 4.11.16 [34]. The tri-rotor configuration reduces the deflection of the rotors allowing a reduction of the amount of oil needed during operation.

The *single screw compressor (SSC)* was originally developed by Zimmern in the 1960s as an oil-injection free compressor. The SSC presents different configurations depending on the type of screw rotor (cylindrical, planar or conical) and the type of gate rotors (cylindrical or planar). The most common configuration is the CP-type (cylindrical-planar) with a central grooved rotor and two gate rotors (or starwheels) that are engaged antisymmetrically with respect to the main rotor rotation axis, as shown in Figure 4.11.17a. The gate rotor teeth sweep the screw grooves to achieve compression which represents a substantial difference with the twin screw compressors. The working process of the SSC is illustrated in Figure 4.11.17b [35]. Similarly to the twin screw compressor, the suction and discharge processes occur without the use of valves. The discharge port presents a triangular shape which follows the trailing edge profile of the groove. The typical grooves to teeth ratio is 6/11, which guarantees smooth engagement (issues arise if a common factor in the engagement ratio is used) and achievable compression ratios. The metallic screw rotor is in contact with high-wear-resistance material gate rotors (typically polyether–ether–ketone or PEEK) in such a manner that the compression process, although requiring cooling, does not necessarily require lubrication. However, the tooth flank is still subjected to wear after several hours of operation. Different meshing pair designs have been proposed to reduce the friction between the tooth flank and the groove. Most of the commercial SSCs adopt a straight-line envelope profile as a contact line between the meshing pair. With the advancements in the machining technology, multistraight line and multicolumn envelope meshing pair have been developed to allow the contact distribution over multiple lines and over a curved surface, respectively [36].

The SSC presents high volumetric efficiency and it does not require significant amounts of oil to seal the high-pressure gas from the low-pressure gas because the land between the trailing edge of a groove and the leading edge of the previous one increases as the pressure increases minimizing the cascade effect. Cooling can be accomplished by the injection of liquid refrigerant.

Due to the fact that the compression process occurs simultaneously in both sides of the main rotor, the swept volume is doubled and the radial loads on the rotor are balanced. The lower bearing loads contribute to lower noise and vibration levels compared to the twin screw compressors, which ultimately leads to an extended bearing life up to 15 years.

Part-load operation is accomplished by means of one or two slide valves on each side of the rotor that control the capacity and the volume ratio.

Screw compressors are made as semi-hermetic compressors, and as open drive compressors. Both twin and SSCs can be utilized to compress a variety of gases. Both designs are amenable to

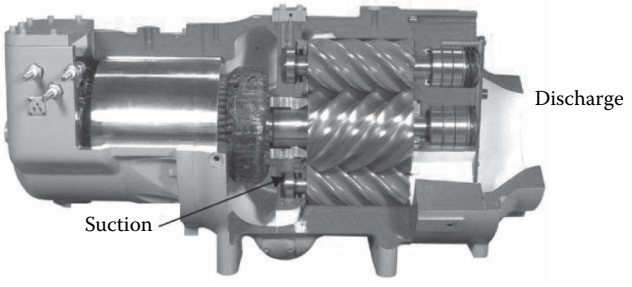


FIGURE 4.11.16 View of a 500 ton (1750 kW) tri-rotor screw compressor with a 8/6/6 lobe configuration for the male rotor and the two female rotors, respectively. (From Jacobs, J. S., Variable speed tri-rotor screw compressor technology, *International Engineering Conference*, Paper 1825, 2006. With permission.)

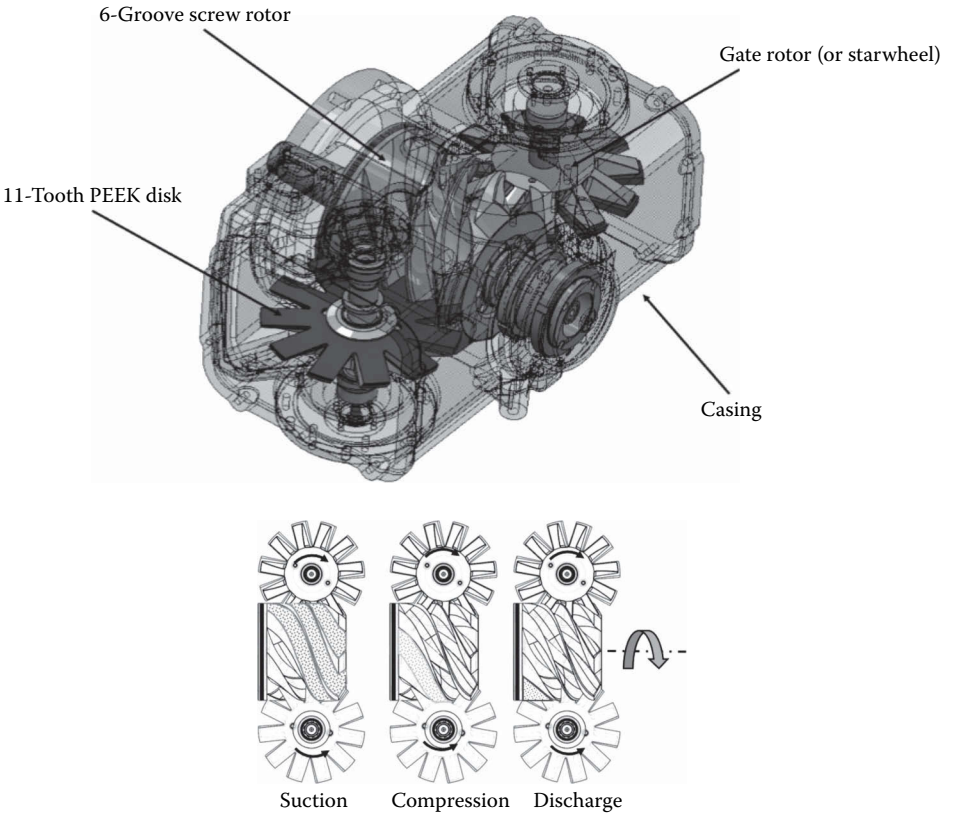


FIGURE 4.11.17 (a) CP-type single-screw compressor assembly and (b) working process of a single-screw compressor. (Adapted from Fig. 1 in Ziviani, D. et al., Update on single-screw expander geometry model integrated into an open-source simulation tool, *IOP Conference Series: Materials Science and Engineering,* 90, 012064, 2015. With permission.)

capacity control by varying shaft speed as well as by use of slide valves to delay the start of compression, thereby reducing the amount of inlet gas available for compression.

When used for air conditioning and refrigeration, screw compressors, similar to scroll compressors, can be designed so that vapor or liquid can be injected in the middle of compression, making it possible to compound the refrigeration cycle with the economizer concept. This is accomplished with a port in the housing that provides access to a point in the compression process.

Screw compressors vary in capacity from 15 to more than 400 tons of refrigeration (53 to more than 1400 kW). Most screw compressors are made in the range of 50–300 tons of refrigeration (175–1050 kW). An advantage of screw compressors with respect to reciprocating compressors above 15 tons of refrigeration (53 kW) is the significantly fewer number of parts and their relatively small size. Screw compressors have also become competitive with centrifugal compressors up to 400 tons of refrigeration (1400 kW) because of their superior part load performance. The part load performance of screw compressors is generally 8%–10% more efficient than for centrifugal technology. Also, screw compressors do not have the surge problems inherent with centrifugal compressors.

Dynamic Compressors

Dynamic compressors belong to the family of turbomachines that include fans, propellers, and turbines. Unlike positive displacement compressors which depend on a volume change for the compression process, centrifugal and axial compressors use the dynamic effect of velocity change,

acceleration followed by deceleration and static pressure recovery, for the compression process. Thus, the technology used is different for centrifugal and axial compressors than for positive displacement compressors.

There are many basic references about the technologies of dynamic compressors. Five references are suggested. The basic theory is covered in Refs. [37,38], and design is covered in Ref. [39]. A reference is suggested for centrifugal compressors [40], and one suggested for axial flow compressors [41].

A key equation used in the analysis of dynamic compressor performance is Euler's pump and turbine equation:

$$W_s = u_2 C_{\theta 2} - u_1 C_{\theta 1} \tag{4.11.1}$$

where W_s is the shaft work input per unit mass of fluid; u, the blade tip velocity; and C_θ, the tangential component of the fluid velocity leaving the blade. This equation, derived from the law of conservation of angular momentum, can be used with the aid of velocity diagrams (see Figure 4.11.20) to show the energy transfer mechanisms in centrifugal and axial flow compressors.

Centrifugal Compressors

Basic Description

In a centrifugal compressor, the suction gas enters the rotating element, or impeller, in the axial direction and is discharged in a radial direction at a higher velocity. The change in diameter through the impeller increases the velocity of the gas flow. This dynamic pressure is then converted to static pressure through a diffusion process, which generally begins within the impeller and ends in a radial diffuser outboard of the impeller. Figure 4.11.18 is a cross-sectional drawing of a typical single-stage centrifugal compressor.

A centrifugal compressor can be a single-stage compressor, having only one impeller, or it can be a multistage compressor, having two or more impellers mounted in the same casing. The suction

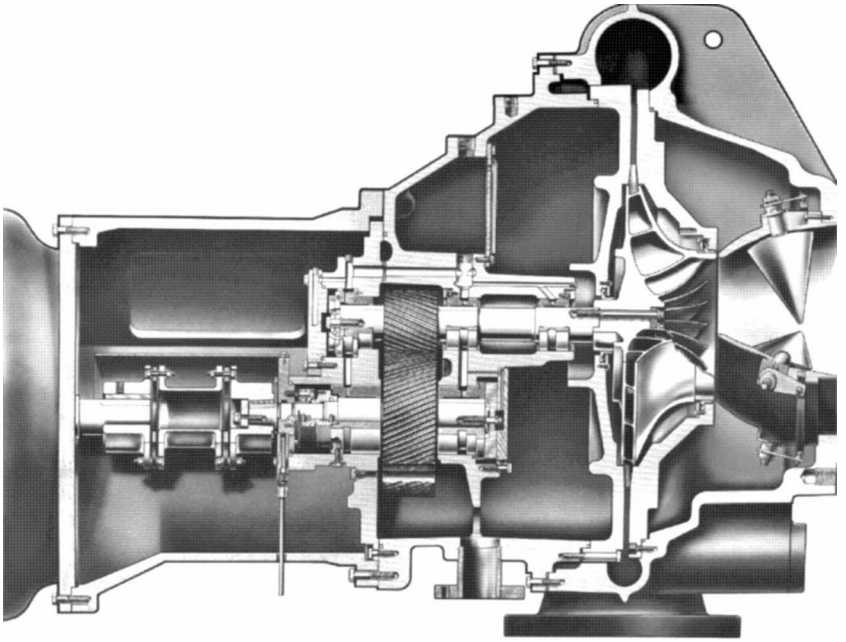


FIGURE 4.11.18 Drawing of a typical single-stage centrifugal compressor. (Courtesy of York International Corporation, York, PA.)

gas generally passes through a set of adjustable inlet guide vanes (prerotation vanes) or an external suction damper before it enters the rotating impeller. The guide vanes (or the suction damper) are used to control the flow rate through the compressor. The high-velocity gas discharging from the impeller enters a radial diffuser, which can have vanes or be without vanes. Diffusers with vanes are typically used in compressors designed to provide high-pressure ratios. These vanes are generally fixed, but they can also be adjustable. Adjustable diffuser vanes can be used for capacity modulation either in lieu of or in conjunction with the inlet guide vanes.

Centrifugal compressors are comparatively simple machines with few moving parts and moderate manufacturing tolerances. Yet they have high efficiency and are very robust. If the compressor's driver allows, the compressor speed can be varied to control the pressure ratio. Centrifugal compressors have a broad flow-handling range. At a given speed, the upper limit on flow is set by stall in the diffuser or impeller passages. The lower limit on flow is marked by surge—an unsteady flow condition marked by increased noise and flow reversal through the machine.

For multistage compressors, the gas discharged from the first stage is cooled and directed to the inlet of the second stage through a return channel. The return channel contains a set of fixed flow-straightening vanes or an additional set of adjustable inlet guide vanes. Once the gas reaches the last stage, it is discharged in a volute or collector chamber. From there, the high-pressure gas flows to the compressor discharge connection.

Design Characteristics

Centrifugal compressors range in flow rate from 150 cfm (0.071 m^3/s) in automobile turbochargers to 25,000 cfm (12 m^3/s) in large industrial process compressors. Pressure ratios per stage range up to 6:1 for air compressors, with higher values attained in research equipment. Figure 4.11.19 shows a typical performance map (head rise vs. flow rate) for such a machine [42]. The horizontal axis is proportional to the suction volume flow rate and the vertical axis is proportional to head (head is the isenthalpic enthalpy increase of the gas flowing through the impeller). The term "head" carries over from pump design nomenclature and is expressed in "feet" or "meters" even for compressors. Head is an index of the pressure ratio across the compressor.

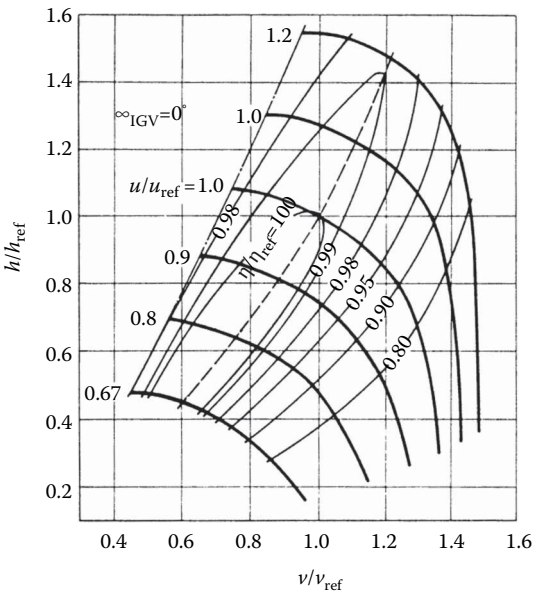


FIGURE 4.11.19 Centrifugal compressor performance map. (From the American Society of Mechanical Engineers, ASME Paper No. 86-GT-127. With permission.)

The plot also shows contours of constant efficiency, η. The head, flow, efficiency, and tip speed, u, all are nondimensionalized in the plot by using reference values. The reference value in each case is the value of the parameter at the design/rating point of the machine.

The right portions of the head-vs.-flow characteristics curves drop because of choking (explained below). The left end of each head-vs.-flow curve is limited by surge, also explained below. The surge limit, shown as a dot-dash line, connects the head/flow curves for different tip speeds. A dashed line in the center of the efficiency contours represents the loci of the "best efficiency points" at all operating speeds.

Performance maps such as the one in Figure 4.11.19 are essential tools for engineers applying compressors in particular applications. An application dictates the flow and head requirements the compressor must satisfy. The map shows the efficiency obtainable from the compressor for this flow and head, and the speed at which it must run.

Most compressors operate over a range of conditions in their applications. The performance map allows the application engineer to determine whether any requirements of the application will cause the compressor to operate too closely to surge or stall limits. If so, speed may have to be changed, prerotation vanes may have to be turned, or a different compressor design may have to be selected.

Operational limits for centrifugal compressors may be set either by structural limits or by flow parameters. For low molecular weight gases such as air, the compressor speed must remain within the stress limits of the impeller materials. Impeller tip speeds commonly fall below 1500 ft/s (460 m/s). For gases with higher molecular weights and lower acoustic velocities, the compressor speed is limited by the compressible flow phenomenon at the impeller and diffuser inlets. Impeller tip speeds commonly fall below a machine's Mach number of 1.5. Mach number here is defined as the impeller tip speed divided by the compressor inlet gas speed of sound. The maximum pressure level for centrifugal compressors is set by the casing design, and can reach several thousand psi (pounds per square inch) for multistage gas compressors used in the process industry.

The bearings in centrifugal compressors can be antifriction bearings in small machines. Plain journal bearings are commonly used in larger machines. Multistage compressors have more complicated bearing systems when operation above the first critical speed is required. The bearings are lubricated by a pressurized oil supply. Suitable seals are provided to minimize or eliminate lubricating oil carryover into the gas being compressed.

Centrifugal compressors are most commonly driven by electric motors with gearboxes to step up the speed to the level required for the compressor's duty. Steam or gas turbines drive some centrifugal compressors. Reciprocating engines are used in a few applications, generally with speed-increasing gears.

The efficiency of centrifugal compressors can be quite high under appropriate application conditions. Isentropic efficiencies of 80%–83% can be attained in single-stage machines. Sources of losses that reduce efficiency are flow friction (due to high gas velocities), separation and mixing losses, windage in clearance spaces, bearing losses, and effects of high Mach number flows (compressibility and local shocks). Losses arising from internal heat transfer generally are negligible in these machines.

Centrifugal compressors continue to be the dominant choice for very large refrigeration sizes, e.g., applications above 1500 kW (427 tons of refrigeration) cooling capacity. This is driven by unchallenged overall system efficiencies in these sizes. Centrifugal compressors are also very simple with few moving parts. They have manufacturing tolerances two orders of magnitude greater than equivalent screw compressors do. Therefore, centrifugal compressors have relatively low manufacturing costs with high reliability. In addition, centrifugal compressors for large-size applications operate within a very narrow band of operating head pressure, thereby minimizing problems associated with surge.

Development Directions

In the future, IPLV (integrated part load value) will become more important, requiring centrifugal compressors to operate efficiently over a broader range of applications. The main challenge for the

next generation of centrifugal compressors will be to maintain the full load efficiency at current levels while meeting the part load challenges of the future.

With the advent of ceramic bearings and hydrostatic bearings coupled with liquid refrigerant used as the lubricant, the elimination of oil management systems and hydrodynamic bearings can be expected. Therefore, oilless compression remains a viable technology for the centrifugal compressor. Indeed, magnetic bearings are offered in some process industry compressors.

Since 1993, centrifugal compressors are being redesigned for CFC refrigerant replacements. Many opportunities exist for increased efficiency and improved thermal aerodynamics. Utilizing the technology of the high-speed, small centrifugal turbines, even smaller sizes of centrifugal compressors are possible. Other opportunities exist for improving reliability and serviceability, for reducing noise and vibration, for increasing the use of electronics for sensing and control, the use of inverters for speed control, and for improving the manufacturing processes to reduce costs.

Flow and Performance Characteristics

Figure 4.11.20 shows the velocity vector diagram for a centrifugal compressor. The velocity diagram indicates that work input to the impeller accelerates the flow to a high velocity at its tip. A static pressure rise may occur across the impeller to an extent dictated by the impeller geometry and gas density change. The diffuser passages downstream of the impeller may be without vanes or may have vanes to decelerate the flow and increase its static pressure. A scroll or volute collects the flow at the radial diffuser exit. A conical diffuser often is located at the collector exit to achieve additional static pressure rise.

Prerotation vanes, fixed or adjustable, may be located upstream of the impeller inlet. These vanes can be used to adjust the impeller work input and inlet flow angles. Also, when turned to large angles, the prerotation vanes act as a throttling device, which may allow the compressor to operate at low mass flow rates where otherwise surging would occur.

The impeller discharge velocity triangle includes the slip factor, *S*. This factor accounts for the mass-averaged deviation of the flow through the impeller. The flow does not leave at the geometric blade angle because of irrotational flow in the absolute reference frame, creating the "slip."

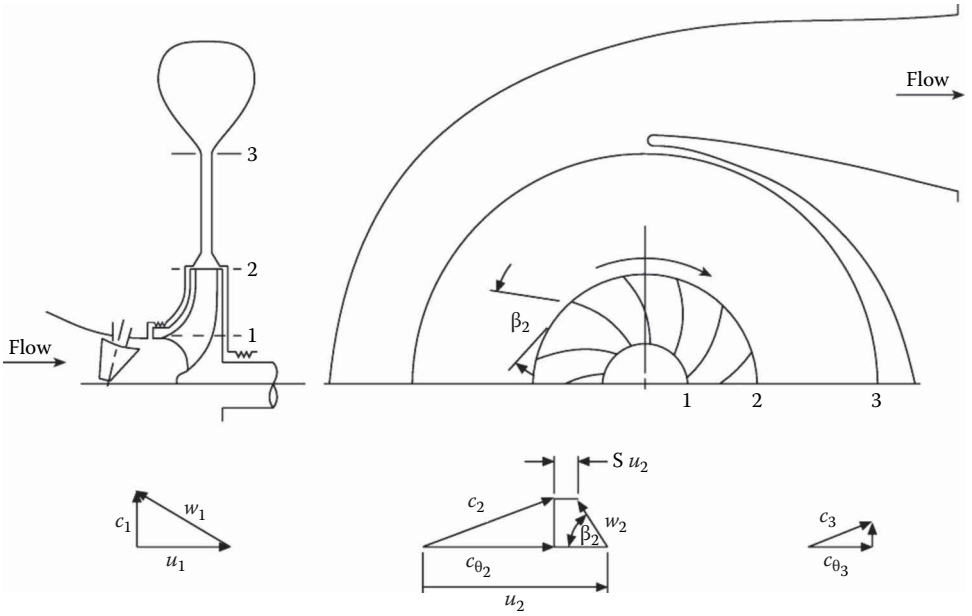


FIGURE 4.11.20 Velocity vectors for a centrifugal compressor.

Equation 4.11.1 shows that the work input to a centrifugal compressor calls for maximizing the product $u_2C_{\theta 2}$. The upper limit on the value of u_2 is set by the strength of the impeller material or by the acoustic velocity of the gas being compressed. Air compressors are limited by impeller strength while refrigerant compressors are limited by flow velocities approaching the sonic limit at the impeller eye (w_1) or at the inlet of vaned diffusers (related to C_2). Note that w is the relative velocity between C and u (see Figure 4.11.20).

The limit for the value of $C_{\theta 2}$ is set by the limit on the value of w_2 and the choice of the impeller discharge blade angle, β_2. Impeller passages act as diffusers, with a maximum limit of 1.4–1.6 set for the value of the diffusion ratio, w_1/w_2, to avoid significant stall losses. The implication of the blade angle is discussed below.

The performance map of a centrifugal compressor for work, head, and losses is illustrated qualitatively in Figure 4.11.21. The compressor losses are at a minimum when flow angles throughout the compressor match the design conditions. At lower flow rates, losses increase as impeller, diffuser, and scroll boundary layer flows stall. Stall with high loss also occurs in the impeller and scroll at high flow rates. Ultimately, surge (low flow) and choke (high flow) limit the compressor's operating range. Surge is a major flow instability in the compressor, accompanied by complete flow reversal, which occurs when stall losses become so great that the compressor no longer can sustain the discharge pressure imposed upon it by the external system. Choking occurs at high flow rates when stall losses within the compressor become larger than the work input can sustain. In high-performance compressors, choking may be affected by added losses from transonic flow disturbances.

Figure 4.11.21 applies to an impeller with "backward leaning" vanes, i.e., $\beta_2 < 90°$. Figure 4.11.20 shows that if $\beta_2 = 90°$, the radial component of C_2 will have no impact on the level of $C_{\theta 2}$. This is because w_2 is radial in direction, if β_2 is equal to 90°. The slip factor, S, essentially remains constant for varying flow rates in a particular impeller. Thus, $C_{\theta 2}$ becomes a maximum that affects the surge limit of a centrifugal compressor and reduces the stable flow range. In Figure 4.11.21, the line representing work input, W, will be horizontal, not sloping downward with increasing flow. With the work input, W, as a horizontal line combined with the loss characteristics shown, the curve representing the head, H, will drop at low flow rates. Thus, surge will occur at higher flow rates for a radial bladed impeller than for a backward leaning one. A radial impeller generates maximum head but exhibits reduced range of flow between surge and choke limits.

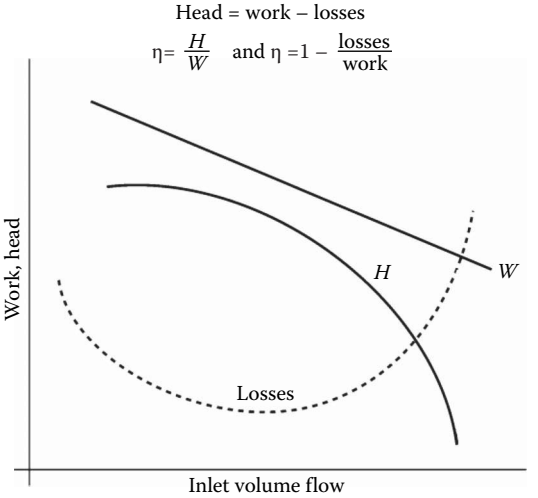


FIGURE 4.11.21 Centrifugal compressor performance at constant speed.

Axial Flow Compressors

Basic Description

In an axial flow compressor, the suction gas enters and the discharge gas leaves the compressor in the axial direction. As shown by the cross section in Figure 4.11.22, axial compressors have a small frontal (inlet) area but are comparatively long in the axial dimension because multiple stages are required for most applications [43].

Axial flow compressors are used primarily as components of gas turbines for power generation or aircraft propulsion. A few are used in process gas compression service where large volume flow rates are required. Typical flow rates for axial flow compressors range from 25,000 cfm (12 m^3/s) in small aircraft engines to 1 million cfm (470 m^3/s) for large stationary plants. The number of stages ranges from 7 to about 17, with the upper limit set by the dynamic considerations of the rotor.

Design Characteristics

Axial compressors are complex machines with high precision required in the manufacturing of the blades. Each blade row is made up of thin, carefully profiled airfoil surfaces, which must have adequate strength to withstand large aerodynamic forces. Consequently, the flow range is narrow for axial compressors as indicated by the typical performance map shown in Figure 4.11.23 [43]. The narrow range of the nearly vertical pressure ratio (or head) vs. flow curves in Figure 4.11.23 contrasts with the broader flow range exhibited by the head vs. flow curves of the centrifugal compressor shown in Figure 4.11.19 [42].

Flow control in axial compressors is achieved by using variable stator vanes and, in engine applications, speed control. The flow range is limited by blade stall on both the high- and low-flow sides of the performance map. Axial compressors can exhibit surge and also rotating stall patterns under low flow conditions—both are unsteady flow conditions. Sometimes, interstage flow bleeds are used to broaden the flow range of these machines.

Compressible flow (Mach number) effects due to the high flow velocities in axial flow compressors commonly limit the pressure ratio that is achieved per stage.

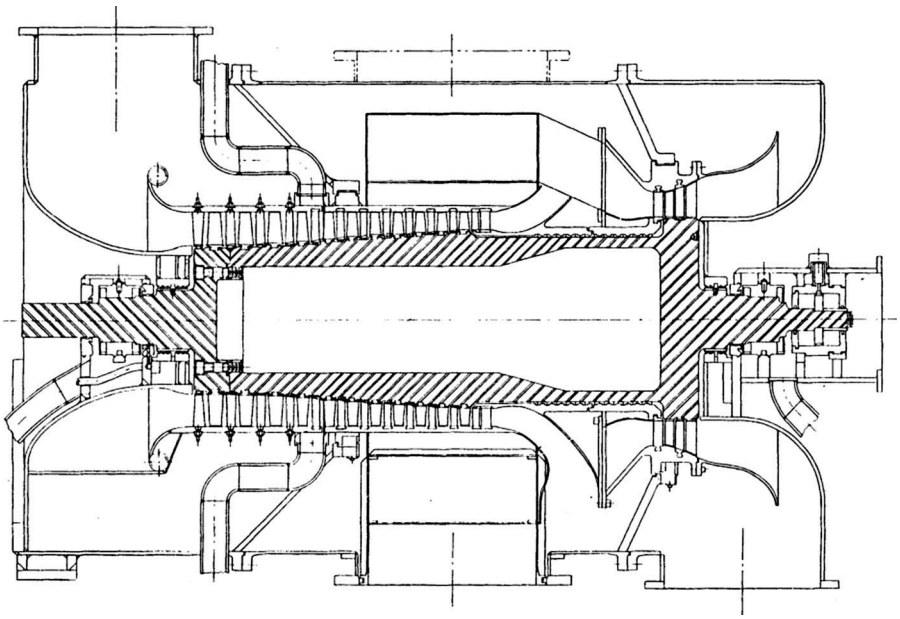


FIGURE 4.11.22 Cross-sectional drawing of an axial-flow compressor. (From the American Society of Mechanical Engineers, ASME Paper No. 58-SA-25. With permission.)

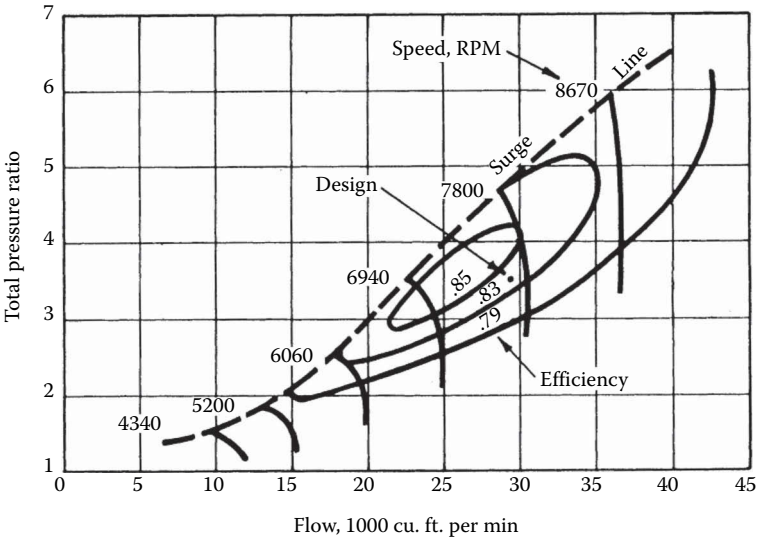


FIGURE 4.11.23 Typical performance map of an axial flow compressor. (From the American Society of Mechanical Engineers, ASME Paper No. 58-SA-25. With permission.)

Tight clearances between rotating blade tips and casing surfaces are required to minimize tip losses. Thus, most axial compressors employ antifriction bearings because of their small clearances. Pressurized lubrication systems are provided, with seals that assure little or no oil carryover into the compressed gas stream.

Gas or steam turbines are the most common drivers for axial compressors. Electric motors with speed-increasing gears sometimes are used in stationary applications.

Axial compressors are employed where high efficiency is essential. Isentropic efficiencies exceed 80% in most applications. Sources of losses which reduce efficiency include flow friction (resulting from high velocities), losses associated with transonic flow, tip clearance losses, flow separation and mixing losses, flow deviation losses in downstream stages, windage losses in disc clearance spaces, and bearing losses. Losses due to internal heat transfer are negligible in axial compressors.

Flow and Performance Characteristics

Figure 4.11.24 shows the velocity vector diagram for an axial flow compressor. The work input in axial compressors occurs as a result of a series of small changes in flow direction taking place at high-velocity levels with no significant change in radius. Small static pressure increases commonly occur across each blade row, downstream of the inlet guide vanes.

Axial compressor blade design practices at first were closely related to aircraft airfoil design practices. In the United States, an NACA airfoil series, and in the United Kingdom, a circular arc airfoil series, were used by compressor designers as jet engines advanced in the 1950s. Wind tunnel tests of cascades of blades provided correlation of lift (angle change) and drag (losses) as a function of blade geometry, pitch/chord ratio, and stagger (blade angle relative to the axis).

Stationary blade rows in axial compressors may be fitted with mechanisms to change their angular position. Inlet guide vanes and the first several stator blade rows may be so equipped to allow the compressor to operate over a wider flow range or range of pressure ratios. Downstream stages sometimes have bleed-off ports to accommodate changes in operating conditions that otherwise might cause subsequent stages to stall.

The basic performance map of a single stage (stator and rotor blade row) of an axial compressor is similar to the performance map of the centrifugal compressor shown in Figure 4.11.21.

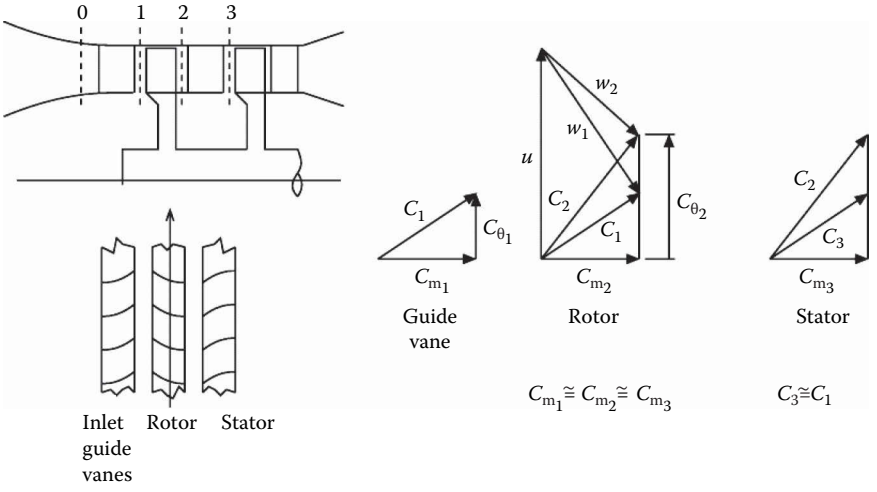


FIGURE 4.11.24 Velocity vectors for an axial flow compressor.

The axial compressor's flow range and work input per stage are much lower than for a centrifugal compressor because of the narrow range of blade inlet flow angles tolerated by the airfoils before stall occurs at positive or negative angles of attack. Axial compressors often run at speeds such that flow velocities approach the velocity of sound in the gas being compressed. Transonic flow can occur at points along the blade rows, increasing the likelihood of stall as flow angles depart from design values.

For axial compressors, Equation 4.11.1 shows that the work input is a result of the product $u(C_{\theta 2} - C_{\theta 1})$ because u, the blade tip velocity, essentially is constant through each blade row. The velocity triangles in Figure 4.11.24 show that both the rotor and stator blade rows act as diffusers. The limits on the value of $(C_{\theta 2} - C_{\theta 1})$ are set by the limits on the amount of diffusion, C_3/C_2 or w_2/w_1 that can be accomplished without high losses caused by diffuser stall.

The velocity, u, has a maximum limit set either by rotor material strength or by the acoustic velocity in the gas being compressed. As values of flow velocities, C_2 and w_1, approach the acoustic velocity, losses increase due to oblique shocks and flow redirection in local regions of transonic flow.

The operating range of a complete multistage axial flow compressor is limited by rotating stall (low flow) and choking (high flow). Rotating stall is a phenomenon that occurs when the stall of one blade (or several blades) on a given blade row causes adjacent blades on the same rotor row to stall. The resulting stall "patch" rotates around the circumference of the blade row. Downstream blade rows are affected by the upstream stall zone and may undergo their own rotating stall, leading to stall for the compressor as a whole. At high flow rates, stall of one or more blade rows creates high losses that prevent the rows from handling increased flow rates against the pressure rise imposed on the compressor by the external system.

Air Compressors

Air compressors can be classified according to whether they are portable or stationary, displacement or dynamic, oilfree or lubricated. An air compressor unit is usually sold including an "airend" (compression module), driver, controls, a cooling system, and a safety system. An enclosure to reduce noise is also available in most cases. A special foundation (or compressor base) is normally not required. "Packaged compressors" are self-contained, unlike refrigeration compressors.

Compressor Types for Various Applications

Factors affecting the selection of an air compressor are initial price, operating cost, reliability, warranty, and the availability of service and parts. Figure 4.11.25 shows the ranges of pressure and capacity typical of different types of compressors used in compressed air systems. The range of different types sometimes overlaps. The reciprocating compressors, such as the one shown in Figure 4.11.26, are used for low cost and low capacity (low flow rate) applications. For industrial applications below about 2500 cfm (1.2 m^3/s) airflow rate, the dominant type is the oil-flooded screw compressor. These are usually purchased as a packaged compressor, such as shown in Figure 4.11.27. For oilfree applications, dry screw or reciprocating compressors are the most popular choices, with

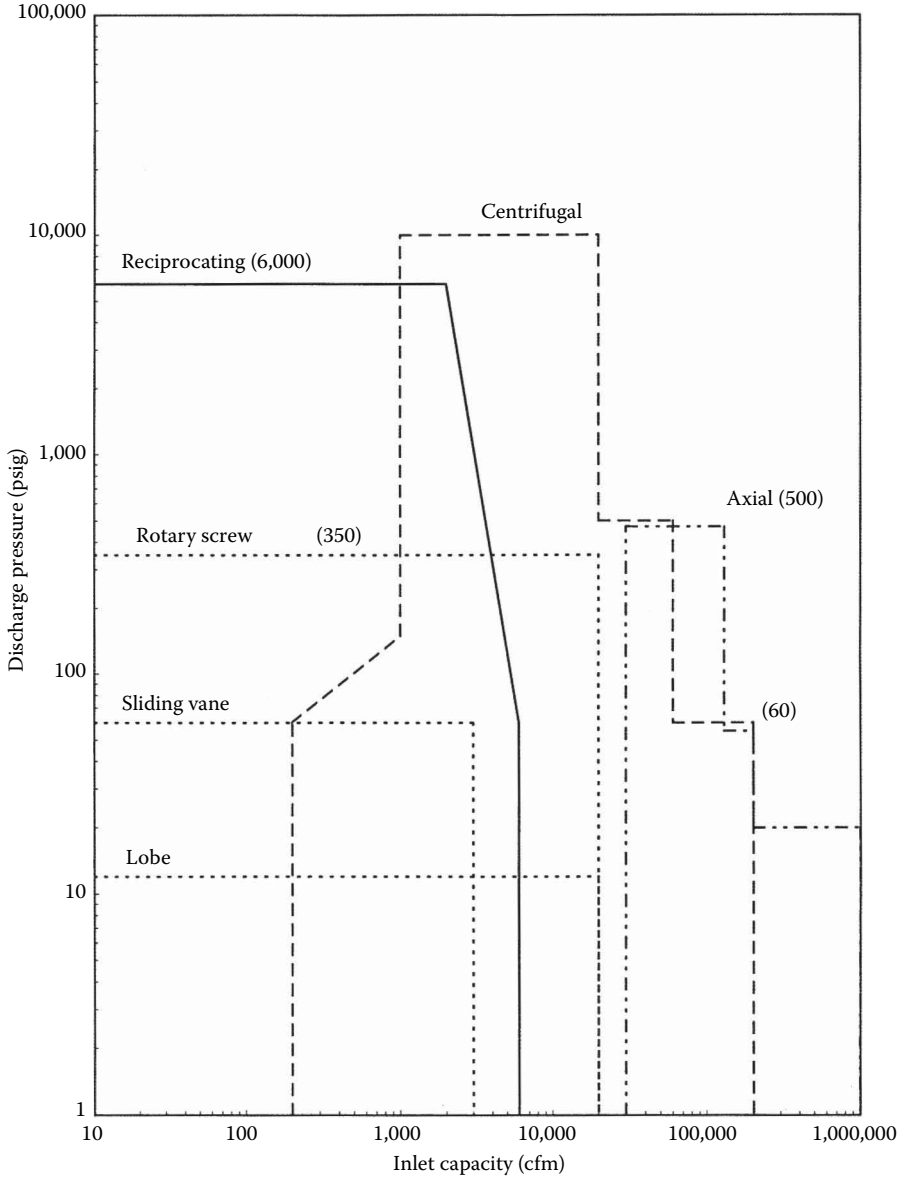


FIGURE 4.11.25 Pressure–capacity chart showing the effective ranges of most air compressors. (Courtesy of Ingersoll-Rand Corp., Woodcliff, NJ.)

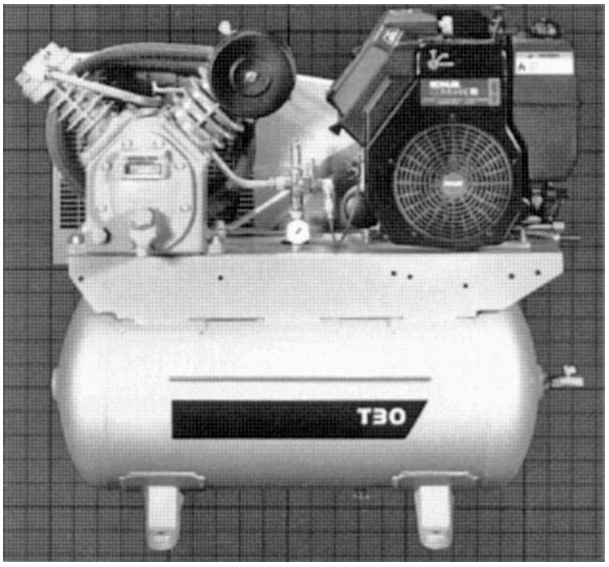


FIGURE 4.11.26 Two-stage reciprocating air compressor package with horizontal receiver tank. (Courtesy of Ingersoll-Rand Corp., Woodcliff, NJ.)

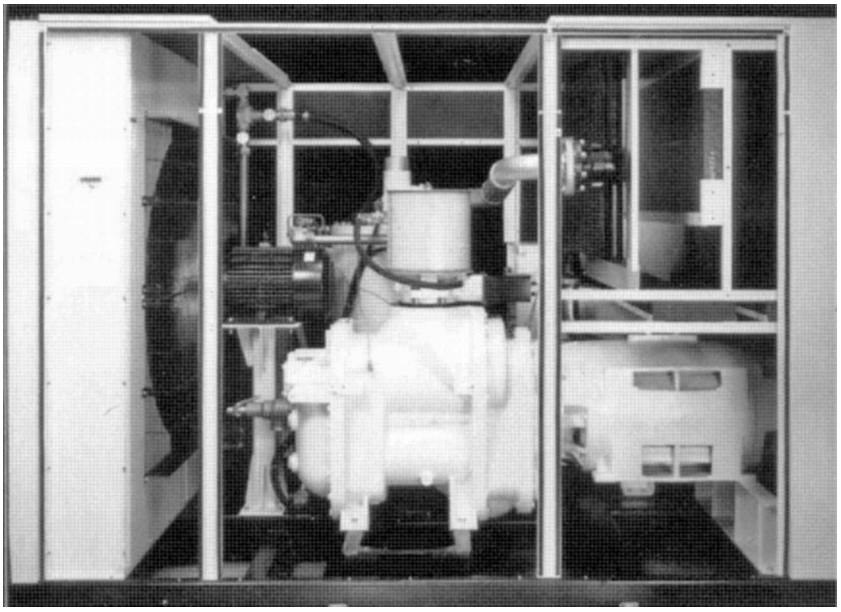


FIGURE 4.11.27 Interior of an air-cooled, two-stage, 200 hp oil-flooded screw compressor package. (Courtesy of Ingersoll-Rand Corp., Woodcliff, NJ.)

dry screw compressors taking an increasingly larger share of the market. Reciprocating compressors are also used where higher pressures are needed. Above 2500 cfm (1.2 m^3/s) airflow rate, centrifugal compressors, such as shown in Figure 4.11.28, dominate, up to the sizes where axial flow compressors become practical.

The smallest portable compressors, i.e., less than 50 cfm (0.0235 m^3/s) airflow rate, are usually built with reciprocating compressors driven by small gasoline engines. Above this range, most

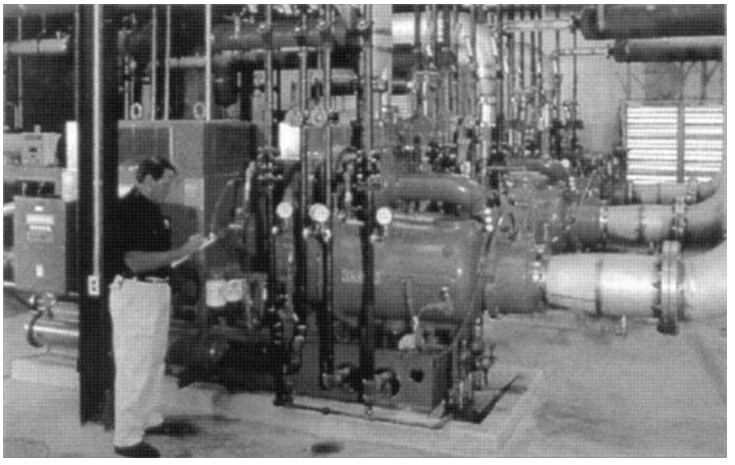


FIGURE 4.11.28 Three 5700 cfm, 1250 hp centrifugal air compressors supplying oilfree air to a textile facility. (Courtesy of Ingersoll-Rand Corp., Woodcliff, NJ.)

portable air compressors utilize an oil-flooded rotary twin screw design driven by a Diesel engine. Low-pressure air compressors with air discharge pressures below 200 psig (1480 kPa) are usually single-stage designs, whereas higher-pressure compressors have multiple stages. Oilfree portable compressors are available with two-stage rotary screw or centrifugal compressors. Most portable compressors have air flow rates from 100 to 1800 cfm (0.047 to 0.85 m^3/s), producing air delivery pressures of 100–150 psig (790–1140 kPa).

Nearly oil-free compressed air can be supplied to the process using an oil-flooded compressor by cooling the air after compression and then passing it through a series of very fine coalescing filters to remove any remnants of the oil aerosols in the compressed air.

Performance Considerations

Ratings

Air compressors are rated on delivered airflow, discharge pressure, and power consumption. The airflow is stated in terms of inlet volumetric flow, sometimes called "free air delivery." It is determined in a test by measuring the mass flow rate at the discharge and dividing by the inlet density. For a bare compressor, the inlet is clearly defined, but for a packaged compressor the inlet can be specified to be at either the package or compression module inlet. The stated flow will usually be lower if the package inlet is used, because the temperature is lower with a higher density. Test codes for measuring compressor performance have been developed by the European Committee of Manufacturers of Compressors, Vacuum Pumps and Pneumatic Tools (PNEUROP), and in the U.S. by the Compressed Air and Gas Institute (CAGI). These test codes have recently been included in ISO 1217, which is gradually becoming the industry standard in Europe and the U.S. This ISO standard specifies a standard package inlet condition of 20°C (68°F), 100 kPa (14.5 psia), and 0% relative humidity. The unit of flow in the ISO system is m^3/s, whereas in the U.S. ft^3/min (cfm) is generally used. Power is normally expressed in kilowatts or horsepower. For portable compressors, fuel consumption is stated instead of power.

Capacity and Pressure Control

Capacity control is an important performance characteristic, since it is a major factor contributing to how efficiently the compressor operates at part load. In addition to on–off control, compressors are usually fitted with an inlet valve. This inlet valve controls the airflow to the compressor by either

opening, throttling, or closing completely in response to pressure-sensing controls. Some compressors are equipped with a check valve that allows the receiver tank to be depressurized during periods of low demand. This reduces unloaded power by reducing the discharge pressure. Speed control is used on engine-driven compressors, and, increasingly, on those driven by electric motors. Dry screw compressors are not throttled at the inlet, since this increases the pressure ratio and thus, the temperature rise across a stage. These compressors are usually equipped with a check valve so the discharge can be depressurized. The inlet valve is either full open (when loaded) or full closed with the discharge at ambient pressure.

Ambient Temperature

Since the heat of compression must be carried away by a cooling system, all compressors have a limiting ambient temperature (LAT) above which the internal temperatures exceed their limits. Higher LAT values are possible with better cooling systems. The engine or motor cooling system also affects the LAT. Portable compressors usually have a temperature switch at the compressor discharge and another for the engine coolant. Exceeding the set point of either switch will cause the compressor to shut down.

Compressors also have low-temperature limits. Typical minimum operating temperatures are 35°F (1.7°C) for stationary compressors, and –10°F (–23.3°C) for portables. A low ambient compressor option is usually available for stationary compressors operating outside in cold climates. The low-temperature limits are due to the potential of the condensate freezing in the pneumatic control system components and in the aftercooler. For portable compressors, engine operation at cold temperatures can be the factor that determines the low-temperature limit. Startup at low temperatures must be carefully controlled to ensure that the compressor receives adequate lubrication. Preheating the oil will sometimes allow the compressor to be started at lower temperatures. In general, dry screw compressors have a more narrow ambient temperature range than oil-flooded screw compressors.

Ambient Pressure and Altitude

Except for dry screw and dynamic compressors, the ambient pressure is usually not a limiting factor from the compression module standpoint. As the altitude increases, the ambient pressure and temperature decrease. As a result, the power requirement of the compressor also decreases. This more than offsets the lower cooling efficiency when operating with less dense ambient air. For portable compressors, the maximum altitude is usually limited by the engine, which loses power capability with altitude. Dry screw and dynamic compressors must operate at lower discharge pressures at higher altitudes, to keep the pressure ratio from exceeding the limit for a particular unit.

Start–Stop Frequency

While on–off operation is an efficient mode of capacity control, it is limited by the maximum start–stop frequency. This is typically six per hour for an electric motor, to prevent overheating the motor starter components. Diesel engines must operate long enough between starts to recharge the battery. In addition, the starter motor and ring gear will wear out prematurely if the start–stop frequency is too high.

Compressor Speeds and Drives

Oil-flooded screw compressors operate most efficiently at a tip speed of about 35 m/s (115 ft/s). At higher tip speeds the efficiency decreases and the noise increases. These “airends” are usually not operated with tip speeds above 50 m/s (164 ft/s). Dry screw compressors must operate at tip speeds near 100 m/s (328 ft/s), to minimize backflow leakage. As a result, they are much noisier than oil-flooded screw compressors and require more acoustical treatment.

In general, a large, slow-running compressor is more efficient and quieter than a small, high-speed compressor. However, the initial cost is higher. As compressor design advances, the trend is to smaller, faster-running compressors, with better silencing techniques and more sophisticated controls to keep operating costs down.

Most portable air compressors are driven by Diesel engines, except those at the lower end of the airflow rates [below 100 cfm (0.047 m^3/s)], where gasoline engines are often used. Electric motor-driven portables are used in some areas. Stationary compressors are driven by electric motors. The majority of them operate at either 1800 rpm with a 60 Hz line frequency or at 1500 rpm where the line frequency is 50 Hz. Very large compressors in process plants are often driven by steam turbines.

Most portable compressors use a drive coupling to transmit the torque from the engine to the compression module, although small compressors are often belt driven. The drive coupling can be a troublesome component and must be carefully selected to minimize operational problems. Ideally, the compressor should be rugged enough to withstand the torque pulses from the engine, so that a torsionally stiff drive coupling can be used. This is often not possible, however, especially with the larger compressors, and a torsionally soft coupling must be used to isolate the "airend" from the engine. In these cases, the engine speed must be maintained well above the torsional critical speed to prevent drive train failures. The torque pulses from an electric motor are mild compared to those from a Diesel engine. As a result, the coupling is often eliminated by directly coupling the compressor and motor. This is accomplished by mounting the compressor drive gear on the motor shaft. If a coupling is used, it can be torsionally stiff.

Maintenance Requirements

One of the primary reasons for the popularity of oil-flooded screw compressors is their ability to operate continuously for many years with only routine maintenance. Filter maintenance is most important, since the incoming air is mixed with the oil and any particles in the incoming air end up in the oil. This same oil is used to lubricate the bearings, and particles in the oil will decrease bearing life. Typically, a 25 μm (0.64 μm) air filter is specified, with cleaning or changing required when the restriction exceeds 25 in. of water (6.2 kPa). The oil system also has a 25 micron (0.64 μm) filter, as well as an oil separation element, which captures particles as small as 1 micron (0.025 μm). Any contamination entering the system ends up in the separator element, thus shortening its life. The use of finer air and oil filters will increase the time between separator element changes. This is usually desirable, since changing the air and oil filters is much easier than changing the separator element.

The coolers must also be kept clean. Dirty or fouled coolers reduce the LAT, sometimes so much that the compressor shuts down because of overheating. Another undesirable consequence of dirty coolers is a shorter bearing life due to higher oil temperatures.

Regular filter maintenance is just as important on dry screw and dynamic compressors. A restricted air inlet reduces the compressor inlet pressure and raises the pressure ratio across it. This can cause overheating of a dry screw compressor and surge in a dynamic compressor.

Thermal Issues

Performance

Assuming adiabatic compression and applying the first law of thermodynamics, the rate of work of compression for a single stage compressor is given by

$$\dot{W} = \dot{m} c_p (T_2 - T_1) \tag{4.11.2}$$

where $\dot{m}$ is the mass flow rate of the compressed air; c_p, the specific heat; T_2, the outlet temperature; and T_1, the inlet temperature. Treating the air as a perfect gas, this becomes

$$\dot{W} = P_1 \dot{V}_1 \frac{k}{k-1} \left[\left(\frac{P_2}{P_1} \right)^{\frac{k-1}{k}} - 1 \right] \frac{1}{\eta} \tag{4.11.3}$$

where k is the specific heat ratio; P_1, the absolute inlet pressure, $\dot{V}_1$, the inlet volumetric flow rate; P_2, the absolute discharge pressure; and η, the adiabatic efficiency. For air, $k = 1.4$ and $c_p = 0.24$ Btu/(lb_m °F) [$c_p = 1.004$ kJ/(kg K)]. In practice, the input power and flow rate for a bare compressor (i.e., without inlet filters or discharge restriction) are measured on a dynamometer, and the adiabatic efficiency is calculated from Equation 4.11.3. This efficiency includes the power lost in the gears and bearings. Positive displacement compressors, including oil-flooded screw compressors, are almost always analyzed using the adiabatic model. Comparing adiabatic efficiencies at different pressure ratios is not a valid procedure. The efficiency should be measured at each pressure ratio of interest.

Examination of Equation 4.11.3 shows that the required power input depends only on the inlet flow rate, the inlet pressure, the pressure ratio, and the k-value. It does not depend on the inlet temperature of the gas or its molecular weight. In two-stage compressors, the power to the second stage can be reduced by intercooling, i.e., cooling the air between the first and second stages. The lower the temperature at the entrance of stage 2, the lower will be V_1 for that stage. By adding even more stages and intercooling before each additional stage, isothermal compression is approached, which requires less power than adiabatic compression. The better the intercooling, the more efficient is the overall compression process. The power increases approximately 1% for every 10°F (5.6°C) rise in inlet temperature to each stage. It is also important to keep the interstage pressure drop as small as possible, as pressure drop increases the pressure ratio for the next stage.

The inlet flow rate depends mainly on the speed, although it decreases slightly as the pressure ratio increases. In oil-flooded screw compressors, the air is heated by hot injection oil, which in turn decreases the inlet density of the air, thus decreasing the mass flow rate. A reasonable approximation for this effect is a 1% decrease in flow per 10°F (5.6°C) rise in the difference between the oil injection and air inlet temperatures. Since the injection oil temperature is held constant by the thermostat in cold weather, this means the flow decreases about 1% for every 10°F (5.6°C) drop in ambient temperature.

Water Condensation

All air contains moisture, the amount being influenced by pressure, temperature, and proximity to oceans, lakes, and rivers. The water content of saturated air (100% relative humidity) is given by

$$W_{sat} = 0.622 \frac{P_{sat}}{P - P_{sat}} \tag{4.11.4}$$

where W_{sat}, is the humidity ratio, in lb_m water vapor/lb_m dry air; P_{sat}, the saturation pressure of water vapor at the temperature of the air; and P, the absolute pressure of the air–water vapor mixture. P_{sat} can be obtained from steam tables. As Equation 4.11.4 indicates, the water-holding capacity of air decreases if the air is compressed, unless P_{sat} is raised by increasing the temperature. The dew point temperature is the temperature at which W_{sat} is equal to the specific humidity of the ambient air. For example, ambient air at 14.7 psia (101 kPa), 80°F (26.7°C), and 60% relative humidity has a saturation pressure of 0.51 psia (3.52 kPa) and a vapor pressure of $0.60 \times 0.51 = 0.31$ psia (2.14 kPa). The humidity ratio is $W = 0.622 \times 0.31/(14.7 - 0.31) = 0.013$ lb_m water/lb_m dry air. If the air is compressed to 200 psig (1480 kPa) and then cooled to 100 °F (37.8°C), $P_{sat} = 0.95$ psia (6.55 kPa) and the water-holding capacity is $W_{sat} = 0.622 \times 0.95/(214.7 - 0.95) = 0.0028$ lb_m water/lb_m dry air. This is much lower than the entering humidity, so the amount of condensation is $0.013 - 0.0028 = 0.010$ lb_m water/lb_m dry air (0.01 kg/kg). At 1000 cfm (472 m^3/s) or 75 $lb_{m,\,air}$/min (34 kg_{air}/min), the condensation rate is $0.010 \times 75 = 0.75$ $lb_{m,\,water}$/min (0.34 kg_{water}/min) or 5 gal_{water}/h (18.9 L_{water}/h). If the compressed air temperature is 180°F (82.2°C), instead of 100°F (37.8°C), the saturation pressure is 7.51 psia (51.8 kPa) and $W_{sat} = 0.022$ lb_m water/lb_m dry air. At this temperature, no condensation takes place. Repeating this calculation for different values of the temperature, one finds that W_{sat} is equal to 0.013 lb_m water/lb_m dry air at 160°F (71.1°C). Thus, the dew point temperature is 160°F (71.1°C).

Condensation in oil-flooded screw compressors is particularly objectionable because the water will mix with the oil and reduce its lubricating qualities. Using a thermostat in the oil system to keep the "airend" discharge temperature above a minimum level prevents condensation. A minimum discharge temperature of 180°F (82°C) is usually adequate for pressures up to 350 psig (2500 kPa). Once the oil has been separated from the air, the air can be cooled in an aftercooler and the condensation removed in a water separator. Since the air is saturated when it leaves the aftercooler, and still above ambient temperature, further condensation will take place in the air line. For this reason, most plants employ an air dryer downstream of the compressor.

Reduction of Energy Use

Compressed air can account for 10%–20% of a manufacturing facility's electricity bill [4]. Energy efficiency for compressed air systems has reached a new height recently with the announcement in 1998 by the U.S. Department of Energy of the Compressed Air Challenge. This initiative is designed to build awareness among users of compressed air about the benefits and approaches for improving and maintaining compressed air system efficiency. Some basic methods for improving efficiency are given by Edwards [44] as follows:

1. Reduce unnecessary demand by eliminating leaks, by using condensate drain valves that open only when a predetermined amount of water has accumulated, by regulating all point-of-use operations at the lowest possible pressure, and by modifying and/or eliminating "blowoffs." Often the output of a low-pressure blower or fan can replace compressed air for "blowoff" operations.
2. Shut off the air supply to "off-line" production equipment. If one point-of-use requires air pressure at a much higher pressure than the rest of the system, consider using a pressure booster or using a separate dedicated system at that point.
3. Use storage tanks to supply air for intermittent events that require large amounts of air for a short time. Storage can eliminate the need for additional compressors to handle peak demands.
4. Reduce pressure drop by having adequate line sizes.
5. Select the compressor sizes and control their operation so that a compressor is either off, or running at rated load. No more than one compressor at a time should operate at part load.
6. Utilize the waste heat.

REFERENCES

1. Rollins, J. P., Ed., *Compressed Air and Gas Handbook*, 5th ed., Compressed Air and Gas Institute, Cleveland, OH, 1989.
2. Soedel, W., Ed., *Proceedings of 1972 Purdue Compressor Technology Conference, 90 Papers*, 568 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 25–27, 1972.
3. Soedel, W., Ed., *Proceedings of 1974 Purdue Compressor Technology Conference, 73 Papers*, 485 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 10–12, 1974.
4. Hamilton, J. F., Ed., *Proceedings of 1976 Purdue Compressor Technology Conference, 81 Papers*, 551 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 6–9, 1976.
5. Hamilton, J. F., Ed., *Proceedings of 1978 Purdue Compressor Technology Conference, 55 Papers*, 400 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 19–21, 1978.
6. Soedel, W., Ed., *Proceedings of 1980 Purdue Compressor Technology Conference, 69 Papers*, 420 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 23–25, 1980.
7. Cohen, R., Ed., *Proceedings of 1982 Purdue Compressor Technology Conference, 63 Papers*, 446 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 21–23, 1982.
8. Cohen, R., Ed., *Proceedings of 1984 International Compressor Engineering Conference Purdue, 91 Papers*, 651 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 11–13, 1984.

9. Hamilton, J. F. and Cohen, R., Eds., *Proceedings of 1986 International Compressor Engineering Conference Purdue, Vol. I, II, and III, 84 Papers*, 1201 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, August 4–7, 1986.
10. Soedel, W., Ed., *Proceedings of 1988 International Compressor Engineering Conference Purdue, Vol. I and II, 89 Papers*, 666 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 18–21, 1988.
11. Soedel, W., Ed., *Proceedings of 1990 International Compressor Engineering Conference Purdue, Vol. I and II, 109 Papers*, 982 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 17–20, 1990.
12. Hamilton, J. F., Ed., *Proceedings of 1992 International Compressor Engineering Conference Purdue, Vol. I, II, III and IV, 160 Papers*, 1558 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 14–17, 1992.
13. Soedel, W., Ed., *Proceeding of 1994 International Compressor Engineering Conference Purdue, Vol. I and II, 127 Papers*, 824 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 19–22, 1994.
14. Soedel, W., Ed., *Proceedings of 1996 International Compressor Engineering Conference Purdue, Vol. I and II, 133 Papers*, 856 pages, Purdue University, Ray W. Herrick Laboratories, West Lafayette, IN, July 23–26, 1996.
15. Soedel, W., Ed., *Proceedings of 1998 International Compressor Engineering Conference Purdue, Vol. I and II, 153 Papers*, 950 pages, Purdue University, Ray W. Herrick Laboratories, July 14–17, 1998.
16. ASHRAE., *ASHRAE Handbook: HVAC Systems and Equipment*, American Society of Heating Refrigerating, and Air Conditioning Engineers, Inc., Atlanta, GA, 1996, Chap. 34.
17. Cohen, R. and Groll, E. A., Update on refrigerant compressors in light of CFC substitutes, *Bull. Int. Inst. Refrig.*, 96(5), LXXVI, 1996.
18. Bell, I. H., *Theoretical and experimental analysis of liquid flooded compression in scroll compressors*, Ph.D. Dissertation, Purdue University, 2011.
19. Bell, I. H., Groll, E. A., Braun, J. E., Horton, W. T., Experimental testing of an oil-flooded hermetic scroll compressor, *Int. J. Refrig.*, 36(7), 1866–1873, 2013.
20. Ramaraj, S., Yang, B., Braun, J. E., Groll, E. A., Horton, W. T., Experimental analysis of oil flooded R410A scroll compressor, *Int. J. Refrig.*, 46, 185–195, 2014.
21. Ribas, F. A., Deschamps, C. J., Fagotti, F. et al., Thermal analysis of reciprocating compressors—A critical review, *International Compressor Engineering Conference*, Purdue University, West Lafayette, IN, Paper 1306, July 14–17, 2008.
22. Adair, R. P., Qvale, E. B., Person, J. T., Instantaneous heat transfer to the cylinder wall on reciprocating compressors. International Compressor Engineering Conference, Purdue University, West Lafayette, IN, pp. 521–526, July 15–18, 1972.
23. Brok, S. W., Touber, S., van der Meer, J. S., Modeling of cylinder heat transfer—Large effort, little effect? International Compressor Engineering Conference, Purdue University, West Lafayette, IN, pp. 43–50, July 20–23, 1980.
24. Hsieh, W. H., Wu, T. T., Experimental investigation of heat transfer in a high-pressure reciprocating gas compressor, *Exp. Therm. Fluid Sci.*, 13, 44–54, 1996.
25. Disconzi, F. P., Pereira, E. L., Deschamps, C. J., Development of an in-cylinder heat transfer correlation for reciprocating compressors*, International Compressor Engineering Conference*, Purdue University, West Lafayette, IN, Paper 1342, July 16–19, 2012.
26. Yang, B., Zhao, Y., Piston ring-cylinder liner lubrication analysis in a CO_2 refrigeration reciprocating compressor. *Proceedings IMechE,* Vol. 225 Part C: J. Mechanical Engineering Science, University of London, London, pp. 2638–2648, September 5–6, 2011.
27. Xin, D. B., Feng, J. M., Ding, L. Q. et al., Experimental investigation of pressure distribution between the piston rings and its formation in reciprocating compressors. *Proceedings IMechE,* Vol. 225 Part C: J. Mechanical Engineering Science, University of London, London, pp. 2701–2712, September 5–6, 2011.
28. Bloch, H. P., *A Practical Guide to Compressor Technology*, John Wiley & Sons, New York, p. 67, 2006.
29. Danny, M., Deffenbaugh, A., Smalley, K., *Advanced Reciprocating Compression Technology*, United States Department of Energy, 2005.
30. Kurtulus, O., Yang, B., Lumpkin, D., Performance and operating characteristics of a novel positive-displacement oil free CO2 compressor, *International Compressor Engineering Conference*, Purdue University, West Lafayette, IN, Paper 1644, July 14–17, 2014.

31. Kinjo, K., Nakano, A., Hikichi, T. et al., Study on CFD considering valve behavior in reciprocating compressor, *International Compressor Engineering Conference*, Purdue University, West Lafayette, IN, Paper 1256, July 12–15, 2010.
32. Stosic, N., Smith, I., Kovacevic, A., *Screw Compressors—Mathematical Modelling and Performance Calculation*, Springer, New York, 2005.
33. Kovacevic, A., Rane, S., Stosic, N., Jiang, Y., Lowry, S., Furmanczyk, M., Influence of approaches in CFD Solvers on performance prediction n screw compressors, *International Compressor Engineering Conference*, Purdue University, West Lafayette, IN, Paper 2252, July 14–17, 2014.
34. Jacobs, J. S., Variable speed tri-rotor screw compressor technology, *International Engineering Conference*, Purdue University, West Lafayette, IN, Paper 1825, July 17–20, 2006.
35. Ziviani, D., Bell, I. H., De Paepe, M., van den Broek, M., Update on single-screw expander geometry model integrated into an open-source simulation tool, *IOP Conference Series: Materials Science and Engineering,* 90, 012064, 2015.
36. Wu, W., Hao, X., He, Z., Li, J., Design of the curved flank for the star-wheel tooth in single screw compressors, *J. Mech. Des.*, 136, 1–5, 2014.
37. Cumpsty, N. A., *Compressor Aerodynamics*, Longman, Harlow, 1989.
38. Shephard, D. G., *Principles of Turbomachinery*, Macmillan, New York, 1956.
39. Wilson, D. G., *The Design of High-Efficiency Turbomachinery and Gas Turbines*, MIT Press, Cambridge, MA, 1984.
40. Ferguson, T. B., *The Centrifugal Compressor Stage*, Butterworth & Co. Ltd., London, 1963.
41. Horlock, J. H., *Axial Flow Compressors*, Butterworth & Co. Ltd., London, 1958.
42. Simon, H., Wallmann, T., and Moenk, T., *Improvements in Performance Characteristics of Single-Stage and Multistage Centrifugal Compressors by Simultaneous Adjustment of Inlet Guide Vanes and Diffuser Vanes*, ASME Paper No. 86-GT-127, American Society of Mechanical Engineers, New York, 1986.
43. Reisweber, R. C., Glessner, J. W., and Shields, J. R., *Design and Development of a Supercharger for a Pressure-Fired Boiler*, ASME Paper No. 58-SA-25, American Society of Mechanical Engineers, New York, 1958.
44. Edwards, P., Basic efficiency steps help total plant electrical costs, *Facil. Eng.*, 25, 9, 1998.

4.12 PUMPS AND FANS

Robert F. Boehm

Introduction

Pumps are devices that impart a pressure increase to a liquid. Fans are used to increase the velocity of a gas, but this is also accomplished through an increase in pressure. The pressure rise found in pumps can vary tremendously, and this is a very important design parameter along with the liquid flow rate. This pressure rise can range from simply increasing the elevation of the liquid to increasing the pressure hundreds of atmospheres. Fan applications, on the other hand, generally deal with small pressure increases. In spite of this seemingly significant distinction between pumps and fans, there are many similarities in the fundamentals of certain types of these machines as well as with their application and theory of operation. The final section of this chapter discusses vacuum pumps in a very brief fashion. Some of these devices are very similar to conventional pumps discussed earlier, and some are considerably different. This portion of the chapter is very much an overview of the topic.

The appropriate use of pumps and fans depends upon the proper choice of device and the proper design and installation for the application. A check of sources of commercial equipment shows that many varieties of pumps and fans exist. Each of these had special characteristics that must be appreciated for achieving proper function. Preliminary design criteria for choosing between different types are given by Boehm (1987). An abbreviated version of this is given in Table 4.12.1 for pumps.

As is to be expected, the wise applications of pumps and fans require knowledge of fluid flow fundamentals. Unless the fluid mechanics of a particular application are understood, the design could be less effective than is desirable. Some ideas from fluid mechanics that are especially relevant to pump and fan operation are reviewed.

Pumps

Raising of water from wells and cisterns is the earliest form of pumping (a very detailed history of early applications is given by Ewbank, 1842). Current applications are much broader, and these find a wide variety of machines in use. Modern pumps function on one of two principles. By far, the majority of pump installations are of the *velocity head* type. In these devices, the pressure rise is achieved by giving the fluid a movement. At the exit of the machine, this movement is translated into a pressure increase by slowing the fluid velocity. The other major type of pump is called *positive displacement*. These devices are designed to increase the pressure of the liquid while essentially trying to compress the volume. A categorization of pump types has been given by Krutzsch and Cooper (2008), and an adaptation of this is shown below.

I. Velocity head
   A. Centrifugal
      1. Axial flow (single or multistage)
      2. Radial flow (single or double suction)
      3. Mixed flow (single or double suction)
      4. Peripheral (single or multistage)

   B. Special effect
      1. Gas lift
      2. Jet
      3. Hydraulic ram
      4. Electromagnetic

TABLE 4.12.1
Data for the Preliminary Selection of Pumps

	Axial Flow	Centrif. Radial	Turbine (Regen)	External Gear	Screw	Sliding Vane	Piston	Diaphr.	Jet
Max P (bars)	350	350	50	350	350	350	1000	350	350
Min T (°C)	–240	–240	–30	–30	–30	–30	–30	–30	–240
Max T (°C)	500	500	250	400	370	270	370	270	500
ΔP/stage (bar)	2	20	35	200	20	150	1500	70	1
Max flow (m^3/s)	5	10	1	0.1	0.1	0.1	0.03	0.006	1
Max viscosity (Pa·s)		0.2	0.1	400	1000	100	400	100	
Efficiency (%)	50–85	50–85	20–40	40–85	40–70	40–85	60–90	40–70	5–20
Relative purchase cost	Low	Low	Mod	Mod	High	Mod	High	Mod	Low

Source: Ulrich, G., Vasude, P., *Chemical Engineering Process Design and Economics: A Practical Guide,* Process Publishing, 2004.

II. Positive displacement
  A. Reciprocating
    1. Piston, plunger
      a. Direct acting (simplex or duplex)
      b. Power (single or double acting, simplex, duplex, triplex, multiplex)
    2. Diaphragm (mechanically or fluid driven, simplex or multiplex)

  B. Rotary
    1. Single rotor (vane, piston, screw, flexible member, peristaltic)
    2. Multiple rotor (gear, lobe, screw, circumferential piston)

Many factors impact the selection of a pump for a particular application, including the flows anticipated (maximum, minimum, and design, and whether they are continuous or variable), the pressures required at these flows, the type of fluids handled, as well as many other factors. A variety of pump types are available to the system designer. Several of these are compared in Table 4.12.1, and the original reference makes several more comparisons.

In the next two sections, two of the more common categories of pumps are described: centrifugal and positive displacement. There are several subcategories under each of these types, but these will not be delineated in the information given here. References are included that should allow the latter type of information to be found.

Centrifugal and Other Velocity Head Pumps

Centrifugal pumps are used in more industrial applications than any other kind of pump. This is primarily because these pumps offer low initial and upkeep costs. Traditionally, pumps of this type have been limited to low-pressure-head applications, but modern pump designs have overcome this problem unless very high pressures are required. Some of the other good characteristics of these types of devices include smooth (nonpulsating) flow and the ability to tolerate nonflow conditions.

A representative set of typical kind of variations for pumps of this type is shown in Figure 4.12.1. This information can be used for proper pump application.

The most important parts of the centrifugal pump are the *impeller* and *volute*. An impeller can take on many forms, ranging from essentially a spinning disk to designs with elaborate vanes.

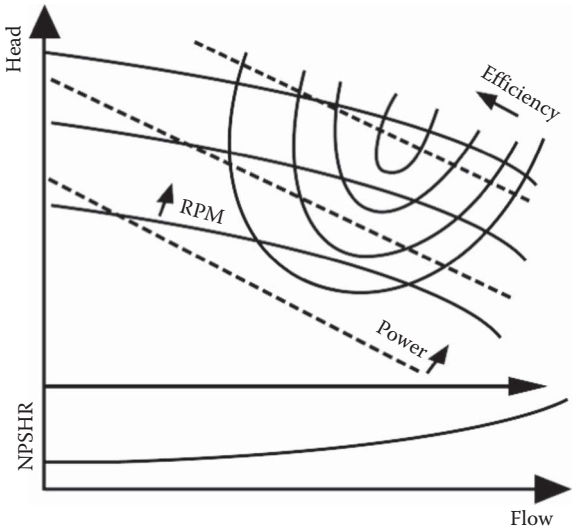


FIGURE 4.12.1 Many pump operating parameters are usually available from the pump manufacturer. An example of typical centrifugal pump parameter variations is shown in this figure.

The latter is usual. Impeller design tends to be somewhat unique to each manufacturer, as well as finding a variety of designs for a variety of applications. This device imparts a radial velocity to the fluid that has entered the pump perpendicular to the impeller. The volute (there may be one or more) performs the function of slowing the fluid and increasing the pressure. A good discussion of centrifugal pumps has been given by Lobanoff and Ross (1992) and in the ASHRAE Handbook (2012).

Very important factors in the specification of a centrifugal pump are the *casing orientation* and *type*. For example, the pump can be oriented vertically or horizontally. Horizontal mounting is most common. Vertical pumps usually offer benefits related to ease of priming and reduction in required net positive suction head (see discussion below). This type also requires less floor space. Submersible and immersible pumps are always of the vertical type. Another factor in the design is the way the casing is split, and this has implications about ease of manufacture and repair. Casings that are split perpendicular to the shaft are called *radially split*, while those split parallel to the shaft axis are denoted as *axially split*. The latter can be *horizontally split* or *vertically split*. The number of *stages* in the pump greatly affects the pump-output characteristics. Several stages can be incorporated into the same casing, with an associated increase in pump output. Multistage pumps are often used for applications with total developed head over 50 atm.

In general, centrifugal pumps become more efficient in larger sizes. This can be seen in the experimental data compiled by Rishel (2000). See Figure 4.12.2.

Whether or not a pump is self-priming can be important. If a centrifugal pump is filled with air when it is turned on, the initiation of pumping action may not be sufficient to bring the fluid into the pump. Pumps can be specified with features that can minimize priming problems.

There are other types of velocity head pumps. *Jet pumps* increase pressure by imparting momentum from a high-velocity liquid stream to a low-velocity or stagnant body of liquid. The resulting flow then goes through a diffuser to achieve an overall pressure increase. *Gas lifts* accomplish a pumping action by a drag on gas bubbles that rise through a liquid.

Positive-Displacement Pumps

Positive-displacement pumps demonstrate high discharge pressures and low flow rates. Usually, this is accomplished by some type of pulsating device. A piston pump is a classic example of

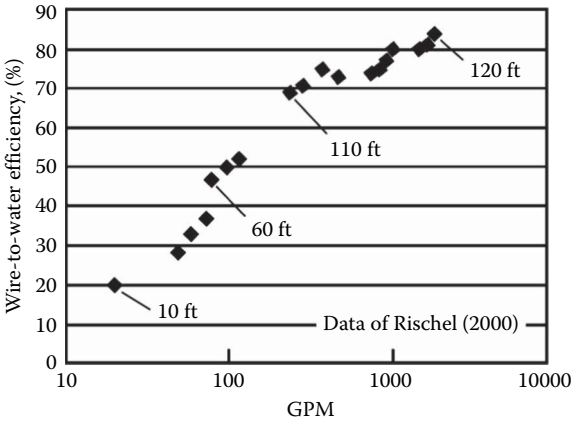


FIGURE 4.12.2 "Wire-to-water" efficiencies for centrifugal pumps shown are based upon the data of Rishel (2000). The head varies throughout the data within the ranges shown.

positive-displacement machines. Rotary pumps are one type of positive-displacement device that do not impart pulsations to the existing flow (a full description of these types of pumps is given by Turton, 1994). Several techniques are available for dealing with pulsating flows, including use of double-acting pumps (usually of the reciprocating type) and installation of pulsation dampeners.

Positive-displacement pumps usually require special seals to contain the fluid. Costs are higher both initially and for maintenance compared with most pumps that operate on the velocity head basis. Positive-displacement pumps demonstrate an efficiency that is nearly independent of flow rate, in contrast to the velocity head type.

Reciprocating pumps offer very high efficiencies, reaching 90% in larger sizes. These types of pumps are more appropriate for pumping abrasive liquids (e.g., slurries) than are centrifugal pumps.

A characteristic of positive displacement pumps that may be valuable is that the output flow is proportional to pump speed. This allows this type of pump to be used for metering applications. Also a positive aspect of these pumps is that they are self-priming, except at initial start-up.

Very high head pressures (often damaging to the pump) can be developed in positive-displacement pumps if the downstream flow is blocked. For this reason, a pressure-relief-valve bypass must always be used on the high-pressure side of positive-displacement pumps.

Pump/Flow Considerations

Performance characteristics of the pump must be considered in system design. Simple diagrams of pumping systems are shown in Figure 4.12.3. First, consider the left-hand figure. This represents a flow circuit, and the pressure drops related to the piping, fittings, valves, and any other flow devices found in the circuit must be estimated using the laws of fluid mechanics. Usually, these resistances (pressure drops) are found to vary approximately with the square of the liquid flow rate. Typical characteristics are shown in Figure 4.12.4. Most velocity head pumps demonstrate a flow vs. pressure rise variation that is a positive value at zero flow and decreases to zero at some larger flow. Positive-displacement pumps, as shown on the right-hand side of Figure 4.12.4, are an exception to this in that these devices usually cannot tolerate a zero flow. An important aspect to note is that a closed system can presumably be pressurized. A contrasting situation and its implications are discussed below.

It is important to maintain the substance inside a pump as a liquid. If conditions exist that convert the flow to a vapor, serious shortcomings in performance can be exhibited that could cause damage to the pump. At the minimum the pump flow will decrease, possibly in a substantial way.

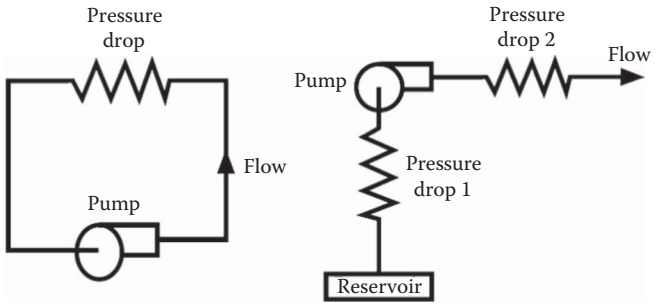


FIGURE 4.12.3 Typical pump applications, either in circuits (see left-hand side above) or once-through systems (right-hand side), can be represented with the frictional drops as shown as resistances. The latter are determined by fluid mechanics analysis.

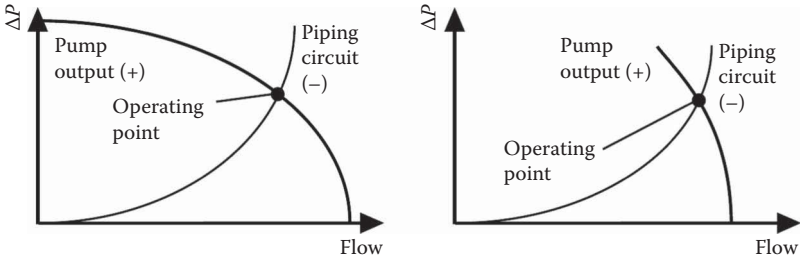


FIGURE 4.12.4 An overlay of the pump flow vs. head curve with the circuit piping characteristics gives the operating state of the circuit. A typical velocity head characteristic is shown on the left, while a positive-displacement pump curve is shown on the right.

The piping diagram shown on the right-hand side of Figure 4.12.3 is a once-through system, another frequently encountered installation. However, the leg of piping through "Pressure Drop 1" shown there can have some very important implications related to *net positive suction head*, often denoted as *NPSH*. In simple terms, NPSH indicates the difference between the local pressure and the thermodynamic saturation pressure at the fluid temperature. If NPSH=0, the liquid can vaporize, and this can result in a variety of outcomes from noisy pump operation to outright failure of components. This condition is called *cavitation*. Cavitation, if it occurs, will first take place at the lowest pressure point within the piping arrangement. Often this point is located at, or inside, the inlet to the pump. Most manufacturers specify how much NPSH is required for satisfactory operation of their pumps. Hence, the actual NPSH (denoted as *NPSHA*) experienced by the pump must be larger than the manufacturer's required NPSH (called *NPSHR*). If a design indicates insufficient NPSH, changes should be made in the system, possibly including alternative piping layout, including elevation and/or size, or use of a pump with smaller NPSH requirements.

The manufacturer should be consulted for a map of operational information for a given pump. A typical form has been shown in Figure 4.12.1. This information will allow the designer to select a pump that satisfied the circuit operational requirements while meeting the necessary NPSH and most-efficient operation criteria.

Several options are available to the designer for combining pumps in systems. Consider a comparison of the net effect between two identical pumps operating in series or operating the same two pumps in parallel. Examples of this are for pumps such as centrifugal units that have characteristics similar to those shown in Figure 4.12.1. It is clear that one way to achieve high pumping pressures with centrifugal pumps is to place a number of units in series. This is a related effect to what is essentially found in *multistage* designs. However, if higher flows are needed at the same pressure rise, similar pumps can be operated in parallel.

Fans

As noted earlier, fans are devices that cause air to move. This definition is broad and can include a flapping palm branch, but the discussion here deals only with devices that impart air movement due to *rotation of an impeller inside a fixed casing*. In spite of this limiting definition, a large variety of commercial designs are included.

Fans find application in many engineering systems. Along with the chillers and boilers, they are the heart of heating, ventilating, and air conditioning (HVAC) systems. When large physical dimensions of a unit are not a design concern (usually the case), centrifugal fans are favored over axial flow units for HVAC applications. Many types of fans are found in *power plants*. Very large fans are used to furnish air to the boiler, as well as to draw or force air through cooling towers and pollution-control equipment. *Electronic cooling* finds applications for small units. Even automobiles have several fans in them. Because of the great engineering importance of fans, several organizations publish rating and testing criteria (see, for example, ASME, 1990).

Generally, fans are classified according to how the air flows through the impeller. These flows may be *axial* (essentially a propeller in a duct), *radial* (conceptually much like the centrifugal pumps discussed earlier), *mixed*, and *cross*. While there are many other fan designations, all industrial units fit one of these classifications. Mixed-flow fans are so named because both axial and radial flows occur on the vanes. Casings for these devices are essentially like those for axial-flow machines, but the inlet has a radial-flow component. On cross-flow impellers, the gas traverses the blading twice.

Characteristics of fans are shown in Figure 4.12.5. Since velocities can be high in fans, often both the total and the static pressure increases are considered. While both are not shown in this figure, the curves have similar variations. Of course the total ΔP will be greater than the static value, the difference being the velocity head. This difference increases as the volume flow increases. At the zero flow (the shutoff point), the static and total pressure difference values are the same. Efficiency variation can show a sharp optimum value at the design point. For this reason, it is critical that fan designs be carefully tuned to the required conditions.

A variety of vane types are found on fans, and the types of these are also used for fan classification. Axial fans usually have vanes of airfoil shape or vanes of uniform thickness. Some vane types that might be found on a centrifugal (radial-flow) fan are shown in Figure 4.12.6.

One of the more important aspects that is an issue in choosing fans for a particular application is fan efficiency. Typical efficiency comparisons of the effect of blade type on a centrifugal fan are shown in Figure 4.12.7. Since velocities can be high, the value of aerodynamic design is clear. Weighing against this is primarily cost but there could be other factors.

An additional aspect that may be important in the choice of fans is noise generation. This may be most critical in HVAC applications. It is difficult to describe noise characteristics in brief terms

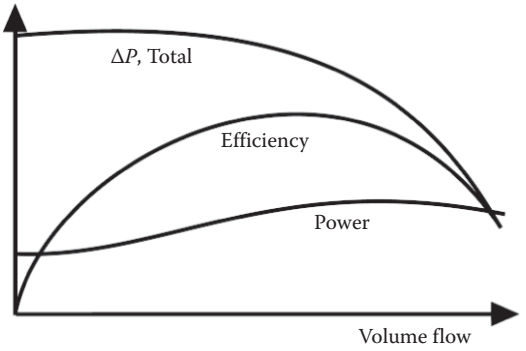


FIGURE 4.12.5 Shown are typical characteristics of a centrifugal fan. The drawbacks of operating away from the optimal conditions are obvious from the efficiency variation.

because of the frequency-dependent nature of these phenomena. However, a comparison of specific sound power level (usually denoted by K_w) shows backward-curved centrifugal fans with aerodynamic blades perform best among the designs. Details of noise characteristics are given elsewhere (ASHRAE, 2011). One point to note is that fan sound is usually at a minimum near the point of maximum fan efficiency.

While each type of fan has some specific qualities for certain applications (some comparisons are given in Table 4.12.2), most installations use centrifugal (radial-flow) fans. A primary exception is for very-high-flow, low-pressure-rise situations, where axial (propeller) fans are used.

Similarities exist between fans and pumps because the fluid density essentially does not vary through either type of machine. Of course, in pumps this is because a liquid can be assumed to be incompressible. In fans, a gas (typically air) is moved with little pressure change. As a result, the gas density can be taken to be constant. Since most fans operate near atmospheric pressure, the ideal gas assumptions can be used in determining gas properties.

Flow control in fan applications, where needed, is a very important design concern. Methods for accomplishing this involve use of dampers (either on the inlet or on the outlet of the fan), variable

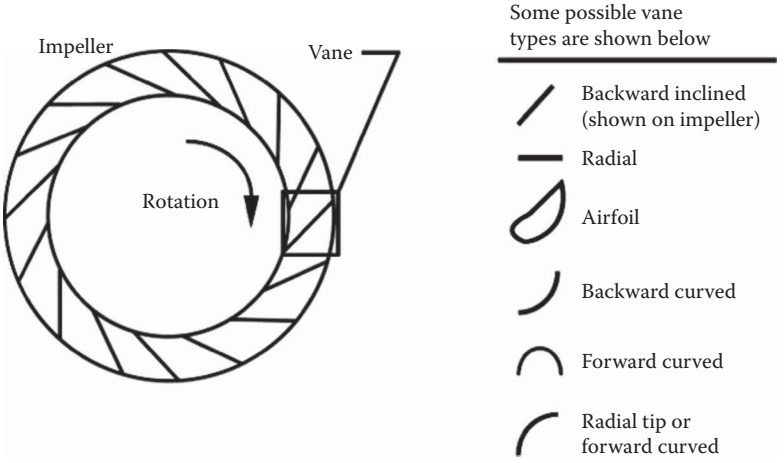


FIGURE 4.12.6 A variety of vane types that might be used on a centrifugal fan are shown.

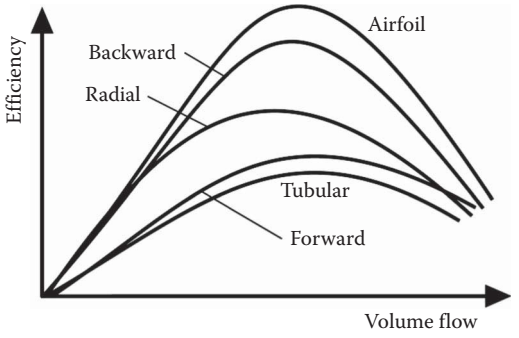


FIGURE 4.12.7 Efficiency variation vs. volume flow rate of centrifugal fans for a variety of vane types is shown.

TABLE 4.12.2
Characteristics of Various Fans

Type of Fan	Specific Speed Range (RPM)	Air Volume (CFM)	Static Pressure	Fan Efficiency	Manufacturing Cost	Size and Weight
Propeller	500,000–100,000	Large	Low	Low	Low	Small
Tubeaxial	300,000–60,000	Large	Low	Medium	Low	Small
VAF, one stage	130,000–50,000	Large	Medium	High	Medium	Small
VAF, two stages	90,000–35,000	Large	Medium	High	Medium	Medium
Mixed flow	80,000–10,000	Medium	Medium	Low	High	Medium
Wide FC	70,000–25,000	Large	High	Medium	Medium	Large
AF, BC, BI	70,000–20,000	Medium	Medium	High	Medium	Medium
Radial tip	65,000–25,000	Medium	Medium	Medium	Medium	Medium
Radial blades	25,000–10,000	Low	Medium	Low	Medium	Medium
Turbo blower	30,000–5,000	Small	High	Medium	Medium	Medium
Multistage turbo blower	8,000–1,000	Small	High	Low	High	Large

Source: Bleier, F.P., *Fan Handbook: Selection, Application, and Design*, McGraw-Hill, New York, 1998, p. 8.5.
Note: VAF, vaneaxial fan; BC, backward curved; FC, forward curved; BI, backward inclined; AF, airfoil.

pitch vanes, or variable speed control. Dampers are the least expensive to install, and also the most inefficient in terms of energy use. Modern solid state controls for providing a variable frequency power to the drive motor is becoming the preferred control method, when a combination of initial and operating costs is considered.

Types of drive motors can be important to the design of a fan installation. Some guidelines for doing this are given here (Adapted from Bleier (1998) p. 16.2).

1. The most frequently used motors whenever a fan requires 1 hp or more are three-phase squirrel-cage types.
2. Sometimes adjustable-speed drives use three-phase wound rotor motors, but these can be expensive.
3. Single-phase, split-phase induction motors, with an auxiliary winding for starting function, are used for applications up to 1/2 hp. While these tend to be less efficient, they are about as good as can be applied for single-phase power. They require a high starting current.
4. For requirements up to 1/3 hp, single-phase, permanent-split-capacitor motors are used. A capacitor is used instead of the auxiliary winding noted in number 3 (above). A low starting current is realized but is associated with a lower efficiency. Many applications find these used in direct drive arrangements.
5. Up to ¾ hp, single-phase, capacitor-start-and-run, capacitor-run motors are used. Because of an auxiliary winding in series with a capacitor, they are able to furnish high starting torques.
6. For high-speed applications up to 1 hp like vacuum cleaners, single-phase universal motors are used. Although they can be used in direct drives, they have the disadvantage of slightly lower efficiencies and brushes that may need adjustment and replacement.

Drives are also important. A few suggestions Adapted from Bleier (1998, p. 16.2) are shown below.

1. Three-phase motors are used with belt drive for most large fans.
2. Single-phase motors with direct drive are used for most small fans.

3. Belt drive by electric motors or by engines is sometimes used for multistage turbo blowers.
4. Gear drives from electric motors running at very high speeds are used to drive single-stage turbo blowers.

Vacuum Pumps

While this is a topic that involves pumps and systems that are related to conventional pumps and fans, many aspects of the field are quite distinct from what is given earlier in this review. Only a limited amount of information will be listed here comparing the general kinds of pumps used in these systems. The reader who wishes more information can consult the many books on this topic including the one by Hablanian (1997). In these types of applications, various kinds of pumps might be applied for varying degrees of vacuum. Approximate vacuum ranges and relative costs of some of these pumps are shown in Figure 4.12.8.

Pumps for vacuum systems range from conventional pumps discussed earlier for "roughing" applications to very specialized pumps for achieving high vacuum. The general categories shown in Figure 4.12.8 will be discussed briefly in what follows. It should be noted, however, that many more categories of pumps can be defined than are given here.

Mechanical pumps are used to perform "roughing" (roughing is where relatively large flows of evacuated gases are found) processes. That is, they are used to remove the initial amount of gas from the region where a vacuum is desired. Any type of conventional compressor could be used to perform this duty. Compressor devices using pistons, rotational elements, etc., can be candidates to perform this function.

Moving to lower pressure applications, the Roots unit is an example of a *mechanical blower.* It was developed in second half of the 1800 by two Roots brothers. It consists of lobes and is generally for lower pressure roughing applications than conventional compressors.

Diffusion pumps are widely used devices in vacuum applications. This type of unit is a vapor jet pump (a name by which it is also known), and it is the most widely used approach in vacuum systems. It transports gas by momentum transfer as a result of a collision with a vapor stream. Typically, the vapor stream is an organic fluid or mercury. Several stages of compression (usually 3–5) are usually used.

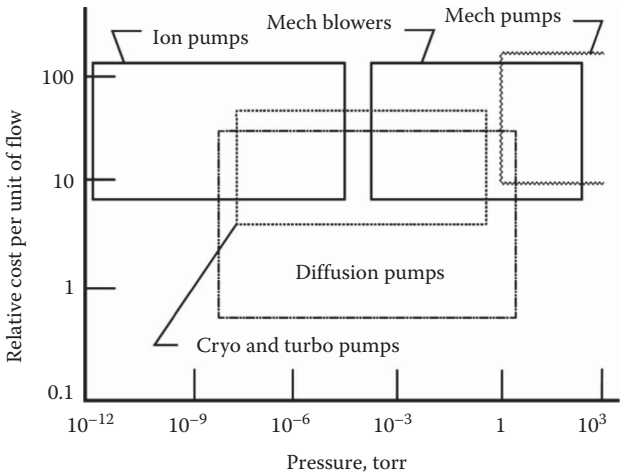


FIGURE 4.12.8 General categories of vacuum pumps are shown with approximate ranges of costs and application pressures. (Adapted from Hablanian, M.H., *High-Vacuum Technology, a Practical Guide*, 2nd ed., Marcel Dekker, New York, 1997.)

A device that acts like a molecular turbine is the basis of the operation of a *turbomolecular pump.* In simple terms, it can be viewed as the special adaptation of an axial flow compressor, as it compresses gas by a momentum transfer from a high-speed rotating element. Several stages of compression are used because the pressure increase per stage is small.

Lowering the temperature below that which will condense a gas is the technical basis of the *cryogenic pump*. These pumps include a very cold porous surface (not the whole pump) that will perform this action. Typically, helium is used as the refrigerant to cool the active surface. This type of pumping is very clean because no foreign fluid comes in contact with the pumped flow.

Several types of devices come under the general name of *ion pumps*. *Gettering pumps* are a form of these configurations where a source of unoxidized titanium (usually in film form) is in contact with the gas that is being pumped. Another approach is to generate ions in the gas being pumped, which will then be energetic and attracted to a surface.

DEFINING TERMS

Cavitation: Local liquid conditions allow vapor voids to form (boiling).

NPSH: Net positive suction head is the difference between the local absolute pressure of a liquid and the thermodynamic saturation pressure of the liquid based upon the temperature of the liquid. Applies to the inlet of a pump.

NPSHA: Actual net positive suction head is the NPSH at the given state of operation of a pump.

NPSHR: Required net positive suction head is the amount of NPSH required by a specific pump for a given application.

REFERENCES

ASHRAE, 2011. *ASHRAE Handbook 2011, HVAC Applications*, American Society of Heating, Refrigerating, and Air Conditioning Engineers, Atlanta, Chapter 48.

ASHRAE, 2012. *ASHRAE Handbook 2012, HVAC Systems and Equipment*, American Society of Heating, Refrigerating, and Air Conditioning Engineers, Atlanta, Chapter 44.

ASME, 1990. *ASME Performance Test Codes, Code on Fans*, ASME PTC 11–1984 (reaffirmed 1990), American Society of Mechanical Engineers, New York.

Bleier, F. P., 1998. *Fan Handbook: Selection, Application, and Design*, McGraw-Hill, New York.

Boehm, R. F., 1987. *Design Analysis of Thermal Systems*, John Wiley and Sons, New York, pp. 17–26.

Ewbank, T., 1842. *A Description and Historical Account of Hydraulic and Other Machines for Raising Water*, 2nd ed., Greeley and McElrath, New York.

Hablanian, M. H., 1997. *High-Vacuum Technology, a Practical Guide*, 2nd ed., Marcel Dekker, NewYork.

Krutzsch, W. C. and Cooper, P., 2008. Chapter preliminaries, in *Pump Handbook*, 4nd ed., I. Karassik et al., Eds., McGraw-Hill, New York, Chapter 1.

Lobanoff, V. and Ross, R., 1992. *Centrifugal Pumps: Design & Application*, 2nd ed., Gulf Publishing Company, Houston, TX.

Rishel, J. B., 2000. 40 years of fiddling with pumps, *ASHRAE Journal*, 42(3), 43–48.

Turton, R. K., 1994. *Rotodynamic Pump Design*, Cambridge University Press, Cambridge, UK.

Ulrich, G. and Vasude, P., 2004. *Chemical Engineering Process Design and Economics: A Practical Guide,* Process Publishing, Durham, NH, pp. 250–251.

Further Information

Dickson, C., 1988. *Pumping Manual*, 8th ed., Trade & Technical Press, Morden, UK.

Dufour, J. and Nelson, W., 1992. *Centrifugal Pump Sourcebook*, McGraw-Hill, New York.

Fans, 2012. In *2012 ASHRAE Handbook, HVAC Systems and Equipment*, American Society of Heating, Refrigerating, and Air Conditioning Engineers, Atlanta, GA, Chapter 21.
Garay, P.N., 1996. *Pump Application Desk Book*, 3rd ed., Fairmont Press, Liburn, GA.
Karassik, Paul, et al., 2007, *Pump Handbook*, 4th ed., McGraw-Hill, New York.
Stepanoff, A.J., 1993. *Centrifugal and Axial Flow Pumps: Theory, Design, and Application* (Reprint Edition), Krieger Publishing Company, Malabar, FL.

4.13 COOLING TOWERS

Anthony F. Mills

Introduction

In a wet cooling tower, water is evaporated into air with the objective of cooling the water stream. Both natural- and mechanical-draft towers are popular, and examples are shown in Figure 4.13.1. Large natural-draft cooling towers are used in power plants for cooling the water supply to the condenser. Smaller mechanical-draft towers are preferred for oil refineries and other process industries, as well as for central air-conditioning systems and refrigeration plant. Figure 4.13.1a shows a natural draft *counterflow* unit in which the water flows as thin films down over a suitable packing, and air flows upward. In a natural-draft tower, the air flows upward due to the buoyancy of the warm, moist air leaving the top of the packing. In a mechanical-draft tower, the flow is forced or induced by a fan. Since the air inlet temperature is usually lower than the water inlet temperature, the water is cooled both by evaporation and by sensible heat loss. For usual operating conditions, the evaporative heat loss is considerably larger than the sensible heat loss. Figure 4.13.1b shows a mechanical draft cross-flow unit. Figure 4.13.2 shows a natural-draft cross-flow tower for a power plant.

Packing Thermal Performance

Counterflow Units

Merkel's method (Merkel, 1925) for calculating the number of transfer units required to cool the water stream, for specified inlet and outlet water temperatures and inlet air condition is (Mills, 1995)

$$N_{\text{tu}} = \frac{g_m S}{\dot{m}_L} = \int_{h_{l,\text{in}}}^{h_{l,\text{out}}} \frac{\mathrm{d}h_L}{h_s - h_G} \tag{4.13.1}$$

$$h_G = h_{G,\text{in}} + (\dot{m}_L / \dot{m}_G)(h_L - h_{L,\text{out}}) \tag{4.13.2}$$

$$h_s(P, T_s) = h_s(P, T_L) \tag{4.13.3}$$

It is imperative that the usual enthalpy datum states be used, namely, zero enthalpy for dry air and liquid water at 0°C. Table 4.13.1 gives enthalpy data for 1 atm pressure. The important assumptions required to obtain this result include

1. A Lewis number of unity;
2. Low mass transfer rate theory is valid;
3. The liquid-side heat transfer resistance is negligible, that is, $T_s \simeq T_L$;
4. The amount of water evaporated is small compared with the water and air flow rates.

The method is accurate up to temperatures of about 60°C; comparisons with more exact results are usually within 3%–5%, and seldom show errors greater than 10%. Notice that the method does not give the outlet state of the air; however, in situations encountered in practice, the outlet air can be assumed to be saturated for the purposes of calculating its density. It is possible to extend Merkel's method to include a finite liquid-side heat transfer resistance, but such refinement is seldom warranted. For typical operating conditions, the bulk liquid temperature is seldom more than 0.3 K above the interface temperature.

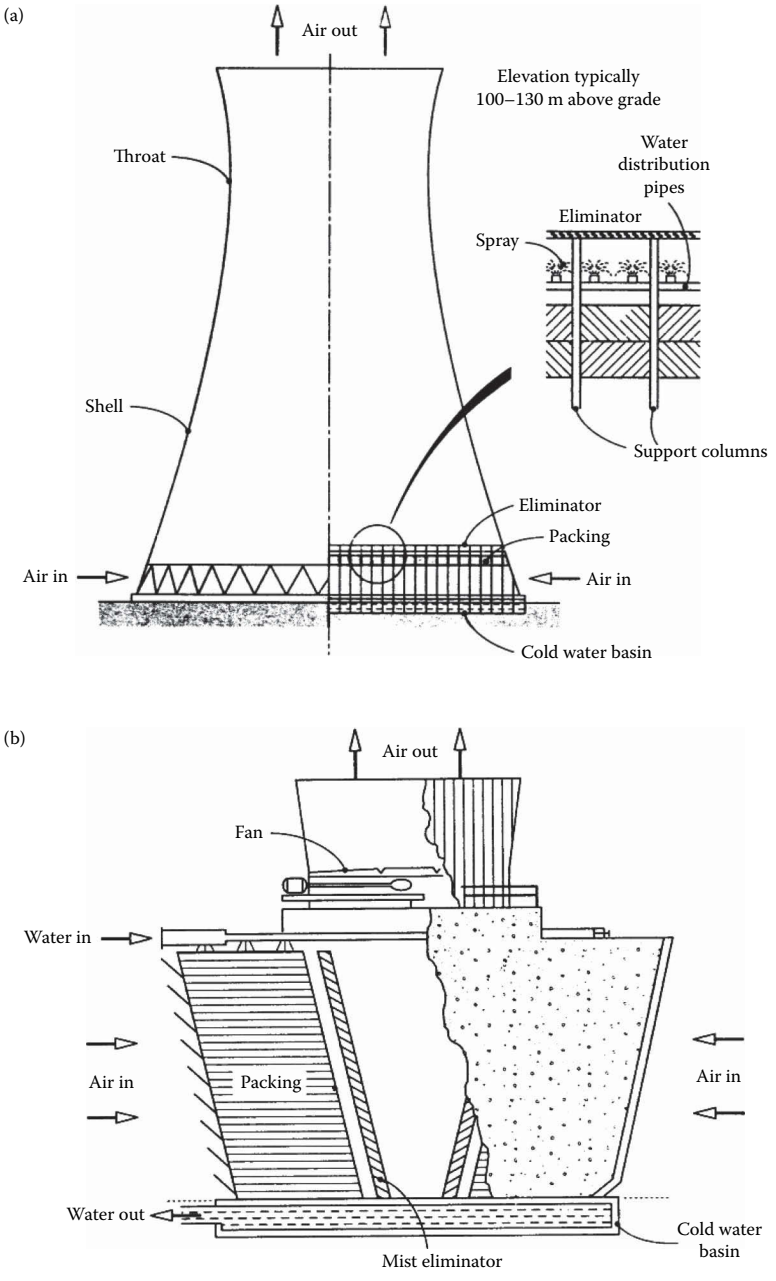


FIGURE 4.13.1 (a) A natural-draft counterflow cooling tower for a power plant. (b) A cross-flow cooling tower for an air-conditioning system.

Cross-Flow Units

Figure 4.13.3 shows a schematic of a cross-flow packing. If we assume that both the liquid and gas streams are unidirectional, and that there is no mixing in either stream, then use of Merkel's assumptions leads to the following pair of differential equations (Mills, 1995):

$$\frac{\partial h_G}{\partial x} = \frac{g_m a}{G}\left(h_s - h_G\right) \tag{4.13.4}$$

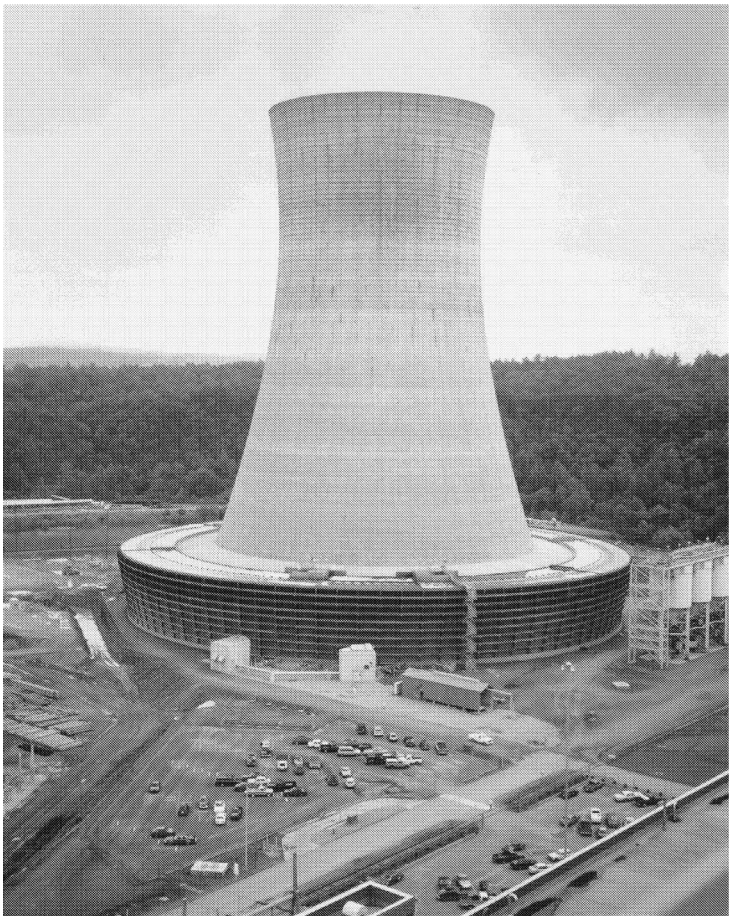


FIGURE 4.13.2 A natural-draft cross-flow cooling tower for a power plant.

$$\frac{\partial h_L}{\partial y} = -\frac{g_m a}{L}\left(h_s - h_G\right) \tag{4.13.5}$$

Also $h_s = h_s(h_L)$ for a negligible liquid-side heat transfer resistance and the required boundary conditions are the inlet enthalpies of both streams. Equations 4.13.4 and 4.13.5 are solved numerically and the solution used to evaluate the average enthalpy of the outlet liquid,

$$\overline{h}_{L,\text{out}} = \frac{1}{X}\int_0^X h_{L,\text{out}}\,\mathrm{d}x \tag{4.13.6}$$

Substituting in an exchanger energy balance on the liquid stream gives the heat transfer as

$$q = \dot{m}_L\left(h_{L,\text{in}} - h_{L,\text{out}}\right) \tag{4.13.7}$$

Sample Calculation

Consider a counterflow unit that is required to cool water from 40°C to 26°C when the inlet air is at 10°C, 1 atm, and saturated. We will calculate the number of transfer units required for balanced flow, that is, $\dot{m}_G/\dot{m}_L = 1$. Equation 4.13.1 is to be integrated numerically, with h_G obtained

TABLE 4.13.1
Thermodynamic Properties of Water Vapor–Air Mixtures at 1 atm

Temp. (°C)	Saturation Mass Fraction	Specific Volume (m^3/kg)		Enthalpy[a,b] (kJ/kg)		
		Dry Air	Saturated Air	Liquid Water	Dry Air	Saturated Air
10	0.007608	0.8018	0.8054	42.13	10.059	29.145
11	0.008136	0.8046	0.8086	46.32	11.065	31.481
12	0.008696	0.8075	0.8117	50.52	12.071	33.898
13	0.009289	0.8103	0.8148	54.71	13.077	36.401
14	0.009918	0.8131	0.8180	58.90	14.083	38.995
15	0.01058	0.8160	0.8212	63.08	15.089	41.684
16	0.01129	0.8188	0.8244	67.27	16.095	44.473
17	0.01204	0.8217	0.8276	71.45	17.101	47.367
18	0.01283	0.8245	0.8309	75.64	18.107	50.372
19	0.01366	0.8273	0.8341	79.82	19.113	53.493
20	0.01455	0.8302	0.8374	83.99	20.120	56.736
21	0.01548	0.8330	0.8408	88.17	21.128	60.107
22	0.01647	0.8359	0.8441	92.35	22.134	63.612
23	0.01751	0.8387	0.8475	96.53	23.140	67.259
24	0.01861	0.8415	0.8510	100.71	24.147	71.054
25	0.01978	0.8444	0.8544	104.89	25.153	75.004
26	0.02100	0.8472	0.8579	109.07	26.159	79.116
27	0.02229	0.8500	0.8615	113.25	27.166	83.400
28	0.02366	0.8529	0.8650	117.43	28.172	87.862
29	0.02509	0.8557	0.8686	121.61	29.178	92.511
30	0.02660	0.8586	0.8723	125.79	30.185	97.357
31	0.02820	0.8614	0.8760	129.97	31.191	102.408
32	0.02987	0.8642	0.8798	134.15	32.198	107.674
33	0.03164	0.8671	0.8836	138.32	33.204	113.166
34	0.03350	0.8699	0.8874	142.50	34.211	118.893
35	0.03545	0.8728	0.8914	146.68	35.218	124.868
36	0.03751	0.8756	0.8953	150.86	36.224	131.100
37	0.03967	0.8784	0.8994	155.04	37.231	137.604
38	0.04194	0.8813	0.9035	159.22	38.238	144.389
39	0.04432	0.8841	0.9077	163.40	39.245	151.471
40	0.04683	0.8870	0.9119	167.58	40.252	158.862
41	0.04946	0.8898	0.9162	171.76	41.259	166.577
42	0.05222	0.8926	0.9206	175.94	42.266	174.630
43	0.05512	0.8955	0.9251	180.12	43.273	183.037
44	0.05817	0.8983	0.9297	184.29	44.280	191.815
45	0.06137	0.9012	0.9343	188.47	45.287	200.980
46	0.06472	0.9040	0.9391	192.65	46.294	210.550
47	0.06842	0.9068	0.9439	196.83	47.301	220.543
48	0.07193	0.9097	0.9489	201.01	48.308	230.980
49	0.07580	0.9125	0.9539	205.19	49.316	241.881

[a] The enthalpies of dry air and liquid water are set equal to zero at a datum temperature of 0°C.

[b] The enthalpy of an unsaturated water vapor–air mixture can be calculated as $h = h_{\text{dry air}} + (m_1/m_{1,\text{sat}})(h_{\text{sat}} - h_{\text{dry air}})$.

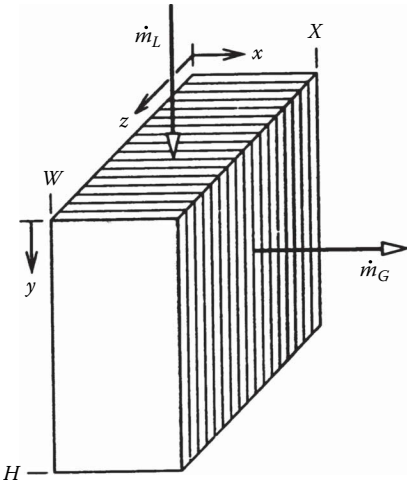


FIGURE 4.13.3 Schematic of a cross-flow cooling tower packing showing the coordinate system.

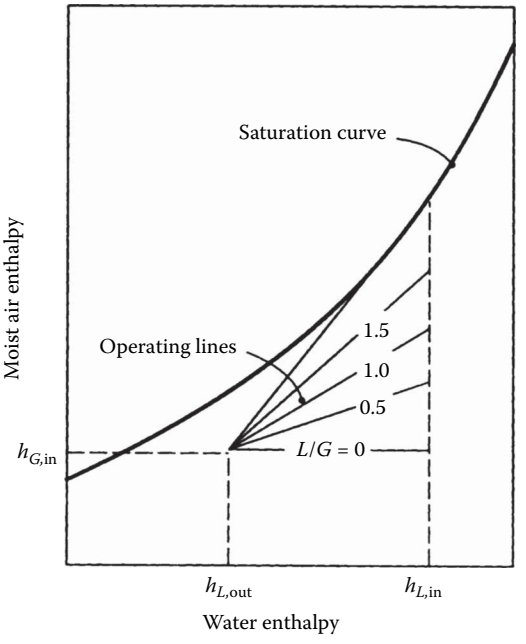


FIGURE 4.13.4 Counterflow cooling tower operating lines for various water-to-air flow-rate ratios shown on an enthalpy chart.

from Equation 4.13.2. The required thermodynamic properties can be obtained from Table 4.13.1. Using Table 4.13.1, $h_{G,\text{in}} = h_{\text{sat}}$ (10°C) = 29.15 kJ/kg, $h_{L,\text{out}} = h_L$(26°C) = 109.07 kJ/kg. Substituting in Equation (4.13.2),

$$h_G = 29.15 + (h_L - 109.07)$$

Choosing 2°C intervals for convenient numerical integration, the table on page 4–865 is constructed, with h_L and $h_s = h_s(T_L)$ also obtained from Table 4.13.1.

Using the trapezoidal rule,

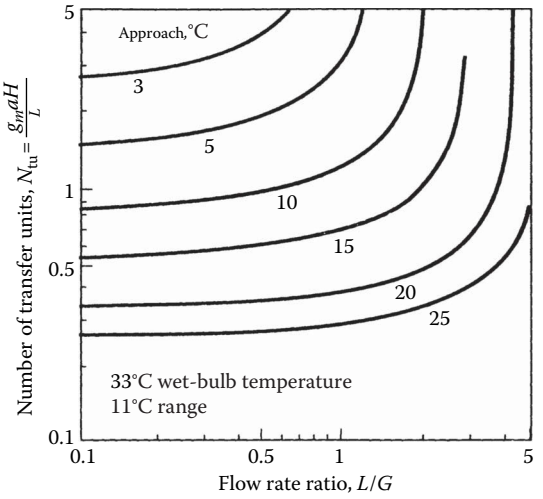


FIGURE 4.13.5 Example of cooling demand curves for a specified wet-bulb temperature and range: N_{tu} vs. flow rate ratio for a fixed approach.

$$\int_{h_{L,\text{out}}}^{h_{L,\text{in}}} \frac{dh_L}{h_s - h_G} = \frac{8.36}{2}[0.02001 + 2(0.01986 + 0.01942 + 0.01871 + 0.01776 + 0.01662 + 0.01536) + 0.01404] = 1.043$$

T_L (°C)	h_L (kJ/kg)	h_G (kJ/kg)	h_s (kJ/kg)	$h_s - h_s$ (kJ/kg)	$\frac{1}{h_s - h_G}$
26	109.07	29.15	79.12	49.97	0.02001
28	117.43	37.51	87.86	50.35	0.01986
30	125.79	45.87	97.36	51.49	0.01942
32	134.15	54.23	107.67	53.44	0.01871
34	142.50	62.58	118.89	56.31	0.01776
36	150.86	70.94	131.10	60.16	0.01662
38	159.22	79.30	144.39	65.09	0.01536
40	167.58	87.66	158.86	71.20	0.01404

From Equation 4.13.1, $N_{tu} = 1.043$. Also, by using Table 4.13.1, $T_{G,\text{out}} = 27.9°$ for saturated outlet air.

Thermal-Hydraulic Design of Cooling Towers

The thermal-hydraulic design of a mechanical-draft cooling tower is relatively straightforward. The flow rate ratio $\dot{m}_L/\dot{m}_G$ can be specified and varied parametrically to obtain an optimal design, for which the size and cost of the packing is balanced against fan power requirements and operating cost. Data are required for mass transfer conductances and friction for candidate packings. Tables 4.13.2a and b give correlations for a selection of packings. In Table 4.13.2b, the mass transfer conductance is correlated as $g_m a/L$, where a is the transfer area per unit volume and $L = \dot{m}_L/A_{fr}$ is the superficial mass velocity of the water flow (also called the *water loading* on the packing). Similarly, we define $G = \dot{m}_G/A_{fr}$. Typical water loadings are 1.8 to 2.7 kg/m^{2}s, and superficial air velocities fall

TABLE 4.13.2A
Packings for Counterflow and Cross-Flow Cooling Towers: Designations and Descriptions

Counterflow Packings

1. Flat asbestos sheets, pitch 4.45 cm
2. Flat asbestos sheets, pitch 3.81 cm
3. Flat asbestos sheets, pitch 3.18 cm
4. Flat asbestos sheets, pitch 2.54 cm
5. 60° angle corrugated plastic, Munters M12060, pitch 1.17 in.
6. 60° angle corrugated plastic, Munters M19060, pitch 1.8 in.
7. Vertical corrugated plastic, American Tower Plastics Coolfilm, pitch 1.63 in.
8. Horizontal plastic screen, American Tower Plastics Cooldrop, pitch 8 in. 2 in. grid
9. Horizontal plastic grid, Ecodyne shape 10, pitch 12 in.
10. Angled corrugated plastic, Marley MC67, pitch 1.88 in.
11. Dimpled sheets, Toschi Asbestos-Free Cement, pitch 0.72 in.
12. Vertical plastic honeycomb, Brentwood Industries Accu-Pack, pitch 1.75 in.

Cross-Flow Packings

1. Doron V-bar, 4×8 in. spacing
2. Doron V-bar, 8×8 in. spacing
3. Ecodyne T-bar, 4×8 in. spacing
4. Ecodyne T-bar, 8×8 in. spacing
5. Wood lath, parallel to air flow, 4×4 in. spacing
6. Wood lath, perpendicular to air flow, 4×4 in. spacing
7. Marley α-bar, parallel to air flow, 16×4 in. spacing
8. Marley ladder, parallel to air flow, 8×2 in. spacing

in the range 1.5 to 4 m/s. No attempt is made to correlate g_m and a separately. The number of transfer units of a packing of height H is then

$$N_{tu} = \frac{g_m S}{\dot{m}_L} = \frac{g_m a H}{L} \tag{4.13.8}$$

The correlations are in terms of dimensionless mass velocities L^+ and G^+, and a *hot water correction* T^+_{HW}. The hot water correction accounts for a number of factors, such as errors associated with Merkel's method, deviations from low mass transfer rate theory at higher values of T_s, and fluid property dependence on temperature. Frictional resistance to air flow through the packings is correlated as a *loss coefficient* $N = \Delta P/(\rho V^2/2)$ per unit height or depth of packing, as a function of L^+ and G^+. The velocity V is superficial gas velocity. No hot water correction is required.

In a natural-draft tower, the thermal and hydraulic performance of the tower are coupled, and the flow rate ratio $\dot{m}_L/\dot{m}_G$ cannot be specified *a priori*. The buoyancy force producing the air flow depends on the state of the air leaving the packing which in turn depends on $\dot{m}_L/\dot{m}_G$ and the inlet air and water states. An iterative solution is required to find the operating point of the tower. The buoyancy force available to overcome the shell and packing pressure drops is

$$\Delta P^B = g(\rho_a - \rho_{G,out})H \tag{4.13.9}$$

where ρ_a is the ambient air density and H is usually taken as the distance from the bottom of the packing to the top of the shell. The various pressure drops are conveniently expressed as

TABLE 4.13.2B
Mass Transfer and Pressure Drop Correlations for Cooling Towers

Packing Number	C_1 (m^{-1})	n_1	n_2	n_3	C_2 (m^{-1})	n_4	n_5
			Counterflow Packings: $L_0 = G_0 = 3.391$ kg/m² s				
1	0.289	–0.70	0.70	0.00	2.72	0.35	–0.35
2	0.361	–0.72	0.72	0.00	3.13	0.42	–0.42
3	0.394	–0.76	0.76	0.00	3.38	0.36	–0.36
4	0.459	–0.73	0.73	0.00	3.87	0.52	–0.36
5	2.723	–0.61	0.50	–0.34	19.22	0.34	0.19
6	1.575	–0.50	0.58	–0.40	9.55	0.31	0.05
7	1.378	–0.49	0.56	–0.35	10.10	0.23	–0.04
8	0.558	–0.38	0.48	–0.54	4.33	0.85	–0.60
9	0.525	–0.26	0.58	–0.45	2.36	1.10	–0.64
10	1.312	–0.60	0.62	–0.60	8.33	0.27	–0.14
11	0.755	–0.51	0.93	–0.52	1.51	0.99	0.04
12	1.476	–0.56	0.60	–0.38	6.27	0.31	0.10
			Cross-flow Packings: $L_0 = 8.135$ kg/m² s, $G_0 = 2.715$ kg/m² s				
1	0.161	–0.58	0.52	–0.44	1.44	0.66	–0.73
2	0.171	–0.34	0.32	–0.43	1.97	0.72	–0.82
3	0.184	–0.51	0.28	–0.31	1.38	1.30	0.22
4	0.167	–0.48	0.20	–0.29	1.25	0.89	0.07
5	0.171	–0.58	0.28	–0.29	3.18	0.76	–0.80
6	0.217	–0.51	0.47	–0.34	4.49	0.71	–0.59
7	0.213	–0.41	0.50	–0.42	3.44	0.71	–0.85
8	0.233	–0.45	0.45	–0.48	4.89	0.59	0.16

Correlations (SI units)

Mass transfer: $\dfrac{g_m a}{L[\mathrm{kg/m^2\ s}]} = C_1 (L^+)^{n1} (G^+)^{n2} (T_{HW}^+)^{n3}$; Pressure drop: $\dfrac{N}{H \text{ or } X} = C_2 (L^+)^n + (G^+)^t$

where $L^+ = \dfrac{L}{L_0}$, $G^+ = \dfrac{G}{G_0}$, $T_{HW}^+ = \dfrac{1.8 T_{L,\mathrm{in}}[^\circ\mathrm{C}] + 32}{110}$

Source: Lowe, H.J. and Christie, D.G., Heat transfer and pressure drop data on cooling tower packings and model studies of the resistance of natural draft towers to in flow. *International Developments in Heat Transfer, Proceedings of the International Heat Transfer Conference*, Boulder, CO, ASME, New York, 1961; Johnson, B.M., Ed., *Cooling Tower Performance Prediction and Improvement*, Vols. 1 and 2, EPRIGS-6370, Electric Power Research Institute, Palo Alto, CA, 1990. With permission.

$$\Delta P_i = N_i \frac{\rho_{Gi} V_i^2}{2} \tag{4.13.10}$$

Where N_i is the loss coefficient and V_i is the air velocity at the corresponding location. The pressure drops are associated with the shell, the packing, the mist eliminators, supports and pipes, and the water spray below the packing. Some sample correlations are given in Table 4.13.3.

Water loadings in counterflow natural-draft towers typically range from 0.8 to 2.4 kg/m² s, and superficial air velocities range from 1 to 2 m/s. The ratio of base diameter to height may be 0.75 to 0.85, and the ratio of throat to base diameter 0.55 to 0.65. The height of the air inlet is usually 0.10 to

TABLE 4.13.3
Pressure Drop Correlations for Cooling Tower Shells, Sprays, Supports, and Mist Eliminators

1.	Shell (natural draft counterflow): $N = 0.167\left(\frac{D_B}{b}\right)^2$ where D_B is the diameter of the shell base and b is the height of the air inlet.
2.	Spray (natural-draft counterflow): $N = 0.526(Z_p[\text{m}]+1.22)(\dot{m}_L/\dot{m}_G)^{1.32}$
3.	Mist eliminators: N=2–4
4.	Support columns, pipes, etc. (natural-draft counterflow): N=2–6
5.	Fan exit losses for mechanical-draft towers (velocity based on fan exit area): N=1.0, forced draft ≃ 0.5, induced draft, depending on diffuser design
6.	Miscellaneous losses for mechanical-draft towers (velocity based on packing crosssectional area): $N \simeq 3$

Source: Lowe, H.J. and Christie, D.G., Heat transfer and pressure drop data on cooling tower packings and model studies of the resistance of natural draft towers to in flow. *International Developments in Heat Transfer Proceedings of the International Heat Transfer Conference*, Boulder, CO, ASME, New York, 1961; Singham, J.R., *Hemisphere Handbook of Heat Exchanger Design*, Section 3.12.3, Hewitt, G.E., Coord. Ed., Hemisphere, New York, 1990. With permission.

Note: N is the loss coefficient defined by Equation 4.9.10, with velocity based on cross-sectional area for air flow underneath the packing in items 1 through 4.

0.12 times the base diameter to facilitate air flow into the tower. In practice, the air flow distribution in natural-draft towers is not very uniform. However, the assumption of uniform air and water flows in our model of counterflow packing is adequate for most design purposes.

Cost-optimal design of cooling towers requires consideration of the complete power or refrigeration system. For refrigeration, the economics are determined mainly by the operating cost of the chiller (Kintner-Meyer and Emery, 1995).

Cooling Tower Behavior

There are a number of computer programs available that use variations of Merkel's method to calculate the cooling tower performance, for example, TEFRI (Bourillot, 1983), VERA2D-84 (Majumdar et al., 1985), CTOWER (Mills, 1995). These programs can be used to perform parametric studies to obtain the response of cooling towers to environmental, duty, and design changes. However, before using such programs, some thought should be given to the important characteristics of cooling tower behavior. For this purpose, it is useful to consider a graphical representation of Merkel's theory for a counterflow tower. Figure 4.13.4 shows a chart with moist air enthalpy plotted vs. water enthalpy (or, equivalently, water temperature) at 1 atm pressure. The *saturation curve* $h_s(T_s)$ is the enthalpy of saturated air. The *operating lines* $h_G(h_L)$ are given by Equation 4.13.2 and relate the air enthalpy to the water enthalpy at each location in the packing. The slope of an operating line is *L/G*. Since the assumption $T_s=T_L$ is made in Merkel's method, vertical lines on the chart connect h_s and h_G at each location in the packing. The driving force for enthalpy transfer, $(h_s - h_G)$, is the vertical distance between the saturation curve and the operating line. The integral in Equation (4.13.1) averages the reciprocal of this distance. By using this chart, a number of observations about cooling tower behavior can be made.

1. Figure 4.13.4 shows the effect of L/G for fixed water inlet and outlet temperatures, and fixed inlet air temperature and humidity. If we imagine L to be fixed as well, we see that as G decreases, the driving forces decrease, and so a larger N_{tu} is required.
2. The minimum N_{tu} required corresponds to $L/G=0$, that is, an infinite air flow rate, for which the operating line is horizontal.
3. Due to the curvature of the operating line, it is possible for the operating line to be tangent to the saturation curve. The indicated N_{tu} is then infinite, which tells us that the air flow rate must be increased in order to achieve the desired water cooling range.
4. For a mechanical-draft tower, the optimal value of L/G lies between the two limits described in items 2 and 3 above. If L/G is large, the required height of packing is large, and the capital cost will be excessive. If L/G is small, the fan power will be excessive (since fan power is proportional to air volume flow rate times pressure drop).

Range and Approach

Cooling tower designers and utility engineers have traditionally used two temperature differences to characterize cooling tower operation. The *range* is the difference between the water inlet and outlet temperatures (also called simply the hot and cold water temperatures). The *approach* is the difference between the outlet water temperature and the wet-bulb temperature of the entering (ambient) air. The approach characterizes cooling tower performance; for a given inlet condition, a larger packing will produce a smaller approach to the ambient wet-bulb temperature, and hence a lower water outlet temperature. The water cannot be cooled below the ambient wet-bulb temperature. The approach concept is useful because the ambient dry-bulb temperature has little effect on performance at usual operating conditions (for a specified wet-bulb temperature).

Cooling Demand Curves

Electrical utility engineers have found it convenient to use charts of *cooling demand curves* to evaluate packing specifications. Figure 4.13.5 is an example of such a chart, on which the required N_{tu}, for a given inlet air wet-bulb temperature and range, is plotted vs. L/G with the approach as a parameter. Such a plot is possible since the inlet air dry-bulb temperature has only a small effect under usual operating conditions. Now, if it is possible to correlate the mass transfer conductance as

$$\frac{g_m a}{L} = C\left(\frac{L}{G}\right)^{-n} \tag{4.13.11}$$

the N_{tu} of a packing of height H is

$$\frac{g_m S}{\dot{m}_L} = \frac{g_m a H}{L} = C\left(\frac{L}{G}\right)^{-n} H \tag{4.13.12}$$

Equation 4.13.12 can also be plotted on the chart to give the *packing capability line*. For a required approach, the *operating point* of the tower is the intersection of the cooling demand curve and packing capability line. Charts of cooling demand curves are available (Cooling Tower Institute, 1967; Kelly, 1976). Correlations of the form of Equation 4.13.11 do not necessarily fit experimental data well. A dependence $g_m a \propto L^{1-n} G^n$ is implied and, in the past, experimental data were often forced to fit such a relation. If the $g_m a$ correlation does not have the form of Equation 4.13.11, the N_{tu} cannot be plotted as a line on a cooling demand chart.

With the almost universal use of computers and the availability of suitable computer programs, one can expect less use of cooling demand charts in the future. The major sources of error in the predictions made by these programs are related to nonuniform air and water flow, and the

correlations of packing mass transfer and pressure drop experimental data. The experimental data are obtained in small-scale test rigs, in which it is impossible to simulate many features of full-size towers—for example, nonuniform flow due to entrance configuration, nonuniform wetting of the packing, and, in the case of counterflow towers, the effect of spray above the packing and rain below the packing. Furthermore, since testing of packings in small-scale test rigs is itself not easy, considerable scatter is seen in such test data. Correlations of the data typically have root mean square errors of 10%–20%.

Legionnaires' Disease

Legionnaires' disease is a form of pneumonia caused by a strain of legionnella bacteria (sero group I). Smokers and sick people are particularly vulnerable to the disease. Major outbreaks have occurred at conventions and in hospitals, for which the source of the bacteria has been traced to cooling towers of air-conditioning systems. The bacteria require nutrients such as algae or dead bacteria in sludge, and thrive if iron oxides are present. However, properly designed, installed, and maintained cooling towers have never been implicated in an outbreak of the disease. Key requirements to be met include the following:

1. Mist (drift) eliminators should be effective.
2. The tower should be located so as to minimize the possibility of mist entering a ventilation system.
3. Corrosion in the tower and water lines should be minimized by use of glass fiber, stainless steel, and coated steel.
4. The design should facilitate inspection and cleaning, to allow early detection and remedy of sludge buildup.
5. Water treatment and filtration procedures should meet recommended standards.

REFERENCES

Bourillot, C. 1983. *TEFRI:* Numerical model for calculating the performance of an evaporative cooling tower, EPRI CS-3212-SR, Electric Power Research Institute, Palo Alto, CA.

Cooling Tower Institute. 1967. *Cooling Tower Performance Curves*, Cooling Tower Institute, Houston.

Kelly, N.W. 1976. *Kelly's Handbook of Cross-Flow Cooling Tower Performance*, Neil W. Kelly and Associates, Kansas City, MO.

Kintner-Meyer, M. and Emery, A.F. 1995. Cost-optimal design of cooling towers, *ASHRAE J.*, April 1, 37, 46–55.

Majumdar, A.K., Singhal, A.K., and Spalding, D.B. 1985. VERA2D-84: A computer program for 2-d analysis of flow, heat and mass transfer in evaporative cooling towers, EPRI CS-4073, Electric Power Research Institute, Palo Alto, CA.

Merkel, F. 1925. Verdunstungskühlung, *Forschungsarb. Ing. Wes.*, no. 275.

Mills, A.F. 1995. *Heat and Mass Transfer*, Richard D. Irwin, Chicago.

Further Information

Baker, D. 1984. *Cooling Tower Performance*, Chemical Publishing Company, New York.

Johnson, B.M., Ed. 1990. *Cooling Tower Performance Prediction and Improvement*, Vols. 1 and 2, EPRI GS-6370, Electric Power Research Institute, Palo Alto, CA.

Singham, J.R. 1990. Natural draft towers, in *Hemisphere Handbook of Heat Exchanger Design*, Section 3.12.3, Hewitt, G.E., Coord Ed., Hemisphere Publishing, New York.

Stoeker, W.F. and Jones, J.W. 1982. *Refrigeration and Air Conditioning*, 2nd ed., McGraw-Hill, New York.

Webb, R.L. 1988. A critical review of cooling tower design methods, in *Heat Transfer Equipment Design*, Shah, R.K., Subba Rao, E.C., and Mashelkar, R.A., Eds, Hemisphere Publishing, Washington, DC.

4.14 PINCH TECHNOLOGY

Santanu Bandyopadhyay and Shankar Narasimhan

Introduction

The human appetite for energy has been growing rapidly, especially since the industrial revolution, and it is now used as one of the indicators of a country's progress and development. According to a UNIDO working paper in 2010 (Upadhyaya, 2010), more than 80% of the world's energy consumption is met by fossil fuels. Out of this, manufacturing accounts for a quarter of the total energy consumption. The increase in energy consumption has concomitantly led to a growing concern about the impact that the burning of fossil fuels has on the environment. While the world continues to seek ways and means of reducing our dependence on fossils fuels, by utilizing renewable energy sources, it is also important to use advanced technological methods to improve the energy efficiency of industrial processes. Pinch technology is a systematic approach that was developed in the mid-1980s to analyze the energy utilization in process industries, and to design integrated energy sub-system of a process which achieves the optimal trade-off between the cost of energy used and the capital investment made (Linnhoff et al., 1982; Klemeš, 2013).

In any design, there exists a trade-off between the capital investment and operating cost. For an energy sub-system, the capital cost is mainly dependent on the number and size of heat transfer equipment (boilers, furnaces, heat exchangers), while the operating cost is primarily a function of the consumption rate of utilities such as fuel, cooling water, refrigerants, steam, etc. Pinch technology effectively captures the trade-off between capital and operating costs, and can be used for the optimal design of the energy sub-system of a process. The method is now widely used to design the heat and power sub-system of processes such as refineries, petrochemical plants, pulp and paper plants, thermal power plants, etc. It is reported that energy savings of up to 30% have been achieved using this technology. Although, pinch technology started by considering heat integration in processes it has now been extended to minimize usage of other resources such as water and to minimize emissions, and has also been rechristened as *pinch analysis* (Linnhoff, 1993). In this chapter, the basic concepts used in pinch technology are introduced and explained. The use of these concepts for the optimum design of a heat exchanger network is also described in detail. The integration of the heat exchanger network with other energy consumers/producers such as distillation columns, pumps, compressors, and turbines is also discussed. There are several books, book chapters, and review articles that the reader can refer to get a more comprehensive understanding of pinch technology. Shenoy (1995) and Kemp (2007) have written books devoted entirely to pinch technology, while the book by Smith (2005) contains several chapters related to pinch technology. A bibliography compiled by Furman and Sahinidis (2002) contains references to more than 400 research papers on HEN synthesis based on pinch technology and mathematical programming methods. Early applications of pinch technology were mainly in refineries and petrochemical plants (Kemp, 2007), but applications to other industries such as pulp and paper (Svensson and Harvey, 2011) and steel (Grip et al., 2013) have been recently reported.

Motivating Example

Before describing the essential concepts of pinch technology and how it can be used to design a heat exchanger network (HEN), a simple example is provided to illustrate the significant role that a HEN structure has on the energy requirements. A simplified example abstracted from a thermal power plant is used to provide the motivation. In a thermal power plant, fuel is burnt in a boiler to generate high-pressure steam, which is subsequently used to generate power. The flue gas from the boiler contains sufficient energy, which can be recovered. The recovered energy can be used to preheat the air used for combustion, or to preheat the water fed to the boiler for generating steam. Figure 4.14.1

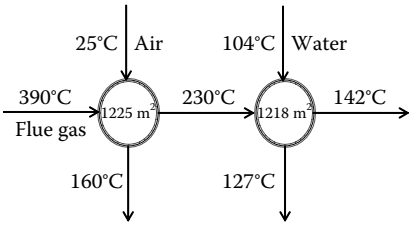


FIGURE 4.14.1 Existing HEN structure.

TABLE 4.14.1
Streams and Their Properties

Stream	Specific Heat Capacity (kJ/kg °C)	Flow Rate (kg/h)
Flue gas	1.1	39,032
Air	1.01	40,072.8
Water (10.5 bar)	4.41	31,824.0

TABLE 4.14.2
Exchangers and Their Properties for Original Configuration

Exchanger	Area (m^2)	Heat Lost by Hot Stream (kW)	Heat Gained by Cold Stream (kW)	Heat Lost to Ambient (kW)	U (W/m^2 K)
HE1	1225.0	1908.23	1517.76	390.47	5.7
HE2	1218.1	1049.53	897.25	152.28	11.3

shows an existing network of two heat exchangers in series used to recover energy from the flue gas. The flue gas is first used to preheat the air in the first exchanger (HE1) and subsequently to preheat pressurized water in the second exchanger (HE2). The flow rate and specific heat capacity of the different streams are given in Table 4.14.1, while the inlet and outlet temperatures of streams from each exchanger are shown in Figure 4.14.1. The heat exchanger areas, heat loads, and heat lost to ambient are given in Table 4.14.2.

From the heat loads of exchangers, it is observed that 2957.76 kW of energy is recovered from the flue gas out of which 542.75 kW is lost (about 20%) to the ambient. From Figure 4.14.1, it is observed that the final exit temperature of the flue gas is 142°C. The question is whether it is possible to extract more energy from the flue gas using the same exchangers. Figure 4.14.2 shows an alternative configuration where the flue gas is first used to preheat water in exchanger HE1 followed by preheating air in exchanger HE2. It is assumed that the overall heat transfer coefficient depends only on the streams exchanging energy. Furthermore, in order to make a fair comparison, it is assumed that heat losses in the exchangers are same as before. Table 4.14.3 shows the heat loads of exchangers for the new configuration. It is observed that the total energy recovered from flue gas is 3375.5 kW out of which the loss to the ambient is almost same as before (542.4 kW). The final exit temperature of the flue gas is 107°C. If this temperature is above the acid dew point, then technically it is possible to extract 10% more energy from the flue gas just by a simple change in the structure of the heat exchanger network (interchange the order in which flue gas exchanges energy with air and water). Although it is not shown here, it is also possible to recover 10% more energy using the same exchangers in a parallel configuration with appropriate apportioning of the flue gas flow rate to the exchangers. This example clearly demonstrates that the heat exchanger network configuration

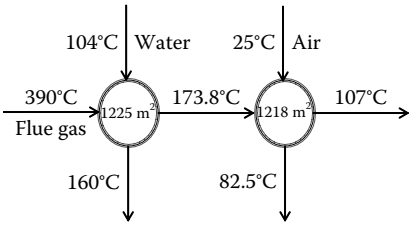


FIGURE 4.14.2 Modified HEN structure.

TABLE 4.14.3
Exchangers and Their Properties for Modified Configuration

Exchanger	Area (m^2)	Heat Lost by Hot Stream (kW)	Heat Gained by Cold Stream (kW)	Heat Lost to Ambient (kW)	U (W/m^2 K)
HE1	1225.0	2578.2	2186.3	391.9	11.3
HE2	1218.1	797.3	645.8	151.5	5.7

plays a significant role in the overall energy consumption of a process. In an actual thermal power plant, several heat exchangers (preheaters, superheaters, economizers, feed heaters, etc.) are used in a series-parallel arrangement. A more complex network of exchangers is used in a refinery to preheat the crude oil using the hot distillate streams of the atmospheric and vacuum columns. The optimal design of the heat exchanger network that can recover the maximum energy by interchanging energy between the process streams, while giving due consideration to the capital investment is a challenging problem. Pinch technology provides an insightful and systematic approach for solving this design problem. Although the description in the rest of the chapter pertains to the design of a HEN for a grassroots design, the concepts of pinch technology can also be used to analyze existing processes and retrofit these designs to recover more energy.

Pinch Technology

One of the important features of pinch technology from a design viewpoint is that it can be used to estimate achievable targets prior to the actual design of a HEN. These estimated targets include the following:

- Minimum external hot and cold utility required for solving a given problem without violating thermodynamic constraints
- Minimum number of heat exchangers required for heat integration
- Minimum overall area of the HEN
- Minimum overall capital cost of the HEN

While the minimum external utility requirements estimate can be shown to be theoretically rigorous, all other target estimates are approximate, but are reasonably good for design purposes. The establishment of such targets prior to design is useful for assessing how closely a proposed design meets these targets. It also indicates the direction in which a proposed design should be modified or evolved for meeting the targets.

Another feature of pinch technology is the extensive use of graphical diagrams such as *composite curves, temperature interval analysis and cascade diagram, source–sink diagram, driving force plots*, etc., which provide significant insights into the energy integration possibilities.

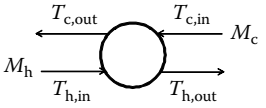


FIGURE 4.14.3 Counter-current heat exchanger.

The computation of the different targets for a given example problem, and the graphical diagrams used in these computations are first explained before the actual design of the HEN.

Interestingly, the mathematical background required for understanding pinch technology is minimal. Essentially, the following three equations that relate the different quantities associated with a heat exchanger are used in the design process. It is assumed that all the heat exchangers are counter-current exchangers, the schematic of one of which is shown in Figure 4.14.3.

$$Q = M_h C_{ph}\left(T_{h,in} - T_{h,out}\right) = M_c C_{pc}\left(T_{c,out} - T_{c,in}\right) \tag{4.14.1}$$

$$Q = UA\Delta T_{ln} \tag{4.14.2}$$

where

$$\Delta T_{ln} = LMTD = \frac{\left(T_{h,in} - T_{c,out}\right) - \left(T_{h,out} - T_{c,in}\right)}{\ln\left(\dfrac{\left(T_{h,in} - T_{c,out}\right)}{\left(T_{h,out} - T_{c,in}\right)}\right)}$$

$$\frac{1}{U} = \frac{1}{h_h} + \frac{1}{h_c} \tag{4.14.3}$$

The first equation relates the heat load of the exchanger (Q) to the heat lost by the hot stream and the heat gained by the cold stream, where for simplicity it is assumed that the specific heat capacities of the streams are constant and do not change with the temperature of the fluid. The second equation is the design equation, which relates the heat load to the overall heat transfer coefficient and area of exchanger, and the third equation relates the overall heat transfer equation to the heat transfer coefficients of the hot and cold streams that are exchanging heat. In general, the stream heat transfer coefficients depend on the geometry of the heat exchanger (such as diameter of the tubes, number of tubes, tube pitch and layout, shell diameter, baffle spacing, etc.). However, for the purpose of designing the HEN, it is assumed that the stream heat transfer coefficients for a given stream are constant and are computed for assumed typical values of the exchanger geometry.

Energy Targeting

Graphical Analysis

Pinch analysis was developed as a systematic physical-insight-based tool for energy conservation in heat exchanger networks (HENs) (Hohmann, 1971; Linnhoff and Flower, 1978; Umeda et al., 1978). The energy targets are obtained via developing temperature-heat duty profiles of heat availability and heat demands of all the process streams. To illustrate the graphical procedure, an illustrative example of two hot streams and two cold streams is considered. Supply and target temperatures of these streams, along with their heat capacity flow rates, are given in Table 4.14.4.

The heat availability profile of all the hot streams (i.e., streams that are to be cooled) is called the hot composite curve. Hot composite curve for this example can be generated by thermodynamically adding two hot streams, H1 and H2. Figure 4.14.4a shows the two hot streams on temperature-heat

TABLE 4.14.4
Stream Data for Illustrative Example

Stream	T_{in} (°C)	T_{out} (°C)	MC_p (MW/°C)	Duty (MW)
H1	210	60	0.10	15
H2	180	30	0.20	30
C3	40	100	0.15	9
C4	60	200	0.30	42

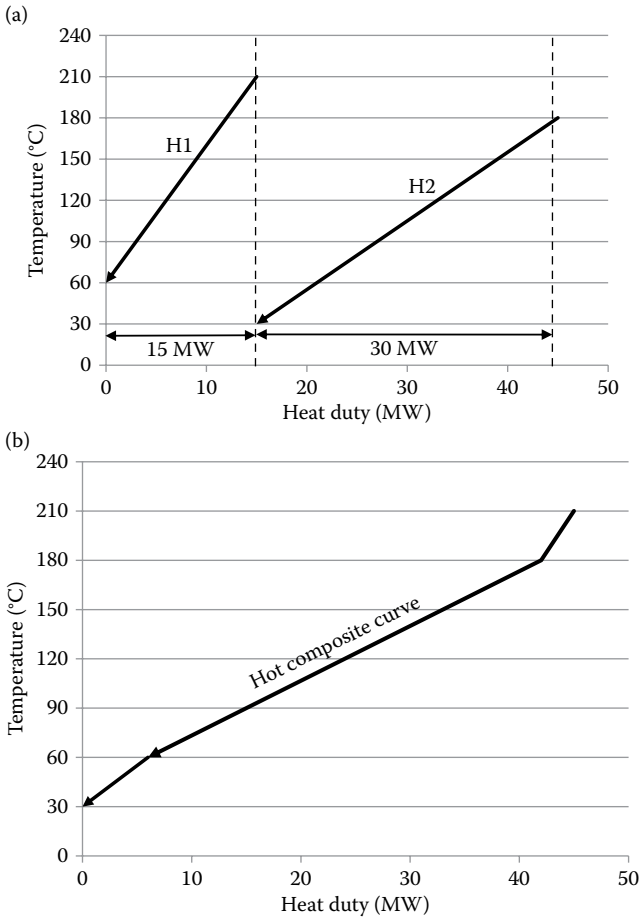


FIGURE 4.14.4 (a) Temperature-heat duty diagrams of hot streams, and (b) Hot composite curve.

duty diagram. The positioning of the streams in the temperature-heat duty diagram is arbitrary on the heat duty scale. However, the horizontal distance for each stream is fixed. This is because the starting value is not important; the overall heat duty of each stream is important. Figure 4.14.4b shows the two hot streams combined together to give a hot composite curve. The addition of heat availability from both the streams can be performed between each temperature interval. This is achieved by combining the duties and heat capacity flow rates of all the streams that have a common temperature range. This can also be done algebraically. Based on the supply and target temperatures of these streams, three temperature intervals can be identified: 30°–60°, 60°–180°, and 180°–210°. Between 30°C and 60°C, heat is available from H2 only. In this temperature interval, the total heat

available is $0.2 \times (60 - 30) = 6$ MW. Between 60°C and 180°C, heat is available from both H1 and H2. In this temperature interval, total heat available is $(0.2 + 0.1) \times (180 - 60) = 36$ MW. Similarly, 3 MW of heat, from the hot stream H1, is available in the temperature interval 180°–210°. For each temperature interval, heat available can be represented as a line segment on a temperature-heat duty diagram. Therefore, the coordinates of the hot composite curve in Figure 4.14.4b are (0, 30), (6, 60), (42, 180), and (45, 210).

Similarly, heat requirement profile of all the cold streams (i.e., streams that are to be heated) is called the cold composite curve. Based on the supply and target temperatures of two cold streams, C3 and C4, three temperature intervals can be identified: 40°–60°, 60°–100°, and 100°–200°. Between 40°C and 60°C, heat is required by C3 only. In this temperature interval, total heat required is $0.15 \times (60 - 40) = 3$ MW. Similarly, 18 and 30 MW of heat are required in the temperature intervals 60°–100° and 100°–200°. By joining these heat requirement line segments, the cold composite curve can be generated. Individual cold streams and the cold composite curve are shown in Figure 4.14.5. The coordinates of the cold composite curve are (0, 40), (3, 60), (21, 100), and (51, 200).

By construction, these composite curves, as a function of temperature, are monotonic in nature. These composite curves are plotted on a single diagram such a way that they are vertically separated

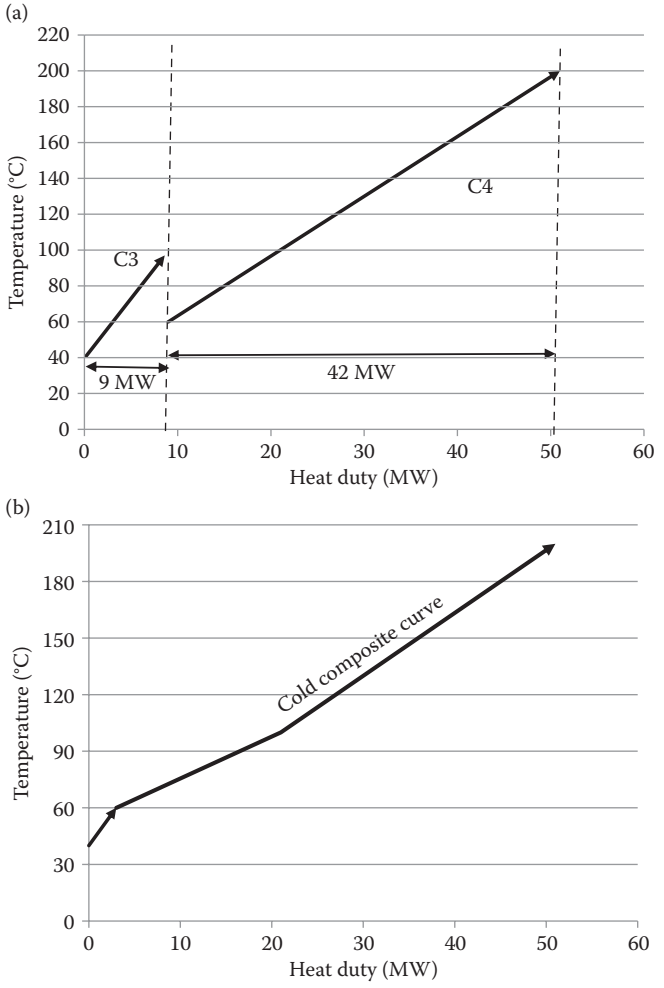


FIGURE 4.14.5 (a) Temperature-heat duty diagrams of cold streams, and (b) Cold composite curve.

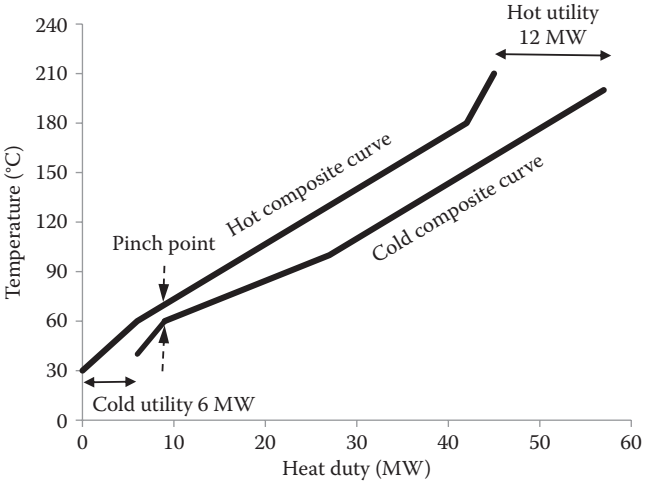


FIGURE 4.14.6 Hot and cold composite curves at ΔT_{min} of 10°C and energy targeting.

by at least a given minimum approach temperature (ΔT_{min}). This vertical temperature difference ensures satisfaction of the second law of thermodynamics and proper heat transfer from the hot streams to the cold streams (see Figure 4.14.6). For this illustrative example, ΔT_{min} is assumed to be 10°C. The overlap of these two composite curves indicates the maximum process heat recovery between hot and cold streams. Overhang of the hot composite curve over the cold composite curve denotes the requirement of the minimum cold utility (to remove excess heat from the system) and similarly, overshoot of the cold composite curve over the hot composite curve indicates the requirement of the minimum hot utility (to supply additional heat to the system). From Figure 4.14.6, the minimum hot and cold requirements are identified to be 12 and 6 MW. The point where ΔT_{min} is observed is known as the pinch point (see Figure 4.14.6). This is the point where the temperature driving force is the least or the composite curves are closest.

Algebraic Analysis

Problem Table Algorithm (PTA), originally proposed by Linnhoff and Flower (1978), is the algebraic tool of pinch analysis for targeting the minimum external utility requirements in a heat recovery network. Over the years, different modifications of the original PTA were proposed. For example, Ozkan and Dincer (2001) combined shifting of stream temperatures and the PTA in one tabular calculation and applied to target utility requirements in a refinery. Salama (2005) introduced numerical calculation procedure based on the reversal procedure of PTA. Based on the geometrical analysis, an algebraic technique was formulated by Salama (2006). An algebraic procedure to target the minimum utilities, known as the modified Problem Table Algorithm (MPTA), was proposed by Bandyopadhyay and Sahu (2010). In this chapter, MPTA is applied for energy targeting. Steps of the proposed algorithm are as follows.

Step 1: Shifting of stream temperatures.

Hot stream temperatures are shifted down by ½ ΔT_{min} and the cold stream temperatures are shifted up by ½ ΔT_{min}. However, it may be noted that any shifting may be permitted as long as the net shifting between the hot and cold stream temperatures are exactly ΔT_{min}. This ensures that there is an adequate driving force (ΔT_{min}) between the hot and cold streams.

Step 2: Determination of sources and demands from shifted stream data.

Inlet of each stream, irrespective of whether it is hot or cold, is considered as a source. Each source produces the heat capacity flow rate (MC_P) at its inlet temperature.

Similarly, the outlet of each stream is considered as a demand that accepts a heat capacity flow rate (MC_P) at respective outlet temperature. Source and demand data for the illustrative example are given in Table 4.14.5.

Step 3: Determination of temperature intervals (column 1)

Temperatures of all sources and demands, as determined in the previous step, are tabulated in decreasing order in the first column. If the value of a particular temperature occurs more than once, the same need not be repeated. The temperature for *k*th row is denoted as T_k. In the first column of Table 4.14.6, temperatures are arranged in descending order for the example.

Step 4: Calculation of heat capacity flows (column 2)

Heat capacity flows (i.e., algebraic sum of heat capacity flows corresponding to any particular temperature) are tabulated in the second column (see Table 4.14.6). Consider flows corresponding to sources as positive and demands as negative. For *k*th row, total heat capacity flow is denoted as F_k.

Step 5: Calculation of net MC_P (column 3)

Cumulative heat capacity flows are tabulated in the third column. Summation of heat capacity flows of all previous rows denotes the cumulative heat capacity flows for any particular row. Due to conservation of mass flow rate and constant specific heat for each

TABLE 4.14.5
Source and Demand Data for Illustrative Example with Shifted Temperatures

	Heat Capacity Flow (MW/°C)	Shifted Temperature (°C)
	Sources	
SH1	0.10	205
SH2	0.20	175
SC3	0.15	45
SC4	0.30	65
	Demands	
DH1	0.10	55
DH2	0.20	25
DC3	0.15	105
DC4	0.30	205

TABLE 4.14.6
Modified Problem Table Algorithm Applied to Illustrative Example

T_{int} (°C)	F_k (kW/°C)	$MC_{p,int}$ (kW/°C)	Q_{int} (kW)	Q_{cas} (kW)	R_{cas} (kW)
205	–0.2	–0.2	0	0	12
175	0.2	0	–6	–6	6
105	–0.15	–0.15	0	–6	6
65	0.3	0.15	–6	–12	0
55	–0.1	0.05	1.5	–10.5	1.5
45	0.15	0.2	0.5	–10	2
25	–0.2	0	4	–6	6

stream, last entry should be zero. Mathematically cumulative heat capacity flow at T_k is given as:

$$\text{CumF}_k = \sum_{l=1}^{k} F_l \tag{4.14.4}$$

Step 6: Calculation of net enthalpy (column 4)

The fourth column represents the enthalpy (Q_k) for each temperature interval. The first entry in the fourth column is kept 0. For all subsequent columns, the difference between the last two temperatures is multiplied by the cumulative heat capacity flows to calculate the enthalpy and tabulated in the fourth column. Mathematically, net enthalpy (Q_k) for each temperature interval can be calculated using the following formula:

$$Q_k = \begin{cases} 0, & \text{for } k = 1 \\ (T_{k-1} - T_k)\left(\sum_{l=1}^{k-1} F_l\right), & \text{for } k > 1 \end{cases} \tag{4.14.5}$$

Net enthalpy, calculated in this column, represents net surplus ($Q_i > 0$) or net deficit ($Q_i < 0$) of enthalpy in each temperature interval.

Step 7: Calculation of cascaded heat duty (column 5)

Cascaded heat duty for each temperature interval can be calculated by summing enthalpies for all previous rows and tabulated in the fifth column. Using Equation 4.14.5, cascaded heat duty for *k*th row may be expressed as:

$$Q_{\text{cas},k} = \left(\sum_{l=1}^{k} Q_l\right) = \begin{cases} 0, & \text{for } k = 1 \\ \sum_{l=1}^{k-1} F_l(T_l - T_k), & \text{for } k > 1 \end{cases} \tag{4.14.6}$$

Step 8: Revised cascade (column 6)

As some of the entries in column 5 may be negative, it suggests an infeasible cascade. The minimum value of Q_{cas} in column 5 is subtracted from each value in the fifth column to obtain the revised feasible cascaded heat duty and tabulated in the sixth column.

$$R_{\text{cas},k} = Q_{\text{cas},k} - \min(Q_{\text{cas}}) \tag{4.14.7}$$

In Table 4.14.6, the minimum value of Q_{cas} in column 5 is 12 MW. Using Equation 4.14.7, revised cascade data are generated and tabulated in column 6.

The first and last values of the sixth column suggest the minimum hot ($Q_{hu,min}$) and the minimum cold ($Q_{cu,min}$) utility requirements, respectively. Based on the data reported in column 6 of Table 4.14.6, the minimum hot and cold requirements are identified to be 12 and 6 MW, identical to the values obtained graphically (see Figure 4.14.6).

The temperature in column 1 that corresponds to zero revised cascaded heat in column 6 is the pinch temperature. For the illustrative example, pinch corresponds to the shifted temperature of 65°C. Therefore, the hot and cold side pinch temperatures are 70° and 60°C.

Golden Rules of Pinch Analysis

The pinch temperature boundary divides the process into two sub-processes, above and below the pinch. The process hot stream segments above pinch temperature should not transfer their energy to cold stream segments below pinch temperature, if a design has to meet minimum utility targets. This is nicely illustrated by the source–sink diagram shown in Figure 4.14.7a. It may be noted that the sub-process above the pinch is a net sink since it requires external hot utility to meet its energy requirements, while the sub-process below the pinch is a net source since it requires external cold utility to transfer the excess energy available. If an amount of energy α is transferred across pinch, then both the hot utility and cold utility requirements increase by the same amount α to satisfy the energy balances above and below pinch. Similarly, if cold utility is used above pinch or hot utility is used below pinch, then both hot and cold utility requirements increase by an identical amount. Figure 4.14.7b–d illustrates this graphically. Therefore, in order to achieve the minimum energy targets for a process, following rules must be obeyed. If these rules are not satisfied, requirement of hot and cold utilities goes up simultaneously, leading to a double energy penalty. These rules are known as golden rules of pinch analysis.

- No heat transfer across the pinch,
- No cold utility above the pinch, and
- No hot utility below the pinch.

Cost Targeting

Once the energy targets are established, the operating cost of the HEN can easily be established. In addition to the operating cost, it is useful to estimate the capital cost of the network. The capital cost of a HEN depends on the number of heat exchangers, number of shells, type of heat exchangers, heat exchange area, material of construction, pressure rating, etc. Capital cost targeting for HEN is briefly discussed.

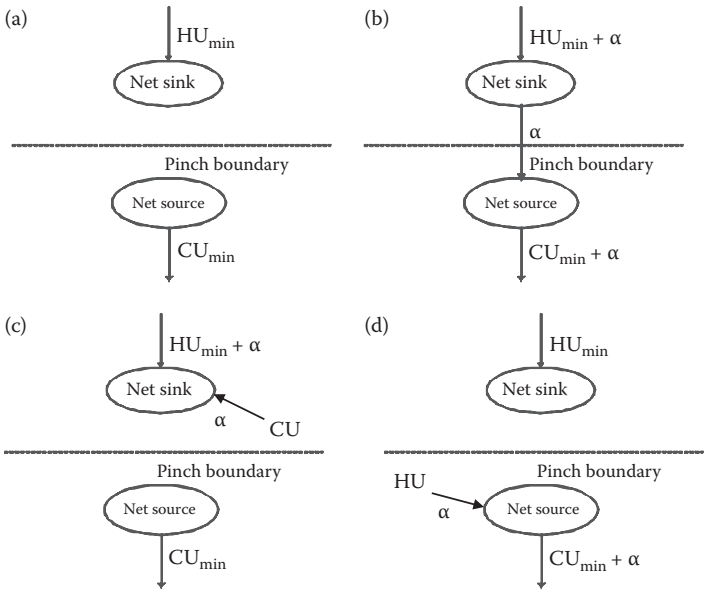


FIGURE 4.14.7 Derivation of golden rules of pinch analysis. (a) No heat transfer across pinch. (b) Heat transfer of α across pinch. (c) Cold utility of α used above pinch. (d) Hot utility of α used below pinch.

Targeting for Number of Heat Exchangers

A minimum number of heat exchangers can be estimated using the mathematical results from graph theory (Linnhoff et al., 1979). The number of heat exchanger in a HEN is related to the number of streams by the following equation:

$$\begin{pmatrix}\text{Number of heat}\\ \text{exchangers}\end{pmatrix}=\begin{pmatrix}\text{Number of streams}\\ \text{including utilities}\end{pmatrix}+\begin{pmatrix}\text{Number of independent}\\ \text{loops in the network}\end{pmatrix}-\begin{pmatrix}\text{Number of independent}\\ \text{components in the network}\end{pmatrix} \tag{4.14.8}$$

Before designing a HEN, a number of independent loops and independent components are not known. For simplicity, it is assumed that there is no loop in the network and there is a single component. These assumptions can simplify Equation 4.14.8 significantly.

$$\begin{pmatrix}\text{Number of heat}\\ \text{exchangers}\end{pmatrix}=\begin{pmatrix}\text{Number of streams}\\ \text{including utilities}\end{pmatrix}-1 \tag{4.14.9}$$

Equation 4.14.9 can be used as a targeting tool to estimate the number of heat exchangers, prior to the design of the HEN. Typically, Equation 4.14.9 is applied for the above pinch and below pinch portion separately to estimate the total number of heat exchangers. For the illustrative example, all four process streams exist above pinch and only hot utility can be used. There are total five streams and hence, four exchangers are expected. Similarly, there are three process streams (H1, H2, and C3), which exist below pinch with cold utility. This leads to a target of three heat exchangers below pinch. The total number of heat exchangers for this example is targeted to be 7.

Targeting for Heat Transfer Area

Heat transfer area may be estimated from the composite curves, after incorporating utility streams. Composite curves may be divided into vertical enthalpy intervals as shown in Figure 4.14.8. Assume a constant overall heat transfer coefficient U for the entire network and countercurrent heat transfer, the heat transfer area for each enthalpy interval may be calculated using the following formula:

$$\begin{pmatrix}\text{Heat transfer area in}\\ \text{an enthalpy interval}\end{pmatrix}=\frac{\text{heat duty of the enthalpy interval}}{U\times\text{LMTD}} \tag{4.14.10}$$

where LMTD stands for the log mean temperature difference of the interval. Different heat transfer coefficient for each stream can easily be incorporated and the following expression may be used for targeting the total heat transfer area for the HEN (Linnhoff and Ahmad, 1990):

$$A=\sum_{k}^{\text{Intervals}}\frac{1}{\text{LMTD}_{\text{k}}}\left[\sum_{i}^{\text{hot streams}}\frac{q_{\text{i,k}}}{h_{\text{i}}}+\sum_{j}^{\text{cold streams}}\frac{q_{\text{j,k}}}{h_{\text{j}}}\right] \tag{4.14.11}$$

It should be noted that Equation 4.14.11 does not predict the true minimum network area for different heat transfer coefficient for each stream. Another difficulty is to estimate heat transfer coefficient for each stream without a detailed design of the heat exchanger.

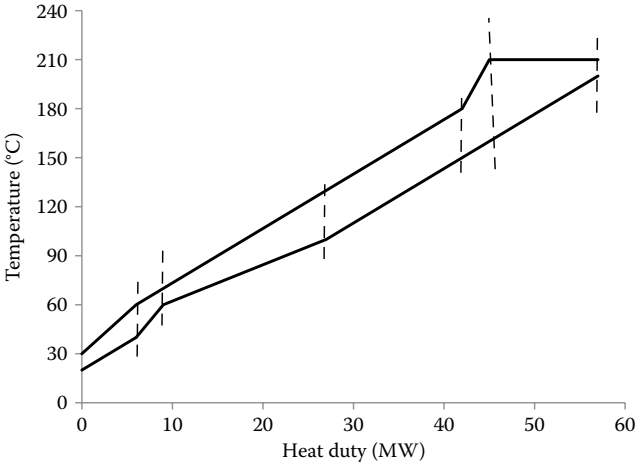


FIGURE 4.14.8 Enthalpy intervals and area targeting.

Cost Optimization

Typically, the cost of a single heat exchanger with surface area A can be estimated using some simplifying relation of the following form:

$$\left(\text{Cost of a Heat Exchanger}\right) = a + bA^c \tag{4.14.12}$$

where a, b, and c are various cost constants that vary according to the materials of construction, pressure rating, and type of exchanger. Using this, a simplified expression may be derived to estimate the capital cost of the entire network (Ahmad et al., 1990).

$$\left(\text{Capital Cost of HEN}\right) = N(a + b(A/N)^c) \tag{4.14.13}$$

where N is the number of heat exchangers.

Once the capital cost is targeted, total annualized cost of the entire network may also be estimated by adding annual operating cost with the annualized capital cost. By varying the total annualized cost of HEN with respect to the $\Delta T_{\min}$, the optimum value of $\Delta T_{\min}$ may be determined (Linnhoff and Ahmad, 1989). A typical capital-energy trade-off, as a function of $\Delta T_{\min}$, is illustrated in Figure 4.14.9. As $\Delta T_{\min}$ increases, the consumption of both hot and cold utilities increases together in order to maintain the energy balances. This leads to increase in operating cost. Typically, energy cost or the operating cost is a piece-wise linear function of $\Delta T_{\min}$. On the other hand, the temperature differences for heat transfer between the hot and cold streams throughout the process become larger as $\Delta T_{\min}$ increases. This means that the heat transfer area decreases, hence decreasing the capital costs. The jumps in the capital cost curve are created due to variation in the expected number of heat exchangers. Combining the energy cost and capital cost allows an optimal value of $\Delta T_{\min}$ to be identified. In practice, the shape of the optimization curve tends to be quite flat and a very precise optimization is often not necessary. As long as the value of $\Delta T_{\min}$ is somewhere in the region of the optimum point, there is little to be lost or gained in terms of the overall cost by small changes in $\Delta T_{\min}$. For chemical processes, a reasonable value of $\Delta T_{\min}$ is usually around 10°C–15°C. A larger value is usually required for refinery processes (typically 15°–30°C). For low-temperature processes, a reasonable value of $\Delta T_{\min}$ is often lower (typically 5°C or less), because of the expense of providing cooling through refrigeration in low-temperature processes.

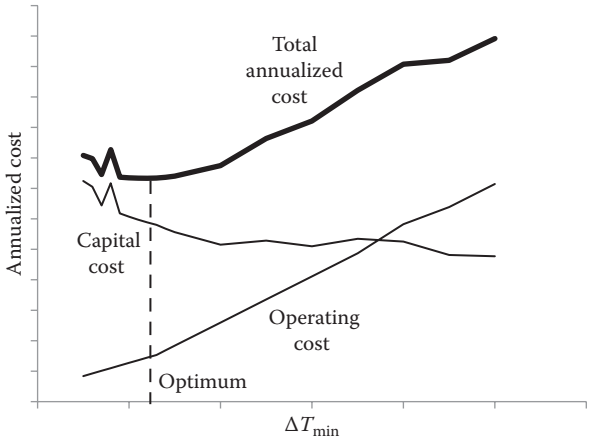


FIGURE 4.14.9 Capital-energy trade-offs and determination of optimum ΔT_{min}.

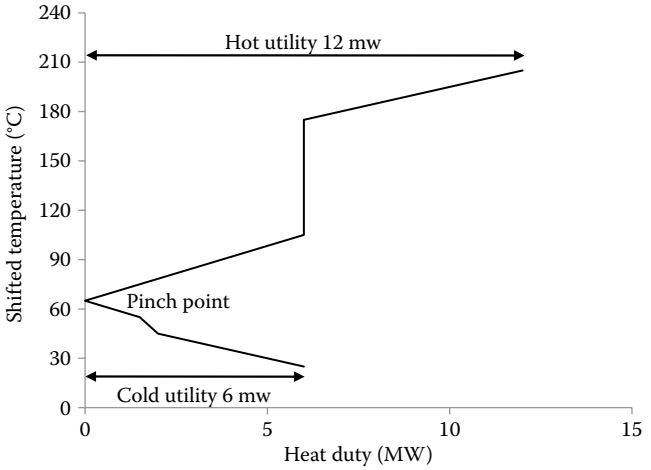


FIGURE 4.14.10 Grand composite curve (GCC).

Targeting for Multiple Utilities

In the previous section, energy targeting is performed with a single hot utility and a single cold utility. However, in many practical situations, multiple utilities are present. To target multiple utilities a piece-wise linear curve, known as the Grand Composite Curve (GCC), may be utilized. GCC may be generated by drawing the shifted temperature (in column 1 of MPTA) against the feasible cascade (in column 6 of MPTA) in a temperature-heat duty diagram. GCC for the example is shown in Figure 4.14.10. The GCC represents the heat demand (segments with positive slope) and supply (segments with negative slope) within each temperature interval. GCC touches the temperature axis at the shifted pinch temperature, signifying no heat transfer across the pinch. The portion of the GCC above the pinch point represents heat demand and requires only hot utilities. Similarly, the portion below the pinch has surplus heat and this surplus heat is rejected through cold utilities.

Figure 4.14.11 illustrates some examples in the use of the grand composite curve. Figure 4.14.11a shows high-pressure steam and cooling water used to satisfy the heating and cooling requirements of a GCC. In Figure 4.14.11b, high-pressure and low-pressure steams are used to satisfy the hot utility requirement of a GCC. Flue gas from a fired heater or from gas turbine exhaust may also be used

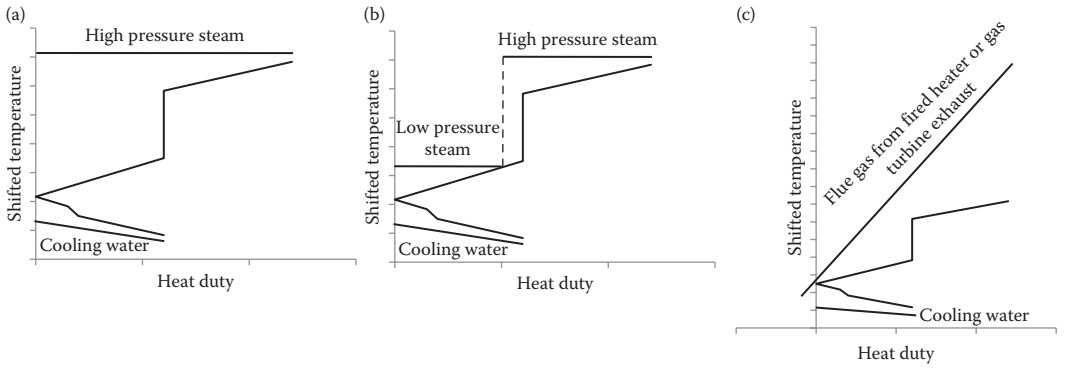


FIGURE 4.14.11 Examples in the use of the GCC: (a) use of high-pressure steam and cooling water, (b) use of multiple steam levels, and (c) use of flue gas as hot utility.

to provide hot utility requirement in a process (see Figure 4.14.11c). Essentially utility profiles are matched against the GCC to target multiple utilities. During profile matching, utility profiles may touch the GCC. This does not imply that ΔT_{min} is zero. It should be noted that the temperature axis in GCC is in shifted scale and the same temperature shift (by $\Delta T_{min}/2$) should also be incorporated in generating the utility profile. Therefore whenever the utility profile touches the GCC, the temperature driving force is exactly equal to ΔT_{min}.

Shenoy et al. (1998) proposed a targeting methodology to determine the optimum distributions of loads for multiple utilities considering the cost tradeoffs. GCC plays an important role in the thermal integration of fired heaters (Varghese and Bandyopadhyay, 2007), integration between various processes (Bandyopadhyay et al., 2010), cogeneration plant (Bade and Bandyopadhyay, 2015), etc.

Network Design and Evolution

In the previous sections, methodologies for setting various targets are discussed. It should be noted that these targets are set prior to the detailed design of the HEN. In this section, a simple methodology, called the pinch design method, to design HEN is discussed (Linnhoff and Hindmarsh, 1983). The starting point is the energy targets and the pinch temperatures. The design procedure starts by setting up a design grid, as shown in Figure 4.14.12. The hot streams are shown at the top of the diagram running in the direction left to right and the cold streams at the bottom of the diagram running right to left. A vertical line at the pinch temperature of 70°C for the hot streams and 60°C for the cold streams divides the grid into two parts (see Figure 4.14.12). To the left of the pinch line is the above pinch portion. To the right of the pinch line is the below pinch portion. It should be noted that

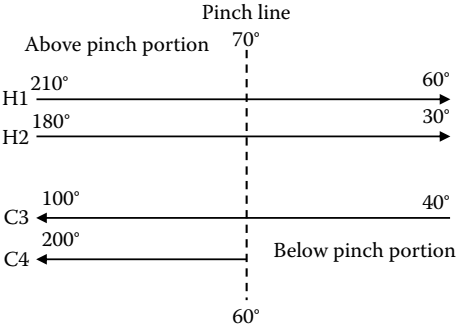


FIGURE 4.14.12 Parallel grid diagram for heat exchanger network design.

all four streams exist above pinch, but only three streams exist in the below pinch portion. Pinch is the most constrained part of the composite curve. Temperature driving force is the minimum at the pinch (see Figure 4.14.6). The design procedure starts at the pinch region and proceeds further.

For the successful design of a HEN, two feasibility criteria have to be met. First feasibility criterion is related to the heat capacity flow rates (MC_p) of the hot and the cold streams. When a hot stream and a cold stream, just above the pinch point, are selected to transfer heat through a heat exchanger, the temperature driving force at the pinch side is exactly ΔT_{min}. The temperature driving force at the other side of the exchanger must be above ΔT_{min}. This is only possible if the MC_p of hot stream is lower than that of the cold stream. The exactly opposite relation has to be obeyed just below the pinch. The MC_p feasibility criterion may be summarized as:

$$\left(MC_P\right)_{\text{hot stream}} \leq \left(MC_P\right)_{\text{cold stream}} \quad \text{just above pinch} \tag{4.14.14}$$

$$\left(MC_P\right)_{\text{hot stream}} \geq \left(MC_P\right)_{\text{cold stream}} \quad \text{just below pinch} \tag{4.14.15}$$

Second feasibility criterion is related to the number of streams present at the pinch. To achieve the energy target, no cold utility to be placed above the pinch. This implies that every hot stream, present just above the pinch, must be brought to the pinch temperature by exchanging heat with a cold stream. Therefore, the number of cold streams, just above the pinch, must be greater than that of the hot streams. The similar argument just below the pinch suggests that the number of hot streams must be greater than that of the cold streams. The number feasibility criterion may be summarized as:

$$\text{Number of hot streams} \leq \text{Number of cold streams} \quad \text{just above pinch} \tag{4.14.16}$$

$$\text{Number of hot streams} \geq \text{Number of cold streams} \quad \text{just below pinch} \tag{4.14.17}$$

Whenever the number criterion is not met, a stream may be split into multiple fractions. It should be remembered that these feasibility criteria are applicable only at the pinch and not away from the pinch. As the temperature driving force is higher, these criteria are not directly applicable.

The next step in the design is to determine the duty of each exchanger. By choosing a large duty, the number of heat exchangers can be minimized. The maximum duty a heat exchanger can have is the lower of the heat duties of the exchanging streams. This is known as tick-off heuristic. These steps can be used to design a HEN. Applications of these steps are demonstrated with the illustrative example.

We start with the below pinch portion (see Figure 4.14.13a). There are three streams below pinch. Both the hot streams start at 70°C, the hot side pinch temperature. The cold stream terminates at 60°C, the cold side pinch temperature. MC_p criterion below pinch (Equation 4.14.15) suggests that C3 with a MC_p of 0.15 can be matched with a hot stream with higher MCp. Only H2 with a MC_p of 0.2 matches that. This also satisfies the number criterion (Equation 4.14.17). Therefore, a match between H2 and C4 is chosen (see Figure 4.14.13b). The heat exchanger in the grid diagram is represented by a vertical line joining two open circles on the streams that are being matched. H2 has 8 MW of heat available and C3 can accept 3 MW of heat. Therefore, the duty of this exchanger is 3 MW. Now, no cold stream is available. The remaining heat of the hot streams must be removed by cold utilities. Two coolers, one on H1 and other on H2, are placed as shown in Figure 4.14.13c. This completes the design of below pinch HEN. It matches with the cold utility target of 6 MW and expected number of heat exchangers of 3.

Now, we start with the above pinch portion (see Figure 4.14.14a). All four streams are present above pinch. Both the hot streams terminate at 70°C, the hot side pinch temperature. Both the cold streams start at 60°C, the cold side pinch temperature. MC_p criterion above pinch (Equation 4.14.14)

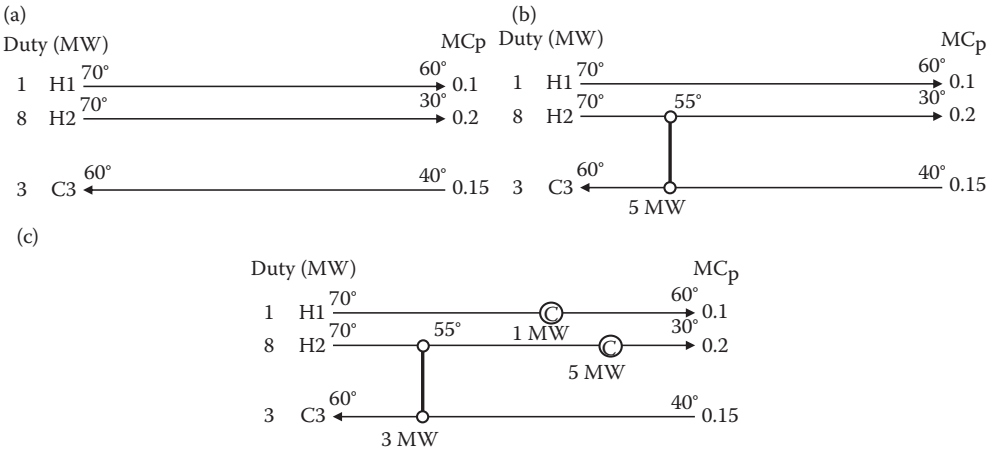


FIGURE 4.14.13 Design of below pinch HEN: (a) below pinch parallel grid diagram, (b) the first match, and (c) placement of coolers.

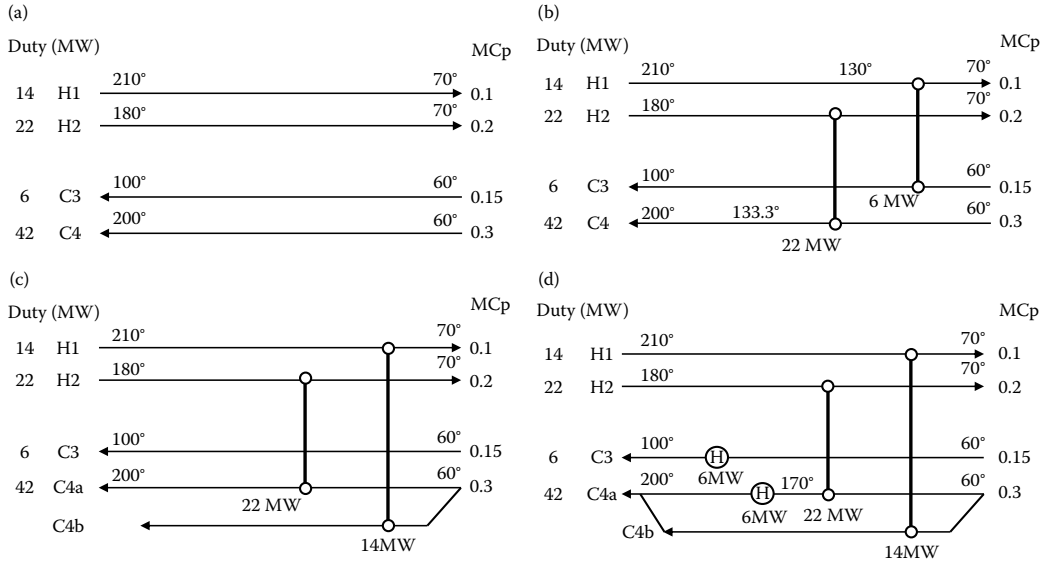


FIGURE 4.14.14 Design of above pinch HEN: (a) above pinch parallel grid diagram, (b) the hidden pinch phenomenon, (c) stream splitting, and (d) placement of heaters.

suggests that C3 with a MC_p of 0.15 can be matched with H1 with a MC_p of 0.1 and C4 with a MC_p of 0.3 can be matched with H2 with a MC_p of 0.2. Number criterion is automatically satisfied as two hot streams are matched with two cold streams. Duties of the H1-C3 and H2-C4 matches are 6 MW (duty of C3) and 22 MW (duty of H2). These matches are shown in Figure 4.14.14b. Now H1 and C4, with a residual duty of 20 MW, are left unsatisfied. However, have a look at the temperatures of the remaining streams. 8 MW of heat is available (remaining H1 stream) from 210° to 130°C and 20 MW of heat is required (remaining C4 stream) from 130.3° to 200°C. This is now impossible to match these two streams without any penalty. This is known as the hidden pinch phenomenon. This is primarily due to the failure of the tick-off heuristic.

One simple way to avoid hidden pinch phenomenon for this example is to avoid using C3 stream. In this case, both H1 and H2 have to match with C4. This leads to two hot streams and a single cold stream, which leads to a violation of the number criterion above pinch (Equation 4.14.16). Therefore,

C4 is split into two (C4a and C4b) and each portion is matched with different hot streams, as shown in Figure 4.14.14c. Duties of the H1-C4b and H2-C4a matches are decided to be 14 MW (duty of H1) and 22 MW (duty of H2). As the MC_p of C4 is exactly sum of the MC_ps of the hot streams, MC_p of C4a and C4b must be 0.2 and 0.1, respectively. This also satisfies the MC_p criterion. Now, no hot stream is available. Remaining heat requirement of the cold streams must be satisfied by hot utilities. Two heaters, one on C3 and other on C4a are placed as shown in Figure 4.14.14d. Two split fractions of C4 are joined and this completes the design of above pinch HEN. It matches with the hot utility target of 12 MW and expected number of heat exchangers of 4. Complete HEN is shown in Figure 4.14.15.

The pinch design method creates a network structure that satisfies the energy and unit targets. The minimum temperature driving force for the network is ΔT_{min}. Based on this initial structure, the network may be evolved and optimized further. Two important concepts are used while evolving a given network: loop and path. A loop is defined by a series of connected matches such that it starts and ends at the same node. One loop is shown in Figure 4.14.16. This loop has the following connections: H(C3), H2-C3, H2-C4, and H(C4). It should be noted that the heaters are connected due to the presence of a single hot utility. It is interesting to note that every loop in a HEN must contain an even number of matches. A path is defined by a series of connected matches such that it starts from a heater and ends at a cooler. One path is shown in Figure 4.14.17. This path has the following connections: H(C3), H2-C3, and C(H2). It is interesting to note that path in a HEN must contain an odd number of matches.

Loops and paths can be used to evolve any network. One classical technique is to identify a loop and distribute the duties of various heat exchanges in the loop. The distribution of the duties may

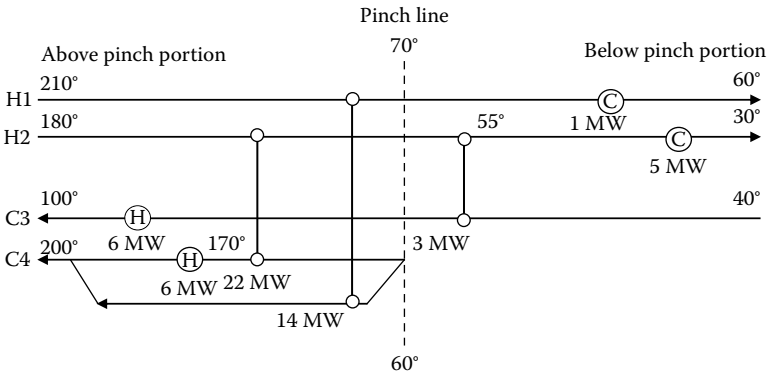


FIGURE 4.14.15 Complete heat exchanger network.

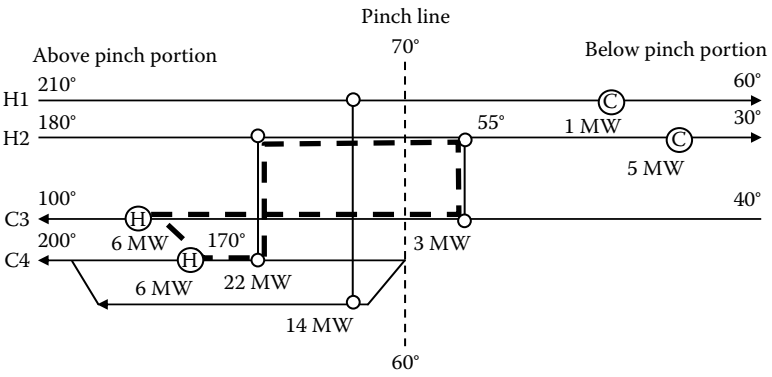


FIGURE 4.14.16 Loop in a heat exchanger network.

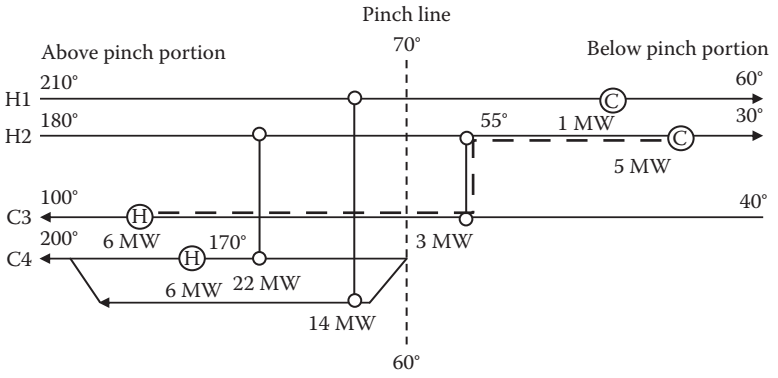


FIGURE 4.14.17 Path in a heat exchanger network.

be performed in such a way that the duty of one exchanger may be reduced to zero while the duties of the other exchangers are positive. This is known as loop breaking. Typically, the heat exchanger with the minimum duty is removed by loop breaking (Linnhoff and Hindmarsh, 1983). By breaking a loop one heat exchanger may be reduced. However, in most cases, this leads to a violation of the minimum approach temperature for some exchangers (certain cases it may become negative). A path is then identified through a heat exchanger where the minimum approach temperature is violated. Heater and cooler duties are increased and duties of other exchangers are adjusted to restore the minimum approach temperature of the network. Distribution of duties of different heat exchangers in a path is known as path relaxation. A combination of loop breaking and path relaxation is performed to evolve a HEN. In general, the breaking of a loop and recovery of temperature violations using a path leads to energy transfer across the pinch and consequent increase in energy requirements. The extra energy penalty incurred can be compared with the decrease in capital cost due to elimination of an exchanger and if the trade-off is favorable, then the evolved network can be accepted. The energy penalty incurred depends on the choice of loop and path used for eliminating an exchanger and there are no guidelines for an appropriate choice of loop and paths for evolutionary synthesis. Alternatively, a rigorous optimization-based formulation proposed by Mehta et al. (2001) using all loops and paths of the process can be formulated which can be used to determine the least energy penalty for eliminating an exchanger. This formulation can also be used to determine whether it is feasible to eliminate an exchanger from a given HEN.

Integrating HEN with Other Energy Devices

In a process, electrical energy is used to drive pumps and compressors. The process may also have a captive power plant for generating electrical power from high-pressure steam. Pinch technology provides clear rules for heat and power integration which can increase the overall reduction in energy utilization.

Figure 4.14.18a shows a process represented as a source–sink diagram with HU_{min} and CU_{min} being the minimum utility requirements. Figure 4.14.18a also shows a heat engine, which is part of the process (such as a turbine) which converts part of the thermal energy it receives (Q_{in}) into useful work W and rejects the remaining energy $Q_{out} = Q_{in} - W$ to the environment. The heat engine operates between the temperatures T_{source} and T_{sink}. The process and heat engine are not heat integrated, and the combined hot utility required by the process and heat engine is $HU_{min} + Q_{in}$, while the combined cold utility required is $HU_{min} + Q_{out}$. The net useful work obtained is W (in the form of electrical energy) which may be used in the process itself or may be exported. Figure 4.14.18b–d conceptually depicts different ways of thermally coupling the heat engine with the process depending on the temperatures, T_{source}, T_{sink} and T_{pinch}. It is observed from these diagrams that if the heat engine is heat integrated across pinch, then there is no net change in the total utility consumption or useful work produced when compared

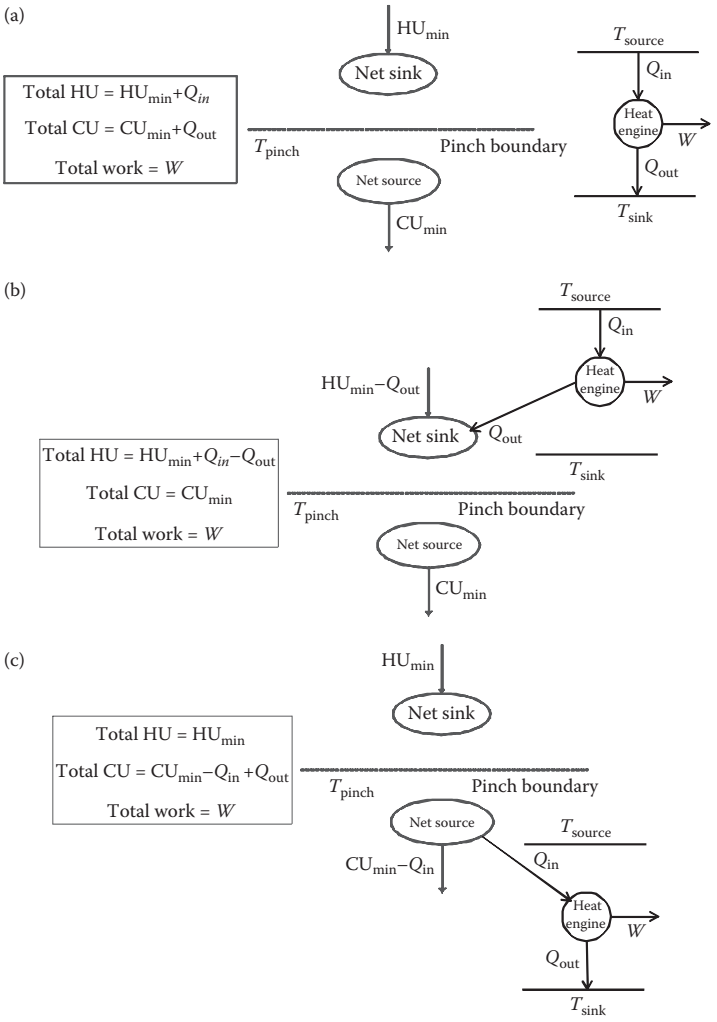


FIGURE 4.14.18 Integration of heat engine with process. (a) Process and heat engine not thermally integrated. (b) Heat engine integrated above pinch. (c) Heat engine integrated below pinch.

to the non-integrated scheme of Figure 4.14.18a. However, if the heat engine is integrated by rejecting heat to the process above pinch or by receiving energy from the process below pinch, then there is net reduction in overall energy consumption of either Q_{out} or Q_{in}, respectively. It may be noted that the diagrams represent the maximum energy reduction that can be obtained by integrating the heat engine with the process. The actual reduction depends on the temperature levels at which the integration is performed, and the available or required process energy at that temperature level.

It may be noted that a distillation column is similar to a heat engine since it takes in heat from a reboiler and rejects heat to the condenser, while performing useful work of separating a multicomponent mixture into products of the desired composition. As an example, consider the process described in Table 4.14.4. A binary distillation column operating at atmospheric pressure is also part of the process. The bottom product has a bubble point temperature of 120°C and the top product has a dew point temperature of 90°C. The reboiler duty is 2.2 MW and condenser duty is 2 MW. One possibility to integrate the column reboiler and condenser with the process HEN is to include the top product stream as a hot stream and bottom product as a cold stream in the stream table. The HEN designed for this extended process will also include the reboiler and condenser of the

distillation column. Alternatively, if the product streams of the column are not included in the process stream table, then we can consider the integration of this column with the background process HEN by treating it like a heat engine. We will assume that a minimum approach temperature of 10°C is required in the reboiler and condenser. Since the condenser temperature is greater than the pinch temperature, the only possibility is to reject heat to the process above pinch by integrating the condenser with the process. In order to satisfy the minimum approach temperature, the cold process stream receiving the condenser duty should be <80°C. From the grand composite curve (Figure 4.14.6), we observe that the net energy deficit from pinch temperature up to 80°C is 1.5 MW. Thus, the maximum amount of energy that can be rejected from the condenser to the background process is 1.5 MW and the remaining energy of 0.5 MW must be met by an external cold utility. The net energy savings by heat integrating the column with the process is 1.5 MW. It is also possible to obtain additional savings by considering the option of increasing the column pressure which will increase both the reboiler and condenser temperatures. The trade-off, in this case, is the increase in capital cost of the column (to withstand the increased pressure) and increased pumping costs as against the reduction in the cost of external utilities.

A similar analysis can be carried out for integrating a heat pump with the HEN. A heat pump takes in energy from a low-temperature source and does useful work on it and rejects it to a high-temperature sink. If the source temperature is below the pinch temperature, then the heat pump can receive energy from the process below pinch. Similarly, if the sink temperature is above pinch, then the heat pump can reject heat to the process above pinch. Thus, it is best if the heat pump is heat integrated across the pinch.

This leads to the following pinch technology rules for heat and power integration:

- Integrate a heat engine above or below pinch, but not across the pinch temperature.
- Integrate a heat pump across pinch (receive heat from below pinch and reject heat above pinch).

Conclusion

Pinch analysis was originally developed as a thermodynamic-based approach for conserving thermal energy through a network of heat exchangers (Linnhoff et al., 1982). It was developed as a holistic tool to analyze overall heat recovery systems for energy conservation in process industries. Problems of pinch analysis can easily be generalized where streams are characterized by both quantity and quality measures. In the case of heat recovery, energy streams are quantified in terms of heat duty (quantity) and temperature (quality). The fundamental problem involves minimization of the externally sourced, high-value resource (e.g., hot utilities in heat recovery pinch analysis) subject to quality constraints (e.g., temperature constraints due to the second law of thermodynamics). It is important to note that a consequence of such an optimization is the simultaneous reduction of unusable waste streams (e.g., rejected heat in heat recovery pinch analysis) (Bandyopadhyay, 2006).

Based on the similarities between heat and mass transfer phenomena, El-Halwagi and Manousiouthakis (1989) proposed the synthesis of mass exchange networks for the efficient use of industrial mass separating agents such as solvents, adsorbents, etc. The primary objective is to recover impurity load from a set of rich streams (process streams containing impurity) to a set of lean streams (process mass separating agent(s) that can remove impurity) and thus reducing the use of external mass separating agent. Advanced problems in a mass exchange network were addressed by Hallale and Fraser (1998).

Further work on mass integration in the 1990s led to the development of pinch analysis techniques for material resource conservation networks including that for industrial water network (Wang and Smith, 1994), refinery hydrogen system (Alves, 1999), and property integration (Kazantzi and El-Halwagi, 2005). In these problems, valuable resources (e.g., water, hydrogen, and other materials) from a set of process sources (stream containing resources) to a set of process sinks (units requiring resources) are recovered and thus, the use of external resources (such as fresh water, fresh hydrogen, etc.) is reduced (El-Halwagi, 2011; Foo, 2012).

Tan and Foo (2007) proposed pinch analysis–based methodology for carbon-constrained energy planning. In carbon-constrained energy planning, energy sources (fossil fuels) are allocated to various energy demands (e.g., regions/areas that require energy), in order to reduce the overall CO_2 emission. Carbon capture and storage planning is one of the important issues for carbon-constrained energy sector planning and applied for country-specific energy sector planning like Ireland (Crilly and Zhelev, 2008), New Zealand (Atkins et al., 2010), China (Chen et al., 2011), India (Krishna Priya and Bandyopadhyay, 2013), etc. Techniques of pinch analysis have been extended to study problems such as to determine the extent of carbon capture to existing power plant that minimizes power loss (Tan et al., 2009; Sahu et al., 2014), to determine the maximum amount of CO_2 to be stored in various storage systems with given capacity and injectivity constraints (Diamante et al., 2014), etc.

There are numerous other extensions of pinch analysis. Other than the examples listed above, pinch analysis has also been applied to aggregate production planning problems (Singhvi and Shenoy, 2002), isolated energy systems (Arun et al., 2007; Bandyopadhyay, 2011), human resource planning, and work scheduling (Foo et al., 2010). Over the years, pinch analysis has established itself as a structural tool for analyzing and conserving resources in numerous diversified applications, such as energy sector planning, financial analysis, supply chain management, isolated energy system design, batch process scheduling, carbon dioxide sequestration, etc. (Linnhoff, 1993; Smith, 2016).

REFERENCES

Ahmad, S., B. Linnhoff, and R. Smith (1990). Cost optimum heat exchanger networks—2. Targets and design for detailed capital cost models. *Computers & Chemical Engineering*, 14(7): 751–767.

Alves, J. (1999). Analysis and design of refinery hydrogen systems. Ph.D. Thesis, University of Manchester Institute of Science and Technology, Manchester, UK.

Arun, P., R. Banerjee, and S. Bandyopadhyay (2007). Sizing curve for design of isolated power systems. *Energy for Sustainable Development,* 11: 21–28.

Atkins, M.J., A.S., Morrison, and M.R. Walmsley (2010). Carbon emissions pinch analysis (CEPA) for emissions reduction in the New Zealand electricity sector. *Applied Energy,* 87: 982–987.

Bade, M.H. and S. Bandyopadhyay (2015). Analysis of gas turbine integrated cogeneration plant: Process integration approach. *Applied Thermal Engineering,* 78: 118–128.

Bandyopadhyay, S. (2006). Source composite curve for waste reduction. *Chemical Engineering Journal,* 125: 99–110.

Bandyopadhyay, S. (2011). Design and optimization of isolated energy systems through pinch analysis. *Asia-Pacific Journal of Chemical Engineering,* 6: 518–526.

Bandyopadhyay, S. and G.C. Sahu (2010). Modified problem table. Algorithm for energy targeting. *Industrial & Engineering Chemistry Research,* 49: 11557–11563.

Bandyopadhyay, S., J. Varghese, and V. Bansal (2010). Targeting for cogeneration potential through total site integration. *Applied Thermal Engineering*, 30(1): 6–14.

Chen, Q., C. Kang, Q. Xia, and D. Guan (2011). Preliminary exploration on low-carbon technology roadmap of China's power sector. *Energy,* 36: 1500–1512.

Crilly, D. and T. Zhelev (2008). Emissions targeting and planning: An application of CO2 emissions pinch analysis (CEPA) to the Irish electricity generation sector. *Energy,* 33: 1498–1507.

Diamante, J.A.R., R.R. Tan, D.C.Y. Foo, D.K.S. Ng, K.B. Aviso, and S. Bandyopadhyay (2014). Unified pinch approach for targeting of carbon capture and storage (CCS) systems with multiple time periods and regions. *Journal of Cleaner Production,* 71: 67–74.

El-Halwagi, M.M. (2011). *Sustainable Design through Process Integration*. Butterworth-Heinemann, Boston, MA.

El-Halwagi, M.M. and V. Manousiouthakis (1989). Synthesis of mass exchange networks. *AIChE Journal,* 35: 1233–1244.

Foo, D.C.Y. (2012). *Process Integration for Resource Conservation*. CRC Press, Boca Raton, FL.

Foo, D.C.Y., N. Hallale, and R.R. Tan (2010). Optimize shift scheduling using pinch analysis. *Chemical Engineering,* 117: 48–52.

Furman, K.C. and N.V. Sahinidis (2002). A critical review and annotated bibliography for heat exchanger network synthesis in the 20th century. *Industrial & Engineering Chemistry Research,* 41(10): 2335–2370.

Grip, C.E., J. Isaksson, S. Harvey, and L. Nilsson (2013). Application of pinch analysis in an integrated steel plant in northern Sweden. *ISIJ International*, 53(7): 1202–1210.

Hallale, N. and D.M. Fraser (1998). Capital cost targets for mass exchange networks a special case: Water minimization. *Chemical Engineering Science,* 53: 293–313.

Hohmann, E.C. (1971). Optimum networks for heat exchange. Ph.D. Thesis, University of Southern California, Los Angeles, CA.

Kazantzi, V. and M.M. El-Halwagi (2005). Targeting material use via property integration. *Chemical Engineering Progress,* 101: 28–37.

Kemp, I.C. (2007). *Pinch Analysis and Process Integration: A User Guide for the Efficient Use of Energy.* Elsevier, Oxford, UK.

Klemeš, J.J. (Ed.) (2013). *Handbook of Process Integration (PI): Minimization of Energy and Water Use, Waste and Emissions*. Woodhead Pub. Ltd., Cambridge, UK.

Krishna Priya, G.S. and S. Bandyopadhyay (2013). Emission constrained power system planning: A pinch analysis based study of Indian electricity sector. *Clean Technologies and Environmental Policy,* 15: 771–782.

Linnhoff, B. (1993). Pinch analysis: A state-of-art overview. *Transactions of the American Institute of Chemical Engineers (Part A),* 71: 503–522.

Linnhoff, B. and S. Ahmad (1989). Supertargeting: Optimum synthesis of energy management systems. *Journal of Energy Resources Technology,* 111(3): 121–130.

Linnhoff, B. and S. Ahmad (1990). Cost optimum heat exchanger networks—1. Minimum energy and capital using simple models for capital cost. *Computers & Chemical Engineering,* 14(7): 729–750.

Linnhoff, B. and J.R. Flower (1978). Synthesis of heat exchanger networks: 1. *AIChE Journal*, 24: 633–642.

Linnhoff, B. and E. Hindmarsh (1983). The pinch design method for heat exchanger networks. *Chemical Engineering Science,* 38(5): 745–763.

Linnhoff, B., D.R. Mason, and I. Wardle (1979). Understanding heat exchanger networks. *Computers & Chemical Engineering,* 3(1–4): 295–302.

Linnhoff, B., D.W. Townsend, D. Boland, G.F. Hewitt, B.E.A. Thomas, A.R. Guy, and R.H. Marshall, (1982). *A User Guide on Process Integration for the Efficient Use of Energy.* Institute of Chemical Engineers, Rugby, UK.

Mehta, R.K.C., S.K. Devalkar, and S. Narasimhan (2001). An optimization approach for evolutionary synthesis of heat exchanger networks. *Transactions of the IChemE,* 79: 143:150.

Ozkan, S. and S. Dincer (2001). Application for pinch design of heat exchanger networks by use of a computer code employing an improved problem algorithm table. *Energy Conversion and Management,* 42: 2043–2051.

Sahu, G.C., S. Bandyopadhyay, D.C.Y. Foo, D.K.S. Ng, and R.R. Tan (2014). Targeting for optimal grid-wide deployment of carbon capture and storage (CCS) technology. *Process Safety and Environmental Protection,* 92: 835–848.

Salama, A.I.A. (2005). Numerical techniques for determining heat energy targets in pinch analysis. *Computers and Chemical Engineering,* 29: 1861–1866.

Salama, A.I.A. (2006). Determination of the optimal heat energy targets in heat pinch analysis using a geometry-based approach. *Computers and Chemical Engineering,* 30: 758–764.

Shenoy, U.V. (1995). *Heat Exchanger Network Synthesis: Process Optimization by Energy and Resource Analysis.* Gulf Publishing, Houston, TX.

Shenoy, U.V., A. Sinha, and S. Bandyopadhyay (1998). Multiple utilities targeting for heat exchanger networks. *Chemical Engineering Research and Design,* 76: 259–272.

Singhvi, A. and U.V. Shenoy (2002). Aggregate planning in supply chains by pinch analysis. *Chemical Engineering Research and Design,* 80: 597–605.

Smith, R. (2005). *Chemical Process Design and Integration*, 2nd Edition. John Wiley & Sons Inc, Chichester, UK.

Smith, R. (2016). *Chemical Process: Design and Integration*. John Wiley & Sons, Chichester, UK.

Svensson, E. and S. Harvey (2011). Pinch analysis of a partly integrated pulp and paper mill. *World Renewable Energy Congress 2011*, May 8–13, Sweden.

Tan, R.R. and D.C.Y. Foo (2007). Pinch analysis approach to carbon-constrained energy sector planning. *Energy,* 32: 1422–1429.

Tan, R.R., D.K.S. Ng, and D.C.Y. Foo (2009). Pinch analysis approach to carbon-constrained planning for sustainable power generation. *Journal of Cleaner Production,* 17: 940–944.

Umeda, T., J. Itoh, and K. Shiroko (1978). Heat exchange system synthesis. *Chemical Engineering Progress,* 74(7): 70–76.

Upadhyaya, S. (2010). Compilation of energy statistics for economic analysis, Working paper 01/2010, UNIDO, Vienna.

Varghese, J. and S. Bandyopadhyay (2007). Targeting for energy integration of multiple fired heaters. *Industrial & Engineering Chemistry Research,* 46(17): 5631–5644.

Wang, Y.P. and R. Smith (1994). Wastewater minimization. *Chemical Engineering Science,* 49: 981–1006.

4.15 AIR-CONDITIONING SYSTEMS

Donald L. Fenton

Introduction

The term *air conditioning* refers to all equipment and processes used in maintaining comfortable and healthy indoor air suitable for human occupancy. This includes the adjustment or conditioning of the air, the distribution of air, air motion in the occupied space, and cleanliness of the air. This section of the handbook treats the thermal processes important in maintaining the temperature and humidity of the spaces occupied by humans. The acronym HVAC&R signifies heating, ventilating, air conditioning, and refrigerating where all these activities fall under the umbrella of air conditioning. Consequently, air conditioning is not restricted to only the cooling and dehumidifying of air.

Maintaining comfortable indoor air conditions requires understanding the behavior of the air and water vapor mixture—a subject called psychrometry. A space occupied by humans enclosed by an insulated structure limits the energy gains in the summer and losses in the winter. These energy gains and losses are referred to as cooling loads and heating loads, respectively. They are the energy transfer rates that the heating and refrigerating equipment must meet to ensure indoor comfort. These loads are based on the determination of the severe, but likely, outdoor weather conditions typical for the specific geographical location. The introduction of conditioned air to the occupied space must be sufficient to offset the summer and winter loads. The heating and cooling equipment is generally located a distance away from the space, necessitating an air duct system for the conveyance of the conditioned air.

The inch-pound unit system is used, reflecting its prevalence in the HVAC&R industry in the U.S. Conversion to SI units may be made using suitable factors.

Properties of Moist Air

The thermodynamic properties of moist air and how these properties may change is the subject generally referred to as psychrometrics. Moist air is a binary mixture of dry air and water vapor. NOAA (2016) and Schlatter (2009) report that the composition of dry air is nitrogen, 78.084%(v/v); oxygen, 20.946%(v/v); argon, 0.9340%(v/v); carbon dioxide, 0.0397%(v/v); neon, 0.001818%(v/v); helium, 0.000524%(v/v); methane, 0.000179% (v/v); sulfur dioxide, 0 to 0.0001%(v/v); hydrogen, 0.00005%(v/v); and other negligible gases such as krypton, xenon, and ozone. Carbon dioxide concentration measurements reported from 1995 to 2015 (McGee, 2016) suggest that the annual increase in carbon dioxide concentration is 0.00020% per year. Water vapor concentration varies from 0.001% to 5.0% depending on local conditions and is not included in the dry air composition (Wallace and Hobbs, 2006). The apparent molecular weight of the dry air gas mixture is 28.9645. The dry air gas constant is computed by

$$R_a = \frac{R}{M} = \frac{1545.32}{28.9645} = 53.352 \frac{\text{ft lbf}}{\text{lbm°R}}, \tag{4.15.1}$$

where the molecular weight value is based on carbon having a molecular weight of exactly 12.

Moist air is a mixture of dry air and water vapor. All other contaminants that may be present in the atmospheric air are excluded from moist air. These contaminants may include smoke, pollen, dust, and gaseous pollutants generally present in the ambient or outside air. The quantity of water vapor in moist air varies from none to a maximum amount that depends on the temperature and pressure of the air. This condition of maximum water vapor content is called saturation, which is an

equilibrium condition between the moist air and condensed water. The molecular weight of water is 18.01528 and its gas constant is

$$R_{H_2O} = \frac{1545.32}{18.01528} = 85.778 \frac{\text{ft lbf}}{\text{lbm°R}}. \tag{4.15.2}$$

The barometric pressure and temperature of atmospheric air changes with increasing altitude above sea level. Standard atmospheric conditions are defined at sea level where the pressure is 29.921 in. of mercury (in.Hg) and the temperature is 59°F. The U.S. Standard Atmosphere defines the variation of pressure and temperature with altitude and is generally used to estimate these properties at different altitudes. The troposphere (lower atmosphere) is taken to consist of dry air behaving as an ideal gas. The acceleration due to gravity is also taken to be a constant, 32.1740 ft/s^2. The pressure and temperature are given by

$$T = 59 - 0.00356616Z, \tag{4.15.3}$$

$$p = 14.696(1 - 6.8754 \times 10^{-6})^{5.2559}, \tag{4.15.4}$$

where Z is the altitude (ft), p is the barometric pressure (in.Hg), and T is the temperature (°F). Equations 4.15.3 and 4.15.4 are accurate from sea level to an altitude of 36,000 ft. More extensive information including higher altitudes, other physical data, and SI units may be obtained from NASA (1976).

Thermodynamic Properties of Saturated Water

The thermodynamic saturated state is discussed in Chapter 3 of this handbook. Relationships giving the thermodynamic properties of saturated water have been developed by Herrmann et al. (2009) over the temperature range from −80°F to 200°F. The determination of these properties is based on an ideal temperature scale (International Temperature Scale of 1990 (ITS-90)) almost identical to the generally used practical temperature scale used for measurements. As an example, the standard boiling temperature of water (at 14.696 psia or 29.921 in.Hg) is 211.95°F rather than the usual value of 212°F on the practical scale. The enthalpy and entropy properties of saturated liquid water have values defined as zero at the triple point, 32.018°F. Water at a temperature between its triple point and critical point may coexist in two states, liquid and vapor. When these phases are in equilibrium, all the physical features of the system remain constant with time and the vapor and liquid are saturated.

Over liquid water, the saturation pressure for the temperature range from 32°F to 392°F is

$$\ln(p_{ws}) = \frac{C_1}{T} + C_2 + C_3 T + C_4 T^2 + C_5 T^3 + C_6 \ln T, \tag{4.15.5}$$

where $C_1 = -1.0440397 \times 10^4$, $C_2 = -1.1294650 \times 10^1$, $C_3 = -2.7022355 \times 10^{-2}$, $C_4 = 1.2890360 \times 10^{-5}$, $C_5 = -2.4780681 \times 10^{-9}$, and $C_6 = 6.5459673$, and where T is the absolute temperature (°R) and the subscript "ws" indicates saturated water. The saturation pressure over ice in terms of temperature for a range from −148°F to 32°F is

$$\ln(p_{ws}) = \frac{C_7}{T} + C_8 + C_9 T + C_{10} T^2 + C_{11} T^3 + C_{12} T^4 + C_{13} \ln T, \tag{4.15.6}$$

where $C_7 = -1.0214165 \times 10^4$, $C_8 = -4.8932428$, $C_9 = -5.3765794 \times 10^{-3}$, $C_{10} = 1.9202377 \times 10^{-7}$, $C_{11} = 3.5575832 \times 10^{-10}$, $C_{12} = -9.0344688 \times 10^{-14}$, and $C_{13} = 4.1635019$. The ASHRAE *Handbook of Fundamentals* (2013) gives the thermodynamic properties of water at saturation.

Thermodynamic Properties of Moist Air

Hyland and Wexler (1983a,b) developed formulas for the thermodynamic properties of moist air based on the "thermodynamic temperature scale," an ideal temperature scale almost identical to the generally used practical scale used for physical measurements. The ASHRAE *Handbook of Fundamentals* (2013) gives the thermodynamic properties of moist air based on the Hyland and Wexler equations. These thermodynamic properties are as follows:

Temperature (°F): T, related to the absolute temperature (°R) by T (°R) = T (°F) + 459.67.
Humidity Ratio: W_s, the condition when the water vapor phase is in thermodynamic equilibrium with the condensed water at a certain temperature and pressure. The humidity ratio for a given moist air condition may have a humidity ratio, W, such that it has a value of zero or greater but less than or equal to W_s.
Specific Volume (ft^3/lb): v_a, dry air specific volume.
Saturated Moist Air Specific Volume (ft^3/lb): v_s, volume of saturated moist air per pound of dry air.
Specific Enthalpy of Dry Air (Btu/lb dry air): h_a, reference value is zero at 0°F and 14.696 psia.
Saturated Moist Air Specific Enthalpy Air (Btu/lb dry air): h_s.
Condensed Water Specific Enthalpy (Btu/lb water): h_w, specific enthalpy of water in equilibrium with saturated air at a certain temperature and pressure. The reference value is zero at the triple point of water and saturation pressure.
Water Vapor Pressure (psia): p_w, the actual partial pressure in moist air exerted by the water vapor.
Saturated Water Vapor Pressure (psia): $p_{s,}$ water vapor pressure in saturated air. Very nearly the same as the saturation vapor pressure of pure water, p_{ws}. Therefore, p_s and p_{ws} may be interchanged for one another.
Barometric Pressure (psia): p, atmospheric pressure.
Specific Entropy of Dry Air (Btu/lb°R): s_a, reference value is zero at 0°F and 14.696 psia.
Saturated Moist Air Specific Entropy (Btu/lb°R): s_s.
Condensed Water Specific Entropy (Btu/lb°R): s_w, specific entropy of water in equilibrium with saturated air. The reference value is zero at the triple point of water and saturation pressure.

Numerical values of the properties of moist air are calculated on the basis of several assumptions. These are as follows:

1. The gaseous phase behaves as a mixture of ideal gases.
2. The liquid and solid phases contain no dissolved gases.
3. The equilibrium between the condensed phase and the gas mixture is not affected by the presence of the other component when the mixture and the condensed phase are at the same pressure and temperature.

Several additional properties and related mathematical relationships are useful when considering calculations relating problems regarding moist air. These are as follows:

Dry-Bulb Temperature (°F): T, temperature of air as indicated by a thermometer.
Thermodynamic Wet-Bulb Temperature (°F): T^*, the temperature of water (may be liquid or solid) evaporating into moist air at a certain dry-bulb temperature T and humidity ratio W, brings the air to saturation adiabatically at the same temperature T^* maintaining the pressure at a constant value (Howell et al., 1998).

Wet-Bulb Temperature (°F): T_{wb}, the equilibrium temperature of water in a wet wick surrounding the bulb of a thermometer placed in an air stream. A psychrometer is used to measure simultaneously the dry-bulb and the wet-bulb temperatures using two thermometers. While this process involves the simultaneous transfer of heat and mass from the bulb, it is different than the adiabatic saturation process defining the thermodynamic wet-bulb temperature. However, the correction needed to convert wet-bulb temperatures to thermodynamic wet-bulb temperatures is small.

Humidity Ratio: the humidity ratio is calculated as the mass of water divided by the dry air mass in a sample of moist air.

$$W = \frac{m_w}{m_a} \tag{4.15.7}$$

or that

$$W = \frac{0.621945\, p_w}{(p - p_w)}. \tag{4.15.8}$$

Manipulation of the psychrometric and thermodynamic properties yields

$$W = \frac{(1093 - 0.556T^*)W_s^* - 0.240(T - T^*)}{1093 + 1.805T - 4.186T^*}, \tag{4.15.9}$$

where W_s^* is the saturated humidity ratio if saturated at the wet-bulb temperature and where T and T^* are in °F.

Degree of Saturation: μ, the ratio of the actual humidity ratio divided by the saturated air humidity ratio W_s at the same pressure and temperature.

$$\mu = \frac{W}{W_s}. \tag{4.15.10}$$

Relative Humidity: ϕ, the ratio of the water vapor mole fraction in a given sample of moist air divided by the water vapor mole fraction of a sample saturated at the same pressure and temperature. Using the assumption that the gas mixture may be considered as a mixture of ideal gases gives

$$\phi = \frac{p_w}{p_{ws}}. \tag{4.15.11}$$

Dew-Point Temperature (°F): T_d, the temperature of a sample of moist air after having become saturated at the same pressure and humidity ratio as the original sample of moist air. Also, the dew-point temperature is the saturation temperature corresponding to the saturation pressure as listed in the thermodynamic properties of saturated water. For the temperature range from 32°F to 200°F, the dew-point temperature is given by

$$T_d = 100.45 + 33.193 \ln p_w + 2.319(\ln p_w)^2 + 0.17074(\ln p_w)^3 + 1.2063(p_w)^{0.1984} \tag{4.15.12}$$

and for temperatures less than 32°F,

$$T_d = 90.12 + 26.142 \ln p_w + 0.8927(\ln p_w)^2 \tag{4.15.13}$$

by ASHRAE (2013).

Other properties include specific volume and enthalpy. The specific volume of moist air is given in units of dry air where the atmospheric pressure is the sum of the dry air and water vapor partial pressures. Thus,

$$v = \frac{R_a T}{(p - p_w)}, \tag{4.15.14}$$

where p is the atmospheric pressure, p_w is the water vapor partial pressure, and R_a is the gas constant for air (53.3504 ft-lbf/lbm°R). Combining the expression for W above gives an alternate relationship for v sometimes useful in calculations,

$$v = 0.370486(T + 459.67)\frac{(1 + 1.6058W)}{p}. \tag{4.15.15}$$

The enthalpy is equal to the sum of the enthalpies of the dry air and the water vapor, or

$$h = h_a + Wh_g, \tag{4.15.16}$$

where h_a is the enthalpy of the dry air and h_g is the specific enthalpy for saturated water vapor at the temperature of the moist air mixture. Approximate values for h_a (Btu/lb) and h_g (Btu/lb) are

$$h_a = 0.240T, \tag{4.15.17}$$

$$h_g = 1061 + 0.444T, \tag{4.15.18}$$

where T is the dry-bulb temperature. The resulting expression for the enthalpy h (Btu/lb) is

$$h = 0.240T + W(1061 + 0.444T). \tag{4.15.19}$$

The above relationships for the moist air mixture are generally adequate for air-conditioning design calculations.

Psychrometric Chart

The psychrometric chart displays simultaneously a number of the thermodynamic properties of moist air. An example psychrometric chart is shown in Figure 4.15.1, where several variables are plotted against temperature and the humidity ratio. Processes involving heat transfer and the addition or removal of moisture are readily shown on the psychrometric chart and so the chart is an aid in visualizing the process. Furthermore, the changes that occur in the moist air properties may be observed from the chart (Figure 4.15.1), particularly changes in dry-bulb temperature, humidity ratio, and enthalpy. Observe that the horizontal movement across the chart involves only a change in dry-bulb temperature or a sensible change. On the other hand, a vertical movement on the chart represents a pure latent change where the dry-bulb temperature remains constant but the humidity ratio changes. Figure 4.15.2 shows a schematic view of the psychrometric chart where the process directions common to air conditioning are shown.

We now consider several examples where the use of the psychrometric chart is demonstrated.

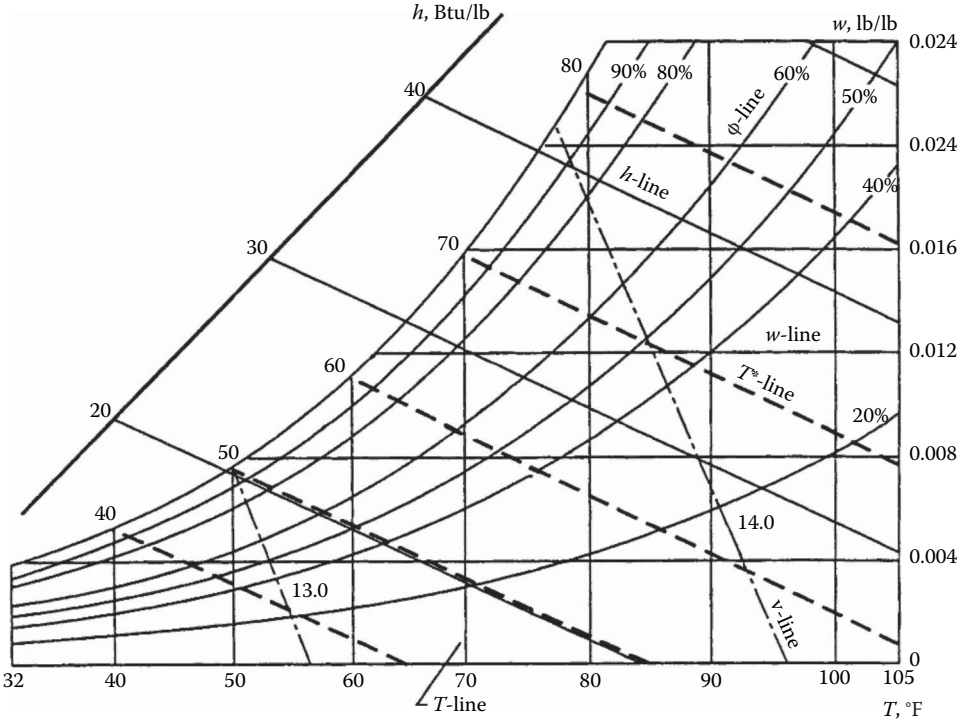


FIGURE 4.15.1 Psychrometric chart. (From Kreith, F., Ed., *The CRC Handbook of Mechanical Engineering*, CRC Press LLC, Boca Raton, FL, 1998. With permission.)

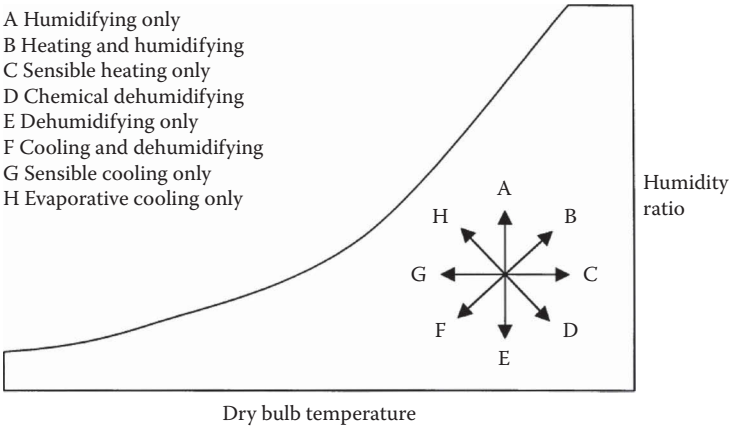


FIGURE 4.15.2 Air conditioning process directions on a psychrometric chart.

Cooling and Dehumidifying Air

Moist air, when cooled to a temperature below its dew point, will condense a portion of the water vapor initially in the mixture. Equipment that could be used to carry out the cooling and dehumidifying process is shown in Figure 4.15.3, where a refrigerant enters and leaves an evaporator coil. The air inlet and exit conditions are defined as well as the condensed water drained from the pan below the coil. If the system operates on a steady-state flow basis (all variables remain constant with time), then the process may be described by the following water mass and energy equations:

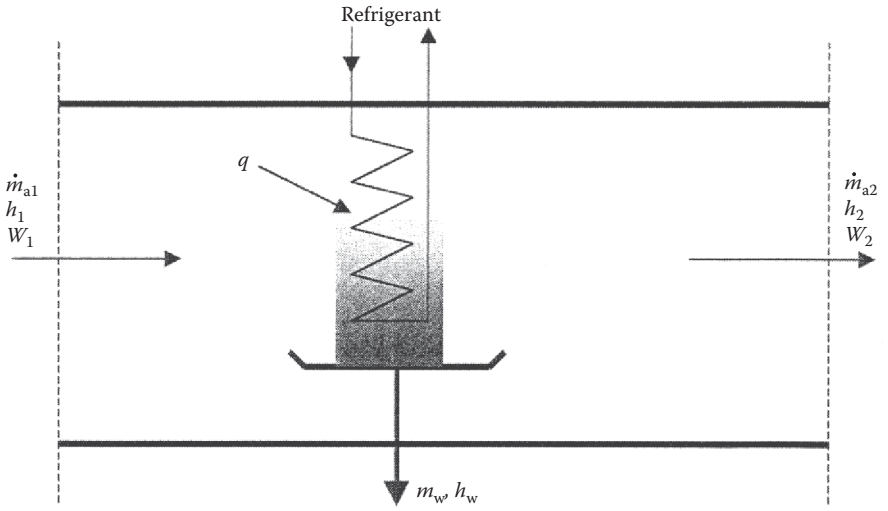


FIGURE 4.15.3 Schematic diagram of cooling and dehumidifying equipment.

$$\dot{m}_a W_1 = \dot{m}_a W_2 + \dot{m}_w, \tag{4.15.20}$$

$$\dot{m}_a h_1 = \dot{m}_a h_2 + \dot{m}_w h_w + q, \tag{4.15.21}$$

where q is the heat transfer from the moist air to the coil. These two equations may be rearranged to give

$$\dot{m}_w = \dot{m}_a (W_1 - W_2), \tag{4.15.22}$$

$$q = \dot{m}_a [(h_1 - h_2) - (W_1 - W_2) h_w], \tag{4.15.23}$$

where the equation for q indicates that the cooling and dehumidifying process includes sensible and latent heat transfer. The sensible and latent heat transfer quantities are

$$q(\text{sensible}) = \dot{m}_a (h_1 - h_2) = \dot{m}_a c_p (T_1 - T_2), \tag{4.15.24}$$

$$q(\text{latent}) = \dot{m}_a (W_1 - W_2) h_{fg}. \tag{4.15.25}$$

Figure 4.15.2 shows the direction of this process on a psychrometric chart where the process is noted to be a line sloping downward to the left. The sensible and latent heat transfers are noted by the horizontal and vertical changes in the condition of the air when passing from the inlet to the outlet, respectively.

Heating and Humidifying Air

The equipment shown in Figure 4.15.4 may be used to heat and humidify moist air during the winter heating season. Water mass and energy balances on the system yield

$$\dot{m}_a W_1 + \dot{m}_w = \dot{m}_a W_2, \tag{4.15.26}$$

$$\dot{m}_a h_1 + q + \dot{m}_w h_w = \dot{m}_a h_2, \tag{4.15.27}$$

which may be solved to give the humidity ratio and enthalpy of the exit air.

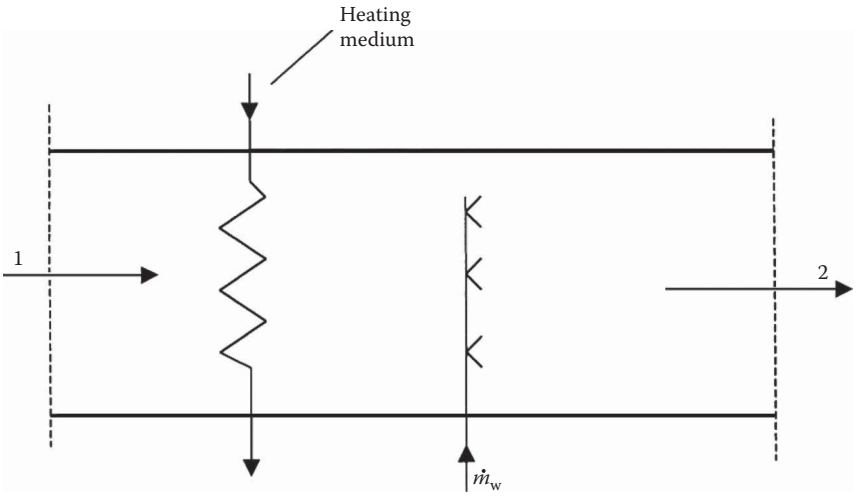


FIGURE 4.15.4 Schematic diagram of typical heating and humidifying equipment.

Adiabatic Mixing of Two Streams

Two streams of moist air converging into one stream are shown in Figure 4.15.5. If there is no heat transfer from the ducts to the surroundings, the mixing process is adiabatic, which is closely approximated by insulated ducts, then the equations describing this process include an air mass balance, a water mass balance, and an energy balance, respectively.

$$\dot{m}_{a1} + \dot{m}_{a2} = \dot{m}_{a3}, \tag{4.15.28}$$

$$\dot{m}_{a1}W_1 + \dot{m}_{a2}W_2 = \dot{m}_{a3}W_3, \tag{4.15.29}$$

$$\dot{m}_{a1}h_1 + \dot{m}_{a2}h_2 = \dot{m}_{a3}h_3. \tag{4.15.30}$$

The above three equations indicate that the leaving humidity ratio and enthalpy are weighted averages of the entering corresponding values. In practice, the approximation is made that the exit temperature is also a weighted average of the entering temperatures where the error, due to the variation

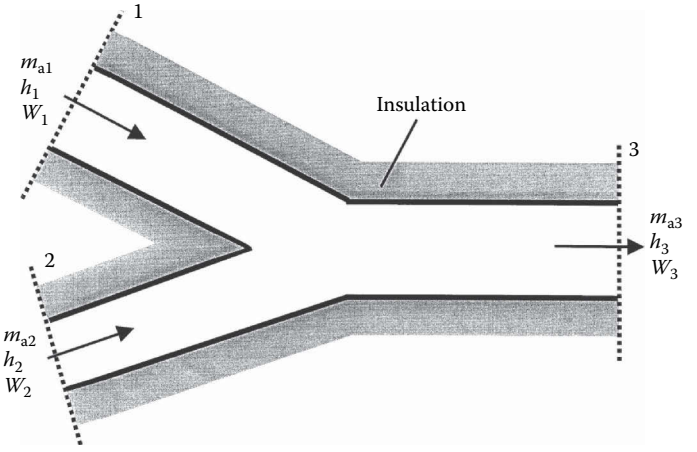


FIGURE 4.15.5 Adiabatic mixing of two streams of moist air.

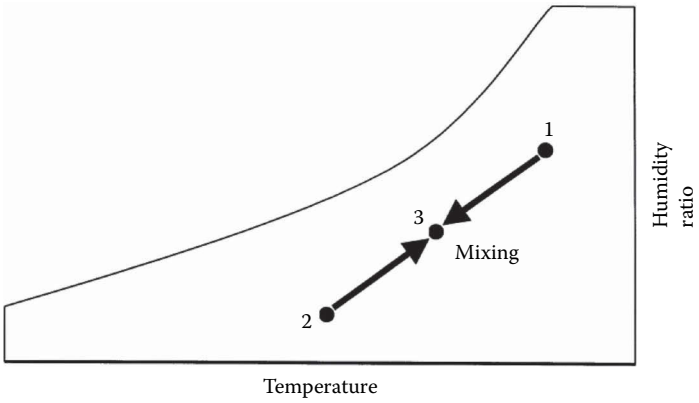


FIGURE 4.15.6 Adiabatic mixing of two streams.

of the specific heats for the moist air, is typically less than 1%. Consequently, the exit condition of the moist air lies on a straight line connecting the two entering conditions, or that

$$\frac{\dot{m}_{a2}}{\dot{m}_{a1}} \cong \frac{\text{distance (1) to (3)}}{\text{distance (2) to (3)}}. \tag{4.15.31}$$

Figure 4.15.6 shows the adiabatic mixing process path on a schematic of the psychrometric chart. The above three equations may also be combined to give

$$\frac{\dot{m}_{a2}}{\dot{m}_{a1}} = \frac{h_3 - h_1}{h_2 - h_3} = \frac{W_3 - W_1}{W_2 - W_3} \cong \frac{T_3 - T_1}{T_2 - T_3}. \tag{4.15.32}$$

Example

Consider the mixing of two streams of moist air where one entering stream has a volumetric flow rate of 5000 cfm with a 40°F dry-bulb temperature and a 35°F wet-bulb temperature. The other inlet stream has a volumetric air flow rate of 15,000 cfm with 75°F dry-bulb temperature and 50% relative humidity. Determine the exit conditions for the leaving air stream.

Solution:
Label the 5000 cfm air stream as "1" and the 15,000 cfm stream as "2." The exit stream is labeled as "3." From the psychrometric chart, Figure 4.15.1, the following information is observed:

$$v_1 \cong 12.65 \text{ ft}^3/\text{lb}, h_1 \cong 14.0 \text{ Btu/lb},$$

$$v_2 \cong 13.68 \text{ ft}^3/\text{lb}, h_2 \cong 28.0 \text{ Btu/lb}.$$

Thus,

$$\frac{\dot{m}_{a1}}{\dot{m}_{a2}} = \frac{\dfrac{\dot{V}_1}{v_1}}{\dfrac{\dot{V}_2}{v_2}} = \frac{\dfrac{5000}{12.65}}{\dfrac{15000}{13.68}} = 0.3605$$

$$0.3605 = \frac{h_3 - h_2}{h_1 - h_3} = \frac{h_3 - 28.0}{14.0 - h_3}$$

$$h_3 = 24.3 \text{ Btu/lb}$$

and from the psychrometric, observe that T_3 is approximately 66°F.

Example

Consider the air-conditioned space shown in Figure 4.15.7, where moist air leaves at 80°F dry-bulb and 65°F thermodynamic wet-bulb temperatures. Due to human occupancy, the moisture gain to the space occurs at the rate of 10.5 lb/h as saturated water vapor at 90°F. The space also has a sensible gain of 20,000 Btu/h. Moist air is introduced into the space at 55°F. Estimate the needed thermodynamic wet-bulb temperature and the volumetric flow rate of the supply air.

Solution:
The sensible heat gain does not involve any heat transfer due to changes in moisture content of the air. The water vapor added to the space adds energy equal to its specific enthalpy. The governing equations involve air, water, and energy balances

$$\dot{m}_{a1} = \dot{m}_{a2} = \dot{m}_a,$$

$$\dot{m}_{a1} W_1 + \dot{m}_w = \dot{m}_{a2} W_2,$$

$$\dot{m}_{a1} h_1 + \dot{m}_w h_w + q_s = \dot{m}_a h_2,$$

where q_s is the added sensible energy, $m_w\, h_w$ is the latent load due to the moisture content change, and $\dot{m}_a$ is the dry air flow rate entering and leaving the space. Rewriting the energy balance gives

$$q_s + \dot{m}_w h_w = \dot{m}_a (h_2 - h_1)$$

and rearranging the water balance yields

$$\dot{m}_w = \dot{m}_a (W_2 - W_1).$$

Compute the ratio:

$$\frac{\text{Total energy added to space}}{\text{Total water added to space}} = \frac{q_s + \dot{m}_w h_w}{m_w}$$

$$= \frac{\dot{m}_a (h_2 - h_1)}{\dot{m}_a (W_2 - W_1)} = \frac{\Delta h}{\Delta W},$$

where the ratio is observed to also equal the enthalpy change divided by the humidity ratio change. A protractor on the psychrometric chart (Figure 4.14.8) shows the slopes associated with

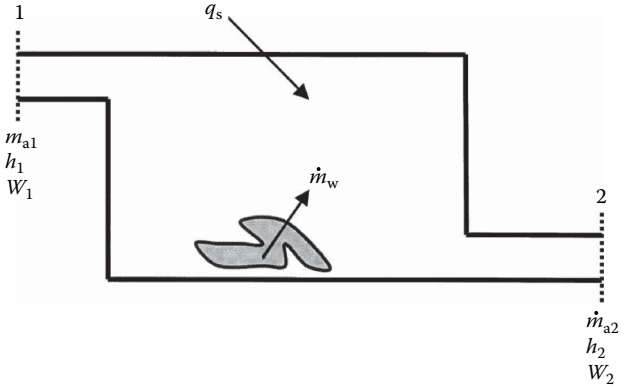


FIGURE 4.15.7 Schematic diagram of an air-conditioned space.

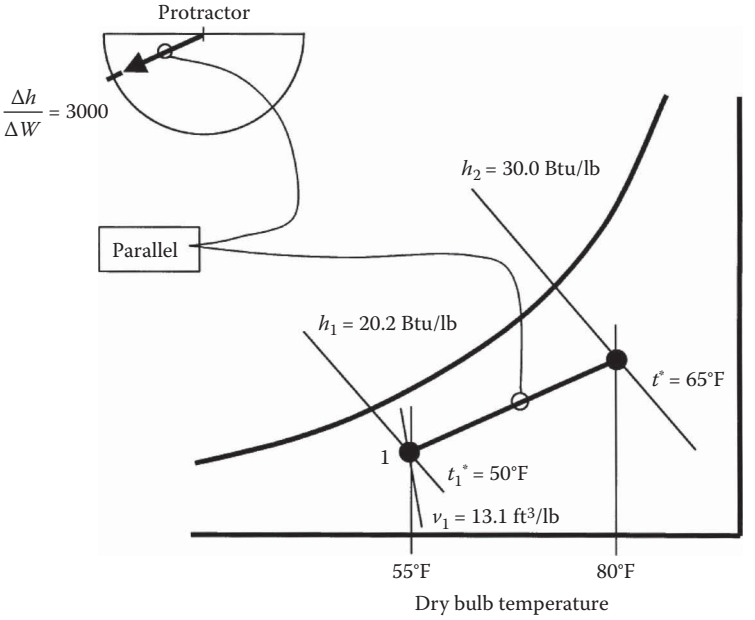


FIGURE 4.15.8 Schematic of the psychrometric chart for an air-conditioned space.

the direction that this process takes on the chart. All the possible supply air conditions lie on this line. From the saturated water thermodynamic property tables, $h_w = 1100.4$ Btu/lb. Thus,

$$\frac{\Delta h}{\Delta W} = \frac{20{,}000 + (10.5)(1100.4)}{10.5} = 3005 \cong 3000.$$

Draw a line through state "2" at a slope equal to $\Delta h/\Delta W = 3000$. Where this line intersects the supply dry-bulb temperature identifies the wet-bulb temperature, which is approximately 50°F. To obtain the volumetric flow rate at state "1," first calculate the mass flow rate:

$$\dot{m}_a = \frac{q_s + \dot{m}_w h_w}{(h_2 - h_1)}.$$

Estimate from the psychrometric chart that $h_1 = 20.2$ Btu/lb (dry air) and $h_2 = 30.0$ Btu/lb (dry air). Thus,

$$\dot{m}_a = \frac{20{,}000 + (10.5)(1100.4)}{(30.0 - 20.2)} = 3220 \frac{\text{lb(dry air)}}{\text{h}} \text{ or } 53.7 \frac{\text{lb(dry air)}}{\text{min}}.$$

The volumetric flow rate is

$$\dot{V}_1 = \dot{m}_a v_1$$

and from the psychrometric chart, $v_1 = 13.1$ ft³/lb (dry air). Thus,

$$\dot{V}_1 = (53.7)(13.1) = 703 \frac{\text{ft}^3}{\text{min}}.$$

Thermal Comfort Conditions

The now widely accepted definition for human comfort is "that condition of mind that expresses satisfaction with the thermal environment" (ANSI/ASHRAE *Standard 55* 2010). Thus, the quantification of human comfort is nebulous in that a person's feelings are important. A number of inside

environmental conditions influence human thermal comfort and include dry-bulb temperature, wet-bulb temperature, dew-point temperature, humidity ratio, air velocity, and the mean radiant temperature. The mean radiant temperature is a representative temperature for all the individual surfaces that exist in the occupied space. The amount and type of clothing affects human comfort. Furthermore, several environmental temperature indices are used to assist with the quantification of thermal comfort. These indices are equivalent temperature, effective temperature, humid operative temperature, the globe temperature, and the resultant temperature, which is applied sometimes in the United Kingdom. As an example of one of these temperature indices, the human operative temperature T_{oh} is the temperature when 100% relative humidity yields the same total heat loss from the skin as would occur for the actual environment (ASHRAE, 2017). Specifically,

$$T_{oh} = T_o + wi_m LR(p_v - p_{oh,s}), \tag{4.15.33}$$

where T_0 is the operative temperature, w is the wet skin fraction, i_m is the moisture permeability index, LR is the Lewis relation, p_v is the partial pressure of water vapor, and $p_{oh,\,s}$ is the saturated vapor pressure at T_{oh}. The operative temperature is given by

$$T_o = \frac{h_r T_{mr} + h_c T_a}{h_r - h_c} \tag{4.15.34}$$

and the Lewis relation is given by

$$LR = \frac{h_e}{h_c}, \tag{4.15.35}$$

where h_c is the convective heat transfer coefficient, h_r is the linearized radiative heat transfer coefficient, h_e is the evaporative heat transfer coefficient, T_{mr} is the mean radiant temperature of surfaces in the space, and T_a is the space air temperature. Note that the operative temperature is the average of the mean radiant and air temperatures weighted by their respective heat transfer coefficients. While these equations and definitions reveal the complexities associated with quantifying human comfort parameters, additional details are needed to make calculations. Refer to the ASHRAE *Handbook of Fundamentals* (2017) for more information.

ANSI/ASHRAE Standard 55 (2010) defines comfort zones, which specify the conditions for reasonable comfort for human occupants either sedentary or slightly active. These comfort zones carefully incorporate the results obtained through research and are different for winter and summer due to the clothing typically worn during these periods. The boundaries of the comfort zones are not sharp, with the upper and lower humidity levels being less precise than the left and right boundaries associated with temperature. Indoor conditions of temperature and humidity for the conditioned space are usually selected to comply with the ANSI/ASHRAE Standard 55 (2010) limits. The specific influence of humidity on human comfort is explored in greater detail by Berglund (1995).

Several models have been reported in the literature that predict the level of thermal comfort that is likely to occur in a given situation and space. Fanger (1982) developed a predictive method based on a steady-state energy balance and Gagge (1971, 1986) devised a two-node model suitable for predicting steady-state human comfort responses. Multisegment models divide the human body into more segments consisting of the head, chest, hands, and feet along with additional tissue layers, so that thermal comfort in spaces with nonuniform conditions may be predicted (Huizenga et al., 2001; Fiala et al. 2003; Zhang et al. 2010 among others). These predictive models are summarized in ASHRAE (2017).

Load Calculations

In the context of HVAC (heating, ventilating, and air conditioning), the term load is applied to a building to designate the thermal energy removal or addition rate needed to maintain the desired

inside air conditions. For summer months, the removal of thermal energy from the space is required, typically performed by a refrigeration system. Spaces during winter months generally require the addition of thermal energy to maintain the desired inside conditions. Both summer cooling and winter heating involve systems accomplishing the conditioning of the inside air. While the phrase "thermal energy" is technically correct for explaining the air warming and cooling processes, it is common in the industry to use the term "heat" to refer to this energy. The discussion here follows the industry practice of using the term heat.

For cooling loads originating from human comfort requirements in buildings, the absorption of solar radiation into the space is significant. Special procedures have been developed by ASHRAE to accurately predict cooling loads. The current procedures (Pederson et al., 1998; Spitler 2009; Spitler 2014) are improvements of earlier methods (ASHRAE, 1975; McQuiston and Spitler, 1992). The cooling load is a transient phenomenon—the heat transfer rate through the space envelope varies with time. An enclosed space where the interior temperature is controlled by a single thermostat is usually referred to as a zone. A large nonresidential building may be composed of many zones sharing walls between them.

The heat balance concept is essentially the application of a control volume around a zone of a building. To be specific, a heat balance zone is an air volume, uniform in temperature, which also includes all the heat transfer and thermal storage surfaces bounding or inside the air volume. While the concept of a zone originates with the thermal behavior of the space, several rooms may comprise a zone. Family residential buildings are usually treated as being one zone. Observe that the wall treatments on the interior surfaces and the furnishings characterize the heat balance zone. The typical zone is a 12-surface volume that is used in determining loads where complex geometry of an actual building is reduced to a simple shape accurately duplicating the zone's thermal features. The important information includes the following:

- Orientation of each surface and configuration area
- Construction features of each surface
- Outside and inside surface environmental conditions

The heat balance approach to load calculations necessitates the use of a digital computer. While it is possible to perform approximate load calculations (more so with the heat load), it is now considered impractical due to the prevalent accessibility of computers.

The migration of moisture through the materials comprising the building's envelope is of concern as well. The accumulation of moisture within the walls of a structure must be avoided to maintain structural integrity, insulating capacity, and indoor air quality. The growth of molds and other biological items are accelerated by the presence of trapped moisture in walls, which in turn degrade the interior air quality. Hosni et al. (1998) have investigated this subject concerning the special concerns that result from high moisture climates such as southern Texas in the U.S. Additionally, Burch and Chi (1997) have developed a moisture migration model, used and partially experimentally verified by Hosni et al. (1998), quantifying the water migration rates in wall structures. The transport of heat, air, and moisture in building material assemblies is further discussed by ASHRAE (2017).

Cooling Load

The cooling load associated with a zone is difficult to accurately predict. For this reason, ASHRAE has continued to improve the available methods (ASHRAE, 2017) for the prediction of cooling load. The use of the heat balance concept, developed by Pederson et al. (1998) and subsequently improved by Spitler (2009, 2014), has simplified the calculation procedures while maintaining adequate accuracy. This is accomplished by performing four heat balances as described below. The cooling load calculations are sufficiently complex that digital computing machines are needed to make the calculations.

Heat Balance on Outside Surface

A heat balance applied to the outer surface of an exterior wall includes three sources of heat: direct and diffuse radiation from the sun (shortwave radiation), radiation from the environment (longwave radiation), and convection heat transfer. The heat balance on the outside surface is

$$q''_{\text{asol}} + q''_{\text{lwr}} + q''_{\text{co}} - q''_{\text{ko}} = 0, \tag{4.15.36}$$

where q''_{asol} is the absorbed direct and diffuse solar radiant heat flux (Btu/h ft^2), $q''_{\text{lwr}}$is the net longwave radiant flux from the air and objects surrounding the outside surface, $q''_{\text{co}}$is the convective heat flux from the air to the surface, and q''_{ko} is the conductive heat flux leaving the surface entering the wall material. All the heat transfer terms are positive with the exception of the conduction term. Inside the wall, the temperature gradient is negative causing the conduction heat flux term to be positive. The solar radiation, longwave radiation, and convection heat fluxes are sometimes combined using a procedure resulting in the sol–air temperature (ASHRAE, 2017). However, ASHRAE's implementation of the heat balance method involves straightforward models for solar, environmental, and outside convection to determine the outside surface heat balance (ASHRAE, 2017).

- q''_{asol} is determined by procedures identified by ASHRAE *(Handbook of Fundamentals,* 2017).
- q''_{lwr} is estimated using the traditional radiant energy exchange models between surfaces. ASHRAE provides guidance giving values for the sky temperature, ground temperature, surface temperature, surface absorptance, and ground view factors.
- q''_{co} is calculated using Newton's law of cooling where the temperature difference between the outside air and outside surface temperature is given by

$$q''_{\text{co}} = h_{\text{co}}(T_{\text{air}} - T_{\text{os}}), \tag{4.15.37}$$

where h_{co} is the outside convection heat transfer coefficient dependent on the velocity of the air next to the surface, and the T_{air} and T_{os} are temperatures of the outside air and outside surface.

The required outdoor conditions are the dry-bulb and the wet-bulb temperatures. The ASHRAE *Handbook of Fundamentals* (2017) shows these temperatures for a number of locations throughout the U.S. and the world. Cooling design day dry- and wet-bulb temperatures for numerous geographical locations are compiled from weather data. Summer climatic data are based on hourly weather observations over at least a 12-year period. The summer month conditions are reported on the basis of 0.04, 1.0, and 2.0 annual percentiles. These annual percentiles represent, respectively, 35, 88, and 175 h from a year of 8760 h. Typically, July 21 in the Northern Hemisphere is taken as the summer design day.

Interior Wall Conduction

Heat conduction through the material comprising the wall may be determined in several ways. Prevalent methods now include numerical finite difference, numerical finite element, transform methods, and time series methods. All of these methods introduce a time-dependent feature into the cooling load calculation where the inside surface and outside surface temperatures and heat transfer rates are time dependent. Direct analysis of the heat transfer has inputs consisting of two temperature functions and outputs consisting of the two heat fluxes. While any of the four mentioned methods yield usable results, Pederson et al. 1998 and Spitler (2014) employ a conduction transfer function (CTF) procedure for reasons of adequate generality and good computation speed. The CTF factors in this formulation may be considered as response factors associated with the wall.

Heat Balance on Inside Surface

A heat balance on the inside surfaces involves four heat transfer terms: conduction through the wall material, convection to the inside air, shortwave radiant absorption and reflection, and longwave radiant exchange. The heat balance on the inside surface is

$$q''_{ki} + q''_{lwx} + q''_{sw} + q''_{lws} + q''_{sol} + q''_{ci} = 0, \tag{4.15.38}$$

where

q''_{ki} = the conduction heat flux to the inside surface from the interior of the wall material
q''_{lwx} = the net longwave radiant flux between the zone surfaces which involves emittance and absorption from all surfaces in the zone space
q''_{sw} = the net shortwave radiant flux to the inside surface from lights
q''_{lws} = the longwave radiant heat flux from the equipment in the zone
q''_{sol} = the transmitted solar radiation flux absorbed at the inside surface
q''_{ci} = the convective heat flux from the surface to the interior air in the zone

A common approximation used in cooling load calculations, also used by Spitler (2014), is to treat the air as completely transparent to longwave radiation.

Furnishings that may be present inside the zone influences the time response characteristics of the space. This influence is complicated to quantify and is under investigation. The furniture in a space adds surface area which participates in the radiant and convective heat exchange processes tending to reduce the time response. On the other hand, the furnishings also add thermal mass to the space, which increases the time response of the space. The inside surface heat balance permits these two features to be accommodated in a reasonable fashion in that both the surface area and the thermal mass are included in the heat transfer processes.

The longwave radiation from internal sources is traditionally estimated by establishing a division between the radiation and the convection heat transfer for the heat that originates from the equipment. The portion assigned to radiation is then distributed among the surfaces in a reasonable manner (ASHRAE, 2017). While this approach is awkward and deviates from the heat balance approach, it is practical in that it avoids the specification of location and temperature of all equipment in the space. The shortwave radiation from the lights and the transmitted solar energy is also distributed to the surfaces in the space in a reasonable manner (ASHRAE, 2017).

The convective heat flux from the inside surface is calculated using Newton's law of cooling:

$$q''_{ci} = h_{ci}(T_a - T_{is}), \tag{4.15.39}$$

where h_{ci} is the inside surface convective heat transfer coefficient, T_a is the zone air temperature, and T_{is} is the inside surface temperature. Numerical values for h_{ci} are available (ASHRAE, 2017), but are based on carefully controlled natural convection experiments. Present-day conditioned spaces are mechanically ventilated, suggesting that the actual inside surface convective heat transfer coefficients may be somewhat different.

Inside Air Heat Balance

The thermal capacitance is generally neglected for the heat balance applied to the air in the space. With this simplification, the air heat balance is performed on a quasi-steady basis for each calculation time increment. The four heat transfers associated with the heat balance are

$$q_{ci} + q_{ce} + q_{iv} + q_{sys} = 0, \tag{4.15.40}$$

where q_{ci} is the heat transfer rate from the surfaces in the zone, q_{ce} is the convection heat transfer rate originating from the internal loads, q_{iv} is sensible load from the infiltration of outside air, and

q_{sys} is the heat transfer to the HVAC system. The convection from the internal surfaces is obtained by summing the contributions from each surface:

$$q_{ci} = \sum_{i=1}^{n} h_{ci} A_i (T_a - T_{is}), \tag{4.15.41}$$

where n is the number of surfaces in the space, h_{cl} is the convection heat transfer for surface "i," A_i is the surface area "i" from which the heat transfer q_{ci} occurs, T_{is} is the temperature of surface "i," and T_a is the inside air temperature. The convection heat transfer from the internal loads q_{ce} is simply added to the heat balance on the air in the space.

This is not strictly correct because the temperatures of these surfaces that produce the internal loads actually add the heat to the air by convection. Unfortunately, this is very difficult to incorporate in the heat balance calculation because the detailed information necessary is usually not obtainable. With the air heat balance, any air that enters the space from the outside is assumed to mix instantaneously with the inside air. The ASHRAE *Handbook of Fundamentals* (2017) presents the latest information on the methods available for estimating the air infiltration rate. Sometimes, the infiltration rate is converted to a parameter equal to the number of air changes per hour (ACH) which would be included in the air heat balance applying the outside air temperature for that time increment (McQuiston et al., 2005).

Heating Load

Air conditioning of a space usually includes the design of a heating system. The heating system provides thermal energy that is then delivered to the space, maintaining the desired space temperature. Additionally, moisture may be added to maintain the relative humidity within acceptable limits.

The actual design of the heating system is dependent on the largest expected heat loss from the space to be heated. If the space is comprised of multiple spaces, each separately controlled by a thermostat, the heat loss is the sum of the largest individual heat losses for each space. Contributions to the heat loss include the following:

- Heat transfer through walls, floor, ceiling, glass, and all other surfaces
- Heat needed to warm the air from the outside entering the heated space
- Heat required to warm or thaw materials brought into the space

The outdoor temperature, wind velocity, and sunlight vary with time which causes the heat loss rate from the space to also vary with time. However, when determining the heat loss rate, or heat load, the outside temperature is selected as a design temperature originating from climatic data for that location. During winter, the coldest days are typically those that involve cloudy and stormy weather where the outdoor temperature has been noted to remain fairly constant. As a consequence, the heat loss rate from the space is also nearly constant. This is especially true when the internal heat gains are small. For this reason, the design heat load is determined on the basis of steady-state heat transfer and one-design outdoor temperature.

The heating system, as regulated by the thermostat, provides only the quantity of heat required to replace the heat that is lost. As the weather conditions change, the amount of heat supplied will vary. The heating system is not designed to withstand the most severe weather conditions ever recorded. To do so would result in a heating system that would rarely operate at full capacity. Rather, it is considered acceptable to not meet the inside temperature requirement for short periods of time. Consequently, the design outdoor temperature and the thermal characteristics of the zone govern the heat load for the space. A related activity is to estimate the energy consumed by the heating system over a year. To do this, simulation studies are made using hourly weather conditions where the heat

loss is then estimated on an hourly basis. The sum of all the hourly heat losses over the heating season is the annual energy consumption needed to heat the space.

Many thermostats incorporate a temperature setback feature to reduce the energy consumption of the heating system. The temperature is lowered several degrees at night and raised during the day when the space is occupied. It turns out that the early morning hours is the time when the heat load is the largest. Additionally, the morning time is typically when the thermostat calls for increasing the temperature. As a consequence, the heating system may need considerable time to meet the indoor temperature demand, during which time the occupants may experience discomfort. The time interval needed to warm the space after thermostat setback may be estimated using the heat balance method (Spitler, 2014).

The determination of the heat load assumes that no solar energy enters the space, internal heat gains are negligible, and the outdoor and indoor temperatures are constant. The calculation of the design heat load from a space is as follows:

1. Choose outdoor design conditions: temperature, humidity, wind direction, and wind speed.
2. Choose indoor design conditions.
3. Determine temperatures in any adjacent unheated spaces, if any.
4. Choose the heat transmission coefficients and calculate the heat loss rates for the walls, floors, ceilings, glass, doors, and all other surfaces where heat is transferred to the outside.
5. Calculate the heat load resulting from the infiltration of air and all other outdoor air entering the space.
6. Sum all the contributions to the heat load from transmission and infiltration.

Outdoor Design Conditions

Heating design day dry- and wet-bulb temperatures along with wind speed for numerous geographical locations are compiled from recent climatic data by ASHRAE's *Handbook of Fundamentals* (2017). Winter climatic data are based on hourly weather observations over at least a 12-year period. Winter month conditions are reported on the basis of 99.6 and 99.0 annual percentiles. These annual percentiles represent, respectively, 35 and 88 h from a year of 8760 h. For example, Wichita, KS has a heating dry-bulb temperature equal to 7.4°F at the 99.6 percentile, which means that for 35 h out of the year the temperature is likely to be below 7.4°F. Other climatic data, if available for a particular condition, suggesting lower dry-bulb temperatures and/or higher wind velocities should be considered in selecting the design conditions.

Indoor Design Conditions

Generally, the design dry-bulb temperature is chosen from the lower portion of the acceptable range, therefore preventing oversizing of the heating system. A heating system operating at conditions less than the design load does so at less efficiency. For this reason, a heating design dry-bulb temperature equal to 70°F is usually chosen. The relative humidity level is typically set at a maximum value of 30% to provide, on the one hand, a healthy environment, while on the other hand assure moisture integrity of the building's envelope.

Heat Transmission Loss

Heat transfer through all the building envelope materials comprising the walls, ceiling, roof, floors, windows, and doors is sensible and is calculated using

$$q = UA(T_i - T_0), \tag{4.15.42}$$

where U is the overall heat transfer coefficient or U-factor (Btu/h ft^2 °F), A is the normal surface area through which the heat transfer passes (ft^2), T_i is the inside design temperature (°F), and T_0 is the outside design temperature (°F). The U and A values are generally different for walls, ceilings,

roofs, and windows and so separate calculations are needed for each. Heat transmission loss through floor slabs and below-grade walls is determined using

$$q = U_s P(T_i - T_0), \tag{4.15.43}$$

where U_s is the overall heat transfer coefficient based on the slab perimeter (Btu/h ft^2 °F) and P is the slab perimeter (ft). Approximate numerical values for U_s are given by ASHRAE (2017).

Infiltration

Every building envelope leaks inside air to the outside due to pressure differences between the inside and outside air. This results in a winter heat loss because the air that is lost is replaced by cold outside air. The load on the heating system is dependent on the amount of heat needed to warm the infiltrated outside air to the inside design temperature and on the moisture added to raise the humidity to the indoor design condition. The sensible heating load is determined by

$$q_s = c_p \dot{V} \rho (T_i - T_0), \tag{4.15.44}$$

where c_p is the specific heat of the air (Btu/lb °F), $\dot{V}$ is the volumetric flow rate of air entering the building (ft^3/h), and ρ is the air density at temperature T_0 (lb/ft^3). The volumetric air flow rate entering the building depends on the wind speed and direction and on the size, among other factors, of the openings through which the infiltration air passes. The ASHRAE *Handbook of Fundamentals* (2017) carefully explains the determination of the volumetric flow due to infiltration. The latent heat loss due to infiltration is given by

$$q_l = \dot{V} \rho (W_i - W_o) h_{fg}, \tag{4.15.45}$$

where ρ is the air density at temperature T_i (lb/ft^3), W_i is the indoor air humidity ratio, W_o is the outdoor air humidity ratio, and h_{fg} is the latent heat of vapor at temperature T_i (Btu/lb).

Infiltration also occurs during summer cooling where warm moist air enters the building providing both sensible and latent loads. Calculation procedures are similar to those applied for winter infiltration.

Air Distribution System Heat Losses

When the ducts conveying warm air pass through spaces outside the conditioned space, heat losses will occur. The heat loss from supply air ducts contributes to the load for the space. While these losses may be reduced by insulating the duct, the heat loss cannot be completely eliminated. The heat loss through a duct may be estimated by

$$q = U A_s \Delta (T_a - T_u)_m, \tag{4.15.46}$$

where U is the overall heat transfer coefficient (Btu/h ft^2°F), A_s is the outside surface area of the duct (ft^2), $\Delta(T_a - T_u)_m$ is the mean temperature difference between the air in the duct, T_a and the temperature of the unconditioned space outside the duct, T_u. For a relatively short duct that is overlaid with 1 or 2 in. of insulation and covered with a reflective material, the mean temperature difference may be approximated by the simple difference between the supply air temperature and the unconditioned space temperature. In the HVAC industry, designers typically assume a duct heat loss equal to 2%–5% of the sensible load of the conditioned space.

Auxiliary Heat Gains

Within a conditioned space, numerous sources of heat may exist. Examples of common heat sources are lights, motors, machinery, and people. Usage patterns of equipment and people occupancy levels

influence the magnitude of the auxiliary heat gain rate. However, as a practical matter, auxiliary heat gains are ignored in those situations where the contribution to the total load is small. An exception is an industrial space where equipment, lighting, and occupancy may result in a significant heat gain. ASHRAE (2017) provides guidance in assessing the contributions from a number of auxiliary sources.

The ASHRAE *Handbook of Fundamentals* (2017) gives specific procedures for space heat load predictions for both residential and nonresidential buildings. The heat losses that occur through residential building slabs, crawl spaces, attics, and the like are well characterized. With nonresidential buildings, ASHRAE provides the latest information on auxiliary heat gains that may be helpful for an accurate prediction of the space heat load.

Refrigeration

The refrigeration system, sometimes called a heat pump, is the equipment in the HVAC system that provides the cooling for the indoor space. Refrigeration is applied to many fields all sharing the need for a source of low temperature. The application of refrigeration to air conditioning involves the removal of heat from the space, thereby maintaining the desired inside temperature. The refrigeration system accomplishes the absorption of heat from the conditioned space by maintaining a surface in the space at a temperature lower than that of the space. Two refrigeration system types are prevalent in air conditioning: the vapor compression system and the absorption system. The vapor compression system is available and popular in sizes applicable for single rooms to those appropriate for large commercial buildings composed of many individual zone spaces.

The capacity of a refrigeration system may be given in terms of kW or Btu/h. However, by tradition dating back to the early 1800s when ice was harvested for use during the summer in the United States, the term "ton of refrigeration" was used. It is still used today. A one-ton refrigeration plant has the capacity to produce one ton of ice per day (24 h). The latent heat of ice is 144 Btu/lb and so the capacity of a one-ton plant is equivalent to 12,000 Btu/h or 200 Btu/min.

A measure of a refrigeration system's performance is called the coefficient of performance or COP, defined by

$$\mathrm{COP} = \frac{\text{useful refrigerating rate}}{\text{net energy input rate}}. \tag{4.15.47}$$

Another measure is the refrigerating efficiency, η_R which is the ratio of the system's actual COP to that system's ideal COP,

$$\eta_R = \frac{\mathrm{COP}}{(\mathrm{COP})_{\mathrm{ideal}}}. \tag{4.15.48}$$

The operating power input of the refrigeration system relative to the refrigerating rate or "effect" is useful when comparing refrigeration equipment of different capacities. A common parameter is the horsepower per ton, given by

$$\mathrm{hp/ton} = \frac{4.715}{\mathrm{COP}}. \tag{4.15.49}$$

Carnot Refrigeration Cycle

The Carnot cycle is a theoretical cycle that is useful because no cycle can be constructed operating under the same temperature conditions with a greater operating efficiency. All the processes in the Carnot cycle are thermodynamically reversible (i.e., all the heat transfers occur at zero temperature difference, no friction occurs in any of the components). This cycle consists of two isothermal processes and two adiabatic processes. The Carnot refrigeration cycle appears as a rectangle

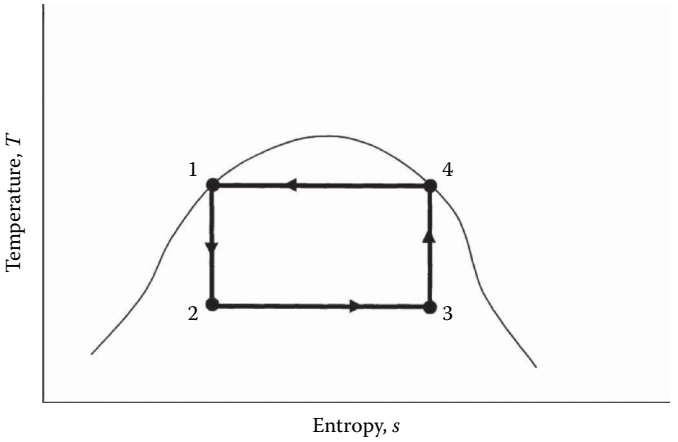


FIGURE 4.15.9 Temperature–entropy diagram for the Carnot refrigeration cycle.

on a temperature-entropy diagram shown in Figure 4.15.9. Observe that the cycle moves through the processes in a counterclockwise direction. Heat absorption occurs during the process occurring between points "2" and "3" and that heat rejection occurs between points "4" and "1." Heat absorption takes place in or adjacent to the space kept cool and heat rejection is made at a location outside the conditioned space. The expansion process identified as process "1" to "2" is adiabatic resulting in a temperature decrease. The compression process, process "3" to "4," is also adiabatic resulting in a temperature increase. Also, note that the expansion process extracts work partially offsetting the compression work. The COP for the Carnot cycle is

$$\text{COP} = \frac{T_c}{T_h - T_c}, \tag{4.15.50}$$

where T_c is the cold temperature and T_h is the hot temperature. The hot temperature is the temperature of the surrounding environment to which the heat is rejected.

Figure 4.15.9 also shows the saturated liquid and vapor lines for the working fluid used in the Carnot refrigeration cycle. The heat absorption and rejection processes represented by horizontal lines in the diagram may be achieved by working fluids undergoing a phase change. This cycle, however, is not practical for two important reasons. The compression process, from "3" to "4," occurs with liquid entering the suction side (inlet) of the compressor. The liquid, if allowed to enter the compressor, will cause very severe damage. The other reason involves the expansion process from "1" to "2." Here, the Carnot cycle requires that work be extracted from the working fluid as it expands, flashing into a mixture of vapor and liquid. One development capable of extracting work from the expansion of the refrigerant is an expander specially developed to operate in the wet expansion region is reported by Brasz (1995). Other approaches to improving the vapor compression cycle efficiency by energy recovery from the expansion process are summarized by Minh et al. (2006). The Carnot cycle remains important as a conceptual cycle in that it serves as a standard by which other cycles may be compared. No refrigeration cycle may have an efficiency greater than that of the Carnot cycle operating between the same cold and hot temperatures.

Vapor Compression Cycle

A practical refrigeration cycle is the vapor compression cycle which consists of the four components shown in Figure 4.15.10. The expander in the Carnot cycle, process "1" to "2," is replaced by an expansion device. Generally, an expansion device is a valve providing a restriction to the flow of

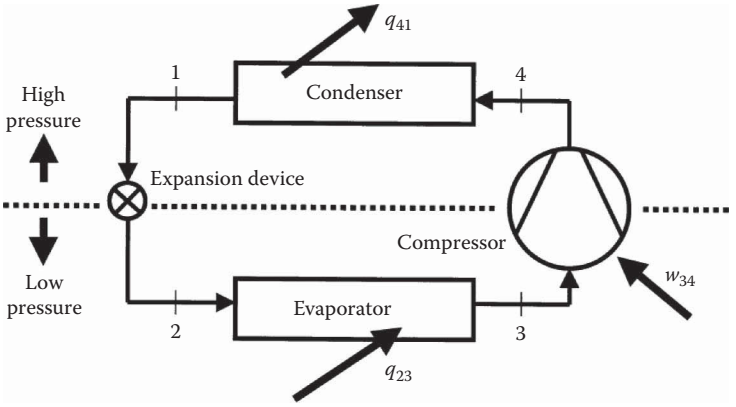


FIGURE 4.15.10 Components for the basic vapor compression refrigeration cycle.

liquid refrigerant. The compressor applied to the process from "3" to "4," operates at a rate sufficient to pump the low-pressure vapor from the evaporator to a high-pressure vapor capable of condensing in the condenser. Thus, the regulation of refrigerant flow to the evaporator by the expansion valve and the pumping action of the compressor involves a cycle with a "low-pressure side" and a "high-pressure side."

An ideal vapor compression cycle operates at the conditions revealed by the pressure-enthalpy diagram shown in Figure 4.15.11. The refrigerant leaving the condenser at state "1" is a saturated liquid. With the sudden pressure drop in the expansion valve, the refrigerant is a saturated mixture of liquid and vapor entering the evaporator. As the liquid refrigerant boils inside the evaporator, it absorbs heat. Only saturated vapor leaves the evaporator and enters the compressor. In this way, no liquid enters the compressor, thereby assuring long life. Ideal or isentropic compression of the refrigerant vapor takes place in the compressor, where the vapor is superheated at the discharge of the compressor. The saturation temperature corresponding to the discharge pressure must be greater than the temperature of the air or water to which the heat from the condenser is rejected. In the condenser, desuperheating occurs first; next, the vapor is condensed to a liquid. In an ideal condenser, no subcooling occurs and so the liquid leaves the condenser as saturated liquid at the condensing pressure. Furthermore, no pressure drop occurs in any of the piping transporting the refrigerant

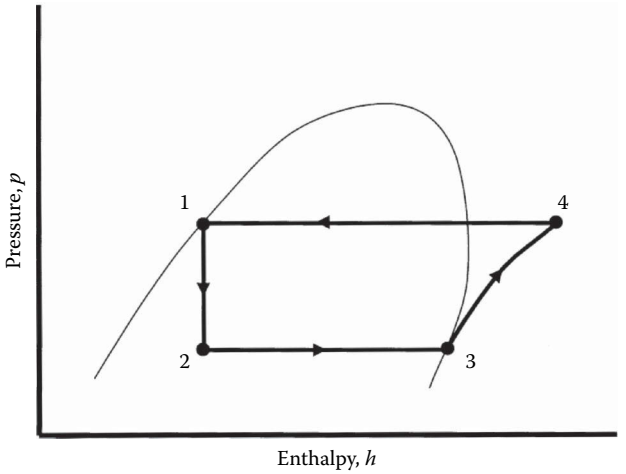


FIGURE 4.15.11 Pressure–enthalpy diagram for the vapor compression refrigeration cycle.

between the components nor in the evaporator and condenser. Only two pressures exist in an ideal vapor compression cycle: low and high pressure. The COP for the vapor compression cycle may be computed using the following relationships:

$$q_{23} = h_3 - h_2, \tag{4.15.51}$$

$$w_{34} = h_4 - h_3, \tag{4.15.52}$$

$$q_{41} = h_1 - h_4, \tag{4.15.53}$$

$$q_{23} + w_{34} + q_{41} = 0, \tag{4.15.54}$$

$$\text{COP} = \frac{h_3 - h_1}{h_4 - h_3}, \tag{4.15.55}$$

where q represents heat transfer, w the input of work to the compressor, and h the thermodynamic property enthalpy.

In reality, there is some pressure drop in the piping connecting the components and in the evaporator and condenser as well. Pressure drop in the piping represents a loss and a departure from the ideal cycle. Pressure drop in the evaporator causes a "glide" or change in the temperature of the boiling liquid as it progresses through the evaporator. Typically, depending on the particular refrigerant in use, the glide may be several degrees Fahrenheit. A similar situation would occur in the condenser. The design of vapor compression refrigeration cycles requires the estimation of the pressure losses in the piping. Information on these pressure drops is given in the ASHRAE *Handbook of Refrigeration* (2014). For the most part, single-component refrigerants (i.e., R-134a) do not exhibit a significant glide in conventional systems. However, refrigerants consisting of blends of chemicals may have larger glides in temperature and may require accommodation of this feature to accurately predict the performance of the cycle.

Estimation of an actual cycle's performance may be based upon consideration of an ideal cycle operating between the same temperatures but having an isentropic efficiency representing actual compressor performance (Stoecker, 1998). Actual compressor isentropic efficiencies typically fall in the range from 70% to 80%. However, the subcooling that may occur in the condenser, the superheat at the exit of the evaporator, and the pressure drops in the piping may be readily accounted for in cycle performance calculations when thermodynamic tables of the refrigerant are available. The ASHRAE *Handbook of Fundamentals* (2017) and NIST (2016) are sources of thermodynamic property data for a number of refrigerants.

Compressors

The compressor is one of the four basic components of the vapor compression cycle. The compressor has two important functions in the cycle: to increase the pressure of the refrigerant vapor from the evaporator so that it will condense, and to maintain the low-side pressure. The evaporator and condenser are heat exchangers facilitating the heat transfer in conjunction with a phase change of the refrigerant. The performance and design of heat exchangers are covered in detail on the heat transfer sections of this handbook.

Compressors fall into two main categories based on their operating principles. These are the positive displacement compressor and the aerodynamic, or more simply, the dynamic compressor. With the positive displacement compressor, the vapor enters the suction port and is subsequently geometrically trapped. The decreasing volume of the trapped vapor due to the expenditure of work causes the compression process to occur. In the dynamic compressor, the vapor enters the suction port whereupon the pressure of the vapor increases as a result of the high rotating velocity which produces a large centrifugal force.

Reciprocating Compressors

The reciprocating compressor is a common positive displacement compressor in the refrigeration industry manufactured in power input capacities from one-tenth of a horsepower to over several hundred horsepower. Compression is achieved by a reciprocating piston inside a cylinder along with the alternate opening and closing of suction and discharge valves. The smaller-capacity compressors are generally single-cylinder units and the large machines use multiple cylinder arrangements where the in-line, *V*, *W*, and radial configurations are available. Figure 4.15.12 is a diagram of a typical single-cylinder reciprocating compressor.

When the compressor is driven by an electric motor, several configurations are available. One is the "open-type compressor" where the crankshaft protrudes out from the compressor's housing. The motor is externally coupled to the compressor shaft. Another configuration is the "hermetically sealed compressor" where the compressor is inside a sealed container. The hermetic compressor is prevalent in small electrically driven units applied to residential air conditioners and smaller commercial units. The hermetic compressor has only the suction line, the discharge line, and the electrical power cable passing through the container's shell or "can." The absence of moving parts provides for a long-term and reliable seal. Another configuration is the "semihermetic compressor," which is identical to the hermetic compressor except that the cylinder heads are accessible for maintenance.

Performance characteristics of the reciprocating compressor include the following:

- *Pressure ratio*: discharge pressure divided by the suction pressure (both absolute pressures)
- *Displacement*: volume swept out by the piston as it moves over the extent of its travel in the cylinder, V_d
- *Displacement rate*: the rate of volume swept by the piston as the compressor operates, $\dot{V}_d$
- *Actual volumetric efficiency*: the actual refrigerant vapor volume flow rate at the suction port divided by the compressor's displacement rate, $\dot{V}_{actual} / \dot{V}_d$
- *Isentropic efficiency*: ratio of the isentropic compression work to the actual compression work

$$\eta_c = \frac{h_{4s} - h_3}{h_4 - h_3}, \tag{4.15.56}$$

where h is the enthalpy and subscripts 3 and 4 correspond to the suction and discharge states shown in Figure 4.15.10.

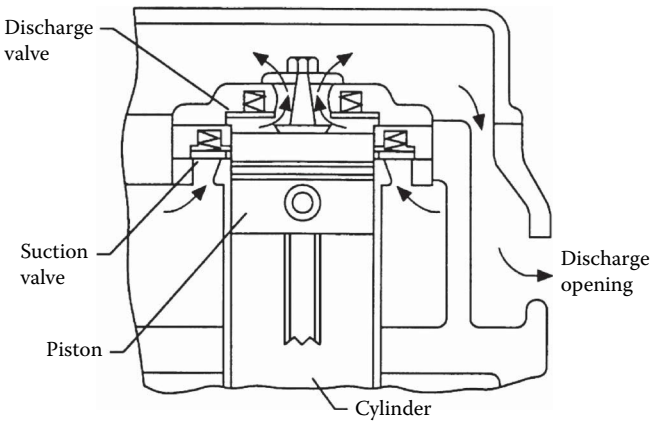


FIGURE 4.15.12 Typical reciprocating compressor cross section. (From Kreith, F., Ed., *The CRC Handbook of Mechanical Engineering,* CRC Press LLC, Boca Raton, FL, 1998. With permission.)

The isentropic efficiency includes the influence of pressure losses across the suction and discharge valves and the heat transfer to the refrigerant vapor from the compressor's internal surfaces. Friction originating from the rubbing of the mechanical parts (e.g., piston rings against the cylinder wall) is also included in the isentropic efficiency. However, leakage is not accounted for in the isentropic efficiency.

Accessories and subcomponents for the reciprocating compressor include a form of capacity control to meet varying refrigerating loads. Several methods are used to achieve capacity control in reciprocating compressors:

- Cycling the compressor on and off to satisfy the required capacity, usually applied to small systems.
- Throttling the suction line, good control but inefficient.
- Returning a portion of the discharge vapor to the suction line, good control, but inefficient and the compressor may operate hot.
- Multicylinder compressors may deactivate cylinders, opening the suction valve and closing the discharge valve, called "cylinder unloading."
- Changing compressor operating speed.

Small compressors are lubricated and cooled with a simple oil splash system. Lubrication oil that is miscible in the refrigerant must be conveyed by the refrigerant throughout the entire cycle. The halocarbon refrigerants use miscible oils and attention must be directed at the sizing of all components and piping to ensure that the velocities are sufficiently high everywhere to prevent the oil from accumulating anywhere in the system.

Scroll Compressor

The newest arrival among positive displacement compressors, it has a widespread application in residential, commercial, and automotive applications. The range of input power that the scroll compressor can accommodate is from about 1 to 60 hp. Refer to Section 4.7 on compressors for details of the scroll compressor.

Capacity control is available using variable-speed motor drives and porting holes in the fixed scroll member to adjust the displacement of the compressor. A control device sequentially opening the ports and venting the refrigerant vapor to the suction port progressively reduces capacity. Capacity control may also be achieved using suction throttling. The isentropic efficiency of the scroll compressor is comparable to the efficiencies of the other compressor types, including the reciprocating compressor. The scroll compressor feature limiting capacity is the size of the scroll. Currently, the largest refrigerating capacity available in a scroll compressor is 180 tons (Emerson, 2013).

Centrifugal Compressor

The centrifugal compressor accomplishes compression by imparting a high velocity to the refrigerant vapor and converting this velocity, actually kinetic energy, to a pressure. The compressor consists of an impeller driven to a high rotating speed. The vapor enters the suction port of the compressor through a set of guide vanes that direct the flow to the center of the impeller. The blades on the impeller are curved in such a way to direct the vapor to flow outward, by centrifugal force, at ever-increasing speeds. Leaving the impeller, the vapor flow enters the diffuser, where a portion of the kinetic energy gained is converted to pressure. Figure 4.15.13 shows a cross-sectional view of a two-stage centrifugal compressor. A two-stage centrifugal compressor utilizes two impellers to perform the compression.

The performance of the centrifugal compressor is obtained by equating the power applied to the impeller to the isentropic work of compression. After rearrangement and simplification (Stoecker, 1982),

$$V_t = \sqrt{g_c J(\Delta h_s)}, \tag{4.15.57}$$

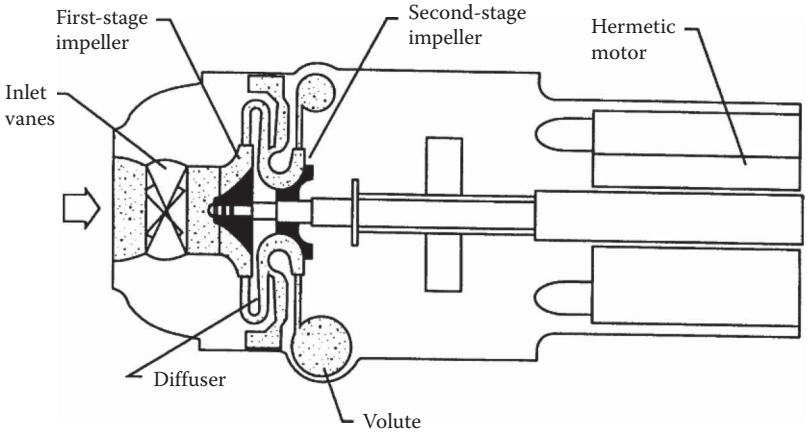


FIGURE 4.15.13 Cross section view of a centrifugal compressor. (From Kreith, F., Ed., *The CRC Handbook of Mechanical Engineering,* CRC Press LLC, Boca Raton, FL, 1998. With permission.)

where V_t is the tangential velocity (ft/s), g_c is a constant equal to 32.17 lb-ft/lbf-s^2, J is a constant equal to 778 ft lbf/Btu, and Δh_s is the enthalpy change for the isentropic work. This expression, while not accurate, does give reasonable approximations. The above expression shows that the refrigerant selected for a centrifugal compressor affects the diameter of the impeller. It turns out that the tip speed of the impeller is limited to approximately 800 ft/s, the actual value depending on the material used for the impeller. Refrigerants with thermodynamic properties similar to those of R-123 are desirable because the diameter of the impeller is smaller. Typically, centrifugal compressors using R-123 require only one compression stage and still have an acceptable tip speed. However, R-123 is not generally used in new equipment due to its toxicity.

A common application of the centrifugal compressor is the chilling of water for a central air-conditioning system. The refrigerating capacity where this compressor is applied ranges from about 50 to 2500 tons. Isentropic efficiencies for centrifugal compressors are usually in the range from 70% to 80%. Capacity control is accomplished by partially closing the inlet guide vanes upstream from the first impeller or varying the rotating speed of the impeller. Other methods of capacity control, such as bypassing the discharge gas, are not used, because they are not efficient.

Rotary Screw Compressor

This compressor is a positive displacement compressor where the compression process occurs in conjunction with the rotation of one or two screws. Details concerning the construction and operation of screw compressors are given in the compressor section of this handbook.

Performance of the screw compressor may be given in terms of the compression volume ratio, V_i, which is defined by

$$V_i = \frac{\text{internal volume when suction port covered}}{\text{internal volume when discharge port uncovered}}, \tag{4.15.58}$$

which gives

$$\text{Pressure ratio} = \frac{p_{dis}}{p_{suc}} = V_i^k, \tag{4.15.59}$$

where k is the ratio of specific heats c_p/c_v. The above equation is only valid when the compression is isentropic. When the compression process is not isentropic, the exponent k is replaced with exponent n which is not equal to k. When the temperature of the vapor remains constant during the compression, the value of n is 1. For actual rotary screw compressors, the value of n lies between k and 1 (Stoecker, 1998).

The discharge pressure for a rotary screw compressor may not match the pressure in the discharge piping. When these pressures are not equal, the adjustment that the vapor undergoes, over- or under-compression, results in losses. Variation of the compressor's volume ratio is possible through the addition of a slide valve that changes the extent of compression that occurs by changing the size of the discharge port. With a variable volume ratio, the over- and under-compression losses are eliminated. Capacity control is achieved by the addition of another slide valve that allows a portion of the suction vapor to bypass back to suction. Oil injection is used to seal the small spaces (approximately 0.0005 in. tolerance) between the lobes and gullies on the screws. Because the oil is heated during compression, it must be cooled to maintain its lubricating properties.

The isentropic efficiency of the screw compressor is about the same as that of the reciprocating compressor. Screw compressors of the same capacity as reciprocating compressors are smaller in size due to their higher rotating speeds. Refrigerating capacities are available from 20 tons to over 1800 tons and are prevalent in the industrial refrigeration field. Recently, screw compressors have been applied to large water-chilling units for air-conditioning applications in commercial buildings.

Absorption Cycle

The absorption refrigeration cycle differs from the vapor compression cycle in that the mechanical work used to compress the vapor is not needed. Rather, the vapor compression process is replaced by a series of processes whereby the refrigerant vapor is first absorbed into a liquid solution, the strong liquid solution is pumped to a higher pressure, and then driven from the generator as a vapor. To accomplish the absorption of the refrigerant and its subsequent release, heat is transferred out of the solution and transferred into the solution. The liquid pump, while requiring the input of work, requires only a small fraction of the work necessary to compress the vapor leaving the evaporator. The diagram in Figure 4.15.14 shows an ideal absorption refrigeration cycle. Observe that three of

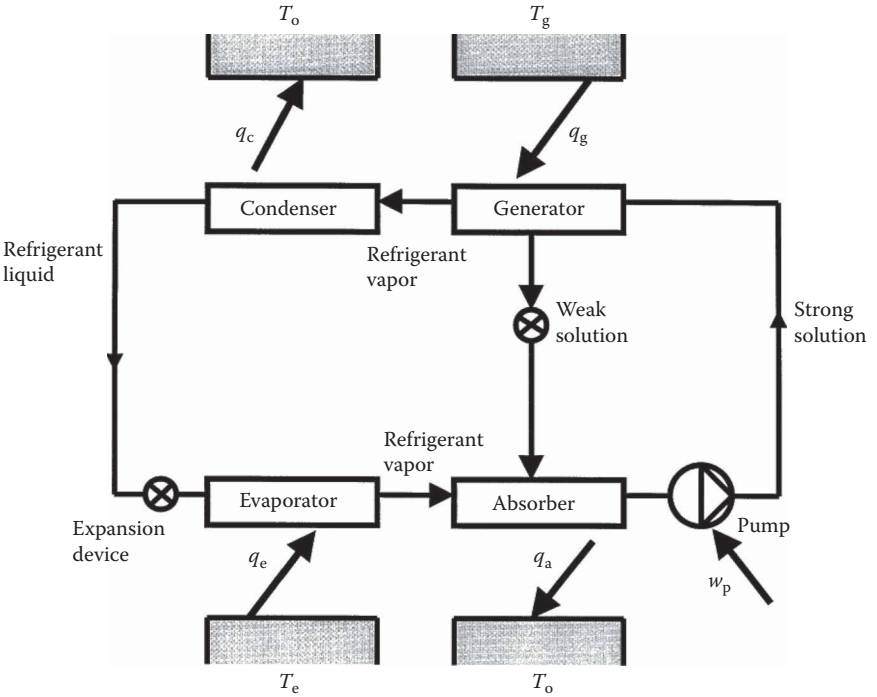


FIGURE 4.15.14 Ideal absorption refrigeration cycle.

the vapor compression cycle's components are also present in the absorption cycle. However, the compressor is replaced by the absorber, liquid pump, and generator which constitutes the path that the liquid solution takes as it circulates through in the cycle.

The working fluid in the absorption cycle is a binary fluid which is a homogeneous mixture of two chemicals. The performance of the absorption cycle depends on the affinity that the two components of the binary mixture have for each other. The pair commonly used for chilling water for air-conditioning applications is lithium bromide and water (Kuehn et al., 1998). In this binary mixture, the refrigerant is water which limits the cycle to temperatures above 32°F. Ammonia–water is another binary pair used in absorption cycles where ammonia is the refrigerant. The ammonia–water pair is typically used in industrial applications where temperatures lower than 32°F must be achieved (Bogart, 1981).

Absorption cycles provide an advantage when waste heat is available. Sources of waste heat may include heat rejected by an electrical power-generating plant, a large internal combustion engine, or an array of solar collectors. Even though the overall performance of the absorption cycle may be significantly less than that of the vapor compression cycle, when waste heat is available it may be competitive on a cost basis. Bosnjakovic (1965) developed the following expression for the maximum COP possible from the ideal absorption cycle shown in Figure 4.15.14:

$$(\mathrm{COP})_{\max} = \frac{T_e(T_g - T_o)}{T_g(T_o - T_c)}. \tag{4.15.60}$$

Refrigerant Selection

Vapor-compression refrigeration cycles used for the air conditioning of commercial buildings, private homes, and automobiles for human comfort have become common in many parts of the world. One factor promoting the widespread application of the vapor-compression cycle was the development of refrigerants composed of halocarbons, all of which are nontoxic and nonflammable. Halocarbons are chemical compounds containing one or more atoms from the halogen family, namely, chlorine, fluorine, and bromine, in conjunction with at least one carbon atom. The technical term for these compounds is chlorofluorocarbons, generally abbreviated to CFCs. This group of refrigerants were called Freon's, soon after their development in the 1930s. Early members of this new group were R-12 and R-11. They have gained prevalent use serving the refrigeration and air-conditioning industry for over 50 years.

During the 1960s, measurements of very small quantities of R-11 were obtained in a number of remote locations on the Earth, including at sea. Questions were raised by many as to whether or not CFC refrigerants present in the atmosphere have an influence on it. The answer became clear as a result of investigations conducted in the 1970s and 1980s. The survival time of a CFC in the atmosphere is very long as a result of its high chemical stability. This permits the CFC molecules to rise to the upper atmosphere and move through the ozone (O_3) layer. In this region of the atmosphere, ultraviolet radiation from the sun is present, which causes the chlorine atom to break free and chemically react with the ozone decreasing its concentration. The chlorine is then free to react again with another ozone molecule, thereby destroying it. This reaction continues for the life of the chlorine atom in the upper atmosphere.

The measurement of significant ozone depletion over Antarctica observed during the spring season of the year and the mounting evidence indicating significant ozone depletion over North America, Europe, and Japan caused a number of countries to make an agreement in Montreal, Canada through the United Nations. This agreement, created in 1987, is called the "Montreal Protocol on Substances that Deplete the Ozone Layer," and has been revised several times since (United Nations, 1994; UNEP, 1995). As a result of this agreement, CFC production has completely ceased. While old systems are still operating with CFC's, new systems are not allowed.

The Montreal Protocol also includes agreements that will result in the future phase-out of another group of refrigerants—the hydrogenated halocarbons referred to as hydrochlorofluorocarbons (HCFCs). A prominent refrigerant in this group is R-22, a refrigerant widely used in air-conditioning systems. R-22 has a modest ozone depletion potential (ODP), but nonetheless, its phase-out has been completed by developed countries. Developing countries are still using R-22 with phase-out having started in 2015 with a 10% reduction per year. R-123, an HCFC, is scheduled to cease its use in new equipment in 2020 and to cease all production in 2030. Another halocarbon class, the hydrofluorocarbons (HFCs), does not contain chlorine and, therefore, is not ozone depleting. HFCs are not regulated by the Montreal Protocol. However, their future is uncertain due to their contribution to global warming.

As a consequence of the Montreal Protocol, R-11 and R-12, which are both CFCs, cannot be used in new refrigeration equipment. In response, the refrigeration industry has developed alternative refrigerants and is conducting research that will lead to additional alternative refrigerants. The selection of suitable refrigerants and the attendant concerns with safety and the environment are decisions that the designer must make.

Selection of Refrigerants

When considering vapor compression refrigeration cycles, two overall system requirements must be considered. These are the temperature of the air-conditioned space and the temperature of the environment to which the heat must be rejected. Generally, the temperature of the boiling refrigerant in the evaporator must be about 20°F below that of the conditioned space. The pressure in the low-side of the system should be greater than atmospheric pressure, preventing air from entering the system due to leaks. Generally, the evaporator pressure is very near the saturation pressure which establishes the temperature of the refrigerant. Refrigerants R-11 and R-123 are exceptions in that evaporator pressures are typically less than 10 psia. The condenser presents an additional system constraint because at the condensing pressure which is very near the saturation pressure, the corresponding saturation temperature of the refrigerant must be about 20°F higher than the environment where the heat is rejected.

Table 4.15.1 lists several physical properties of refrigerants used in air conditioning. The critical pressure is given for the refrigerants. The condensing pressure in a conventional vapor compression refrigeration system must be somewhat less so that the refrigerant vapor may condense at a nearly constant pressure in the condenser. Table 4.15.2 is extracted from ASHRAE (2017) and shows the relative performance of some refrigerants when operating under the same conditions in ideal vapor compression cycles. From the COP values reported in Table 4.15.2, observe that most are about the same. R-410A is the lowest at 10.4 and R-123 is the highest at 11.4. Compressor discharge temperature is highest for ammonia and represents a disadvantage of ammonia particularly when evaporator temperatures are low and the compression ratio is high. However, R-123 should be excluded from the comparison as this refrigerant is uniquely suited to the centrifugal compressor due to their large specific volume resulting in reasonable wheel tip speeds. Otherwise, ammonia has the highest COP at 11.2, which is due to its high latent heat.

The environmental impact of some refrigerants is given in Table 4.15.3 where two characteristics have been selected: ODP and global warming potential (GWP). ODP is the relative capability of the refrigerant to destroy ozone in the atmosphere of the earth. The ODP for R-11 is defined as 1.00. GWP is the capability of the refrigerant to warm the atmosphere of the Earth relative to that caused by carbon dioxide over a period of 100 years. Chemicals in the atmosphere have a unique average residence time which in turn influences the GWP.

Table 4.15.3 indicates that CFCs have high ODP, HCFCs have significantly lower ODP, and HFCs have zero ODP. Halocarbon refrigerants should not be vented to the atmosphere. GWP values indicate that refrigerants, including the HFCs, represent an important segment of the chemicals that contribute to global warming due to the high values. In the future, restrictions on the use of refrigerants high in GWP may occur.

TABLE 4.15.1
Physical Properties of Refrigerants Used for Air Conditioning and Refrigeration (ASHRAE, 2017)

ASHRAE Designation	Name	Chemical Formula	Molecular Weight	Boiling Pt. (°F) at 14.7 psia	Freezing Pt. (°F)	Critical Temp. (°F)	Critical Press. (psia)
R-11 (CFC)	Trichloromonofluoromethane	CCl_3F	137.4	74.67	−166.8	388.3	639.3
R-12-(CFC)	Dichlorodifluoromethane	CCl_2F_2	120.91	−21.55	−25.69	233.5	599.9
R-22 (HCFC)	Chlorodifluoromethane	$CHClF_2$	86.468	−41.468	−251.356	723.7	32.7
R-32 (HFC)	Difluoromethane	CH_2F_2	52.024	−60.972	172.589	838.6	26.5
R-123 (HCFC)	Dichlorotrifluoroethane	$CHCl_2CF_3$	152.93	82.08	−160.87	362.624	531.1
R-134a (HFC)	Tetrafluoroethane	CF_3CH_2F	102.03	−14.933	−153.94	213.908	588.8
R-143a (HFC)	Trifluoroethane	CH_3CF_3	84.041	−53.034	−169.258	162.8726	545.5
R-152a (HFC)	Difluoroethane	CHF_2CH_3	66.051	−11.241	−181.462	235.868	655.1
R-404A (Blend)	R-125/143a/134a(44/52/4)		97.604	−51.200		161.6828	540.8
R-407C (Blend)	R-32/125/134a(23/25/52)		86.204	−46.529		186.8612	671.5
R-410A (Blend)	R-32/125(50/50)		72.585	−60.597		160.4444	711.1
R-507A (Azeotrope)	R-125/143a(50/50)		98.859	−52.134		159.1106	537.4
R-170 (Hydrocarbon)	Ethane	C_2H_6	30.07	−127.476	−297.01	89.924	706.6
R-290 (Hydrocarbon)	Propane	C_3H_8	44.096	−43.805	−305.72	206.13	616.6
R-717 (Inorganic)	Ammonia	NH_3	17.03	−27.989	−107.779	270.05	1643.7
R-718 (Inorganic)	Water	H_2O	18.015	211.953	32.018	705.11	3200.1
R-729 (Inorganic)	Air		28.959	−317.65		−221.062	549.6
R-744 (Inorganic)	Carbon dioxide	CO_2	44.01	−109.12[a]	−69.8044[b]	87.7604	1070.0

[a] Sublimes.
[b] At 76.4 psia.

TABLE 4.15.2
Performance Characteristics of Several Refrigerants Used for Air Conditioning per Ton of Refrigeration (ASHRAE, 2017)[a]

Refrigerant	Evap. Press. (psia)	Cond. Press. (psia)	Compression Ratio	Net Refrig Effect (Btu/lb)	Mass Flow Rate (lb/min per ton)	Suction Vapor Specific Volume (ft^3/lb)	Compressor Displ. (gal/min)	Power Consump. (hp)	Compr. Disch. Temp. (^oF)	COP
R-12	56.3	107.9	1.92	54.6	3.67	0.719	2.64	0.429	91.6	11.004
R-22	90.8	172.9	1.9	73.5	2.72	0.604	1.64	0.433	104.5	10.885
R-32	147.7	279.6	1.89	112.2	1.78	0.577	1.03	0.445	116.4	10.602
R-123	6.5	15.9	2.44	66.9	2.99	5.3	15.85	0.414	86.0	11.397
R-134a	54.7	111.7	2.04	69.2	2.89	0.868	2.51	0.433	90.6	10.903
R-407C	92.8	183.7	1.98	74.7	2.68	0.588	1.57	0.443	102.7	10.655
R-410A	145.0	273.6	1.89	75.2	2.66	0.416	1.11	0.455	103.7	10.379
R-290	85.3	156.5	1.84	130.7	1.53	1.26	1.92	0.439	90.7	10.743
R-717	81.0	169.3	2.09	484.9	0.41	3.61	1.49	0.421	137.4	11.186

[a] 45^oF evaporating temperature, 86^oF condensing temperature.

TABLE 4.15.3
Flammability and Environmental Characteristics of Several Refrigerants (NIOSH, 2016, ANSI/ASHRAE Std. 34, 2010)

Refrigerant	Chemical Formula	ODP	GWP (CO2 at 100 yr)	Atmos. Life (yr)	TLV (ppm)[a]	LFL (%)	ASHRAE Safety Group
R-11	Trichloromonofluoromethane	1	4,750	45	1,000	None	A1
R-12	Dichlorodifluoromethane	1	10,900	100	1,000	None	A1
R-22	Chlorodifluoromethane	0.055	1,810	12	1,000	None	A1
R-32	Difluoromethane	0	675	4.9	1,000	None	A2
R-123	Dichlorotrifluoroethane	0.02	77	1.3	50	None	B1
R-134a	Tetrafluoroethane	0	1,430	14	1,000	None	A1
R-143a	Trifluoroethane	0	4,470	52	1,000	None	A2
R-152a	Difluoroethane	0	124	1.4	1,000	None	A2
R-404A	R-125/143a/134a(44/52/4)	0	3,900	40.4	1,000	None	A1
R-407C	R-32/125/134a(23/25/52)	0	2,100	15.7	1,000	None	A1
R-410A	R-32/125(50/50)	0	2,100	17	1,000	None	A1
R-507A	R-125/143a(50/50)	0	4,000	40.5	1,000	None	A1
R-170	Ethane	0	~10	12±3	1,000	3.0	A3
R-290	Propane	0	~20	12±3	1,000	2.1	A3
R-717	Ammonia	0	<1	<0.02	25	14.8	B2L
R-744	Carbon Dioxide	0	1	~ 30,000	5,000	None	A1

[a] TLV, "threshold limit value" defined as personal exposure limit over a time-weighted average concentration for a normal eight-hour work day and a 40 hour work week.

ASHRAE Classification of Refrigerants

ASHRAE, through Standard 34–2013 "Designation and Safety Classification of Refrigerants," publishes refrigerant classifications as determined by the standard and is periodically updated to include new developments. The ASHRAE classification is based on two refrigerant characteristics: toxicity and flammability. Two levels, "A" (nontoxic) and "B" (toxic), comprise the toxicity classifications. A refrigerant is classified as toxic when exposed to humans at a concentration equal to 400 ppm (v/v) in air causes toxicity (Gopalnarayanan, 1998). Flammability is divided into three classes: highly flammable, moderately flammable, and nonflammable which are designated "3," "2," and "1," respectively. The term flammability refers to the capability of a mixture of the refrigerant vapor and air to propagate a flame. Table 4.15.3 includes the ASHRAE safety classification for several refrigerants.

The prefix "R" is used to designate a refrigerant. As mentioned earlier, letter designations (i.e., CFC) are also applied to halocarbon refrigerants. The numbering of halocarbon refrigerants and some hydrocarbons is dependent on the atoms that make up the molecule. The first digit on the right is the number of fluorine atoms. The second digit from the right is one more than the number of hydrogen atoms, and the third digit from the right is one less than the number of carbon atoms. When the third digit is zero, it is not written. With inorganic compounds serving as refrigerants, the first number on the left is "7." The next digits on the right are equal to the molecular weight of the compound. An example is R-744 (carbon dioxide). When the first digit is "4," the refrigerant is a blend (or zeotrope) of two or more compounds. The two digits to the right provide the constituents that comprise the mixture, and the letter the order in which the designation was made. When a "5" is designated, the refrigerant is a mixture of two compounds which form an azeotrope, which is a special mixture in that neither constituent can be recovered by distillation alone. The digits to the right of "5" indicate the historical order of that azeotrope's commercial introduction.

Table 4.15.3 also lists values for the refrigerant's vapor toxicity via the threshold limit value or "TLV." It is defined as the time-weighted average exposure that a worker can have to a substance for a normal 40-h work week without any adverse health effects.

Exergy Analysis

The ideal vapor compression refrigeration cycle has undergone exergy analyses (Kotas, 1995; Bejan, 1996, 2006; Li, 1996). These analyses are based on the second law of thermodynamics where the property exergy is used to track the losses that occur in the cycle. The second law efficiency for the cycle, defined in the thermodynamics section of this handbook, is relatively small. Calculations applied to idealized cycles show that the greatest exergy loss occurs in the condenser, the next largest loss being in the expansion device. The compressor, considered isentropic, has by far the smallest exergy loss. However, in an actual cycle, the exergy losses in the compressor, condenser, and evaporator are roughly the same at typical air-conditioning temperatures. The least exergy loss occurs in the expansion device in actual cycles (Kotas, 1995). Further considerations regarding the importance of exergy analysis are given in the discussions of Kotas (1995) and Bejan (2006).

Energy Distribution Systems

After the cooling and heating loads are determined for the building or application, the air distribution system must be designed resulting in the selection of equipment. The psychrometric processes discussed earlier play an important role in determining the magnitude of the sensible and latent loads for the space. For example, if winter heating is required, the air temperature and moisture content must be high enough to absorb the entire heating load for the space. The air is heated and humidified before being ducted to the occupied space where it is cooled and dehumidified, which are basic psychrometric processes. For commercial buildings, fresh outdoor air is generally required by national codes and local jurisdictions. To provide the fresh air, the return air is usually mixed with outdoor air before it enters the heating and humidifying equipment. For summer cooling, similar psychrometric processes occur where the air is conditioned by cooling and dehumidification. In both situations, the outdoor air, the return air, and the space load are all important in establishing the needed psychrometric processes.

All-Air Systems

Air-conditioning systems that psychrometrically prepare the air and transport that air through air ducts are called all-air systems. These systems are classified into two types by ASHRAE's *Handbook of HVAC Systems and Equipment* (2012): single-duct systems and dual-duct systems. Single-duct systems have one air flow path in which all the heating and cooling coils are inserted in series. These systems have a common air duct distribution system resulting in one temperature supplied to all the terminal outlets. In contrast, the dual-duct system has the main heating and cooling coils positioned in parallel. Two variations of the dual-duct system are used. One is to maintain the separation of the cool and warm air using two air duct distribution systems—one for the cool air and the other for the warm air. This is a traditional dual-duct system. The other system is to have a dedicated supply air duct for each zone with the mixing of the supply air to the needed conditions occurring at a main mixing damper array. This setup is called a multizone air duct system.

Single-Duct Systems

With single-duct systems, two basic categories characterize most systems: constant air volume and variable air volume (VAV).

Constant Volume

The supply air temperature changes in response to the space load while the air flow rate is maintained constant. The simplest version would be used for a single-zone system where the supply air equipment provides air directly or through air ducts to the space. For more complex systems, multiple-zone reheat is one version that accommodates the requirements of several spaces, each controlled by a separate thermostat. Additionally, multiple-zone systems can simultaneously provide heating and cooling to the zones as needed. Reheat of the air, either preconditioned primary air or recirculated air, occurs with these systems.

Small multiple-zone systems incorporate only low air pressure drops in the ducts and locate the reheat coils in the ducts near each zone. When the system is more complex (i.e., many zones), the primary distribution air ducts have higher pressures to reduce size and cost (McQuiston et al., 2005). Pressure reduction equipment is installed to maintain constant air flow volume for each zone. These systems use conditioned air from centrally located equipment, usually at a constant cold temperature (ASHRAE, 2012). However, to reduce the quantity of reheat required and conserve energy, the cold temperature may be adjusted by the controls.

A modification to the constant volume reheat system incorporates a bypass box replacing the reheat coil. The change becomes a constant volume primary system with a VAV secondary system. To meet the space load, the amount of zone supply air is adjusted by dumping the excess air to the return and simply bypassing the zone altogether. Bypass systems are limited in application to small multiple-zone systems, typically when energy consumption is less important.

Variable Air Volume

The supply air temperature remains constant while the amount of air supplied adjusts to the changing requirements of the space by means of a VAV terminal device. In operation, the air supply temperature is fairly constant, causing the VAV system to be well suited to perimeter and interior zones (ASHRAE, 2012). An opportunity for energy savings exists when using VAV systems for perimeter zones where solar heating results in a reduction of the supply air flow rate to make temperature changes. VAV systems may use individual or common fans, individual or common zone temperature control, and may or may not include auxiliary heating equipment.

Humidity control with VAV systems is more difficult than with constant volume systems. Approaches sometimes taken are to set a lower limit on the air flow rate, provide reheat to prevent overcooling with reasonable relative humidity, and provide ventilation using outdoor air. A typical lower limit on the air flow rate is 50% of the original flow rate.

The VAV system typically uses reheat when the heating occurs at the terminal unit a small distance upstream from the space. This system establishes the following control hierarchy: first, reduce the air flow volume, and second, add heat at the reheat coil. Relative to constant volume reheat, VAV systems consume less energy by reducing the primary air that must be cooled and by reducing the secondary air that needs heating. The terminal unit is adjusted to operate at a predetermined throttling ratio which is arrived at by establishing the lowest air flow rate sufficient to:

- Meet the heating load
- Limit the relative humidity (superseded by independent ventilation system)
- Provide acceptable air movement
- Provide required ventilation air (superseded by independent ventilation system)

Variations in the terminal units include provision for the induction of air from the space and have the advantage of recovering a portion of the heat from the lights (ASHRAE, 2012). The terminal unit may also be fan-powered in either a series or parallel flow arrangement. The advantage with fan-powered units is the larger circulation rate that is possible through the space at lower loads.

Dual-Duct Systems

All the supply air is conditioned in centrally located equipment which in turn is conveyed to the individual spaces by two main parallel ducts. Cold air is conveyed in one duct and warm air in the other. In a dual-duct box near the space, the warm and cold air flows are mixed by a valve in proportion to satisfy the space load. An advantage of the dual-duct system is that no pipes are required to feed the reheat coils, avoiding the possibility of leaks into the conditioned space. However, energy consumption for the dual-duct system is greater than that of the VAV single-duct system. Dual-duct systems may be either constant volume or VAV.

Constant Volume

Single-fan, no reheat resembles a single-duct system except for the face-and-bypass damper at the cooling coil (ASHRAE, 2012). This damper bypasses a mixture of recirculated and outdoor air in response to changes in the heat load. A disadvantage of this configuration is the imprecise control temperature and humidity in the occupied space. An alternative is the single-fan reheat configuration where the single reheat coil is placed in the central unit.

Variable Air Volume

VAV systems utilizing dual-ducts mix the cold and warm air yielding different volume combinations (ASHRAE, 2012). One configuration employs a single fan capable of handling both the cold and hot ducts. Typically, the cold deck is controlled to hold a constant temperature while the temperature of the hot deck is allowed to vary depending on the outside conditions. The hot deck temperature is increased when the outside temperature is low and the humidity high so that the flow over the cold deck increases to improve dehumidification. Another version is the dual fan VAV dual-duct system where each supply fan is independently controlled. The cold deck temperature is maintained constant by a refrigeration unit in conjunction with providing minimum fresh air, or by an economizer when the outside temperature is less than the setpoint temperature for the cold deck. The hot deck may still obtain heat from the recirculated air or from the heating coil when the recirculated air is too cold.

Air-and-Water Systems

Air-and-water systems use both water and air to convey the energy to and from the conditioned space. The supplied air originates from a central room where the heating and cooling equipment is placed and is called "primary air." The supplied water is called "secondary water." This system, as with the systems described earlier, requires a terminal unit for each zone. The terminal unit may be a fan coil device or an induction unit. Sometimes, an ordinary air outlet in combination with a radiant panel may serve as a terminal unit (ASHRAE, 2012).

The purpose of the primary air for heating is to provide fresh air to comply with ventilation requirements. When needed, water may be added to the warm air to improve comfort. For cooling, the primary air is dehumidified in order to achieve comfortable temperature and humidity in the space. The amount of primary air supplied to each space or zone is dependent on the fresh air ventilation requirement and the sensible cooling load at design conditions. Thus, the space humidity is controlled by the primary air. The secondary water is transported to the terminal units to supply a coil with either warm or cold water. When properly operating during the cooling season, the coil is dry, which greatly reduces odors and bacterial growth, while at the same time extending the life of the coil. However, condensate drains are recommended for all air-and-water systems.

Secondary water is supplied using a pump and piping arrangement. At the central plant, the water is typically cooled by heat exchange with the central chilled water plant. Air-and-water systems are classified as two-pipe, three-pipe, or four-pipe systems. All these piping arrangements accomplish the goal of supplying chilled or warmed water to the coil.

- A two-pipe system consists of one supply pipe and one return pipe for either warm or cold water supply temperatures. The seasonal changeover of these systems is relatively difficult.
- A three-pipe system consists of one cold water supply, one warm water supply, and a common return pipe. The mixing of the warm and cold water at the terminal unit generates significant energy losses.
- A four-pipe system consists of a cold water supply and return and a warm water supply and return. This arrangement is relatively efficient and changeover between seasons is convenient.

All-Water Systems

Another system is the all-water system, which may also be employed to heat and cool a space depending on the outdoor conditions. The air in the space is heated and cooled by any combination of the three basic heat transfer mechanisms: conduction, convection, and radiation. Applications include fan-coil units, panels on the floor, walls, or ceiling, and baseboard radiation. For further information on this subject, consult the ASHRAE *Handbook of HVAC Systems and Equipment* (2012).

REFERENCES

Anon. 1995. Assessment of the UNEP Montreal Protocol on substances that deplete the ozone layer. *Refrigeration, Air Conditioning and Heat Pumps Technical Options Committee, Parties to the Montreal Protocol*, United Nations, New York.

ASHRAE. 1975. *Procedure for Determining Heating and Cooling Loads for Computerizing Energy Calculations, Algorithms for Building Heat Transfer Subroutines,* American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

ASHRAE. 2013. Designation and safety classification of refrigerants, *ANSI/ASHRAE Standard 34-2013*, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

ASHRAE. 2012. *Handbook of HVAC Systems and Equipment,* American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

ASHRAE. 2013. *Handbook of Fundamentals,* American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

ASHRAE. 2014. *Handbook of Refrigeration,* American Society of Heating, Refrigeration, and Air-Conditioning Engineers, Atlanta, GA.

ASHRAE. 2017. *Handbook of Fundamentals,* American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

ANSI/ASHRAE. 2010. *Standard 55. Thermal Environmental Conditions for Human Occupancy*, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

Berglund, L.G. 1995. Comfort criteria: Humidity and standards. *Proceedings of Pan Pacific Symposium on Building and Urban Environmental Conditioning in Asia*, Vol. 2, pp. 369–382. University of Nagoya, Japan.

Bejan, A. 1996. *Entropy Generation Minimization,* CRC Press, Boca Raton, FL.

Bejan, A. 2006. *Advanced Engineering Thermodynamics*, 4th Edition, John Wiley and Sons, New York.

Bogart, M. 1981. *Ammonia Absorption Refrigeration in Industrial Processes,* Gulf Publishing, Houston, TX.

Bosnjakovic, F. 1965. *Technical Thermodynamics,* Holt, Reinehart & Winston, New York.

Brasz, J.J. 1995. Improving the refrigeration cycle with turbo-expanders. *IIR Congress*, The Hague, the Netherlands.

Burch, D.M., and Chi, J. 1997. Moist: A PC program for predicting heat and moisture transfer in building envelopes. *Release 3.0, NIST Special Publication 917,* U.S. Government Printing Office, Washington, D.C.

Emerson Climate Technologies. 2013. *Emerson Climate Technologies Launching World's Largest Scroll Compressor,* St. Louis, MO. http://www.emersonclimate.com/en-us/about_us/news/news_releases/pages/emerson_climate_technologies_launching_worlds_largest_scroll_compressor.aspx.

Fanger, P.O. 1982. *Thermal Comfort*, Kreiger Publishing, Malabar, FL.

Fiala, D., Lomas, K., and Stohrer, M. 2003. First principles modeling of thermal sensation responses in steady state and transient boundary conditions. *ASHRAE Trans.* 109(Part 1), 179–186.

Gagge, A.P., Fobelets, A.P., and Berglund, L.G. 1986. A standard predictive index for human response to the thermal environment. *ASHRAE Trans.,* 92(1), 709–731.

Gagge, A.P., Stolwijk, J., and Nishi, Y. 1971. An effective temperature scale based on a simple model of human physiological regulatory response. *ASHRAE Trans.,* 77(1), 247–262.

Gopalnarayanan, S. 1998. Choosing the right refrigerant. *Mechanical Engineering Magazine,* Falcioni, J.G. Ed., American Society of Mechanical Engineers, New York, 120(10), 92–95.

Herrmann, S., Kretzschmar, H.-J., and Gatley, D. 2009. Thermodynamic properties of real moist air, dry air, steam, water, and ice (RP-1485), Rademacher, R. Ed., *HVAC&R Res.*, 15, (5) 961–986.

Howell, H.H., Sauer, H.J. Jr., and Coad, W.J. 1998. *Principles of Heating, Ventilating and Air Conditioning,* ASHRAE, Atlanta, GA.

Huizenga, C., Zhang, H., and Arens, E. 2001. A model of human physiology and comfort for assessing complex thermal environments. *Build. Environ.* 36(6), 691–699.

Hyland, R.W., and Wexler, A. 1983a. Formulations for the thermodynamic properties of the saturated phases of H_2O from 173.14 K to 473.15 K. *ASHRAE Trans.,* 89(2A), 500–519.

Hyland, R.W., and Wexler, A. 1983b. Formulations for the thermodynamic properties of dry air from 173.15 K to 372.15 K, and of saturated moist air from 173.15 K to 372.15 K, at pressures to 5 MPa. *ASHRAE Trans.,* 89(2A), 520–535.

Hosni, M.H., Sipes, J.M., and Wallis, M.H. 1998. Controlling moisture in walls exposed to hot and humid climates. Final Report, ASHRAE Res. Project 864-RP, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

Hwang, Y., Ohadi, M., and Radermacher, R. 1998. Natural refrigerants. *Mechanical Engineering Magazine,* Falcioni, J.G. Ed., American Society of Mechanical Engineers, New York, 120(10), 96–99.

Kotas, T.J. 1995. *The Exergy Method of Thermal Plant Analysis,* Krieger Publishing, Malabar, FL.

Kuehn, T.H., Ramsey, J.W., and Threlkeld, J.L. 1998. *Thermal Environmental Engineering,* 3rd Edition, Prentice-Hall, Englewood Cliffs, NJ.

Li, K.W. 1996. *Applied Thermodynamics: Availability Method and Energy Conversion,* Taylor & Francis, Washington, DC.

McGee, M. 2016. CO2 Earth. https://www.co2.earth/.

McQuiston, F.C., and Spitler, J.D. 1992. *Cooling and Heating Load Calculation Manual,* American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

McQuiston, F., Parker, J., and Spitler, J. 2005. *Heating, Ventilating, and Air Conditioning Analysis and Design,* 6th Edition, John Wiley and Sons, New York (ISBN 0-471-47015-5).

Minh, O., Hewitt, N., and Eames, P. 2006. *Improved Vapour Compression Refrigeration Cycles: Literature Review and Their Application to Heat Pumps,* University of Ulster, Belfast, Northern Ireland, United Kingdom.

NASA. 1976. U.S. Standard Atmosphere. National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration, and the U.S. Air Force, Superintendent of Documents, U.S. Government Printing Office, Washington, DC.

NIST. 2016. NIST Reference fluid thermodynamic and transport properties database (REFPROP Version 9.1. 2016. National Institute of Standards and Technology (NIST), Gaithersburg, MD. http://www.nist.gov/srd/nist23.cfm.

NOAA. 2016. Earth System Research Laboratory. http://www.nasa.gov/mission_pages/sunearth/science/mos-upper-atmosphere.html.

Pederson, C.O., Fisher, D.E., Spiltler, J.D., and Liesen, R.J. 1998. *Load Calculation Principles,* American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

Schlatter, T. W. 2009. Atmospheric composition and vertical structure, Pub. eae319MS-1. Earth System Research Laboratory, NOAA, U.S. Dept. of Commerce.

Spitler, J. D. 2009. *Load Calculation Applications Manual*, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

Spitler, J. D. 2014. *Load Calculation Applications Manual, 2nd Edition, IP Version*. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

Stoecker, W.F. 1982. *Refrigeration and Air Conditioning,* McGraw-Hill, New York.

Stoecker, W.F. 1998. *Industrial Refrigeration Handbook,* McGraw-Hill, New York.

United Nations. 1994. Report of the Refrigeration, Air Conditioning and Heat Pumps Technical Options Committee, Montreal Protocol on Substances that Deplete the Ozone Layer, United Nations Environmental Program, New York.

Wallace, J. M., and Hobbs, P. V. 2006. *Atmospheric Science; An Introductory Survey*, 2nd Edition, Elsevier, Amsterdam. ISBN 987-0-12-732951-2.

Zhang, H., Arens, E., and Han, T. 2010. Thermal sensation and comfort models for non-uniform and transient environments. Parts 1, 2, and 3. *Build. Environ.*, 45(2), 380–410.

4.16 HEAT TRANSFER ENHANCEMENT

Raj M. Manglik

Introduction

The imperatives of energy conservation, materials and manufacturing input savings, and in recent years the concerns for environmental conservation, along with associated economic incentives, have led the efforts to produce more efficient heat exchange equipment. The objectives for improving and optimizing the thermal and hydrodynamic performance are generally the following: reduce the size of a heat exchanger required for a specified heat duty, or increase or upgrade the capacity of an existing heat exchanger, or reduce the approach temperature difference (or the inlet temperature difference) for the process streams, and/or reduce the pumping (or blower or fan) power requirements for moving the flow stream.

Heat transfer *enhancement* [or *enhanced* heat transfer, or for that matter enhanced mass transfer (Bergles et al., 1983; Manglik and Bergles, 2004)] refers to the study of techniques for improving the thermal performance of a heat (or mass) exchange device or system. It is sometimes also referred to as heat transfer *augmentation* or *intensification*. In general, this entails an increase in the heat transfer coefficient and encapsulates the broader science and engineering of methods for producing higher convective heat/mass transfer coefficients, reducing frictional losses, and increasing the overall thermal–hydrodynamic efficiencies of exchangers. Attempts to increase "normal" heat transfer coefficients have been recorded for more than 150 years in modern history (Bergles and Manglik, 2013) and indeed date back several thousand years in antiquity (Manglik and Jog, 2009). As a result, there is a very large store of information that has been documented and disseminated in several periodic surveys (Manglik and Bergles, 2004; Bergles and Manglik, 2013; Manglik et al., 2013). The literature comprising of technical publications, excluding patents and manufacturers' literature, has expanded rapidly since 1955 and approximately 350–450 papers and reports on the subject are now published annually.

Methods for enhancement of heat and/or mass transfer can be classified either as *passive techniques*, which require no direct application of external power, or as *active techniques*, which require external power. A listing of the different methods and their classification is given in Table 4.16.1, and a few representations of passive methods in current usage are depicted in Figure 4.16.1. The effectiveness of both types of techniques is strongly dependent on the mode of heat transfer, which may range from single-phase free convection to dispersed-flow film boiling. Brief descriptions of these methods are as follows:

- *Treated surfaces* involve fine-scale alteration of the surface finish or coating (continuous or discontinuous). They are used for boiling and condensing applications; the roughness height is below that, which affects single-phase heat transfer.
- *Rough surfaces* are produced in many configurations ranging from random sand grain-type roughness to discrete protuberances (see Figure 4.16.1a). The configuration is generally chosen to disturb the viscous sublayer rather than to increase the heat transfer surface area. Application of rough surfaces is directed primarily toward single-phase flow.
- *Extended surfaces* are routinely employed in many heat exchangers (see Figure 4.16.1b). Work of special interest to enhancement is directed toward improvement of heat transfer coefficients on extended surfaces by geometrically altering the fin or extended surfaces.
- *Displaced enhancement devices* are inserted into the flow channel so as indirectly to improve energy transport at the heated surface. They are used with forced flow (see Figure 4.16.1c).

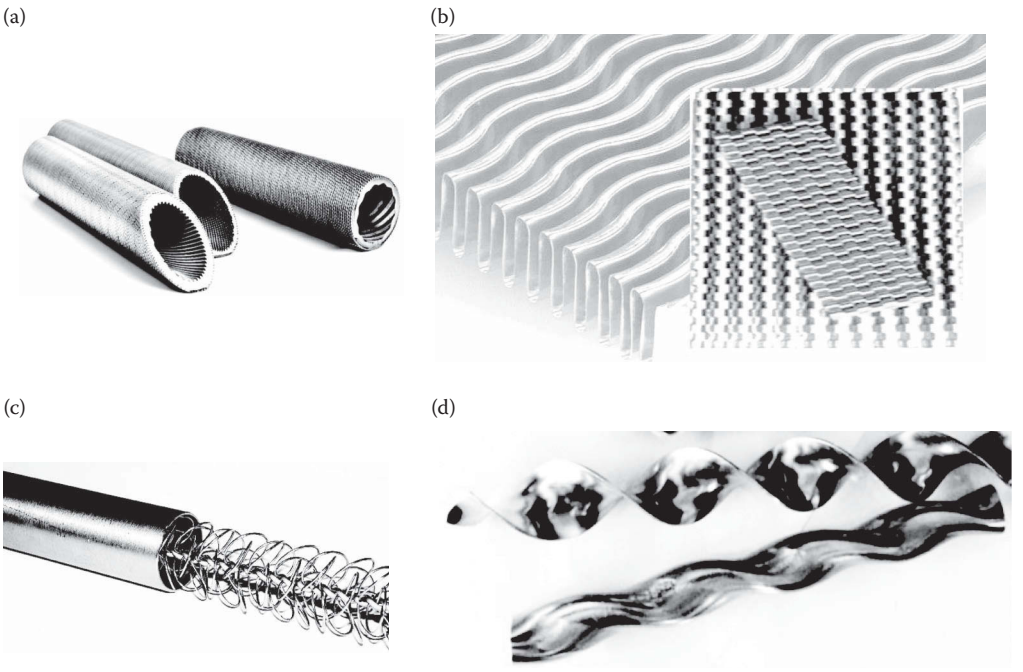


FIGURE 4.16.1 Enhanced surfaces that are typically used in forced convection heat transfer: (a) spirally ribbed and structured surface roughness tubes [photo courtesy Wieland-Werke AG], (b) wavy plate fins and offset-strip fins, (c) displaced enhancement tube insert, and (d) twisted-tape insert and twisted oval tube.

TABLE 4.16.1
Classification of Different Heat Transfer Enhancement Techniques

Passive Techniques	**Active Techniques**
Treated surfaces	Mechanical aids
Rough surfaces	Surface vibration
Extended surfaces	Fluid vibration
Displaced enhancement devices	Electrostatic fields
Swirl flow devices	Injection
Coiled tubes	Suction
Surface tension devices	Jet impingement
Additives for liquids	
Additives for gases	
Compound Enhancement	
Two or more passive and/or active techniques used together	
Examples:	
Rough-surface tube with twisted-tape insert (two passive techniques)	
Rotating internally finned tubes (an active and a passive technique)	

- *Swirl-flow devices* include a number of geometric arrangements or tube inserts for forced flow, which induce rotating and/or secondary flow and promote greater fluid mixing. Examples include axially twisted tubes (see Figure 4.16.1d), coiled tubes, inlet vortex generators, twisted-tape inserts (see Figure 4.16.1d), and axial-core inserts with a screw-type winding.

- *Surface-tension devices* consist of wicking or grooved surfaces to direct the flow of liquid in boiling and condensing.
- *Additives for liquids* include solid particles and gas bubbles in single-phase flows and liquid trace additives for boiling systems.
- *Additives for gases* are liquid droplets or solid particles, either dilute phase (gas–solid suspensions) or dense phase (fluidized beds).
- *Mechanical aids* involve stirring the fluid by mechanical means or by rotating the surface. Surface "scraping," widely used for batch processing of viscous liquids in the chemical process industry, is applied to the flow of such diverse fluids as high-viscosity plastics and air. Equipment with rotating heat exchanger ducts is found in commercial practice.
- *Surface vibration* at either low or high frequency has been used primarily to improve single-phase heat transfer.
- *Fluid vibration* is the practical type of vibration enhancement because of the mass of most heat exchangers. The vibrations range from pulsations of about 1 Hz to ultrasound. Single-phase fluids are of primary concern.
- *Electrostatic fields* (DC or AC) are applied in many different ways to dielectric fluids. Generally speaking, electrostatic fields can be directed to cause greater bulk mixing or fluid or disruption of fluid flow in the vicinity of the heat transfer surface, which enhances heat transfer.
- *Injection* is utilized by supplying gas to a stagnant or flowing liquid through a porous heat transfer surface or by injecting similar fluid upstream of the heat transfer section. Surface degassing of liquids can produce enhancement similar to gas injection. Only single-phase flow is of interest.
- *Suction* involves vapor removal, in nucleate or film boiling, or fluid withdrawal, in single-phase flow, through a porous heated surface.

Two or more of the aforementioned techniques may also be utilized simultaneously to produce an enhancement that is larger than either of the techniques operating separately. Such usage is referred to as *compound enhancement* (see Table 4.16.1 for an example).

Extended descriptions along with many other examples can be found in several surveys and monographs (Bergles, 1998; Manglik, 2003; Webb and Kim, 2005). It should be emphasized here that one motivation, among some others, for studying enhanced heat transfer is to assess the effect of an inherent condition on heat transfer. Some practical examples include roughness that is naturally produced by standard manufacturing, degassing of liquids with high gas content, surface vibration resulting from rotating machinery or flow oscillations, fluid vibration resulting from pumping pulsation, and electrical fields present in electrical equipment (e.g., electric transformers).

The enhanced surfaces and tubular inserts represented in Figure 4.16.1 have been used for both single-phase and two-phase (boiling or evaporation and condensation) heat transfer enhancement. The emphasis here is on effective and cost-competitive (either proven or potential) techniques that have made the transition from the laboratory to commercial heat exchangers; the effectiveness of a method is determined by the relative trade-offs between increased heat transfer coefficient and frictional penalty.

Single-Phase Free Convection

With the exception of the familiar technique of providing extended surfaces (or what are popularly known as finned surfaces), the passive techniques have little to offer in the way of enhanced heat transfer for free convection. This is because the buoyancy-driven flow velocities are usually too low to cause flow separation, or secondary flow, or other such significant flow perturbation that would enhance the convective heat transport.

Serrated, or perforated, or slotted, or similarly segmented or interrupted finned surfaces promote enhanced heat transfer. The periodic restarting of thermal boundary layers over the interrupted extended surfaces is the primary enhancement mechanism that substantially increases the heat transfer coefficient, and this more than compensates for the lost surface area. Pin fins, of either circular or rectangular cross section, and arrayed in staggered or in-line arrangements, are some other extended surface geometries that find wide usage with natural convection, particularly in electrical and electronic cooling (Kraus et al., 2001; Kreith et al., 2011).

Mechanically aided thermal processing, which may entail both heat and mass transfer, is a standard technique in the chemical and food industries when viscous liquids are involved. The predominant geometry for surface vibration has been the horizontal cylinder, vibrated either horizontally or vertically. Heat transfer coefficients can be increased 10-fold for both low-frequency/high-amplitude and high-frequency/low-amplitude situations. It is, of course, equally effective and more practical to provide steady forced flow. Furthermore, the mechanical designer is concerned that such intense vibrations could result in equipment failures.

Since it is usually difficult to apply surface vibrations to practical equipment, an alternative technique is utilized whereby vibrations are applied to the fluid and focused toward the heated surface. With proper transducer design, it is also possible to improve heat transfer to simple heaters immersed in gases or liquids by several hundred percent.

Electric fields are particularly effective in increasing heat transfer coefficients in free convection. Dielectrophoretic or electrophoretic (especially with ionization of gases) forces cause greater bulk mixing in the vicinity of the heat transfer surface. Heat transfer coefficients may be improved by as much as a factor of 40 with electrostatic fields up to 100 kV. Again, the equivalent effect could be produced at lower capital cost and without the voltage hazard by simply providing forced convection with a blower or fan.

Single-Phase Forced Convection

While there are many different geometric manifestations of heat exchangers used in industry (Shah and Sekulic, 2003; Wang et al., 2007), the present discussion emphasizes enhancement of heat transfer *inside* tubes and ducts that are primarily of circular cross section. Representative data for some of the extended surfaces (offset-strip fins and wavy-plate fins) (Kays and London, 1998) and tube with structured roughness (Ravigururajan and Bergles, 1996) are graphed in Figure 4.16.2, along with reference results for plain fins (or unembellished fins) and smooth tube. The smooth tube results are predictions of the following Nusselt number and Fanning friction factor correlations, respectively:

$$\mathrm{Nu}_s = \frac{0.5 f \mathrm{Re}_{D_i} \mathrm{Pr}}{1 + 12.7 \left(0.5 f\right)^{0.5} \left(\mathrm{Pr}^{0.667} - 1\right)} \tag{4.16.1}$$

and

$$f_s = \left(1.58 \ln \mathrm{Re}_{D_i} - 3.28\right)^{-2} . \tag{4.16.2}$$

The subscript *s* in these equations denotes a smooth or unembellished plain tube or surface, and the characteristic length scale for Nu and Re is the inner diameter D_i of the tube. As seen from these plots, while both enhancement devices produce a sizable elevation in the Colburn *j*-factor (or dimensionless heat transfer coefficient) at constant Reynolds number, or velocity, there is also an accompanying large increase in the friction factor. Balancing this penalty to obtain net thermal benefits thus becomes an optimization exercise, and for which several evaluation figures of merit have been devised (Bergles et al., 1974; Bergles, 1998; Manglik, 2003). A representative example is discussed in the subsequent subsection.

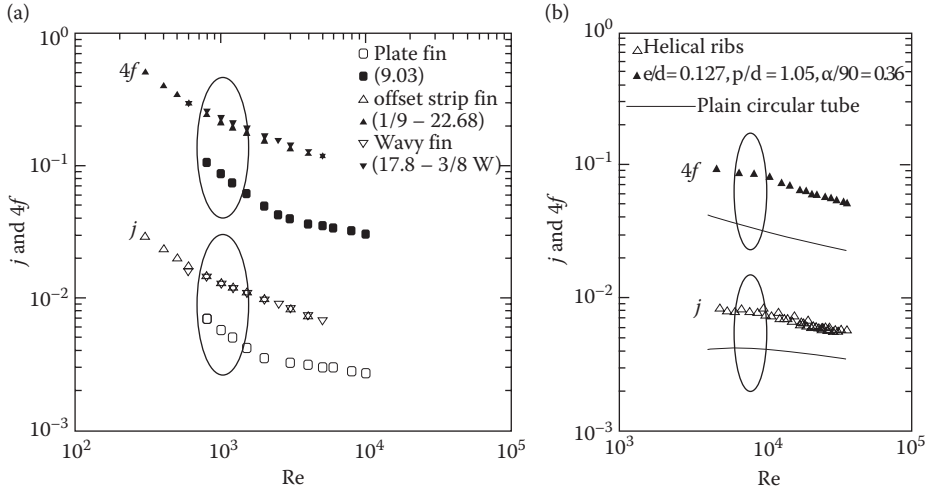


FIGURE 4.16.2 Typical data and results for heat transfer (Colburn *j*-factor) and Fanning friction factor *f* for passive enhancement techniques: (a) plate-fin geometries or extended surfaces (wavy fins and offset-strip fins; data from Kays and London, 1998) and (b) tubes with internal structured roughness (helically ribbed tubes; data from Ravigururajan and Bergles, 1996).

Surface roughness has been used extensively to enhance forced convection heat transfer. Integral or structured roughness may be produced by the traditional manufacturing processes of machining, forming, casting, or welding. Various inserts (spiral wire-coil inserts) can also provide surface protuberances. In view of the infinite number of possible geometric variations, it is not surprising that, even after decades of research and hundreds of reported studies, no completely satisfactory unified treatment of structured rough surfaces is available.

Notwithstanding the increased friction loss that requires balancing in the eventual thermal–hydrodynamic optimization of a heat exchanger (Bergles, 1998; Manglik, 2003), the improvements in heat transfer coefficient with turbulent flow in rough tubes (based on nominal surface area) are as much as 100% and more (Ravigururajan and Bergles, 1994; 1996). In efforts to provide predictive design tools, analogy solutions for sand grain-type roughness and for square-repeated-rib roughness have been proposed. A statistical correlation was also developed subsequently (Ravigururajan and Bergles, 1996), for both the heat transfer coefficient and Fanning friction factor, which was based on new experiments as well as a very large data set available in the literature. These predictive equations are recommended for tubes that have structured roughness made up of transverse- or helical-repeated ribs (the latter are shown Figure 4.16.1a) with turbulent flow and are as follows:

$$\mathrm{Nu}_{\mathrm{D_i},a} = \mathrm{Nu}_s \left[1 + \left\{ 2.64 \mathrm{Re}^{0.036} \left(\frac{e}{D_i} \right)^{0.212} \left(\frac{p}{D_i} \right)^{-0.21} \left(\frac{\alpha}{90} \right)^{0.29} \mathrm{Pr}^{-0.024} \right\}^7 \right]^{1/7}, \tag{4.16.3}$$

$$f_a = f_s \left[\left\{ 29.1 \mathrm{Re}_{\mathrm{D_i}}^{a1} \left(\frac{e}{D_i} \right)^{a2} \left(\frac{p}{D_i} \right)^{a3} \left(\frac{\alpha}{90} \right)^{a4} \left(1 + \frac{2.94 \sin \beta}{n} \right) \right\}^{15/16} \right]^{16/15}, \tag{4.16.4a}$$

where

$$a1 = \left[0.67 - 0.06\left(\frac{p}{D_i}\right) - 0.49\left(\frac{\alpha}{90}\right)\right]; \quad a2 = \left[1.37 - 0.157\left(\frac{p}{D_i}\right)\right],$$
$$a3 = \left[-1.66\times10^{-6}\,\mathrm{Re} - 0.33\left(\frac{\alpha}{90}\right)\right]; \quad a4 = \left[4.59 + 4.11\times10^{-6}\,\mathrm{Re} - 0.15\left(\frac{p}{D_i}\right)\right]. \tag{4.16.4b}$$

In these equations, the characteristic length scale is the smooth or empty inner diameter of the tube, D_i, and the subscript a refers to the enhanced tube, whereas the subscript s refers to the smooth tube. The other special symbols are given as follows: e = protuberance height; p = repeated-rib pitch; α = spiral angle in degrees for helical ribs; n = number of sharp corners facing the flow; and β = contact angle in degrees for the rib profile. Also, the smooth tube Nusselt number and Fanning friction factor are, respectively, given by Equations 4.16.1 and 4.16.2.

Much work has been done to obtain the enhanced heat transfer of parallel angled ribs in short rectangular channels, simulating the interior of gas turbine blades. A variety of shapes and rib patterns have been considered with the focus on improving internal blade cooling (Wright and Han, 2014). In several industrial materials processing applications, jets (gas or liquid) are frequently used for carrying out heating, cooling, and drying operations. For such cases, a number of studies have reported that roughness elements, particularly of the transverse-repeated-rib type, mitigate the deterioration in heat transfer downstream of the jet stagnation region.

In a recent new development (Champagne and Bergles, 2011), a novel technique to render the inside surface of a circular tube with variable-pitch-structured roughness of fixed height, (e/D_i), has been proposed and tested. This is achieved by inserting wire coils of fixed diameter that are made of a shape-memory alloy (SMA); the coil geometry, or spiral tightness, is designed to change in response to fluid flow temperature. Functionally, upon being heated, the wire-coil insert that has fixed (e/D_i) changes from its original compressed shape, which occupies a small fraction of the tube length, to an expanded shape that has the desired or "trained" roughness pitch (p/D_i) and helix pitch $(\alpha/90)$.

Extended surfaces can be considered "old technology" as far as most applications are concerned. The real interest now is in increasing heat transfer coefficients on the extended surface itself, and this is often achieved by a variety of geometrical alterations of the fins so as to produce what would be termed as enhanced extended surfaces. Present-day compact heat exchangers of the plate-fin or tube-and-fin variety use several different such enhanced fins: offset-strip fins (Figure 4.16.1b), louvered fins, perforated fins, wavy or corrugated fins (Figure 4.16.1b), and some newer configurations that are still undergoing laboratory testing and modeling (Huzayyin et al., 2010). Heat transfer coefficients that are several hundred percent above the plain fin values can be achieved without significant loss of surface area; however, the pressure drop is also substantially increased, and, as is the case with all enhancement devices, an appropriate balancing and thermal–hydrodynamic optimization is required for a given application. Common usage includes automobile radiators and oil coolers, aircraft recuperators, evaporators and condensers for refrigeration and air-conditioning units, among many others.

For the case of rectangular offset-strip fins (Figure 4.16.1b), the following correlations are recommended for calculating the Colburn factor j and Fanning friction factor f for varying flow and offset-fin-geometry characteristics (Manglik and Bergles, 1995):

$$j_h = 0.6522\,\mathrm{Re}_h^{-0.5403}\alpha^{-0.1541}\delta^{0.1499}\gamma^{-0.0678}\left[1 + 5.269\times10^{-5}\,\mathrm{Re}_h^{1.340}\alpha^{0.504}\delta^{0.456}\gamma^{-1.055}\right]^{0.1}, \tag{4.16.5}$$

$$f_h = 9.6243\,\mathrm{Re}_h^{-0.7422}\alpha^{-0.1856}\delta^{0.3053}\gamma^{-0.2659}\left[1 + 7.669\times10^{-8}\,\mathrm{Re}_h^{4.429}\alpha^{0.920}\delta^{3.767}\gamma^{0.236}\right]^{0.1}. \tag{4.16.6}$$

Here j_h, or the Colburn j-factor, is defined as $\left(\mathrm{Nu_h}/\mathrm{Re_h}\mathrm{Pr}^{1/3}\right)$, and $\mathrm{Nu_h}$, $\mathrm{Re_h}$, and f_h, are based on the hydraulic diameter of the inter-fin passage given by

$$D_h = \frac{4shl}{\left[2\left(sl + hl + th\right) + ts\right]}. \tag{4.16.7}$$

The special symbols used in these expressions represent the following: α=aspect ratio s/h, δ=ratio t/l, γ=ratio t/s, s=lateral spacing of strip fin, h=strip fin height, l=length of one offset module of strip fins, and t=fin thickness. These equations are based on experimental data for 18 different offset-strip-fin geometries, and they represent the data continuously in the laminar, transition, and turbulent flow regions.

Internally finned circular tubes are typically available in aluminum and copper (or copper alloys), among some other metal alloys, for commercial usage. Correlations (for heat transfer coefficient and friction factor) are available for laminar flow as well as the turbulent regime, for both straight and spiral continuous fins. For internally finned tubes with straight or spiral fins and laminar flows, Watkinson et al. (1975) have proposed the following equations, based on the hydraulic diameter D_h of the tube, which include a common expression for the isothermal Fanning friction factor but two different equations for the respective Nusselt numbers:

$$f_h = \left(\frac{16.4}{\mathrm{Re_h}}\right)\left(\frac{D_h}{D_i}\right)^{1.4}. \tag{4.16.8}$$

Straight-fin tubes:

$$\mathrm{Nu_h} = \frac{1.08 \times \log \mathrm{Re_h}}{n^{0.5}\left(1 + 0.01\mathrm{Gr_h^{1/3}}\right)} \mathrm{Re_h^{0.46}} \mathrm{Pr}^{1/3} \left(\frac{L}{D_h}\right)^{1/3} \left(\frac{\mu_w}{\mu_b}\right)^{0.14}. \tag{4.16.9a}$$

Spiral-fin tubes:

$$\mathrm{Nu_h} = \frac{8.533 \times \log \mathrm{Re_h}}{\left(1 + 0.01\mathrm{Gr_h^{1/3}}\right)} \mathrm{Re_h^{0.26}} \mathrm{Pr}^{1/3} \left(\frac{t}{p}\right)^{0.5} \left(\frac{L}{D_h}\right)^{1/3} \left(\frac{\mu_w}{\mu_b}\right)^{0.14}. \tag{4.16.9b}$$

In these equations, n is the number of fins, L the fin height, t the fin thickness, and p is the spiral pitch; the subscripts w and b, respectively, refer to properties evaluated at the tube-wall temperature and the bulk fluid temperature. Turbulent flow in tubes with straight or helical fins was correlated by Carnavos (1979) as

$$\mathrm{Nu_h} = 0.023\mathrm{Pr}^{0.4}\mathrm{Re_h^{0.8}} \left[\frac{A_c}{A_{ci}}\right]^{0.1} \left[\frac{A_{s,i}}{A_s}\right]^{0.5} \left(\sec\alpha\right)^3, \tag{4.16.10}$$

$$f_h = 0.046\mathrm{Re_h^{-0.2}} \left[\frac{A_c}{A_{c,i}}\right]^{0.5} \left(\sec\alpha\right)^{0.75}, \tag{4.16.11}$$

where $A_{c,i}$ is based on the maximum inside (envelope) flow area, $A_{s,i}$ is based on the maximum inside (envelope) surface area, and α the spiral angle in degrees for helical fins.

A numerical analysis of turbulent flow in tubes with idealized straight fins was reported. The necessary constant for the turbulence model was obtained from experimental data for air. Further improvements in numerical techniques are expected, so that a wider range of geometries and fluids can be handled without resort to extensive experimental programs. Moreover, for use of fins or extended surfaces, whether attached to tubular flow channels or in plate-fin-type ducts, the issue of

fin surface area thermal effectiveness (based on fin efficiency) has to be taken into account in the heat exchanger design. Guidelines for such considerations have been provided in a recent extended analysis (Manglik et al., 2011), and these are generally applicable to all plate-fin geometries as well as other extended surfaces.

Many proprietary surface configurations have been produced by deforming the basic tube. They are often referred to as "convoluted," "corrugated," "spiral," or "spirally fluted" tubes (Bergles, 1998; 1999; Zimparov et al., 2012), and they have multiple-start spiral corrugations along the tube length. The corrugations, or tube surface flutes, are substantially large and deep enough so as to add "extended" area to the otherwise plain tube surface. A systematic survey of the single-tube performance of condenser tubes indicates up to 400% increase in the nominal inside heat transfer coefficient (based on diameter of a smooth tube of the same maximum inside diameter); however, pressure drops on the water side are about 20 times higher.

Displaced enhancement devices are typically in the form of inserts, within elements arranged to promote transverse mixing (e.g., the looped wire insert of Figure 4.16.1c). They are used primarily for viscous liquids, to promote either heat transfer or mass transfer; the wire-loop inserts have been used for enhancement of both laminar and turbulent flow. Displaced-flow promoters, manufactured as bent-strip-type turbulators with a variety of geometric features (Bergles et al., 1991), are also used to enhance the radiant heat transfer in high-temperature applications. In the flue tube of a hot-gas-fired hot water heater, there is a trade-off between radiation and convection (Junkhan et al., 1985).

Another type of displaced insert generates vortices that sweep the heat transfer surface downstream of the element, and thus enhance the downstream flow. Delta wing and rectangular wing promoters, both corotating and counterrotating, have been studied in a variety of geometric attributes (Fiebig, 1995; Bergles and Manglik, 2013). One manifestation of this is to punch a wing-shaped protrusion into the fins of a compact heat exchanger, and this can be achieved in both plate-fin and tube-fin configurations (Bergles and Manglik, 2013). However, no general correlations for f and Nu have been developed to facilitate the implementation in and design of heat exchangers; this is perhaps also a reflection of the challenges of scaling and correlating a large number of parameters that come to play in their usage.

Swirl-flow devices are made up of tube inserts, geometrical surface variations, and tangential fluid entry arrangements that produce swirling and vortex-induced mixing along the flow direction. They typically include twisted-tape inserts, periodic tangential fluid injection, and helically twisted tubes, among a few others (Manglik and Bergles, 2002). Of these, twisted-tape inserts have been widely used to improve heat transfer in both laminar and turbulent flow inside circular tubes. Correlations for predicting Fanning friction factors and Nusselt numbers for laminar flow as well as turbulent flow in tubes have been developed (Manglik and Bergles, 1993a; 1993b).

In the case of laminar flows, the helical swirl generation in the tube fitted with a twisted-tape insert has been found to scale with the following dimensionless swirl parameter (Manglik and Bergles, 1993a):

$$\mathrm{Sw} = \left(\mathrm{Re}_s \big/ \sqrt{y}\right). \tag{4.16.12a}$$

Here y is the tape-twist ratio (180° helical twist pitch divided by tube inside diameter; the geometrical attributes are identified in Figure 4.16.3), and based on the mass flux G (mass flow rate per unit cross-sectional area) the swirl velocity V_s and Re_s are

$$\mathrm{Re}_s = \left(\rho V_s D_i/\mu\right), V_s = \left(G/\rho\right)\left[1+\left(\pi/2y\right)^2\right]^{1/2}. \tag{4.16.12b}$$

Based on this scaling, the isothermal Fanning friction factor is given by

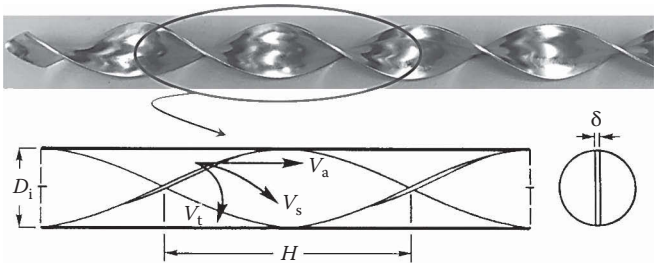


FIGURE 4.16.3 A typical twisted-tape insert and its geometrical descriptors.

$$f_s = \frac{15.767}{\mathrm{Re}_s}\left[\frac{\pi + 2 - 2(\delta/D_i)}{\pi - 4(\delta/D_i)}\right]^2 \left(1 + 10^{-6}\,\mathrm{Sw}^{2.55}\right)^{1/6}, \tag{4.16.13a}$$

where δ is the tape thickness, and f_s is based on the effective swirl velocity and swirl-flow length:

$$f_s = \left(\frac{\Delta p D_i}{2\rho V_s^2 L_s}\right), \text{ and } L_s = L\left[1 + \left(\frac{\pi}{2y}\right)^2\right]^{1/2}. \tag{4.16.13b}$$

For laminar forced convection heat transfer in tubes with twisted-tape inserts and walls maintained at a uniform temperature, the following correlation (Manglik and Bergles, 1993a) is recommended:

$$\mathrm{Nu}_m = \underbrace{4.612\left(\mu_b/\mu_w\right)^{0.14}}_{\text{fully developed flow}}\left[\left\{\underbrace{\left(1 + 0.0951\mathrm{Gz}^{0.894}\right)^{2.5}}_{\text{thermal entrance}} + \underbrace{6.413\times10^{-9}\left(\mathrm{Sw}\cdot\mathrm{Pr}^{0.391}\right)^{3.835}}_{\text{swirl flows}}\right\}^2 + \underbrace{2.132\times10^{-14}\left(\mathrm{Re}_a\cdot\mathrm{Ra}\right)^{2.23}}_{\text{free convection}}\right]^{0.1}. \tag{4.16.14}$$

The interplay of the different groupings (convective effects) on the twisted-tape-induced heat transfer performance is shown in Figure 4.16.4. These correlations have been found to predict very well the data available in the literature (Manglik et al., 2001; Manglik and Bergles, 2013).

In the turbulent flow regime ($\mathrm{Re} \geq 10^4$), the scaling with Sw is inapplicable and instead the Fanning friction factor correlates as

$$f = \left(\frac{0.0791}{\mathrm{Re}^{0.25}}\right)\underbrace{\left[\frac{\pi}{\pi - (4\delta/D_i)}\right]^{1.75}\left[\frac{\pi + 2 - (2\delta/D_i)}{\pi - (4\delta/D_i)}\right]^{1.25}}_{\text{tube partitioning}}\underbrace{\left(1 + \frac{2.752}{y^{1.29}}\right)}_{\text{swirl}}. \tag{4.16.15}$$

For turbulent heat transfer with $\mathrm{Re} \geq 10^4$, the recommended Nusselt number correlation is

$$\mathrm{Nu} = \left(0.023\mathrm{Re}^{0.8}\mathrm{Pr}^{0.4}\right)\underbrace{\left[\frac{\pi}{\pi - (4\delta/D_i)}\right]^{0.8}\left[\frac{\pi + 2 - (2\delta/D_i)}{\pi - (4\delta/D_i)}\right]^{0.2}}_{\text{tube partitioning}}\underbrace{\left[1 + \frac{0.769}{y}\right]}_{\text{swirl}}\phi, \tag{4.16.16}$$

where ϕ is the temperature-dependent property correction factor and

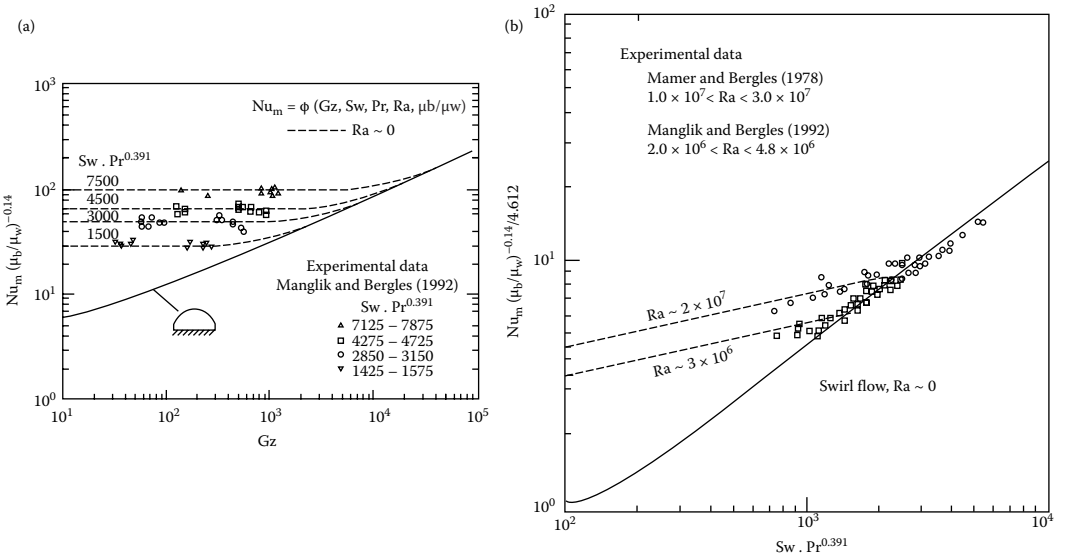


FIGURE 4.16.4 Influence of twisted-tape-generated swirl on the laminar forced convection in circular tubes with uniform wall temperature; comparison of prediction with experimental data. (a) Swirl flow and thermal entrance region and (b) swirl flow with superimposed free convection effects. (Data reported by Marner, W.J., and Bergles, A.E., Augmentation of tubeside laminar flow heat transfer by means of twisted-tape inserts, static-mixer inserts, and internally finned tubes. *Proceedings of Sixth International Heat Transfer Conference*, August 7–11, 1978, Toronto, Hemisphere, Washington DC, pp. 583–588, 1978; Manglik, R.M., and Bergles, A.E., *Wärme-und Stoffübertragung*, 27(4), 249–257, 1992.).

$$\phi = \left(\mu_b/\mu_w\right)^n \text{ or } \left(T_b/T_w\right)^m .$$

$$n = \begin{cases} 0.18 & \text{liquid heating} \\ 0.30 & \text{liquid cooling} \end{cases} \quad \text{and} \quad m = \begin{cases} 0.45 & \text{gas heating} \\ 0.15 & \text{gas cooling} \end{cases}$$

Again, these correlations find generalized vetting with much of the available data, and extended discussions on their development, validity, as well as design and applications issues are given in (Manglik et al., 2001; Manglik and Bergles, 2002).

Another attractive swirl-flow technique is to use noncircular ducts that are helically twisted along their flow axis. Two such examples are twisted oval tubes and twisted rectangular tubes as shown in Figure 4.16.5. With air flows inside constant surface temperature twisted oval tube, up to 2.0–2.5 times higher Nu are obtained, compared to straight oval tubes of same cross-sectional aspect ratio and same pumping power (Bishara et al., 2013). The relative enhancement, however, is less than that with a comparable twisted-tape insert. In the case of laminar viscous liquid ($5 \le Pr \le 100$) flows through twisted rectangular tube, and depending upon the cross-sectional aspect ratio and the helical twist ratio, up to 2.4–13 times higher heat transfer rates can be sustained on a constant-pumping-power basis compared to that in a straight duct (Manglik et al., 2012; Patel et al., 2012). Moreover, when considering the constraint of fixed heat load and pressure drop, as much as 50%–90% reduction in surface area can be obtained in a new heat exchanger.

Performance Evaluation Criteria for Single-Phase Forced Convection in Tubes

Numerous, and sometimes conflicting, factors enter into the ultimate decision to use an enhancement technique: heat duty increase or area reduction that can be obtained, initial cost, pumping

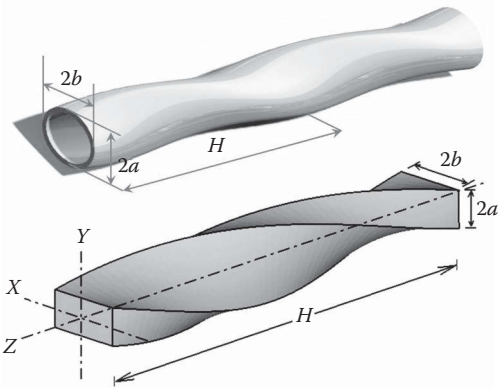


FIGURE 4.16.5 Swirl-producing twisted tubes of oval and rectangular cross sections.

power or operating cost, maintenance cost (especially cleaning), safety, and reliability, among others. These factors are difficult to quantify in a common comparative basis, and a generally acceptable selection criterion may not exist. It is possible, however, to suggest some performance criteria for preliminary design guidance. Several application-objective-based figures of merit have been developed to this end (Bergles, 1998), and a functional listing is given in Table 4.16.2. These seek to provide optimization metrics for the trade-off between improved convection heat transfer and the concomitant frictional penalty.

As an example, consider the criterion FG-2a, in which the basic heat exchanger geometry, approach temperature difference, and the pumping power fixed, and the application objective is to increase the heat transfer or the heat load capacity. This can be expressed as the following ratio of heat transfer rate for the tube with insert q_a and that for the plain or empty tube q_s:

TABLE 4.16.2
Performance Evaluation Criteria or Enhancement Optimization Figures of Merit for Single-Phase Forced Convection in Tubular Heat Exchangers

		Fixed				
Case	**Geometry**	$\dot{m}$	P	q	ΔT_i	**Objective**
FG-1a	N, L	✓			✓	$q \uparrow$
FG-1b	N, L	✓		✓		$\Delta T_i \downarrow$
FG-2a	N, L		✓		✓	$q \uparrow$
FG-2b	N, L		✓	✓		$\Delta T_i \downarrow$
FG-3	N, L			✓	✓	$P \downarrow$
FN-1	N		✓	✓	✓	$L \downarrow$
FN-2	N	✓		✓	✓	$L \downarrow$
FN-3	N	✓		✓	✓	$P \downarrow$
VG-1	–	✓	✓	✓	✓	$(NL) \downarrow$
VG-2a	(NL)	✓	✓		✓	$q \uparrow$
VG-2b	(NL)	✓	✓	✓		$\Delta T_i \downarrow$
VG-3	(NL)	✓		✓	✓	$P \downarrow$

N, number of tubes; L, length of tubes; $\dot{m}$, mass flow rate; P, pumping power; q, rate of heat transfer; and ΔT_i, inlet or approach temperature difference.

$$\text{FG-2a} = \frac{q_a}{q_s} = \left(\frac{\text{Nu}_a}{\text{Nu}_s}\right)_{N,L,D_i,P,\Delta T_i} . \tag{4.16.17}$$

Here, P=pumping power, ΔT_i=inlet or approach bulk temperature difference, N=number of tubes in a shell-and-tube exchanger, and L=length of the tubes.

With the pumping power per tube (neglecting entrance and exit losses) given as

$$P = \frac{\dot{m}\Delta p}{\rho} = \left(\frac{\pi\mu^2 L}{2\rho^2 D_i^2}\right)\left(f\text{Re}^3\right), \tag{4.16.18a}$$

and with a specified $\left(L/D_i\right)$, the fixed pumping power demand thus imposes the following restriction:

$$\left(f\text{Re}^3\right)_a = \left(f\text{Re}^3\right)_s . \tag{4.16.18b}$$

Typical results for tubes with twisted-tape inserts and single-phase convection in the laminar flow regime (Yerra et al., 2006) are presented in Figure 4.16.6. As much as 3.5–5 times higher heat transfer rates are seen to be sustained in heat exchangers, and the enhancement tends to be much higher with larger Re but smaller y and Pr.

Similar evaluations for most passive techniques can be carried out, though it would have different optimization objectives based on the specific heat exchanger application.

Active and Compound Techniques for Single-Phase Forced Convection

Under active techniques, *mechanically aided heat transfer* in the form of either surface rotation or surface scraping can increase forced convection heat transfer. The latter is widely used in the thermal processing of viscous food products. Surface scraping rejuvenates the heat transfer surface with unprocessed media, thereby removing the high-resistance processed media as well as effectively reducing surface fouling; these mechanisms improve convection as well as prevent its deterioration. A novel offshoot of this technique has recently been devised (Solano et al., 2011; Bergles and

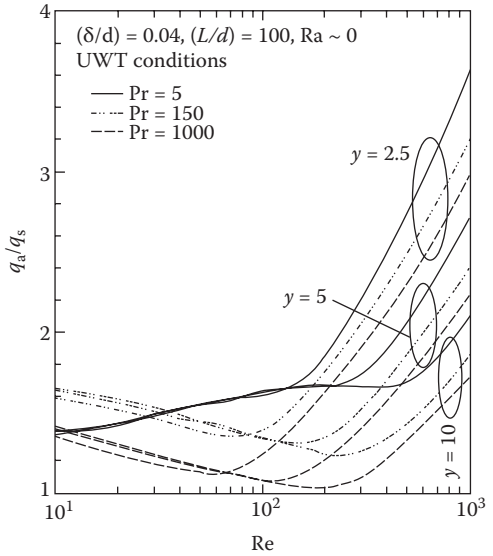


FIGURE 4.16.6 Enhanced heat transfer performance in laminar flows inside circular tubes with twisted-tape inserts for fixed pumping power and heat exchanger size (or criterion FG-2a).

Manglik, 2013); a tube-side self-cleaning scheme for shell-and-tube exchangers, which consists of a reciprocating rod and surface scraper elements.

Surface vibration has been demonstrated to improve heat transfer to both laminar and turbulent duct flow of liquids. In newer approaches to fruitfully employing this technique, as an example, piezoelectric actuators have been used to augment the natural convection heat transfer from air-cooled finned heat sinks. Likewise, *fluid vibration* has been extensively studied for both air (using acoustic fields, e.g., loudspeakers and sirens) and liquids (using flow interrupters, pulsations inducing actuators, and ultrasonic transducers). Pulsations are relatively simple to apply to low-velocity liquid flows, and improvements of several hundred percent can be realized. However, in some recent work, ultrasonic vibrations have also been applied to obtain enhanced subcooled boiling and not single-phase convection; see (Bergles and Manglik, 2013).

Some very impressive enhancements have been recorded with *electrical fields* (DC or AC), particularly in the laminar-flow region and when applied to dielectric liquids. Improvements of at least 100% were obtained when voltages in the 10-kV range were applied to transformer oil. It is found that even with intense electrostatic fields, the heat transfer enhancement disappears as turbulent flow is approached in a circular tube with a concentric inner electrode. In essence, electrostatic fields tend to promote greater bulk fluid mixing or flow disruption and thereby enhance heat transfer.

Compound techniques are a slowly emerging area of enhancement that holds promise for practical applications, since heat transfer coefficients can usually be increased above each of the several techniques acting alone. Some examples that have been studied (Manglik and Bergles, 2004; Bergles and Manglik, 2013) are as follows: circular tubes with structured roughness (ribs and/or corrugations) and twisted-tape inserts, a rough cylinder subject to acoustic vibrations, internally finned tube with twisted-tape inserts, finned tubes submerged in fluidized beds, externally finned tubes subjected to vibrations, rotating channels with internal ribbed roughness, gas–solid suspension with an electrical field, fluidized bed with pulsations of air, and a rib-roughened channel with longitudinal vortex generation.

In an extended recent study (Zimparov et al., 2012), the compound usage of twisted tapes by inserting them inside corrugated tubes (or tubes with structured roughness that in some cases increase area as well), have been considered in turbulent water flows ($3.5 \times 10^3 < \mathrm{Re} < 5.0 \times 10^4$). Several variations of tube corrugation and tape geometries are explored, and the data suggest that, for the specified experimental conditions, the best thermodynamic performance is obtained with a corrugated tube with ($e/d = 0.057$; $p/e = 6.77$) and a tape-twist ratio of $y = 4.7$. In yet another example (Neshumayev and Tiikma, 2007), a twisted-tape insert with a spirally wound strip over its surface has been used for enhanced convective and radiative heat recovery from flue gases inside tubes of fire-tube boilers. The radiation contribution was ascertained to be 20% of the total heat flux.

Pool Boiling

Selected passive and active enhancement techniques have been shown to be effective for pool boiling and flow boiling/evaporation. Most techniques apply to nucleate boiling; however, some techniques are applicable to transition and film boiling.

It should be noted that phase-change heat transfer coefficients are relatively high. The main thermal resistance in a two-fluid heat exchanger often lies on the non-phase-change side. Fouling of either fluid-side surface can represent the dominant thermal resistance and alter this consideration (Bergles and Somerscales, 1995). For this reason, the emphasis is often on enhancement of single-phase flow. On the other hand, the overall thermal resistance may then be reduced to the point where significant improvement in the overall performance can only be achieved by enhancing the two-phase flow stream. Two-phase enhancement would also be important in double-phase-change (boiling/condensing) heat exchangers.

As discussed elsewhere in considering the fundamental aspects of heterogeneous phase change (Bergles, 1988), surface material and finish have a strong effect on nucleate and transition pool boiling.

However, reliable control of nucleation on plain surfaces is not easily accomplished. Accordingly, since the earliest days of boiling research, there have been attempts to relocate the boiling curve through use of relatively gross modification of the surface. For many years, this was accomplished simply by area increase in the form of low helical fins. The subsequent tendency was to structure surfaces to improve the nucleate boiling characteristics by a fundamental change in the boiling process. Many of these advanced surfaces are being used in commercial shell-and-tube boilers.

Several manufacturing processes have been employed: machining, forming, layering, and coating. A few examples of such manufactured surfaces are given in Figure 4.16.7, where the characteristic features of surface modifications produced by different machining methods are highlighted. In Figure 4.16.7a, a tunnel-and-pore arrangement produced by rolling, upsetting, and brushing is depicted. An alternative modification is to roll, split, and reroll very low fins or knurl and roll them into the different T-shaped structures shown in Figure 4.16.7b. In one of the earliest developments in this arena was to sinter or braze metallic particles on the surface to form a porous matric. These surfaces have been found to enhance pool boiling considerably (Bergles, 1997; Manglik, 2006). Wall superheat reductions of up to a factor of ten are common with these surfaces. The advantage is not only a high nucleate boiling heat transfer coefficient, but the fact that boiling can take place at very low-temperature differences.

These structured boiling surfaces, which span the domain of *rough surfaces* and *extended surfaces*, and which were developed for refrigeration and process applications, have also been used as "heat sinks" for immersion-cooled microelectronic chips. For applications in large-scale heat exchangers, shell-and-tube exchangers, for example, the behavior of tube bundles is found to often be different with structured-surface tubes. Enhanced nucleate boiling dominates in such cases, and the convective boiling enhancement, found in plain tube bundles, does not occur.

Some recent noteworthy development has been in the use *additives for liquids* to enhance pool boiling heat transfer. These include solid particles (in recent usage referred to as "nano" particles), liquid traces, soluble substances (e.g., surface-active agents and polymers), and gas bubbles, among some others, to form suspensions or colloidal solutions and liquid mixtures. The addition of surfactants or certain surface-active polymers to water in weak concentrations has been shown to significantly enhance pool boiling from a cylindrical heater (Wasekar and Manglik, 2000; 2002; Zhang and Manglik, 2004; 2005a; 2005b). Up to ~20% enhancement was reported (Athavale et al., 2012) with weak aqueous solutions of the reagent-like polymer HEC QP-300; in stronger solutions with higher polymer concentrations, however, heat transfer degrades at low heat fluxes (incipience and partial boiling) but is enhanced by ~45% at higher heat fluxes (fully developed nucleate boiling). The complex interplay of reagent-induced liquid–solid surface wetting, along with dynamic (time-dependent) changes in the vapor–liquid interfacial tension as well as the shear-dependent shear-thinning (in polymers) has been attributed to the molecular mobility (adsorption-desorption, physisorption, and electrokinetics) of the additive for effecting the enhanced boiling behavior

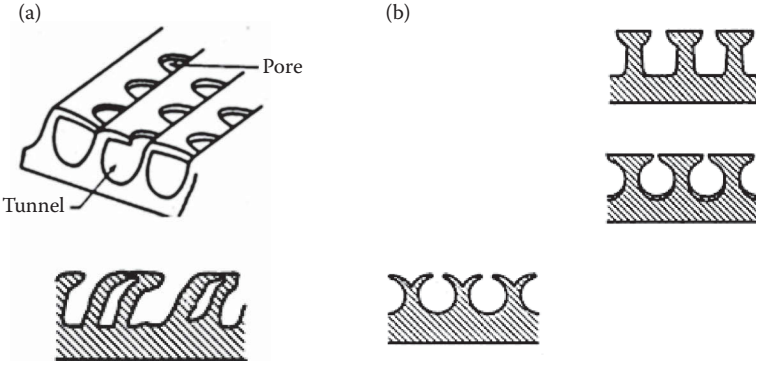


FIGURE 4.16.7 Some examples of structured boiling surfaces produced by machining and forming: (a) deeply knurled surface, or Hitachi Thermoexcel™ surface and (b) low-fin surfaces, or Wieland Gewa™ surfaces.

(Manglik, 2011); Marangoni convection, consisting of both thermocapillary and diffusocapillary convection might also be at play in these situations (Wasekar and Manglik, 2003).

Active enhancement techniques include heated surface rotation, surface wiping, surface vibration, fluid vibration, electrostatic fields, and suction at the heated surface. Although active techniques are effective in reducing the wall superheat and/or increasing the critical heat flux, the practical applications are very limited, largely because of the difficulty of reliably providing the mechanical or electrical effect.

Compound enhancement, which involves two or more techniques applied simultaneously, has also been studied. Electrohydrodynamic enhancement was applied to a finned tube bundle, resulting in nearly a 200% increase in the average boiling heat transfer coefficient of the bundle, with a small power consumption for the field. Another example is the use of a wire mesh to wrap the surface of a micro-finned tube to promote compound pool boiling enhancement with refrigerants (Bergles and Manglik, 2013). The intent is to augment the extended surface with nucleation sites (provided by the wire mesh) so as to aid the nucleate phase-change process.

Convective Boiling/Evaporation

The structured surfaces described in the previous section are generally not used for in-tube vaporization, because of the difficulty in manufacturing. One notable exception is the high-flux surface in a vertical thermosiphon reboiler. The considerable increase in the low-quality, nucleate boiling coefficient is desirable, but it is also important that more vapor is generated to promote circulation.

Structured *rough surfaces*, comprising of helical-repeated ribs and helically coiled wire inserts have been used to increase vaporization coefficients and the dry-out heat flux in once-through boilers.

Numerous tubes with internal fins (*extended surface*), either integral or attached, are available for refrigerant evaporators. Original configurations were tightly packed, copper, offset strip fin inserts soldered to the copper tube or aluminum, star-shaped inserts secured by drawing the tube over the insert. Examples are shown in Figure 4.16.8. Average heat transfer coefficients (based on surface area of smooth tube of the same diameter) for typical evaporator conditions are increased by as much as 200%. A cross-sectional view of a typical "microfin" tube is included in Figure 4.16.8. The average evaporation boiling coefficient is increased 30%–80%. The pressure drop penalties are less; that is, lower percentage increases in pressure drop are frequently observed.

Twisted-tape inserts, or *swirl-flow devices*, are generally used to increase the burnout heat flux for subcooled boiling at high imposed heat fluxes $10–10^2$ MW/m^2, as might be encountered in the cooling of fusion reactor components. Increases in burnout heat flux of up to 200% were obtained at near atmospheric pressure. In the bulk boiling region, the helically flowing liquid tends to be centrifuged to the wall so that a liquid film is maintained during the evaporation process, and thereby delaying dryout until a very high quality. The remaining droplets are then again centrifuged to the wall, which reduces the temperature excursion. To provide additional performance and predictive details, a mechanistic evaluation of subcooled as well as bulk boiling is delineated elsewhere (Manglik and Bergles, 2013). In fact, twisted tapes have been shown to enhance convective vaporization of water

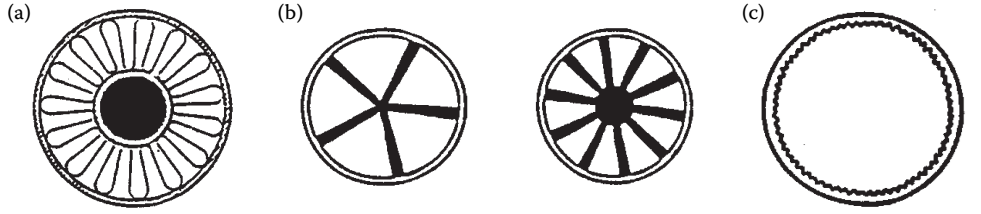


FIGURE 4.16.8 Inner-fin tubes for refrigerant evaporators: (a) strip-fin inserts, (b) star-shaped inserts, and (c) microfin.

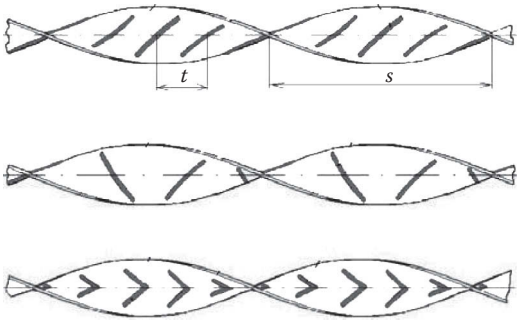


FIGURE 4.16.9 Compound enhancement using twisted-tape inserts that are wire-wound or have wire-laced ribs on their surface in different array configurations. (Adapted from Tarasevich, S.E., Yakovlev, A.B., Giniyatullin, A.A., and Shishkin, A.V., Heat and mass transfer in tubes with various twisted tape inserts. *ASME 2011 IMECE*, Denver, CO. Denver, CO: ASME, New York, Paper No. IMECE2011-62088, 2011).

(low pressure and high pressure), refrigerants (R-11, R-113, R-12, and R-22), a cryogenic fluid (liquid nitrogen), and liquid metals (cesium, mercury, potassium, rubidium, and sodium).

In a recent study on a novel *compound enhancement* usage of twisted tapes that have either wire wraps or laces that form "ribs" on the tape surface (Tarasevich et al., 2011), it has been experimentally shown that two-phase flow behavior is improved. A sampling of some of the modified tapes used in the study is presented in Figure 4.16.9. In the annular flow regime, the liquid accumulation on the tape surface tends to reduce. The small-scale tape-surface wire ribs are perhaps instrumental in disrupting liquid film formation and thereby reducing "dry spots" at the tube surface. While the hydrodynamic behavior has been observed visually, quantification of the heat transfer enhancement has yet to be established (Manglik and Bergles, 2013); though ~2.5 times higher Nusselt number with a 10-fold increase in the frictional resistance has been reported for single-phase flows.

Vapor-Space Condensation

As discussed elsewhere, condensation can be either filmwise or dropwise. In a sense, dropwise condensation is enhancement of the normally occurring film condensation by surface treatment. The only real application is for steam condensers, because nonwetting coatings are not available for most other working fluids. Even after much study, little progress has been made in developing permanently hydrophobic coatings for practical steam condensers. The enhancement of dropwise condensation is pointless, because the heat transfer coefficients are already so high.

Surface extensions are widely employed for enhancement of condensation. The integral low fin tubing (with a tapered narrow-tip to wide-base profile), used for kettle boilers, is also used for horizontal tube condensers. With proper spacing of the fins to provide adequate condensate drainage, the average coefficients can be several times those of a plain tube with the same base diameter. These fins are normally used with refrigerants and other organic fluids that have low condensing coefficients, but which drain effectively, because of low surface tension.

The fin profile can be altered according to mathematical analysis to take full advantage of the Gregorig effect, whereby condensation occurs mainly at the tops of convex ridges. Surface tension forces then pull the condensate into concave grooves, where it runs off. The average heat transfer coefficient is greater than that for an axially uniform film thickness. The initial application was for condensation of steam on vertical tubes used for reboilers and in desalination. According to numerical solutions, the optimum geometry is characterized by a sharp fin tip, gradually changing curvature of the fin surface from tip to root, wide grooves between fins to collect condensate, and periodic condensate strippers. Figure 4.16.10 schematically shows the configuration, as that employed by Mori et al. (1981).

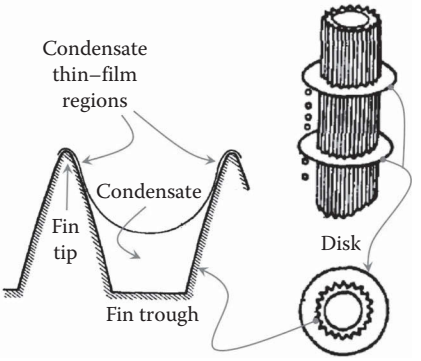


FIGURE 4.16.10 Recommended flute profile and schematic of condensate strippers [adapted from Mori et al. (1981)].

Recent interest has centered on 3-D surfaces for horizontal-tube condensers. The considerable improvement relative to low fins or other 2-D profiles is apparently due to multidimensional drainage at the fin tips. Other 3-D shapes include circular pin fins, square pins, and small metal particles that are bonded randomly to the surface. Another focus area has been to produce hydrophobic surfaces to promote dropwise condensation, which yields order of magnitude higher heat transfer coefficient compared to that in film condensation (Rose, 2015). Plasma ion implantation has been employed in one instance to create high hydrophobicity on surfaces of stainless steel tubes.

Convective Condensation

This final section on enhancement of the various modes of heat transfer focuses on in-tube condensation. The applications include horizontal kettle-type reboilers, moisture separator reheaters for nuclear power plants, and air-conditioner condensers.

Internally grooved or knurled tubes, deep spirally fluted tubes, random roughness, and conventional inner-fin tubes have been shown to be effective for condensation of steam and other fluids.

The microfin tubes mentioned earlier have also been applied successfully to in-tube condensing. As in the case of evaporation, the substantial heat transfer improvement is achieved at the expense of a lesser percentage increase in pressure drop. By testing a wide variety of tubes, it has been possible to suggest some guidelines for the geometry, for example, more fins, longer fins, and sharper tips; however, general correlations are not yet available. Fortunately for heat-pump operation, the tube that performs best for evaporation also performs best for condensation.

Twisted-tape inserts result in rather modest increases in heat transfer coefficient for complete condensation of either steam or refrigerant. The pressure drop increases are large because of the large wetted surface. Coiled tubular condensers provide a modest improvement in average heat transfer coefficient.

REFERENCES

Athavale, A.D., Manglik, R.M., and Jog, M.A. (2012). An experimental investigation of nucleate pool boiling in aqueous solutions of a polymer. *AIChE J.*, 58(3), 668–677.

Bergles, A.E. (1988). Fundamentals of boiling and evaporation. In Kakaç, S., Bergles, A.E., and Fernandes, E.O. (Eds.), *Two-Phase Flow Heat Exchangers: Thermal-Hydraulic Fundamentals and Design* (pp. 159–200). The Netherlands: Kluwer.

Bergles, A.E. (1997). Enhancement of pool boiling. *Int. J. Refrig.*, 20(8), 545–551.

Bergles, A.E. (1998). Techniques to enhance heat transfer. In Rohsenow, W.M., Hartnett, J.P., and Cho, Y.I. (Eds.), *Handbook of Heat Transfer* (3rd ed., pp. 11.11–11.76). New York: McGraw-Hill.

Bergles, A.E. (1999). Enhanced heat transfer: Endless frontier, or mature and routine? *J. Enhanced Heat Transfer*, 6(2–4), 79–88.

Bergles, A.E., and Manglik, R.M. (2013). Current progress and new developments in enhanced heat and mass transfer. *J. Enhanced Heat Transfer*, 20(1), 1–15.

Bergles, A.E., and Somerscales, E.F.C. (1995). The effect of fouling on enhanced heat transfer equipment. *J. Enhanced Heat Transfer*, 2(1–2), 157–166.

Bergles, A.E., Bunn, R.L., and Junkhan, G.H. (1974). Extended performance evaluation criteria for enhanced heat transfer surfaces. *Lett. Heat and Mass Transfer*, 1(2), 113–120.

Bergles, A.E., Jensen, M.K., Somerscales, E.F.C., and Manglik, R.M. (1991). Literature review of heat transfer enhancement technology for heat exchangers in gas-fired applications, Report No. GRI 91–0146. Gas Research Institute, Chicago, IL.

Bergles, A.E., Nirmalan, V., Junkhan, G.H., and Webb, R.L. (1983). Bibliography on augmentation of convective heat and mass transfer – II, Heat Transfer Laboratory Report No. HTL-31, ISU-ERI-Ames-84221. Iowa State University, Ames, IA.

Bishara, F., Jog, M.A., and Manglik, R.M. (2013). Heat transfer enhancement due to swirl effects in oval tubes twisted about their longitudinal axis. *J. Enhanced Heat Transfer*, 20(4), 289–304.

Carnavos, T.C. (1979). Cooling air in turbulent flow with internally finned tubes. *Heat Transfer Eng.*, 1(2), 41–46.

Champagne, P.R., and Bergles, A.E. (2011). Development and testing of a novel, variable-roughness technique to enhance, on demand, heat transfer in a single-phase heat exchanger. *J.Enhanced Heat Transfer*, 8(5), 341–352.

Fiebig, M. (1995). Vortex generators for compact heat exchangers. *J. Enhanced Heat Transfer*, 2(1–2), 43–61.

Huzayyin, O.A., Jog, M.A., and Manglik, R.M. (2010). Low Reynolds number air-flow heat transfer in trapezoidally corrugated perforated plate-fin ducts. *ASHRAE Trans.*, 116(2), 339–346.

Junkhan, G.H., Bergles, A.E., Nirmalan, V., and Ravigururajan, T.S. (1985). Investigation of turbulators for fire tube boilers. *J. Heat Transfer*, 107(2), 354–360.

Kays, W.M., and London, A.L. (1998). *Compact Heat Exchangers* (3rd ed.). Malabar, FL: Krieger Publishing Company.

Kraus, A.D., Aziz, A., and Welty, J. (2001). *Extended Surface Heat Transfer.* New York: Wiley.

Kreith, F., Manglik, R.M., and Bohn, M.S. (2011). *Principles of Heat Transfer* (7th ed.). Stamford, CT: Cengage Learning.

Manglik, R.M. (2003). Heat transfer enhancement. In Bejan, A., and Kraus, A.D. (Eds.), *Heat Transfer Handbook* (pp. 1029–1130) (Chapter 14). Hoboken, NJ: Wiley.

Manglik, R.M. (2006). On the advancements in boiling, two-phase flow heat transfer, and interfacial phenomena. *J. Heat Transfer*, 128(12), 1237–1242.

Manglik, R.M. (2011). Molecular-to-macro-scale control of interfacial behavior in ebullient phase change in aqueous solutions of reagents. *Int. J. Transp. Phenom.*, 12(3–4), 229–243.

Manglik, R.M., and Bergles, A.E. (1992). Heat transfer enhancement and pressure drop in viscous liquid flows in isothermal tubes with twisted-tape inserts. *Wärme-und Stoffübertragung*, 27(4), 249–257.

Manglik, R.M., and Bergles, A.E. (1993a). Heat transfer and pressure drop correlations for twisted-tape inserts in isothermal tubes: Part I-laminar flows. *J. Heat Transfer*, 115(4), 881–889.

Manglik, R.M., and Bergles, A.E. (1993b). Heat transfer and pressure drop correlations for twisted-tape inserts in isothermal tubes: Part II-transition and turbulent flows. *J. Heat Transfer*, 115(4), 890–896.

Manglik, R.M., and Bergles, A.E. (1995). Heat transfer and pressure drop correlations for the rectangular offset-strip-fin compact heat exchanger. *Exp. Therm. Fluid Sci.*, 10(2), 171–180.

Manglik, R.M., and Bergles, A.E. (2002). Swirl flow heat transfer and pressure drop with twisted-tape inserts. In Hartnett, J.P., Irvine, T.F., Cho, Y.I., and Greene, G.A. (Eds.), *Advances in Heat Transfer* (Vol. 36, pp. 183–266). New York: Academic.

Manglik, R.M., and Bergles, A.E. (2004). Enhanced heat and mass transfer in the new millennium: A review of the 2001 literature. *J. Enhanced Heat Transfer*, 11(2), 87–118.

Manglik, R.M., and Bergles, A.E. (2013). Characterization of twisted-tape-induced helical swirl flows for enhancement of forced convective heat transfer in single-phase and two-phase flows. *J. Therm. Sci. Eng. Appl.*, 5(2), 021010 (021011–021012). doi:10.1115/1.4023935.

Manglik, R.M., and Jog, M.A. (2009). Molecular-to-large-scale heat transfer with multiphase interfaces: current status and new directions. *J. Heat Transfer*, 131(12), 121001(121001–121011).

Manglik, R.M., Bergles, A.E., Dongaonkar, A.J., and Rajendran, S. (2013). Limitations of compiling the global literature on enhanced heat and mass transfer. *J. Enhanced Heat Transfer*, 20(1), 83–92.

Manglik, R.M., Huzayyin, O.A., and Jog, M.A. (2011). Fin effects in flow channels of plate-fin compact heat exchanger cores. *J. Therm. Sci. Eng. Appl.*, 3(4), 041004 (041001–041009).

Manglik, R.M., Maramraju, S., and Bergles, A.E. (2001). The scaling and correlation of low reynolds number swirl flows and friction factors in circular tubes with twisted-tape inserts. *J. Enhanced Heat Transfer*, 8(6), 383–395.

Manglik, R.M., Patel, P., and Jog, M.A. (2012). Swirl-enhanced forced convection through axially twisted rectangular ducts-Part 2, heat transfer. *J. Enhanced Heat Transfer*, 19(5), 437–450.

Marner, W.J., and Bergles, A.E. (1978). Augmentation of tubeside laminar flow heat transfer by means of twisted-tape inserts, static-mixer inserts, and internally finned tubes. *Proceedings of Sixth International Heat Transfer Conference*, August 7–11, 1978, Toronto, Hemisphere, Washington DC, Vol. 2, pp. 583–588.

Mori, Y., Hijikata, K., Hirasawa, S., and Nakayama, W. (1981). Optimized performance of condensers with outside condensing surfaces. *J. Heat Transfer*, 103(1), 96–102.

Neshumayev, D., and Tiikma, T. (2007). Radiation heat transfer of turbulator inserts in gas-heated channels. *Proceedings of the Baltic Heat Transfer Conference*, September 19–21, 2007 Saint Petersburg State Polytechnical University, Russia, Vol. 2, pp. 525–531.

Patel, P., Manglik, R.M., and Jog, M.A. (2012). Swirl-enhanced forced convection through axially twisted rectangular ducts-Part 1, fluid flow. *J. Enhanced Heat Transfer*, 19(5), 423–436.

Ravigururajan, T.S., and Bergles, A.E. (1994). Visualization of flow phenomena near enhanced surfaces. *J. Heat Transfer*, 116(1), 54–57.

Ravigururajan, T.S., and Bergles, A.E. (1996). Development and verification of general correlations for pressure drop and heat transfer in single-phase turbulent flow in enhanced tubes. *Exp. Therm. Fluid Sci.*, 13(1), 55–70.

Rose, J.W. (2015). Personal reflections on fifty years of condensation heat transfer research. *J. Enhanced Heat Transfer*, 22(2), 89–120.

Shah, R.K., and Sekulic, D.P. (2003). *Fundamentals of Heat Exchanger Design*. New York: Wiley.

Solano, J.P., Garcia, A., Vicente, P.G., and Viedma, A. (2011). Performance evaluation of a zero-fouling reciprocating scraped-surface heat exchanger. *Heat Transfer Eng.*, 32(3–4), 331–338.

Tarasevich, S.E., Yakovlev, A.B., Giniyatullin, A.A., and Shishkin, A.V. (2011). Heat and mass transfer in tubes with various twisted tape inserts. *ASME 2011 IMECE*, Denver, CO. Denver, CO: ASME, New York, Paper No. IMECE2011-62088.

Thome, J.R. (1990). *Enhanced Boiling Heat Transfer.* New York: Hemisphere.

Wang, L., Sundén, B., and Manglik, R.M. (2007). *Plate Heat Exchangers: Design, Applications and Performance*. Southampton: WIT Press.

Wasekar, V.M., and Manglik, R.M. (2000). Pool boiling heat transfer in aqueous solutions of an anionic surfactant. *J. Heat Transfer*, 122(4), 708–715.

Wasekar, V.M., and Manglik, R.M. (2002). The influence of additive molecular weight and ionic nature on the pool boiling performance of aqueous surfactant solutions. *Int. J. Heat and Mass Transfer*, 45(3), 483–493.

Wasekar, V.M., and Manglik, R.M. (2003). Short-time-transient surfactant dynamics and marangoni convection around boiling nuclei. *J. Heat Transfer*, 125(5), 858–866.

Watkinson, A.P., Miletti, D.C., and Kubanek, G.R. (1975). Heat transfer and pressure drop of internally finned tubes in laminar oil flows. Paper No. 75-HT-41. ASME, New York.

Webb, R.L., and Kim, N.-H. (2005). *Principles of Enhanced Heat Transfer* (2nd ed.). Boca Raton, FL: Taylor & Francis.

Wright, L.M., and Han, J.-C. (2014). Heat transfer enhancement for turbine blade internal cooling. *J. Enhanced Heat Transfer*, 21(2–3), 111–140.

Yerra, K.K., Manglik, R.M., and Jog, M.A. (2006). Optimization of heat transfer enhancement in single-phase tubeside flows with twisted-tape inserts. *Int. J. Heat Exchangers*, 8(1), 117–138.

Zhang, J., and Manglik, R.M. (2004). Effect of ethoxylation and molecular weight of cationic surfactants on nucleate boiling in aqueous solutions. *J. Heat Transfer*, 126(1), 34–42.

Zhang, J., and Manglik, R.M. (2005a). Additive adsorption and interfacial characteristics of nucleate pool boiling in aqueous surfactant solutions. *J. Heat Transfer*, 127(7), 684–691.

Zhang, J., and Manglik, R.M. (2005b). Nucleate pool boiling of aqueous polymer solutions on a cylindrical heater. *J. Non-Newtonian Fluid Mech.*, 125(2–3), 185–196.

Zimparov, V., Petkov, V.M., and Bergles, A.E. (2012). Performance characteristics of deep corrugated tubes with twisted-tape inserts. *J. Enhanced Heat Transfer*, 19(1), 1–11.

Further Information

This section gives some indication as to why heat transfer enhancement is one of the fastest growing areas of heat transfer. Many techniques are available for improvement of the various modes of heat transfer. Fundamental understanding of the transport mechanism is growing, but, more importantly, design correlations are being also established. Many effective and cost-competitive enhancement techniques have made the transition from the laboratory to commercial heat exchangers. Most prominent applications of such technology transfer include airborne recuperators and regenerations, automotive heat exchangers, electronic cooling heat sinks, steam condensers (water cooled and air cooled), steam generators, and petroleum-product processing heat exchangers, among some others.

Moreover, new and emerging applications demand newer variations or manifestations of the enhancement techniques, which perhaps even require advanced manufacturing techniques. This is particularly the case with microscale exchangers, and which might even evolve into much smaller-scale variants. The associated research is also directed more toward characterization of the underlying enhancement mechanisms, so as to develop more phenomenologically scaled or flow-physics-based predictive correlations.

Broad reviews of developments in enhanced heat transfer are available (Thome, 1990; Bergles et al., 1991; Bergles, 1997; 1998; Manglik, 2003; Manglik and Bergles, 2004; Webb and Kim, 2005; Bergles and Manglik, 2013). Also, several journals, especially *Journal of Enhanced Heat Transfer*, which is entirely devoted to the dissemination of new research and developments in enhancement methods, along with *Heat Transfer Engineering*, *Journal of Heat Transfer*, *Journal of Thermal Science and Engineering Applications*, and *International Journal of Refrigeration*, among the numerous other publications, more often feature this technology.

4.17 HEAT PIPES

Sameer Khandekar

Introduction

The term "heat pipe" is a rather interesting way to describe a "synergistic engineering structure" which is equivalent to a material having a thermal conductivity greatly exceeding to that of any known material (Grover, 1966)—for all practical purposes, what it means is that a *heat pipe* is a thermal superconductor. We know that materials around us have different thermal conductivities, that is, the ease with which it transports thermal energy through it when subjected to a temperature gradient. For example, good thermal conductors commonly known are copper (with $k \sim 398$ W/mK) and aluminum (with $k \sim 185$ W/mK), while common insulating materials are wood (with $k \sim 0.1$ W/mK) or Teflon (with $k \sim 0.35$ W/mK). Materials like stainless steel and brass have intermediate values. Exotic materials such as diamond may have very high thermal conductivities, usually exceeding about 2000 W/mK. Coming back to a "heat pipe" and its definition—it sounds like a "pipe" which, of course, carries heat and works like a superconductor; this, necessarily means that its thermal conductivity should be very large. How large?—is a valid question to ask at this point. To be of some worth, it must at least be an order of magnitude greater than, say the common engineering materials used for heat transfer applications, such as copper or aluminum. If such a synergistic engineering structure can indeed be achieved, at a "low" cost, probably also lower weight, and if it can operate "passively," then it certainly sounds a very reasonable proposition to pursue. Indeed, the idea of such a device was conceived in the mid-1960s, which we now know as a "heat pipe."

Before we proceed further, certain terms and concepts associated with heat pipes need some further clarification. First, heat transfer from a hot object (source) to a cold object (sink) can be achieved "actively" or "passively." By "active" method, we mean that some external power, such as an electrically driven fan, is used for flow of a coolant to achieve the required heat transfer. This is usually true for forced convection systems, which may or may not involve any phase-change phenomena. Passive thermal systems, on the other hand, do not need any additional external power source for their operation. The heat that needs to be transferred itself usually acts as the "driver" to help facilitate its own transfer. For example, under natural convection heat transfer, a density gradient is initiated due to the imposed temperature gradient, which leads to a free-convective motion, resulting in passive heat transfer. In a similar fashion, a heat pipe is a passive device. The associated thermal-hydraulic phenomena are driven by the imposed thermal gradient itself. Industrial applications tend to prefer passive systems for obvious advantages, primarily because of reliability and redundancy considerations. Second, it must be noted that a heat pipe is not a "cooling" device, it is a pure "heat transfer" device. A heat pipe is therefore not a "refrigerator," in a thermodynamic sense. It transfers heat in the natural direction only, from high-temperature source to low-temperature sink—but does it so efficiently that it can be termed as a thermal superconductor. There is no external "work transfer" crossing the boundary of a heat pipe.

The reverse of thermal conductivity is thermal resistance—superconductors have exceedingly low resistances. Drawing an analogy of heat conduction with an electrical resistance system, following Ohm's law, wherein the "emf" is the driving potential for the flow of electronic charge through a medium, we can write Fourier's law of heat conduction as

$$\dot{Q} = k \cdot A \left(\frac{\Delta T}{\Delta x} \right) = \left(\frac{\Delta T}{\left(\frac{\Delta x}{kA} \right)} \right) = \frac{\Delta T}{R_{\text{th}}} \tag{4.17.1}$$

where

$$\Delta T = T_{\text{source}} - T_{\text{sink}}\text{, which is analogous to,} \tag{4.17.2}$$

$$I = \Delta V / R_{\text{elec}} \tag{4.17.3}$$

As in the case of electronic transport under an applied voltage, the objective of a heat pipe, or for that matter, any thermal conductor, is to transport maximum amount of heat at a given operating thermal potential difference, thus operating at the lowest possible resistance. The reverse is true for an insulator, wherein the objective function is to minimize the net heat transport when the system is subjected to a thermal gradient.

At this stage, we have understood that a heat pipe is an excellent thermal conductor of heat. Hence, the ensuing net thermal resistance of engineering systems that incorporate heat pipes ought to be rather small. Another important aspect of understanding a heat pipe thermal system is to appreciate the difference between "heat load" and "operating temperature." By heat load, we mean the amount of heat which is handled by the system or which is being dissipated/expected to be dissipated by a given system, say for example, a power electronic component that generates heat due to I^2R_{elec} heating. In principle, the heat load can be removed from this component at any operating temperature which we desire. For a specified heat load and geometry of the system, as the temperature of the sink goes up or down (for whatever reasons, by design or during its operation), the source temperature will keep going up or down, respectively, to satisfy Equation 4.17.1. Hence, under such conditions, the operating temperature of the device (the source temperature) depends on the sink temperature. In this context, it should be appreciated here that heat load and the operating temperature are two distinct entities, linked to each other via the sink temperature and the effective thermal conductivity of the medium connecting the source and the sink. In this background, there are a few pertinent questions to ask ourselves (i) is the component safe and reliable at the operating temperature? (ii) will the heat pipe (or any other thermal conductor) used to transport heat, change its performance with changing operating temperatures? (iii) is there a maximum heat load beyond which the heat pipe will stop working as a superconductor thereby rendering itself to be practically useless—alternately, are there any limitations on the superconducting nature of the device? Naturally, the maximum allowable operating temperature of any system with specified heat dissipation is intrinsically linked to its thermo-physical material properties, functionality, and desirable operating characteristics. For example, typical semiconductor-based microelectronic systems fail, or start underperforming, at the die-level operating temperatures exceeding about 85°C. In the following sections, we will explore the answers to these questions.

What Is a Conventional Heat Pipe?

Let us now explore the physical construction of a basic (conventional) heat pipe and see how thermal superconductivity is realized in a practical device, viz. a "synergistically operating engineering structure." In realizing a standard heat pipe with the goal of creating a superconducting system, the term synergy refers to the collaborative interaction of at least four different concepts of thermophysics: (1) evaporation and/or boiling, (2) condensation, (3) porous media fluid flow, and (4) surface tension driven capillary pumping.

In general, a typical conventional heat pipe utilizes the continuous evaporation/condensation of a suitable working fluid for two-phase heat transport, utilizing latent heat in a closed system, as shown in Figure 4.17.1. A conventional heat pipe consists of three sections: (1) the evaporator section, wherein the heat load to be dissipated is attached; heat comes inside the heat pipe system from here, and hence, the existence of this section is mandatory (otherwise why are we here for anyway!), (2) an adiabatic section, where there is no thermal interaction of the pipe with the environment.

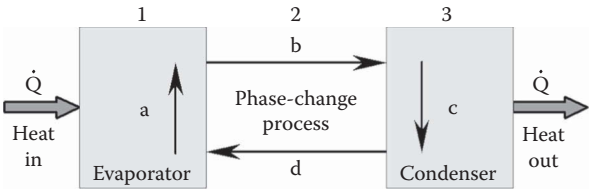


FIGURE 4.17.1 Block diagram showing the components of heat pipe and their respective functions.

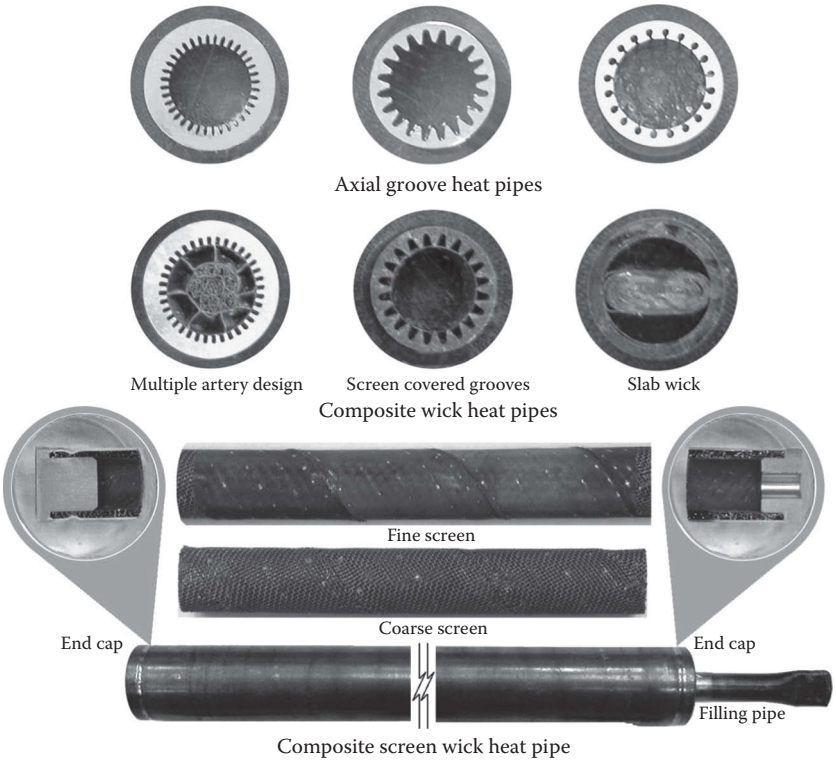


FIGURE 4.17.2 Various types of wick structures (top) and typical construction of a conventional heat pipe using screen wicks (bottom).

Depending on the system, this section may or may not be there. The existence of an adiabatic section, is, in principle, not essential for a heat pipe to operate, and (3) a condenser section, which serves as the heat sink—this is certainly an essential element of the system, as shown in Figure 4.17.1. The physical components of a conventional heat pipe, as shown in Figure 4.17.2, are as follows:

- A container, usually made of metal; glass, ceramics, and polymers can also be employed. For the process of heat transfer to be materialized, the very first resistance that needs to be overcome is the thermal resistance of the container material. Hence, unless there is a special requirement or purpose, good conducting materials are always preferable, at least from a thermal perspective.
- A capillary wick structure, made of sintered metal powders, screens wick or axial grooves cut longitudinally along the length of the heat pipe; woven fiberglass or metals; metal and carbon fibers can also be employed. Various composite wicks have also been used to enhance the performance; for example, axial grooves can be combined with screen wicks. Some designs of the internal wick structures are shown in Figure 4.17.2.

- A small amount of working fluid (which is sufficient to saturate the wick structure), which depends on the heat pipe operating temperature range (can vary from 4 to 2200 K).

For filling the heat pipe, the container is evacuated and then an appropriate amount of working fluid is administered into it. The container is then sealed and the heat pipe is ready for operation. Referring back to Figure 4.17.1, heat (Q) is added to the evaporator section (1) and thereby the working fluid, which is in equilibrium with its own vapor, gets vaporized (a). Driven by the pressure difference between the hot evaporator section and the cold condenser section, the vapor flows (b) to the condenser section (3) and is condensed there (c). During the phase-change processes, the fluid absorbs or releases, respectively, the latent heat of vaporization. The liquid returns (d) through the capillary structure from the condenser to the evaporator driven by capillary forces. By virtue of the capillary pumping action, the heat pipe can be operated in a microgravity field (as in satellites) or against gravity (on the ground with evaporator section located above the condenser section).

Needless to say that there should be good chemical compatibility between all the material components of a heat pipe, that is, container, wick structure, endcaps, etc. lest it will start a chemical reaction between them (say, e.g., the working fluid and the wick or the container), and in the process, unwanted chemical products/noncondensable gases (NCGs) will form inside the system. Additionally, for most heat pipe systems, a thorough sealing is required once the working fluid has been filled inside the evacuated pipe on both ends of the pipe. It is imperative that the soldering/brazing/welding material used for achieving the sealing is also compatible with the working fluid. Purity of the working fluid itself needs to be ensured so that foreign elements do not enter into the system; throughout the active life of the heat pipe, the working fluid inside must remain as a pure single-component two-phase system. (As will be introduced later, there are some specialized heat pipe systems in which a noncondensing gas is intentionally introduced for achieving the desired purpose; here, we are only discussing a standard conventional heat pipe in which NCGs are not tolerated under normal operation.)

It is also worthwhile to mention here that whenever a working fluid is introduced inside an evacuated pipe/container, it "flashes," that is, it undergoes an isenthalpic expansion process from the pressure at which it is stored. During this expansion process (which is a nonequilibrium process), a certain amount of vapor gets automatically generated until the final pressure inside the pipe, once the filling valve is shut, corresponds to the equilibrium pressure corresponding to the temperature at which the heat pipe is stored. If the entire heat pipe, after filling, is relocated to a chamber wherein the temperature is higher or lower, then the fluid pressure inside the pipe will change, and will always be corresponding to the saturation pressure corresponding to the temperature at which the heat pipe is stored. Remember, the phase equilibrium is achieved when the specific Gibbs energy of both the phases equalizes. The phase-change process continues till this condition is achieved. The very idea of the heat pipe operation emanates from the fact that we create a potential imbalance between the evaporator and the condenser, which the fluid cannot tolerate due to thermodynamic constraints and hence, continuously strives to bring the internal system to an equilibrium, thereby achieving our purpose of efficient and passive heat transfer.

The dimensions of conventional heat pipes range from a few centimeters in length and millimeters in hydraulic diameter, up to lengths of more than 10 m and hydraulic diameters of the order of many centimeters. Accordingly, the maximum heat handling capacity (before something drastic occurs because of which the heat pipe stops working and no more behaves as a superconductor) can range from μW to several kW. We will talk about these drastic effects, that is, the maximum heat-carrying capacity of heat pipes, or conversely, the heat transfer limitations in heat pipe operation a bit later.

Brief History of Heat Pipe Development

The discussion on the history of "conventional heat pipes," wherein, as stated earlier, surface tension forces are responsible for internal fluidic transport, must commence from the earlier development

of wickless, gravity driven, closed two-phase thermosyphons, which were patented more than 100 years ago. Such first generation passive devices, as shown in Figure 4.17.3a, are the simplest form of efficient heat transfer devices incorporating a liquid–vapor phase-change system. The system consists of a plain and simple tube (without any wick material inside), filled partially with a working fluid after evacuation, and subsequently sealed. As there is no dedicated and special mechanism for liquid condensate return, such as a capillary porous wick, these thermosyphons work under gravity conditions only. Hence, they are also referred to as "gravity-assisted heat pipes." The first closed two-phase thermosyphon design was patented in 1830 by Jacob Perkins (1766–1849, USA), the first in a row of four generations of ingenious engineers, and widely used in baking ovens to achieve uniform baking—in fact, they were referred to as "Perkins tube." Subsequently, the concept of looped thermosyphons, which are also referred to as "separate type gravity heat pipes" in contemporary times, as shown in Figure 4.17.3b, were patented by Ludlow Patton Perkins, together with W. E. Buck, in 1892, the former being the great-grandson of Jocob Perkins (Groll, 2013).

The story of the invention of conventional wicked heat pipes is a bit more interesting. It was R. S. Gaugler of General Motors Company who first described the use of wick structures, to transport liquid in a closed tube against gravity, in his patent of 1944 (Gaugler, 1944). However, his idea was never implemented by the company and the patent went into oblivion. In another independent study in 1962, Trefethen from Tufts University in Medford, MA, in an internal report for General Electric's Missile and Space Vehicle Department, wrote a rather detailed description on the feasibility of surface tension driven systems for space applications. This report also remained out of public access until 1995 (Groll, 2013). The name "heat pipe" eventually appeared in yet another independent (re-)invention due to George M. Grover at the Los Alamos National Laboratory (LANL, then called as Los Alamos Scientific Laboratory), USA, wherein the idea to employ the pumping action due to surface tension forces, for passive heat transport in an evaporation–condensation heat transfer device which could be of particular interest for application in space reactors, that is, in a microgravity environment, was described. Subsequently, the first scientific paper appeared in 1965, wherein a conventional heat pipe, driven by sodium as the working fluid, was described (Grover

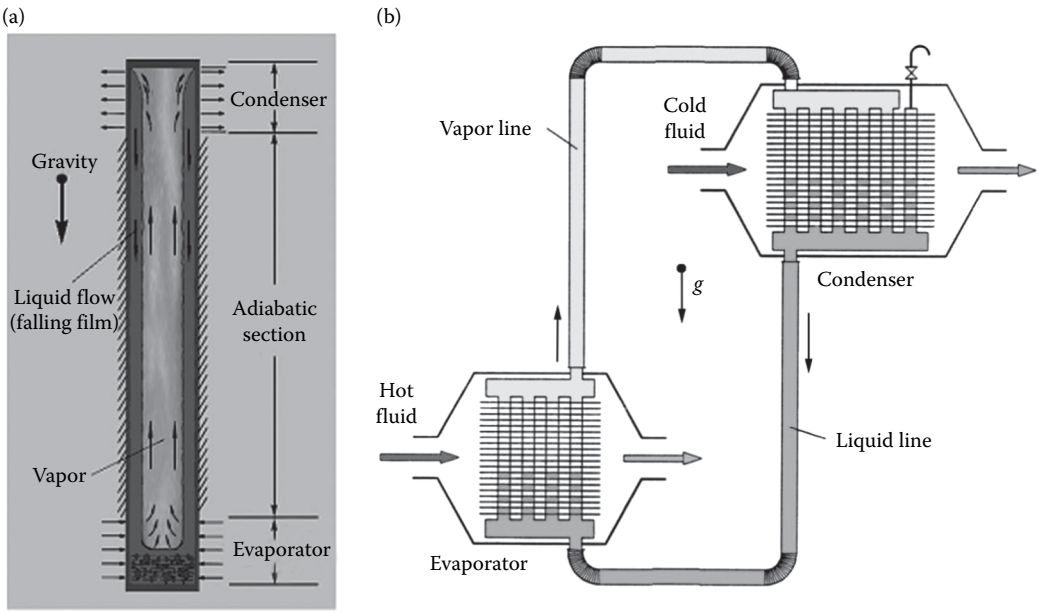


FIGURE 4.17.3 A classical gravity driven two-phase closed thermosyphon (a) and a separate type gravity-assisted heat pipe in which liquid–vapor interface shear is avoided (b).

et al., 1964). In the years which followed, the rapid development of the technology took place, with the primary focus on applications in Space and Nuclear energy sector. Percolation to terrestrial systems, employing two-phase thermosyphons and wicked heat pipes, soon followed with pioneering work carried out by NASA laboratories on the one hand, and several thermal and nuclear systems laboratories in Europe, especially United Kingdom, Italy, Germany, and erstwhile USSR. Today, several different varieties of heat pipes, in various forms, shapes, lengths, wick structures, working fluids, and special features have penetrated in a gamut of applications ranging from space, military, avionics, commercial electronics, medical devices, buildings and structures, power plants, railways and automobiles, deicing and permafrost stabilization, solar heating systems, isothermal furnace liners, blackbody radiators, etc. (Faghri, 1995, 2012, 2014; Groll, 2013; Vasiliev and Kakaç, 2013).

Other Types of Heat Pipes

In the preceding section, we have described a conventional wick-based heat pipe and its predecessor, a gravity-assisted heat pipe or a closed two-phase thermosyphon. In this section, we describe the nuances of the other two major types of heat pipes in wide use and development, that is, (1) loop heat pipes (LHPs) and (2) pulsating heat pipes (PHPs). Finally, we also touch upon two other types of heat pipes, that is, microheat pipes and variable-conductance heat pipes.

Loop Heat Pipes

While the conventional wicked heat pipes are excellent heat transfer devices, their thermal performance gets strongly influenced by the operating orientation. Of course, space application is comparable to horizontal operation on earth-gravity. One of the reasons for inferior performance is the fact that the wick structure is present throughout its length and this presents a substantial pressure drop for liquid return flow. Furthermore, due to the counter-current flow of the liquid and the vapor phase, interfacial interaction also adds to the total pressure penalty. In order to overcome these limitations, LHPs were invented in the early 1980s (Gerasimov et al., 1974, 1975; Maydanik, 2005). Approximately during the same time period, capillary pumped loop (CPL) designs were also evolving in the United States (Faghri, 1995). CPLs and LHPs are qualitatively similar in design; however, their dynamical response is somewhat different (Bazzo and Riehl, 2003; Pouzet et al., 2004; Joung et al., 2010; Boubaker et al., 2015). In this chapter, we will confine our discussion to LHPs, which are comparatively more popular than CPLs. LHPs also use the capillary action to remove heat passively from the evaporator to the condenser. However, in contrast to a conventional heat pipe, the main driving capillary wick structure is present only in the evaporator section. All other components (condenser, liquid return line, and vapor feeder line to the condenser) are made from simple smooth tubing/piping. Hence, net liquid-return pressure drop reduces drastically in comparison to the conventional heat pipes, and the additional viscous interaction between liquid and vapor phases is also absent in LHPs, as both phases move in independent tube sections. With these features and the possibility of using enhanced porous evaporator wicks, LHPs are proving to be substantially better in terms of thermal performance than conventional wicked heat pipes. Even under the adverse gravity conditions, LHPs outperform conventional heat pipes of comparable physical dimensions (Maydanik, 2005; Faghri, 2012, 2014).

A schematic diagram of an LHP is shown in Figure 4.17.4. It shows an evaporator, a condenser, an accumulator volume (compensation chamber), a vapor line, and a liquid line. The evaporator contains the porous wick cartridge. As heat is supplied to the evaporator, liquid vaporizes and menisci form at the liquid–vapor interface in the evaporator wick. Similar to a conventional heat pipe, it develops additional vapor pressure to push the vapor to the condenser through the vapor line. In turn, a capillary driving pressure is created across it which then passively drives the fluid inside the loop. The vapor condenses in the condenser and the capillary forces enable the liquid to return back to the evaporator. The evaporator is used not only as a heat sink to remove the heat, but also as a source to provide the total capillary pumping pressure, thus a highly efficient evaporator is the key to passive action and enhanced thermal performance of LHPs.

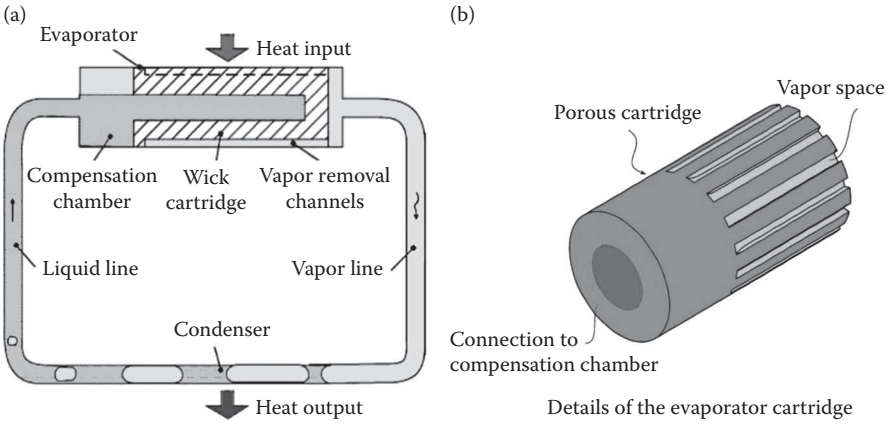


FIGURE 4.17.4 A schematic of an LHP (a) and a typical evaporator cartridge for a cylindrical evaporator (b).

Finely engineered porous wick structures are the key part of the LHP evaporator. This wick structure may be in the form of a cartridge, shown schematically in Figure 4.17.4. The combination of various characteristics like porosity, permeability, pore size, and the effective thermal conductivity of the wick control the thermal performance of LHPs. Typical ranges of porosity, permeability, and pore size of LHPs are 35%–75%, 10^{-12}–$10^{-13}\,m^2$, and 1–20 μm, respectively. Some of the required properties for the porous material of LHPs are: (1) small pore size diameter to ensure a large capillary driving force, (2) large porosity and permeability for low hydraulic resistance, and (3) low effective thermal conductivity to minimize the heat leak from the evaporator to the accumulator. Porous materials used in LHPs include pure metals, alloys, and polymers. Among these, polymers (polyethylene, polypropylene, PTTE) and plastics have the advantage of low-thermal conductivity, but maximum porosity is limited to about 50%. Wicks made of metal or alloys have comparatively higher porosity range but thermal conductivities are also relatively higher. Nickel is the most commonly used metal for evaporator wicks in LHPs because of its lower thermal conductivity and ability to be formed in fine pore sizes with high porosity. LHPs with nickel as the porous wick and ammonia as the working fluid are one of the most suitable combinations for satellite cooling applications. Some of the other metals or alloys used in LHPs include stainless steel, titanium, copper, brass, iron oxide and few composites wicks like nickel–chromium, nickel–copper. The use of copper for LHP capillary wick is also under active consideration in many terrestrial applications, although its thermal conductivity is quite high. Pore size plays an important role in LHP operation. Development of biporous wicks has also gained some ground as these wicks contain pore sizes of two different length scales. Fine pores provide higher capillary force and coarse pores provide higher permeability, which eventually reduces the hydraulic flow resistance. Heat transfer coefficient for biporous wicks is about two to three times more than the monoporous wicks. The most common technique for making LHP wick is via the sintering route; however, other possibilities have been pursued (Launay et al., 2007; D'Entremont and Ochterbeck, 2008; Vasiliev et al., 2009; Samanta et al., 2011; Adoni et al., 2012; Ambirajan et al., 2012; Maydanik et al., 2014).

Suitability for miniaturization, flexibility and modularity of design, possibility of using multiple evaporators, separation of vapor and liquid interfacial interaction, extremely high capillary performance, possibility of providing a flexible coupling between source and sink, and robustness, are some of the positive features of LHPs which qualify them as a clear choice in space missions (Gilmore, 2002; Hoang and Ku, 2002; Kaya and Ku, 2003; Ku et al., 2012). In fact, the aerospace industry is also seriously considering its deployment for thermal management of commercial aircrafts (Bugby et al., 1998; Wang et al., 2008; Sarno et al., 2013). With an improvement in the ability to control the porous wick characteristics through various manufacturing routes, LHPs are expected to cater to a large segment of aerospace thermal management as well as terrestrial and defense

related applications. Adequate literature on modeling is also evolving, although this remains a topic of active research (Siedel et al., 2015). Combined LHP and PHP designs are also under active consideration (Vasiliev et al., 2009; Pastukhov and Maydanik, 2013).

Pulsating Heat Pipes

A PHP, introduced by Akachi (1990, 1996), consists of a plain meandering tube of capillary dimensions with many U-turns as shown in Figure 4.17.5. In this design, in contrast to a conventional heat pipe, there is no additional capillary structure inside the tube. There are two ways to arrange the tube, (1) open loop and (2) closed loop, configuration. As the names suggest, in a closed-loop structure, the tube is joined end to end, as detailed in Figure 4.17.5. The tube is first evacuated and then filled partially with a working fluid, which distributes itself spontaneously in the form of liquid–vapor plugs and slugs inside the capillary tube. One end of this tube bundle receives heat, transferring it to the other end, by the pulsating action of the liquid–vapor/bubble-slug system. There may exist an optional adiabatic zone in between. Also, one or more flow direction control check valves may be introduced at suitable locations to augment the performance (Khandekar, 2004). PHPs can also be developed as flat structures, on to which, electronic circuits can be directly mounted. PHPs have also been embedded on planar metallic/polymer space radiator structures to enhance the effective thermal conductivity of the resulting composite structure; such planar structures have been extensively considered for various applications (Khandekar et al., 2002; Yang et al. 2009; Hemadri et al., 2011; Karthikeyan et al., 2014; Ayel et al., 2015). Figure 4.17.6 shows a range of industrial prototypes of PHPs which have been developed in recent times.

A PHP is a complex heat transfer device with a strong thermo-hydraulic coupling governing its performance. It is essentially a nonequilibrium heat transfer device. The successful performance of the device is primarily due to the continuous maintenance or sustenance of these nonequilibrium conditions within the system. The liquid and vapor slug transport is because of the pressure pulsation caused in the system. Since these pressure pulsations are thermally driven, because of the inherent construction of the device, there is no external mechanical power source required for the fluid transport. In an actual working PHP, there exists a temperature gradient between the evaporator and the condenser section. Further, inherent perturbations are always present in real

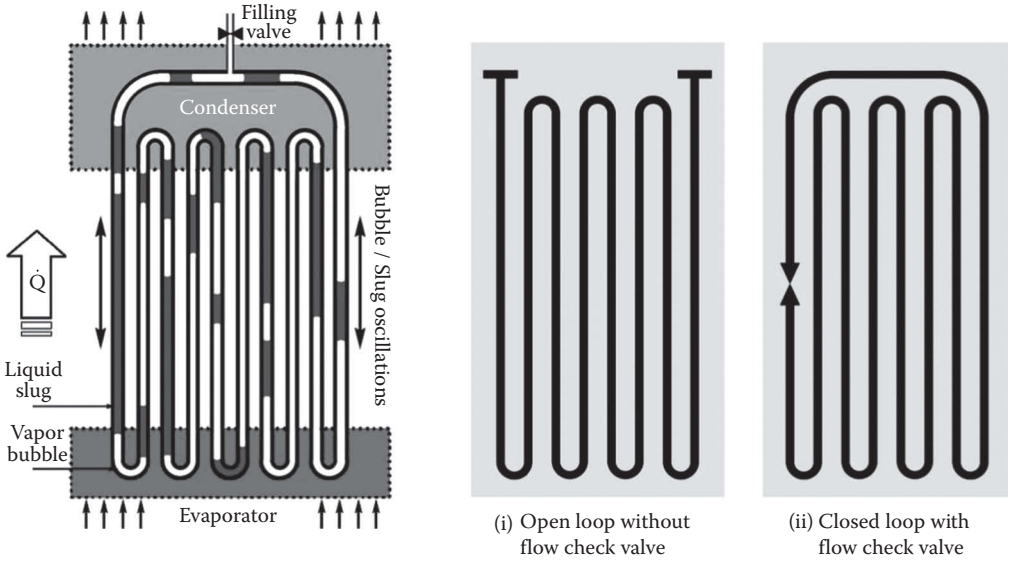


FIGURE 4.17.5 Schematic of a closed-loop PHP without any check valve (left) and two typical operating configurations (right (i) and (ii)).

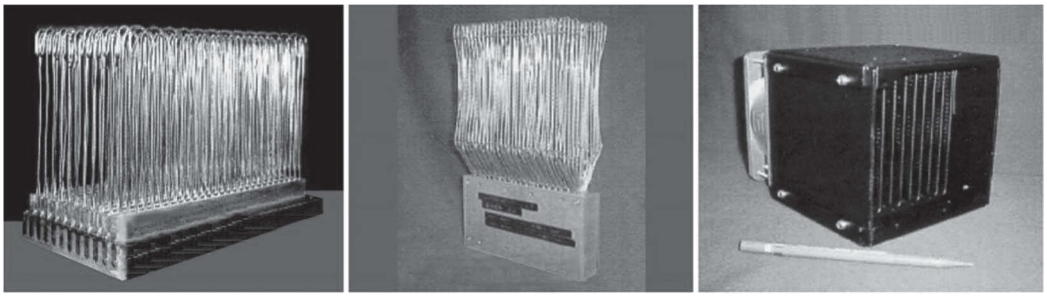


FIGURE 4.17.6 Different engineering prototypes of PHPs which have been designed and tested in recent times.

systems (Khandekar et al., 2003). These perturbations are due to: (1) pressure fluctuations within the evaporator and condenser sections due to local nonuniform heat transfer always occurs in real systems, (2) asymmetrical liquid–vapor distributions causing uneven void fraction in the tubes, and (3) the presence of approximately triangular or sawtooth alternating component of pressure drop superimposed on the average pressure gradient in a capillary slug flow due to the presence of vapor bubbles (Khandekar, 2004).

The net effect of these temperature gradients within the system is to cause nonequilibrium pressure condition which, as stated earlier, is the primary driving force for thermo-fluidic transport in a PHP. A sustained nonequilibrium state exists between the driving thermal potentials and the natural causality which tries to equalize the pressure in the system. Thus, a self-sustained thermally driven oscillating flow is obtained in a PHP. Hence, it is noted that there is no "classical steady state" possible in a PHP operation. Instead, the pressure waves and pulsations are generated in each of the individual tubes which interact with each other possibly leading to secondary and ternary reflections with perturbations. From a force balance of gravity and surface tension leading to the definition of Bond number, the theoretical maximum tolerable inner diameter (D_{crit}) of a PHP capillary tube can be calculated as $\mathrm{Bo}_{crit} \leq 2$, where the Bond number is defined as

$$\mathrm{Bo} = D\sqrt{g\left(\rho_l - \rho_v\right)/\sigma} \tag{4.17.4}$$

where ρ is the density of the involved phases and σ is the liquid–vapor surface tension. At diameters below this value, there is a tendency of surface tension forces to dominate and this assists in the formation of the stable liquid slugs under nonoperating conditions, a prerequisite for PHP operation.

Numerous experimental investigations have been done with the aim of a parametric study of a PHP. These studies indicate the following main variables affecting the PHP performance (Charoensawan et al., 2003; Khandekar et al., 2003; Khandekar, 2004; Khandekar and Groll, 2004a):

- *Geometric variables*: Overall length of the PHP, diameter/size and shape of the tube, length, and number of turns of evaporator/condenser/adiabatic section.
- *Physical variables*: Quantity of the working fluid (filling ratio), physical properties of the working fluid, tube material.
- *Operational variables*: Open-loop or closed-loop operation, heating and cooling methodology, orientation of the PHP during operation, use of check valves, etc.

It is evident that there are multiple variables that simultaneously affect the operation and performance of PHPs. Further, so far as capillary slug flow exists inside the entire device, it has been demonstrated that the latent heat will not play a significant role in the device performance. This is a major difference as compared to a conventional heat pipe that primarily transfers only the latent

heat component of the working fluid. Nevertheless, in PHPs, bubbles are certainly needed for self-sustained thermally driven oscillations. On the other hand, under certain operating conditions, flow may transit to annular flow regime during the PHP operation. The probability of such an event is high with a combination of high Bond number, high heat flux, and a comparatively low liquid filling ratio (~50% or lower). If the flow regime changes from slug to annular, the respective roles of latent and sensible heat transport mechanism may considerably change (Khandekar, 2004; Khandekar and Groll, 2004b; Zhang and Faghri, 2008). Further, the dry-out mechanisms of PHPs are not similar to the conventional heat pipes; this remains a topic of continued research (Yang et al., 2008; Khandekar et al., 2009; Tang et al., 2013; Han et al., 2016).

It is clear that the performance not only depends on a large number of parameters but also on the flow pattern. This makes it all the more difficult to undertake mathematical modeling using conventional techniques. Thus, in order to discern the flow physics, it is logical to split the PHP into different simpler well-posed problems, study their individual behavior and then make a coupling to have a comprehensive understanding of the complex PHP thermo-physics. A few of these simple thermal-hydraulics problems that give insight to the PHP operation are summarized here:

- Isolated studies of vapor–liquid interfaces in confined geometries, including local phase-change mechanisms at the interfaces (Rana et al., 2014; Srinivasan et al., 2015a,b).
- Hydrodynamics of isolated Taylor bubble and continuous Taylor bubble-train flow of two-component two-phase systems can be investigated in minichannels, during upward, downward, and horizontal configurations. No heat transfer and phase change are involved in this analysis; such studies provide the relevant information of frictional pressure drops, local flow patterns inside liquid slugs, interfacial shapes and ensuing enhanced transport properties (Bajpai and Khandekar, 2012; Majumder et al., 2013).
- External pulsations of different frequencies are added to continuous Taylor bubble train, which will make the Taylor bubble system to pulsate and effect of pulsations on hydrodynamics is studied (Mehta and Khandekar, 2014a,b).
- After studying hydrodynamics, the parametric study of heat transfer in these systems is analyzed without the phase-change process being taken into account (Mehta and Khandekar, 2014b).
- Finally, phase change can be incorporated in the Taylor bubble flow, under both unidirectional and oscillating conditions, to map the flow with the functioning of real-time PHPs (Bonnenfant et al., 2012; Rao et al., 2013, 2015; Chauris et al., 2015).

This approach of splitting the problem has enriched the understanding of several new parameters that were not previously accounted for in the modeling of PHP, some of which are:

- *Contact angle hysteresis*—The additional pressure drop associated with the hysteresis phenomenon plays a major role in the hydrodynamics of the system.
- *Vortex circulation inside confined liquid Taylor plugs*—Due to the presence of the two interfaces, local 3-D flow is induced. These internal vortices enhance the ensuing heat transfer.
- *Wettability*—The ability to synthesize different wetting surfaces, give way to different hydrodynamics and heat transfer behavior, enhancing design possibilities of PHPs.
- *Wet flows and dry flows*—Depending on the existence of liquid-film surrounding the vapor phase, the flow can be categorized as wet or dry. In the wet flow, the liquid-thin film accounts for most of the heat transfer, due to its low-thermal resistance. In dry flows, the three-phase contact line motion accounts for much of the dissipation, and the pressure drop is higher in this case than that in the wet flows. The fact that PHPs have both wet and dry flows occurring simultaneously during its working makes its mathematical modeling challenging.

Thus, several insights into some of the fundamental issues of flows in mini/micro channel are revealed and understood, which has not only facilitated the mathematical description of the PHP internal processes, but also other engineering systems which rely on Taylor bubble flows, such as flow boiling heat exchangers for electronics cooling, lab-on-chip devices, micro-reactors, and fuel cells (Khandekar et al., 2010). Mathematical modeling of PHPs through "first" principles is a contemporary problem which remains quite elusive. Simplifications and assumptions made in all the modeling approaches developed so far render them unsuitable for engineering design because the idealizations contained therein are not in sync with the real time transport processes taking place in the device. At present, there are no theoretical models or correlations available that can predict the complex thermo-hydrodynamic transport behavior of PHPs. This prevents their widespread use in industrial applications. Several important conclusions can be made based on the recent literature from the last decade, some of which are elaborated below:

Thin film evaporation and condensation plays a crucial role in the overall dynamics. Unless this is included in the global model, realistic predictions cannot be made. Corollary to this is the fact that unless mass addition/removal from the vapor space is considered, the amplitude of oscillations observed in actual experiments cannot be predicted. Treating the vapor as an ideal gas with only sensible heating/cooling is not sufficient (Mameli 2012; Rao et al., 2013, 2015; Mameli et al. 2014).

The frictional pressure drop required to sustain the flow, especially in the liquid plugs with small *L/D* ratios, may be substantially higher than what is predicted by assuming a simple Poiseuille flow inside them. The fluid flow pattern in the oscillating two-phase "unit-cell" system (i.e., one Taylor bubble with adjoining liquid plug) is rather complex leading to high viscous dissipation. Moreover, three-phase contact lines tend to form inside the system, especially in the evaporator subsection, which lead to very high local dissipation (Janeček and Nikolayev, 2013; Srinivasan et al., 2015a,b). As *L/D* ratio increases, classical Poiseuille friction is sufficient to describe the flow.

Inertia forces play a significant role here. Sudden acceleration and deceleration of the "unit-cell" and its interaction with surface and capillary forces lead to complex flow phenomena, including breakage of liquid plugs into smaller fragments. While most hydrodynamic theories of visco-capillary flows are valid for low capillary number situations (less than ~10^{-3}) and high wettability conditions, engineering scale PHP devices deviate appreciably from such ideal situations. Partial wetting and relatively large capillary numbers are routinely encountered (Khandekar et al., 2010; Rana et al., 2014; Mehta and Khandekar, 2014a).

The radial heat transfer to/from the liquid–vapor system must happen via the capillary tube material. Unlike a conventional heat pipe that attains a quasi-steady state, the transient oscillatory flow patterns inside a PHP manifests itself as a conjugate (conduction–convection) problem. Hence, the interplay of diffusional and convective timescales of the system strongly affects the ensuing thermal-hydraulic response. In effect, thermal properties of the tube material are vital in predicting the overall system dynamical response (Nikolayev, 2013, 2016; Rao et al. 2013, 2015).

It has been conjectured that nonequilibrium states of liquid and vapor coexist during the normal operation of PHPs. Conclusive experimental evidence is yet to come; however, initial results do point toward this fact. It requires very precise and repeatable dynamic measurements of the local vapor and fluid film temperature so as to ascertain nonequilibrium condition and/or metastability. At one extreme is the assumption that vapor and liquid in the Taylor bubbles are always under thermodynamic equilibrium. This essentially means that the relaxation timescales of the phase change by heat addition/subtraction are much faster than the typical characteristics timescales of the internal fluid flow and heat transfer through the walls in the PHP system. However, the essence of the entire PHP thermal phenomena lies in its transient thermo-fluidic characteristics. While one can talk about a "quasi-steady" state in a global sense, nothing is steady on the local level inside an operating PHP. Hence, the assumption of thermodynamic equilibrium between adjacent vapor bubbles and liquid plugs may not be justified. The rapid inertial transient flow, vis-à-vis evaporation/condensation timescales, may lead to situations wherein superheated vapor and/or subcooled liquid coexists for finite timescales (Khandekar 2004; Nikolayev, 2013; Rao et al., 2015).

It is clear from the above discussion that PHPs continue to be a subject of active research, both from experimental as well as theoretical standpoint (Groll and Khandekar, 2002b; Zhang and Faghri, 2008; Khandekar et al., 2010; Tang et al., 2013; Han et al., 2016).

Micro Heat Pipes

Stringent cooling requirements required in the electronics industry concerning miniaturization and higher heat fluxes have led to the introduction of a micro heat pipe by Cotter (1984), as shown in Figure 4.17.7. Micro heat pipes are very small, wickless heat pipes having internal geometries containing sharp angled corner regions responsible for creating the required capillary pressure and also serve as liquid arteries. There are two supplementary definitions of a micro heat pipe. According to the original definition of Cotter (1984), the hydraulic radius R_h of a micro heat pipe is of the order of (greater than or equal to) the capillary radius R_c of the vapor–liquid interface. Subsequent definition of Chen et al. (1992) reflects not only the geometry, but also the physical behavior—it defines a micro heat pipe as a heat pipe for which the Bond number is small typically less than 2. Typical cross-sectional dimensions of micro heat pipes range from some 10 to 100s of μm with heat transport capabilities of some 10s to 100s of mW (Itoh and Polášek, 1990).

A sizable amount of experimental results on individual micro heat pipes as well as micro heat pipe arrays are available. In addition, quite a few papers that refer to "micro" heat pipes really do not correspond to the true "micro" heat pipes as defined here. There exist various hybrid and transitional heat pipe designs/structures in the "meso" or "mini" range too. Some of the variants are without a wick structure, thereby fully relying on the sharp-angled corners for generating capillary pumping, while others can have some form of wicks/microstructured surfaces (Groll and Khandekar, 2002a). Over the years since 1984, periodic reviews on micro heat pipes have appeared covering critical aspects of operations, performance characteristics and limitations, applications and mathematical

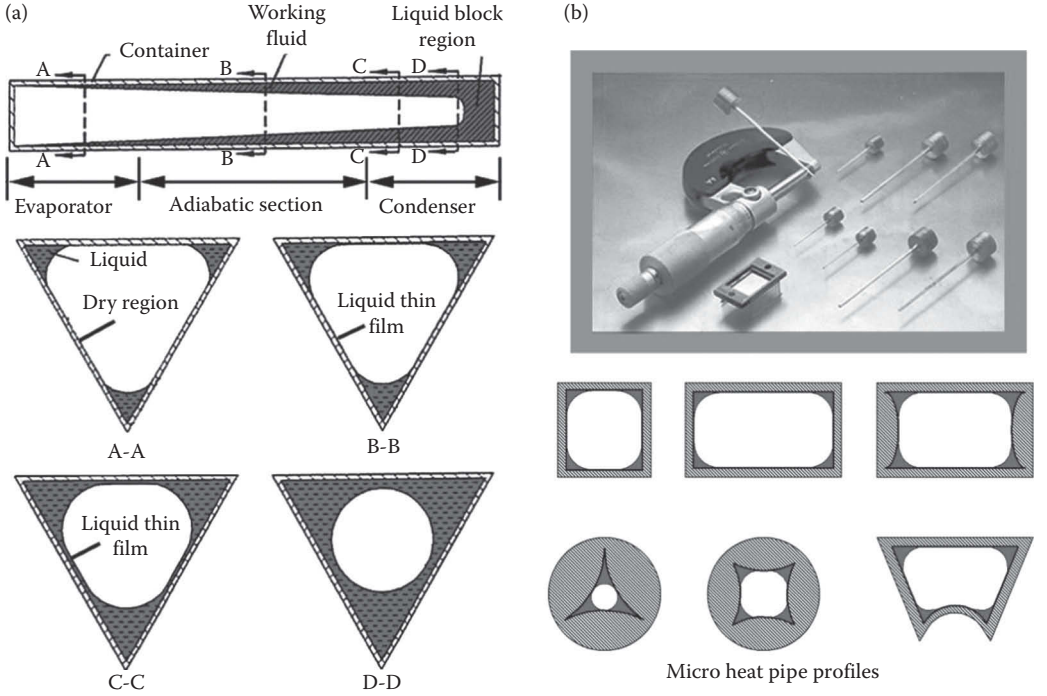


FIGURE 4.17.7 Concept of a micro heat pipe, showing axial and radial cross sections (a) and typical application (b); some possible cross-sectional shapes which provide capillary action due to the presence of sharp angled corners is also depicted below.

modeling (Groll and Rösler, 1992; Cao et al., 1993; Khrustalev and Faghri, 1994; Peterson, 1996, 1997; Groll et al., 1998; Ochterbeck and Peterson, 1999). In the last two decades, there has been a steady decline in the research activity on micro wickless heat pipes, primarily due to the fact that its overall heat handling capacity is rather limited. However, there is indeed sizable amount of fundamental research work which has been carried by various groups (e.g., see Wayner, 1991; Das Gupta et al., 1995; Karthikeyan et al., 1998; Zheng, 2002; Panchamgam et al., 2008; Chatterjee et al. 2011) to understand the microscale evaporation and condensation on a single bubble trapped in a confined tube, which essentially is analogous to a micro heat pipe operation.

The first model of a micro heat pipe was proposed by Cotter (1984) which, although based on simplified assumptions, was quite effective to predict the general qualitative aspects, however, the maximum heat transport capacity was severely over-predicted, as compared to the then experimental results (e.g., Babin et al., 1990; Chen et al., 1992; Swanson and Peterson, 1995; Ha and Peterson, 1998). It was concluded that apart from simple numerical factors and the extreme difference in the typical size, the design considerations for wickless micro heat pipes are essentially the same as those for conventional wicking limited heat pipes. This model led to subsequent refined models, as described by Groll and Khandekar (2002a). Subsequently, Gerner et al. (1992) and Longtin et al. (1994) developed a scaling analysis to investigate the operating limits of a micro heat pipe. Two distinct phenomena were identified, capillary limitation (similar to conventional heat pipes) and dry-outs attributable to interfacial instabilities. Porous media approach with Darcy's equation for two-phase flow imposed by Laplace's capillary equation has been adopted by Chen et al. (1992), Wang et al. (1994), and Tio et al. (2000). This approach has also given satisfactory predictions. On similar lines, Shukla (2009) has also presented a micro heat pipe model.

Variable Conductance Heat Pipes

In many applications, temperature control of the evaporator section (the thermal load) is desirable, irrespective of the variations in the input heat conditions or the thermal conditions existing in the condenser (sink) section. Such applications are frequently encountered in space vehicle applications and satellite thermal management (Karam, 1998; Gilmore, 2002). A conventional heat pipe can be suitably modified to achieve precise temperature control, with or without an active feedback control loop (Bienert and Brennen, 1971; Kirkpatrick and Marcus, 1971; Kirkpatrick, 1973; Groll and Kirkpatrick, 1976). The innovative idea in such devices, generically termed as variable conductance heat pipes (VCHPs) is to change the effective thermal conductivity of the structure, as per the demand. The simplest, nonfeedback controlled, VCHP is as shown in Figure 4.17.8a—such devices were developed as early as 1965 and the space industry gave enormous impetus for their development in the seventies (Gilmore, 2002; Groll, 2013). In such devices, the aim is to maintain the vapor temperature (and thereby essentially also the evaporator temperature) at a constant level, irrespective of heat input and boundary (heat sink) conditions. A standard VCHP consists of inclusion of an NCG buffer inside the system, attached to the condenser section. During normal operation, this NCG tends to collect in the condenser area where the internal fluid pressure is minimum. Depending on the location of the NCG front inside the condenser, the exposed internal area available for vapor condensation is determined. Under normal heat loads, the heat pipe is designed in such a way that the NCG front covers approximately half the condenser area. If thermal load at the evaporator increases, the NCG front is pushed toward the reservoir due to increased vapor pressure. This exposes more area available for vapor condensation, thereby improving the performance of the heat pipe. Hence, the additional thermal load gets accommodated without a corresponding increase in the vapor temperature. Alternately, decrease of heat load pushes the NCG front deeper inside the condenser and blocks vapor condensation. This increases the effective thermal resistance of the heat pipe and hence, for a constant sink temperature, the vapor temperature tends to remain constant. In other words, the reduced condenser area will become just sufficient to maintain the same difference between vapor (condenser) temperature and the sink temperature, and thus keep the vapor temperature constant. In this way, a nonfeedback loop VCHP can be designed and operated

to maintain the vapor temperature within a prescribed limit. While the vapor temperature can, to some degree be kept constant, the more important heat source temperature (i.e., temperature of the component which has to be cooled) will inevitably rise with the increasing heat input due to thermal conductance and resistance between component and heat pipe. To maintain the heat source temperature constant, the increasing temperature difference (for increasing heat load) between heat source and heat pipe must be compensated by a corresponding reduction of the vapor (evaporator surface) temperature. To accomplish this, a feedback control system, active or passive, has to be employed, as shown in Figure 4.17.8b. In the active method, the temperature of the heat sink is sensed, which is linked with an electric heater attached to the wicked reservoir. The heater is activated for decreasing or deactivated for increasing heat input (shut down for maximum heat input). Thereby, the NCG blocks the condenser to such a degree that the required vapor temperature is established to maintain the source temperature constant, viz. large temperature difference between heat source and vapor, ΔT_{HS-v}, for high heat load and small temperature difference for low heat load. Other passive thermal control mechanisms and heat pipe diode designs are also available (Faghri, 1995; Gilmore, 2002).

Other Types of Heat Pipes

Depending on the driving force exploited to run the passive phase-change thermal system, various concepts of "heat pipes" have been proposed and indeed incorporated in real-time thermal systems. As we have seen, while classical heat pipes described earlier make use of capillary forces for liquid

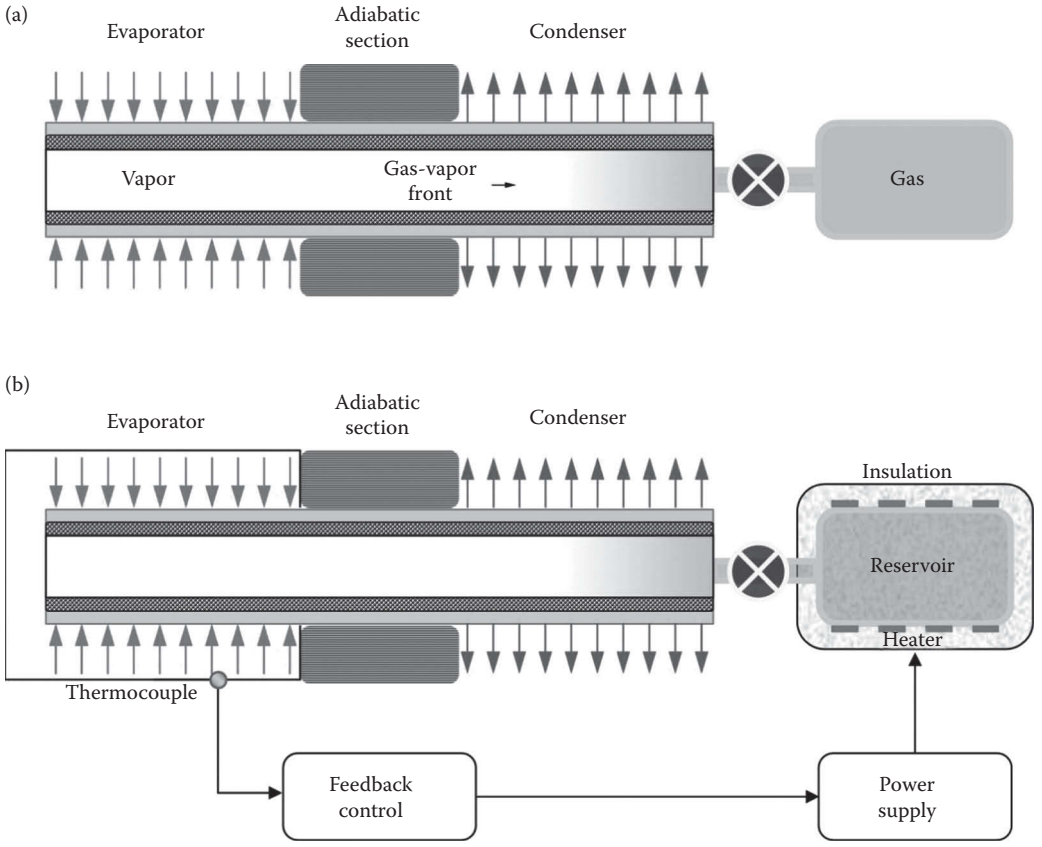


FIGURE 4.17.8 VCHP without feedback system (a) and VCHP with feedback control system (b). Depending on the heat input, the gas–vapor interface position is controlled to vary the exposed available area in the condenser.

return, wickless thermosyphons need gravity force field. When rotating systems needs thermal managements, such as electric motors or power-turbines, centrifugal forces can also be used to make "rotating" heat pipes. Such systems do not require a capillary wick (Bertossi, 2012). If an external electric or magnetic field is used for driving the fluid, this gives rise to electrohydrodynamic flow. Such external fields can also be coupled with capillary forces for enhanced heat transfer. Heat pipe action can also be coupled with sorption systems (Vasiliev and Vasiliev, 2005; Vasiliev and Kakaç, 2013). Concentration gradients can also be effectively harnessed to form electroosmotic heat pipes. In recent years, there is also an effort to incorporate nanofluids (small to moderate concentration of nanometric size metallic/nonmetallic/ceramic powders stably dispersed in conventional working fluids, to get enhanced or tailor-made thermo-physical properties) in heat pipes. At present, however, there are conflicting results regarding the efficacy of such fluids in heat pipe and thermosyphon systems; more focused research in this area needs to be undertaken for clarity (Khandekar et al., 2008; Buschmann, 2013; Sureshkumar et al., 2013).

Conventional Heat Pipe Design and Limitations

We will restrict our attention to the design and limitations of conventional wicked heat pipes in this section. In Section 1, we asked ourselves a question—What are the transport limitations of heat pipes? Obviously, a given heat pipe, with a certain type and geometry of wick structure, a certain length, internal diameter and working fluid, cannot maintain its "superconductivity" at all operating temperatures and/or input heat fluxes. In other words, at given operating temperature of a heat pipe, as we continue to increase the input heat flux to the pipe, there comes a stage wherein its proper operation of a "superconductor" will either degrade or completely cease. Indeed, this stage it is termed as a "dry-out" condition of heat pipe. There are two questions that are relevant for heat pipe dry-out conditions (1) Whether the maximum heat transport possible from a given heat pipe (dry-out heat power or dry-out flux) depends on the operating temperature? and, (2) Which thermo-physical phenomena is responsible for transforming a satisfactorily performing heat pipe, to one that degrades/stops its desired heat transfer operation? Intuitively, one can argue that since the thermo-physical properties of a working fluid depend on the temperature, with nearly all properties substantially varying between the freezing point and the critical point, dry-out conditions must depend on the operating temperature. For example, the viscous stresses scale with dynamic viscosity, the capillary pumping action scales with surface tension and phase velocities will depend on the respective fluid densities. Hence, the maximum heat flux ought to depend on the operating temperature of the heat pipe. As regards the type of phenomena responsible dry-out, there are quite a few different factors responsible for it. Interestingly, which factor predominates the occurrence of a dry-out condition is also a function of the operating temperature. The most common ones are: (1) viscous limitation, (2) sonic limitation, (3) capillary limitation, (4) entrainment limitation, and (5) boiling limitation. We will describe these "phenomena" or limitations briefly hereunder, giving an analytical expression to estimate the dry-out flux:

1. *Viscous limitation*: Imagine a situation when the net pressure drop required to transport the vapor from the evaporator end of the heat pipe to the condenser end becomes comparable to the absolute value of vapor pressure in the evaporator section. Such a situation may occur typically at low operating temperatures, for those working fluids whose saturation pressures at the operating temperature is rather low. In such a situation, it will be impossible for vapor to flow in the heat pipe because the absolute pressure available in the evaporator cannot overcome the dissipation required to send the vapor to the condenser. In other words, the condenser pressure cannot fall below absolute zero pressure! This is called as vapor pressure limitation or viscous limitation of a heat pipe. It can be calculated by using the expression (Dunn and Reay, 1982):

$$\dot{Q}_{\max} = \left(\frac{\pi}{12}\right)\left(\frac{h_{fg}\,\rho_{v,e}\,P_{v,e}}{\mu_{v,e}}\right)\left(\frac{r_v^4}{l_{eff}}\right) \tag{4.17.5}$$

where r_v is the cross-sectional radius of the vapor core (m), h_{fg} is the latent heat of vaporization (J/kg), $\rho_{v,e}$ is the vapor density in the evaporator (kg/m^3), $P_{v,e}$ is the vapor pressure in the evaporator (Pa), and $\mu_{v,e}$ is the vapor viscosity in the evaporator (Pa·s). l_{eff} is the effective length of the heat pipe (m) equal to $l_{eff}=0.5(l_e+2l_a+l_c)$.

2. *Sonic limitation*: The sonic velocity in a gas/vapor scales with the square root of the temperature. At low temperatures, the speed of sound in the fluid medium is correspondingly low. As the input heat flux increases, especially during startups at low operating temperatures, the velocity of vapor flowing in the vapor core of the heat pipe increases. If, under the operating conditions, the vapor velocity reaches sonic speeds, the heat pipe undergoes a "choking" condition. In a sense, the steady-state operation of a normally operating heat pipe can be compared with a gas flow in a nozzle-diffuser arrangement. In the evaporator section, the vapor flow accelerates, attains constant velocity in the adiabatic section and then, as condensation proceeds, there is deceleration of the fluid. Drawing an analogy, the evaporator acts as a nozzle, while the condenser section corresponds to a typical diffuser. When sonic limitation comes due to sufficiently high mass flow rate of the vapor, choking conditions will be experienced in the adiabatic condition, restricting the ability of the pipe to transfer more heat. Sonic limitation is given as (Dunn and Reay, 1982)

$$\dot{Q}_{\max} = 0.474\left(A_v\,h_{fg}\right)\left(\rho_v P_v\right)^{1/2} \tag{4.17.6}$$

where A_v is the cross-sectional area of the vapor core (m^2).

3. *Entrainment limitation*: The vapor and liquid phases flow in opposite directions in a conventional heat pipe, which creates interfacial shear stress. This additional shear is over and above the internal viscous shear in the respective phases. The interfacial shear increases as the input heat flux to the heat pipe increases due to the enhanced relative velocity of the two fluids. A situation may arise while operation when the interfacial shear is so high that small droplets of liquid are "scooped" off or "entrained" by the high vapor flow, depleting the evaporator of the returning liquid. Such a situation may lead to liquid deprived evaporator thereby leading to a total dry-out or a partial dry-out condition. Entrainment limitation is a manifestation of vapor inertia forces exceeding the stabilizing effects of the surface tension on the liquid–vapor interface, the corresponding nondimensional scaling these two effects being the Weber number. The entrainment limitation can be estimated from (Dunn and Reay, 1982)

$$\dot{Q}_{\max} = A_v\,h_{fg}\sqrt{\frac{\rho_v\sigma_l}{2r_{cap,ave}}} \tag{4.17.7}$$

where σ_l is the surface tension (N/m) and $r_{cap,ave}$ is the average capillary radius of the wick. Note that for many applications, $r_{cap,ave}$ is often approximated by $r_{cap,e}$.

4. *Capillary limitation*: This is one of the most important limitation of a heat pipe. As has been discussed, the capillary pumping action provided by the wick structure is at the heart of successful heat pipe operation. This capillary "pump" has to overcome all the dissipative pressure drops which are encountered by the circulating fluid so as to maintain the heat pipe action. The primary dissipative components include the frictional pressure drop required to sustain the flow of liquid in the porous wick and the frictional pressure drop required to be overcome for the vapor flowing in the core. In addition, as has been explained earlier, since liquid and vapor flow in the opposite direction during heat pipe operation, there is interfacial shear stress due to interaction of the phases at the interface.

This additional interfacial shear may be important in the case of axially grooved heat pipes where the interface is exposed, while in the conventional sintered powder or screen wicks, the impact is somewhat less due to interface being protected by the porous wick. The acceleration pressure drop of vapor in the evaporator is usually neglected as there is some partial recovery of this loss in the condenser area due to deceleration effects. Due to the orientation of the heat pipe (in terrestrial conditions), gravitational body forces may also be present, helping the capillary pump when the evaporator is below the condenser and vice versa. The capillary limitation thus occurs, when the capillary pump fails, or in other words, is not able to overcome the net losses due to fluid motion. Since frictional pressure drop in the system scales with flow velocity, at a certain critical heat flux, the corresponding flow velocity of the fluid is such that it makes the capillary pump ineffective. Under these conditions, the heat pipe is said to have "dried out," due to capillary limitation. For most heat pipes, the maximum heat transfer rate due to the capillary limitation, considering simple linear theory, can be expressed as (Dunn and Reay, 1982; Faghri, 1995):

$$\dot{Q}_{\max} = \left[\frac{\sigma \rho_l h_{fg}}{\mu_l}\right]\left[\frac{A_w K}{l_{eff}}\right]\left[\frac{2}{r_{cap,e}} - \left[\frac{\rho_l}{\sigma}\right] g L_t \sin\beta\right] \tag{4.17.8}$$

where K is the wick permeability (m^2), A_w is the wick cross-sectional area (m^2), ρ_l is the liquid density (kg/m^3), μ_l is the liquid viscosity (Pa-s), $r_{c,e}$ is the wick capillary radius in the evaporator (m), g is the acceleration due to gravity ($9.8\,m/s^2$), and L_t is the total length of the pipe (m). The quantity in the first parenthesis on the RHS has the units of heat flux and is called as the "Figure of merit," as explained later in Section 5.1. The estimation of permeability K is not so straightforward for heat pipe wicks. The interdependency of permeability on the porosity ε is highly nonlinear depending on the type and morphology of the porous structure. Although porosity is a geometrical concept, and analytical expressions are indeed available for simple screen wicks or axial grooved wicks. Often, it needs to be experimentally determined. Similarly, use of analytical expressions of permeability K, as a function of porosity $K=f(\varepsilon)$ can be done, if available, with caution. The values strongly depend on the way the wick structure is made, fabricated and eventually integrated and assembled with the heat pipe container. Moreover, cross-flow porosity and axial flow porosity, and hence respective permeability values, tend to be different for most porous structures used in heat pipes. Hence, a certain degree of heuristic knowledge of realistic in situ values of K and ε, is needed for successful modeling and comparison of the estimated values with the experimental data. For more information on analytical expressions of porosity and permeability of wick structures and their complex interdependency, readers are referred to Kaviany (1991) and Faghri (1995).

5. *Boiling limitation*: Liquid to vapor phase change, via the evaporation process occurring at the liquid–vapor interface, is the preferred mode of normal heat pipe operation. The wall superheat levels at the evaporator should be such that nucleation phenomena inside the wicks is avoided. If the superheat increases, it may lead to formation of bubbles inside the porous structure (Carey, 1992). This is detrimental to the heat pipe operation. Nucleation can be avoided if the wettability of the liquid is good, both with the container and the wick, the wick is properly in contact with the container, the container has high thermal conductivity and heat capacity and finally, the working fluid is clean, without any contamination. Boiling limitation usually occurs at higher end of the operating temperature spectrum and can be estimated by) (Chi, 1976):

$$\dot{Q}_{\max} = \frac{4\pi l_{eff} k_{eff} T_v \sigma_v}{h_{fg} \rho_l ln\left(\frac{r_i}{r_v}\right)}\left(\frac{1}{r_n} - \frac{1}{r_c}\right) \tag{4.17.9}$$

If these limitations are plotted, as a function of the operating temperature and for a given conventional heat pipe, then we will get a plot which is schematically shown in Figure 4.17.9. As can be seen, at operating temperatures T_1–T_5, the heat pipe operation gets limited by different limitations, as shown. Plotting such a graph is usually the first step while designing the heat pipe. If the heat handling capacity required from the heat pipe, comes within the design range, then it is acceptable, after incorporating a suitable factor of safety. However, if the requirement falls outside the design space, it means that the design will fail due to some limitation or the other, and hence suitable modifications must be made in the heat pipe geometry, wick structure, etc. to meet the requirements. As mentioned previously, for most conventional heat pipes operating at median values of their permissible range of temperature (between freezing point temperature and critical temperature), the capillary limit is the most dominant and important limit to estimate. This is also highlighted in Figure 4.17.9.

Choice of Working Fluids

The choice of a working fluid is primarily dictated by the desired operating temperature of the heat pipe operation. Several working fluids have been used in heat pipe operation, ranging from cryogenic fluids such as neon, low temperature fluids such as acetone and ammonia, to high-temperature fluids such as lithium or sodium. Needless to say, water is also an excellent fluid for many terrestrial application which are typically above its freezing point. What makes a working fluid appropriate for a given application? From a thermodynamic standpoint, we can argue that it should render a high capillary pumping power on one hand while requiring low dissipative power losses for its own circulation during operation. Therefore, high surface tension is desirable, as this will intrinsically enhance the available pumping power of the heat pipe. Low dynamic viscosity is another requirement as this will lower the frictional flow resistance of the fluid. As heat pipes primarily work on latent heat transfer mode, it is natural to choose a fluid with high latent heat of vaporization. In addition, the liquid density of working fluid must be high. This lowers the net inventory of the fluid per unit heat transferred. In other words, the flow velocity of the two phases per unit heat transferred by the heat pipe will reduce if the density of the fluid is high. This, in turn, lowers the frictional pressure drop. Combining all these attributes, one can get the desired "Figure of Merit," of a working fluid, given as

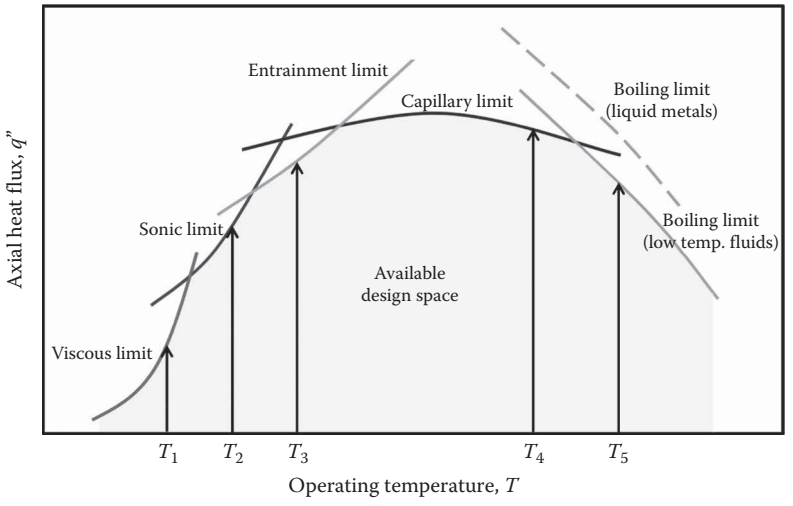


FIGURE 4.17.9 Different limitations of a conventional heat pipe and the available design space in which heat pipe will effectively operate.

$$M = \frac{\sigma \cdot h_{fg} \cdot \rho_l}{\mu_l} \tag{4.17.10}$$

M has the units of heat flux, that is, W/m^2. In addition, M is an implicit function of temperature. Hence, the Figure of merit M scales with the heat flux carrying capacity of a working fluid. Figure 4.17.10 shows the Figure of merit of a few working fluids, as a function of temperature. The typical "inverted U" shape of M is obvious, attributed to the dependence of M on the temperature. A high figure of merit means a corresponding high heat flux handling capacity, subject to heat pipe not getting limited by any thermal-hydraulic phenomena, as detailed earlier.

Genealogy of Heat Pipe Family

Having understood the basic operation, principle and types of heat pipes, we will quickly outline the subtle differences in the operation of these passive thermal systems so that nuances of their design methodology and eventual limitations, are well appreciated. Figure 4.17.11 shows the genealogy of closed passive two-phase heat transfer systems. While the "family tree" is certainly not exhaustive, most relevant systems for the present discussion are included. As will be evident shortly, although all systems shown here have "similar" working principles, there are inherent differences that significantly alter the fundamental heat transfer mechanism and thus the mathematical analyses (Khandekar, 2004).

The presence of a capillary wick for fluid transport is the fundamental characteristics of the first group (Group A) of the family, consisting of conventional and LHPs. While the former has countercurrent flow of the two phases, the latter design avoids it. For analysis of such systems, leading to the determination of maximum heat transport, based on capillary limit, as has been explained earlier (neglecting pressure drops associated with the phase-change process), we use (Faghri, 1995)

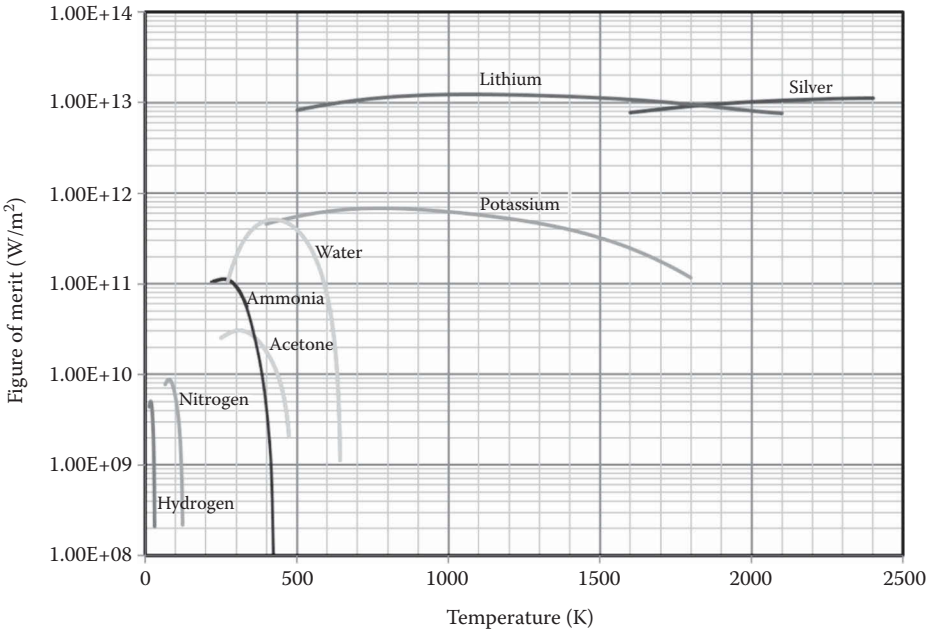


FIGURE 4.17.10 Figure of merit, M, plotted for different fluids, from cryogenic fluids, to medium temperature fluids and liquid metals.

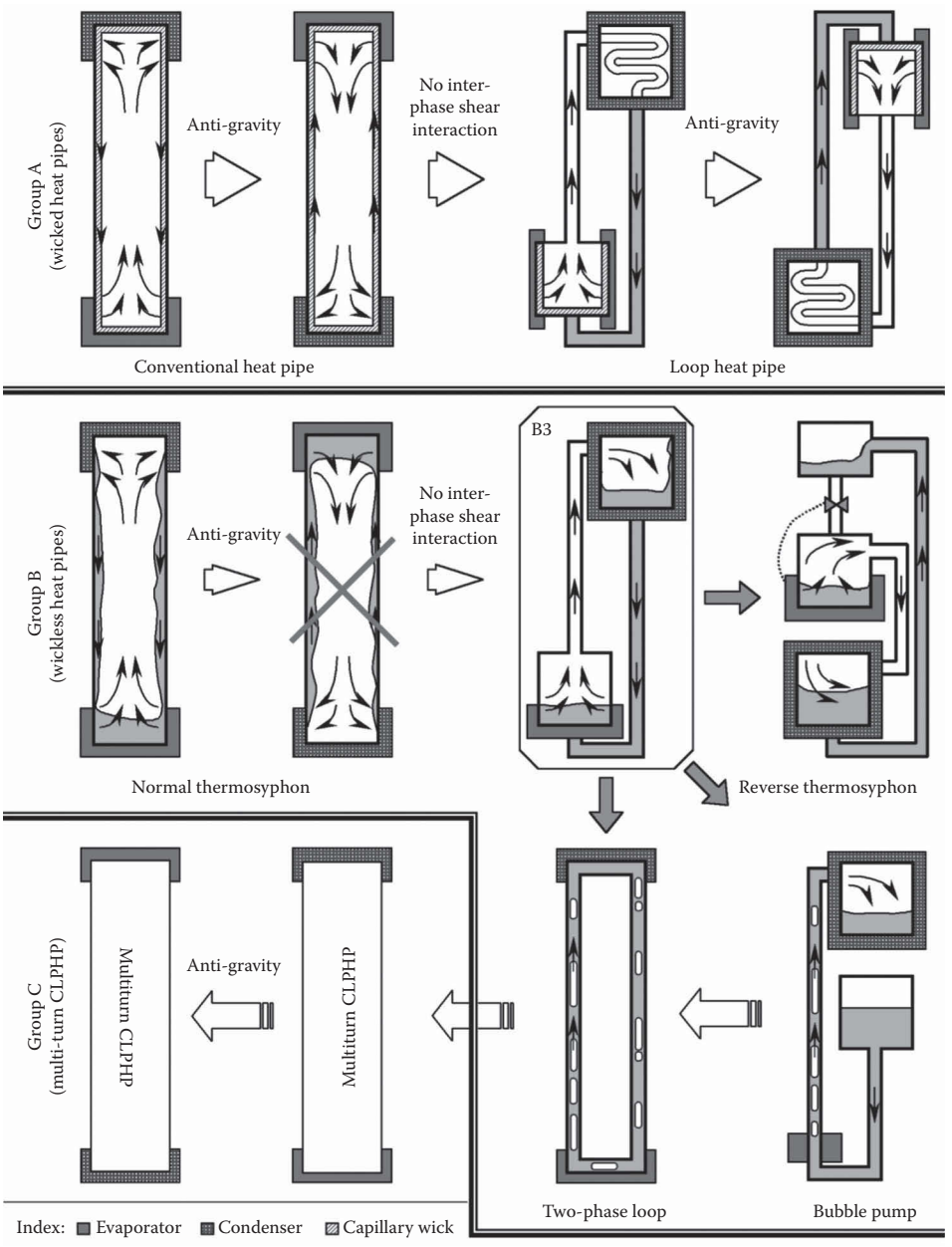


FIGURE 4.17.11 Genealogy of the heat pipe family showing three distinct classes of passive device. Although this classification is not exhaustive, it does include the major types of heat pipes in service.

$$(\Delta P)_{cap} \pm (\Delta P)_{gra} \geq (\Delta P)_{liq} + (\Delta P)_{vap} + (\Delta P)_{liq/vap} \tag{4.17.11}$$

$$(\Delta P)_{gra} = (L_{eff}^{liq} \cdot \rho_{liq} \cdot g \cdot \sin\beta) - (L_{eff}^{vap} \cdot \rho_{vap} \cdot g \cdot \sin\beta) \tag{4.17.12}$$

The R.H.S. of Equation 4.17.11 contains "dissipative" terms that are always positive while the L.H.S. contains the "driving" terms that must be at least equal to the R.H.S. to keep the system running. Thus, all devices in this family can function in antigravity mode if the capillary pumping action

is greater than the gravitational head. The interfacial shear interaction represented by $(\Delta P)_{\text{liq/vap}}$ is avoided in LHPs by the geometrical construction of the device rendering it to be a better alternative to conventional designs.

If the capillary pumping of the first group of devices vanishes, we obtain the second group in the family (Group B). This is represented by the gravity-assisted thermosyphons. Here, the design equation reduces to

$$(\Delta P)_{\text{gra}} \geq (\Delta P)_{\text{liq}} + (\Delta P)_{\text{vap}} + (\Delta P)_{\text{liq/vap}} \tag{4.17.13}$$

Thus, in this group, only the gravity head provides the driving potential. Naturally, since the L.H.S. of Equation 4.17.13 must always be positive for a viable system, antigravity (or zero gravity) operation is not feasible. Usually, $(\Delta P)_{\text{liq/vap}}$ limits the design, leading to counter-current flow limitation (Tien and Chung, 1978; Groll and Rösler, 1992). Such failures can be avoided in design-B3, where the two phases flow in separate channels, as was described in Section 3. This design represents a genealogical crossroad that will lead us to Group C devices including PHPs. All designs discussed so far are primarily based on latent heat transfer. Although sensible heating/cooling does take place, a near isothermal (or isobaric) assumption within the system is justified and hence routinely employed. This assumption is quite realistic for Group A and B devices.

One variation of design-B3 can be obtained by reducing the size of fluid transport channels and the evaporator, to capillary dimensions. One common example of such a case is the standard bubble pump in which, instead of single-phase vapor flow from the evaporator to the condenser, the flow of two-phase mixture results; a direct consequence of slug flow regime being established in the up-header (Benhmidene et al., 2010). In both systems, the construction of the device forces a combination of sensible as well as latent heat transfer. This marks the first significant shift from the previously discussed cases. The second critical aspect is the fact that the gravity head and the dissipative terms are no more based on single-phase calculations but are governed by two-phase pressure drops (effective mixture density and viscosity for given boundary conditions including consideration for the flow regime). This is because, there is no clear stratification of the phases. In addition, depending on the size of the capillary tubes, some additional capillary forces may have to be accounted for depending on their magnitude as compared to the other relevant heads. What complicates the situation further is the inherent presence of two-phase flow instabilities in such systems (Carey, 1992).

Another variation of design B3 comes in the form of an antigravity reverse thermosyphon by incorporating a suitable remote operated valve (e.g., see Sasin et al., 1995; Fantozzi et al., 2002; Filippeschi, 2006). Since there is neither a capillary pumping action nor positive gravity head, the driving potential must come from elsewhere. At the commencement of liquid heating in the evaporator chamber, the remote operated valve is kept closed. Boiling proceeds under nearly constant volume conditions (since there is always some liquid present in the condenser) thereby increasing the local vapor pressure and temperature. After a while, the evaporator pressure becomes high enough to push the liquid out of the condenser, against gravity, toward the accumulator. The evaporator gradually empties while the accumulator starts filling up. A stage is reached when all the evaporator liquid is boiled off. This stage is immediately followed by a rise in heater temperature and a drop in evaporator chamber pressure as the remaining vapor condenses. The remote valve can now be triggered to open by either of these signals creating pressure equalization in the accumulator and evaporator by gravity-assisted liquid return. The cycle then repeats itself. Thus, we observe that a slight design variation necessitates a change in the driving potential; the applicable design criterion now becomes

$$(\Delta P)_{\text{sat}}^{\text{e-acc}} - (\Delta P)_{\text{gra}} \geq (\Delta P)_{\text{liq}} + (\Delta P)_{\text{vap}} \quad \text{where} \tag{4.17.14}$$

$$(\Delta P)_{\text{sat}}^{\text{e-acc}} = (\Delta P)_{\text{sat}}\Big|_{\text{e}} - (\Delta P)_{\text{sat}}\Big|_{\text{acc}} \text{ and} \tag{4.17.15}$$

$$(\Delta P)_{\text{gra}} = (L_{\text{eff}}^{\text{liq}} \cdot \rho_{\text{liq}} \cdot g \cdot \sin\beta) - (L_{\text{eff}}^{\text{vap}} \cdot \rho_{\text{vap}} \cdot g \cdot \sin\beta) \tag{4.17.16}$$

With the preceding discussion, we conclude that the "driving" terms in this family of devices are $(\Delta P)_{gra}$, $(\Delta P)_{cap}$ or $(\Delta P)_{sat}$, acting individually or in a combination as the case may be. The most complicated situation will arise if all the "driving" terms are simultaneously present and two phases are not flowing separately but as a mixture; the design equation then takes the following form (again neglecting the pressure drops associated with phase change in the evaporator and the condenser):

$$\sum_{1}^{2}(\Delta P)_{cap} + \sum_{1}^{2}(\Delta P)_{sat} \pm \sum_{1}^{2}(\Delta P)_{gra}^{\text{two-phase}} \geq \sum_{1}^{2}(\Delta P)_{\text{two-phase}} \tag{4.17.17}$$

where 1 and 2 are specified locations in the system. The RHS includes two-phase friction as well as acceleration terms, if any. Working relationships for Equation 4.17.17 for devices like PHPs are still not available.

Literature on Heat Pipes

Literature on heat pipes is available from multiple sources. Several books and monographs have been written on the subject and some have also been updated in the recent past (Chi, 1976; Dunn and Reay, 1982; Ivanovskii et al., 1982; Peterson, 1994; Faghri, 1995; Reay et al., 2013; Vasiliev and Kakaç, 2013). The material is also available in the form of reports and data compilation. Several well-known journals attract contributions on heat pipes and related technology, both from the stand point of fundamental understanding as well as applications. In recent times, nearly 250 articles appear annually in Web-of-Science recognized journals. For example, over 800 articles have been published in more than 200 journals between 2012 and 2014 (source: Web of Science); major contributions are summarized in Figure 4.17.12 (Lips et al., 2016). In the recent past, research emphasis is on LHPs and PHPs. In parallel, the proceedings of the International Heat Pipe Conference series have been a comprehensive source of information, since the inception of this series of conferences in 1973 at Stuttgart, Germany; the conference has already been held for 17 times, the jubilee edition of which was hosted at the campus of Indian Institute of Technology Kanpur, India, in autumn of 2013. Comprehensive review papers are also available, e.g., see Ochterbeck (2003), Riffat and Ma (2007), Launay et al. (2007), Vasiliev (2008), Khandekar et al. (2010), Maydanik et al. (2014), Siedel et al. (2015), Shabgard et al. (2015), and Chan et al. (2015).

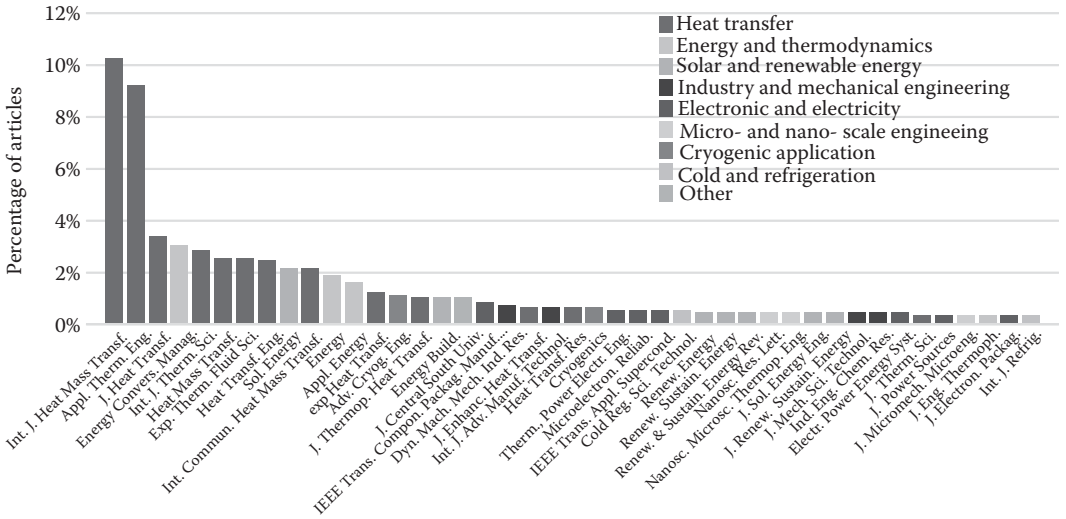


FIGURE 4.17.12 Statistics on journal publications in the field of heat pipe science and technology, taken from Web of Science. (From Lips S., Sartre V., Lefèvre, F., Khandekar S., and Bonjour J., *Interfac. Phenom Heat Transfer*, 4, 33–53, 2016. With permission.)

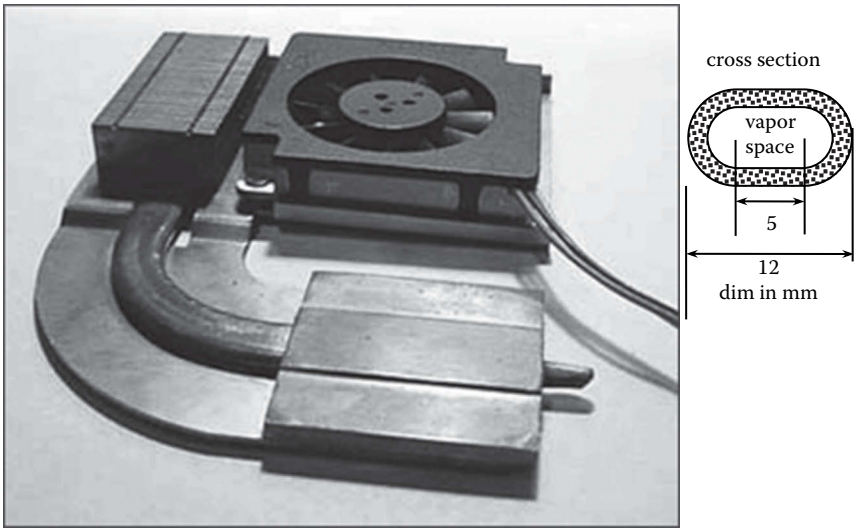


FIGURE 4.17.13 A typical heat pipe system of a laptop computer: example problem.

Example Problem

A copper water heat pipe is used to transfer heat from a microprocessor of a laptop to the ambient. A typical heat pipe-based module, with its cross-sectional shape, is shown in Figure 4.17.13. The evaporator section is 40 mm long while the condenser section is 60 mm long. The total length of the heat pipe is 160 mm. The heat pipe uses a porous capillary structure, 1.0 mm thick, having permeability 1×10^{-10} m^2, and average pore radius of 100 μm. The additional frictional resistance in the liquid due to the bend in the heat pipe is 30% of the total liquid pressure drop. The operating temperature is 70°C. Find the maximum heat transport capability under horizontal operation.

Solution:

Given:

Length of the evaporator section $= l_e = 40$ mm
Length of the condenser section $= l_c = 60$ mm
Length of the adiabatic section $= l_a = 60$ mm
Permeability of wick $= K = 1\times10^{-10}$ m^2
$r_e = r_c = r_{wick}$ = Pore radius = 100 μm
Effective contact angle in the evaporator = 0°
Effective contact angle in the condenser = 90°
The effective length of the heat pipe can be estimated as: $l_{eff} = l_a + 0.5\,(l_c + l_e) = 110$ mm

$$(\Delta P)_{\text{bending}} = 0.3(\Delta P)_{\text{liq}}$$

Thermophysical properties of water can be obtained from the property tables (refer Section 9). At 70°C, the relevant properties are estimated as

$$\rho_l = 977.74 \text{ kg/m}^3;$$

$$\rho_v = 0.197 \text{ kg/m}^3;$$

$$\mu_l = 3.9889\times10^{-4} \text{ Pa-s};$$

$$\mu_v = 1.0860\times10^{-5} \text{ Pa-s};$$

$$\sigma = 0.0644 \text{ N/m};$$

$$h_{fg} = 2.3277 \times 10^{6} \text{ J/kg}$$

Using the given geometric parameters

$$\text{Area of cross section for vapor flow} = A_{vap} = 44.63 \times 10^{-6} \text{ m}^2$$

$$\text{Area of cross section for liquid flow} = A_{liquid} = 28.85 \times 10^{-6} \text{ m}^2$$

$$D_{hyd,vap} = \frac{4A_{vap}}{P_{inner}} = 6.9 \times 10^{-3} \text{ m} = 6.9 \text{ mm}$$

Pressure due to capillary effect of wick (Driving Pressure)

$$(\Delta p)_{cap} = \frac{2 \cdot \sigma \cdot \left[\cos(\theta_{cap,c}) - \cos(\theta_{cap,e})\right]}{r_{wick}}$$

Pressure drop for the in liquid flow in the porous wick

$$(\Delta P)_l = \frac{\mu_l . l_{eff} u_l}{K}$$

$$\dot{Q} = \dot{m} \cdot h_{fg} = (\rho_l u_l A_l) h_{fg} = (\rho_v u_v A_v) h_{fg}$$

$$(\Delta P)_l = \frac{\mu_l l_{eff}}{K} \left(\frac{\dot{Q}}{\rho_l A_l h_{fg}} \right)$$

$(\Delta P)_{vap}$: Pressure drop in vapor which flows in the inner cross section

$$(\Delta P)_{vap} = \frac{f . l_{eff} u_v^2}{2 D_{hyd,vap}}$$

where

- *Case 1*: when the vapor flow is laminar, then

$$f = \frac{64}{\text{Re}_v} \quad \left(\text{Laminar flow fiction factor}\right)$$

- *Case 2*: when the vapor flow is turbulent

$$f = 0.079 \text{Re}^{-1/4} \quad \left(\text{Turbulent flow fiction factor}\right)$$

$$\text{Re}_v = \frac{\rho_v u_v D_{hyd,v}}{\mu_v} = \frac{D_{hyd,v}}{\mu_v} \left(\frac{\dot{Q}}{A_v h_{fg}} \right)$$

To proceed further, let us first assume that the vapor flow is laminar for 70°C operating temperature. We will solve the problem and then verify whether this assumption is true. We get

$$(\Delta P)_{vap} = \frac{64}{\left(\frac{\rho_v u_v D_{hyd,v}}{\mu_v} \right)} \left[\frac{\frac{1}{2} l_{eff} u_v^2}{D_{hyd,v}} \right] = \frac{32 \mu_v l_{eff} . \dot{Q}}{\rho_v^2 h_{fg} A_v D_{hyd,v}^2} = (0.196) \cdot \dot{Q} \text{ Pa}$$

or $(\Delta P)_{vap}$ (by the Hagen–Poiseuille equation)

$$(\Delta P)_l = \frac{\mu_l l_{eff}}{K}\left(\frac{\dot{Q}}{\rho_l A_l h_{fg}}\right) = (6.682)\dot{Q}\ \text{Pa}$$

$$(\Delta p)_{cap} = \frac{2\cdot\sigma\cdot\left[\cos(\theta_{cap,c}) - \cos(\theta_{cap,e})\right]}{r_{cap}} = 1.289\times10^3\ \text{Pa}$$

For finding the capillary limit, we equate the available pumping head, that is, the capillary pressure drop with the net dissipation in the system:

$$(\Delta P)_{cap} = (\Delta P)_l + (\Delta P)_v + (\Delta P)_{bending}$$

$$= 1.3(\Delta P)_l + (\Delta P)_v;\ \text{as}\ (\Delta P)_{bending} = 0.3(\Delta P)_l$$

$$(\Delta P)_{cap} = \left[\left[1.3\times\left\{\left(\frac{\mu_l l_{eff}}{KA_l\rho_l h_{fg}}\right)\right\}\right] + \frac{32\mu_v l_{eff}}{\rho_v^2 h_{fg} A_v D_{hyd,v}^2}\right]\dot{Q}$$

From this, $\dot{Q} = 145.1\,\text{W}$

For this $\dot{Q}$, calculating the vapor Reynolds number we get, $\text{Re}_v = 909.2$.

Hence, the laminar flow assumption for the vapor flow is correct, as $\text{Re}_v < 2300$. Also, for these conditions, the vapor pressure drop is negligible as compared to the liquid pressure drop. In general, using the above set of equations, the maximum heat capacity or capillary limit is calculated by varying wick thickness and results are plotted in Figure 4.17.14.

Thermophysical Property Data

Reliability of the thermophysical data used for the design and development of heat pipes is extremely important. In this context, the design data provided by Brennen and Kroliczek (1979) is very reliable and easy to implement in a computer program. We provide here analytical expressions for various properties in the form of either a power law or exponential fit (as indicated in the table below) for some working fluids. This can be directly used in computer subroutines for calling from the main program.

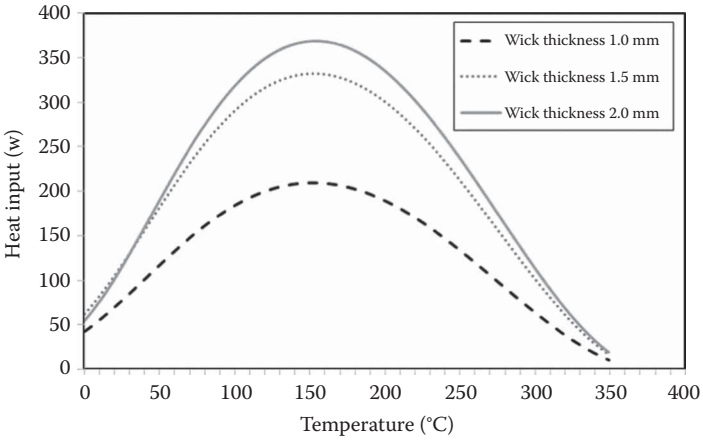


FIGURE 4.17.14 Capillary limit of the heat pipe discussed in the example problem, showing the parametric effect of wick thickness.

** Quantities follow the natural log – law fit, given by : ln(property value)

$$= a_0 + (a_1 * T) + (a_2 * T^2) + (a_3 * T^3) + (a_4 * T^4) + (a_5 * T^5)$$

All other properties follow the simple power-law fit, given by : property value

$$= a_0 + (a_1 * T) + (a_2 * T^2) + (a_3 * T^3) + (a_4 * T^4) + (a_5 * T^5)$$

Property variation based on Kelvin temperature scale

Ammonia (data validity range = 200–405 K)

Property	a_0	a_1	a_2	a_3	a_4	a_5
P_{sat} (Pa) **	–0.48050e2	0.68546	–0.33219e–2	0.87447e–5	–0.12019e–7	0.67905e–11
ρ_{liq} (kg/m^3)	0.27909e5	–0.49211e3	0.35315e1	–0.12540e–1	0.21993e–4	–0.15260e–7
μ_{liq} (Pa-s)	–0.10850e–1	0.20793e–3	–0.14481e–5	0.47962e–8	–0.76779e–11	0.47859e–14
ρ_{vap} (kg/m^3)**	–0.17401e3	0.27918e1	–0.18562e–1	0.63024e–4	–0.10743e–6	0.73141e–10
μ_{vap} (Pa-s)	–0.14764e–2	0.26963e–4	–0.19375e–6	0.68806e–09	–0.12066e–11	0.83651e–15
σ (N/m)	–0.12139	0.33113e–2	–0.23485e–4	0.77093e–7	–0.12620e–9	0.83185e–13
h_{fg} (J/kg)	0.48869e8	–0.86867e6	0.63176e4	–0.22752e2	0.40498e–1	–0.28572e–4
k_{liq} (W/mK)	0.19277e1	–0.15745e–1	0.93709e–4	–0.32309e–6	0.55219e–9	–0.37408e–12

Water (data validity range = 273–643 K)

Property	a_0	a_1	a_2	a_3	a_4	a_5
P_{sat} (Pa) **	–0.56346e2	0.50859	–0.16094e–2	0.27634e–5	–0.24744e–8	0.90907e–12
ρ_{liq} (kg/m^3)	0.60281e4	–0.64604e2	0.32699	–0.81081e–3	0.98069e–6	–0.46678e–9
μ_{liq} (Pa-s)	0.50668e2	–0.57958	0.23688e–2	–0.49807e–5	0.53004e–8	–0.22662e–11
ρ_{vap} (kg/m^3) **	–0.68620e2	0.52894	–0.17525e–2	0.31666e–5	–0.30147e–8	0.11981e–11
μ_{vap} (Pa-s)	–0.52621e–3	0.65779e–5	–0.32142e–7	0.77779e–10	–0.92704e–13	0.43559e–16
σ (N/m)	–0.14890e–1	0.13684e–2	–0.69541e–5	0.16145e–7	–0.19327e–10	0.92992e–14
h_{fg} (J/kg)	0.19664e8	–0.20880e6	0.10134e4	–0.24457e1	0.29101e–2	–0.13754e–5
k_{liq} (W/mK)	0.32183e1	–0.40434e–1	0.22075e–3	–0.55327e–6	0.65931e–9	–0.30706e–12

Methanol (data validity range = 273–503 K)

Property	a_0	a_1	a_2	a_3	a_4	a_5
P_{sat} (Pa) **	–0.94426e2	0.11137e1	–0.51015e–2	0.12471e–4	–0.15649e–7	0.79292e–11
ρ_{liq} (kg/m^3)	0.27598e5	–0.37094e3	0.20444e1	–0.55970e–2	0.76002e–5	–0.41010e–8
μ_{liq} (Pa-s)	0.59706e–1	–0.70041e–3	0.33431e–5	–0.80401e–8	0.96919e–11	–0.46707e–14
ρ_{vap} (kg/m^3) **	–0.19440e3	0.23581e1	–0.12009e–1	0.31490e–4	–0.41683e–7	0.22125e–10
μ_{vap} (Pa-s)	0.11817e–4	–0.14244e–6	0.10151e–8	–0.29020e–11	0.40919e–14	–0.22760e–17
σ (N/m)	0.28443	–0.30420e–2	0.14372e–4	–0.33905e–7	0.38809e–10	–0.17362e–13
h_{fg} (J/kg)	0.381476e8	–0.52614e6	0.29710e4	–0.82953e1	0.11440e–1	–0.62559e–5
k_{liq} (W/mK)	–0.14846e1	0.24079e–1	–0.13242e–3	0.35549e–6	–0.47110e–9	0.24639e–12

Benzene (data validity range = 280–553 K)						
Property	a_0	a_1	a_2	a_3	a_4	a_5
P_{sat} (Pa) **	–0.42991e2	0.42877	–0.14269e–2	0.25936e–5	–0.24811e–8	0.98520e–12
ρ_{liq} (kg/m^3)	0.15383e5	–0.18810e3	0.97455	–0.25091e–2	0.31951e–5	–0.16131e–8
μ_{liq} (Pa-s)	0.47188e–1	–0.51986e–3	0.23287e–5	–0.52489e–8	0.59235e–11	–0.26704e–14
ρ_{vap} (kg/m^3) **	–0.96643e2	0.10111e1	–0.44839e–2	0.10464e–4	–0.12482e–7	0.60224e–11
μ_{vap} (Pa-s)	–0.40125e–3	0.52693e–5	–0.27117e–7	0.69433e–10	–0.88142e–13	0.44497e–16
σ (N/m)	0.74267e–1	–0.74353e–4	–0.11568e–5	0.53611e–8	–0.99779e–11	0.67184e–14
h_{fg} (J/kg)	0.12604e8	–0.15941e6	0.82907e3	–0.21353e1	0.27172e–2	–0.13701e–5
k_{liq} (W/mK)	0.22155	–0.30098e–3	0.37104e–6	–0.12214e–8	0.18210e–11	–0.10066e–14

Summary and Outlook

Heat pipes deserve a special mention in the broad field of thermal management. These simple yet highly efficient class of passive devices have served the industry very well. Their attractiveness lies in the simplicity, excellent performance, possibilities to operate in a wide temperature range, long-term durability, nearly no maintenance and flexibility to cater to a range of applications ranging from space thermal management to terrestrial systems. Heat pipes are also available on several length scales, with submillimeter micro heat pipes to large-scale deployment in, power plants, permafrost stabilization and geothermal energy recovery units. The wave of miniaturization in the commercial electronics industry in the last three decades has provided new impetus and possibilities for the development of frugal heat pipe technology solutions. In this context, copper-water heat pipes have become ubiquitous in PC and laptop industry. Similarly, contemporary satellite operations without heat pipes are unthinkable. In the last two decades, several niche application areas have also emerged, from biomedical to nuclear industry. Starting their journey from simple gravity-assisted thermosyphons applied in bakeries, heat pipes have indeed come a long way, with state-of-the-art high-performance LHPs/CPLs, to deceptively simple, yet one of the most theoretically challenging systems manifested as PHPs.

With a robust history of applications behind them, theoretical understanding of various designs of heat pipes has also progressed. However, several aspects of complex and interlinked thermo-physics of heat pipes still remain to be fully understood. Two glaring examples worth mentioning here are (1) the fundamental understanding of the boiling/evaporation phenomena in LHP evaporator wicks and (2) mathematical description of intricate and coupled thermal-hydraulic mechanisms of heat transfer in a PHP. In addition, even for conventional heat pipe designs, the transient models continue to perform unsatisfactorily as compared to steady-state models. In the last decade, significant developments have been witnessed in material science, composites, polymers, nanotechnology, and indeed on the broader science of interfacial phenomena, wetting, and hierarchical modeling of complex transport physics. The synergistic structure which the heat pipe is, its development cannot remain isolated from these broader developments. There is a need to comprehensively integrate and infuse these developments and progress of several stake-holding topics into the realm of heat pipe technology.

ACKNOWLEDGMENTS

I was introduced to the art and science of heat pipes by Prof. Manfred Groll, my doctoral advisor, in the year 2000. Over the last 40 years, he has trained and inspired a series of students and researchers throughout the world, providing unmatched contribution to this field of engineering science. I dedicate this work to him as he continues his journey of lifelong achievements. I must also mention the special place which Prof. Yuri Maydanik holds in the heat pipe community, not only as an inventor of LHPs, but also as an inspiring entrepreneur. Special thanks to Prof. Jocelyn Bonjour and Prof. Frederic Lefèvre, my collaborators from INSA-Lyon, France. I would like to express my gratitude

to all funding agencies who have supported my research ideas. The contribution of my institute, which has provided excellent atmosphere for professional development, cannot be understated. Last, but not the least, I wish to thank all my graduate students and research associates who have indeed taught me many things!

NOMENCLATURE

Roman symbols

A	cross section (m^2)
D, d	diameter (m)
g	gravitational constant (m/s^2)
h_{fg}	latent heat of vaporization (J/kg)
I	electric current (A)
K	permeability (m^2)
k	thermal conductivity (W/(mK))
L, l	length (m)
M	figure of merit (W/m^2)
P	pressure (Pa)
$\dot{Q}$	heat flow, thermal power (W)
q	heat flux (W/m^2)
R	resistance
r	radius (m)
T	temperature (°C)
t	time (s)
V	electric potential, emf (V)
V, u	velocity (m/s)

Greek symbols

β	inclination with horizontal (°)
ϵ	porosity (–)
μ	dynamic viscosity (Pa s)
ρ	density (kg/m^3)
σ	surface tension (N/m)
θ	wetting angle (°)

Subscripts

acc	acceleration
a, ad	adiabatic
ave	average
ax	axial
c	condenser
cha	characteristic
crit	critical
cap	capillary
e, **E**	evaporator
eff	effective

elec	electrical
gra	gravitational
h, **hyd**	hydraulic
i	inner
l, **liq**	liquid
max	maximum
n	nucleation
rad	radial
s, **sat**	saturation
t	total
th	thermal
v, **vap**	vapor

Nondimensional groups

Bo	Bond number, $\mathrm{Bo}=D\,(g(\rho_l-\rho_v)/\sigma)^{0.5}$
Re	Reynolds number, $\mathrm{Re}=(\rho\cdot v\cdot d_h)/\mu$
We	Weber number, $\mathrm{We}=(\rho_v\cdot v^2\cdot l_{cha})/\sigma$

REFERENCES

Adoni A.A., Vaidya J.S., Ambirajan A., Kumar D., Krishnaswamy B., and Dutta, P., Evaporation heat transfer coefficient in a capillary pumped loop and loop heat pipe for different working fluids, *Heat Transfer Eng.*, 33 (9), pp. 765–774, 2012.

Akachi H., A Structure of a Heat Pipe, US Patent No. 4921041, 1990.

Akachi H., L-type heat pipe, US Patent No. 5490558, 1996.

Ambirajan A., Adoni A.A., Vaidya J.S., Rajendran A.A., Kumar D., and Dutta, P., Loop heat pipes: A review of fundamentals, operation and design, *Heat Transfer Eng.*, 33 (4–5), pp. 387–405, 2012.

Ayel V., Araneo L., Scalambra A., Mameli M., Romestant C., Piteau A., Marengo M., Filippeschi S., and Bertin, Y., Experimental study of a closed loop flat plate pulsating heat pipe under a varying gravity force, *Int. J. Therm. Sci.*, 96, art. no. 4429, pp. 23–34, 2015.

Babin B.R., Peterson G.P., and Wu D., Steady-state modeling and testing of a micro heat pipe, *ASME J. Heat Transfer*, 112, pp. 595–601, 1990.

Bajpai A.K., and Khandekar S., Thermal transport behavior of a liquid plug moving inside a dry capillary tube, *Heat Pipe Sci. Technol.*, 3 (2–4), pp. 97–124, 2012.

Bazzo E., and Riehl R.R., Operation characteristics of a small-scale capillary pumped loop, *Appl. Therm. Eng.*, 23 (6), pp. 687–705, 2003.

Benhmidene A., Chaouachi B., and Gabsi S., A review of bubble pump technologies, *J. Appl. Sci.*, 10 (16), pp. 1806–1813. 2010.

Bertossi R., Guilhem N., Ayel V., Romestant C., and Bertin Y., Modeling of heat and mass transfer in the liquid film of rotating heat pipes, *Int. J. Therm. Sci.*, 52 (1), pp. 40–49, 2012.

Bienert W., and Brennan P.J., Transient performance of electrical feedback-controlled variable conductance heat pipes, ASME Paper 71-Av-27, 1971.

Bonnenfant J.-F., Benselama A.M., Ayel V., and Bertin, Y., Evaporating bubble in a heated capillary: Effect of passage on the temperature field of the external wall, *J. Phys.: Conf. Ser.*, 395 (1), art. no. 012172, 2012.

Boubaker R., Platel V., Berges A., Bancelin M., and Hannezo E., Dynamic model of heat and mass transfer in an unsaturated porous wick of capillary pumped loop, *Appl. Therm. Eng.*, 76, pp. 1–8, 2015.

Brennen P.J., and Kroliczek E.J., *Heat Pipe Design Handbook*, NASA Contract Number NAS5-23406, 1979.

Bugby D., Nguyen T., Kroliczek E.J., Ku J., Swanson T., Cullimore B., and Baumann, J., Development and testing of a cryogenic capillary pumped loop flight experiment, *33rd Intersociety Engineering Conference Energy Conversion*, Colorado Springs, CO, August 2–6, 1998.

Buschmann M., Nanofluids in thermosyphons and heat pipes: Overview of recent experiments and modelling approaches, *Int. J. Therm. Sci.*, 72, pp. 1–17, 2013.

Cao Y., Faghri A., and Mahefkey E., Micro/miniature heat pipes and operating limitations, *Proceedings of 29th ASME National Heat Transfer Conference, HTD*, Atlanta, GA, Vol. 236, pp. 55–62, August 8–11, 1993.

Carey V. P., *Liquid-Vapor Phase-Change Phenomena: Introduction to the Thermophysics of Vaporization and Condensation in Heat Transfer Equipment*, Series in Chemical and Mechanical Engineering, Taylor and Francis, New York, 1992.

Chan C.W., Siqueiros E., Ling-Chin J., Royapoor M., and Roskilly A.P., Heat utilization technologies: A critical review of heat pipes, *Renewable Sustainable Energy Rev.*, 50, pp. 615–627, 2015.

Charoensawan P., Khandekar S., Groll M., and Terdtoon P., Closed loop pulsating heat pipes—Part A: Parametric experimental investigations, *Appl. Therm. Eng.*, 23 (16), pp. 2009–2020, 2003.

Chatterjee A., Plawsky J.L., and Wayner Jr., P.C., Disjoining pressure and capillarity in the constrained vapor bubble heat transfer system, *Adv. Coll. Interface Sci.*, 168 (1–2), pp. 40–49, 2011.

Chauris N., Ayel V., Bertin Y., and Romestant, C., Evaporation of a liquid film deposited on a capillary heated tube: Experimental analysis by infrared thermography of its thermal footprint, *Int. J. Heat Mass Transfer*, 86, art. no. 11816, pp. 492–507, 2015.

Chen H., Groll M., and Roesler S., Micro heat pipes: Experimental investigation and theoretical modelling, *Proceedings of 8th International Heat Pipe Conference*, Beijing, pp. 396–400, September 14–18, 1992.

Chi S.W., *Heat Pipe Theory and Practice*, Hemisphere Publishing Corporation, Washington, DC, 1976.

Cotter T.P., Principles and prospects for micro heat pipes. *Proceedings of 5th International Heat Pipe Conference*, Tsukuba, Vol. 1, pp. 328–335, May 14–18, 1984.

Das Gupta S., Plawsky J.L., and Wayner Jr., P.C., Interfacial force field characterization in a constrained vapor bubble thermosyphon, *AIChE*, 41 (9), pp. 2140–2149, 1995.

D'Entremont B.P., and Ochterbeck J.M., Performance of low-conductivity ceramic wicks in capillary evaporators, *Proceedings of 40th AIAA Thermophysics Conference*, Seattle, WA, art. no. 2008–3925, June 23–26, 2008.

Dunn P., and Reay D.A., *Heat Pipes* (3rd ed.). Pergamon Press, Oxford, 1982.

Faghri A., *Heat Pipe Science and Technology*. Taylor & Francis, Washington, DC, 1995.

Faghri A., Review and advances in heat pipe science and technology, *ASME J. Heat Transfer*, 134 (12), art. no. 123001, 2012.

Faghri A., Heat pipes: Review, opportunities and challenges, *Front. Heat Pipes*, 5, 013001, 2014.

Fantozzi F., Filippeschi S., Sasin V.J., and Savchenkova N.M., Heat transport device based on pulsating thermosyphons with forced fluctuation of pressure, *Proceedings of 12th International Heat Pipe Conference*, pp. 469–474, Moscow, Russia, May 19–24, 2002.

Filippeschi S., On periodic two-phase thermosyphons operating against gravity, *Int. J. Thermal Sci.*, 45 (2), pp. 124–137, 2006.

Gaugler R.S., Heat transfer device, US Patent 2350348, 1944.

Gerasimov Y.F., Maidanik Y.F., Shchegolev G.T., Filippov, G.A., Starikov L.G., Kiseev V.M., and Dolgirev Y.E., Heat pipe, USSR Inventors Certificate 449 213, 1974.

Gerasimov Y.F., Maidanik Y.F., Shchegolev G.T., Filippov, G.A., Starikov L.G., Kiseev V.M., and Dolgirev Y.E., Low-temperature heat pipes with separate channels for vapor and liquid, *J. Eng. Phys.*, 28 (60), pp. 683–685, 1975.

Gerner F.M., Longtin J.P., Henderson H.T., Hsieh W.M., Ramdas P., and Chang W.S., Flow limitations in micro heat pipes, *Proceedings of 28th ASME National Heat Transfer Conference*, San Diego, USA, August 9–12, 1992.

Gilmore D.G. (Editor), *Spacecraft Thermal Control Handbook* (2nd ed.). The Aerospace Press, AIAA, Inc., El Segundo, 2002.

Groll M., Heat pipe science and technology: A historical review. *Proceedings of 17th International Heat Pipe Conference*, Kanpur, India, October 13–18, 2013.

Groll M., and Khandekar S., Heat transfer and fluid flow in microchannels: Micro heat pipes, *Heat Exchanger Design Handbook (Update)*, Chapter on Microscale Boiling and Condensation, Vol. 9, No. 1/2, Begell House, Danbury, CT, 2002a.

Groll M., and Khandekar S., Pulsating heat pipes: A challenge and still unsolved problem in heat pipe science, *Arch. Thermodyn.*, 23 (4), pp. 17–28, 2002b.

Groll M., and Kirkpatrick J.P., Heat pipes for spacecraft temperature control—An assessment of the state-of-the-art. *Proceedings of 2nd IHPC*, Bologna (ESA Report SP112), Vol. 1, 1976.

Groll M., and Rösler, S., Operation principles of heat pipes and closed thermosyphons, *J. Non-Equilib. Thermodyn.*, 17, pp. 91–151, 1992.

Groll M., Schneider M., Sartre V., Zaghdoudi M.C., and Lallemand M., Thermal control of electronic equipment by heat pipes, *Rev. Gén. Therm.*, 37, pp. 323–352, 1998.

Grover G.M., Evaporation-condensation heat transfer device, US Patent 3229759, 1966.
Grover G.M., Cotter T.P., and Erikson G.F., Structures of very high thermal conductivity, *J. Appl. Phys.*, 35, pp. 1990–1991, 1964.
Ha J.M., and Peterson G., The maximum heat transport capacity of micro heat pipes, *ASME J. Heat Transfer*, 120, pp. 1064–1071, 1998.
Han, X., Wang, X., Zheng, H., Xu, X., and Chen, G., Review of the development of pulsating heat pipe for heat dissipation, *Renewable Sustainable Energy Rev.*, 59, pp. 692–709, 2016.
Hemadri V. A., Gupta A., and Khandekar S., Thermal radiators with embedded pulsating heat pipes: Infra-red thermography and simulations, *Appl. Therm. Eng.*, 31, pp. 1332–1346, 2011.
Hoang T.T., and Ku, J., Advanced loop heat pipes for spacecraft thermal control, *8th AIAA/ASME Joint Thermophysics and Heat Transfer Conference*, St. Louis, MO, June 24–26, 2002.
Itoh A., and Polasek F., Development and application of micro heat pipes, *Proceedings of 7th International Heat Pipe Conference*, Minsk, pp. 295–310, May 21–25, 1990.
Ivanovskii M.N., Sorokin V.P., and Yagadkin I.V., *The Physical Principles of Heat Pipes*. Clarendon Press, Oxford, 1982.
Janeček V., and Nikolayev V.S., Apparent-contact-angle model at partial wetting and evaporation: Impact of surface forces, *Phys. Rev. E*, 87 (1), pp. 012404, 2013.
Joung W., Hwang H., and Lee J., Experimental study on the operating characteristics of a capillary pumped loop with a flat evaporator, *Int. J. Heat Mass Transfer*, 53 (1–3), pp. 268–275, 2010.
Karam R.D., *Satellite Thermal Control for Systems Engineers, Progress in Astronautics and Aeronautics*, Vol. 181. AIAA, Inc., Reston, VA, 1998.
Karthikeyan M., Huang J., Plawsky J., and Wayner Jr., P.C., Experimental study and modeling of the intermediate section of the non-isothermal constrained vapor bubble, *J. Heat Transfer*, 120 (1), pp. 166–173, 1998.
Karthikeyan V.K., Khandekar S., Pillai B.C., and Sharma P., Infrared thermography of pulsating heat pipe: Flow regimes and multiple steady states, *Appl. Therm. Eng.*, 62 (2), pp. 470–480, 2014.
Kaviany M., *Principles of Heat Transfer in Porous Media* (1st ed.), Series in Mechanical Engineering. Springer, New York. Reprint, 1991.
Kaya T., and Ku J., Thermal operational characteristics of a small-loop heat pipe, *J. Thermophys. and Heat Transfer*, 17 (4), pp. 464–470, 2003.
Khandekar S., Thermo-hydrodynamics of closed loop pulsating heat pipes, PhD thesis, University of Stuttgart, 2004. Available at: http://elib.uni-stuttgart.de/opus/volltexte/2004/1939/.
Khandekar S., and Groll M., An insight into thermo-hydraulic coupling in pulsating heat pipes, *Int. J. Therm. Sci.*, 43 (1), pp. 13–20, 2004a.
Khandekar S., and Groll M., Pulsating heat pipes: Attractive entrants in the family of closed passive two-phase systems, *J. Energy Heat Mass Transfer*, 26, pp. 99–115, 2004b.
Khandekar S., Dollinger N., and Groll M., Understanding operational regimes of pulsating heat pipes: An experimental study, *Appl. Therm. Eng.*, 23 (16), 23 (6), pp. 707–719, 2003.
Khandekar S., Gautam A.P., and Sharma P., Multiple quasi-steady states in a closed loop pulsating heat pipe, *Int. J. Therm. Sci.*, 48 (3), 535–546, 2009.
Khandekar S., Joshi Y., and Mehta B., Thermal performance of closed two-phase thermosyphon using nanofluids, *Int. J. Therm. Sci.*, 47 (6), pp. 659–667, 2008.
Khandekar S., Panigrahi P.K., Lefèvre F., and Bonjour J., Local hydrodynamics of flow in a pulsating heat pipe: A review, *Front. Heat Pipes*, 1, pp. 1–20, 2010.
Khandekar S., Schneider M., Schaefer P., Kulenovic R., and Groll M., Thermofluiddynamic study of flat plate closed loop pulsating heat pipes, *Microscale Thermophys. Eng.*, 6 (4), pp. 303–318, 2002.
Khrustalev D., and Faghri A., Thermal analysis of a micro heat pipe, *ASME J. Heat Transfer*, 116, pp. 189–198, 1994.
Kirkpatrick J.P., Variable conductance heat pipes—From the laboratory to space. *Proceedings of 1st IHPC*, Stuttgart, October 15–17, 1973.
Kirkpatrick J.P. and Marcus, B.D., A variable conductance heat pipe experiment. AIAA Paper 71–411, 1971.
Ku J., Ottenstein L., Douglas D., and Hoang T., Technology overview of a multi-evaporator miniature loop heat pipe for spacecraft applications, *J. Spacecraft and Rockets*, 49 (6), pp. 999–1007, 2012.
Launay S., Sartre V., and Bonjour J., Parametric analysis of loop heat pipe operation: A literature review, *Int. J. Therm. Sci.*, 46 (7), pp. 621–636, 2007.
Lips S., Sartre V., Lefèvre, F., Khandekar S., and Bonjour J., Overview of heat pipe studies during the period 2010-2015, *Interfac. Phenom Heat Transfer*, 4, pp. 33–53, 2016.

Longtin J.P., Badran B.F., and Gerner M.A., One-dimensional model of a micro heat pipe during steady-state operation, *ASME J. Heat Transfer*, 116, pp. 709–715, 1994.

Majumder A., Mehta B., and Khandekar S., Local Nusselt number enhancement during gas-liquid Taylor bubble flow in a square mini-channel: An experimental study, *Int. J. Therm. Sci.*, 66, pp. 8–18, 2013.

Mameli M., Pulsating heat pipes: Numerical modeling and experimental assessment, PhD thesis, University Bergamo, Italy, 2012.

Mameli M., Marengo M., and Khandekar S., Local heat transfer measurement and thermo-fluid characterization of a pulsating heat pipe, *Int. J. Therm. Sci.*, 75, pp. 140–152, 2014.

Maydanik Y.F., Loop heat pipes, *App. Therm. Eng.*, 25 (5–6), pp. 635–657, 2005.

Maydanik Y.F., Chernysheva M.A., and Pastukhov V.G., Review: Loop heat pipes with flat evaporators, *Appl. Therm. Eng.*, 67 (1–2), pp. 294–307, 2014.

Mehta B., and Khandekar S., Taylor bubble-train flow and heat transfer in the context of pulsating heat pipes, *Int. J. Heat Mass Transfer*, 79, pp. 279–290, 2014a.

Mehta B., and Khandekar S., Measurement of local heat transfer coefficient during gas-liquid Taylor bubble train flow by infrared thermography, *Int. J. Heat and Fluid Flow*, 45, pp. 41–52, 2014b.

Nikolayev V.S., Oscillatory instability of the gas-liquid meniscus in a capillary under the imposed temperature difference, *Int. J. Heat Mass Transfer*, 64, pp. 313–321, 2013.

Nikolayev V.S., Effect of tube heat conduction on the single branch pulsating heat pipe start-up, *Int. J. Heat Mass Transfer*, 95, pp. 477–487, 2016.

Ochterbeck J.M., Chapter 16—Heat transfer handbook, *Heat Pipes*. Wiley, Hoboken, NJ, 2003.

Ochterbeck J.M., and Peterson G.P., Chapter 7—Modeling of heat transfer in heat pipes. In: *Modeling of Engineering Heat Transfer Phenomena*, B. Sunden and M. Faghri (eds.), WIT Press, UK, pp. 175–212, 1999.

Panchamgam S.S., Chatterjee A., Plawsky J.L., and Wayner Jr., P.C., Comprehensive experimental and theoretical study of fluid flow and heat transfer in a microscopic evaporating meniscus in a miniature heat exchanger, *Int. J. Heat Mass Transfer*, 51 (21–22), pp. 5368–5379, 2008.

Pastukhov V.G., and Maydanik, Y.F., Combined LHP and PHP based heat-transfer system, *Int. J. Therm. Sci.*, 74, pp. 81–85, 2013.

Peterson G.P., *An Introduction to Heat Pipes: Modeling, Testing and Applications*. Wiley, New York, 1994.

Peterson G.P., Modeling, fabrication and testing of heat pipes: An update, *Appl. Mech. Rev.*, 49 (10), pp. 175–183, 1996.

Peterson G.P., Chapter 8. *Microscale Energy Transport*, Tien C-L, Majumdar A., and Gerner F. (Eds.), Series in Chemical and Mechanical Engineering, Taylor and Francis, CRC Press, Washington, DC, 1997.

Pouzet E., Joly J.-L., Platel V., Grandpeix J.-Y., and Butto C., Dynamic response of a capillary pumped loop subjected to various heat load transients, *Int. J. Heat Mass Transfer*, 47 (10–11), pp. 2293–2316, 2004.

Rana G.R., Sikarwar B.S., Khandekar S., and Panigrahi P. K., Hydrodynamics of a confined meniscus in a square capillary tube at low capillary numbers, *Front. Heat Pipes*, 5 (1), 013007(1–12), 2014.

Rao M., Lefèvre F., Khandekar S., and Bonjour J., Understanding transport mechanism of a self-sustained thermally driven oscillating two-phase system in a capillary tube, *Int. J. Heat Mass Transfer*, 65, pp. 451–459, 2013.

Rao M., Lefèvre F., Khandekar S., and Bonjour J., Heat and mass transfer mechanisms of a self-sustained thermally driven oscillating liquid-vapor meniscus, *Int. J. Heat Mass Transfer*, 86, pp. 519–530, 2015.

Reay D., McGlen, R., and Kew, P., *Heat Pipes: Theory, Design and Applications* (6th ed.). Butterworth-Heinemann, Amsterdam, 2013.

Riffat S., and Ma X., Recent developments in heat pipe technology and applications: A review, *Int. J. Low-Carbon Technol.*, 2 (2), pp. 162–177, 2007.

Samanta S.K., Sharma B.B., Das P., and Lohar A.K., Development of tubular Ni wick used in LHP for space applications, *Front. Heat Pipes*, 2, 043004 (1–5), 2011.

Sarno C., Tantolin C., Hodot R., Maydanik Y., and Vershinin S., Loop thermosyphon thermal management of the avionics of an in-flight entertainment system, *Appl. Therm. Eng.*, 51 (1–2), pp. 764–769, 2013.

Sasin V.J., Borodkin A.A., and Feodorov V. N., Experimental investigation and analytical modeling of auto-oscillation two-phase loop, *Proceedings of 9th International Heat Pipe Conference*, Los-Alamos, USA, May 1–5, 1995.

Shabgard H., Allen M.J., Sharifi N., Benn S.P., Faghri A., and Bergman T.L., Heat pipe heat exchangers and heat sinks: Opportunities, challenges, applications, analysis, and state of the art, *Int. J. Heat Mass Transfer*, 89 pp. 138–158, 2015.

Shukla K.N., Heat transfer limitation of a micro heat pipe, *ASME J. Electron. Packag.*, 131 (1–3), pp. 0245032, 2009.

Siedel B., Sartre V., and Lefèvre, F., Literature review: Steady-state modelling of loop heat pipes, *Appl. Therm. Eng.*, 75, pp. 709–723, 2015.

Srinivasan V., Khandekar S., Bouamrane N., Lefèvre F., and Bonjour J., Motion of an isolated liquid plug inside a capillary tube: Effect of contact angle hysteresis, *Exp. Fluids*, 56, 14 (6 pages), 2015a.

Srinivasan V., Marty-Jourjon V., Khandekar S., Lefèvre F., and Bonjour J., Evaporation of an isolated liquid plug moving inside a capillary tube, *Int. J. Heat Mass Transfer*, 89, 176–185, 2015b.

Sureshkumar R., Mohideen S.T., and Nethaji N., Heat transfer characteristics of nanofluids in heat pipes: A review, *Renewable Sustainale Energy Rev.*, 20, 397–410, 2013.

Swanson L.W., and Peterson G.P., The interfacial thermodynamics of micro heat pipes, *ASME J. Heat Transfer*, 115, pp. 195–201, 1995.

Tang X., Sha L., Zhang H., and Ju Y., A review of recent experimental investigations and theoretical analyses for pulsating heat pipes, *Front. Energy*, 7 (2), pp. 161–173, 2013.

Tien C.L., and Chung K.S., Entrainment limits in heat pipes, *Proceedings 3rd International Heat Pipe Conference*, Palo Alto, AIAA Paper 78–382, May 22–24, 1978.

Tio K., Liu C., and Toh K., Thermal analysis of micro heat pipes using a porous-medium model, *Heat Mass Transfer*, 36, pp. 21–28, 2000.

Vasiliev L.L., Micro and miniature heat pipes—Electronic component coolers, *Appl. Therm. Eng.*, 28 (4), pp. 266–273, 2008.

Vasiliev L.L., and Kakaç S. (Editor), *Heat Pipes and Solid Sorption Transformations: Fundamentals and Practical Applications*. CRC Press, Hoboken, NJ, 2013.

Vasiliev L., and Vasiliev Jr. L., Sorption heat pipe-a new thermal control device for space and ground application, *Int. J. Heat Mass Transfer*, 48 (12), pp. 2464–2472, 2005.

Vasiliev L., Lossouarn D., Romestant C., Alexandre A., Bertin Y., Piatsiushyk Y., and Romanenkov V., Loop heat pipe for cooling of high-power electronic components, *Int. J. Heat Mass Transfer*, 52 (1–2), pp. 301–308, 2009.

Wang C., Groll M., Roesler S., and Tu C.J., Porous medium model for two-phase flow in mini channels with applications to micro heat pipes, *Heat Recovery Syst. & CHP*, 14 (4), pp. 377–389, 1994.

Wang G., Mishkinis D., and Nikanpour D., Capillary heat loop technology: Space applications and recent Canadian activities, *Appl. Therm. Eng.*, 28 (4), pp. 284–303, 2008.

Wayner Jr., P.C., The effect of interfacial mass transport on flow in thin liquid films, *Colloids Surf.*, 52, pp. 71–84, 1991.

Yang H., Khandekar S., and Groll M., Operational limit of closed loop pulsating heat pipes, *Appl. Therm. Eng.*, 28 (1), pp. 49–59, 2008.

Yang H., Khandekar S., and Groll M., Performance characteristics of pulsating heat pipes as integral thermal spreaders, *Int. J. Therm. Sci.*, 48 (4), pp. 815–824, 2009.

Zhang, Y. and Faghri, A., Advances and unsolved issues in pulsating heat pipes, *Heat Transfer Eng.*, 29 (1), pp. 20–44, 2008.

Zheng L., Study of microscale transport processes and the stability of the thin film in a loop constrained vapor bubble, PhD thesis, Rensselaer Polytechnic Institute, New York, 2002.

4.18 LIQUID ATOMIZATION AND SPRAYING

Mario F. Trujillo and Rolf D. Reitz

Sprays are involved in many practical applications, including in the process industries (e.g., spray drying, spray cooling, powdered metals); in treatment applications (e.g., humidification, gas scrubbing); in coating applications (e.g., surface treatment, spray painting, and crop spraying); in spray combustion (e.g., burners, furnaces, rockets, gas turbines, diesel and port fuel injected engines); and in medicinal and printing applications. To be able to describe sprays, it is necessary to obtain a detailed understanding of spray processes.

In the simplest case, the liquid to be sprayed is injected at a high velocity through a small orifice. Atomization is the process whereby the injected liquid is broken up into droplets. Atomization has a strong influence on spray vaporization rates because it increases the total surface area of the injected liquid greatly. Fast vaporization may be desirable in certain applications, but undesirable in others, where the liquid is required to impinge on a target. The trajectories of the spray drops are governed by the injected momentum of the drop, drag forces, and interactions between the drops and the surrounding gas. Control of these and other spray processes can lead to significant improvements in performance and in quality of product, and to reduction of emission of pollutants.

Spray Characterization

Practical atomizers generate sprays with a distribution of drop sizes, with average sizes in the diameter range from a few microns ($1\,\mu m = 10^{-6}\,m$) to as large as 0.5 mm. It is important to quantify the details of the distribution depending on the application. For example, the smaller drops in a spray vaporize fast, and this is helpful to control ignition processes in some combustion systems. On the other hand, the large drops carry most of the mass and momentum of the injected liquid and these drops are able to penetrate into the high-pressure gases in engine combustion chambers. Typical average drop sizes for broad classes of sprays are shown schematically in Figure 4.18.1. It should be noted that the terminology used to describe sprays in Figure 4.18.1 is qualitative and is not universally agreed upon.

Methods for characterizing the size distribution of spray drops are discussed in ASTM (1988) and by Mugele and Evans (1951). A probability distribution function, $F(D)$, is introduced that represents the fraction of drops per unit diameter range about the diameter, D, as shown in Figure 4.18.2. The spray drop sizes span a range from a minimum diameter, D_{min}, to a maximum diameter, D_{max}. It is also convenient to introduce a mean or average drop diameter instead of having to specify the complete drop size distribution. The number median drop diameter represents that drop whose diameter is such that 50% of the drops in the spray have sizes less than this size. Spray drop size distribution data can also be represented as a volume (or mass) distribution function, $F(D^3)$; this gives more weight to the large drops in the distribution. In this case, a volume median diameter or a mass median diameter can also be defined, as indicated in Figure 4.18.2.

Various other mean diameters are also in common use. These are summarized using the standard notation of Mugele and Evans (1951) as

$$(D_{jk})^{j-k} = \frac{\int_{D_{min}}^{D_{max}} D^j f(D)\,\mathrm{d}D}{\int_{D_{min}}^{D_{max}} D^k f(D)\,\mathrm{d}D} \tag{4.18.1}$$

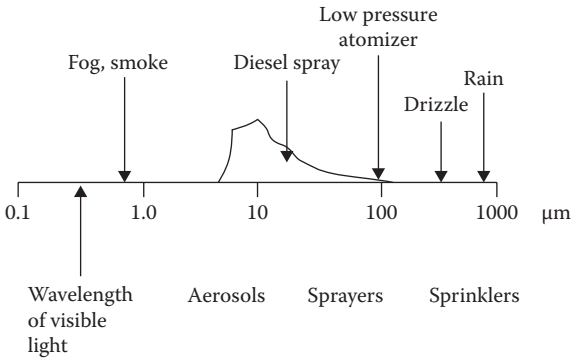


FIGURE 4.18.1 Typical average spray drop sizes for various classes of sprays. A representative size distribution is depicted for the diesel spray.

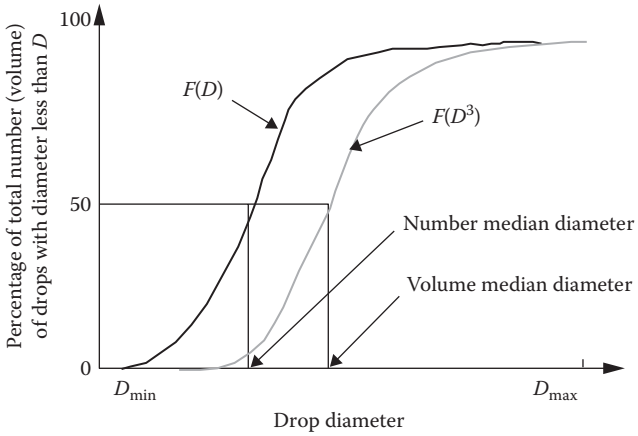


FIGURE 4.18.2 Cumulative spray drop number and volume distributions.

where $f(D) = \mathrm{d}F(D)/\mathrm{d}D$ is the drop size probability density function (usually normalized such that $\int_{D_{\min}}^{D_{\max}} f(D)\,\mathrm{d}D = 1$). Commonly used mean diameters are D_{10} (i.e., $j=1$, $k=0$, sometimes called the length mean diameter (Lefebvre, 1989) and D_{32} (i.e., $j=3$, $k=2$, called the Sauter mean diameter [SMD]). The SMD has a useful physical interpretation in combustion applications since drop vaporization rates are proportional to the surface area of the drop. It represents the size of that drop that has the same volume-to-surface area ratio as that of the entire spray.

Several distribution functions have been found to fit experimental data reasonably well. Among these are the Nukiyama–Tanasawa and the Rosin–Rammler distributions that have the general form (Lefebvre, 1989) $f(D) = aD^p \exp(\{-bD\}^q)$, where the constants a, p, b, and q characterize the size distribution. The higher the parameter, q, the more uniform the distribution, and typically $1.5 < q < 4$. Other distributions have been proposed which consist of logarithmic transformations of the normal distribution, such as $f(D) = a\exp(-y^2/2)$, where $y = \delta \ln(\eta D/(D_{\max} - D))$, and α, δ, and η are constants. In this case, the smaller δ, the more uniform the size distribution. It should be noted that there is no theoretical justification for any of these size distributions. Spray drop size distributions can be measured nonintrusively by using optical laser diffraction and phase/Doppler instruments. A discussion of these techniques and their accuracy is reviewed by Chigier (1983).

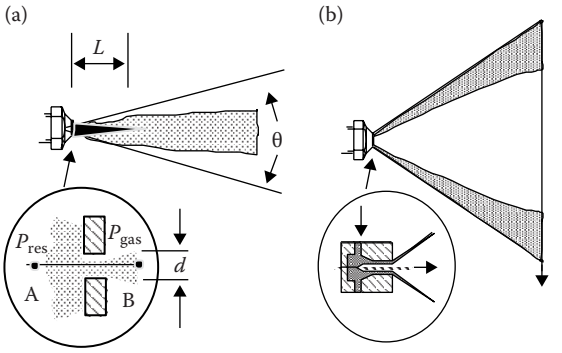


FIGURE 4.18.3 Schematic diagram of (a) solid cone and (b) hollow cone pressure atomizer sprays.

Atomizer Design Considerations

Atomization is generally achieved by forcing a liquid or a liquid–gas mixture through a small hole or slit under pressure to create thin liquid sheets or jets moving at a high relative velocity with respect to the surrounding ambient gas. Desirable characteristics of atomizers include the ability to atomize the liquid over a wide range of flow rates, low power requirements, and low susceptibility to blockage or fouling. In addition, atomizers should produce consistent sprays with uniform flow patterns in operation.

Atomizers can be broadly characterized as those producing hollow cone or solid cone sprays, as depicted in Figure 4.18.3. In solid cone (or full cone) sprays, the spray liquid is concentrated along the spray axis, Figure 4.18.3a. These sprays are useful in applications requiring high spray penetration, such as in diesel engines. In hollow cone sprays, the axis region is relatively free of drops, giving wide spray dispersal, Figure 4.18.3b. These sprays are often used in furnaces, gas turbines, and spray-coating applications.

Many different atomizer designs are found in applications. Common atomizer types include pressure, rotary, twin-fluid (air-assist, air-blast, effervescent), flashing, electrostatic, vibratory, and ultrasonic atomizers, as discussed next.

Atomizer Types

In *pressure* atomizers, atomization is achieved by means of a pressure difference, $\Delta P = P_{res} - P_{gas}$, between the liquid in the supply reservoir pressure, P_{res}, and the ambient medium pressure, P_{gas}, across a nozzle. The simplest design is the plain orifice nozzle with exit hole diameter, d, depicted in Figure 4.18.3a. The liquid emerges at the theoretical velocity $U = \sqrt{2\Delta P/\rho_{liquid}}$, the (Bernoulli) velocity along the streamline A–B in Figure 4.18.3a, where ρ_{liquid} is the density of the liquid. The actual injection velocity is less than the ideal velocity by a factor called the discharge coefficient, C_D, which is between 0.6 and 0.9 for plain hole nozzles. C_D accounts for flow losses in the nozzle.

Four main jet breakup regimes have been identified, corresponding to different combinations of liquid inertia, surface tension, and aerodynamic forces acting on the jet, as shown in Figure 4.18.4. At low injection pressures, the low-velocity liquid jet breaks up due to the unstable growth of long-wavelength waves driven by surface tension forces (Rayleigh regime). As the jet velocity is increased, the growth of disturbances on the liquid surface is enhanced because of the interaction between the liquid and the ambient gas (the first and second wind-induced breakup regimes). At high injection pressures, the high-velocity jet disintegrates into drops immediately after leaving the nozzle exit (atomization regime). Criteria for the boundaries between the regimes are available (Chigier and Reitz, 1996). Aerodynamic effects are found to become very important relative to inertial effects when the jet Weber number, $We_j > 40$, where $We_j = \rho_{gas} U^2 d/\sigma$, σ_{gas} is the gas density, and σ is the liquid surface tension.

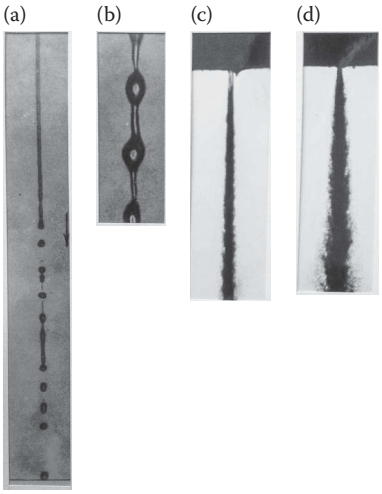


FIGURE 4.18.4 (a) Rayleigh breakup. Drop diameters are larger than the jet diameter. Breakup occurs many nozzle diameters downstream of nozzle. (b) First wind-induced regime. Drops with diameters of the order of jet diameter. Breakup occurs many nozzle diameters downstream of nozzle. (c) Second wind-induced regime. Drop sizes smaller than the jet diameter. Breakup starts some distance downstream of nozzle. (d) Atomization regime. Drop sizes much smaller than the jet diameter. Breakup starts at nozzle exit.

Experiments show that the unstable growth of surface waves is aided by high relative velocities between the liquid and the gas, and also by high turbulence and other disturbances in the liquid and gas flows, and by the use of spray liquids with low viscosity and low surface tension.

Liquid breakup characteristics such as the spray drop size, the jet breakup length, and the spray angle have been related to the unstable wave growth mechanism. The wavelengths and growth rates of the waves can be predicted using results from a linear stability analysis with (Reitz, 1988)

$$\frac{\Lambda}{2} = 9.02\frac{(1+0.45Z^{0.5})(1+0.4T^{0.7})}{(1+0.87\mathrm{We}_2^{1.67})^{0.6}} \tag{4.18.2a}$$

$$\Omega\left(\frac{\rho_l a^3}{\sigma}\right)^{0.5} = \frac{0.34+0.38\mathrm{We}_2^{1.5}}{(1+Z)(1+1.4T^{0.6})} \tag{4.18.2b}$$

where Λ is the wavelength, Ω is the growth rate of the most unstable surface wave, and a is the liquid jet radius. The maximum wave growth rate increases, and the corresponding wavelength decreases with increasing Weber number, $\mathrm{We}_2 = \rho_{gas}U^2a/\sigma$, where U is the relative velocity between the liquid and the gas. The liquid viscosity appears in the Ohnesorge number, $Z = \mathrm{We}_l^{1/2}\mathrm{Re}_l$. Here, the Weber number We_l is based on the liquid density, the Reynolds number is $\mathrm{Re}_l = Ua/\nu_l$, ν_l is the liquid viscosity, and the parameter $T = Z\mathrm{We}_2^{1/2}$. The wave growth rate is reduced and the wavelength is increased as the liquid viscosity increases.

The size of the drops formed from the breakup process is often assumed to be proportional to the wavelength of the unstable surface waves in modeling studies (Reitz, 1988). However, the drop sizes in the primary breakup region near the nozzle exist have also been found to be influenced by the length scale of the energy-containing eddies in the turbulent liquid flow (Wu et al., 1995). There is uncertainty about atomization mechanisms since spray measurements are complicated by the high optical density of the spray in the breakup region (e.g., see Figure 4.18.4d). As the drops penetrate into the ambient gas, they interact with each other through collisions and coalescence, and the spray drop size changes dynamically within the spray as a result of secondary breakup and vaporization

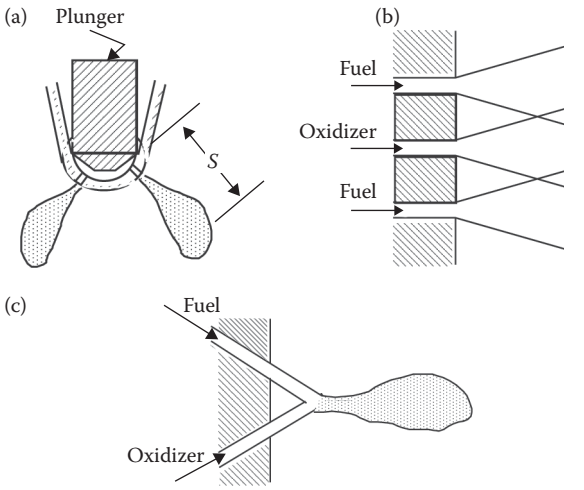


FIGURE 4.18.5 (a) Diesel injector multihole spray nozzle, (b) showerhead, and (c) doublet impingement nozzles.

effects. The drop trajectories are determined by complex drop drag, breakup, and vaporization phenomena, and by interactions with the turbulent gas flow (Reitz, 1988). Recent highly resolved simulations (HRSs) (Deshpande et al., 2015) have reported a drastic difference between the length scales predicted by linear stability analysis, for example, from Equations 4.18.2, and those observed as being responsible for primary atomization.

High-pressure diesel sprays are intermittent and are required to start and stop quickly without dribble between injections. This is accomplished by means of a plunger arrangement that is actuated by a cam and spring system in mechanical "jerk" pump systems (Figure 4.18.5). Modern electronic injectors include electromagnetic solenoids that permit the duration and injection pressure to be varied independently of each other and of engine speed. Experiments on diesel-type injector nozzles show that the penetration distance, S, of the tip of the spray at time, t, after the beginning of the injection is given by (Hiroyasu and Arai, 1978)

$$\begin{aligned} S &= 0.39Ut(\rho_{\text{liquid}}/\rho_{\text{gas}})^{1/2} \quad \text{for} \quad t < t_b \\ S &= 2.46\sqrt{Udt}\,(\rho_{\text{liquid}}/\rho_{\text{gas}})^{1/4} \quad \text{for} \quad t > t_b \end{aligned} \tag{4.18.3}$$

where the "breakup time" is $t_b = 40.5d(\rho_{\text{liquid}}/\rho_{\text{gas}})^{1/2}/U$. The jet breakup length (see Figure 4.18.3a), $L = Ut_b$ is independent of the injection velocity. On the other hand, for low-speed jets, or for jets injected into a low-gas-density environment, $t_b = 1.04C\,(\rho_{\text{liquid}}d^3/\sigma)^{1/2}$, where C is a constant typically between 12 and 16 and σ is the surface tension. In this case, L increases with the injection velocity (Reitz and Bracco, 1986). The functional form of the above jet breakup time and length correlations can be derived for an inviscid liquid in the limits of large and small Weber number, We_2 from the unstable wave growth rate in Equation 4.18.2 with $t_b \sim \Omega^{-1}$.

For high-speed diesel-type jets in the atomization regime, the resulting spray diverges in the form of a cone with cone angle, θ, that is usually in the range from 5° to 20°. θ increases with gas density following $\tan\theta = A(\rho_{\text{gas}}/\rho_{\text{liquid}})^{1/2}$, where A is a constant that depends on the nozzle passage length and (weakly) on the injection velocity (Reitz and Bracco, 1986). Very high injection pressures are required to produce small drops. In diesel engines, ΔP is typically as high as 200 MPa, and drops are produced with mean diameters of the order of 10 μm (see Figure 4.18.1). Drop size correlations have been proposed for plain-orifice sprays, such as that presented in Table 4.18.1 (Lefebvre, 1989). Note, however, that these correlations do not account for the fact that the spray drop size varies with

TABLE 4.18.1
Representative Drop Size Correlations for Various Spray Devices (Dimensional Quantities Are in SI Units, kg, m, s)

Device	Correlation	Notes
Plain orifice	$\text{SMD} = 3.08\nu_l^{0.385}(\rho_{\text{liquid}}\sigma)^{0.737}\rho_{\text{gas}}^{0.06}\Delta P^{-0.54}$	Use SI units
Fan spray	$\text{SMD} = 2.38d_h(\sigma\mu_{\text{liquid}}^2/\rho_{\text{gas}}d_h^3\Delta P^2)^{0.25}$ $+ 0.26d_h(\sigma\rho_{\text{liquid}}/\rho_{\text{gas}}d_h\Delta P^2)^{0.25}$	d_h = nozzle hydraulic diameter
Rotary atomizer	$\text{SMD} = 0.119Q^{0.1}\sigma^{0.5}/Nd^{0.5}\rho_{\text{liquid}}^{0.4}\mu_{\text{liquid}}^{0.1}$	N = rotational speed (rev/s), Q = volumetric flow rate, A_{inj} U
Pressure swirl	$\text{SMD} = 4.52(\sigma\mu_{\text{liquid}}^2/\rho_{\text{gas}}\Delta P^2)^{0.25}(t\cos\theta)^{0.25}$ $+ 0.39(\sigma\rho_{\text{liquid}}/\rho_{\text{gas}}\Delta P)^{0.25}(t\cos\theta)^{0.75}$ $t = 0.0114A_{\text{inj}}\rho_{\text{liquid}}^{1/2}d\cos\theta$	t = film thickness; θ = cone angle, d = discharge orifice diameter
Twin-fluid/air-blast	$\text{SMD} = 0.48d\left(\sigma/\rho_{\text{gas}}U^2d\right)^{0.2}(1+1/\text{ALR})^{0.4}$ $+ 0.15d(\mu_{\text{liquid}}^2/\sigma\rho_{\text{liquid}}d)^{0.5}(1+1/\text{ALR})$	ALR = air-to-liquid mass ratio
Prefilming air-blast	$\text{SMD} = (1+1/\text{ALR})\left[0.33d_h(\sigma/\rho_{\text{gas}}U^2d_p)^{0.6}\right.$ $\left.+ 0.068d_h(\mu_{\text{liquid}}^2/\sigma\rho_{\text{liquid}}d_p)^{0.5}\right]$	d_h = hydraulic diameter, d_p = prefilmer diameter, Figure 4.18.9
Ultrasonic	$\text{SMD} = \left(4\pi^3\sigma/\rho_{\text{liquid}}\omega^2\right)^{1/3}$	ω = vibration frequency

Source: Lefebvre, A.H., *Atomization and Sprays*, Hemisphere Publishing, New York, 1989. With permission.

time, and from place to place in the spray. Moreover, experimental correlations often do not include some parameters that are known to influence spray drop sizes, such as the nozzle passage length and its entrance geometry. Therefore, overall drop size correlations should only be used with caution.

The plain-orifice design is also used in twin-fluid-type liquid rocket engines in showerhead and doublet designs (Figures 4.18.5b and c). In the case of doublet nozzles, shown in Figure 4.18.6c, the impinging jets create unstable liquid sheets which break up to produce the sprays. Drop size correlations are available for liquid sheets such as those formed by discharging the liquid through a rectangular slit (see fan spray, Table 4.18.1). Thin liquid sheets or slits lead to the production of small drops. The breakup mechanism of liquid sheets is also thought to involve the unstable growth of surface waves due to surface tension and aerodynamic forces (Chigier and Reitz, 1996).

In rotary atomizers, centrifugal forces are used to further enhance the breakup process. In this case, the liquid is supplied to the center of a spinning disk and liquid sheets or ligaments are thrown off the edges of the disk. The drop size depends on the rotational speed of the disk, as indicated in Table 4.18.1.

A spinning wheel or cup (turbobell) is used in some spray-painting applications. The spray shape is controlled by supplying a coflowing stream of "shaping air."

Centrifugal forces also play a role in the breakup mechanism of pressure swirl atomizers (simplex nozzles). These atomizers give wider spray cone angle than plain orifice nozzles, and are available in hollow cone and solid cone designs. As depicted in Figure 4.18.3b, the spray liquid enters a swirl chamber tangentially to create a swirling liquid sheet. The air-core vortex within the swirl chamber plays an important role in determining the thickness of the liquid sheet or film at the nozzle exit. This type of nozzle produces relatively coarse sprays. A representative SMD correction

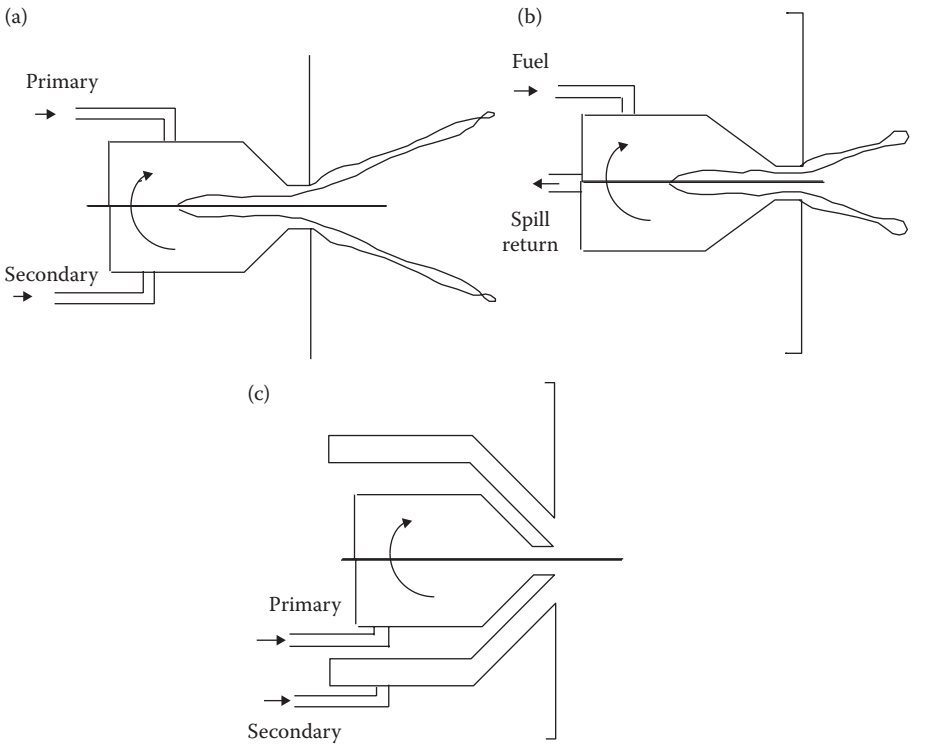


FIGURE 4.18.6 (a) Duplex, (b) dual-orifice, and (c) spill-return-type nozzle designs.

is listed in Table 4.18.1. The spray cone angle depends on the ratio of the axial and tangential liquid velocity components at the exit of the nozzle. This type of atomizer is not well suited for use in transient applications because it tends to dribble at start-up and to shut down when the air-core is not fully formed.

The basic drawback of all pressure atomizers is that the flow rate depends on the square root of ΔP. The volumetric flow rate is $Q=A_{inj}U$, where A_{inj} is the liquid flow area at the nozzle exit, so that a factor of 20 increase in flow rate (a typical turndown ratio from idle to full load operation of a gas turbine engine) requires a factor of 400 increase in injection pressure.

This difficulty has led to the so-called wide-range atomizer designs such as those shown in Figure 4.18.6. The duplex nozzle features two sets of tangential swirl ports; the primary (or pilot) supplies fuel at low flow rates, while the secondary ports become operational at high flow rates. Another variation is the dual-orifice nozzle which is conceptually two simplex nozzles arranged concentrically, one supplying the primary flow and the other supplying the secondary flow. The spill-return nozzle is a simplex nozzle with a rear passage that returns fuel to the injection pump. In this design, the flow rate is controlled by the relative spill amount, and there are no small passages to become plugged. However, the fuel is always supplied at the maximum pressure which increases the demands on the injection pump. But high swirl is always maintained in the swirl chamber and good atomization is achieved even at low flow rates.

In twin-fluid injectors, atomization is aided by a flow of high-velocity gas through the injector passages. The high-velocity gas stream impinges on a relatively low-velocity liquid either internally (in internal mixing nozzles, Figure 4.18.7) or externally (in external mixing designs, Figure 4.18.8). The liquid and gas flows are typically swirled in opposite directions by means of swirl vanes to improve atomization. Air-assist refers to designs that use a relatively small amount of air at high (possibly sonic) velocities. Air-blast refers to designs that use large quantities of relatively

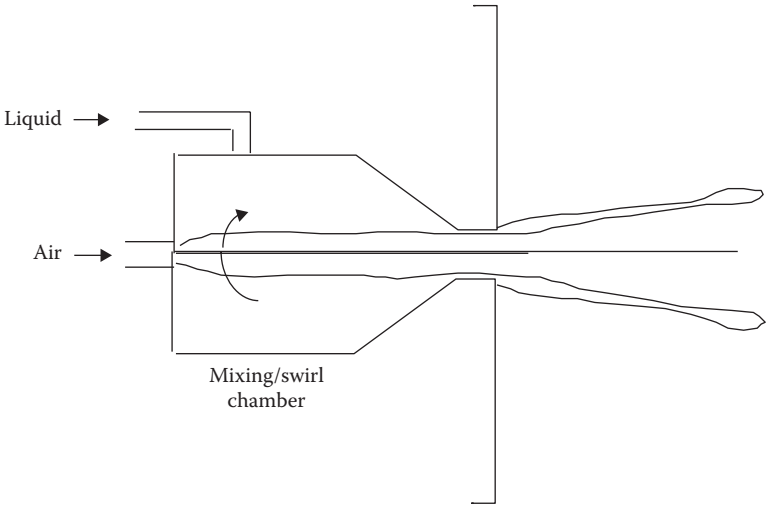


FIGURE 4.18.7 Internal-mixing twin-fluid injector design.

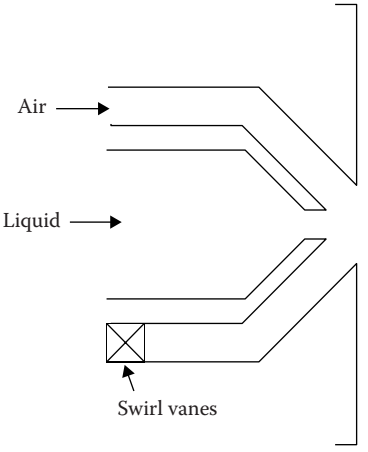


FIGURE 4.18.8 External-mixing twin-fluid injector design.

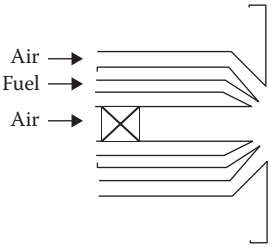


FIGURE 4.18.9 Prefilming air-blast atomizer.

low-velocity air which often supplies some of the air to help decrease soot formation in combustion systems (Lefebvre, 1989) (Figure 4.18.9).

In flashing and effervescent atomizers, a two-phase flow is passed through the injector nozzle exit. In the former, the bubbles are generated by means of a phase change which occurs as the liquid, containing a dissolved propellant gas or vapor, undergoes the pressure drop through the nozzle.

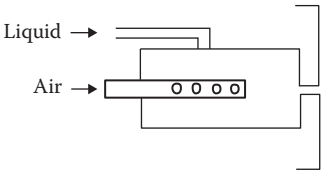


FIGURE 4.18.10 Internal-mixing, effervescent atomizer.

This process is exploited in many household spray cans, but has the disadvantage of releasing the propellant gas required for atomization into the atmosphere. In the so-called effervescent atomizer, air bubbles are introduced into the liquid upstream of the exit orifice, as depicted in Figure 4.18.10. The spray quality is found to depend weakly on the air bubble size and is independent of the nozzle exit diameter. This makes internal mixing, air-assist atomizers very attractive for use with high-viscosity fluids and slurries where nozzle plugging would otherwise be a problem (Lefebvre, 1989).

In electrostatic atomizers, the spray liquid is charged by applying a high-voltage drop across the nozzle. The dispersion of the spray drops is increased by exploiting electrical repulsive forces between the droplets. An electrostatic charge on the drops is also helpful in spray-coating applications, such as in automotive spray painting using electrostatic turbobell sprayers, since the charged drops are attracted to an oppositely charged target surface.

Other atomizer types include vibratory and ultrasonic atomizers (or nebulizers), where the drops are formed by vibrating the injector nozzle at high frequencies and at large amplitudes to produce short-wavelength disturbances to the liquid flow. Ultrasonic atomizers are used in inhalation therapy where very fine sprays (submicron sizes) are required, and an available representative drop size correlation is also listed in Table 4.18.1.

Modeling and Simulating Sprays

The numerical representation of sprays is often accomplished by describing the gas phase in terms of a continuum and the droplets as discrete entities that interact with the gas phase through exchanges of mass, momentum, and energy. This combined liquid/gas phase description is referred to in the literature as the Lagrangian–Eulerian methodology (Balachandar and Eaton, 2010). The most common approach used in practical spray calculations is to employ a Reynolds-averaged–Navier–Stokes to describe the gas phase. For instance, the following equations for mass, chemical species (i), momentum, and energy are solved:

$$\frac{\partial r}{\partial t} + \nabla \cdot (\rho \boldsymbol{u}) = \dot{\rho}_s; \tag{4.18.4a}$$

$$\frac{\partial [\rho Y_i]}{\partial t} + \nabla \cdot (\rho \boldsymbol{u} Y_i) = \nabla \cdot (\mathfrak{D} \nabla Y_i) + \dot{S}_{Y_i} + \dot{C}_{Y_i}; \tag{4.18.4b}$$

$$\frac{\partial \rho \boldsymbol{u}}{\partial t} + \nabla \cdot (\rho \boldsymbol{u}\boldsymbol{u}) = -\nabla \mathrm{P} + \nabla \cdot (\boldsymbol{T}) + \rho \boldsymbol{g} + \dot{S}_{\mathrm{M}}; \tag{4.18.4c}$$

$$\frac{\partial (\rho h)}{\partial t} + \nabla \cdot (\rho \boldsymbol{u} h) = \nabla \cdot (\alpha_{\mathrm{eff}} \nabla \mathrm{h}) + \frac{DP}{Dt} + \dot{S}_h. \tag{4.18.4d}$$

In these expressions, the terms, $\dot{\rho}_s$, $\dot{S}_{Y_i}$, $\dot{C}_{Y_i}$, $\dot{S}_{\mathrm{M}}$, and $\dot{S}_h$ represent, respectively, the source term for droplet evaporation, chemical species production from either spray and/or combustion, momentum source term from the spray, and gas energy deficit associated with spray vaporization. The viscous and turbulent contributions to the stress tensor are included in $\boldsymbol{T}$. The resulting system of unknowns consists of density, chemical species, velocity, pressure, and temperature. To close this

system of equations, an equation of state (EoS) is incorporated into the solution. Under ambient or near-ambient pressure conditions, the ideal EoS provides adequate predictions for the various thermodynamic quantities (Trujillo et al., 2004), but under higher pressures, a more elaborate treatment is necessary consisting of a more general EoS (Qiu and Reitz, 2015). These set of governing equations are generally coupled to equations governing the turbulent kinetic energy and dissipation rate, which also exhibit source terms stemming from the spray.

Theoretically, when the Reynolds-averaging procedure is applied consistently to both liquid and gas phases, the resulting expression for the liquid phase is an evolution equation for the number density ($f(\boldsymbol{x},\boldsymbol{v},r;t)$) of the spray droplets distributed over physical space ($\boldsymbol{x}$), velocity space ($\boldsymbol{v}$), and droplet radius space (r) (for the sake of expedience, we are explicitly skipping the terms associated with departures from sphericity and temperature). This equation was originally introduced by Williams (1958), and more formally derived by Subramaniam (2001). It is given by the following expression:

$$\frac{\partial(f)}{\partial t}+\nabla_x\cdot(f\boldsymbol{v})+\nabla_V\cdot(f\boldsymbol{a})+\frac{\partial}{\partial r}(f\dot{R})=f_{\text{coll}}+f_{\text{break}}\,, \tag{4.18.5}$$

where the gradients over physical and velocity space are, respectively, denoted as ∇_x and ∇_V. The droplet acceleration is given by $\mathbf{a}$, and the rate of change of droplet radius is represented by $\dot{R}$. The right-hand-side terms are sources or sinks due to droplet–droplet collisions (f_{coll}) or hydrodynamic breakup of droplets (f_{break}).

While the expression above does in fact reproduce the evolution of the spray number density, in practical applications it is never solved. The main reason is that this partial differential equation has seven independent variables, discounting temperature and departure from sphericity, and thus represents an almost insurmountable computational challenge. The common alternative, that is frequently adopted, is inspired by direct numerical simulations (DNSs) of dispersed two-phase flow (Elghobashi and Truesdell, 1993), where the solution of droplet position and velocity is performed for every droplet in the computational domain. Rather than tracking each individual droplet in spray modeling efforts, a simplification is often employed, where a given group of droplets, denoted as a parcel, is tracked in exactly the same manner as a droplet. This methodology is referred to as the superparticle method by Ferry and Balachandar (2001).

Under the superparticle methodology, the parcel represents a group of N_{p} identical droplets having the same size, velocity, temperature, position, and other dependent quantities. Thus, for instance, the mass of a given parcel is given by $N_{\text{p}}\rho_{\text{L}}\frac{4}{3}\pi r_{\text{p}}^3$ and a set of relevant evolution equations for parcel position, velocity, and temperature are given as

$$\frac{\mathrm{d}\boldsymbol{x}_{\text{p}}}{\mathrm{d}t}=\boldsymbol{V}_{\text{p}}; \tag{4.18.6a}$$

$$m_{\text{p}}\frac{\mathrm{d}\boldsymbol{V}_{\text{p}}}{\mathrm{d}t}=\frac{\pi D_{\text{p}}^2}{8}\rho_{\text{G}}C_{\text{D}}[\boldsymbol{u}(x_{\text{p}},t)-\boldsymbol{V}_{\text{p}}]\left|\boldsymbol{u}(x_{\text{p}},t)-\boldsymbol{V}_{\text{p}}\right|+\rho_{\text{L}}\boldsymbol{g}\frac{\pi D_{\text{p}}^3}{8}; \quad \text{and} \tag{4.18.6b}$$

$$\frac{\mathrm{d}T_{\text{p}}}{\mathrm{d}t}=\frac{1}{C_{\text{p,p}}m_{\text{p}}}\left[\dot{Q}_{\text{p}}+\dot{m}_{\text{p}}h_{\text{fg}}\right]. \tag{4.18.6c}$$

The sensible heat transfer and heat transfer from phase change are shown, respectively, in the first and second right-hand-side terms of Equation 4.18.6c.

Other more rigorous interpretations for N_{p} based on the droplet distribution function or spray number density are provided in the work of Subramaniam (2001, 2013). The key idea communicated here is that the statistical representation of the spray in terms of parcels must converge to the same statistics as if all of the droplets were individually tracked. In this sense, N_{p} really becomes a

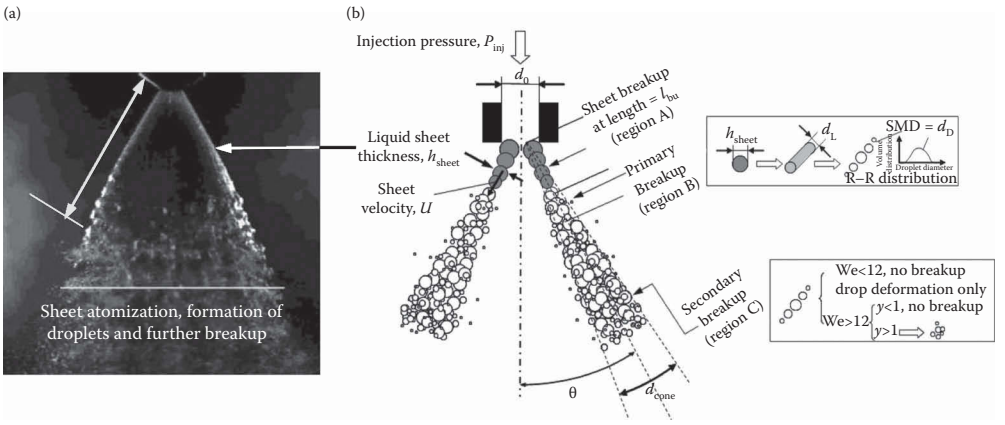


FIGURE 4.18.11 Qualitative comparison between an experimental hollow cone spray (a) and its modeling counterpart (b). Region A in the spray modeling consists of numerical parcels representing the intact liquid sheet. In region B, these initial parcels undergo breakup as dictated by primary atomization models, and in region C, they undergo further breakup governed by secondary atomization models. The authors thank Chris Brown and Vince McDonell for providing the photo image.

statistical weight for the parcel and care must be taken to ensure that as the spray evolves in time, the statistics from the parcel representation remain consistent with the physical spray.

One of the major drawbacks from the Lagrangian–Eulerian treatment is the manner in which primary atomization is treated. To illustrate this, Figure 4.18.11a shows a picture of a hollow-cone spray and its counterpart modeling representation in Figure 4.18.11b. In the modeling procedure, droplets of a specified size are introduced at the nozzle orifice, and an instability model is employed to determine their subsequent breakup. This treatment leads to a size distribution of droplets as shown in one of the encapsulated boxes on the far right of the figure. Additionally, if the resulting droplets are aerodynamically unstable, a second breakup procedure is employed (for secondary atomization), where the corresponding droplet Weber number is used as a discriminating parameter for judging instability.

If we compare this procedure to the actual breakup process as conveyed in Figure 4.18.11a, the most apparent feature is that for a substantial distance from the nozzle orifice, the sheet has remained largely intact; nevertheless, its corresponding modeling representation, at the same locations, is given in terms of droplets. This implies that the predicted momentum interaction with the surrounding gas phase is expected to deviate from experimental findings, since the extent of the gas–liquid interface and its shape is so drastically different between a liquid sheet and a group of droplets. In fact, it is precisely this discrepancy that leads to significant deviations in modeling attempts as presented in Figure 4.18.12, where the experimental image of the spray (Figure 4.18.12a) is compared to the modeling results (Figure 4.18.12b). The red dashed line corresponds to the centerline axis in experiments, and the yellow line is the respective computational axis. The difference between them is substantial and directly affects the downstream spray evolution.

Efforts to address experimental/modeling discrepancies often lead to an optimization of modeling constants for a specific spray configuration. For instance, an optimized set of modeling constants can be used for fuel injection under diesel spray conditions with reasonable good agreement among various diesel spray data sets. However, exercise of these modeling constants for a completely different spray, such as the hollow water spray shown in Figures 4.18.11 and 4.18.12 can potentially, and does in fact, lead to significant discrepancies. While another optimization exercise can be conducted to arrive at better constants for this type of hollow-cone spray, a need exists for generalizing the liquid/gas phase treatment in the near-field and reducing the existing level of empiricism. To

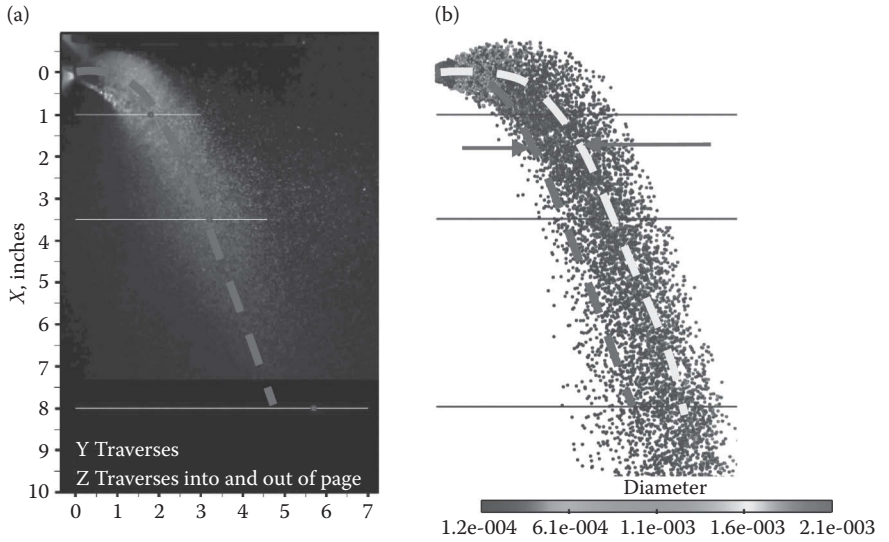


FIGURE 4.18.12 Comparison between the experimental spray axis trajectory, shown in red (a), and the computed axis shown in yellow (b). The liquid injection velocity is 16.6 m/s and the spray is introduced into a cross flow of air at 53 m/s. The initial sheet thickness is 222 μm. Chris Brown and Vince McDonell are gratefully acknowledged for providing the photo image.

address this need, an alternative procedure has been introduced over the last decade, where the goal is to compute from first principles the atomization and spray forming process. This is discussed in the following section.

HRS Strategy

The strategy of performing large-scale numerical simulations of sprays has gained considerable momentum over the last decade (Gorokhovski and Herrmann, 2008; Jiang et al., 2010; Jarrahbashi and Sirignano, 2014; Shinjo et al., 2015). This simulation strategy aims to resolve the injection and the ensuing atomization process directly by solving the underlying equations at the required level of spatial and temporal resolution. Consequently, the Navier–Stokes are solved accounting for changes in density and viscosity between the liquid and gas phases, and incorporating surface tension effects (Chang et al., 1996). This approach is commonly referred in the literature as the DNS methodology; however, it should be kept in mind that due to progressively decreasing length and timescales near the point of hydrodynamic breakup, it is almost impossible to provide adequate level of numerical resolution for all liquid structures present under practical spray conditions. Furthermore, it is not sufficient to adequately resolve the process spatially, but also temporally, and this is particularly difficult since the stable time step size decreases in proportion to $\Delta x^{3/2}$ (Galusinski and Vigneaux, 2008), which again imposes significant computational limitations. Hence, in this document, we employ the term highly resolved simulations, that is, HRS, to more accurately reflect the nature of this type of simulations.

An illustration of HRS approach is contrasted with the traditional modeling methodology in Figure 4.18.13. A key advantage of the HRS is its ability to capture the physics occurring in the near-field of the nozzle. This is where the greatest momentum interaction between the liquid and gas phase takes place, and is the region responsible for initiating much of the turbulence that is responsible for the subsequent mixing of the gas and fuel vapor.

Even though the level of resolution in HRS may be not as high as desirable, a relatively broad spectrum of length scales for liquid structures can still be obtained. To demonstrate this qualitatively, results of the atomization of a liquid jet (diameter = 200 μm) being injected into a gas

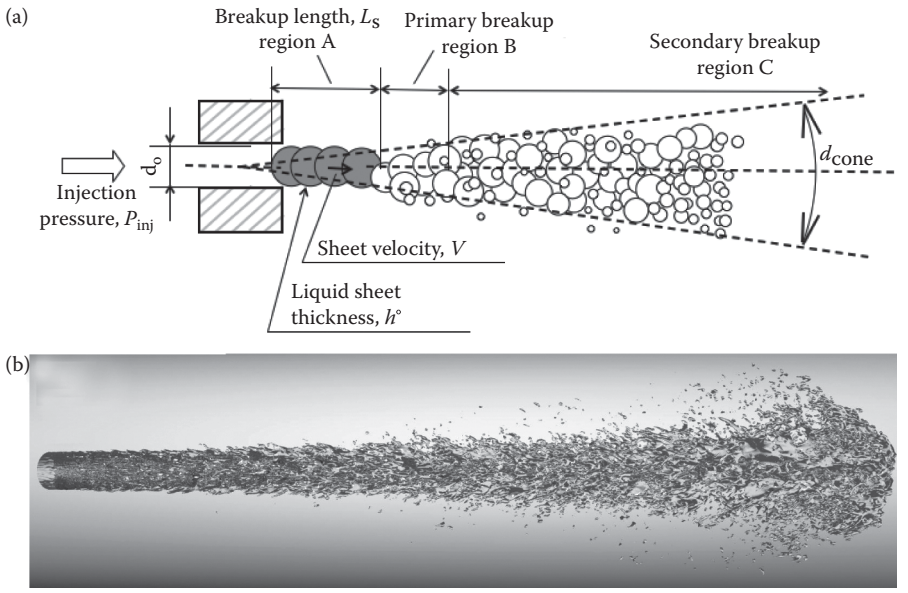


FIGURE 4.18.13 Illustration of the differences between (a) modeling sprays and (b) HRSs of sprays. In modeling, a spectrum of the spray phenomena, for instance, the primary atomization region as depicted in (a), is solved based on a much simpler mathematical representation, whereas in simulations, the actual underlying equations are numerically solved without resorting to approximations.

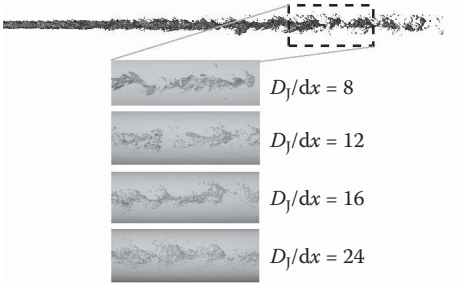


FIGURE 4.18.14 The resolution of the breakup of the intact liquid core is noticeable in these simulations performed under increasing levels of resolution. Whereas the large-scale structures are captured even in coarse-grid simulations, the small-scale droplet population requires an intensive level of refinement.

environment ($\rho_G = 58$ kg/m^3) at 200 m/s are shown in Figure 4.18.14. The upper portion of this figure displays a view of the jet extending 70 diameters downstream of the orifice. The four images below consist of close-up views of the region where the intact liquid core undergoes breakup. The calculations are performed at four different levels of spatial resolution from $D_J \Delta x = 8$ (coarsest) to $D_J / \Delta x = 25$ (finest), where D_J is the jet diameter and Δx is the local grid size. This value for Δx is used uniformly used throughout the primary atomization region with much coarser levels away from the spray. As depicted in Figure 4.18.14, the large-scale characteristics of the breakup are captured relatively well even at $D_J / \Delta x = 12$; however, the population of small droplets requires quite a bit more resolution.

Analogous to the single-phase turbulence problem, where the smallest flow scales put the most demanding requirements on the degree of grid resolution, in sprays, it is the smallest droplets which impose the greatest demands on resolution, and consequently are computationally the most expensive to capture. A popular method for circumventing this problem is the use of adaptive and local

grid refinement strategies adaptive mesh refinement (AMR). Unfortunately, for spray flows at very high injection velocities, for example, diesel, the entire primary atomization region is filled with a dense cloud of droplets, largely counterbalancing any significant benefits from AMR. This does not mean that AMR is not completely ineffective in reducing the computational cost, but only that its benefits are far from those obtained in much cleaner two-phase flow problems.

Due to the complexity involved in capturing the highly dynamic and evolving gas–liquid interface (Γ), the choice of numerical methods that can adequately be employed is confined to implicit interface capturing methodologies. In contrast to explicit surface tracking solvers, where computational nodes are placed directly on Γ, in implicit methods, a characteristic function is described over the entire 3D space from which a reconstruction of Γ can be made. The explicit methodology works very well for various free surface problems, but faces significant difficulties when used under conditions that lead to complex and frequent topological changes. For the spray problem, frequent and complex topological changes are the norm.

The most common implicit numerical strategies are the level set method (Osher and Fedkiw, 2003) and the volume-of-fluid (VoF) method (Tryggvason et al., 2011). Both of these references provide a self-contained and thorough introduction to their corresponding topic. Over the last few years, significant developments have been made in implicit solution strategies, namely, improvements in mass conservation for level set methods (Olsson and Kreiss, 2005; Kees et al., 2011), consistent mass and momentum transport with large density ratios (Raessi and Pitsch, 2012), implementation of particle-based methods for reduction in numerical diffusion in the level set advection (Enright et al., 2005), the introduction of a spectrally refined level set solution (Desjardins and Pitsch, 2009), the use of ghost-fluid cells for improving the treatment of jump conditions at the interface (Desjardins et al., 2008), implementation of an adaptive mesh refinement technique (Fuster et al., 2009), and the introduction of Hermite polynomials combined with semi-Lagrangian schemes resulting in compact numerical stencils (Nave et al., 2010). This continues to be an active area of research, where present efforts are aimed at addressing the broader physics of the spray problem including cavitation and internal nozzle numerical resolution.

One persistent computational issue has been the stringent demands of numerical resolution, particularly at high injection velocities. At present, a common held view in the spray community is the preference of mass or volume conservative methods, for instance, a VoF strategy or a hybrid VoF-level set, over a pure level set implementation. The common complaint is the loss of mass in level set methods. This loss of mass effect is essentially caused by a lower than required resolution of Γ, for instance, liquid bodies characterized by length scales that are well below the local Δx. To illustrate this problem, consider the aftermath of a Rayleigh–Plateau instability where a liquid bridge undergoing breakup results in the formation of two parent droplets and a number of satellite droplets as depicted in Figure 4.18.15.

The VoF representation accurately maintains the mass of the entire liquid system as it evolves in time, even in regions where the details of Γ are well below grid resolution. However, the shape is poorly resolved in comparison to the level set scheme. Specifically, if we consider an interfacial cell, that is, a computational cell that contains a partition of Γ, the gas–liquid interface does not usually conform to a plane or to some other linear reconstruction. Additionally, Γ and its derivatives are continuous in space, not discontinuous at the boundary between computational cells. In VoF methods, some of the most popular reconstruction schemes are those that are explicitly linear, or if they are not, have difficulty maintaining the smoothness of Γ between adjacent interfacial computational cells. Other popular VoF methods, such as the algebraic VoF procedure of interFoam (Deshpande et al., 2012), which is distributed within the open-source toolbox OpenFOAM does not perform a subgrid reconstruction of the interface. It relies on clever procedures for limiting the degree of numerical diffusion, and thus is able to contain Γ to within a band having a thickness of approximately two to three computational cells.

In spite of the high numerical accuracy of advanced level set methods and their reinitialization strategies for transporting Γ (Desjardins and Pitsch, 2009; Nave et al., 2010; Anumolu and Trujillo,

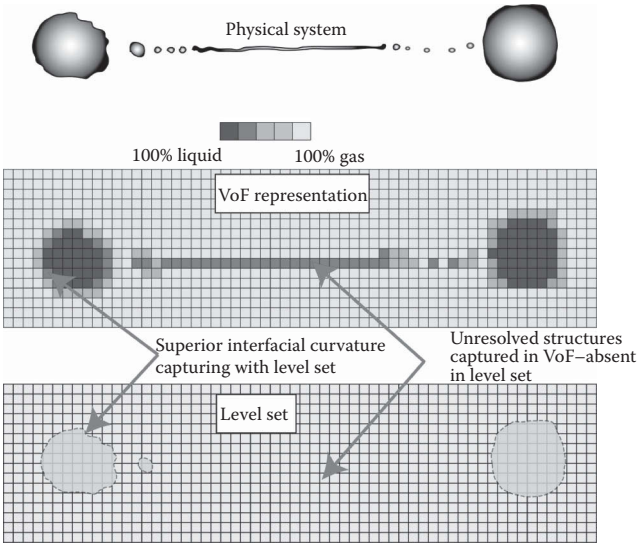


FIGURE 4.18.15 The underlying representation afforded by the VoF and level set methods are contrasted corresponding to the formation of ligaments and satellite droplets following a Rayleigh–Plateau instability. The level set scheme provides much more accurate Γ representation, but suffers in regions where the liquid structures are below grid resolution. The VoF method remains mass conservative, but does not do well in representing the shape of Γ.

2013), realistic spray problems often result in the formation of such small droplets and associated liquid structures, that the required resolution is simply too high to enforce. Consequently, the choice for VoF or other mass conserving methods is often made, particularly in cases of inadequate resolution. Under these demanding conditions, the symptoms of numerical errors for VoF are generally erroneous predicted shapes for Γ, whereas for level set methods, a common symptom is the complete elimination of underresolved liquid structures. Since this is a much more impactful consequence, it is much easier to accept the VoF errors. Future work in dealing with the resolution issue is currently being invested in the design of massively parallel computations leveraging the largest computational resources on the globe.

REFERENCES

American Society for Testing and Materials (ASTM). (1988). Data criteria and processing for liquid drop size analysis. Standard E799.

Anumolu, L., and M.F. Trujillo. (2013) Gradient augmented reinitialization scheme for the level set method. *International Journal for Numerical Methods in Fluids*, 73:1011–1041.

Balachandar, S., and J.K. Eaton. (2010) Turbulent dispersed multiphase flow. *Annual Review of Fluid Mechanics*, 42:111–133.

Chang, Y.C., T.Y. Hou, B. Merriman, and S. Osher. (1996) A level set formulation of Eulerian interface capturing methods for incompressible fluid flows. *Journal of Computational Physics*, 124:449–464.

Chigier, N.A. (1983). Drop size and velocity instrumentation. *Progress in Energy and Combustion Science*, 9:155–177.

Chigier, N., and R.D. Reitz. (1996) Regimes of jet breakup, in *Progress in Astronautics and Aeronautics Series*, K. Kuo, Ed., AIAA, New York, Chapter 4, pp. 109–135.

Deshpande, S. S., L. Anumolu, and M.F. Trujillo. (2012) Evaluating the performance of the two-phase flow solver interfoam. *Computational Science & Discovery*, 5:1–36.

Deshpande, S.S., S.R. Gurjar, and M.F. Trujillo. (2015) A computational study of an atomizing liquid sheet. *Physics of Fluids*, 27:1–26.

Desjardins, O., and H. Pitsch. (2009) A spectrally refined interface approach for simulating multiphase flows. *Journal of Computational Physics*, 228(5):1658–1677.

Desjardins, O., V. Moureau, and H. Pitsch. (2008) An accurate conservative level set/ghost fluid method for simulating turbulent atomization. *Journal of Computational Physics*, 227(18):8395–8416.

Elghobashi, S., and G.C. Truesdell. (1993) On the two-way interaction between homogeneous turbulence and dispersed solid particles. I: Turbulence modification. *Physics of Fluids A*, 5:1790–801.

Enright, D., F. Losasso, and R. Fedkiw. (2005) A fast and accurate semi-Lagrangian particle level set method. *Computers and Structures*, 83(6–7):479–490.

Ferry, J., and S. Balachandar. (2001) A fast Eulerian method for dispersed two-phase flow. *International Journal of Multiphase Flow*, 27: 1199–1226.

Fuster, D., A. Bagué, T. Boeck, L. Le Moyne, A. Leboissetier, S. Popinet, P. Ray, R. Scardovelli, and S. Zaleski. (2009) Simulation of primary atomization with an octree adaptive mesh refinement and VoF method. *International Journal of Multiphase Flow*, 35:550–565.

Galusinski, C., and P. Vigneaux. (2008) On stability condition for bifluid flows with surface tension: application to microfluidics. *Journal of Computational Physics*, 227:6140–6164.

Gorokhovski, M., and M. Herrmann. (2008) Modeling primary atomization. *Annual Review of Fluid Mechanics*, 40:343–366.

Hiroyasu, H., and M. Arai. (1978) Fuel spray penetration and spray angle in diesel engines. *Transactions of JSAE*, 34, 3208.

Jarrahbashi, D., and W.A. Sirignano. (2014) Vorticity dynamics for transient high-pressure liquid injection. *Physics of Fluids*, 26:1–52.

Jiang, X., G.A. Siamas, K. Jagus, and T.G. Karayiannis. (2010) Physical modeling and advanced simulations of gas-liquid two-phase jet flows in atomization and sprays. *Progress in Energy and Combustion Science*, 36:131–167.

Kees, C.E., I. Akkerman, M.W. Farthing, and Y. Bazilevs. (2011) A conservative level set method suitable for variable-order approximations and unstructured meshes. *Journal of Computational Physics*, 230:4536–4558.

Lefebvre, A.H. (1989) *Atomization and Sprays*, Hemisphere Publishing, New York.

Mugele, R., and Evans, H.D. (1951) Droplet size distributions in sprays, *Industrial & Engineering Chemistry*, 43:1317–1324.

Nave, J.-C., R.R. Rosales, and B. Seibold. (2010) A gradient-augmented level set with an optimally local, coherent advection scheme. *Journal of Computational Physics*, 229:3802–3827.

Olsson, E., and G. Kreiss. (2005) A conservative level set method for two phase flow. *Journal of Computational Physics*, 210(1):225–246.

Osher, S., and R. Fedkiw. (2003) *Level Set Methods and Dynamic Implicit Surfaces*. Applied Mathematical Sciences, Vol. 153, Springer, New York.

Qiu, L., and Reitz, R.D. (2015) An investigation of thermodynamic states during high-pressure fuel injection using equilibrium thermodynamics. *International Journal of Multiphase Flow* 72:24–38.

Raessi, M., and H. Pitsch. (2012) Consistent mass and momentum transport for simulating incompressible interfacial flows with large density ratios using the level set method. *Computers and Fluids*, 63:70–81.

Reitz, R.D. (1988) Modeling atomization processes in high-pressure vaporizing sprays. *Atomization and Spray Technology*, 3:309–337.

Reitz, R.D., and Bracco, F.V. (1986) Mechanisms of breakup of round liquid jets, in *The Encyclopedia of Fluid Mechanics*, Vol. 3, N. Cheremisnoff, Ed., Gulf Publishing, Houston, TX, Chapter 10, pp. 233–249.

Shinjo, J., and A. Umemura. (2011) Surface instability and primary atomization characteristics of straight liquid jet sprays. *International Journal of Multiphase Flow*, 37:1294–1304.

Shinjo, J., J. Xia, and A. Umemura. (2015) Droplet/ligament of local small-scale turbulence and scalar mixing in a dense fuel spray. *Proceedings of the Combustion Institute*, 35:1595–1602.

Subramaniam, S. (2001) Statistical modeling of sprays using the droplet distribution function. *Physics of Fluids*, 13:624–642.

Subramaniam, S. (2013) Lagrangian-Eulerian methods for multiphase flows. *Progress in Energy and Combustion Science*, 39:215–245.

Trujillo, M.F., D.J. Torres, and P.J. O'Rourke. (2004) High-pressure multicomponent liquid sprays: Departure from ideal behavior. *International Journal of Engine Research*, 5:229–246.

Tryggvason, G., R. Scardovelli, and S. Zaleski. (2011) *Direct Numerical Simulations of Gas-Liquid Multiphase Flows*. Cambridge University Press, New York.

Williams, F.A. (1958) Spray combustion and atomization. *Physics of Fluids* 1:541.

Wu, P.-K., Miranda, R.F., and Faeth, G.M. (1995) Effects of initial flow conditions on primary breakup of nonturbulent and turbulent round liquid jets, *Atomization and Sprays*, 5:175–196.

Further Information

Information about recent work in the field of atomization and sprays can be obtained through participation in the Institutes for Liquid Atomization and Spraying Systems (ILASS-Americas, -Europe, -Japan, -Korea). These regional ILASS sections hold annual meetings. An international congress (ICLASS) is also held biennially. More information is available on the ILASS-Americas homepage at http://ucicl.eng.uci.edu/ilass. Affiliated with the ILASS organizations is the Institute's Journal publication *Atomization and Sprays* published by Begell House, Inc., New York.

4.19 HEAT TRANSFER IN PLASMA SPRAYS

Milind A. Jog

Introduction

Thermal plasmas are increasingly used for manufacturing and materials processing applications which include plasma spraying, thermal plasma chemical vapor deposition, plasma sintering, plasma synthesis, plasma spheroidization, decomposition (waste-destruction), microelectronic manufacturing, and net shape manufacturing (Ayyaswamy and Cohen, 2002; Boulos et al., 1994; Graves, 1989; Rojas et al., 2009; Smith et al., 1989; Verdelle et al., 2015). The high temperature sustained by plasmas facilitates energy transport and the free electrons and ions provide an environment for chemical reactions (Mostaghimi and Boulos, 2015). In many of these applications, the heat transport from plasma to particles is critical in determining the quality of the products. For example, in plasma spheroidization, irregular fine powder particles are injected in a thermal plasma. Heat transfer from the plasma results in melting the particles. Under ideal conditions, the molten droplet thus formed becomes near spherical under the action of surface tension. Upon quenching, the droplets solidify as spheres. The spherical particles can be compacted to form near net-shape products. This method has been used to form carbide–metal composites. The heat transfer from the plasma determines the temperature, the degree of melting, and therefore the shape, eventual porosity, and quality of the product. In plasma spraying, powder of the coating material is introduced in plasma flow. Heat transfer from the plasma results in melting the powder. The molten droplets thus formed then impinge on a surface to be coated. Subsequent spreading and solidification forms a thin coat. Such techniques are used for example in aerospace industries for thermal barrier coating of turbine blades and in biomaterials industries to coat artificial bones. The quality and uniformity of the coating depends upon the velocity, temperature, and shape of the drop as it impacts the surface (Fukai et al., 1993; Liu et al., 1994a,b; Trapaga and Szekely, 1991; Trapaga et al., 1992; Vardelle et al., 1982, 2015). More recently, nano-scale coatings have been achieved by plasma spraying of suspensions and solutions (Fauchais et al., 2015). It is evident from these examples that the plasma-particle heat transport is of immense importance in achieving the desired product quality in materials processing applications. The prediction and control of the plasma-particle heat transfer is therefore critical to the improvements of these manufacturing methods and in the design of the required plasma systems (Chen et al., 1997). This chapter summarizes the state-of-the-art in heat transfer analysis and provides design correlations useful in applications. The drag force acting on the particle and the rate of heat transfer from plasma are the two important quantities of interest. These are usually expressed in dimensionless form in terms of drag coefficient C_D and Nusselt number Nu. These are defined as:

$$C_D = \frac{F_D}{\frac{1}{2}\rho U_\infty^2 A_F}, \tag{4.19.1}$$

where F_D is the drag force, ρ is the gas density, U_∞ is the free stream velocity, and A_F is the frontal area. For a spherical particle, the frontal area is $\frac{\pi d^2}{4}$, where d is the sphere diameter.

$$\mathrm{Nu} = \frac{hd}{k}, \tag{4.19.2}$$

where h is the heat transfer coefficient, d is the particle diameter, and k is the gas thermal conductivity.

Defining Features of Plasma Heat Transfer

Computational modeling as well as experimental measurements of particle–plasma heat transfer are challenging due to a number of unique features of thermal plasma flows. The plasma velocities, temperatures and consequently plasma electrical–thermophysical properties can change by orders of magnitude over a short distance. Furthermore, the charged species production and recombination processes have a strong nonlinear dependence on electron temperature and number densities. The transport in plasma, the charged particle processes, and the electron–ion recombination at the particle surface must all be considered to determine the heat transfer from plasma. To study the heat transfer from plasma, the hydrodynamic, the thermal, the electrical effects, and the surface phenomena must all be considered. Effects of large property variations, electric field effects, coupled with nonspherical particles and multiparticle interactions are elaborated below.

Effect of Electron–Ion Transport on Particle Heat Transfer

The analysis and modeling of heat transfer from plasma to a solid/liquid surface is significantly more complicated than that for a nonionized gas by the fact that it involves, over and above the hydrodynamic and the thermal boundary layers, an electrical sheath (sometimes referred to as the electrical boundary layer) close to the surface. This is caused by the low electron and ion number densities in the vicinity of a solid surface due to their recombination at surfaces (Kuryshev and Sakhin, 2001; Pfender, 1989; Zhang, 2005). As a result, the electric field near a surface can be significant even if the far away plasma is field free. The processes that take place in the electric sheath region govern the motion of electrons and ions, their recombination, and contribute to the heat transfer from the plasma to the surface. The relative sizes of the hydrodynamic and thermal boundary layers with respect to the electric sheath, and their interaction determines the flow field close to the boundary and therefore determines the heat/mass transport to the boundary.

Large Variations in Thermophysical Properties

Thermal conductivity, mass diffusivity, density, and viscosity are strongly dependent on temperature for gases and plasmas. The thermal conductivity, diffusivity, and viscosity increase with temperature with a roughly square root dependence on temperature for an ideal gas at low pressures. Once ionization and dissociation are taken into consideration, the plasma properties have a more complex dependence on temperature. Furthermore, the particle surface temperature is close to its melting temperature (~1000–3000 K) (Streibl et al., 2006; Xiong et al., 2006), and away from the particle surface, within a relatively short distance, temperature may reach the free stream temperature (>10,000 K) (Schein et al., 2015). Consequently, the thermophysical properties may change by an order of magnitude in the flow field around a particle. As a result, the flow and temperature variation around a particle injected in thermal plasma flow can be quite different from that of a particle injected in a gas flow where thermophysical property variation is relatively small. Detailed thermophysical and electrical properties of plasmas are available in Wen (2004), Brown (1967), and Devoto (1963).

Noncontinuum Effects

Powder particle injected in plasma sprays can range from a micrometer to hundreds of micrometers. In many applications, low-pressure plasmas may be used. Additionally, the high temperatures present in plasma spraying systems lead to low density of gases. A combination of these effects can give rise to Knudsen numbers ($\mathrm{Kn} = \frac{\lambda_m}{d}$, λ_m = mean free path) in the noncontinuum range. These effects have to be considered in plasma systems as well.

Nonspherical Particle Shapes

Powder particles typically injected in plasma spray system may not be spherical. Based on the size of the particle, the flow Reynolds number may be in the intermediate regime. Once the particles get heated and begin to melt, their shape deforms. The molten drops thus formed may attain an

TABLE 4.19.1
Correlations for Drag Coefficient for Plasma Flow over an Isolated Sphere in Plasma

Correlations (Hossain et al., 2007; Proulx et al., 1985)	Applicable Range of Reynolds Number	Equation Number
$C_{Df} = \frac{24}{Re}$	$Re \le 0.2$	4.19.3
$C_{Df} = \frac{24}{Re}\left(1+\frac{3}{16}Re\right)$	$0.2 < Re \le 2.0$	4.19.4
$C_{Df} = \frac{24}{Re}(1+0.11Re^{0.81})$	$2.0 < Re \le 21.0$	4.19.5
$C_{Df} = \frac{24}{Re}(1+0.189Re^{0.62})$	$21.0 < Re \le 200$	4.19.6

oblate-like or a prolate-like shape. Clift et al. (1978) provide regime map of droplet shapes in gas flow based on the Weber number and Bond number. Hence the effect on heat transfer due to nonspherical particle shape with large thermophysical property variation in the flow field needs to be taken into consideration for plasma flows.

Multiparticle Interactions

In plasma spraying process, due to the injection of a large number of particles into the plasma flow, the flow field and heat transfer around a given particle gets significantly affected by the proximity of neighboring particles. Therefore, the appropriate Nusselt number and drag coefficients for closely spaced particles are likely to be considerably different from that for an isolated sphere, and are needed to incorporate the effects due to interparticle interactions in plasma spray simulations. Simulations of particle/plasma flow and heat transfer in plasma torch have typically employed the particle-source-in-cell method (Proulx et al., 1985, 1991; Rojas et al., 2009; Sato et al., 2004). In this method, for each time step, the continuity, momentum conservation, and the energy conservation equations for the continuous phase are solved first. The drag force and heat transfer to the particles are considered as momentum/energy source terms in the corresponding equations for the continuous phase. The source terms for a computational cell are determined by summing the contribution of individual particles in that cell by using drag force and Nusselt number correlations for an isolated particle. The particle trajectories are then determined based on the calculated flow field. In addition to plasma sprays, fuel sprays in combustion systems have dense sprays where interparticle effects have been addressed. These studies are discussed in detail in a monograph by Sadhal et al. (2012) and a review by Ayyaswamy (1995).

Modeling, Data, and Design Correlations

Heat Transfer to a Single Particle

Spherical Particle

Before addressing heat transfer to a spherical particle from a plasma flow, it is worthwhile considering flow over a particle from a nonionized gas. Monographs by Clift et al. (1978) and Sadhal et al. (2012) provide extensive reviews of experimental, numerical, and analytical work in Newtonian flows while Chhabra (2007) covers non-Newtonian flows. In typical spray applications, the gas flow may be turbulent. However, the particle sizes are small and Reynolds number

based on the particle diameter is in the laminar regime. At low Re (0 < Re < ~20), the flow tends to conform to the particle surface and the local heat transfer decreases monotonically from the front stagnation point to the rear. For intermediate Re (20 < Re < ~140), there is flow separation near the rear stagnation point. The recirculatory toroidal vortex thus formed brings fluid from the free stream closer to the particle surface and the local Nusselt number increases near the rear stagnation point. From Reynolds number of about 140 onward, the flow begins to exhibit unsteady character eventually leading to vortex shedding above Re of 400. Well-established correlations are available for drag coefficient and Nusselt number for flow of nonionized gas over a spherical particle. These are often modified with appropriate correction factors to account for property variation and noncontinuum effects.

A number of models have been recommended in the literature to account for large property variations. Joshi et al. (1986) have outlined a method for calculating mass transfer in plasma flow by using integral conductivity as $I(T) = \int_{T_r}^{T} \chi(T)\,\mathrm{d}T$ and an effective conductivity as $\chi_{\text{average}} = \dfrac{I(T_\infty) - I(T_s)}{T_\infty - T_s}$. They showed that when average properties are evaluated using this integral approach, the constant property correlations for heat and mass transfer can be used in terms of effective properties. Other researchers have proposed corrections factors based on ratios of properties evaluated at the particle surface temperature and the far-field temperature. This approach of introducing correction factors to constant property drag and heat transfer correlations (Table 4.19.1) has been extensively employed in the literature for its simplicity and accuracy. This is outlined below.

Correction factor for large property variations:

$$f_1 = \left(\frac{\rho_\infty \mu_\infty}{\rho_s \mu_s}\right)^{-0.45}. \tag{4.19.7}$$

Correction factor for noncontinuum effects:

$$f_2 = \left[1 + \left(\frac{2 - a_c}{a_c}\right)\left(\frac{\gamma}{1+\gamma}\right)\frac{4}{\mathrm{Pr}_s}\mathrm{Kn}\right]^{-0.45} \quad \text{for} \quad 0.01 < \mathrm{Kn} < 1, \tag{4.19.8}$$

where a_c is the accommodation coefficient, Pr is the Prandtl number, γ is the heat capacity ratio, the subscript s indicates particles surface conditions while the subscript ∞ indicates far-field conditions.

The overall drag coefficient is calculated using the correction factors as:

$$C_D = C_{Df} f_1 f_2 \tag{4.19.9}$$

For heat transfer to an isolated spherical particle (Chyou and Pfender, 1989; Hossain et al., 2007; Young and Pfender, 1987):

$$\mathrm{Nu}_f = 2.0 + 0.6\mathrm{Re}_f^{1/2}\mathrm{Pr}_f^{1/3}, \tag{4.19.10}$$

where the subscript f denotes that the quantities are evaluated at the film temperature.

Clift et al. (1978) recommend the below correlation that covers a large range of Prandtl numbers:

$$\mathrm{Nu}_f = 1 + \left(1 + \frac{1}{\mathrm{Re}_f \mathrm{Pr}_f}\right)\mathrm{Re}_f^{0.41}\mathrm{Pr}_f^{1/3} \quad \text{for} \quad 1 \le \mathrm{Re} \le 400, \quad 0.25 \le \mathrm{Pr} \le 100. \tag{4.19.11}$$

Correction factor for property variations:

$$f_3 = \left(\frac{\rho_\infty \mu_\infty}{\rho_s \mu_s}\right)^{0.6} \left(\frac{c_{p,\infty}}{c_{p,s}}\right)^{0.38}. \tag{4.19.12}$$

Correction factor for noncontinuum effects:

$$f_4 = \left[1 + \left(\frac{2 - a_c}{a_c}\right)\left(\frac{\gamma}{1+\gamma}\right)\frac{4}{\mathrm{Pr}_s}\mathrm{Kn}\right]^{-1} \quad \text{for} \quad 0.001 < \mathrm{Kn} < 1, \tag{4.19.13}$$

$$\mathrm{Nu} = \mathrm{Nu}_f f_3 f_4, \tag{4.19.14}$$

where a_c is the accommodation coefficient, Pr is the Prandtl number, γ is the heat capacity ratio, the subscript s indicates particles surface conditions while the subscript ∞ indicates far-field conditions. Recommended value of the accommodation coefficient $a_c = 0.8$.

Nonspherical Particles

Heat transfer to nonspherical particles in plasma flow was addressed by Wen and Jog (2005). Reynolds number of 20, 50, and 100 were considered for prolate and oblate particles from disk-like (axis ratio = 1.6) to cylinder-like (axis ratio = 0.4). Particle surface temperatures between 2000–3000 K and far-field temperatures of 5,000–12,000 K were considered. Their predictions show that there is flow separation at the rear of the particle with the streamlines showing a toroidal vortex flow pattern. For a spherical particle, the flow separation begins at Re of around 20 near the rear stagnation point, while it is delayed for a prolate shape and is present at a lower Re for an oblate particle. Moreover, the size of the recirculation zone is larger for an oblate particle as it provides the largest obstruction to the flow. For a prolate particle, the gradients of velocity and temperature are the highest near the front stagnation point. However, for an oblate particle, the gradients are the highest near the lateral midpoint along the surface. Figure 4.19.1 compares the local Nusselt number for three different aspect ratios. For a prolate particle, the heat transfer coefficient is the highest near the front stagnation point and decreases monotonically along the particle surface. The heat transfer between plasma and the surface decrease plasma temperature as it moves along the particle surface. This in turn reduces the temperature gradient close to the surface and the corresponding heat transfer coefficient. A similar trend in heat transfer coefficient is evident for a sphere in the front half. However, a recirculatory vortex in the rear of the particle brings hot plasma gas close to the particle surface and increases the heat transfer coefficient. For an oblate particle, the local heat transfer coefficient is not the highest at the front stagnation point but occurs around a location about halfway (Hader and Jog, 1998b).

For plasmas, viscosity and thermal conductivity are strong functions of temperature with both properties increasing with temperature. If we compare flow field and temperature field results obtained with constant property assumptions, it is found that the variable property computations predict a smaller recirculating zone as well as higher velocity and temperature gradients compared to the constant property flow. When the Nusselt number is based on thermal conductivity evaluated at the far-field temperature, the Nusselt number values are low with higher plasma temperatures.

The effect of particle aspect ratio on the overall Nusselt number is nonlinear but monotonic. Moreover it is not a function of Reynolds number for $20 < \mathrm{Re} < 100$. The best fit for the overall Nusselt number for aspect ratios from 0.4 to 1.6 with variable property flow is given in Wen and Jog (2005) as

$$\frac{\mathrm{Nu}}{\mathrm{Nu}_{\mathrm{sphere}}} = \exp\left[-0.255\left(\frac{b}{a} - 1\right)\right], \tag{4.19.15}$$

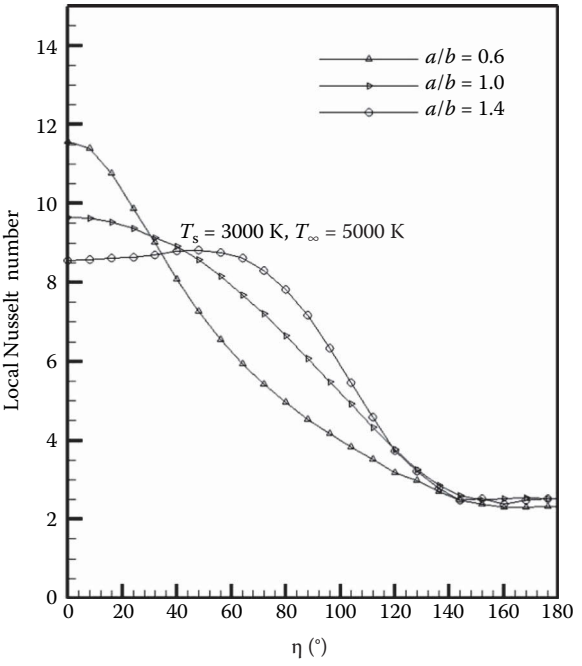


FIGURE 4.19.1 Variation of local Nusslet number for nonspherical particles. Length of the axis along the flow is 2*b* and that perpendicular to the flow is 2*a*. (Based on data from Wen, Y., and M. A. Jog, *Int. J. Heat and Fluid Flow,* 26(5), 780–791, 2005.)

where 2*b* and 2*a* are the lengths of the axis along the flow and perpendicular to the flow, respectively. This correlation deviates less than ±3% from the numerical solutions.

Heat Transfer to an Array of Particles

To determine the rate of heat transfer to particles introduced in plasma spraying, computer simulations have typically relied on source-in-cell methods (Crowe et al., 1977; Sripada et al., 1996). The continuity, momentum conservation, and the energy conservation equations as well as ionized species conservation equations are first solved in the continuous phase. The drag force and heat transfer to the particles are accounted for via source terms in the momentum and energy conservation equations. These source terms for each cell must be recalculated at each time step by summing the contribution of individual particles in that cell by using drag force and Nusselt number correlations for an isolated particle. The particle trajectories are then determined based on the calculated flow field. Similar methods have also been employed in the analysis of fuel sprays in combustion systems. To determine multiparticle interactions, a cylindrical cell model (Sripada et al., 1996; Tal et al., 1983) has been employed to describe the hydrodynamics of particle assemblages. An array of uniformly spaced spheres is analyzed (see Figure 4.19.2). By using symmetry and periodic structure of the particles in an array, the problem is reduced to one of spheres in tandem. The domain around the particles in tandem is considered axisymmetric.

Plasma flow over an array of spheres was considered by Verma and Jog (2000). They showed that the drag force and Nusselt number for the downstream spheres are lower compared to those for the first sphere. The values of C_D and Nu for the second and the third sphere are nearly equal and remain essentially the same thereafter. Because the drag coefficient and the Nusselt number for the second and the successive spheres are nearly the same, the results for the second sphere can be used for any interior sphere of the array. Verma and Jog have presented the results for drag coefficient and Nusselt number for such an interior particle normalized by those for an isolated

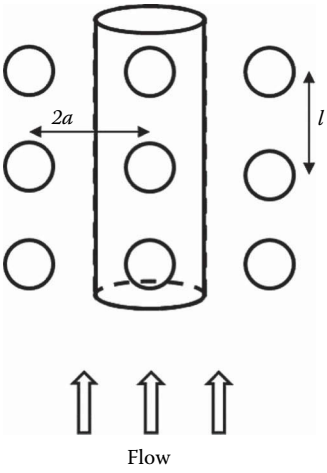


FIGURE 4.19.2 Schematic of the cylindrical cell model.

sphere. The results for normalized drag coefficient are shown in Figure 4.19.3, whereas those for the normalized Nusselt number are shown in Figure 4.19.4. When the longitudinal distance between successive particles is low, the wake region of the forward particle creates a low-pressure region in front of the next particle and the pressure drag decreases. Furthermore, the recirculating flow in the wake region can cause reduction or reversal in the direction of local shear stress in the front surface of the following sphere and reduces friction drag. Therefore, the total drag coefficient decreases with the decrease in longitudinal spacing. When longitudinal spacing is held constant and lateral spacing is decreased, it results in an increased drag coefficient. The flow

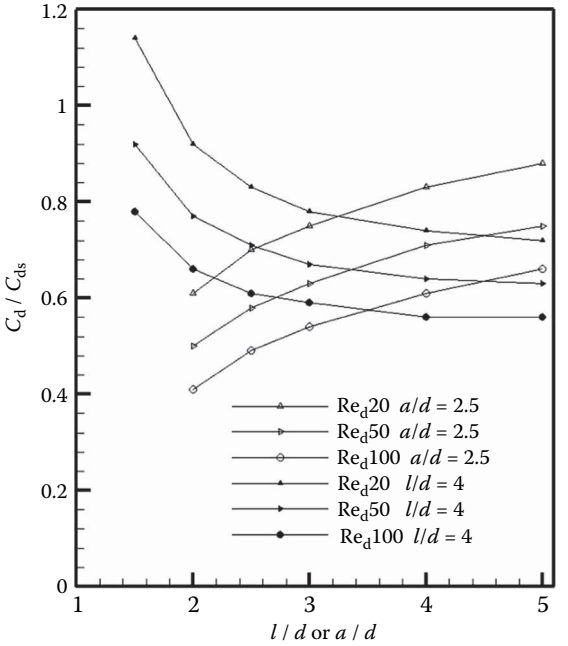


FIGURE 4.19.3 Normalized drag coefficient for particle arrays. (Based on data from Verma, A. and M. A. Jog, *Int. J. Heat and Mass Transfer,* 43, 101–112, 2000.)

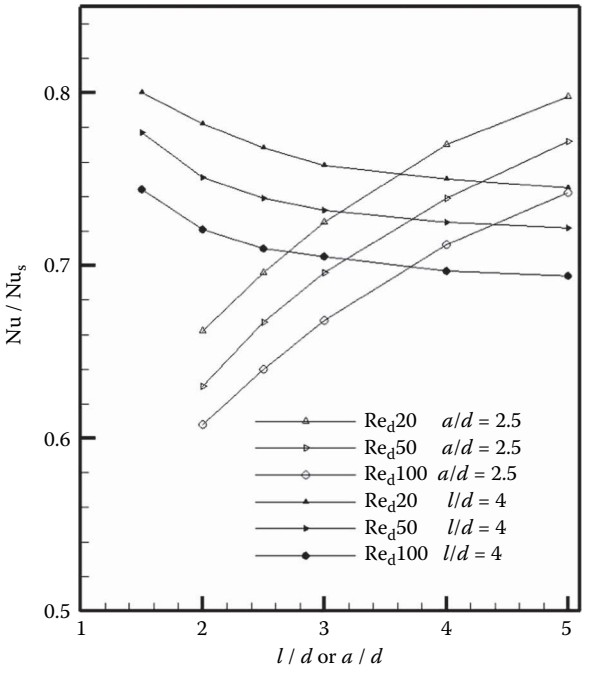


FIGURE 4.19.4 Normalized Nusselt number for particle arrays. (Based on data from Verma, A. and M. A. Jog, *Int. J. Heat and Mass Transfer,* 43, 101–112, 2000.)

velocity between adjacent spheres increases as the particles are brought close to each other. This gives rise to higher velocity gradients and increases the drag force. Verma and Jog have correlated their numerical results for the normalized drag coefficient for an interior particle for the range of Reynolds number from 20 to 100 with $2 \le (a/d, l/d) \le 5$ as

$$\frac{C_D}{C_{D,\text{sphere}}} = 1-(0.86+0.006\text{Re}^{0.5})(l/d)^{(-1.3+0.173\ln\text{Re})}+5.41\text{Re}^{-0.4}\exp(-a/d), \qquad (4.19.16)$$

where $2a$ is the lateral spacing and l is the longitudinal spacing between adjacent particles as shown in Figure 4.19.2. The above correlation fits the numerical results within ±3%.

The results for the variation of normalized overall Nusselt number of an interior sphere with lateral and longitudinal spacing are shown in Figure 4.19.4. Reduction in lateral spacing between adjacent spheres gives rise to higher flow velocity in the region between the adjacent spheres. This increases the convective heat transport and the corresponding Nusselt number. However, the normalized overall Nusselt numbers for an interior sphere is always lower than that for an isolated sphere at the same Reynolds number. For instance, at $l/d = a/d = 2.5$, the overall Nusselt number for the interior sphere is about 31%–36% less than an isolated sphere for Reynolds number range of 20–100. Verma and Jog (2000) provide a correlation for the overall Nusselt number for an interior sphere for $20 < \text{Re} < 100$, $\text{Pr} = 0.7$ and $2 \le l/d, a/d \le 5$, as

$$\frac{\text{Nu}}{\text{Nu}_{\text{sphere}}} = 1-\left[0.571-\frac{1.08}{\text{Re}}\right]\left(\frac{l}{d}\right)^{(-0.544+0.0011\text{Re})}+0.25\exp\left(-\frac{a}{d}\right), \qquad (4.19.17)$$

where $2a$ is the lateral spacing and l is the longitudinal spacing between adjacent particles as shown in Figure 4.19.2.This correlation fits their numerical results within ±3%.

Electric Field Effects

When a particle is injected in a plasma flow, more electrons impact on its surface and the particle gets negatively charged. This is because of the large difference in mobilities of ions and electrons. The negatively charged particle repels electrons and attracts ions. Eventually, the flow of the ions and electrons becomes equal at the surface and the surface is then at the floating potential (Khrapak et al., 2007; Seshadri, 1973). An electric sheath of the order of Debye length, with steep gradients in electron/ion densities and temperatures as well as large gradients in electric potential, is formed around the particle. The regimes of plasma heat transfer can be addressed in terms of the relative magnitude of the appropriate length scales. The length scales of importance are d (length scale of the particle), λ_D (the Debye length), λ_m (the smallest mean free path), δ (hydrodynamic boundary layer thickness), and δ_T (thermal boundary layer thickness). Note that the boundary layer thicknesses are introduced only to give a physical picture of the regions of importance near the surface.

In thermal plasma spraying applications, the overall plasma flow is collision-dominated (Boulos et al., 1994). However, even for a collision-dominated plasma, the mean free path need not be the smallest length scale in the problem. At higher temperatures and higher charged particle densities, the Debye length can be smaller than the mean free path. A comparison of the length scales shows the following regimes:

1. Continuum thinsheath $d \gg \delta, \delta_T > \lambda_D \gg \lambda_m$ (4.19.18)

2. Continuum moderate sheath $d \gg \lambda_D \geq \delta, \delta_T \gg \lambda_m$ (4.19.19)

3. Continuum thick sheath $\lambda_D \geq d \gg \delta, \delta_T \gg \lambda_m$ (4.19.20)

4. Noncontinuum sheath $d \gg \delta, \delta_T > \lambda_m \gg \lambda_D$ (4.19.21)

Note that the Prandtl number for the neutral gas varies with temperature. In general, the value is O(1), therefore δ and δ_T are of the same order. In the first three regimes, the electric sheath is collisional, whereas in the last regime it is collisionless. These regimes would correspond to different operating conditions. For example, the last regime is relevant for DC plasma spraying and the first regime would be appropriate for RF plasma spraying. The heat transfer characteristics are different in each regime as the dominant governing mechanisms are different in each regime. For example, in the last regime the sheath is very small and therefore the charged particle number densities are likely to be high close to the surface. As a result, the recombination rate of the charged particles and the corresponding heat transfer effect may be high. In the third regime, the transport from the neutral gas may be dominant and the electrical effects may be small.

Early studies of the interaction of a solid body with plasma were carried out in the context of electrostatic probes (Chung et al., 1975). The asymptotic analyses of Cohen (1963) and Su and Lam (1963) provide the current–voltage characteristics for a spherical probe in a weakly ionized plasma. The theory was extended by Barad and Cohen (1974) to include moderate ionization by including charge–charge and charge–neutral collisions. These analyses were limited to isothermal and stationary plasma. A review of the probe theories and experiments is available in Chung et al. (1975). Lam (1964) has discussed the probe characteristics for electrostatic probes in flowing plasmas. He suggested that the flow field can be divided into three regions, an electric sheath, a viscous boundary layer, and a quasi-neutral region. However, his analysis was restricted to stagnation point probe characteristics. These studies predict the variations of electron and ion number densities and their rate of recombination close to a solid surface but do not provide plasma-surface heat transfer due to electron ion recombination.

The electrical charge effects on heat transfer were analyzed for an isolated stationary sphere in plasma by Jog (1995) and Jog and Huang (1996). It was found that the charged particle recombination

and conduction contribute significantly to heat transport (Jog, 1995; Jog and Huang, 1996; Yeh, 2007; Yeh and Wei, 2004). The contribution of ions and electrons was a function of plasma temperature and the particle surface temperature. Increase in heat transport due to electron–ion recombination at the surface for a particle in rarefied plasma has also been reported (Gnedovets and Uglov, 1992). Meyer and Pfender (1973) have carried out detailed heat flux measurements for plasma flow over cylindrical wires. A significant part of the total heat transport was attributed to recombination of ions and electrons. These studies show the importance of surface phenomena on heat transfer.

To determine the heat transfer to particles from electrons, ions, and neutral gas, conservation equations for mass, momentum, and energy are solved along with the conservation equations for the electrons and ions densities, and an equation for the electric field. The neutral velocity field can be determined first at low degree of ionization. The electron ion diffusion, drift, and convection are then calculated. At low degree of ionization, the neutral velocity can be considered as the overall velocity of plasma. The electron and ion velocities will be different from neutral velocities due to diffusion and drift effects. At the higher degree of ionization, the charge–charge collisions may become comparable to charge–neutral collisions. The electron and ion flow is obtained by solution of the continuum conservation equations for electron and ion densities and temperatures (Mitchner and Kruger, 1973). A model for production and recombination of electrons and ions in the main body of the plasma (Jog et al., 1991, 1992) as well as recombination at the particle surface must be included in the formulation. In the electric sheath region, the complete governing equations need to be solved. However, away from the particle surface, as the ion and electron densities are expected to be large, the governing equations may be simplified (Jog, 1995). The solutions of electron, ion densities, temperatures, and the electric field in the sheath region must be matched with the solutions in the main body of the plasma. A computational method to achieve such matching has been outlined in Zhang (2005).

The heat transport at the interface is determined by accounting for all the forms of energy transport at the surface (Ecker, 1981). The conduction by electrons, ions, and neutrals, as well as energy deposited by recombination of electrons and ions at the particle surface needs to be accounted for. The drag force on the particle can be calculated by evaluating the pressure and the viscous forces. Instantaneous velocity of the particle can be determined by the balance of drag, inertia, gravity, and buoyancy forces. The governing equations are as follows:

$$\nabla \cdot (\rho \vec{u}) = 0 \tag{4.19.22}$$

$$\rho \vec{u} \cdot \nabla \vec{u} = -\nabla p + \nabla \vec{\tau} \tag{4.19.23}$$

$$\rho c_p \vec{u} \cdot \nabla T_n = \nabla \cdot (\chi_n T_n) \tag{4.19.24}$$

$$\nabla \cdot [\vec{u} N_e + \Gamma_e] = P - R \tag{4.19.25}$$

$$\nabla \cdot [\vec{u} N_i + \Gamma_i] = P - R \tag{4.19.26}$$

$$\nabla \cdot \vec{E} = -\frac{e}{\varepsilon_0}(N_i - N_e) \tag{4.19.27}$$

$$\Gamma_e = -\frac{\mu_e}{e}\nabla(N_e k T_e) + \mu_e N_e \nabla V \tag{4.19.28}$$

$$\Gamma_i = -\frac{\mu_i}{e}\nabla(N_i k T_i) - \mu_i N_i \nabla V \tag{4.19.29}$$

$$\vec{q}_e = \frac{5}{2} k T_e \Gamma_e - \chi_e \nabla T_e \tag{4.19.30}$$

$$P - R = \gamma N_e \left[\frac{2 g_i N_n}{g_n} \left(\frac{2\pi m_e k T_e}{h_P^2} \right)^{3/2} \exp\left(-\frac{e V_i}{k T_e} \right) - N_e N_i \right], \tag{4.19.31}$$

where $\gamma = 1.09 \times 10^{-20} T_e^{-9/2} \text{m}^6/\text{s}$. In the above, u is velocity, p is pressure, χ is thermal conductivity, subscripts e, i, and n are for electrons, ions, and neutral gas, respectively. E is electric field, V is electric potential, N is used for number densities, q is heat flux, Γ denotes charge species flux, μ_e and μ_i are electron and ion mobilities, k_B is Boltzmann constant, e is the amount of charge on a single electron, and ε_0 is the permittivity of the free space, P and R are rates of production and recombination of electrons and ions, g_i is the statistical weight of ions, g_n is the statistical weight of the ground state of neutrals, m_e is the mass of an electron, h_P is Planck's constant, and V_i is ionization voltage.

In plasma sprays, the particle sizes are typically less than 100 μm and the corresponding Reynolds number less than 100 (Vardelle et al., 2015). Hader and Jog (1998a) have considered heat transfer to a spherical particle from a plasma flow with Re < 100 and have analyzed the heat transfer from electron–ion recombination on the particle surface relative to the convective heat transfer. Figure 4.19.5 shows the variation of Nusselt number for electron–ion recombination, neutral gas, and the total heat transfer as a function of Reynolds number based on particle diameter. It is evident that all contributions increase with increasing Reynolds number. This is to be expected as a stronger flow at higher Reynolds number results in thinner hydrodynamic, thermal, as well as electric boundary layer over majority of the sphere surface. Furthermore, the charged species contribution is temperature dependent and increases significantly with increasing far-field temperature.

Figure 4.19.5 compares their predictions of Nu for two values of far-field temperature and a value of surface temperature close to reported data (Fincke et al., 1990). While correlations are not available to account for these effects, the results shown in Figure 4.19.5 provide a useful measure of the extent of contributions from the charged species recombination.

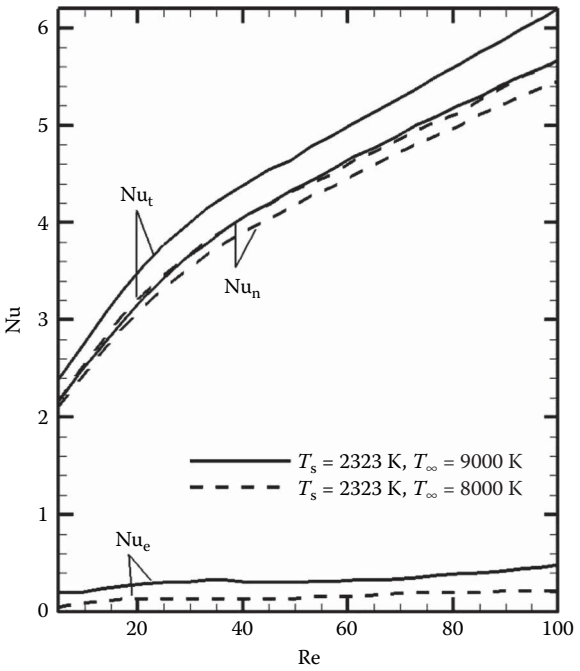


FIGURE 4.19.5 Variation of Nusselt numbers for electron–ion recombination and neutral gas. (Based on data from Hader, M. A., and M. A. Jog, *Phys. Plasmas,* 5(4), 902–909, 1998a.)

Summary

Plasma spraying is employed in a variety of applications from depositing nano-scale to macro-scale coating to net-shape manufacturing. Particles are injected in plasma jet in plasma spraying process and heat transfer from plasma to particles determines the efficacy of the process. The analysis of heat transfer is complicated by unique features of plasma particle interactions that include large electrical–thermophysical properties variations over a short distance, electrical boundary layer and electron–ion recombination on the particle surface, nonspherical particle shape, noncontinuum conditions, and multiparticle interactions. These features are discussed here and data and design correlations are provided to determine drag and heat transfer interactions between particles and plasma flow in plasma spraying process.

REFERENCES

Ayyaswamy, P. S. 1995. Direct contact heat transfer with drops, in *Advances in Heat Transfer*, J. P. Hartnett and T. F. Irvine (eds.), Cambridge: Academic Press, Vol. 26, pp. 1–104.

Ayyaswamy, P. S., and I. M. Cohen. 2002. *Annual Reviews of Heat Transfer*, New York: Hemisphere Publishing, Vol. 12, pp. 27–78.

Barad, M. S., and I. M. Cohen. 1974. Continuum theory of spherical electrostatic probes in stationary, moderately ionized plasma. *Physics of Fluids* 17: 724–734.

Boulos, M. I., P. Fauchais, and E. Pfender. 1994. *Thermal Plasmas: Fundamentals and Applications*. New York: Plenum Press.

Brown, S. C. 1967. *Basic Data of Plasma Physics*. Cambridge: MIT Press.

Chen, H. C., E. Pfender, and J. Heberlein. 1997. Improvement of plasma spraying efficiency and coating quality. *Plasma Chemistry and Plasma Processing* 17: 93–105.

Chhabra, R. P. 2007. *Bubbles, Drops, and Particles in Non-Newtonian Fluids*. Boca Raton, FL: CRC Press.

Chung, P. M., L. Talbot, and K. J. Touryan. 1975. *Electrostatic Probes in Stationary and Flowing Plasmas: Theory and Application*, New York: Springer-Verlag.

Chyou, Y. P., and E. Pfender. 1989. Behavior of particulates in thermal plasma flows. *Plasma Chemistry and Plasma Processing* 9: 45–71.

Clift, R., J. R. Grace, and M. E. Weber. 1978. *Bubbles, Drops, and Particles*, Cambridge: Academic Press.

Cohen, I. M. 1963. Asymptotic theory of spherical electrostatic probes in a slightly ionized, collision dominated plasma. *Physics of Fluids* 6: 1492–1499.

Crowe, C. T., M. P. Sharma, and D. E. Stock. 1977. The particle-source-in cell (PSI-cell) model for gas-droplet flows. *Journal of Fluids Engineering* 99: 325–332.

Devoto, R. S. 1963. Transport coefficients of ionized argon. *Phys. Fluids* 6: 616–623.

Ecker, G. 1981. Arc discharge electrode phenomena. *Electrical Breakdown and Discharges in Gases*, E. E. Kunhardt and L. H. Luessen (eds.), New York: Plenum Press, pp. 167–179.

Fauchais, P., M. Vardelle, S. Goutier, and A. Vardelle. 2015. Key challenges and opportunities in suspension and solution plasma spraying. *Plasma Chemistry and Plasma Processing* 35: 511–525.

Fincke, J. R., W. D. Swank, and C. L. Jeffery. 1990. Simultaneous measurement of particle size, velocity, and temperature in thermal plasma. *IEEE Transactions on Plasma Science* 18: 948–957.

Fukai, J., Z. Zhao, D. Poulikakos, C. M. Megaridis, and O. Miyatake. 1993. Modeling of the deformation of a liquid droplet impinging on a flat surface. *Physics of Fluids* 5: 2588–2599.

Gnedovets, A. G., and A. A. Uglov. 1992. Heat transfer to a non-spherical particle in rarefied plasma flow. *Plasma Chemistry and Plasma Processing* 12: 371–382.

Graves, D. B. 1989. Plasma processing in microelectronic manufacturing. *AIChE Journal* 35: 1–28.

Hader, M. A., and M. A. Jog. 1998a. Continuum plasma flow past a sphere. *Physics of Plasmas* 5(4): 902–909.

Hader, M. A., and M. A. Jog. 1998b. Effect of drop deformation on the heat transfer to a drop suspended in an electric field. *Journal of Heat Transfer* 120(3): 682–689.

Hossain, M. M., Y. Yao, and T. Watanabe. 2007. A numerical study of plasma-particle energy exchange dynamics in induction thermal plasmas for glassification. In Numerical, Mathematical, and Physical Modeling Tools for Materials Processing, *Proceedings of Material Science and Technology*, Detroit, MI, September 16–20.

Jog, M. A., I. M. Cohen, and P. S. Ayyaswamy. 1991. Breakdown of a wire-to-plane discharge. *Physics of Fluids B* 3: 3532–3536.

Jog, M. A., I. M. Cohen, and P. S. Ayyaswamy. 1992. Electrode heating in a wire-to-plane arc. *Physics of Fluids B* 4: 465–472.

Jog, M. A. 1995. Analysis of heat transfer to a spherical particle from continuum plasma. *Journal of Applied Physics* 78: 1424–1429.

Jog, M. A., and L. Huang. 1996. Transient heating and melting of particles in plasma coating process. *Journal of Heat Transfer* 118: 471–477.

Joshi, S. V., J. Y. Park, P. R. Taylor, and L. S. Richardson. 1986. Particle-gas mass transfer under plasma conditions. *International Journal of Heat and Mass Transfer* 29(10): 1565–1573.

Khrapak, S. A., B. A. Klumov, and G. E. Morfill. 2007. Ion collection by a sphere in a flowing highly collisional plasma. *Physics of Plasma* 14: 034502.

Kuryshev, A. P., and V. V. Sakhin. 2001. The thermal state of a small spherical body in an air plasma. *Technical Physics* 46(11): 1453–1456.

Lam, S. H. 1964. A general theory for the flow of weakly ionized gases. *AIAA Journal* 2: 256–262.

Liu, H., E. J. Lavernia, and R. H. Rangel. 1994a. Numerical investigation of micro-pore formation during substrate impact of molten droplets in plasma spray processes. *Atomization and Sprays* 4: 369–384.

Liu, H., E. J. Lavernia, and R. H. Rangel. 1994b. Numerical investigation of substrate impact and freezing of drops in plasma spray processes. *Journal of Physics D: Applied Physics* 26: 1900–1908.

Meyer, T. N., and E. Pfender. 1973. Experimental and analytical aspects of plasma heat transfer. *Warme- und Stoffübertragung* 6: 25–31.

Mitchner, M., and C. H. Kruger. 1973. *Partially Ionized Gases*, New York: Wiley.

Mostaghimi, M., and M. I. Boulos. 2015. Thermal plasma sources: How well are they adopted to process needs? *Plasma Chemistry and Plasma Processing* 35: 421–436.

Pfender, E. 1989. Particle behavior in thermal plasma. *Plasma Chemistry and Plasma Processing* 9: 167S–193S.

Proulx, P., J. Mostaghimi, and M. I. Boulos. 1985. Plasma-particle interaction effects in induction plasma modeling under dense loading conditions. *International Journal of Heat Mass Transfer* 28: 1327–1355.

Proulx, P., J. Mostaghimi, and M. I. Boulos. 1991. Radiative energy transfer in induction plasma modelling. *International Journal of Heat Mass Transfer* 34: 2571–2578.

Rojas, J. R., M. A. Cruchaga, D. J. Celentano, M. El Ganaoui, and B. Pateyron. 2009. Numerical simulation of the melting of particles injected in a plasma jet. *Ingeniare: Revista chilena de ingeniería* 17(3): 299–308.

Sadhal, S. S., P. S. Ayyaswamy, and J. N. Chung. 2012. *Transport Phenomena with Bubbles and Droplets*, New York: Springer-Verlag.

Sato, T., O. P. Solonenko, and H. Nishiyama. 2004. Evaluations of ceramic spraying processes by numerical simulation. *Materials Transactions* 45(6): 1874–1879.

Schein, J., K. Hartz-Behrend, S. Kirner, M. Kühn-Kauffeldt, B. Bachmann, and E. Siewert. 2015. New methods to look at an old technology: Innovations to diagnose thermal plasmas. *Plasma Chemistry and Plasma Processing* 35: 437–453.

Seshadri, S. R. 1973. *Fundamentals of Plasma Physics*, New York: American Elsevier Publishing Company.

Smith, R. W., D. Wei, and D. Apelian. 1989. Thermal plasma materials processing—Applications and opportunities. *Plasma Chemistry and Plasma Processing* 9: 135S–165S

Sripada, S. S., P. S. Ayyaswamy, L. J. Huang. 1996. Condensation on a spray of water drops: A cell model study: I. Flow description. *International Journal of Heat Mass Transfer* 39: 3781–3790.

Streibl, T., A. Vaidya, M. Friis, V. Srinivasan, and S. Sampath. 2006. A critical assessment of particle temperature distributions during plasmas: Experimental results for YSZ. *Plasma Chemistry and Plasma Processing* 26(1): 73–102.

Su, C. H., and S. H. Lam. 1963. Continuum theory of spherical electrostatic probes. *Physics of Fluids* 6: 1479–1491.

Tal, R., D. N. Lee, W. A. Sirignano. 1983. Hydrodynamics and heat transfer in sphere assemblages: Cylindrical cell models. *International Journal of Heat Mass Transfer* 26: 1265–1273.

Trapaga, G., E. F. Matthys, J. J. Valencia, and J. Szekely. 1992. Fluid flow, heat transfer and solidification of molten metal droplets impinging on substrates – Comparison of numerical and experimental results, *Metallurgical Transactions B: Process Metallurgy* 23: 701–718.

Trapaga, G., and J. Szekely. 1991. Mathematical modeling of the isothermal impingement of liquid droplets in spraying processes. *Metallurgical Transactions B: Process Metallurgy* 22: 901–914.

Vardelle, A., M. Vardelle, and P. Fauchais. 1982. Influence of velocity and surface temperature of alumina particles on the properties of plasma sprayed coatings. *Plasma Chemistry and Plasma Processing* 2: 255–291.

Vardelle, A., C. Moreau, N. J. Themelis, and C. Chazelas. 2015. A perspective on plasma spray technology. *Plasma Chemistry and Plasma Processing* 35:491–509.

Verma, A. and M. A. Jog. 2000. Plasma flow over an array of particles. *International Journal of Heat and Mass Transfer* 43: 101–112.

Wen, Y. 2004. Effect of temperature dependent properties and particle shape on heat transfer in plasma flow. MS Thesis, University of Cincinnati.

Wen, Y., and M. A. Jog. 2005. Variable property, laminar, continuum plasma flow over non-spherical particles. *International Journal of Heat and Fluid Flow* 26(5): 780–791.

Xiong, H-B, L.-L. Zheng, and T. Streibl. 2006. A critical assessment of particle temperature distributions during plasma spraying: Numerical studies for YSZ. *Plasma Chemistry and Plasma Processing* 26(1): 53–72.

Yeh, F. B. 2007. A semi-analytical method to study the temperature evolutions of a slab and a semi-infinite target for plasma immersion ion implantation. *International Journal of Heat and Mass Transfer* 50: 789–798.

Yeh, F. B., and P. S. Wei. 2004. Plasma energy transport to an electrically biased surface. *International Journal of Heat and Mass Transfer* 47: 4019–4029.

Young, R. M. and E. Pfender. 1987. Nusselt number correlations for heat transfer to small spheres in thermal plasma flows. *Plasma Chemistry and Plasma Processing* 7: 211–227

Zhang, Q. 2005. Kinetic theory analysis of heat transfer to a sphere from a stationary ionized gas. PhD Thesis, University of Cincinnati.

4.20 THERMAL PROCESSING AND PRESERVATION OF FOODS

Prabhat Kumar and K.P. Sandeep

Introduction

Thermal processing is one of the most common methods used to preserve foods for consumption at a later time (days, weeks, months, or years). The degree of thermal processing required depends on various product factors and the storage conditions (ambient, refrigerated, or frozen). Some of the product factors include pH, water activity, physical property (density), rheological properties (such as viscosity for a Newtonian fluid, and consistency coefficient and flow behavior index for power-law fluids), thermal properties (such as thermal conductivity and specific heat), dielectric properties (such as dielectric constant and loss factor) for electromagnetic (microwave and radio frequency) heating, and electrical property (electrical conductivity) for ohmic heating. Some of the above product characteristics dictate the types of microorganisms (both pathogenic and spoilage) that could potentially grow in the product depending on the amount of air in the product (aerobes, anaerobes, microaerophiles), the storage conditions (psychrotrophs/psychrophiles, mesophiles, and thermophiles), and special factors such as low water activity (xerophiles), high salt content (halophiles), and high sugar content (osmophiles). The steps involved in thermally processing the product are to determine the target organism (the most thermally resistant organism that could potentially be of concern in the product), identify a surrogate organism (a nonpathogenic organism that has similar kinetics of inactivation in the product under consideration), establish an optimized thermal processing protocol (time–temperature combination or *F*-value required to ensure product safety and quality), and identify the thermal processing equipment (retort with or without agitation, tubular/shell-and-tube/scraped-surface/electromagnetic/ohmic) and associated system and process parameters that can be used to deliver the required thermal treatment. Depending on the product characteristics and desired storage conditions, the product may be subjected to blanching, pasteurization, hot filling, retorting, or aseptic processing.

Product Considerations

The choice of a thermal processing system is based on several product considerations including the characteristics of the product (pH, water activity, etc.) and properties of the product (physical properties, rheological properties, thermal properties, dielectric properties for electromagnetic heating, and electrical properties for ohmic heating).

pH

The extent of thermal treatment given to a food product depends on whether the food product is an acid product, an acidified product, or a low-acid product. An acid food product is defined as the one with a natural pH of less than 4.6. pH is a measure of acidity and is defined as the negative logarithm of H^+ ion concentration. pH normally ranges from 0 to 14 with 7 being regarded as neutral pH (pH of pure water). Acid food products include apple juice, orange juice, ketchup, etc. An acidified food product is the one with a finished equilibrium pH of less than 4.6 and a water activity (a_w) greater than 0.85. Example of acidified foods include peppers treated in an acid brine, pickled foods, etc. Acidified food products are typically treated at 90°C–95°C for a period of 30–90 s to inactivate microorganisms. A low-acid food product is any food other than alcoholic beverages with a natural equilibrium pH greater than 4.6 and a water activity greater than 0.85. These food products include butter, cheese, fresh eggs, pears, papaya, and raisins (Skudder, 1993). Low-acid food products are capable of sustaining the growth of *Clostridium botulinum* spores which is capable of producing a potent neurotoxin. Foodborne botulism is a severe type of food poisoning caused by the ingestion of

foods containing the potent neurotoxin formed during the growth of *C. botulinum*. The spores of *C. botulinum* must be destroyed or effectively inhibited to avoid germination and subsequent production of the deadly toxin which causes botulism.

Water Activity

Water present in a food product may be broadly classified as free water (if it is available to participate in various biochemical reactions) and bound water (if it is not available to participate in various biochemical reactions). Water activity is a measure of the amount of free water in a product. It is determined as the ratio of the vapor pressure of water in that product to the vapor pressure of pure water at that same temperature. It is also sometimes defined as equilibrium relative humidity divided by 100. Water activity is determined using a hygrometer. Most microorganisms require water for their growth and hence grow faster in high water activity foods. Accordingly, lowering water activity is one of the common techniques used to preserve many types of foods. Since most bacteria do not grow rapidly at water activity levels below 0.85, that is used as a cutoff value for foods likely to be shelf-stable with minimal or no thermal processing. Between a water activity of 0.80 and 0.85, molds are generally the organisms of concern. Between water activity values of 0.5 and 0.8, osmophilic yeasts, xerophilic molds, and halophilic bacteria are the organisms of concern.

The magnitude of water activity may be adjusted to control various reactions such as microbial growth, enzyme activity, Maillard browning, lipid oxidation, and vitamin degradation. It may also be used to control caking/clumping, rehydration, and texture. Its magnitude also enables quantification of moisture migration within a product. At low water activity, lipid oxidation is the main reaction that takes place. As water activity increases, Maillard browning begins to take place. Beyond a certain level of water activity, caking/clumping, crystallization, and enzymatic activity start to take place. As water activity approaches 0.80, mold growth starts to take place. After that, yeasts and bacteria begin to grow.

As far as food products are concerned, fresh fruits and vegetables, milk, juices, and raw meats have a water activity value generally higher than 0.95. At the intermediate range, products such as cooked bacon, chocolate syrup, and sweetened condensed milk have water activity values between 0.7 and 0.9. Products such as dried fruits and honey fall under the lower end of the spectrum of water activity, with values between 0.5 and 0.7.

Physical Properties

The physical property of interest in thermal processing is density. Density (ρ) is defined as mass per unit volume with the units of kg/m^3. Materials with more compact molecular arrangement have higher densities. Specific gravity is defined as the ratio of the density of a material to the density of water at the same temperature. There are three types of densities for food materials—solid density, particle density, and bulk density. Solid density does not take into account the internal pores in the food materials. Particle density is defined as the ratio of the actual mass of a particle to its actual volume. Particle density accounts for the internal pores in the food materials. Bulk density is defined as the mass of the food material per unit volume. Bulk density accounts for the void spaces between the particles. The void space in food materials is described by porosity which is the ratio of the void volume to the total volume of the material.

Rheological Properties

Rheology is defined as the science of flow and deformation of materials. The main rheological property is viscosity. Viscosity (μ) is defined as the fluid property which denotes the resistance between the internal layers of a fluid to movement. It is also called absolute or dynamic viscosity. Its unit is Pa s (N s/m^2) or poise (dyne s/cm^2) and is given by:

$$\mu = \frac{\sigma}{\dot{\gamma}},$$

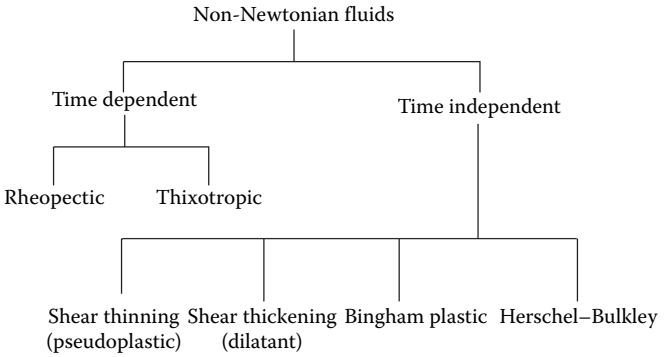


FIGURE 4.20.1 Classification of non-Newtonian fluids.

where σ is the shear stress and $\dot{\gamma}$ is the shear rate. Force per unit area is defined as stress and when the force is applied parallel to the surface, the stress is called shear stress. Shear rate is the relative change in velocity divided by the distance. Liquids that follow the above equation exhibit a direct proportionality between shear rate and shear stress and are called Newtonian fluids. Examples of Newtonian fluids include water, fluid milk, honey, and fruit juices. Fluids that do not have a direct proportionality between shear rate and shear stress are known as non-Newtonian fluids. Non-Newtonian fluids can be classified as time-independent and time-dependent (Figure 4.20.1). For time-dependent fluids, the shear stress vs. shear rate graph is not the same as a function of time. For example, for a constant shear rate, the shear stress may increase or decrease as a function of time.

The relationship between shear stress and shear rate for Newtonian and non-Newtonian fluids is shown in Figure 4.20.2.

Two important types of time-independent non-Newtonian fluids are shear thinning (pseudoplastic) and shear thickening (dilatant). The difference between these two can be understood on the basis of apparent viscosity. An apparent viscosity is calculated by assuming that the non-Newtonian fluid obeys Newton's law of viscosity at a selected shear rate. The slope of the straight line from the origin to a point on the curve at a given shear rate gives the value of apparent viscosity. For a shear thinning fluid, the apparent viscosity decreases as the shear rate increases. Some examples of shear thinning fluids are condensed milk, fruit purees, mayonnaise, mustard, and vegetable soups. For a shear thickening fluid, the apparent viscosity increases as the shear rate increases. Some examples of shear thickening fluids are 40% raw corn starch and some types of honey.

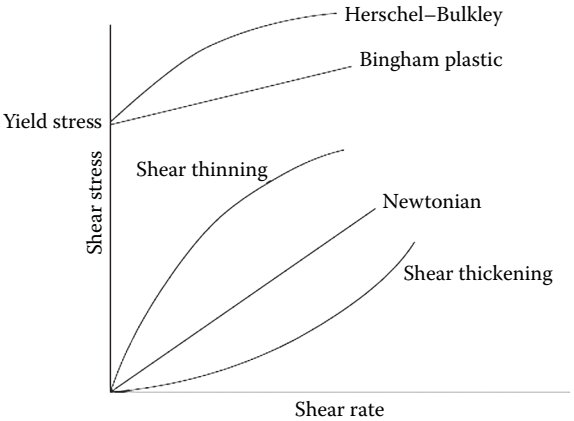


FIGURE 4.20.2 Relationship between shear stress and shear rate for Newtonian and non-Newtonian fluids.

Another class of non-Newtonian fluids requires an application of a stress prior to any flow. This stress is known as the yield stress (σ_0). After the application of yield stress, those fluids which behave like a Newtonian fluid are known as Bingham plastic. If the fluids behave as shear thinning fluids, then they are known as Herschel–Bulkley fluids.

Time-dependent non-Newtonian fluids attain a constant value of apparent viscosity only after some time has elapsed after the application of shear stress. The fluid is known as thixotropic if its viscosity decreases over time and rheopectic if the viscosity increases over time.

Thermal Properties

Thermal properties include specific heat, thermal conductivity, and thermal diffusivity. They have a significant effect on the rate of heat transfer into the food product. Study of thermal properties of the food product plays an important role in the design of thermal processing systems.

Specific heat (c_p) is the amount of heat that is gained or lost by a unit mass of product for a unit change in temperature without changing state and is given by the following equation:

$$c_p = \frac{Q}{m(\Delta T)},$$

where Q is the heat gained or lost (kJ), m is the mass (kg), and ΔT is the temperature change (°C). For processes where a change of state takes place (freezing or thawing), an apparent specific heat is used. Specific heat plays an important role in the thermal analysis of food processing equipment used in heating or cooling of foods. Comprehensive databases (Singh, 1994), predictive equations, or experimental methods are generally used to obtain the values of specific heat.

Thermal conductivity (k) is the quantity of heat (Q) conducted per unit area (A) through a unit length (L) of material due to a unit temperature difference (ΔT) and is given by the equation:

$$k = \frac{QL}{A\ (\Delta T)}.$$

The units of thermal conductivity are W/m K. Thermal conductivity of a food plays an important role in calculations involving the rate of heat transfer by conduction. Thermal conductivities of water and air at 20°C are 0.597 and 0.0251 W/m-K, respectively (Singh and Heldman, 2001). The thermal conductivity of food materials is bounded by water and air on the upper and lower end, respectively. Comprehensive databases (Singh, 1994; Krokida et al., 2001), predictive equations, or experimental methods are generally used to obtain the values of thermal conductivity.

Thermal diffusivity (α) is the rate at which temperature at one point in a material is transmitted to another point. It is the ratio of thermal conductivity to volumetric heat capacity and is given as:

$$\alpha = \frac{k}{\rho c_p}.$$

The units of thermal diffusivity are m^2/s. Thermal diffusivity can be calculated by determining the values of thermal conductivity, density, and specific heat. However, this approach requires considerable time and elaborate instrumentation. Another approach is the direct measurement of thermal diffusivity by experimental methods.

Dielectric Properties

Interaction of electromagnetic (microwave and radio frequency) energy with materials depends on their dielectric properties. Metals are not heated significantly by electromagnetic waves as they are excellent reflectors. Dielectric materials are better absorbers and transmitters of electromagnetic waves. Dielectric materials are heated mainly through the absorption of electromagnetic energy and the absorption depends on composition and temperature of the materials along with the frequency

of the electromagnetic wave. Dielectric properties determine the heating behavior of a material when subjected to an electromagnetic field. Therefore, knowledge of dielectric properties is important for the design of a heating system based on electromagnetic energy.

Dielectric properties consist of dielectric constant (ε') and dielectric loss factor (ε''). Dielectric constant is a measure of the ability of a material to store electromagnetic energy, whereas dielectric loss factor is a measure of the ability of a material to convert electromagnetic energy to heat (Metaxas and Meredith, 1983). Dielectric constant and dielectric loss factor can be defined in terms of complex permittivity (ε). The complex permittivity (ε) is composed of a real part (ε', relative dielectric constant) and an imaginary part (ε'', effective relative dielectric loss factor) and is given by the following equation (Saltiel and Datta, 1999):

$$\varepsilon = \varepsilon_0(\varepsilon' - j\varepsilon''),$$

where $j=(-1)$ and, ε_0 is permittivity of free space (8.86×10^{-12}F/m).

Electrical Properties

Understanding of electrical properties (electrical conductivity) is an important parameter for the design of Ohmic heating system. Electrical conductivity is the ability of a material to transport an electric charge. Electrical conductivity (σ) is the reciprocal of electrical resistance (R) through a unit cross-sectional area (A) over a distance (L), given by the following equation:

$$\sigma = \frac{L}{AR}.$$

The electrical resistance of a sample is determined as the ratio of voltage (V) across the sample divided by current (I).

Microorganisms

The microorganisms important in thermal processing are bacteria and fungi because they can multiply in foods and cause spoilage or public health issues. There are various types of bacteria and fungi and each has a specific range of conditions (temperature, oxygen, pH, water activity, type of nutrients, etc.) under which they can grow rapidly.

Bacteria are a large group of unicellular prokaryotic microorganisms. They are found in a wide range of shapes such as spheres (cocci), rods, and spiral. Bacteria reproduce asexually through the process of binary fission. Under favorable growth conditions, bacteria can grow and divide rapidly. The growth cycle of bacteria can be divided into four phases: lag, log, stationary, and death. During the lag phase, bacteria adapt to the new surrounding and multiple slowly. During the log phase, bacteria multiply at a very rapid exponential rate. During the stationary phase, growth rate slows down again and eventually they stop multiplying resulting in their death in the death phase. Some species of rod-shaped bacteria can form highly resistant structures known as spores. Spores can survive extreme stress conditions such as high heat and pressure. Bacteria in this dormant state may remain viable for thousands of years (Tucker and Featherstone, 2011).

Fungi are a group of eukaryotes that include yeasts and molds. Fungi are neither plants nor animals as they possess properties similar to plants (cell wall) as well as animals (absence of chlorophyll). Yeasts are unicellular fungi which derive their energy from organic compounds and do not require sunlight to grow. Yeasts are larger than bacteria, but smaller than molds. Yeasts reproduce asexually by budding in which a small bud forms on the parent cell and gradually enlarges into another yeast cell. Yeasts are either obligate aerobes or facultative anaerobes. They grow best in a neutral or a slightly acidic medium and are generally destroyed above a temperature of 50°C. Yeasts are used in the food industry for leavening of bread and production of alcohol. However, their ability

to grow at low pH and water activity (a_w) make them organisms of concern for spoilage in fruit products such as juices and jams (Tucker and Featherstone, 2011).

Molds are multicellular fungi that grow in the form of hyphae (multicellular tubular filaments). Molds reproduce both sexually and asexually by means of spores produced on specialized structures or in fruiting bodies. All molds are aerobic and some can grow in low oxygen conditions. They can also grow at extreme conditions (high acid, high salt, low temperature). Molds are used in the food industry for production of cheese and soy sauce. Molds are also capable of consuming acids, which can remove the acidic condition that inhibit the growth of *C. botulinum*. Spoilage of food products by mold is primarily due to mycotoxins (aflatoxin, ocratoxin, patulin, etc.), which are secondary metabolites produced by molds (Tucker and Featherstone, 2011).

Classification Based on Gram Staining

Bacteria can be classified as gram-positive or gram-negative based on gram staining. Gram staining is a method based on the chemical and physical properties of bacterial cell wall. The four basic steps of gram staining include applying a primary stain (crystal violet), addition of iodine, decolorization with alcohol (ethanol), and counterstaining with safranin. After gram staining, gram-positive organisms appear purple, whereas gram-negative organisms appear pink. Gram-positive is more resistant to changes in environment because of its thick peptidoglycan layer (Tucker and Featherstone, 2011).

Classification Based on Oxygen

Based on oxygen requirement, microorganisms can be classified into aerobes, anaerobes, facultative anaerobes, and microaerophiles. Aerobes grow in the presence of atmospheric oxygen, whereas anaerobes grow in the absence of atmospheric oxygen. Facultative anaerobes are in between these two extremes and can grow in either the presence or absence of atmospheric oxygen. Microaerophiles require a small amount of oxygen to grow.

Classification Based on Temperature

Based on the response to temperature, microorganisms can be classified into psychrophilic, psychrotrophic, mesophilic, and thermophilic. Psychrophilic microorganisms have an optimum growth temperature between 12°C and 15°C, but can grow up to 20°C. Psychrotrophic microorganisms have an optimum growth temperature between 20°C and 30°C, but can grow up to 0°C. Mesophilic microorganisms have an optimum growth temperature between 30°C and 42°C, but can grow between 15°C and 47°C. Thermophilic microorganisms have an optimum growth temperature between 55°C and 65°C, but can grow between 40°C and 90°C.

Classification Based on pH, Water Activity, Salt, and Osmotic Pressure

Most microorganisms grow best at neutral pH and only a few are able to grow at a pH value of less than 4.0. Bacteria are more selective about pH requirements than yeasts and molds which can grow over a wide range of pH. Microorganisms which can withstand low pH are known as aciduric. Bacteria require higher a_w for growth as compared to that required by yeasts and molds. Gram-negative bacteria cannot grow at a_w less than 0.95 whereas most gram-positive bacteria cannot grow at a_w less than 0.90. However, *Staphylococcus aureus* can grow at a_w value as low as 0.85 and halophilic bacteria can grow at a minimum a_w value of 0.75. Halophilic microorganisms require a high salt concentration for growth. Most yeasts and molds can grow at a minimum a_w value of 0.88 and 0.80, respectively. Xerophilic (microorganisms which can grow in low a_w conditions) molds and osmophilic (microorganisms which can grow in high solute concentration) yeasts can grow at a_w values as low as 0.61.

Target Organism

The goal of "processing" a formulated product is to use one or more mechanisms to either inactivate microorganisms (such as by the use of heat) or prevent their growth (by use of appropriate chemicals) to render the product safe for consumption. This includes all microorganisms that may be of public

health significance or may spoil the product. The target organism is the most resistant organism to the method of control of the organisms. If heat is used as the means of inactivating microorganisms, the most thermally resistant organism under the prevailing conditions would be the target organism. For example, for producing a shelf-stable product which has a pH greater than or equal to 4.6 and a water activity greater than 0.85 by thermal processing, the target organism is spores of type A *C. botulinum*.

Surrogate Organism

Thermal process validation involves assurance that the established thermal processing protocol (time–temperature combination based on theoretical considerations) would result in a safe product. This could potentially be achieved by incorporating the target organism at the appropriately high level into the product, thermally processing the product, and conducting microbial analysis to ensure that an adequate level of destruction of the target organism was achieved. However, this is a very risky proposition as it would potentially expose operators to the pathogenic target organism and also contaminate the system and facilities. Thus, a safer proposition would be to identify a nonpathogenic organism that has a similar inactivation kinetics as the target organism and use it as a substitute for the target pathogen. This nonpathogenic organism is called the surrogate organism and is used in process validation in place of the target organism.

Kinetics of Reactions

Kinetics of reactions refers to the rates of various reactions that take place within a food product under various processing and storage conditions. These may be reactions associated with microbial destruction, enzymatic inactivation, and changes in quality/nutritional attributes such as texture/color/flavor/vitamin change. Most of these reactions can be described by the following equation:

$$-\frac{\mathrm{d}c}{\mathrm{d}t} = k_n c^n,$$

where "c" is the concentration of the reacting species (it may be the number of microorganisms, enzyme activity level, vitamin content, etc.), "k_n" is the rate of the reaction, and "n" is the order of the reaction. In order to determine the rate and order of a reaction, a graph of ln ($-\mathrm{d}c/\mathrm{d}t$) is plotted on the y-axis vs. ln (c) on the x-axis. The slope of this graph is the order of the reaction and the intercept is "ln (k_n)". Orders of reactions may take whole number values ($n=0$ for zero-order reaction, $n=1$ for first-order reaction, $n=2$ for second-order reaction) or fractional values ($n=1.3$ for color change in a certain vegetable or $n=0.7$ for a certain flavor change).

D Value

When a homogeneous microbial population is subjected to a constant temperature, T, the rate of microbial destruction follows a first-order reaction kinetics and is given by (David et al., 1996):

$$\frac{-\mathrm{d}N}{\mathrm{d}t} = K_T N,$$

where N is the number of microbes surviving after time t (s) and K_T is the reaction rate (s^{-1}). Integration of the above equation from time 0 to time t yields:

$$\frac{N}{N_0} = e^{-K_T t},$$

where N_0 is the number of viable microorganisms at time $t=0$.

The above equation can be rewritten as:

$$\log_{10}\left(\frac{N}{N_0}\right) = \frac{-t}{D},$$

where $D=2.303/K_T$. D (decimal reduction time) is the time required to reduce the number of surviving microbial population by 90%. D value is a measure of heat resistance of microorganisms. Microorganisms with a higher D value have a higher heat resistance. D value determined at a reference temperature (T_{ref}) is denoted by D_{ref}.

z Value

The effect of temperature on D value is generally described by the following expression (David et al., 1996):

$$\log_{10}\left(\frac{D_T}{D_{ref}}\right)=\frac{T_{ref}-T}{z},$$

where D_T is the D value at temperature T and z is the change in temperature (°C) required to change the D value by 90%. D and z values are the basis of thermal process calculations and are commonly used to design a thermal process.

F Value

The ratio of D_{ref} to D is the lethal rate (L_r). Thermal death time or F value of a process is defined as the process time at a given temperature required for total destruction of a microbial population or the time required for destruction of microorganisms to an acceptable level. F value can be expressed as a multiple of D value for first order microbial kinetics. The F value required for a process depends on the nature of food (pH and water activity), storage conditions after processing (refrigerated vs. room temperature), target organism, and initial population of microorganisms. F value is usually expressed with a superscript denoting z value and a subscript denoting temperature. F value can be computed in terms of lethal rate as:

$$F_{T_{ref}}^{z}=\int_0^t L_r\,dt=\int_0^t\left(10^{\frac{T-T_{ref}}{z}}\right)dt=-D_{ref}\log\frac{N}{N_0}.$$

For a constant temperature process, the above equation for F value reduces to:

$$F_{T_{ref}}^{z}=10^{\frac{T-T_{ref}}{z}}t.$$

The F value at a reference temperature of 121.1°C (250°F) and a z value of 10°C (18°F) is referred to as the F_0 value. The main microorganism of concern for low acid foods is *C. botulinum*, which has a $D_{121.1}$ value of 0.21 min. For processes where *C. botulinum* is the target organism, F_0 value represents the process time necessary to achieve a 12 log reduction (12D) in microbial population of *C. botulinum* at 121.1°C. A thermal process designed to reduce the probability of survival of spores of *C. botulinum* to an acceptable level (12D process) is referred to as the botulinum cook. The F_0 value for a botulinum cook is 2.52 (12×0.21) min. An F_0 value of 2.52 min indicates that the process is equivalent to a heat treatment of 2.52 min at 121.1°C (250°F). Thus, many combinations of time and temperature can yield an F_0 value of 2.52 min. F value can be written in terms of F_0 value as:

$$F_{T_{ref}}^{z}=F_0\times L_r=F_0\times 10^{\frac{T-T_{ref}}{z}},$$

The ratio of F_0 value of the process to the F_0 required for commercial sterility is known as lethality. Thus, lethality must be at least 1 for commercial sterility of the product (David et al., 1996).

C Value

The destruction of nutrients and inactivation of enzymes follow similar first order kinetics to that of microbial destruction. Destruction of nutrients in food products is quantified by the term cook value (*C*), which has been defined as:

$$C = \int_0^t 10^{\frac{T-T_{\text{ref}}}{z_c}} \, dt$$

The *C* value at a reference temperature of 100°C (212°F) and a z_c value of 33.1°C (59.6°F) is referred to as the C_0 value.

Process Optimization

The objective of a food processor is to produce a safe product which retains nutritional and quality attributes at an acceptable level. Therefore, the appropriate combination of time and temperature used for processing is based on factors such as nutrient retention and enzyme inactivation in addition to safety. D_c and z_c values for destruction of nutritional and quality attributes are generally larger than those of microorganisms. This implies that the rate of destruction of microorganisms at a higher temperature will be much higher than the rate of destruction of nutritional and quality attributes. Thus, thermal processing of food products at higher temperature can achieve commercial sterility with better retention of nutritional and quality attributes (David et al., 1996).

Heating Equipment

Retorts

A retort (usually a large cylindrical vessel) works on the principle of cooking under pressure. Pressurized hot water (water immersion, cascading water, or water spray) or steam in the retort is used to raise the temperature of a product already in a package (metal can, glass, metallized pouch, or plastic) to approximately 120°C–125°C. This high of a temperature is needed to inactivate spores of *C. botulinum*, which is the target organism for producing shelf-stable foods that have a pH of 4.6 or higher. In most types of retorts, several (hundreds) containers of food are placed in crates/baskets with or without perforations and one or more of these crates are loaded into the retort for processing. Generally, the older style retorts are oriented vertically, whereas newer retorts are oriented horizontally to facilitate movement of product in and out of the retort easily.

1. Classification based on continuity of process
   Based on continuity of the process, retorts are classified as batch or continuous. Batch retorts may be used by small or large processors. However, continuous retorts are used only by large processors.
   a. Batch
      In a batch retort, containers are loaded into the retort, thermally processed, cooled, and then taken out of the retort. Batch retorts may be oriented horizontally or vertically. Horizontal retorts involve easy loading of crates by means of a rail or track on which multiple crates can be pushed into the retort. In vertical retorts, a crane mechanism has to be used to hoist crates and load them into the retort. The heating medium for batch retorts may be hot water (in the form of containers being immersed in hot water, flowing/cascading over the containers or being sprayed by water).

      Large water immersion retorts generally consist of two horizontal cylinders, one on top of another. The bottom cylinder contains the packaged food product in crates, whereas the top cylinder contains hot water under pressure. When the containers are ready for thermal processing, the preheated water from the upper cylinder is

quickly flooded into the bottom cylinder via large ports connecting the two cylinders. Preheating the water in the upper cylinder shortens the come-up time of the retort (time taken to get all points within the retort to the required processing temperature). If there were only one cylinder and steam was used to heat a large volume of water in that cylinder, it would take a long time for the water to attain the required temperature.

In a cascading water retort, rows of containers in a crate are separated from containers in the row above by means of perforated divider plates. Pressurized hot water then cascades over containers through the perforations, thereby ensuring that all containers receive relatively uniform flow of water over it.

In a water spray retort, pressurized hot water is sprayed over containers of food via multiple nozzles placed at specific locations within the retort. The orientation of the nozzles and the spacing between them are critical in ensuring uniform thermal treatment of all containers of food.

b. Continuous

In a continuous retort, containers are continuously loaded into the retort at a certain rate at one end and continuously taken out of the retort at the same rate at the other end. Continuous retorts increase the throughput and lower the manpower cost. Continuous retorts can be classified as static (hydrostatic) and rotary (hydrolock, sterilmatic, reel and spiral, etc.) retort.

Hydrostatic retorts are vertical systems which uses a water leg as a steam valve. A tall leg of water results in an increase in pressure at the bottom of the leg. At this point, steam can be used to increase the temperature of the water to values as high as 130°C (or higher, if needed), depending on the height of the water. The hydrostatic retort is well suited for products that require long cook and cool times along with higher throughput. The containers of food are transported through this leg of water by means of chains containing slots for the containers to be placed. As the containers travel through the retort, they are subjected to increasing temperatures as they move from top to bottom. Once they receive adequate heat treatment, they pass through legs of water wherein cold water is used to cool off the products.

A continuous rotary retort is a fully automated system designed for high throughput, lower energy consumption, and uniform product quality. These systems require a cylindrical container with limited variation in can diameter and height (Weng, 2005). Containers of food are placed in slots on the outer part of a helical reel mounted around a horizontal cylindrical vessel that is jacketed. Steam is used to heat the products in cans as they are transported in a spiral manner on the reel from one end of the retort to the other.

2. Classification based on mixing

Retorts are also classified based on the nature of mixing induced within the product as static, rotary, and reciprocating. The container size, type of product, and quality considerations generally dictate the type of mixing imparted to the containers of food.

a. Static

As the name suggests, in a static retort, there is no movement of the containers. When there is no movement, the only mechanism for equalization of temperatures within the food in containers is natural (free) convection. Thus, for products that are not viscous or do not contain large particles, this type of retort may be applicable. Additionally, it may be suited in situations where the container size is small enough so that the temperature gradient within the product is not large.

Generally, we associate static retorts with small-scale operations. However, static retorts may be small or large and may or may not contain crates/baskets. A crateless static retort is one type of static retort used by large processors. In this, containers of food are transported on a conveyor belt and loaded from the top into a large vessel

containing water to cushion the fall of containers into the retort. Once the containers are loaded, steam is introduced into the retort to begin the cooking process. Once thermal processing is completed, cold water is introduced into the retort to cool the product. Once cooling is completed, the bottom of the retort opens up and containers fall into a flume of water (cushion) and are transported via a conveyor belt for labeling and palletization.

b. Rotary
When dealing with viscous and particulate food products or large containers, the temperature gradient within a food product may be too large, thereby compromising the quality of the product towards the center of the container. In these cases, a rotary motion is imparted to the containers to mix the product, and equalize the temperature distribution within the product. The rotation may be imparted in an axial or end-over-end manner.

In axial rotation, containers are rotated about their central axis. This causes the air bubble within the container to continuously move along the periphery of the container. In a batch retort, this type of rotation is usually imparted to containers when they are placed in L-shaped slots attached to a reel that rotates within the retort. In a continuous retort, this type of rotation is imparted when the containers are placed on ledges/slots on the periphery of a reel that rotates.

In end-over-end rotation, containers are subjected to flipping action across their ends. This causes the air bubble within the container to continuously move from the top to the curved surface to the bottom. Unlike in axial rotation wherein individual containers rotate about their axis, in end-over-end rotation, generally, a group of containers (usually, one or more crates/baskets) are rotated as one entity. For this to be accomplished, the containers are packed tightly within the crate/basket and held in place by some mechanism to ensure that they do not move around during rotation. This type of agitation is conducive to only batch retorts.

c. Reciprocating
In a reciprocating retort, one or more crates/baskets are placed within the retort and the entire system of crates (with containers) is moved back and forth at a preset speed. This reciprocating action imparts a random mixing action within the product. Studies have shown that the time of processing decreased and product quality increased when using this type of retort. However, one of the concerns with this type of retort is the relatively rapid wear and tear of moving parts within the retort.

Tubular Heat Exchanger

A tubular heat exchanger consists of two or more concentric tubes. The product and heating/cooling media flow in alternate tubes in the same or opposite directions. The double tube heat exchanger (two concentric tubes) is the most common type of tubular heat exchanger. Triple tube heat exchangers (three concentric tubes) are used occasionally, whereas quadruple tube heat exchangers (four concentric tubes) are not commonly used.

1. Double tube
The most common noncontact type of heat exchanger is the double tube heat exchanger. It consists of a tube located concentrically inside another tube of larger diameter. The two fluids flow in the inner and outer tube respectively (Figure 4.20.3).

Double tube heat exchangers can be classified according to the path of fluid flow through the heat exchanger. The three basic flow configurations are cocurrent, countercurrent, and cross-flow. In cocurrent heat exchangers, both the fluid streams enter simultaneously at one end and leave simultaneously at the other end as shown in Figure 4.20.3. In countercurrent heat exchangers, both of the fluid streams flow in opposite directions. In cross-flow heat

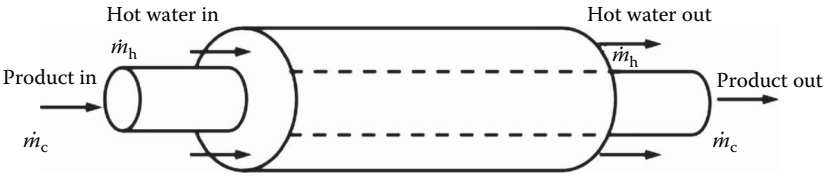


FIGURE 4.20.3 Schematic of a double tube heat exchanger.

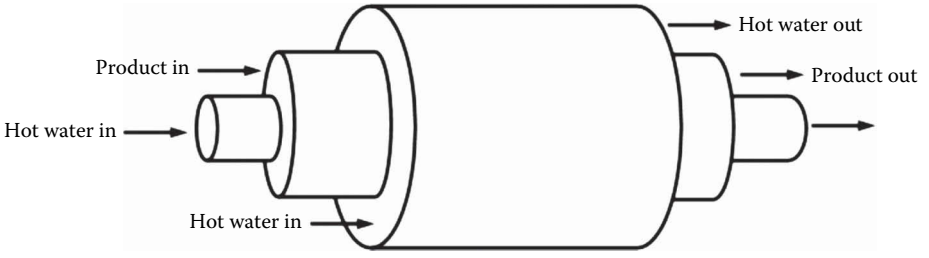


FIGURE 4.20.4 Schematic of a triple tube heat exchanger.

exchangers, one fluid flows through the heat transfer surface at right angles to the flow path of the other fluid (Kakac and Liu, 2000). When both fluid streams are simultaneously in cocurrent, countercurrent, and multipass cross-flow, then the arrangement is called mixed flow. Generally, a countercurrent configuration is preferred since it results in a higher amount of heat transfer. However, when rapid initial cooling of a heat-sensitive product is desired (to minimize quality losses occurring at a greater extent at high temperatures), a cocurrent configuration may be used. In this case, a cocurrent configuration may be used for the initial cooling and a countercurrent configuration may be used for the latter stage of cooling.

2. Triple tube
   A slightly modified version of the double tube heat exchanger is the triple tube heat exchanger. It consists of three concentric tubes. Product flows in the inner annular space and the heating/cooling medium flows in the inner tube and outer annular space (Figure 4.20.4).

   Advantages of a triple tube heat exchanger over a double tube heat exchanger include larger heat transfer area per unit length and higher overall heat transfer coefficients due to higher fluid velocities in the annular regions. Batmaz and Sandeep (2004) developed an expression for effective overall heat transfer coefficient for a triple tube heat exchanger to facilitate comparison of a triple tube heat exchanger to an equivalent double tube heat exchanger.

Shell and Tube

Shell and tube heat exchanger is another common noncontact type of heat exchanger used in the food industry. In this type, there is a small or large shell that houses multiple tubes within it. When the shell is small, it is often called a multitube heat exchanger and when the shell is large, it is the conventional shell-and-tube heat exchanger.

1. Conventional shell and tube
   This consists of a bundle of tubes connected in parallel and enclosed in a large cylindrical shell. A schematic of a shell-and-tube heat exchanger is shown in Figure 4.20.5.

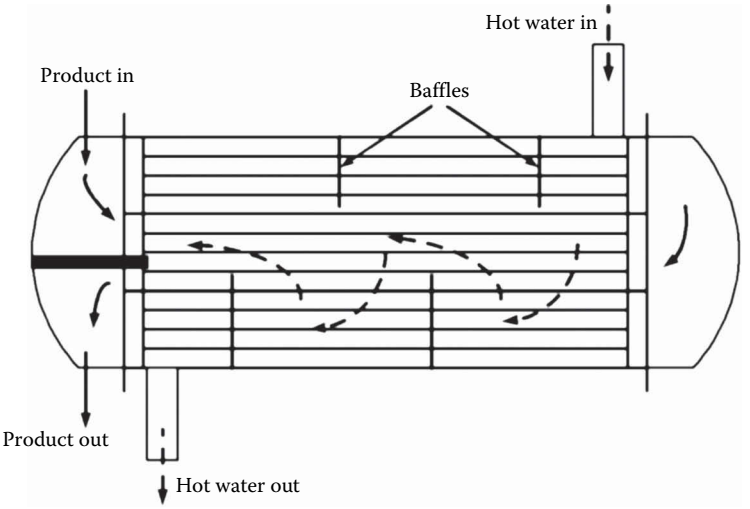


FIGURE 4.20.5 Schematic of a shell-and-tube heat exchanger.

One of the fluids flows inside the tubes while the other fluid flows over the tubes through the shell. To ensure the flow of shell-side fluid over all of the tubes, baffles are used in these heat exchangers. The segmented baffle is the most commonly used baffle and it causes the shell-side fluid to flow in a zigzag manner across the tube bundle. This enhances turbulence on the shell side of the heat exchanger and thus improves the rate of heat transfer. However, improperly designed segmented baffles cause low local heat transfer coefficients in flow stagnation regions where the baffles are attached to the shell wall. Baffles also result in excessive pressure drop by separating the flow at the edges of the baffles. Helical baffles are an alternative to segmented baffles as they offer the following advantages: improved shell-side heat transfer, lower pressure drop for a given shell-side flow rate, reduced shell-side fouling, and prevention of flow-induced vibrations. However, helical baffles are associated with fabrication and manufacturing difficulties (Stehlik et al., 1994). The tubes may be arranged to permit one or more passes of the product through the tubes.

2. Helical shell and tube
   In this type of heat exchanger, tubes are coiled and placed in a large cylindrical shell. The ratio of the diameter of the tube to the diameter of the coil and the vertical spacing between successive coils are important design parameters. The main difference in the flow in straight tubes and helical tubes relates to secondary flow. For flow of a Newtonian fluid under laminar flow conditions in a straight tube of circular cross section, a parabolic velocity profile exists under fully developed flow conditions. In this case, flow takes place only in the axial direction. However, in a coiled tube, flow takes place predominantly in the axial direction (primary flow), but also partly in the radial direction (secondary flow). This is called the Dean effect. The strength of primary flow is generally much greater than that of secondary flow. Nevertheless, the secondary flow creates radial movement of fluid elements and hence causes mixing of fluid elements. This increases the rate and uniformity of temperature distribution within the product.

Scraped Surface Heat Exchanger

Scraped surface heat exchangers are used for viscous products (diced fruit preserves and soups) containing particles of sizes up to approximately 15 mm. Particle concentrations of up to 40% can be accommodated by these heat exchangers. These heat exchangers consist of a jacketed cylinder enclosing scraping blades on a rotating shaft. The rotating action of the blades prevents fouling on

the heat exchanger surface and improves the rate of heat transfer and mixing. A summary of studies on heat transfer in scraped surface heat exchangers has been presented by Rao and Hartel (2006).

Electromagnetic Heating

Electromagnetic heating refers to the use of electromagnetic waves (generated when an electric field couples with a magnetic field) to heat a food product. While infrared, radio frequency, and microwaves can cause heating of a food product, infrared heating is not often used as it results predominantly in heating only the surface of a food product. However, radio frequency waves and microwave can cause heating within a food product, and hence they have been successfully used for various food processing applications.

1. Radio frequency
   Radio frequency (RF) waves are a part of the electromagnetic spectrum and have a frequency between 3 kHz and 300 MHz. Their low frequency, in comparison to that of microwaves, results in a greater depth of penetration of the waves into a product. RF systems are typically equipped with electrodes, the polarity of which is reversed at a preset frequency. The commonly used frequencies with RF are 13.56, 27.12, and 40.68 MHz, with the latter being the most common. They may be used in a conveyor belt configuration for treating discrete solid materials or as a flow-through tube configuration for pumpable materials.

   RF has been used in the timber industry for drying glued joints and in the textile industry for drying heat-sensitive fabrics. In the food industry, it has been used predominantly for the post-bake stage drying (final drying) of cookies and crackers. The main benefit of using RF over hot air for this is that hot air heats up all parts of the cookies and crackers, thereby having a tendency to crack them, while RF targets only the polar component within the product (mainly water), thereby preventing the tendency to cause cracking.
2. Microwave
   Microwaves are a part of the electromagnetic spectrum and have a frequency between 300 MHz and 300 GHz. They lie between the radio (3 kHz–300 MHz) and infrared (300 GHz–400 THz) frequencies of the electromagnetic spectrum. Microwave radiation has the ability to heat materials by penetrating and dissipating heat in them. The important advantages of microwave heating as compared with conventional heating include instant start-up, faster heating, and energy efficiency. The main disadvantage associated with microwave heating is nonuniform heating of the food products. In the U.S., only four microwave frequencies (915 ± 13; $2{,}450 \pm 50$; $5{,}800 \pm 75$; and $24{,}150 \pm 125$ MHz) are permitted by the Federal Communications Commission (FCC) for industrial, scientific, and medical applications.

   The power absorption for volumetric heating using microwave is given as (Coronel et al., 2008):

$$P = 2\pi f \varepsilon_0 \varepsilon'' E^2,$$

   where P is power absorbed (W/m^3) per unit volume, f is the microwave frequency (Hz), and E is the electric field strength (V/m). Power penetration depth (δ_p), often used in microwave heating applications, is the distance within the product at which power drops to e^{-1} of its value at the surface of the material and is given by the following equation (Nelson and Datta, 2001):

$$\delta_p = \frac{\lambda}{2\pi\sqrt{2\varepsilon'\left[\sqrt{1+\left(\frac{\varepsilon''}{\varepsilon'}\right)^2}-1\right]}},$$

where λ is the wavelength of the microwave in free space. Power penetration depth is used to calculate the tube diameter for a continuous flow microwave heating system.

Continuous flow microwave heating systems have been developed to rapidly and uniformly heat various viscous and particulate foods. The resulting product has much better quality and nutritional attributes than the equivalent product processed using conventional thermal processing technologies.

Ohmic Heating

Ohmic heating, also known as Joule heating, electric resistance heating, and electroconductive heating is a process in which an alternating current is passed directly through a conductive food product. The heat is generated internally due to the resistance of the food product to the applied electric current. The heating is rapid due to volumetric generation of heat. The heat generated is given as:

$$P = E^2\sigma,$$

where E is the electric field strength and σ is the electrical conductivity. Electric field strength (V/m) can be varied by adjusting the electrode gap or the applied voltage. Electrical conductivity is the measure of how well a substance conducts electricity and is expressed in the units of Siemens per meter (S/m). The efficiency of Ohmic heating is dependent on the electrical conductivity (σ) of the product (Coronel et al., 2008).

Ohmic heating has been used successfully to thermally process various food products, including various types of soups, especially in Europe. For some products, soaking the product in salt solution, to increase its electrical conductivity, may be needed.

Process Considerations

The most common methods of thermal processing include blanching, pasteurization, hot filling, and sterilization. These methods have been described below.

Blanching

Blanching is a mild heat treatment commonly applied to fruits and vegetables prior to freezing, drying, or canning. Blanching is performed to inactivate enzymes, enhance drying and rehydration, remove tissue gases, enhances color of green vegetables, and reduce the microbial population. The effectiveness of blanching is usually evaluated by enzyme activity (peroxidase and catalase). Blanching is usually accomplished by bringing the product in contact with hot water, hot air, or steam for certain period of time. Water blanching can be conducted as a batch operation by dipping a batch of product in hot water for the required amount of time. Continuous hot water blanching can be accomplished using a screw-type, drum-type, or pipe-type blancher. The screw-type blancher consists of a trough fitted with a helical screw. The drum-type blancher consists of a perforated drum fitted internally with a helical screw. The pipe-type blancher can be used for solid products which can be pumped with water. Similar to water blanching, steam blanching can also be accomplished as a batch or continuous process. The heating time necessary to accomplish blanching depends on the type and size of fruit/vegetable, the method of heating, and the temperature of the heating medium. Typical blanching times at 100°C for commercial blanching range from 1 to 5 min (Lund, 1975).

Pasteurization

Pasteurization refers to heat treatment of a product to kill all vegetative pathogenic microorganisms in it. The time–temperature combination for the pasteurization of milk is 63°C for 30 min, which is referred to as a low-temperature long time process and 72°C for 15 s which is referred to as a high-temperature short time process. The heat treatment in pasteurization is not sufficient to inactivate

all spoilage-causing vegetative cells or heat-resistant spores. Therefore, the shelf life of pasteurized low-acid products such as milk and dairy products is approximately 2–3 weeks under refrigerated conditions. Ultrapasteurization refers to pasteurization at temperatures of 138°C or above for at least 2 s either before or after packaging. This process further extends the shelf life of the product. Ultrapasteurization results in the destruction of a greater proportion of spoilage microorganisms, leading to an extended shelf life of about 6–8 weeks. This process has been used for flavored milks and nondairy creamers in portion pack cups (David et al., 1996).

The choice of heating system for pasteurization depends on the characteristics (rheological and thermal properties) of the product, potential for fouling, ease of cleaning, and cost.

Direct-type heating system (steam injection and steam infusion) is used for homogeneous and high-viscosity products and is particularly suited for shear-sensitive products such as creams, dessert, and sauces. In a steam injection heating system, liquid product is heated by injection of culinary steam into the product. Rapid heating by steam combined with rapid methods of cooling can yield a high-quality product. A steam infusion heating system, similar to steam injection, involves infusing a thin film of liquid product into an atmosphere of steam which provides rapid heating. Direct heating system (steam injection or steam infusion) adds water to the product (10% water addition per 56°C increase in product temperature) due to the condensing steam. The amount of added water should be either accounted for in the product formulation or removed by flash cooling in a vacuum chamber.

There are four main types of indirect-type heating systems: tubular, shell and tube, plate, and scraped surface heat exchangers. Tubular heat exchangers are used for homogeneous and high-viscosity products (soups and fruit purees) containing particles of sizes up to approximately 10 mm. The simplest tubular heat exchanger is a double-pipe heat exchanger consisting of two concentric pipes. Shell and tube heat exchangers consist of a shell (typically cylindrical in shape) with one or more sets of tubes inside it. The tubes may be coiled in a helical manner or arranged in a trombone fashion. This type of heat exchanger is used when a greater degree of mixing than that achieved in a tubular heat exchanger is desired. Plate heat exchangers are used for homogeneous and low-viscosity (<5 Pa s) products (milk, juices, and thin sauces) containing particle sizes up to approximately 3 mm. These heat exchangers consist of closely spaced parallel plates pressed together in a frame. These heat exchangers provide a rapid rate of heat transfer due to the large surface area for heat transfer and turbulent flow characteristics. Scraped surface heat exchangers are used for viscous products (diced fruit preserves and soups) containing particles of sizes up to approximately 15 mm. These heat exchangers consist of a jacketed cylinder housing with scraping blades on a rotating shaft. The rotating action of the scraping blades prevents fouling on the heat exchanger surface and improves the rate of heat transfer. This type of heat exchanger is the best choice for viscous products containing particulates (Skudder, 1993). Apart from tubular, shell and tube, plate, and scraped surface heat exchangers, pasteurization can also be accomplished in a vat- or tank-type heat exchanger. In a tank-type heat exchanger, product is pumped into a jacketed vat or tank, heated to pasteurization temperature, held for required time, and pumped from the vat to the cooling section (Mitten, 1963).

Volumetric heating systems such as microwave and ohmic heating can provide very rapid heating which is desirable for aseptic processing. However, it is challenging to maintain a uniform temperature distribution within the product. Microwave heating systems apply rapidly changing electromagnetic field to the product. Movement or vibration of the molecules of the product due to the changing electromagnetic field generates heat. An ohmic heating system operates by directly passing the electric current through a product. The electrical resistance of the product to the passing electric current generates heat (Coronel et al., 2008).

Hot Fill

Acid/acidified products such as juices and beverages packed in hermetically sealed containers using an appropriate hot filling process yields a commercially sterile shelf-stable product. Hot filling, also

known as "hot fill and hold," refers to filling unsterilized containers with a sterilized acid/acidified food product that is hot enough to render the container commercially sterile. A hermetically sealed container is a container that is designed and intended to be secure against the entry of microorganisms and thereby maintain the commercial sterility of its contents after processing. Commercial sterility means the condition achieved either by (1) the application of heat which renders the food free of microorganisms capable of reproducing in the food under normal nonrefrigerated conditions of storage and distribution and viable microorganisms (including spores) of public health significance, or by (2) the control of water activity and the application of heat, which renders the food free of microorganisms capable of reproducing in the food under normal nonrefrigerated conditions of storage and distribution. Commercially sterile food products are shelf-stable with a longer shelf life (1–2 years) (David et al., 1996; Anderson et al., 2011).

Sterilization

Sterilization refers to killing of all living microorganisms including spores in the food product. Food products are never completely sterilized, instead they are rendered commercially sterile. Sterile product means that there is a complete absence of viable microorganisms in it. A viable microorganism is one that is able to grow when exposed to optimum conditions. Low-acid food products are commercially sterilized since they are capable of sustaining the growth of *C. botulinum* spores (Lund, 1975; David et al., 1996). Commercial sterility can be achieved by in-container sterilization and in-flow sterilization. In-container sterilization generally refers to the retorting process, whereas the in-flow sterilization refers to aseptic processing.

1. Retorting
   Traditionally, retorting has been used to process low-acid food products to ensure destruction of *C. botulinum* spores. Conventional retorting involves filling of the product in metal cans, glass jars, retortable semirigid plastic containers or retortable pouches; double-seaming or heat sealing; followed by heating, holding, and cooling in a pressurized batch or continuous retort. Retorting of foods in cans, invented by Nicholas Appert in the early 1800s, still remains the preferred choice for preservation of foods. Retorts can be operated in either a batch or continuous mode of operation. Batch retort is the most versatile sterilization system with the ability to handle different products (conduction heating and convection heating) and package types. Batch retort can further be classified into still/static (horizontal, vertical, or crateless) retort and agitating/rotary (end over end or axial rotation) retort. When steam is used as the heating medium, it should be introduced into the retort with care such that all the air in the retort is displaced. Removal of air by steam is also known as venting. Cooling is accomplished by shutting off steam and introducing cold water into the retort. Overpressure is often used to prevent internal pressure inside the container from destroying its integrity. The rotary retort agitates the product inside the container by the movement of the air bubble created by the headspace, resulting in enhanced heat transfer in the container. A larger headspace results in faster heating of a product due to efficient mixing. Different heating media and heating methods used in various batch retorts include steam, water, steam–air, water cascading, water spray, or water immersion (Lund, 1975; Weng, 2005).
2. Aseptic processing
   Aseptic processing offers an alternative to conventional retorting to meet the demand for convenient and high-quality foods. Aseptic processing of foods is a process in which the product and the package are sterilized separately and brought together in a sterile environment. It involves sterilization of a food product, followed by holding it for a specified period of time in a holding tube, cooling it, and packaging it in a sterile container. Aseptic processing uses high temperature for a short period of time, yielding a high-quality (nutrients, flavor, color, or texture) product as compared to that by conventional canning.

Some of the other advantages associated with aseptic processing include longer shelf life (1–2 years at ambient temperature), flexible package size and shape, less energy consumption, less space requirement, eliminating the need for refrigeration, easy adaptability to automation, and need for fewer operators. However, some of the disadvantages of aseptic processing include slower filler speeds, higher overall initial cost, need for better quality control of raw ingredients, better trained personnel, better control of process variables and equipment, and stringent validation procedure. Some of the products which are aseptically processed include fruit juices, milk, coffee creamers, purees, puddings, soups, baby foods, and cheese sauces (David et al., 1996). Sterilization of products in aseptic processing can be accomplished using tubular, shell and tube, scraped surface, or volumetric heating (microwave and ohmic) systems.

Outlook for the Future

Thermal processing has been the most common technique used for inactivating microorganisms. Though nonthermal techniques such as high-pressure processing have been successfully used for certain special categories of foods, thermal processing will likely continue to be the predominant technique used for inactivating microorganisms in food. Some of the advancements expected in the area of thermal processing of foods include improved design of conventional heat exchangers for heating foods, improvements in volumetric heating technologies for rapidly and uniformly heating foods, development of rapid cooling techniques, and development of miniature sensors that can be implanted within food particles to determine the time–temperature history at the cold spot within a food so as to facilitate process validation. Additional work on synergistic effects of heat and other hurdles (such as antimicrobials, redox potential, and nonthermal technologies) is also expected to be an area of research and development of interest to food processors.

REFERENCES

Anderson, N.M., Larkin, J.W., Cole, M.B., Skinner, G.E., Whiting, R.C., Gorris, L.G.M., Rodriguez, A., Buchanan, R., Stewart, C.M., Hanlin, J.H., Keener, L., and Hall, P.A. (2011). Food safety objective approach for controlling *Clostridium botulinum* growth toxin production in commercially sterile foods. *Journal of Food Protection* 74(11), 1956–1989.

Batmaz, E. and Sandeep, K.P. (2004). Calculation of overall heat transfer coefficients in a triple tube heat exchanger. *Heat and Mass Transfer* 41, 271–279.

Coronel, P.M., Sastry, S., Jun, S., Salengke, S., and Simunovic, J. (2008). Ohmic and microwave heating. In: *Engineering Aspects of Thermal Food Processing*. (ed. R. Simpson), pp. 73–92. CRC Press, Inc., Boca Raton, FL.

David, J.R.D., Graves, R.H., and Carlson, V.R. (eds) (1996). *Aseptic Processing and Packaging of Food: A Food Industry Perspective*. CRC Press, Inc., Boca Raton, FL.

Kakac, S. and Liu, H., (2000). *Heat Exchangers: Selection, Rating, and Thermal Design*, 2nd ed., pp.1–501. CRC Press LLC, Boca Raton, FL.

Krokida, M.K., Panagiotou, N.M., Maroulis, Z.B., and Saravacos, G.D. (2001). Thermal conductivity: Literature data compilation for foodstuffs. *International Journal of Food Properties* 4, 111–137.

Lund, D.B. (1975). Heat processing. In: *Principles of Food Science Part II: Physical Principles of Food Preservation*. (eds M. Karel, O.R. Fennema and D.B. Lund), pp. 31–92. Marcel Dekker, New York.

Metaxas, A.C. and Meredith, R.J. (1983). *Industrial Microwave Heating*, pp. 26–69. Peter Peregrinus, Ltd., London.

Mitten, H.L. (1963). Heater-coolers, heat-exchanger equipment, and milk storage tanks. In: *Engineering for Dairy and Food Products*. (ed. A.W. Farrall), pp. 268–296. John Wiley & Sons, Inc., Hoboken, NJ.

Nelson, S.O. and Datta, A.K. (2001). Dielectric properties of food materials and electric field interactions. In: *Handbook of Microwave Technology for Food Applications*. (eds A.K. Datta and R.C. Anantheswaran), pp. 69–114. Marcel Dekker, Inc., New York.

Rao, C.S. and Hartel, R.W. (2006). Scraped surface heat exchangers. *Critical Reviews in Food Science and Nutrition* 46, 207–219.

Saltiel, C. and Datta, A.K. (1999). Heat and mass transfer in microwave processing. *Advances in Heat Transfer* 33, 1–94.

Singh, R.P. (1994). *Food Properties Database*. Version 2. CRC Press, Boca Raton, FL.

Singh, R.P. and Heldman, D.R. (2001). Heat transfer in food processing. In: *Introduction to Food Engineering*. 3rd ed. (eds R.P. Singh and D.R. Heldman), pp. 216–221. Academic Press, Oxford.

Skudder, P.J. (1993). Ohmic heating. In: *Aseptic Processing and Packaging of Particulate Foods*. (ed. E.M.A Willhoft), pp. 74–89. Blackie Academic & Professional, London.

Stehlik, P., Nemcansky, J., Kral, D., and Swanson, L.W. (1994). Comparison of correction factors for shell-and-tube heat exchangers with segmented or helical baffles. *Heat Transfer Engineering* 15(1), 55–65.

Tucker, G. and Featherstone S., ed. (2011). *Essentials of Thermal Processing*. Blackwell Publishing Ltd, Oxford.

Weng, Z.J. (2005). Thermal processing of canned foods. In: *Thermal Food Processing New Technologies and Quality Issues*. (ed. D.W. Sun), pp. 335–362. CRC Press, Inc., Boca Raton, FL.

4.21 THERMAL CONDUCTION IN ELECTRONIC MICROSTRUCTURES AND NANOSTRUCTURES

Sanjiv Sinha, Krishna Valavala and Jun Ma

Structures with physical dimensions ranging from a few nanometers to hundreds of micrometers are the essential building blocks of electronic integrated circuits. Current flow in circuits causes Joule heat dissipation in transistors as well as in interconnects (Waldrop, 2016; Li et al., 2009; Pop et al., 2006). Heat transfer engineering across multiple length scales (Sinha and Goodson, 2005) is necessary to limit the resulting temperature rise. Typically, the highest temperatures in a circuit occur inside the transistors (Su et al., 1994). For design purposes, the distribution of transistor temperatures is ignored in favor of a single temperature for the circuit that represents the maximum value. This is commonly referred to as the junction temperature. Increased junction temperatures affect nearly all key figures of merit of a very large-scale integrated (VLSI) circuit. In particular, temperature affects circuit speed, lifetime, power dissipation, and power plane integrity (Haensch et al., 2006; Hanson et al., 2006). Effective thermal management is necessary across the gamut of modern electronic circuits, from established applications such as microprocessors, power semiconductors, semiconductor lasers, sensors, actuators, and light-emitting diodes to emerging and related areas such as phase change memory (PCM), flexible electronics, and spintronics. Figure 4.21.1 shows a few examples of semiconductor devices that form the key elements of diverse electronic circuits in which heat transfer issues arise.

Across decades, advancements in electronics largely relied on the continued miniaturization of the transistor, in keeping with the famous Moore's law (Moore, 1965). The road map followed thus far by the industry has targeted shrinking the physical dimensions of the transistor to achieve faster circuits and more devices per area, which in turn have enabled even larger-scale integration. Underlying this road map is the principle of "device scaling" (Dennard et al., 1974). The essential idea is to shrink the distance that current must travel in the transistor (approximately equal to the length of the gate terminal) while maintaining the same electric field. In practice, this requires ticking off a complex engineering checklist that spans expertise in semiconductor fabrication, materials physics and chemistry, device physics, circuit design, and computer architecture to name a few. While heat transfer did not feature near the top of the design checklist for many decades of the evolution of Moore's law, power dissipation has always been a major consideration (Mahajan et al., 2000). In the early years of microelectronics, a shift from bipolar transistors to complementary metal-oxide-semiconductor (CMOS) field-effect transistors (FETs) was necessary to limit power dissipation. With the introduction of CMOS, power ideally would be dissipated only during the process of switching the state of a device. Individual transistors became more energy efficient with shrinkage, but the chip power density crept up with increased transistor count per unit area (Mahajan et al., 2002; Samson et al., 2005). In the last decade, transistors have approached a size regime where quantum mechanical tunneling introduces leakage currents even during the off state where ideally no power should be dissipated (Haensch et al., 2006). Figure 4.21.2 shows the power density for both active and the passive states versus gate length. As dimensions reach below 100 nm, both active and passive power contribute equally to the total power density. The industry has responded with many innovations to limit leakage power, including three-dimensional transistors (Topol et al., 2006). However, increased power density with successive generations of integrated circuits has ultimately put a cap on the switching speed (or frequency). And further miniaturization of the most cutting-edge semiconductor devices would make them atomistic scales where altered transport physics stem gains from miniaturization. Even with the expected demise (Waldrop, 2016) of the conventional form of Moore's law, there remain significant opportunities for innovation such as three-dimensional circuits, where heat transfer engineering continues to play an important role.

Figure 4.21.3 shows the schematic of a packaged microprocessor and illustrates the heat flow path for silicon microprocessors (Prasher, 2006). Heat generation occurs in the transistors and

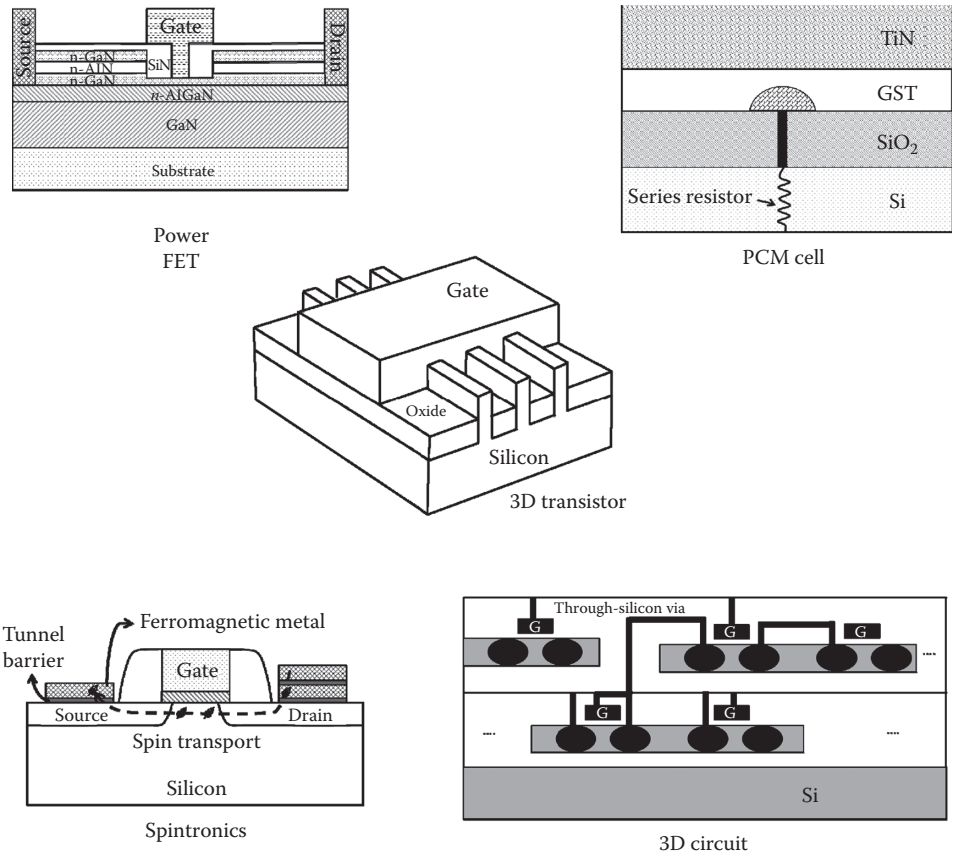


FIGURE 4.21.1 Examples of the diversity of semiconductor devices. In power conversion, GaN FET handle high voltages and power densities. A PCM cell uses heat to change the electrical resistance of phase change material. Advanced three-dimensional silicon transistors are nanoscale with sub-20 nm gate lengths (approximately equal to the distance between the source and drain terminals). The electron spin state is employed as the fundamental signal in a spintronic transistor. Three-dimensional circuits integrate multiple device layers and require innovative heat flow paths.

interconnects. The major heat flow path is toward the heat sink from where heat is transferred to the ambient through convection. At steady state, the thermal circuit shown in Figure 4.21.3 applies. The key temperature nodes are those at the junction (T_J), the case (T_C), the sink (T_S), and the ambient (T_O). The key thermal resistances are due to the die (R_{DIE}), the first-level thermal interface material (R_{TIM1}), the heat spreader ($R_{SPREADER}$), the second-level thermal interface material (R_{TIM2}), and the heat sink (R_{SINK}). Heat produced in the chip flows across the first-level thermal interface material (TIM1) to a heat spreader that is typically integrated with the chip. The heat spreader aims to address nonuniformity in heat production across the surface of the chip. Computational cores, for example, produce more heat than memory caches. The second-level TIM (TIM2) and the heat sink help to reject heat to the ambient.

Thermal management of an integrated circuit spans several distinct engineering disciplines. Figure 4.21.4 shows the multiple scales involved in thermal engineering of integrated circuits. Device level thermal considerations form the first level of thermal management where the effect of elevated temperatures on the performance and reliability of transistors must be understood. The level of detail goes down to the microscopic physics of electron scattering. Circuit level thermal considerations form the second level. It is now common for VLSI circuit designers to incorporate thermal modeling in circuit models to more accurately predict power dissipation, reliability, circuit

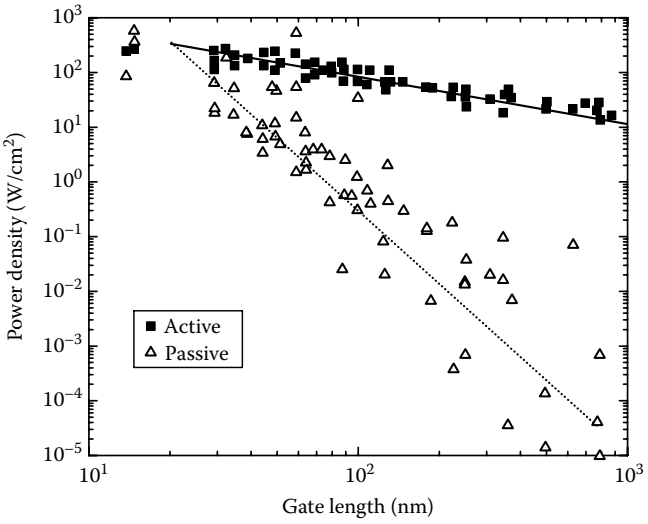


FIGURE 4.21.2 With shrinking gate length, the power density expended in useful functions is equaled by the passive power density, placing greater requirements on thermal engineering down to the circuit level. (Data from Haensch, W. et al., *IBM J. Res. Dev.,* 50 (4/5), 339–61, 2006).

timing characteristics, and supply voltage variations among others. At a higher level, software techniques such as dynamic thread migration and temperature aware task scheduling also find use in limiting temperature rise. On the hardware side, the focus is on cooling technologies (Garimella et al., 2006; Shakouri, 2006), TIMs (Prasher, 2006), and thermal characterization.

Based on the physical length scales involved, the heat transfer problem is typically separated into three hierarchies: chip level that ranges from the transistor to the circuit to the die; package level that includes TIMs and heat spreaders integrated with the die; and finally, system or platform level that includes external heat pipes, heat sinks, fans as well as other electronic components on the printed circuit board. This chapter is mainly concerned with the first two levels of hierarchy. The only mode of heat transfer available is conduction. However, the orders of magnitude in the length scales involved bring significant complexity to the conduction physics. Further, interconnection between charge and heat transport at the smallest length scales in the problem plays an important role in heat transfer engineering. The following material delves first into fundamental physics and then discusses relevant modeling and experimental techniques. Data for a wide range of materials relevant to electronics follow. Application of the techniques to semiconductor devices is discussed at the end of the chapter.

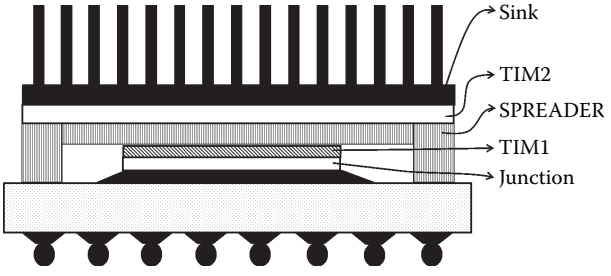


FIGURE 4.21.3 Schematic of desktop microprocessor packaging. The junction temperature is the highest temperature in the silicon die. Heat flows to the ambient toward the top. The path includes the TIM1, an integrated spreader, TIM2, and a heat sink. Nonuniformity of heat generation in the die complicates thermal solutions.

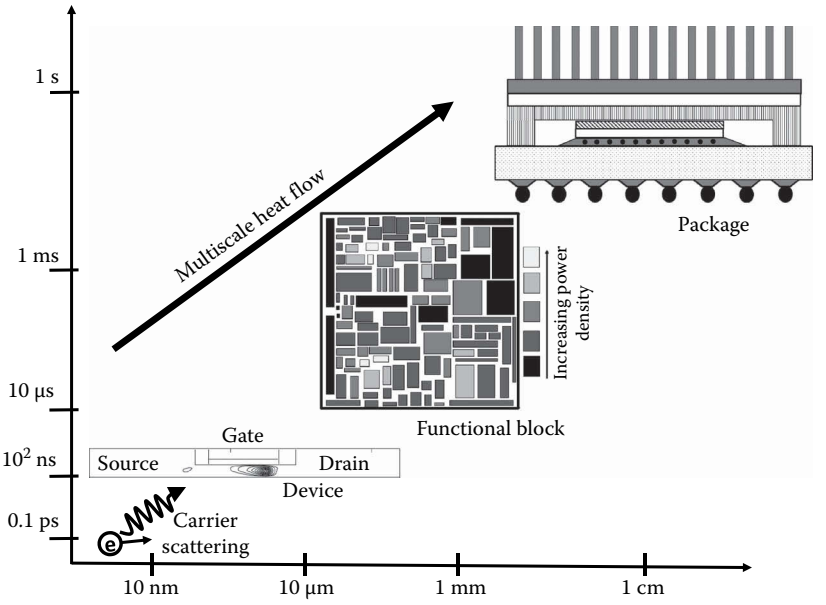


FIGURE 4.21.4 Heat flow in semiconductor electronics involves several orders of magnitude in length and timescales. Carrier scattering that dissipates Joule heat occurs at ~10 nm length scales over 0.1 ps timescales. Heat diffusion out of the transistor occurs over ~100 ns. Interconnects also experience severe heating. At the scale of functional blocks of the integrated circuit, heat production is highly nonuniform. Ultimate removal occurs through the package.

Fundamental Heat Conduction Physics

The Fourier law for heat conduction (Incropera, 2006) in a material medium is a phenomenological equation that relates heat flux to the temperature gradient in the material,

$$q = -\ddot{\kappa}\nabla T, \tag{4.21.1}$$

where $\boldsymbol{q}$ is the heat flux, ∇T is the temperature gradient, and $\ddot{\kappa}$ is the thermal conductivity tensor of the medium. For an isotropic medium, a scalar thermal conductivity κ appears in the equation above. The value of κ can be further interpreted by considering the microscopic mechanisms responsible for energy transport that depend on the medium. In gases, for instance, molecular motion transports heat. The kinetic theory provides a simple relation between thermal conductivity and molecular motion as

$$\kappa = \frac{1}{3} C\bar{v}\Lambda, \tag{4.21.2}$$

where C is the volumetric heat capacity, $\bar{v}$ is the mean molecular speed, and Λ is the mean free path of gas molecules between collisions. Even in solids where the microscopic mechanism for heat conduction arises in the transport of energy by electrons and quantized lattice vibrations (or phonons), the above expression can be reinterpreted as the conductivity of a gas of electrons or phonons.

Further considering conservation of energy in an infinitesimal volume of the material leads to the heat diffusion equation. In an isotropic medium without any heat source,

$$\frac{\partial T}{\partial t} = \alpha_T \nabla^2 T, \tag{4.21.3}$$

where α_T is the thermal diffusivity of the medium and is defined as the ratio of the thermal conductivity to the heat capacity per unit volume. The heat diffusion equation above assumes the presence of continuum. Continuum refers to a description of matter in which it is infinitely divisible. Provided material properties are known, the above equation can be readily solved in complicated geometries to obtain spatial and temporal temperature distributions.

In electronic circuits, however, the relevant length scales are often in micrometers and nanometers where the continuum description of matter may break down. It then becomes necessary to consider the physical process of heat conduction in more detail. Internal energy is transported in solids via energy carriers, typically electrons and phonons (Ziman, 1960; Ashcroft and Mermin, 1976). Both electrons and phonons are quantum particles. The term phonon refers to the energy quantum of the displacement field of a crystal lattice. Electrons dominate heat transport in metals and phonons in dielectrics. The kinetic theory of matter provides a framework for considering energy transport by carriers such as electrons and phonons.

The theory aims to describe the motion of carriers through a statistical distribution function in three-dimensional physical space, three-dimensional velocity (or momentum) space and time. Consider a system of gas particles that may represent either electrons or phonons in a solid. A six-dimensional infinitesimal volume is formed by the volume in physical space between position vectors $\boldsymbol{r}$ and $\boldsymbol{r}+\mathrm{d}\boldsymbol{r}$ as well as the volume in velocity space between $\boldsymbol{v}$ and $\boldsymbol{v}+\mathrm{d}\boldsymbol{v}$. The distribution function f provides the number of particles in the volume through

$$\mathrm{d}N = f\left(\boldsymbol{r},\boldsymbol{v},t\right)\mathrm{d}r\mathrm{d}v. \tag{4.21.4}$$

The distribution function is not stationary but changes due to drift of the particles in physical space as well change in their velocities due to external forces and collisions (or scattering). Detailed consideration of different mechanisms for collisions is often necessary to obtain a reliable answer. Balancing particle numbers leads to an equation for the distribution function called the Boltzmann transport equation (BTE) that is the cornerstone of modeling transport. The BTE written for a particle under an external force is

$$\frac{\partial f}{\partial t} + v\cdot\nabla f + a\cdot\nabla_v f = \left.\frac{\partial f}{\partial t}\right]_{\mathrm{SCATTER}}, \tag{4.21.5}$$

where $\boldsymbol{a}$ is the acceleration of the particle under any external force, and the term on the right-hand side is the net rate of change in f due to scattering. In principle, the solution of the distribution function can provide the heat flux in the medium. At any location, this requires integrating over particles of all velocities as

$$q\left(\boldsymbol{r},t\right) = \int \boldsymbol{v} f\left(\boldsymbol{r},\boldsymbol{v},t\right) E\left(\boldsymbol{v}\right)\mathrm{d}\boldsymbol{v}, \tag{4.21.6}$$

where $E(\boldsymbol{v})$ is the energy of the particle as a function of its velocity. At thermodynamic equilibrium, the net transport of all quantities such as particle flux, mass, momentum, and energy must be zero at all locations and times. The corresponding distribution function, f_0, depends on the type of particle. For electrons, the equilibrium distribution is the Fermi–Dirac (FD) distribution, whereas for phonons, the equilibrium distribution is the Bose–Einstein (BE) distribution. The equilibrium distribution depends on the temperature of the medium, the chemical potential, and the energy of the particle. Figure 4.21.5 plots the distribution functions for FD electrons and BE phonons. Both distributions converge to the classical Maxwell–Boltzmann distribution at high energies.

The scattering term in Equation 4.21.5 is complicated in general. The full solution to the BTE is both computationally expensive and fairly complex and has only been attempted in limited cases.

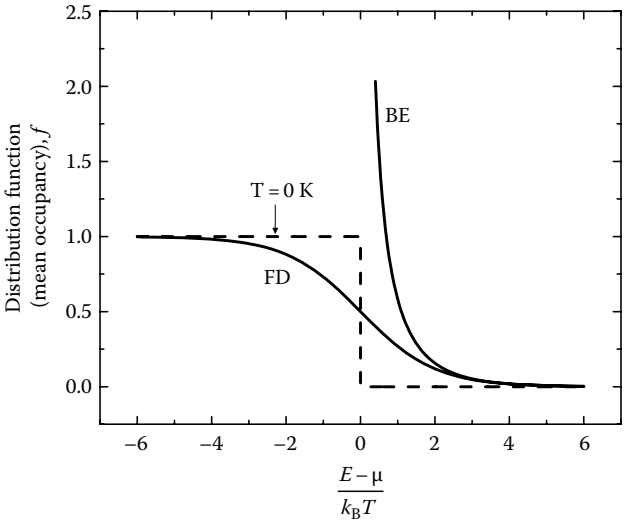


FIGURE 4.21.5 The distribution functions for FD and BE gases. The dashed line shows the FD distribution at 0 K. The occupation is a function of the energy of the particle in relation to the thermal energy, $k_B T$. (Figure prepared by Dhruv Gelda.)

It is difficult to proceed with the BTE without further simplification of the scattering term. The simplest approach is termed the relaxation time approximation. The scattering term on the right-hand side of Equation 4.21.5 is simplified to

$$\left.\frac{\partial f}{\partial t}\right]_{\text{SCATTER}} = -\frac{f - f_o}{\tau}, \tag{4.21.7}$$

where f_0 is the distribution function at thermal equilibrium and τ is a relaxation time. Here the intention is to capture the basic relaxation of the system toward a thermal equilibrium and is especially useful in calculating transport coefficients.

As an example, consider the problem of calculating the thermal conductivity of a medium where external forces are absent and a temperature gradient is present. Since the timescale of interest is much larger than the relaxation time of carriers (typically of the order of picoseconds to nanoseconds at room temperature), steady state is assumed. Using Equations 4.21.7 and 4.21.5 simplifies to

$$v \cdot \nabla f = -\frac{f - f_0}{\tau}. \tag{4.21.8}$$

Typically, the gradient of temperature is small enough that temperature varies very little over the length scale $v\tau$. Therefore, we expect f to be only slightly different from the local f_0 and the difference to be linear in the perturbing gradient. Discarding terms nonlinear in the perturbing gradient renders the equation explicit. This simplification is called "linearizing the BTE." Assuming a temperature gradient in the x-direction, the linearized BTE under the relaxation time approximation reduces to

$$v_x \frac{\partial f_0}{\partial x} = -\frac{f - f_0}{\tau}, \tag{4.21.9}$$

and yields $f = f_o - \tau v_x \frac{\partial f_o}{\partial x}$. Since f_0 is a function of temperature and particle energy, its spatial gradient arises only from any gradient in temperature and thus, $\frac{\partial f_o}{\partial x} = \frac{\partial f_o}{\partial T}\frac{\partial T}{\partial x}$. Using Equation 4.21.6 to describe the heat flux along x,

$$q_x(x) = \int v_x f_o(T(x), v) E(v)\, d\boldsymbol{v} - \int v_x^2 \tau \frac{\partial f_o}{\partial T}\frac{\partial T}{\partial x} f_o E(v)\, d\boldsymbol{v}, \tag{4.21.10}$$

where the first integral vanishes since f_0 is isotropic with respect to velocity. The temperature gradient inside the second integral can be taken out since it is independent of velocity. Comparing the expression with the Fourier law of Equation 4.21.1 yields the thermal conductivity of the medium as

$$\kappa = \int v_x^2 \tau \frac{\partial f_o}{\partial T} f_o E(v)\, d\boldsymbol{v}. \tag{4.21.11}$$

Evaluation of the integral above requires knowledge of two quantities: the variation of the energy E with v and the scattering time τ. These depend on the type of carrier, whether electron or phonon, and are also specific to the material, rendering κ a material property.

Since both electrons and phonons are quantum particles, they can also be described as waves. The wave dispersion relation provides the variation of energy with particle velocity or equivalently, particle momentum. In the case of electrons, the relation is commonly stated between E and the wave number k instead of the velocity or momentum. The momentum of a free electron is $\hbar k$ from the de Broglie relation, where $\hbar$ is the reduced Planck constant. In the case of phonons, the relation is stated between the frequency ω and k. The energy of a phonon is related to its frequency through $E = \hbar\omega$. Obtaining accurate dispersion relation for electrons or phonons requires detailed considerations of these particles. The scattering time is even more involved with multiple scattering mechanisms at play. A common approach is to consider individual scattering mechanisms in as much detail as possible, while considering each to be completely independent of others. The overall scattering time is then given by Matthiessen's rule, which is $\tau^{-1} = \sum \tau_{\text{individual}}^{-1}$. Prominent mechanisms for scattering depend on the type of particle. Electrons prominently scatter with impurity atoms, crystal imperfections, ionic impurities, phonons, plasmons, and surfaces. Similarly, phonons prominently scatter with other phonons, electrons, surfaces, impurities, and crystal imperfections.

Properties of Electrons in Metals

Most properties of electrons in metals require consideration of quantum mechanics. The simplest Sommerfeld theory mainly includes the Pauli exclusion principle leading to the use of the FD distribution instead of the classical Maxwell–Boltzmann distribution. Postulating electrons in the metal to be quantum particles in a box without any forces and to be independent of each other, the Schrödinger wave equation describes such one electron states as plane waves with momentum $\hbar k$ and energy $\frac{\hbar^2 k^2}{2m_e}$, where m_e is the electron rest mass and k is the wave number. The wave numbers are discrete in each direction of wave vector space, for example, $k_x = n_x 2\pi/L_x$, where n_x is an integer and L_x is the size of the crystal along the x-direction. Therefore, each allowed state or discrete k-point in wave vector space has a volume $(2\pi)^3/(L_x L_y L_z)$ or $(2\pi)^3/V$, where V is the physical volume of the crystal.

The properties of the free electron gas in the metal are an average of the value across all states weighed by the probability of occupation of each state. Since states are closely spaced and many, it is more convenient to convert the summation over states to an integral. Further since most properties of interest are independent of the direction in wave vector space and instead depend only on the energy associated with the state, it is useful to average over energy instead. Converting the integral

over k to an integral over E requires knowledge of the density of states, $D(E)$, which is a function of energy describing the number of states between energy E and $E+dE$ per unit physical volume of the crystal and per unit energy interval dE. The density of states for free electrons in three dimensions is given by

$$D(E) = \frac{m_e}{\hbar^2 \pi^2} \sqrt{\frac{2m_e E}{\hbar^2}}. \tag{4.21.12}$$

The density of states function changes with confinement of the electron to fewer dimensions.

The probability of occupation of a state depends on its energy and is given by the FD distribution function as

$$f(E) = \frac{1}{1 + e^{\frac{E - E_F}{k_B T}}}, \tag{4.21.13}$$

where E_F is the Fermi energy and k_B is the Boltzmann constant. At 0 K, electrons occupy all states up to the Fermi energy and leave all states above E_F unoccupied. The Fermi energy is a material property providing a measure of the average energy of the electron (0.6 E_F) and ranges between 5 and 10 eV for most metals.

The thermal properties of the electron gas can be readily calculated using the above framework. The internal energy per unit volume, u_E, for instance, is given by

$$u_E = \int_0^\infty E f(E) D(E) \mathrm{d}E \cong u_E(T=0) + \frac{\pi^2}{6} (k_B T)^2 D(E_F), \tag{4.21.14}$$

where the integral is evaluated using the Sommerfeld expansion. The heat capacity of electrons per unit volume is then

$$C_e = \left. \frac{\partial u}{\partial T} \right|_V = \frac{\pi^2}{3} (k_B)^2 D(E_F) T = \gamma T, \tag{4.21.15}$$

where the constant γ is a material property. Except at low temperatures ($\lesssim$10 K), the heat capacity of electrons contributes little to the overall heat capacity of solids. However, the heat capacity is important in evaluating the thermal conductivity of electrons using the expression in Equation 4.21.2. For electrons in metals, the mean speed is the Fermi speed, v_F that is of the order of 10^6 m/s in most metals. Table 4.21.1 lists the electronic properties of some common metals. A convenient way to estimate the thermal conductivity of metals is to take advantage of the Wiedemann–Franz Law that relates their thermal conductivity to their electrical conductivity as

$$\kappa = L_o \sigma T, \tag{4.21.16}$$

where L_o is a proportionality constant called the Lorenz number and is theoretically evaluated to be 2.44×10^{-8} WΩK^{-2}, and σ is the electrical conductivity.

Properties of Electrons in Semiconductors

Semiconductors are distinct from metals in possessing a gap in the allowed energy bands of electrons. Carriers can be thermally excited across the bandgap such that the electrical conductivity changes with temperature. Semiconductors can be categorized into intrinsic and extrinsic.

TABLE 4.21.1
Properties of Electrons in Metals

Metal	Electron Concentration (10^{28}/m^3)	Fermi Energy (10^{-19} J)	Fermi Velocity (10^6 m/s)	Electrical Conductivity (10^5 /Ωm)	Thermal Conductivity at 300 K (W/mK)	Electron Mean Free Path at 300 K (nm)
Rb	1.15	2.96	0.81	0.80	58	16
Ag	5.85	8.78	1.39	6.21	429	49
Au	5.90	8.83	1.39	4.55	317	32
Cu	8.45	11.2	1.57	5.88	401	26
In	11.49	13.8	1.74	1.14	82	1.5
Pb	13.20	15.0	1.82	0.48	35	0.3
Al	18.06	18.6	2.02	3.65	237	2.9

Source: Frank et al., Applications. In *CRC Handbook of Thermal Engineering*, CRC Press, 1999.

An intrinsic semiconductor is chemically pure and only possesses thermally excited charge carriers. The electrical conductivity at room temperature can be similar to that of an insulator. In an extrinsic semiconductor, a small amount of impurity atoms are added in a process known as doping to drastically improve the electrical conductivity. Based on carrier polarity, extrinsic semiconductors can be *p*-type or *n*-type, where holes and electrons are, respectively, the majority charge carriers. The activation energy for electrons to be donated by or accepted to impurity states is usually low enough that most remain activated at room temperature. The concentration of majority charge carriers is almost identical to the concentration of impurities. Electrical conductivity can be made to vary several orders of magnitude depending on the doping concentration.

The Sondheimer model does not extend well to semiconductors. A more detailed band structure must be taken into consideration to properly account for the bandgap. Further, charge carriers contributed by foreign dopants shift the Fermi level. The effect of the periodic potential imposed by the lattice ions on conduction electrons is important. The quantum mechanical behavior of electrons in a periodic potential can be addressed through an effective mass, m^* that is different from the rest mass. The effective mass is a property of the crystal and can be anisotropic and energy dependent (and even negative). Charge transport usually involves carriers at the bottom of the conduction band and at the top of the valence band. The effective mass of the carrier depends on the curvature of the E–k relationship as $m^* = \dfrac{\hbar^2}{d^2E/dk^2}$.

Charge transport can be analyzed using the linearized BTE under the relaxation time approximation as

$$\mathbf{v} \bullet \left[\nabla f_0 + q\varepsilon \frac{\partial f_0}{\partial E} \right] = \frac{f_0 - f}{\tau}, \tag{4.21.17}$$

where $\boldsymbol{v}$ is electron velocity, f_0 and f are the equilibrium and the nonequilibrium distribution, respectively, ε is the electrical field, and τ is relaxation time. If a temperature gradient is also present, the new distribution function becomes

$$f = f_0 - \tau \mathbf{v} \bullet \left[q\varepsilon - \frac{E - E_F}{T} \nabla T \right] \frac{\partial f_0}{\partial E}. \tag{4.21.18}$$

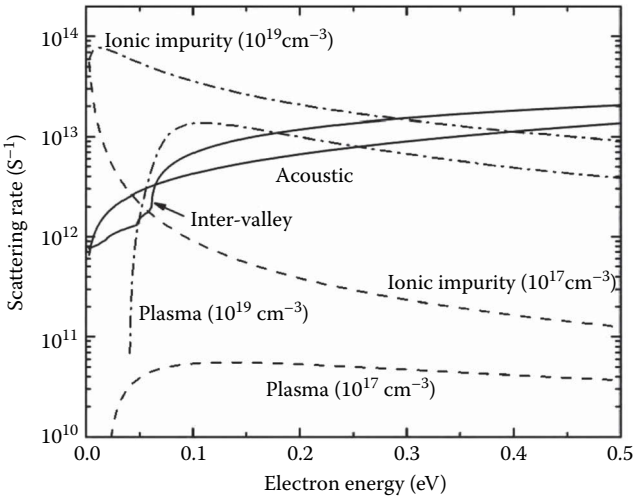


FIGURE 4.21.6 Scattering rates of electrons in silicon at 300 K as a function of electron energy. The solid lines represent scattering mechanisms insensitive to doping. Dashed and dot–dashed lines represent doping levels of 10^{17} cm^{-3} and 10^{19} cm^{-3}, respectively. (Adapted from Ma, J. and Sinha, S., *J. Appl. Phys.*, 112, 073719, 2012. With permission. Copyright 2012 American Institute of Physics.)

The current density along the direction of the electric field is

$$J_Z = \int_{\varphi=0}^{2\pi} \frac{1}{4\pi}\, \mathrm{d}\varphi \int_{\theta=0}^{\pi} \sin\theta\cos\theta\, \mathrm{d}\theta \int_{E=0}^{\infty} fD(E)qv\, \mathrm{d}E, \tag{4.21.19}$$

where θ and φ are the polar and the azimuthal angles in spherical coordinates. The electrical conductivity is

$$\sigma \equiv \frac{J_Z}{\varepsilon_z} = -\frac{2q^2}{3m_C^*} \int_{E=0}^{\infty} \frac{\partial f_0}{\partial E} D(E)E\tau\, \mathrm{d}E, \tag{4.21.20}$$

where m_C^* is the conductivity effective mass and $D(E)$ is electron density of states. The relaxation time τ can be calculated assuming Matthiessen's rule. Key scattering mechanisms include electron–phonon (acoustic and intervalley), ionic impurity, and plasmon scattering (Fischetti, 1991). Figure 4.21.6 plots key scattering rates for silicon to illustrate their relative contribution at 300 K. Two different doping concentrations are considered for impurity scattering.

Properties of Phonons in Crystals

Atoms located at the lattice sites of a crystal undergo small oscillations about their equilibrium positions at all temperatures. The resulting atomic displacement field in the crystal stores and transports energy. Except in metals, they are the dominant contributors to thermal transport in solids. Quantum mechanically, the atomic displacement field may be described either as an infinite number of distinguishable, quantized oscillators, or as a gas of indistinguishable particles called phonons. Consider an atom with position vector $\boldsymbol{R}$, displaced from its equilibrium position $\boldsymbol{R}_0$ in the lattice by $\boldsymbol{u}$, where $\boldsymbol{R}=\boldsymbol{R}_0+\boldsymbol{u}$. The amplitude of motion is small, on the order of 10^{-3} nm compared to the interatomic spacing that is a few angstroms. The potential energy of the crystal can be expanded in a Taylor series in terms of the small displacements as

$$\Phi(\boldsymbol{R}) = \Phi(\boldsymbol{R}_o) + \sum_{\beta}\left[\left.\frac{\partial\Phi}{\partial u_\beta}\right|_{\boldsymbol{R}_o} u_\beta\right] + \frac{1}{2}\sum_{\beta,\beta'}\left[\left.\frac{\partial^2\Phi}{\partial u_\beta\,\partial u_{\beta'}}\right|_{\boldsymbol{R}_o} u_\beta u_{\beta'}\right] + \cdots, \tag{4.21.21}$$

where the subscript β is an index over the directions (x, y, or z). The first term in the expansion is constant. The derivative appearing in the second term represents net forces and must vanish when evaluated at the equilibrium position, $\boldsymbol{R}_o$. The derivative in the third term is a harmonic force constant assuming restoring forces are linear to the displacement. The series is typically truncated at this term assuming harmonic springs and the force constants can be used to obtain vibrational modes of the crystal. The equation of motion of an atom in the lattice along the direction β is

$$m\frac{d^2u_\beta}{dt^2} = -\frac{\partial\Phi}{\partial u_\beta} = -\frac{\partial}{\partial u_\beta}\left[\Phi(\boldsymbol{R}_o) + \frac{1}{2}\sum_{\beta,\beta'}\left[\left.\frac{\partial^2\Phi}{\partial u_\beta\,\partial u_{\beta'}}\right|_{\boldsymbol{R}_o} u_\beta u_{\beta'}\right]\right] = -\sum_{\beta'}\left.\frac{\partial^2\Phi}{\partial u_\beta\,\partial u_{\beta'}}\right|_{\boldsymbol{R}_o} u_{\beta'}, \tag{4.21.22}$$

where m is the mass of the atom.

To illustrate the idea, consider the simple case of motion along one direction (longitudinal) in a linear chain of atoms. The equation of motion for the pth atom in the chain can be written as

$$m\frac{\mathrm{d}^2u_p}{\mathrm{d}t^2} = \sum_j K_j\left(u_{p+j} - u_p\right), \tag{4.21.23}$$

where K_j is the force constant between the pth and the $(p+j)$th atom in the chain. The above equation takes plane waves as solutions of the form $u_p = Ae^{i(pka-\omega t)}$, where A, k, and ω are the amplitude, the wave number, and the frequency of the vibrational wave, respectively, and a is the interatomic spacing (or lattice constant). Inserting the solution into Equation 4.21.23 and solving yields the wave dispersion relation between frequency and wave number. Interatomic forces typically weaken beyond the first few neighbors. Assuming only the first neighboring atoms to have a nonzero force constant, the dispersion relation simplifies to

$$\omega = \sqrt{\frac{\sum_{j>0}\left[2K_j\left(1-\cos(jka)\right)\right]}{m}} = \sqrt{\frac{\left[2K_l\left(1-\cos(ka)\right)\right]}{m}} = \sqrt{\frac{4K_l}{m}}\left|\sin\left(\frac{ka}{2}\right)\right|, \tag{4.21.24}$$

where K_l is the longitudinal force constant between nearest neighbors. The above relation points out some of the prominent features of the phonon dispersion relation: Stiffer springs and lighter masses lead to higher frequency modes. The dispersion is nonlinear at large k's but linear in the limit of small k's. The speed at which a phonon of frequency ω transports energy is its group speed, given by the derivative of ω with respect to k,

$$v_{\mathrm{ph}}(k) = \left.\frac{\partial\omega}{\partial k}\right|_k. \tag{4.21.25}$$

In a three-dimensional crystal, there are several possible dispersion relations, corresponding to different types of motion. The lowest frequencies are associated with acoustic phonons that in the limit of long wavelength, travel at the speed of sound in the crystal. Acoustic phonons may be longitudinally or transversely polarized. A crystal lattice may have more than one atom at each lattice point. In crystals with such polyatomic bases, optical phonons exist in addition to acoustic phonons. In the limit of long wavelength, an optical phonon mode has atoms in the basis moving in opposing directions. This gives them a polar character and facilitates interaction with light. However,

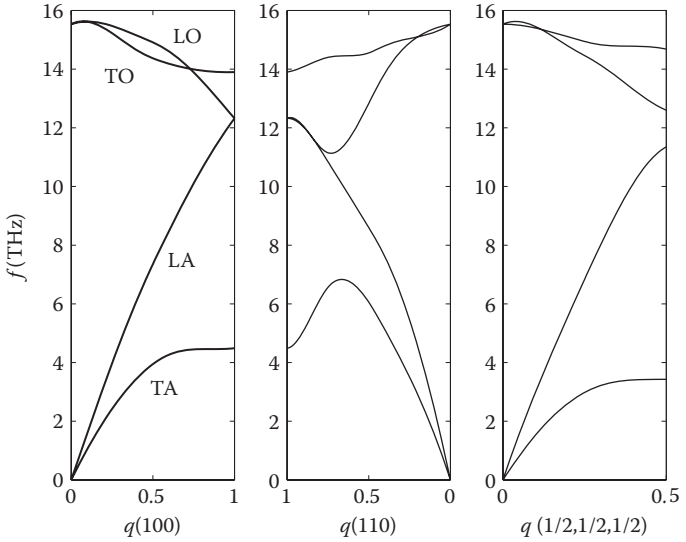


FIGURE 4.21.7 Realistic phonon dispersion relation for silicon along different crystallographic directions of high symmetry. The branches are longitudinal optical (LO), transverse optical (TO), longitudinal acoustic (LA), and transverse acoustic (TA).

optical phonons are higher in frequency and are not sufficiently excited at room temperature to dominate thermal transport. Their contribution, though small at room temperature, may still need to be considered for accurate predictions of thermal conductivity in certain crystals. Figure 4.21.7 plots the dispersion curves for phonons in silicon. The relation is different along different directions in the crystal.

The above description of the modes of vibration of a crystalline lattice is completely classical. To take into consideration quantum behavior, it is further necessary to treat the modes as independent quantum harmonic oscillators. For a three-dimensional lattice with N atoms, there are $3N$ possible modes corresponding to the degrees of freedom. The contribution of a particular mode (k, $\omega_s(\boldsymbol{k})$) to the total energy can only be discrete values of $\left(\overline{n}_{\boldsymbol{k},s}+1/2\right)\hbar\omega_s\left(\boldsymbol{k}\right)$, where s is an index for the phonon branch (whether acoustic or optical and whether longitudinal or transverse), $\overline{n}_{\boldsymbol{k},s}$ is the occupation of the mode at thermodynamic equilibrium given by the BE distribution. The BE distribution function for phonons is the same as the Planck distribution,

$$\overline{n}_{\boldsymbol{k},s}=\frac{1}{1-e^{\frac{\hbar\omega_s(\boldsymbol{k})}{k_BT}}}. \tag{4.21.26}$$

The internal energy density is then

$$u_{\text{ph}}=\frac{1}{V}\sum_{\boldsymbol{k},s}\hbar\omega\left(\overline{n}_{\boldsymbol{k},s}+\frac{1}{2}\right)+u_{\text{EQUILIBRIUM}}, \tag{4.21.27}$$

with the heat capacity of phonons per unit volume of the crystal being

$$C_{\text{ph}}=\left.\frac{\partial u_{\text{ph}}}{\partial T}\right|_{\text{V}}=\frac{1}{V}\frac{\partial}{\partial T}\sum_{\boldsymbol{k},s}\left[\frac{\hbar\omega_s\left(\boldsymbol{k}\right)}{1-e^{\frac{\hbar\omega_s(\boldsymbol{k})}{k_BT}}}\right]=\frac{\partial}{\partial T}\sum_{s}\int\frac{\mathrm{d}\boldsymbol{k}}{\left(2\pi\right)^3}\frac{\hbar\omega_s\left(\boldsymbol{k}\right)}{1-e^{\frac{\hbar\omega_s(\boldsymbol{k})}{k_BT}}}, \tag{4.21.28}$$

where the conversion from the sum to the integral relies on the fact that k points are closely spaced in wave vector space and that the volume per k-point is $V/(2\pi)^3$. Evaluation of the integral in Equation 4.21.28 correctly reproduces the heat capacity of solids as a function of temperature, with the classical Dulong–Petit law emerging in the limit of high temperatures. Figure 4.21.8 plots the heat capacity of phonons in platinum as a function of temperature. The heat capacity of electrons in the same metal is also plotted for the sake of comparison. At low temperatures, the phonon heat capacity varies as $\sim T^3$, a fact that is often used to interpret the thermal conductivity of novel materials.

Since the integrand in Equation 4.21.24 is a function of frequency, it is useful to integrate over frequency rather than $\boldsymbol{k}$. Similar to electrons, the conversion requires the use of the density of states. For phonons this is defined as the function $D(\omega)$ and counts the states between frequencies ω and $\omega+d\omega$ per unit volume and per unit frequency interval. For historical reasons, a simplified dispersion model is often used to evaluate the heat capacity. The Debye model of phonon dispersion assumes a linear, isotropic dispersion for acoustic phonons similar to photons in vacuum. All three phonon branches, corresponding to one longitudinal and two transverse polarizations, have the same dispersion relation, $\omega = c_D k$, where c_D is the Debye speed. However, unlike photons, the total number of phonon modes in the crystal is limited to $3N$ as discussed previously. To reflect this, a cutoff wave number, k_D is chosen such that the total number of states within a sphere of radius k_D is the total number of modes, $3N$. Despite the simplifications, the Debye model is surprisingly accurate in predicting the heat capacity of many crystals. The cutoff frequency corresponding to the cutoff wave number is called the Debye frequency, ω_D. The Debye temperature, θ_D is proportional to the Debye frequency and is related as $\theta_D = \hbar\omega_D / k_B$. At temperatures comparable to or below the Debye temperature, the quantum behavior of phonons is important. At higher temperatures, phonons behave classically as the normal modes of vibrations of a crystal lattice. The phonon density of states in three-dimensions, assuming Debye dispersion is useful in evaluating the integral in Equation 4.21.28 and is given by

$$D(\omega) = \frac{3}{2\pi^2}\frac{\omega^2}{c_D^3}, \omega \le \omega_D . \tag{4.21.29}$$

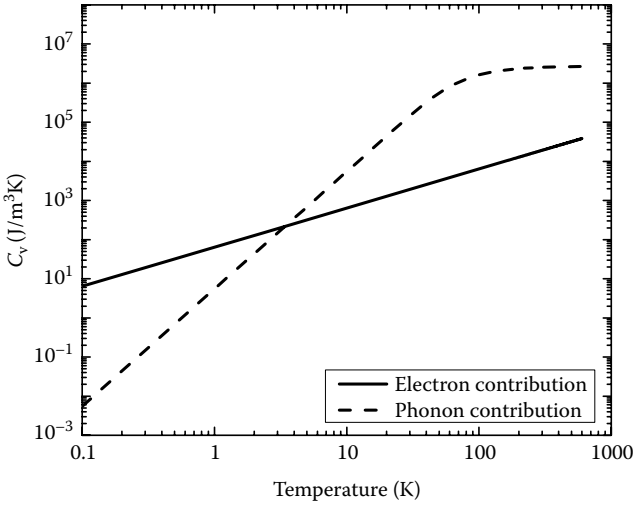


FIGURE 4.21.8 Comparison of the contribution of electrons (solid line) and phonons (dashed line) to the volumetric heat capacity of a solid (metal) as a function of temperature. The electronic contribution is significant only at low temperatures ($\lesssim$10 K). (Figure prepared by Dhruv Gelda.)

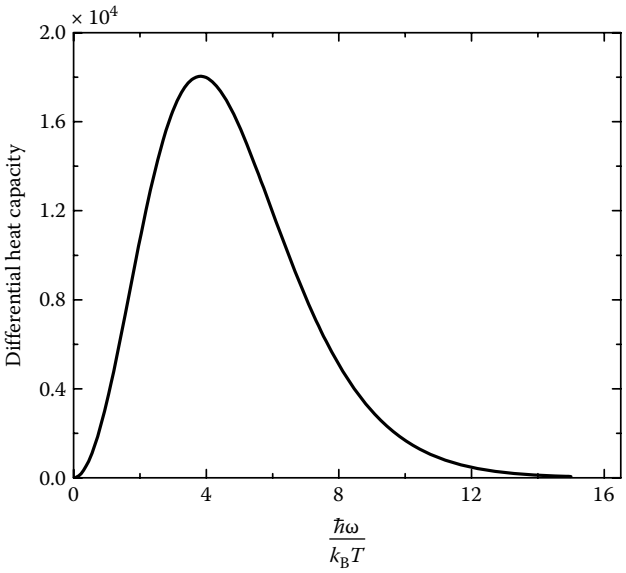


FIGURE 4.21.9 Contribution of different frequencies to the heat capacity of silicon as per the Debye model. The peak occurs close to 4 $k_B T$. The heat capacity is the area under the curve up to the Debye frequency.

Figure 4.21.9 plots the spectral contribution to the heat capacity using the Debye model. Table 4.21.2 lists properties related to phonons for some common solids.

Theory of Interfacial Thermal Transport

Thermal transport across interfaces holds special significance in thermal transport involving electronic micro- and nanostructures. With ever-increasing materials integration, modern integrated circuits are brimming with material interfaces that have significant impact on heat flow. The theory of interfacial thermal transport offers qualitative insight but is usually not quantitatively accurate. Quantitative theories invariably involve atomistic models that remain unwieldy for practical purposes. Generally, good experimental data on thermal transport across interfaces are an important part of device and circuit thermal design. Thermal interfaces fall in two broad categories: Atomically jointed interfaces where a chemical bond exists at the interface and mechanically jointed interfaces

TABLE 4.21.2
Properties of Phonons in Solids

Solid	Atomic Density (10^{28} m^{-3})	Mass Density (10^3kg/m3)	Longitudinal Phonon Velocity (km/s)	Transverse Phonon Velocity (km/s)	Debye Temperature Θ_D (K)
Si	5.00	2.33	8.97	5.33	645
C (diamond)	17.6	3.51	17.5	12.8	2230
GaAs	6.62	2.66	6.09	4.10	492
Al	6.02	2.70	6.24	3.04	428
Cu	8.45	8.93	4.91	2.50	343
Au	5.90	19.3	3.39	1.29	165
Ag	5.85	10.5	3.78	1.74	225

Source: Frank et al., Applications. In *CRC Handbook of Thermal Engineering*, CRC Press, 1999.

where the materials are forced together by mechanical forces. Microscopic theory is necessary to understand the first, whereas continuum heat conduction and contact mechanics provide insight into the second. The focus of this chapter is on the microscopic theory of interfacial thermal transport.

When heat flows across an atomically perfect interface between two solids, a temperature jump ΔT arises at the interface corresponding to a thermal resistance at the interface. Initially discovered by Kapitza (Swartz and Pohl, 1989) while investigating the superfluidity of helium at low temperatures, the interfacial resistance between liquid helium and a solid (and often between two dielectric/semiconducting solids) is also called the Kapitza resistance. The theoretical picture of interfacial thermal transport involves the net radiation of energy across the interface via the crossing of energetic particles. For a more detailed picture, it is necessary to consider the dominant energy carriers on the two sides of the interface. Figure 4.21.10 illustrates the two broad classes of interfaces. In an interface between two metals, thermal transport involves the transmission of electrons across the interface. In an interface between a metal and a dielectric, energy must be transferred from the electron gas in the metal to the phonon gases in the metal and the dielectric. For an interface between dielectrics/semiconductors, only phonon transport across the interface needs to be considered.

Consider a dielectric–dielectric interface with a differential temperature drop δT across it. Ignoring any electron–phonon and inelastic scattering processes, the problem reduces to the net radiation of phonons across the interface. The net heat current from side 1 to 2 is the difference between the currents crossing the interface from either side and can be written as

$$q_{\text{net}} = \frac{1}{2}\left[\begin{array}{c} \sum_s \int \frac{d\boldsymbol{k}_1}{(2\pi)^3} \left|\boldsymbol{v}(\omega,\boldsymbol{k}_1)\cdot\hat{\boldsymbol{n}}_1\right| \hbar\omega(\boldsymbol{k}_1)\alpha_{12}(\omega,\boldsymbol{k}_1)\left[\bar{n}(\omega, T+\delta T)\right]- \\ \sum_s \int \frac{d\boldsymbol{k}_2}{(2\pi)^3} \left|\boldsymbol{v}(\omega,\boldsymbol{k}_2)\cdot\hat{\boldsymbol{n}}_2\right| \hbar\omega(\boldsymbol{k}_2)\alpha_{21}(\omega,\boldsymbol{k}_2)\left[\bar{n}(\omega, T)\right] \end{array}\right], \tag{4.21.30}$$

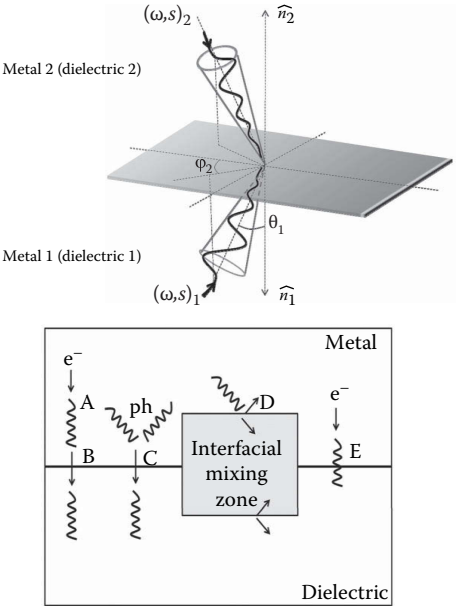


FIGURE 4.21.10 The geometry for considering thermal interface conductance between two metals or dielectrics (or semiconductor) is shown on top. The consideration for metal–dielectric (or semiconductor) is more complex involving several possible channels for energy transfer across the interface. (Adapted from Singh, P. et al., *Appl. Phys. Lett.*, 102, 181906, 2013. With permission.)

where $\hat{\boldsymbol{n}}_i$ is the unit normal at the interface into side i, α_{ij} is the probability that a carrier incident from side i crosses into j, and the factor 1/2 is included to reflect the fact that only half the phonons near the interface on either side are incident toward the interface. The integration is over all possible $\boldsymbol{k}$ on each side of the material. Note that the two integrals do not have the same k space for integration.

Each of the integrands can be written out in spherical coordinates and the integration can be performed over all frequencies and solids angles. The geometry is shown in Figure 4.21.10. For example, the current at the interface from side 1 can be written as

$$q_1 = \sum_s \int_{\omega=0}^{\omega_{\max,1}} \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi/2} \alpha_{12} \frac{1}{4\pi} \hbar\omega v_1 \cos\theta D_1(\omega) \bar{n}(\omega, T+\delta T) \sin\theta \, d\theta \, d\varphi \, d\omega, \tag{4.21.31}$$

where α and v are, in general, functions of $(\omega, \theta, \varphi)$. The two transmission coefficients α_{12} and α_{21} are related as follows. At thermal equilibrium, the net current must be zero. The two integrals in Equation 4.21.30 can be merged into one by performing the integration to the higher of the two cutoff frequencies and explicitly noting that the density of states function in the material with the smaller cutoff frequency will be zero past the cutoff frequency. For the new integral to vanish, the condition is

$$\alpha_{12}(\omega,\theta,\varphi) v_1(\omega,\theta,\varphi) D_1(\omega) = \alpha_{21}(\omega,\theta,\varphi) v_2(\omega,\theta,\varphi) D_2(\omega), \tag{4.21.32}$$

where the subscripts for v and D indicate that they are different functions for the two sides. Thus, knowing the function α from either side is sufficient to calculate the integral in Equation 4.21.30.

To obtain an expression for the interfacial conductance, consider the limit of small δT to Taylor expand the function $\bar{n}(\omega, T+\delta T)$. From Equations 4.21.30 through 4.21.32, the interfacial thermal conductance is

$$G = \sum_s \int_{\omega=0}^{\omega_{\max,1}} \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi/2} \alpha_{12} \frac{1}{4\pi} \hbar\omega v_1 \cos\theta D_1(\omega) \frac{\partial \bar{n}(\omega, T)}{\partial T} \sin\theta \, d\theta \, d\varphi \, d\omega, \tag{4.21.33}$$

where we have assumed that side 1 has the higher cutoff frequency. If side 2 has the higher cutoff frequency, the integration limit for frequency in Equation 4.21.33 would be $\omega_{\max,2}$, with the functions v_1 and D_1 set to zero for frequencies between the two cutoff frequencies. Equivalently, the calculation can proceed from side 2 if α_{21} is known instead of α_{12}, again with the integration limit for frequency set to the higher of the two cutoff frequencies.

The transmission function can be fairly complicated in reality since it not only depends on the bulk phonon properties of the two materials but also on the chemical (for example, bonding) and physical (for example, roughness or dislocations) state of the interface. Extant models only consider ideal interfaces and make further simplifying assumptions. The two popular models are the acoustic mismatch model (AMM) and the diffuse mismatch model (DMM). Neither is quantitatively reliable when compared against experimental data. Their utility lies more in providing physical insight into temperature dependence, for instance. Due to their simplicity, they continue to be popular in the literature.

The AMM (Khalatnikov, 1952) theorizes that the resistance arises due to the difference in acoustic properties of two media. At low temperatures, only long wavelength phonons (corresponding to sound waves in the crystal) are excited. Due to differences in acoustic speeds, phonon momentum is different in the two media. (In the limit $k \to 0$, the phonon momentum is $\hbar k$). In general, it is

difficult to satisfy both momentum and energy conservation for a phonon traversing the interface, leading to a resistance. In terms of acoustics, this can be thought of as an acoustic impedance mismatch at the interface. Khalatnikov's original formulation considers the interface between liquid He and copper but the formulation can be revised to consider the more complicated interface of two elastic solids. In either case, it is possible to show that the thermal conductance at the interface is finite and proportional to T^3. The transmission coefficient for longitudinal phonons at the interface between elastic solids is

$$\alpha_{12} = \frac{4\dfrac{\rho_2 c_{l2}}{\rho_1 c_{l1}}\dfrac{\cos\theta_2}{\cos\theta_1}}{\left(\dfrac{\rho_2 c_{l2}}{\rho_1 c_{l1}} + \dfrac{\cos\theta_2}{\cos\theta_1}\right)^2}, \tag{4.21.34}$$

where ρ is the mass density and c_l is the longitudinal speed of sound in the medium. The function depends on the angle θ. The angle on the other side of the interface may be eliminated from the expression during calculations by using the acoustic Snell's law for refraction, $c_{l2}\sin\Theta_1 = c_{l1}\sin\Theta_2$. While the theory explains the qualitative behavior of the interfacial resistance versus temperature, it is not quantitatively reliable. Several modifications have been suggested over the decades since Khalatnikov's work but a quantitatively accurate general-purpose theory at low temperatures is still lacking.

The DMM (Swartz and Pohl, 1989) theorizes that the resistance arises due to the mismatch in the phonon density of states. The key idea is that, at room temperature, phonon wavelengths contributing to heat conduction are comparable to the mean surface roughness height. Therefore, scattering should be diffusely at the interface and randomize the wave vector and polarization of the phonon. The principle of detailed balance and the fundamental definition of diffuse scattering are cleverly employed to obtain the transmission coefficient of phonons. The Debye approximation is used in the original formulation but more accurate dispersion can be inserted instead to evaluate the transmission coefficient. The model predicts a ~T^3 behavior for the thermal conductance but provides a different quantitative value than the AMM. The transmission coefficient for DMM assuming the Debye dispersion model for the two materials is

$$\alpha_{12} = \frac{\sum_s c_{2,s}^{-2}}{\sum_s c_{1,s}^{-2} + \sum_s c_{2,s}^{-2}}, \tag{4.21.35}$$

where the summations are over the phonon polarizations. The coefficient is independent of phonon frequency and incidence angle.

Within the Debye model, the interfacial thermal conductance for AMM (DMM) simplifies to

$$G_{\text{AMM (DMM)}} = \left(\frac{k_B T}{\hbar}\right)^3 \frac{k_B}{2}\frac{3}{\pi c_D^2}\Gamma_{\text{AMM (DMM)}} \int_0^{\Theta_D/T} \frac{z^4 e^z \mathrm{d}z}{\left(e^z - 1\right)^2}, \tag{4.21.36}$$

where the function Γ depends on the choice of the model and is given by

$$\Gamma_{\text{AMM}} = \int_0^{\min\left(\frac{\pi}{2},\Theta_{\text{MAX}}\right)} \alpha(\theta)\sin\theta\cos\theta\ \mathrm{d}\theta;\ \Gamma_{\text{DMM}} = \frac{\alpha_{\text{DMM}}}{2}. \tag{4.21.37}$$

The upper limit of the integral for AMM must be carefully considered for total internal reflection when the transmission is into the material with the higher speed of sound. The upper limit in this case will be less than $\pi/2$.

Another important benchmark for interfacial thermal conductance is the radiation limit (Stoner and Maris, 1993). In this limit, all phonons incident from the side with the lower density of states have a transmission coefficient of unity. At temperatures large compared to the Debye temperature, this simplifies to

$$G_{\mathrm{RAD}} = \frac{k_{\mathrm{B}}\omega_{\mathrm{D,1}}^3}{24\pi^2}\sum_s\left(\frac{1}{c_{s,2}^2}\right), \tag{4.21.38}$$

where 1 is the side with the lower Debye frequency.

Heat flow across metal–dielectric interfaces are complicated by the involvement of the energy transfer between electrons and phonons. While data suggest that the coupling between electrons and phonons plays an important role, it is still unclear which electron–phonon couplings are relevant: with phonons in the metal, with phonons in the dielectric, or with interfacial phonon modes. One approach is to consider the electron–phonon conductance $G_{\mathrm{e-ph}}$ and the phonon–phonon $G_{\mathrm{ph-ph}}$ conductance in series (Majumdar and Reddy, 2004). The DMM provides the phonon–phonon conductance $G_{\mathrm{ph-ph}}$. The net conductance is

$$G_{\mathrm{net}} = \frac{G_{\mathrm{e-ph}}G_{\mathrm{ph-ph}}}{G_{\mathrm{e-ph}} + G_{\mathrm{ph-ph}}}. \tag{4.21.39}$$

Obtaining $G_{\mathrm{e-ph}}$ presents some difficulty. An empirical approach is to assume a small departure from equilibrium. An energy balance based on the Fourier law yields an electron–phonon thermal conductance, $G_{\mathrm{e-ph}} = \sqrt{Bk_{\mathrm{p}}}$ (Majumdar and Reddy, 2004), where B is the electron cooling rate and k_{p} is the phonon thermal conductivity of the metal. Another approach is to obtain $G_{\mathrm{e-ph}}$ directly from the Bloch–Boltzmann–Peierls formula for electron–phonon interaction (Singh et al., 2013). With this, the expression is

$$G_{\mathrm{e-ph}} = \frac{D_{\mathrm{ac}}^2 m^{*2} k_{\mathrm{B}} T_{\mathrm{e}} L}{4\pi^3\hbar^3\rho(T_{\mathrm{e}} - T)}\int_0^{q_{\mathrm{D}}} q^3\left(\left(e^{\frac{\hbar v_{\mathrm{s}} q}{k_{\mathrm{B}}T}} - 1\right)^{-1} - \left(e^{\frac{\hbar v_{\mathrm{s}} q}{k_{\mathrm{B}}T_{\mathrm{e}}}} - 1\right)^{-1}\right)\left(\ln\left(\frac{1 + e^{\frac{E_{\mathrm{F}}}{k_{\mathrm{B}}T_{\mathrm{e}}}} e^{-\left(\frac{\hbar^2 q^2}{8m^*} + \hbar v_{\mathrm{s}} q\right)/k_{\mathrm{B}}T_{\mathrm{e}}}}{1 + e^{\frac{E_{\mathrm{F}}}{k_{\mathrm{B}}T_{\mathrm{e}}}} e^{-\frac{\hbar^2 q^2}{8m^* k_{\mathrm{B}}T_{\mathrm{e}}}}}\right)\right)\mathrm{d}q, \tag{4.21.40}$$

where D_{ac} is the deformation potential for electron scattering by acoustic phonons, v_{s} is the speed of longitudinal waves, q is the phonon wave vector, ρ is the density of the metal, and L is the characteristic length scale for energy transfer. A suitable choice is $L = 3\Lambda_{\mathrm{e}}$, where Λ_{e} is the electron mean free path in the metal. This corresponds to ~95% energy loss. The electron gas is assumed to be in equilibrium at temperature T_{e} and the phonons to be in equilibrium at temperature T.

Heat flow across metal–metal interfaces is important in technologies such as heat-assisted magnetic recording, spintronics, and magnetic sensors. Recent measurements (Wilson and Cahill, 2012) on metal multilayers (Al/Cu and Pd/Ir) have experimentally validated the Wiedemann–Franz law for metal–metal interfaces,

$$\frac{G_{\mathrm{e-e}}AR}{T} = L_{\mathrm{o}}, \tag{4.21.41}$$

where $G_{\mathrm{e-e}}$ is the interfacial thermal conductance per unit area due to electron transport, AR is the specific electrical resistance of the interface, and L_{o} is the Lorenz number estimated from the

Sommerfeld theory discussed above. Further, the value of G can also be obtained from a DMM for electrons, similar to the consideration above. The expression for G in this case is

$$G_{e-e} = \frac{Z_{e1} Z_{e2}}{4\left(Z_{e1} + Z_{e2}\right)}, \tag{4.21.42}$$

where $Z_e = \gamma T v_F$ is the product of the electron heat capacity per unit volume and the Fermi velocity for each side, respectively.

Modeling Techniques for Materials, Devices, and Circuits

The basic theory described above provides the framework for simulating the thermal behavior of electronic materials, devices, circuits, and packaging. The most detailed models in terms of physics are at the materials level. At the device and circuit levels, the interplay with electrical behavior is important and hence, the emphasis is on electrothermal simulations. In contrast, at the package level, mechanical integrity becomes important and hence, thermomechanical modeling forms the focus. The implementation of thermal models in circuit simulations and packaging is discussed later. The basic thermal equation in these cases is the heat diffusion equation. Here, we discuss detailed physical models of heat conduction at the materials and device levels.

Since phonon wavelengths that dominate heat conduction are approximately 0.1–2 nm at room temperature, a full quantum treatment of heat conduction is rarely warranted at room temperature unless the characteristic dimension approaches nanometers rather than several tens of nanometers. Generally, semiclassical and classical methods for modeling phonon transport are adequate. Semiclassical modeling involves either the use of the phonon BTE (Sinha and Goodson, 2005) with varying details of phonon dispersion or the Monte Carlo (MC) (Liu, 2001; Péraud et al., 2014) technique to statistically simulate phonon transport. Another choice is molecular dynamics (MD), which though classical in origin can be extended to a semiclassical treatment, and involves the use of the linear response theory to compute transport coefficients. MD can also be used to compute physical parameters needed by the BTE (Allen and Tildesley, 1987). Figure 4.21.11 summarizes the hierarchy of semiclassical thermal modeling and highlights the main features of each approach.

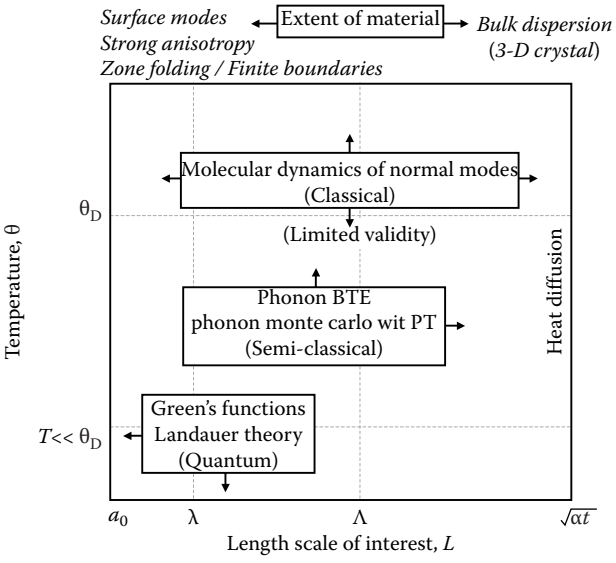


FIGURE 4.21.11 Regimes for subcontinuum thermal modeling relevant to electronics. (Adapted from Sinha, S. and K.E. Goodson, *Int. J. Multiscale Comput. Eng.*, 3, 107–33, 2005. With permission.)

The phonon BTE (also called the Peierls BTE) (Peierls, 1955) is the fundamental semiclassical equation to describe phonon heat conduction in solids. The fundamental assumption in deriving the phonon BTE is that there exists a distribution function, $n_{q,s}(\boldsymbol{r},t)$, which provides the average occupation of the phonon mode ($\boldsymbol{q}$, s), in the neighborhood of a location $\boldsymbol{r}$ at time t. The equation assumes the simultaneous prescription of phonon position and momentum with arbitrary precision. However, in quantum mechanics, these quantities correspond to noncommuting operators and hence, obey the uncertainty principle. The semiclassical BTE for phonons resolves this contradiction by treating phonons as wave formed from a superposition of the normal modes of the crystal such that each wave packet has a small spread δq in the wave vector and is localized in space in a region of size δr such that $\delta q \delta r \sim 1$. Figure 4.21.12 illustrates a spatially Gaussian phonon wave packet (Sinha et al., 2005).

In the semiclassical model, the number of phonons becomes a function not only of the wave vector $\boldsymbol{q}$ and polarization s but also of the spatial coordinate $\boldsymbol{r}$. This evolves in time in a six-dimensional phase space due to motion in real space, due to anharmonic scattering with other phonons, and due to scattering with impurities and boundaries. The solution to the phonon BTE requires the evaluation of the collision term. Though Peierls obtained the formal expression as an integral in wave-vector space over all allowed scattering processes, it is too complicated to be directly evaluated in practice. Instead, the relaxation time approximation has been widely used, with explicit solutions available in the literature for deriving the thermal conductivity. The derivations of the relaxation times involve time-dependent perturbation theory from quantum mechanics (see Kroemer (1994), for example).

Klemens (Han and Klemens, 1993) and others (Callaway, 1959; Holland, 1963) have developed semiempirical expressions for the relaxation rates that model thermal conductivity data over a wide range of temperatures. It is important to note that the original expressions are evaluated using simplified phonon dispersion relationship. Mass difference (isotope) and Umklapp scattering are of the form $\tau_m^{-1} = A\omega^4$ and $\tau_U^{-1} = BT\omega^2 e^{-C/T}$. The value of A depends on the isotope concentration. A commonly used value is (Holland, 1963) $A = 1.32 \times 10^{-45}$ s^3. Using values from Holland (1963) as a starting point, the values of B, C, and the cutoff frequency ω_C can be iteratively determined to simultaneously fit the high-temperature thermal conductivities of bulk silicon and silicon nanostructures, respectively (Li, 2002). Following this procedure, the values are $B = 1.6 \times 10^{19}$ s/K, $C = 152$ K, and $\omega_C = 4.3 \times 1013$ rad/s. The relaxation rate for boundary scattering is available from several models. Figure 4.21.13 plots the scattering rates for phonons in silicon as a function of frequency.

In the simplest Casimir limit, phonons are assumed to scatter diffusely at the boundaries without any other scattering mechanisms present in the bulk of the crystal. The rate is given by $\tau_B^{-1} = v / L_C$, where L_C is the limiting (smallest) size of the crystal and v is the phonon speed. If other scattering

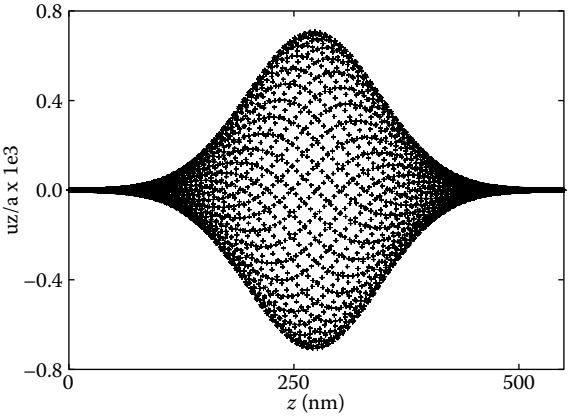


FIGURE 4.21.12 The spatial distribution of energy in an LO phonon wave packet. (Adapted from Sinha, S. and K. E. Goodson, *Int. J. Multiscale Comput. Eng.*, 3, 107–33, 2005. With permission.)

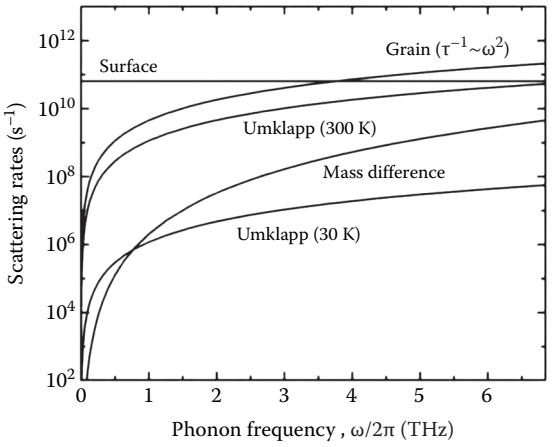


FIGURE 4.21.13 Frequency-dependent phonon scattering rates calculated for silicon. Rates for Umklapp scattering are shown at 30 and 300 K. The characteristic dimension for the surface as well as grain boundary scattering is assumed to be 100 nm in each case. (Adapted from Ma, J. et al., *Nano Lett.*, 13 (2), 618–24, 2013. With permission.)

mechanisms may be assumed to have an overall scattering time τ, an analytical solution to the BTE proposed by Fuchs and Sondheimer (Fuchs, 1938; Sondheimer, 1952) applies. Figure 4.21.14 plots the reduction in thermal conductivity from the value in the bulk as a function of the ratio of the crystal size to the mean free path. Scattering at the boundary may be fully diffuse or may have a partially specular character. This is modeled using a single parameter termed the specularity, p. The specularity parameter reflects the physics that a part of the phonon spectrum with long wavelength should be specular and the part with wavelengths comparable to the surface roughness height should be diffuse. Boundary scattering at room temperature is typically assumed to be diffuse since phonon wavelengths contributing to thermal conduction are a few nanometers and comparable to the surface roughness heights. This corresponds to $p=0$ with $p=1$ being fully specular.

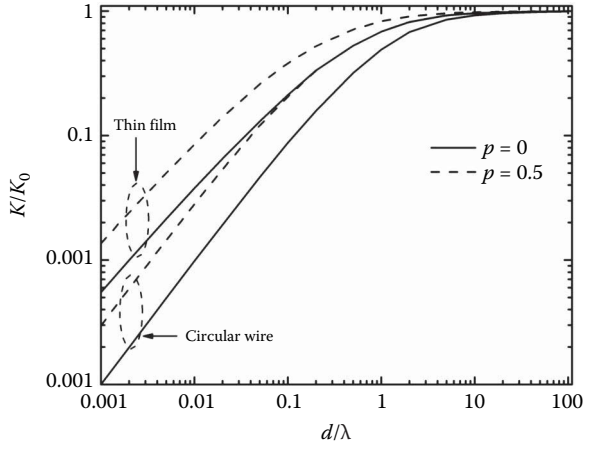


FIGURE 4.21.14 The ratio of thermal conductivity in a finite-size crystal to that in the bulk based on the Fuchs–Sondheimer model (Sondheimer, 1952). The ratio is shown for circular wires and thin films. The x-axis is the ratio of the characteristic length (wire diameter or film thickness) normalized by the MFP. Solid and dashed curves represent specularity, p of 0 and 0.5, respectively.

A key conceptual problem in using the relaxation time approximation is the requirement for a thermodynamic temperature that governs the scattering rate. Since phonons are not in an equilibrium distribution during transport, the temperature cannot be defined in a rigorous sense. The usual practice in such nonequilibrium problems is to define an *ad hoc* "equivalent" temperature based on the local energy (Majumdar, 1993; Chen, 1996, 2001; Sverdrup et al., 2000). We note that this equivalent temperature does not represent the local phonon population but rather the local energy density.

Recent state-of-the-art simulations of thermal conductivity (Ward and Broido, 2011; Esfarjani et al., 2011) have used *ab initio* calculations of the quantum mechanical Schrödinger wave equation to compute interatomic force constants including anharmonic terms. This allows a direct computation of phonon–phonon scattering rates and hence, the thermal conductivity of a bulk material. Figure 4.21.15 shows the computed contributions of different phonon wavelengths and mean free paths to the thermal conductivity of silicon at 277 K. Phonon mean free paths span a range from 10 nm to 10 μm, with substantial contributions from phonons with mean free paths in the micrometers. The range of wavelengths is from 0.1 to ~1 nm.

Solution techniques for the BTE span a range, from MC simulations with detailed scattering rates but relatively small and simple spatial domains to discrete ordinates methods that are suitable for radiative transfer-type formulations of the BTE. The MC method is a statistical sampling technique that can be applied to a wide variety of problems, from simulating a complex physical system such as neutrons in a reactor to solving a mathematical problem such as an integration in multidimensional space. In the MC method, a physical process may be simulated directly without requiring knowledge of the underlying governing equations. The sole requirement is that the process be describable in terms of probability density functions (pdfs). The simulation proceeds by sampling from the pdfs using a computer-generated sequence of random numbers. The MC technique can be applied to simulating a distribution of phonons by tracking the trajectories of individual phonons in phase space, provided the pdfs associated with all scattering events are known. The probabilities for phonon–phonon scattering are particularly difficult to evaluate rigorously in a phonon MC scheme. Instead, the relaxation time approximation is typically employed to make calculations feasible.

The linearized solution to the phonon BTE is of limited usability in modeling transport where the departure from equilibrium is significant. The equation for phonon radiative transfer (EPRT) (Majumdar, 1993) mimics the radiative transfer equation for photons for which solution techniques

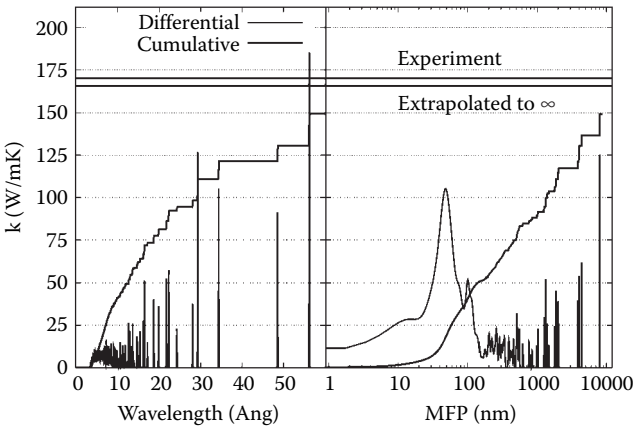


FIGURE 4.21.15 Cumulative contributions of phonons to the thermal conductivity at 277 K, showing contribution based on wavelength (left) and mean free paths (right). Both differential and cumulative thermal conductivities are shown. For comparison, the extrapolated (to infinite k-point mesh) and experimental κ are also shown with horizontal lines at 166 and 174 W/mK, respectively. (Adapted from Esfarjani, K. et al., *Phys. Rev. B*, 84, 085204, 2011. With permission. Copyright 2011 American Physical Society.)

are well developed. The EPRT is written in terms of the phonon intensity, I_ω, with the one-dimensional form being

$$\frac{1}{v}\frac{\partial I_\omega}{\partial t}+\mu\frac{\partial I_\omega}{\partial x}=-\frac{I_\omega-I_\omega^o\left(T\left(x\right)\right)}{v\tau}, \tag{4.21.43}$$

where μ is the direction cosine and v is the (Debye) velocity. The phonon intensity is given as

$$I_\omega=\sum_s vn\left(\omega\left(q,s\right),x,t\right)\hbar\omega D\left(\omega\right). \tag{4.21.44}$$

The EPRT reproduces the expected radiative behavior in the acoustically thin limit and conforms to the Fourier heat flux law in the continuum limit. The EPRT has been used to investigate transport in thin films (Majumdar, 1993; Joshi and Majumdar, 1993; Zeng and Chen, 2000), interfacial transport (Majumdar, 1991), subcontinuum heat sources (Narumanchi et al., 2003), and ballistic conduction from nanoparticles (Chen, 1996).

Another modification of the phonon BTE is the ballistic-diffusive transport equations (BDEs) (Chen, 2001), in which the form of the distribution function is assumed *a priori*. The BDE divides the distribution function, n, at any point into a ballistic, n_b and a diffusive part, n_m. The ballistic part originates from the boundary and experiences out-scattering only, while the diffusive part originates inside the domain and evolves through in-scattering or through a source term. The equations are formally written as

$$\begin{aligned}\frac{1}{|\boldsymbol{v}|}\frac{\partial n_b}{\partial t}+\hat{\boldsymbol{\Omega}}\cdot\nabla n_b&=-\frac{n_b}{|\boldsymbol{v}|\tau}\\ \frac{\partial n_m}{\partial t}+v\cdot\nabla n_m&=-\frac{n_m-\overline{n}}{\tau},\end{aligned} \tag{4.21.45}$$

where $\hat{\boldsymbol{\Omega}}$ is the unit vector along the direction of propagation. Under the gray body approximation, the solution to the BDE matches the solution to the phonon BTE reasonably well. A formulation that is conceptually similar to the BDE but has been developed specifically for semiconductor devices is the split-flux model (Sinha and Goodson, 2005).

The application of the BTE to device problems has focused on the "phonon hotspot" effect. The solution to the BTE for phonon transport in the vicinity of a nanoscale heat source (Chen, 1996) with dimensions smaller than the mean free path reveals a region of nonequilibrium next to the heat source. Phonons emitted from the source do not scatter immediately outside, as is implicitly assumed under the Fourier law. Rather, they travel ballistically over long distances compared to the dimensions of the source. From the perspective of the Fourier law, however, transport is diffusive. The difference between the two temperature profiles is shown for a one-dimensional heat source in Figure 4.21.16. The BTE predicts the effective temperature rise to be larger than the Fourier law. Effectively, this can be thought of as an added thermal resistance to the thermal resistance predicted by the Fourier law.

Fundamental thermal transport in different materials can also be simulated using MD. In new materials where relaxation times are difficult to calculate or obtain, molecular simulations present a better path to understanding heat conduction physics. MD refers to the solution of classical equations of motion (Newton's laws) for a set of molecules. In the context of phonon transport in solids, this amounts to solving the dynamics of atoms in a crystal lattice. For a system of p atoms interacting through a potential, Φ, the equation of motion for the ith atom is

$$m_i\ddot{\boldsymbol{r}}_i=-\nabla_{r_i}-\Phi, \tag{4.21.46}$$

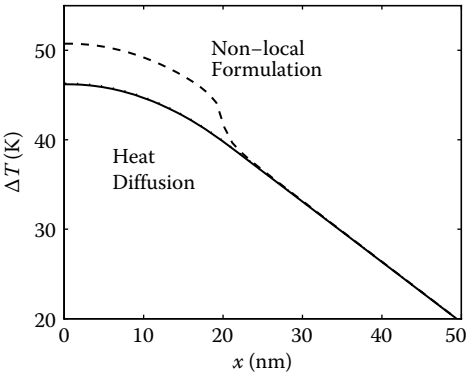


FIGURE 4.21.16 When the characteristic dimension becomes comparable to the energy carrier mean free path, subcontinuum phenomena outside the Fourier law can be observed. Temperatures near a subcontinuum heat source can be elevated compared to predictions from heat diffusion. (Adapted from Sinha, S., and K. E. Goodson, *Int. J. Multiscale Comput. Eng.*, 3, 107–33, 2005. With permission.)

where $\boldsymbol{r}$ is the position of the atom. The interatomic potential has the general form

$$\Phi = \sum_i \Phi_1(\boldsymbol{r}_i) + \sum_i \sum_{j>i} \Phi_2(\boldsymbol{r}_i, \boldsymbol{r}_j) + \sum_i \sum_{j>i} \sum_{k>j>i} \Phi_3(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k) + \cdots, \tag{4.21.47}$$

where Φ_1 is the self-potential, Φ_2 is the contribution from pairs of atoms, and Φ_3 is the contribution from triplets of atoms. A potential that includes terms only till Φ_2 is called a pair potential. A three-body potential includes terms up to Φ_3. Finite difference algorithms (Allen and Tildesley, 1987) such as the Verlet or the Gear predictor–corrector are used to solve the system at each time step, which is usually less than a femtosecond. Computational constraints dictate that a trade-off be made between the size of the system and the time for the simulation. Since the maximum phonon wavelength is given by the size of the system, long wavelength phonons are cut off in an MD simulation. Use of periodic boundary conditions and a careful choice of the system size offset this limitation. Parallelization permits simulations of systems comprising tens of thousands of atoms. MD has generally been used in thermal modeling to compute phonon relaxation times, the thermal conductivity of materials, and the thermal conductance of interfaces. *Ab initio* simulations can be further combined with MD (Ward and Broido, 2011; Esfarjani et al., 2011) to compute the interatomic forces and consequently phonon relaxation times and mean free paths. These simulations represent the state of the art in the thermal modeling of materials.

In simulations of semiconductor devices, charge transport is the primary focus. When a thermal part is included, the simulations are referred to as electrothermal simulations. Electrothermal transport in a semiconductor device involves charge transport by electrons and holes and energy transport by phonons, with the two systems interacting strongly. The charge carriers gain energy from the applied electric field and subsequently lose it to the lattice while restoring thermodynamic equilibrium through scattering. The process may proceed in the reverse direction as well since electrons and holes may absorb lattice phonons and gain energy. The latter occurs near the device source where phonon absorption enables the less energetic carriers to overcome the potential barrier. At the drain terminal, electrons shed energy gained from the field by emitting phonons. In addition, charge carriers also interact with photons during generation and recombination.

Semiclassical electron transport in a device is governed primarily by three sets of equations (Muller and Kamins, 1977; Blotekjaer, 1970). The first is a set of equations for conservation of particles. The second set relates the particle currents to the field and the particle gradients, and finally, the Poisson equation provides the electrostatic potential inside the device due to charged immobile

dopant atoms and mobile carriers. This basic set comprises what is known as the drift-diffusion model (Blotekjaer, 1970). The drift-diffusion model does not include an energy conservation equation but instead assumes isothermal conditions. This can lead to inaccuracies in the diffusion term since temperature dependence is ignored. The next model in the hierarchy is the hydrodynamic model (Wachutka, 1990), which includes an energy equation. The constitutive equations for continuity and momentum are derived by taking moments of the Boltzmann equation for electrons. Heat generation is derived rigorously from irreversible thermodynamics using Onsager's relations. Higher in the hierarchy are MC device simulations (Fischetti and Laux, 1993) that track the trajectory of electrons inside the device. The MC method implements a quantum mechanical description of electron–phonon scattering.

Thermometry Techniques

Measurements are a key component in the thermal engineering of electronic micro- and nanostructures, offering insight and data where the guidance of theory is often qualitative. Thermometry at micro- and nanoscale dimensions requires sophisticated miniaturized probes that often need to be custom designed and fabricated in semiconductor cleanrooms. In almost all cases, the experimental structure requires cleanroom fabrication. Like macroscopic measurements, measurements at microscopic scales may also employ thermocouples. However, there exist several options beside thermocouples: Electrical resistance thermometry relies on the change in electrical resistance of metals or semiconductor diodes with temperature. Thermoreflectance relies on the change in reflectance (originating in the change in the complex dielectric constant) with temperature. Raman thermometry relies on the shift in the frequency of the Raman line with temperature. Infrared thermometry relies on measuring emission characteristics in the infrared. Thermal expansion of a polymer is used as the transducer in Joule expansion thermometry. Length scales approaching a few nanometers and timescales approaching a few picoseconds can be measured, though not simultaneously. Temperature changes as small as 100 mK are routinely measured.

Among experimental thermometry techniques intended for measuring material properties, the more important ones are the 3-omega (3-ω) method and frequency domain thermoreflectance for micrometer scale thin films, time-domain thermoreflectance (TDTR) for nanometer scale thin films, suspended bridge platform for one-dimensional and two-dimensional nanostructures. Other techniques such as infrared thermometry, Raman thermometry, and scanning probe methods are described in the literature (Cahill et al., 2014).

The 3-omega measurement (Cahill, 1990) is widely used in measuring the thermal conductivity of thin films and has been extensively discussed in the literature (Dames and Chen, 2005). A schematic of measurement device configuration is illustrated in Figure 4.21.17. A thin metal line, acting as a heater and a thermometer, is deposited on the film. During the measurement, a sinusoidal current at a frequency ω through the metallic heater sets up an oscillatory temperature rise,

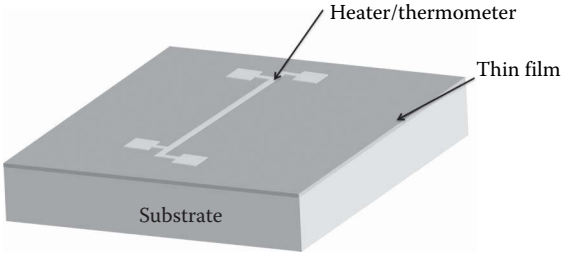


FIGURE 4.21.17 Schematic of a device for 3-ω measurement of thin film thermal conductivity. The metal line serves as the heater/thermometer. It is usually deposited through a shadow mask or patterned lithographically.

ΔT at the harmonic frequency 2ω. This results in a voltage across the metal line that oscillates at 3ω. Measurement of the 3ω voltage via a lock-in amplifier provides ΔT as

$$\Delta T = 2\frac{V_{3\omega}}{I_0}\left(\frac{\mathrm{d}R}{\mathrm{d}T}\right)^{-1}, \tag{4.21.48}$$

where I_0 is the 1ω current and $\mathrm{d}R/\mathrm{d}T$ is the temperature coefficient for heater resistance.

In the simple case when the thermal penetration depth $\sqrt{\alpha_T / 2\omega}$ >> film thickness d, ΔT has a constant contribution from the film and a frequency-dependent component from the semiinfinite substrate:

$$\Delta T = \frac{p}{\pi l k_s}\left(\frac{1}{2}\ln\frac{\alpha_T}{b^2} - \frac{1}{2}\ln(2\omega) + 0.923 - i\frac{\pi}{4}\right) + \frac{pd}{2blk_f}, \tag{4.21.49}$$

where p is the heating power, l and $2b$ are length and width of the heater, k_s and k_f are the thermal conductivities of the substrate and the thin film, respectively. By measuring the 3ω voltage at different frequencies, it is possible to decouple the two components and extract the thermal conductivity of both the thin film and the substrate.

In a more general case when a multilayer system is involved, an N-layer heat diffusion model (Borca-Tasciuc et al., 2001) can be used. The temperature rise is

$$\Delta T = \frac{-p}{\pi l k_1}\int_0^\infty \frac{1}{A_1 B_1}\frac{\sin^2(\lambda b)}{\lambda^2 b^2}\mathrm{d}\lambda, \tag{4.21.50}$$

where

$$A_{i-1} = \frac{A_i\dfrac{k_i B_i}{k_{i-1}B_{i-1}} - \tanh(\phi_{i-1})}{1 - Ai\dfrac{k_i B_i}{k_{i-1}B_{i-1}}\tanh(\phi_{i-1})}, i = 2..N$$
$$\mathrm{Bi} = \left(\lambda^2 + \frac{i2\omega}{\alpha_{Ti}}\right)^{1/2}, \phi_i = B_i d_i \tag{4.21.51}$$

In the above expression, the layer number count starts at the topmost film and properties are assumed to be isotropic. The thermal diffusivity α_T and the individual thicknesses d_i appear in the above equations. Figure 4.21.18 shows a typical plot for a frequency sweep for ΔT. The sample is a silicon nanostructure on a bulk silicon substrate. The fit to the data yields the thermal conductivity. The fitting process works best when the only unknown is the thermal conductivity of the thin film or when the thin film contributes most of ΔT.

TDTR (Cahill et al., 2014) is an optical, noncontact technique that can be used to measure the thermal conductivity of materials. TDTR can be applied to bulk materials, thin films, and individual interfaces. A schematic of the typical experimental setup used is shown in Figure 4.21.19. A Ti:sapphire laser oscillator produces a series of sub-picosecond optical pulses (λ=785 nm) at a high repetition rate, typically 80 MHz. The output of the laser is split into a "pump" beam and a "probe" beam with their relative optical paths being adjusted using a mechanical delay stage. The intensity of the pump beam is modulated using an electro-optic modulator at a frequency, f=9.8 MHz. A part of the reflected beam from the surface is deflected into the CCD camera using a beam splitter. This creates a dark field image in the camera that helps to focus the beams accurately on the sample surface.

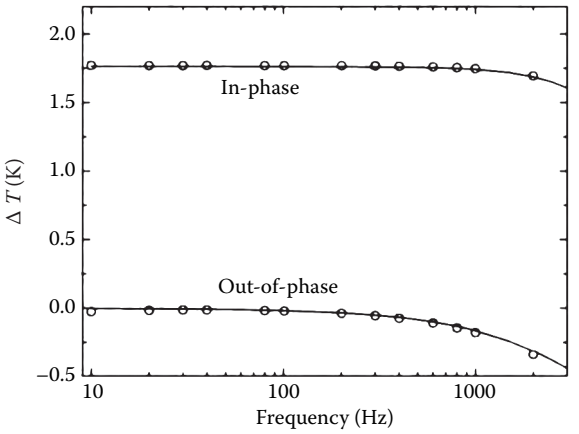


FIGURE 4.21.18 The in-phase and out-of-phase components of the temperature rise measured at different frequencies in a 3-ω measurement.

The pump beam produces ultrafast heating and the probe beam measures the change in the temperature of the surface through the dependence of the optical reflectivity R on temperature. A thin film of aluminum is deposited on the sample and acts as transducer due to its large coefficient of thermoreflectance, $dR/dT \approx 10^{-4}\ K^{-1}$ near room temperature and at the wavelength used. The changes in the reflected probe intensity caused by the pump beam occur at the modulation frequency f and are extracted using lock-in amplifier. The lock-in signal has an in-phase and an out-of-phase component, $V(t) = V_{in}(t) + iV_{out}(t)$. Typically, the in-phase signal is used in the analysis but the out-of-phase signal can be used to correct for the nonidealities in the experiment. Factors such as defocusing of the pump beam and change in the overlap between the

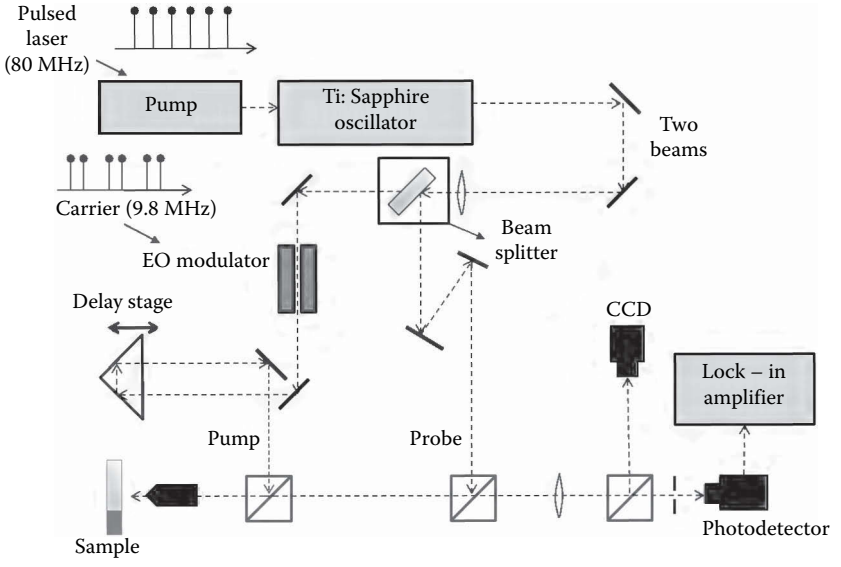


FIGURE 4.21.19 Schematic of the TDTR experimental setup. A Ti:sapphire laser oscillator produces a series of sub-picosecond optical pulses. The output of the laser is split into a "pump" and a "probe" path using a beam splitter. The pump beam is modulated using an electro-optic modulator. The relative path difference between the pump and the probe beam is adjusted using a variable delay stage. (Figure prepared by Dhruv Gelda.)

pump and the probe beams affect both the $V_{in}(t)$ and $V_{out}(t)$ similarly. Therefore, the ratio of the two components $(V_{in}(t)/V_{out}(t))$ is a more robust measurement (Cahill et al., 2014). This ratio is compared with a numerical calculation of heat transport in multilayer structures. The numerical model solves the heat diffusion equation in frequency domain assuming periodic heating from the pulsed laser. The unknown thermal properties, e.g., the thermal conductivity of a thin film, are treated as fitting parameters and adjusted to minimize the difference between the model and the experimental data.

One of the major limitations of TDTR is that the technique is sensitive to the cross-plane thermal conductivity of the material. Except in special cases such as highly anisotropic material, it is difficult to estimate the in-plane thermal conductivity. The penetration depth, δ of thermal waves at a modulation frequency f is given by $\sqrt{\alpha_T / \pi f}$, where α_T is the thermal diffusivity. If the sample consists of a thin film of thickness, $d(< \delta)$, TDTR typically cannot separate the thermal conductance of the film from the thermal conductance of the film–substrate interface.

A microelectromechanical-fabricated device is widely used to measure the thermal conductivity of one-dimensional structures (Shi et al., 2003). Figure 4.21.20 shows the device that includes two suspended membranes supported by long, thermally resistive silicon nitride (SiN_x) beams. These membranes are designated as heating and sensing membranes. Each membrane has a platinum (Pt) coil sandwiched between two SiN_x films (Ghossoub et al., 2013) that serve as resistive heaters/thermometers. The platinum coils (with electrical resistances R_H and R_S) are connected to the contact pads via Pt leads (resistance R_L each) that pass through the supporting beams. Figure 4.21.21 shows a silicon nanowire suspended across the membranes that was transferred from its fabrication wafer. The entire device is kept at bath temperature T_O in a high-vacuum cryostat.

Once the entire device equilibrates to its bath temperature T_O, a DC current, I is passed through the Pt coil in the heating membrane. Joule heat is generated in the coil ($Q_H = I^2 R_H$) and two of the leads ($2 \times Q_L = 2 \times I^2 R_L$) connecting the coil to the contact pads. Half of the heat generated in the

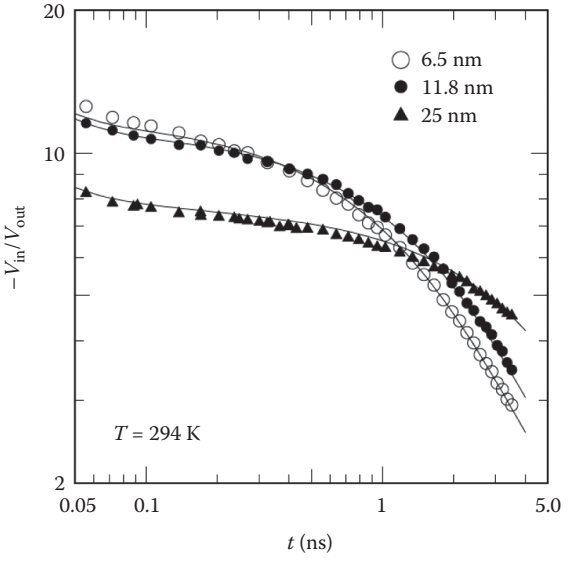


FIGURE 4.21.20 TDTR data acquired at room temperature for $TiN/SiO_2/Si$ samples. Data points are labeled by the SiO_2 thickness. The ratio of the in-phase to out-of-phase signals at the 9.8 MHz modulation frequency is plotted as a function of the delay time t between pump and probe. The solid lines are fits with the thermal conductivity of SiO_2 as the one free parameter. (Adapted from Costescu, R. M. et al., *Phys. Rev. B* 67, 054302, 2003. With permission. Copyright 2003 American Physical Society.)

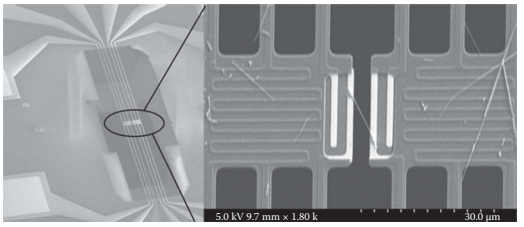


FIGURE 4.21.21 Scanning electron micrographs of a device used to measure the thermal conductivity of one-dimensional structures. The device has two suspended membranes with platinum heater/thermometer lines sandwiched between silicon nitride films. Each membrane is supported by six long thermally resistive silicon nitride beams. The sample to be measured is placed across the membranes and temperature of one of the membranes is raised by Joule heating.

leads reaches the membrane and the other half flows to the substrate at temperature T_O. The total heat in the heating membrane $Q_H + Q_L$ now flows to the substrate at T_O through two pathways. One pathway is directly through the beams supporting the heating membrane at temperature T_H. This pathway is represented by the lower branch of the thermal circuit in Figure 4.21.22. The upper branch in the thermal circuit represents the second heat conduction pathway through the sample, the sensing membrane at temperature T_S and its beams.

By conservation of energy, the heat flowing through the first pathway $Q_1 = G_B(T_H - T_O)$ and the second pathway $Q_2 = G_B(T_S - T_O)$ adds up to the total heat in the heating membrane $Q_H + Q_L$. Assuming the thermal conductance of the beams supporting the heating and the sensing membranes to be equal (G_B), the thermal conductance of the beams is

$$G_B = \frac{Q_H + Q_L}{(T_H - T_O) + (T_S - T_O)}. \tag{4.21.52}$$

Finally, the heat flowing through the sample $G_S(T_H - T_S)$ is the same as the heat flowing between the sensing membrane and the substrate $G_B(T_S - T_O)$. This yields the sample conductance G_S as

$$G_S = G_B \frac{(T_S - T_O)}{(T_H - T_S)}. \tag{4.21.53}$$

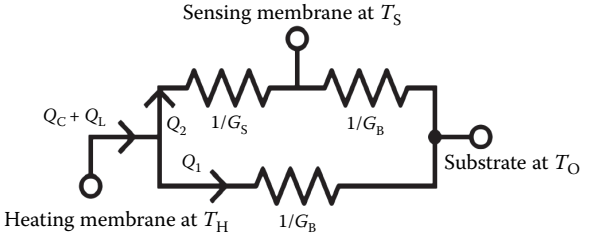


FIGURE 4.21.22 Thermal circuit showing the flow of the generated heat $Q_H + Q_L$ from the heating membrane at T_H to the substrate at T_O. Thermal conductance of the sample is obtained by measuring steady-state temperatures of the membranes and the heat flowing through the sample.

The steady-state temperatures, T_H and T_S, of the two membranes are obtained by measuring the coil resistances, R_H and R_S. The heat flowing through the heating membrane $Q_H + Q_L$ is obtained as $(R_H + R_L) I^2$ by measuring $R_H + R_L$. To measure these resistances, small AC currents at different frequencies are passed through the coils on the heating and sensing sides, and the corresponding voltage drops measured using lock-in amplifiers (Shi et al., 2003). The resistance data can then be converted to a temperature rise using the calibration curve for R_H and R_S. Figure 4.21.23 shows a typical temperature rise measurement at various DC heating currents.

The sample conductance obtained using this device includes the contact conductance between the membranes and the sample. Care should be taken to minimize this contact resistance. Furthermore, for samples with conductance lower than 100 pW/K, the temperature rise on the sensing side may be comparable to the noise equivalent temperature rise, resulting in high error margins.

The above techniques lend themselves well to the measurements of thermal conductivity and conductance in micro- and nanostructures. In addition, measurements on individual transistors, circuit blocks, and complete integrated circuits are of critical importance in applications. Device scale thermometry may employ electrical resistance measurements in contacts to the device or noncontact electroluminescence. Temperatures of circuit blocks can be measured by measuring the frequency shifts of ring oscillators or using electrical resistance thermometers. Finally, infrared cameras and point-by-point thermoreflectance can produce thermal images of circuits and entire chips with spatial resolution on the order of ~10 μm. These are discussed later in the chapter.

Thermal Conductivity/Conductance of Materials/Interfaces

The main effect of the small dimensions of structures used in electronics is to alter the thermal conductivity of the medium due to scattering of carriers at the surfaces and interfaces. In most cases, the heat diffusion equation still applies provided the altered size-dependent thermal conductivity and accurate interfacial thermal conductance are taken properly into account. While theory provides physical insight into altered properties, experimental data are necessary in most cases.

The thermal conductivity of silicon (Toberer et al., 2012) is the most studied and best understood. Figure 4.21.24 shows the data for thermal conductivity of silicon at 300 K as a function of limiting crystal size. Boundary scattering of phonons serves to reduce thermal conductivity as discussed previously. The Casimir limit is shown for comparison. While the data for thin

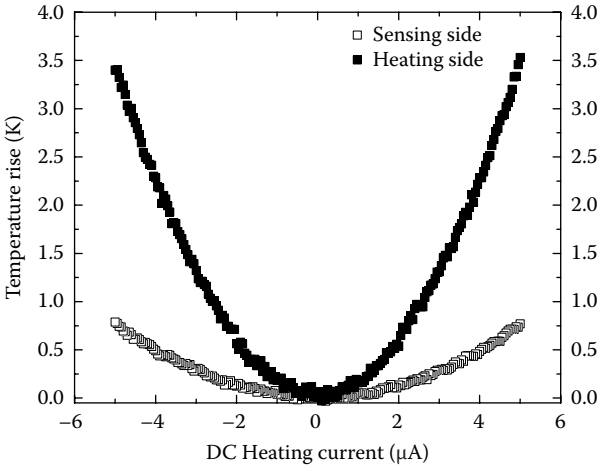


FIGURE 4.21.23 Temperature rise in the membranes at different DC heating currents during a measurement.

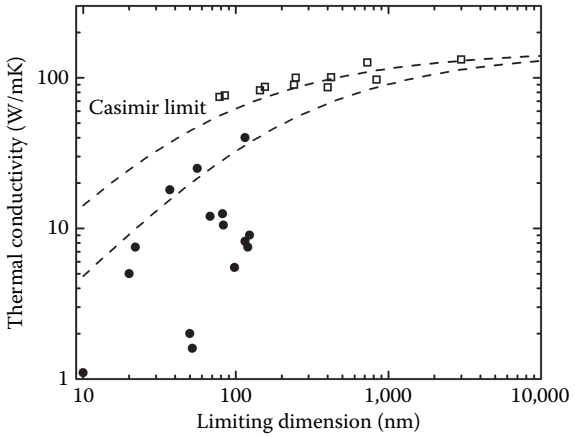


FIGURE 4.21.24 Summary of room temperature thermal conductivity data of silicon nanostructures plotted against the limiting dimension for surface scattering. The closed and open symbols represent nanowires (Boukai et al., 2008; Hochbaum et al., 2008; Li et al., 2003; Ghossoub et al., 2013) and thin films (Asheghi et al., 1998; Ju and Goodson, 1999), respectively. A prediction of thermal conductivity assuming boundary scattering at the Casimir limit of completely diffuse scattering is also plotted following Mingo's modified Callaway model (Mingo, 2003). The geometric factor F for boundary scattering is used with values: $F=1$ for cylindrical wire and $F=4$ for thin films (Wang and Mingo, 2011).

films and smooth surface nanowires appear close to the Casimir limit, the data for rough surface nanowires fall significantly below the Casimir limit. The measurement technique for nanowires differs significantly from that for thin films. There has been speculation whether some of the data for nanowires are unduly influenced by contact resistance (Shi, 2012) in some measurements. Figure 4.21.25 further plots data for doped silicon films at 300 K. At large doping concentrations in excess of ~10^{19} cm^{-3}, thermal conductivity is reduced possibly due to strong electron–phonon scattering.

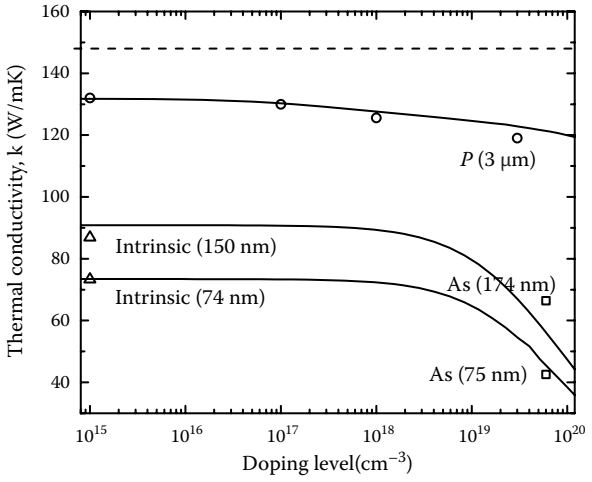


FIGURE 4.21.25 Thermal conductivity as a function of doping concentration at 300 K. The open circles, triangles, and squares are different measurements (Ju and Goodson, 1999; Asheghi et al., 2002; Liu et al., 2006). The curves are calculations for P-doped (3 μm thick), As-doped (174 nm thick), and As-doped (74 nm thick) single crystal silicon films, respectively. (Adapted from Ma, J. and Sinha, S. *Appl. Phys.*, 112, 073719, 2012. Copyright 2012 American Institute of Physics. With permission.)

A multitude of dielectric films are critical in electronics. Figure 4.21.26 presents thermal conductivity data for silicon dioxide and nitride films with various thicknesses. Since these films are, in many cases, in close proximity to heat-producing regions of semiconductor devices, any changes in their thermal conductivity with thickness must carefully be taken into account in simulations. Figure 4.21.27 presents thermal conductivity data for thin films of other dielectrics that find current use in electronics. The theory of minimum thermal conductivity (Cahill and Pohl, 1987) provides a convenient comparison point for the thermal conductivity of amorphous materials. The model assumes the Einstein model for phonons where neighboring atoms are completely uncorrelated. Thus, the mean free path equals the lattice spacing. The minimum value is given by

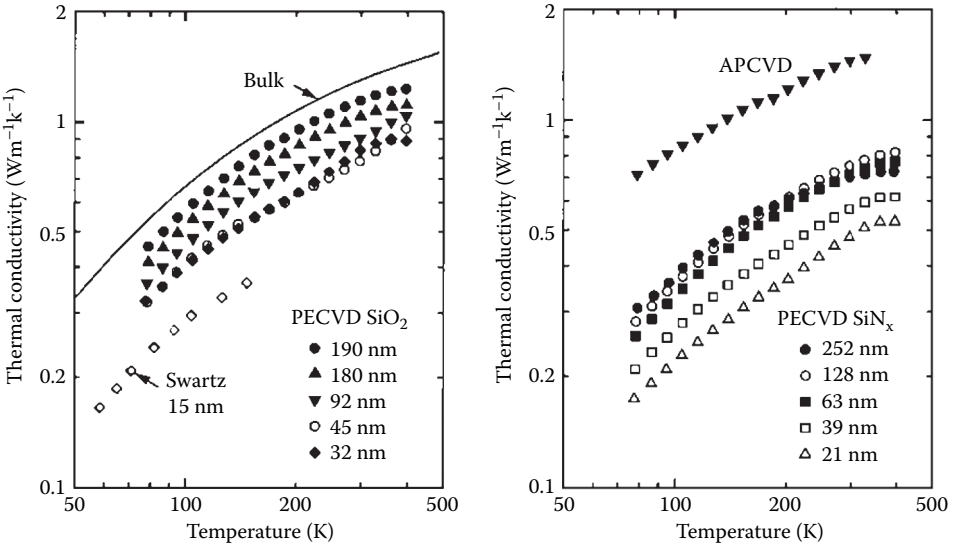


FIGURE 4.21.26 Thermal conductivity of SiO_2 and SiN_x at different film thicknesses. (From Lee, S.-M. and D. G. Cahill, *J. Appl. Phys.*, 81, 2590–5, 1997. Copyright 1997 American Institute of Physics.)

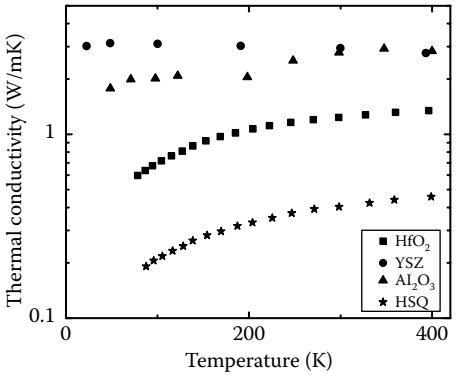


FIGURE 4.21.27 Thermal conductivity vs temperature data for HfO_2 (Lee et al., 1995), α-Al_2O_3 deposited by low-temperature ALD (Cappella et al., 2013), 3 mol% polycrystalline yttria-stabilized zirconia (YSZ) of density 6 Mg/cm^3 (Schlichting et al., 2001), and Dow Corning's flowable oxide FOx HSQ (Costescu et al., 2002). The room temperature thermal conductivity of HfO_2 grown by ALD is reported as 0.48, 0.81, and 0.95 W/mK for 5.6, 11.8, and 20 nm film thicknesses, respectively (Panzer et al., 2009). DC sputtered α-Al_2O_3 show room temperature thermal conductivity of 1.6 W/mK (Lee et al., 1995).

$$k_{\mathrm{MIN}} = \left(\frac{\pi}{6}\right)^{\frac{1}{3}} k_{\mathrm{B}} n_{\mathrm{a}}^{\frac{2}{3}} \sum_{i=1}^{3} v_i \left(\frac{T}{\theta_{\mathrm{D},i}}\right)^2 \int_0^{\theta_{D,i}} \frac{x^3 e^x \mathrm{d}x}{\left(e^x - 1\right)^2}, \tag{4.21.54}$$

where i is an index specifying all three acoustic phonon polarizations, $\theta_{\mathrm{D},i}$ is the Debye temperature associated with polarization i, v_i is the speed of sound, and n_{a} is the number density of atoms. In the high-temperature limit, $k_{\mathrm{MIN}} = \frac{1}{2}\left(\frac{\pi}{6}\right)^{\frac{1}{3}} k_{\mathrm{B}} n_{\mathrm{a}}^{\frac{2}{3}} \sum_{i=1}^{3} v_i$.

Besides silicon, III–V materials are of critical importance in optoelectronics and power electronics. Figure 4.21.28 plots the temperature-dependent thermal conductivities of leading III–V bulk materials. The values should be used with caution since thermal properties in III–V devices can be significantly altered due to the presence of multilayer structures that introduce interfaces, dislocations, and impurities.

The search for electronic materials beyond silicon has driven interest in the thermal properties of almost every potential candidate. Among these, one- and two-dimensional materials/structures attract great interest. Figure 4.21.29 shows the temperature-dependent thermal conductivity of several one-dimensional material structures: single-wall carbon nanotubes (CNTs), multiwall CNTs, smooth and rough silicon nanowires, Si/SiGe superlattice nanowires, InAs nanowires, and SiGe nanowires. Figure 4.21.30 further plots the temperature-dependent thermal conductivity of exfoliated single-layer graphene supported on silicon dioxide as well as suspended five-layer hexagonal boron nitride. For low-dimensional materials/structures, the low-temperature slope of thermal conductivity provides insight into the scattering processes. At low temperatures (~5–30 K), boundary scattering of phonons is the dominant scattering process governing heat conduction in dielectrics. From Equation 4.21.2, the thermal conductivity is $\kappa \sim Cv\Lambda$. In the Casimir limit, the mean free path at low temperatures is proportional to the size of the crystal and is temperature independent. The temperature dependence in thermal conductivity arises only from that in the heat capacity. The temperature dependence of the heat capacity of phonons is $\sim T^d$, where d is the dimension of phonon modes. Thus, at low temperatures, the thermal conductivity too should follow $\sim T^d$. Low-temperature data are thus invaluable in checking the dimensionality involved in phonon heat conduction as well as the applicability of the Casimir limit.

Two-dimensional materials such as graphene are considered attractive for future applications and fall under the category of emerging materials. Figure 4.21.30 plots the thermal conductivity of single-layer graphene supported on oxide as well as suspended five-layer hexagonal boron

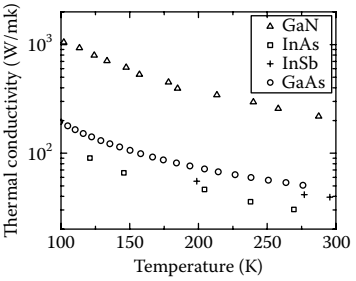


FIGURE 4.21.28 Temperature dependence of the bulk thermal conductivity for III-V semiconductors. Data are shown for GaAs (Inyushkin et al., 2003) crystal with natural composition of gallium isotopes (open circles), *n*-type GaN (Jeżowski et al., 2003) crystals grown under high pressure (open triangles), InSb (plus marks) (Busch and Schneider, 1954), and undoped single crystals of InAs (open squares) (Le Guillou and Albany, 1972).

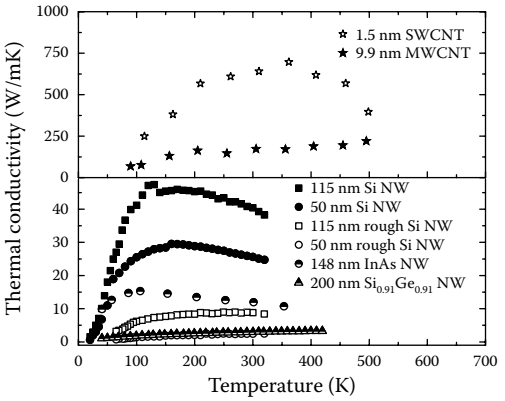


FIGURE 4.21.29 Thermal conductivity as a function of temperature for various one-dimensional materials. The data shown in the top panels are for 1.5 nm diameter single-walled CNT (open stars) and 9.9 nm diameter CNTs with five shells (solid stars) (Pettes and Shi, 2009). Bottom panel shows the data for smooth (Li et al., 2003) and rough (Hochbaum et al., 2008) silicon nanowires, Si/SiGe superlattice nanowires (Li et al., 2003), InAs nanowires (Zhou et al., 2011), and SiGe nanowires (Kim et al., 2010).

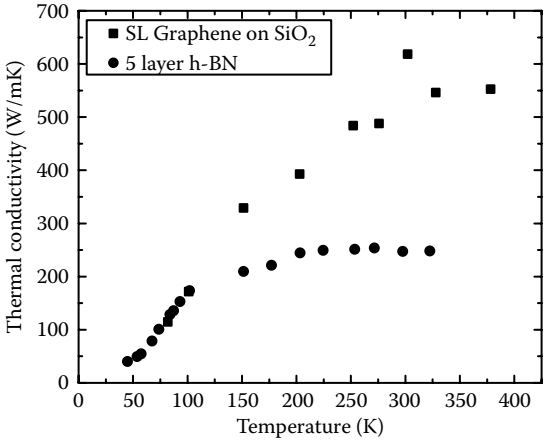


FIGURE 4.21.30 Thermal conductivity of exfoliated single-layer graphene supported on silicon dioxide film (solid squares) (Seol et al., 2010) and suspended five-layer h-BN (solid circles) (Jo et al., 2013) as a function of temperature.

nitride. While thermal conductivity data have also been reported for suspended single-layer graphene, consensus values do not exist. The earlier measurements appear to have larger experimental uncertainties than reported (Shi, 2012). Data for suspended graphene should therefore be used with caution.

Data for interfacial thermal conductance are equally important in the thermal engineering of electronics. While there is a wealth of data on the thermal conductance of material interfaces, the objective is often to measure ideal interfaces and compare against theory. Thus, in many cases, the material combinations chosen are motivated by the availability of high-quality surfaces. Figure 4.21.31 compiles thermal conductance measured across several different solid–solid interfaces (Lyeo and Cahill, 2006). At room temperature, the thermal conductance of good quality interfaces ranges between the lowest and the highest values by a factor of ~60 with high values approaching 500 W/m^{2}K. Besides interfaces formed from bulk materials, interfaces formed by

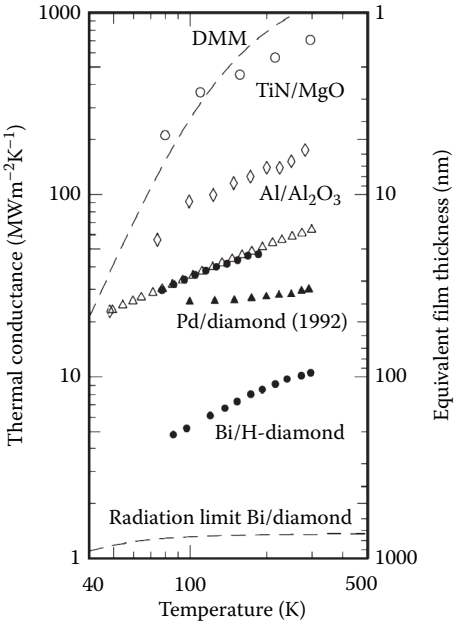


FIGURE 4.21.31 Compilation of the thermal conductance of solid–solid interfaces. The right-hand axis gives the equivalent thickness of a thin film with thermal conductivity 1 W/mK that has the same thermal resistance as these single interfaces. The dashed line labeled "DMM" is a theoretical calculation of the TiN/MgO interface using a modification of the diffuse mismatch model. (Adapted from Lyeo, Ho-Ki and David G. Cahill, *Phys. Rev. B* 73, 144301, 2006. With permission. Copyright 1997 American Institute of Physics.)

low-dimensional materials such as CNT and graphene are important for device applications. Figure 4.21.32 plots the temperature dependence of the interfacial thermal conductance for carbon materials, including single-layer graphene, multilayer graphene, and CNT.

While significant advances have been made in understanding interfacial thermal conductance, a more important question from an engineering perspective is whether the conductance can be altered through engineering. Experiments involving thermal transport across a self-assembled monolayer

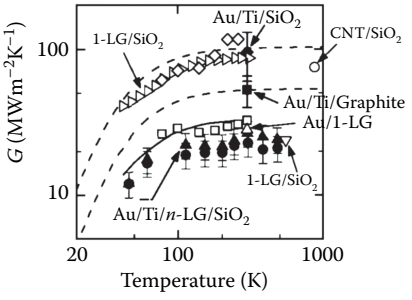


FIGURE 4.21.32 Temperature dependence of the interface thermal conductance (G) for interfaces involving carbon materials: single-layer graphene (1-LG), multilayer graphene (n-LG), graphite, CNT. Data are shown for Au/Ti/1-LG/SiO_2 (solid triangles), Au/Ti/3-LG/SiO_2 (solid circles), Au/Ti/SiO_2 (solid diamond), Au/Ti/graphite (solid square), 1-LG/SiO_2 (open right triangles and open triangle down), Au/1-LG (open triangle up), SWCNT/SiO_2 (open circle), Au/graphite (open squares), and Ti/graphite (open diamonds) interfaces. The dashed lines are DMM calculations for G of Au/Ti/SiO_2 (top) and Au/Ti/graphite (bottom) interfaces. The solid line is the prediction using DMM calculations of $G_{Au/Ti/graphite}$ and measurements of $G_{1\text{-}LG/SiO2}$. (Reprinted from Koh et al., *Nano Lett.*, 10, 4363–8, 2010. With permission.)

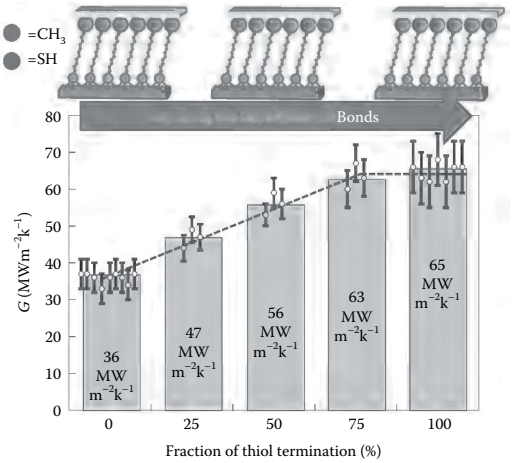


FIGURE 4.21.33 Tuning interfacial thermal conductance. Plot of interfacial thermal conductance (G) as a function of the methyl:thiol end-group ratio for 0%, 25%, 50%, 75%, and 100% thiol end groups. Duplicated structures for each ratio were measured. Error bars represent uncertainty in TDTR data fitting (Reprinted from Losego et al., *Nat. Mater.*, 11 (6), 502–6, 2012. With permission. Copyright 2012 Macmillan Publishers Limited.)

of molecules sandwiched between quartz and Au showed that the interfacial thermal conductance increases significantly as the density of covalent bonds (contributed by thiol end groups) is increased between samples. Figure 4.21.33 shows the data where the conductance nearly doubles from 36 to 65 MW/m^{2}K.

Power Dissipation and Heat Flow in Devices and Circuits

The experimental and theoretical techniques discussed above provide insight into thermal transport in semiconductor devices, circuits, and chips. We now discuss important aspects of thermal management in semiconductor applications. As described in Figure 4.21.4, the heat flow spans multiple length and timescales. Challenges in thermal engineering are unique to the level of consideration, whether device, circuit, chip, or package level.

At the device level, power dissipation is dictated by the microscopic physics of electron–phonon scattering. Electrons accelerated by the electric field in the channel of the transistor relax to equilibrium by emitting phonons. In the device physics literature, such electrons are termed hot electrons. Scattering is typically stronger with high-energy optical phonons and occurs over timescales of ~0.1 ps (Pop et al., 2006). The relaxation of nonequilibrium phonons occurs over ~ps (optical to acoustic) to ~ns (acoustic to acoustic) timescales. Figure 4.21.34 shows the region of heat generation in an 18 nm gate length silicon-on-insulator device (Sinha and Goodson, 2005). The peak volumetric heat generation is 6×10 W/m and the region of intense heat production is on the order of a few nanometers, much smaller than the mean free path of phonons in the lattice bath. A phonon "hotspot" effect can arise if phonons from the heat source travel without scattering over distances much larger than the size of the source. However, the presence of interfaces in close proximity can serve to relax nonequilibrium phonons.

Heat conduction from the transistor occurs through the electrical contacts that act as extended surfaces for conducting heat away from the device. The body of the device serves to spread heat further. The presence of numerous thin film materials and material interfaces complicates the analysis of heat conduction. The thermal conductivity of materials and the conductances of the interfaces must be accurately known to predict temperature distributions reliably. The temperature rise in individual transistors has been measured using numerous techniques. In particular, the electrical

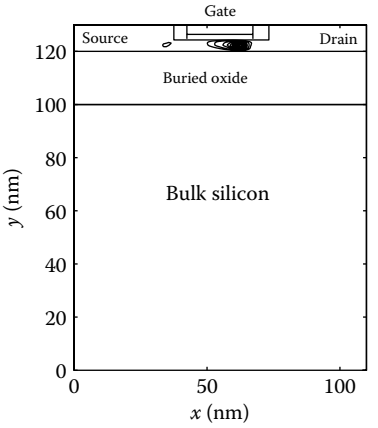


FIGURE 4.21.34 The region of Joule heat dissipation in a thin-body SOI transistor with a gate length of 18 nm. The contours correspond to the spatial distribution of Joule heat and are equi-spaced at 0.5 W/μm^3 with a peak of 60 W/μm^3. (Adapted from Sinha, S. and E. Kenneth, *Int. J. Multiscale Comput. Eng.*, 3, 107–33, 2005. With permission.)

resistance of the gate terminal (Su et al., 1994) has been used as a thermometer to obtain an average temperature for the current-carrying channel. For example, in silicon-on-insulator transistors with 360 nm buried oxide, the measured temperature at ~6 mW power was in the range 60°C–100°C depending on the thickness of the channel. In comparison, the temperature was 40°C in a bulk device at similar power. Besides steady-state measurements, time-resolved electroluminescence (Polonsky and Jenkins, 2004) measurements have yielded temporal temperatures. From the data, the thermal time constant of a 110 nm *n*-type silicon-on-insulator FET is ~107 ns, consistent with a lumped capacitance model.

Circuit level thermal modeling and measurements (Pedram and Nazarian, 2006) are the next level in the hierarchy. The total power dissipation in a CMOS VLSI circuit is the sum $P_{\text{dynamic}}+P_{\text{short circuit}}+P_{\text{static}}$, where the first term represents the dynamic power consumption that occurs when the output signal of a CMOS logic cell changes; the second term represents the power dissipated when both *n*- and *p*-transistors conduct simultaneously to create a short between the supply and the ground; and the third term is mainly due to leakage currents through the gate or subthreshold conduction. While in ~μm size silicon transistors, the dynamic power density is six orders of magnitude larger than the leakage power density, the two power densities become comparable in ~10 nm size transistors (Haensch et al., 2006). Both the short-circuit power and the static power are temperature dependent. Therefore, heat conduction in the chip must be accounted for in circuit models to accurately predict power dissipation. Thermal models in circuit simulations take advantage of the analogy between electrical and thermal circuits. Figure 4.21.35 shows a representative thermal circuit (Pedram and Nazarian, 2006) where the heat flow from a heat source is the current source. The use of resistances, capacitances, and a voltage source for the ambient temperature follows the usual practice. It is critical to consider heat flow along the metal interconnects in detail (Lin et al., 2007a,b). Electrothermal solutions can take two approaches. Thermal and electrical circuits can be solved simultaneously or alternately, they can be solved one at a time and fed to the other. The latter is less computationally intensive and sometimes favored for full-chip simulations. Measurement techniques at the circuit level involve thermometry through embedded temperature sensors or through imaging using infrared imaging as well as thermoreflectance (Altet et al., 2006).

The third problem in the hierarchy is the more practical one. Eventually, heat produced in the circuit needs to be rejected to the ambient. Packaging technologies have introduced numerous innovations to enable advancements through Moore's law (Mahajan et al., 2000). The heat conduction path

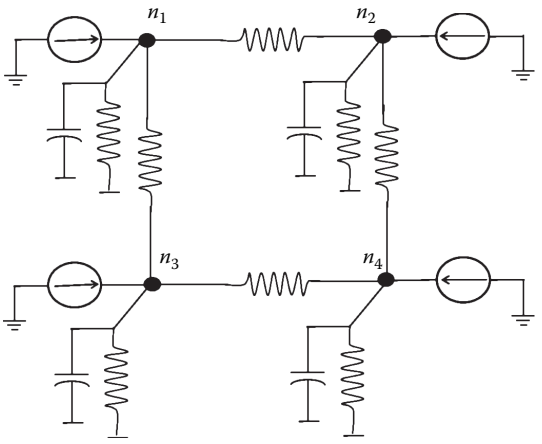


FIGURE 4.21.35 Example of a thermal circuit used in electrothermal modeling of circuits. Temperature is computed at the nodes, *n*. The heat flow is a current source with the thermal resistances and capacitances as indicated. The ground is at the ambient temperature.

from the chip to the ambient goes through the TIM1, the heat spreader, the TIM2, and the heat sink. The heat flow pathway is not one-dimensional since the initial heat distribution on the chip is highly nonuniform. The nonuniformity (Prasher, 2006) is taken into account through a density factor (*DF*). The total thermal resistance, R_{ja}, can be written as the sum $R_{ja}=DF\times\Psi_{jc}+R_{cs}+R_{sa}$, where the subscripts *j*, *a*, *c*, and *s* denote the junction, the ambient, the case (package), and the sink. The symbol Ψ denotes a thermal resistance per unit area. The *DF* has units of cm^{-2} and is numerically 1 for a 1 cm^2 chip heated uniformly. Typically, $DF > 1$ due to nonuniform heating. Within the package, the resistance due to TIM1 is a critical component. Table 4.21.3 lists the thermal conductivities of some important TIMs (Prasher, 2006) employed in electronics along with their thermal properties. The microscopic level theory and measurements discussed above have direct utility in improving thermal interface materials through improved understanding of heat conduction across interfaces.

Summary

Thermal engineering plays a vital role in the advancement of nanoelectronics, with impact at all levels from nanoscale devices to entire packages. Multiple physical length and timescales in transport, the close interplay with charge transport, and the multitudes of materials involved combine to

TABLE 4.21.3
Thermal Conductivity of TIMs

TIM Type	Examples	Range of Thermal Conductivity (W/mK)
Greases	AIN1, Ag, ZnO, Silicon oil	3–5
Gels	Al, Ag, silicone oil, olefin	3–4
Phase change materials	Polyolefins, epoxies, polyesters, acrylics, BN, alumina, Al, CNTs	0.5–5
Phase change metallic alloys	Pure In, In/AG, Sn/Ag/Cu, In/Sn/Bi	30–50
Solders	Pure In, In/AG, Sn/Ag/Cu, In/Sn/Bi	30–50

Source: Samson, Eric C. et al., *Intel Tech. J.*, 9, 75–86, 2005.

create a challenging but interesting problem. With a wide variety of experimental tools and more under development, and powerful theory to provide detailed insight, innovative solutions will continue to emerge. Engineering in this complex environment requires knowledge spanning multiple disciplines. This chapter provides a glimpse into some of the most important aspects.

ACKNOWLEDGMENTS

The authors thank Manjunath Rajagopal and Dhruv Gelda for assistance with the manuscript. S.S. acknowledges financial support from the National Science Foundation (NSF-CAREER and NSF-CBET). This chapter benefited from his interactions over the years with several key colleagues. They are, in chronological order, Ken Goodson, Eric Pop, Ravi Prasher, Ravi Mahajan, Murli Tirumala, Arun Majumdar, Pramod Reddy, and David Cahill.

REFERENCES

Allen, M. P. and D. J. Tildesley. 1987. *Computer Simulation of Liquids*. Oxford: Clarendon Press.

Altet, J., W. Claeys, S. Dilhaire, and A. Rubio. 2006. Dynamic surface temperature measurements in ICs. *Proc. IEEE* 94 (8):1519–33.

Ashcroft, N. W. and N. D. Mermin. 1976. *Solid State Physics*. New York: Rinehart and Winston.

Asheghi, M., K. Kurabayashi, R. Kasnavi, and K. E. Goodson. 2002. Thermal conduction in doped single-crystal silicon films. *J. Appl. Phys.* 91 (8):5079–88.

Asheghi, M., M. N. Touzelbaev, K. E. Goodson, Y. K. Leung, and S. S. Wong. 1998. Temperature-dependent thermal conductivity of single-crystal silicon layers in SOI substrates. *J. Heat Transfer* 120 (1):30–6.

Blotekjaer, K. 1970. Transport equations for electrons in two-valley semiconductors. *IEEE Trans. Electron Devices* 17 (1):38–47.

Borca-Tasciuc, T., A. R. Kumar, and G. Chen. 2001. Data reduction in 3ω method for thin-film thermal conductivity determination. *Rev. Sci. Instrum.* 72 (4):2139–47.

Boukai, A. I., Y. Bunimovich, J. Tahir-Kheli, J.-K. Yu, W. A. Goddard Iii, and J. R. Heath. 2008. Silicon nanowires as efficient thermoelectric materials. *Nature* 451 (7175):168–71.

Busch, G. and M. Schneider. 1954. Heat conduction in semiconductors. *Physica* 20 (7):1084–6.

Cahill, D. G. 1990. Thermal conductivity measurement from 30 to 750 K: The 3ω method. *Rev. Sci. Instrum.* 61 (2):802–8.

Cahill, D. G., P. V. Braun, G. Chen, D. R. Clarke, S. Fan, K. E. Goodson, P. Keblinski et al. 2014. Nanoscale thermal transport. II. 2003–2012. *Appl. Phys. Rev.* 1 (1):011305.

Cahill, D. G., and R. O. Pohl. 1987. Thermal conductivity of amorphous solids above the plateau. *Phys. Rev. B* 35 (8):4067–73.

Callaway, J. 1959. Model for lattice thermal conductivity at low temperatures. *Phys. Rev.* 113 (4):1046–51.

Cappella, A., J.-L. Battaglia, V. Schick, A. Kusiak, A. Lamperti, C. Wiemer, and B. Hay. 2013. High temperature thermal conductivity of amorphous Al_2O_3 thin films grown by low temperature ALD. *Adv. Eng. Mater.* 15 (11):1046–50.

Chen, G. 1996. Nonlocal and nonequilibrium heat conduction in the vicinity of nanoparticles. *J. Heat Transfer* 118 (3):539–45.

Chen, G. 2001. Ballistic-diffusive heat-conduction equations. *Phys. Rev. Lett.* 86 (11):2297–300.

Costescu, R. M., A. J. Bullen, G. Matamis, K. E. O'Hara, and D. G. Cahill. 2002. Thermal conductivity and sound velocities of hydrogen-silsesquioxane low-k dielectrics. *Phys. Rev. B* 65 (9):094205.

Costescu, R. M., M. A. Wall, and D. G. Cahill. 2003. Thermal conductance of epitaxial interfaces. *Phys. Rev. B* 67 (9):054302.

Dames, C. and G. Chen. 2005. 1ω,2ω, and 3ω methods for measurements of thermal properties. *Rev. Sci. Instrum.* 76 (12):124902.

Dennard, R. H., F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. LeBlanc. 1974. Design of Ion-implanted MOSFET's with very small physical dimensions. *IEEE J. Solid-State Circuits* 9 (5):256–68.

Esfarjani, K., G. Chen, and H. T. Stokes. 2011. Heat transport in silicon from first-principles calculations. *Phys. Rev. B* 84 (8):085204.

Fischetti, M. V. 1991. Effect of the Electron-plasmon interaction on the electron mobility in silicon. *Phys. Rev. B* 44 (11):5527–34.

Fischetti, M. V. and S. E. Laux. 1993. Monte Carlo study of electron transport in silicon inversion layers. *Phys. Rev. B* 48 (4):2244–74.

Frank, Kreith, Timmerhaus Klaus, Lior Noam, Shaw Henry, K. Shah Ramesh, J. Bell Kenneth, R. Diller Kenneth et al. 1999. Applications. In *CRC Handbook of Thermal Engineering*. CRC Press.

Fuchs, K. 1938. The conductivity of thin metallic films according to the electron theory of metals. *Proc. Cambridge Philos. Soc.* 34 (1):100–8.

Garimella, S. V., V. Singhal, and Liu Dong. 2006. On-chip thermal management with microchannel heat sinks and integrated micropumps. *Proc. IEEE* 94 (8):1534–48.

Ghossoub, M. G., K. V. Valavala, M. Seong, B. Azeredo, K. Hsu, J. S. Sadhu, P. K. Singh, and S. Sinha. 2013. Spectral phonon scattering from sub-10 nm surface roughness wavelengths in metal-assisted chemically etched Si nanowires. *Nano Lett.* 13 (4):1564–71.

Haensch, W., E. J. Nowak, R. H. Dennard, P. M. Solomon, A. Bryant, O. H. Dokumaci, A. Kumar, X. Wang, J. B. Johnson, and M. V. Fischetti. 2006. Silicon CMOS devices beyond scaling. *IBM J. Res. Dev.* 50 (4/5):339–61.

Han, Y. J. and P. G. Klemens. 1993. Anharmonic thermal resistivity of dielectric crystals at low temperatures. *Phys. Rev. B* 48 (9):6033–42.

Hanson, S., B. Zhai, K. Bernstein, D. Blaauw, A. Bryant, L. Chang, K. K. Das, W. Haensch, E. J. Nowak, and D. M. Sylvester. 2006. Ultralow-voltage, minimum-energy CMOS. *IBM J. Res. Dev.* 50 (4.5):469–90.

Hochbaum, A. I., R. Chen, R. D. Delgado, W. Liang, E. C. Garnett, M. Najarian, A. Majumdar, and P. Yang. 2008. Enhanced thermoelectric performance of rough silicon nanowires. *Nature* 451 (7175):163–7.

Holland, M. G. 1963. Analysis of lattice thermal conductivity. *Phys. Rev.* 132 (6):2461–71.

Incropera, F. P. 2006. *Fundamentals of Heat and Mass Transfer*. New York: John Wiley & Sons.

Inyushkin, A. V., A. N. Taldenkov, A. Y. Yakubovsky, A. V. Markov, L. Moreno-Garsia, and B. N. Sharonov. 2003. Thermal conductivity of isotopically enriched 71 GaAs crystal. *Semicond. Sci. Technol.* 18 (7):685.

Jeżowski, A., B. A. Danilchenko, M. Boćkowski, I. Grzegory, S. Krukowski, T. Suski, and T. Paszkiewicz. 2003. Thermal conductivity of GaN crystals in 4.2–300 K range. *Solid State Commun.* 128 (2–3):69–73.

Jo, I., M. T. Pettes, J. Kim, K. Watanabe, T. Taniguchi, Z. Yao, and L. Shi. 2013. Thermal conductivity and phonon transport in suspended few-layer hexagonal boron nitride. *Nano Lett.* 13 (2):550–4.

Joshi, A. A. and A. Majumdar. 1993. Transient ballistic and diffusive phonon heat transport in thin films. *J. Appl. Phys.* 74 (1):31–9.

Ju, Y. S. and K. E. Goodson. 1999. Phonon scattering in silicon films with thickness of order 100 nm. *Appl. Phys. Lett.* 74 (20):3005–7.

Khalatnikov, I. M. 1952. Teploobmen Mezhdu Tverdym Telom I Geliem-Ii. *Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki* 22 (6):687–704.

Kim, H., I. Kim, H.-J. Choi, and W. Kim. 2010. Thermal conductivities of Si1−xgex nanowires with different germanium concentrations and diameters. *Appl. Phys. Lett.* 96 (23):233106.

Koh, Y. Kan, M.-H. Bae, D. G. Cahill, and E. Pop. 2010. Heat conduction across monolayer and few-layer graphenes. *Nano Lett.* 10 (11):4363–8.

Kroemer, H. 1994. *Quantum Mechanics for Engineering, Materials Science and Applied Physics*. Upper Saddle River, NJ: Prentice Hall.

Le Guillou, G. and H. J. Albany. 1972. Phonon conductivity of InAs. *Phys. Rev. B* 5 (6):2301–8.

Lee, S.-M. and D. G. Cahill. 1997. Heat transport in thin dielectric films. *J. Appl. Phys.* 81 (6):2590–5.

Lee, S.-M., D. G. Cahill, and T. H. Allen. 1995. Thermal conductivity of sputtered oxide films. *Phys. Rev. B* 52 (1):253–7.

Li, D. 2002. *Thermal Transport in Individual Nanowires and Nanotubes*. Berkeley, CA: University of California.

Li, D., Y. Wu, P. Kim, L. Shi, P. Yang, and A. Majumdar. 2003. Thermal conductivity of individual silicon nanowires. *Appl. Phys. Lett.* 83 (14):2934–6.

Li, H., C. Xu, N. Srivastava, and K. Banerjee. 2009. Carbon nanomaterials for next-generation interconnects and passives: Physics, status, and prospects. *IEEE Trans. Electron Devices* 56 (9):1799–821.

Lin, S., G. Chrysler, R. Mahajan, V.K. De and K. Banerjee. 2007a. A self-consistent substrate thermal profile estimation technique for nanoscale ICs-Part I: Electrothermal coupling and full-chip package thermal model. *IEEE Trans. Electron Devices* 54 (12):3342–50.

Lin, S., G. Chrysler, R. Mahajan, V. K. De and K. Banerjee. 2007b. A self-consistent substrate thermal profile estimation technique for nanoscale ICs-Part II: Implementation and implications for power estimation and thermal management. *IEEE Trans. Electron Devices* 54 (12):3351–60.

Liu, J. S. 2001. *Monte Carlo Strategies in Scientific Computing*. New York: Springer-Verlag.

Liu, W., K. Etessam-Yazdani, R. Hussin, and M. Asheghi. 2006. Modeling and data for thermal conductivity of ultrathin single-crystal SOI layers at high temperature. *IEEE Trans. Electron Devices* 53 (8):1868–76.

Losego, M. D., M. E. Grady, N. R. Sottos, D. G. Cahill, and P. V. Braun. 2012. Effects of chemical bonding on heat transport across interfaces. *Nat. Mater.* 11 (6):502–6.

Lyeo, H.-Ki and D. G. Cahill. 2006. Thermal conductance of interfaces between highly dissimilar materials. *Phys. Rev. B* 73 (14):144301.

Ma, J., B. R. Parajuli, M. G. Ghossoub, A. Mihi, J. Sadhu, P. V. Braun, and S. Sinha. 2013. Coherent phonon-grain boundary scattering in silicon inverse opals. *Nano Lett.* 13 (2):618–24.

Ma, J. and S. Sinha. 2012. Thermoelectric properties of highly doped n-type polysilicon inverse opals. *J. Appl. Phys.* 112 (7):073719.

Mahajan, R., K. Brown, and V. Atluri. 2000. The evolution of microprocessor packaging. *Intel Tech. J. Q* 3:2000.

Mahajan, R., R. Nair, V. Wakharkar, J. Swan, J. Tang, and G. Vandentop. 2002. Emerging directions for packaging technologies. *Intel Technol. J.* 6 (2):62.

Majumdar, A. 1991. Effect of interfacial roughness on phonon radiative heat conduction. *J. Heat Transfer* 113 (4):797–805.

Majumdar, A. 1993. Microscale heat conduction in dielectric thin films. *J. Heat Transfer* 115 (1):7–16.

Majumdar, A. and P. Reddy. 2004. Role of electron–phonon coupling in thermal conductance of metal–nonmetal interfaces. *Appl. Phys. Lett.* 84 (23):4768–70.

Mingo, N. 2003. Calculation of Si nanowire thermal conductivity using complete phonon dispersion relations. *Phys. Rev. B* 68 (11):113308.

Moore, G. 1965. Cramming more components onto integrated circuits. *Electronics* 38(8):114–17.

Muller, R. S., and T. I. Kamins. 1977. *Device Electronics for Integrated Circuits*. New York: John Wiley & Sons.

Narumanchi, S. V. J., J. Y. Murthy, and C. H. Amon. 2003. Simulation of unsteady small heat source effects in sub-micron heat conduction. *J. Heat Transfer* 125 (5):896–903.

Panzer, M. A., M. Shandalov, J. A. Rowlette, Y. Oshima, Y. W. Chen, P. C. McIntyre, and K. E. Goodson. 2009. Thermal properties of ultrathin hafnium oxide gate dielectric films. *IEEE Electron Device Lett.* 30 (12):1269–71.

Pedram, M. and S. Nazarian. 2006. Thermal modeling, analysis, and management in VLSI circuits: Principles and methods. *Proc. IEEE* 94 (8):1487–501.

Peierls, R. E. 1955. *Quantum Theory of Solids*. Oxford: Clarendon Press.

Péraud, J.-P.M, C.D. Landon, and N.G. Hadjiconstantinou, 2014. Monte Carlo methods for solving the Boltzmann transport equation. *Annu. Rev. Heat Transfer* 17:205–65.

Pettes, M. T. and L. Shi. 2009. Thermal and structural characterizations of individual single-, double-, and multi-walled carbon nanotubes. *Adv. Funct. Mater.* 19 (24):3918–25.

Polonsky, S. and K. A. Jenkins. 2004. Time-resolved measurements of self-heating in SOI and strained-silicon MOSFETs using photon emission microscopy. *IEEE Electron Device Lett.* 25 (4):208–10.

Pop, E., S. Sinha, and K. E. Goodson. 2006. Heat generation and transport in nanometer-scale transistors. *Proc. IEEE* 94 (8):1587–601.

Prasher, R. 2006. Thermal interface materials: Historical perspective, status, and future directions. *Proc. IEEE* 94 (8):1571–86.

Samson, E. C., S. V. Machiroutu, J.-Y. Chang, I. Santos, J. Hermerding, A. Dani, R. Prasher, and D. W. Song. 2005. Interface material selection and a thermal management technique in second-generation platforms built on Intel® Centrino™ mobile technology. *Intel Tech. J.* 9 (1):75–86.

Schlichting, K. W., N. P. Padture, and P. G. Klemens. 2001. thermal conductivity of dense and porous yttria-stabilized zirconia. *J. Mater. Sci.* 36 (12):3003–10.

Seol, J. H., I. Jo, A. L. Moore, L. Lindsay, Z. H. Aitken, M. T. Pettes, X. Li et al. 2010. Two-dimensional phonon transport in supported graphene. *Science* 328 (5975):213–6.

Shakouri, A. 2006. Nanoscale thermal transport and microrefrigerators on a chip. *Proc. IEEE* 94 (8):1613–38.

Shi, L. 2012. Thermal and thermoelectric transport in nanostructures and low-dimensional systems. *Nanoscale Microscale Thermophys. Eng.* 16 (2):79–116.

Shi, L., D. Li, C. Yu, W. Jang, D. Kim, Z. Yao, P. Kim, and A. Majumdar. 2003. Measuring thermal and thermoelectric properties of one-dimensional nanostructures using a microfabricated device. *J. Heat Transfer* 125 (5):881–8.

Singh, P., M. Seong, and S. Sinha. 2013. Detailed consideration of the electron-phonon thermal conductance at metal-dielectric interfaces. *Appl. Phys. Lett.* 102 (18):181906.

Sinha, S., and K. E. Goodson. 2005. Review: Multiscale thermal modeling in nanoelectronics. *Int. J. Multiscale Comput. Eng.* 3 (1):107–33.

Sinha, S., P. K. Schelling, S. R. Phillpot, and K. E. Goodson. 2005. SCATTERING of g-process longitudinal optical phonons at hotspots in silicon. *J. Appl. Phys.* 97 (2):023702.

Sondheimer, E. H. 1952. The mean free path of electrons in metals. *Adv. Phys.* 1 (1):1–42.

Stoner, R. J. and H. J. Maris. 1993. Kapitza conductance and heat flow between solids at temperatures from 50 to 300 K. *Phys. Rev. B* 48 (22):16373–87.

Su, L. T., J. E. Chung, D. A. Antoniadis, K. E. Goodson, and M. I. Flik. 1994. Measurement and modeling of self-heating in SOI nMOSFET's. *IEEE Trans. Electron Dev.* 41 (1):69–75.

Sverdrup, P. G., Y. S. Ju, and K. E. Goodson. 2000. Sub-continuum simulations of heat conduction in silicon-on-insulator transistors. *J. Heat Transfer* 123 (1):130–7.

Swartz, E. T. and R. O. Pohl. 1989. Thermal boundary resistance. *Rev. Mod. Phys.* 61 (3):605–68.

Toberer, E. S., L. L. Baranowski, and C. Dames. 2012. Advances in thermal conductivity. *Annu. Rev. Mater. Res.* 42 (1):179–209.

Topol, A. W., D. C. L. Tulipe, L. Shi, D. J. Frank, K. Bernstein, S. E. Steen, A. Kumar et al. 2006. Three-dimensional integrated circuits. *IBM J. Res. Dev.* 50 (4/5):491–506.

Wachutka, G. K. 1990. Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* 9 (11):1141–9.

Waldrop, M. M. 2016. The chips are down for Moore's Law. *Nature* 530 (7589):144–7.

Wang, Z. and N. Mingo. 2011. Absence of Casimir regime in two-dimensional nanoribbon phonon conduction. *Appl. Phys. Lett.* 99 (10):101903–13.

Ward, A. and D. A. Broido. 2011. Intrinsic phonon relaxation times from first-principles studies of the thermal conductivities of Si and Ge. *Phys. Rev. B* 81:085205.

Wilson, R. B. and D. G. Cahill. 2012. Experimental validation of the interfacial form of the Wiedemann-Franz Law. *Phys. Rev. Lett.* 108 (25):255901.

Zeng, T. and G. Chen. 2000. Phonon heat conduction in thin films: Impacts of thermal boundary resistance and internal heat generation. *J. Heat Transfer* 123 (2):340–7.

Zhou, F., A. L. Moore, J. Bolinsson, A. Persson, L. Fröberg, M. T. Pettes, H. Kong et al. 2011. Thermal conductivity of indium arsenide nanowires with wurtzite and zinc blende phases. *Phys. Rev. B* 83 (20):205416.

Ziman, J. M. 1960. *Electrons and Phonons*. Oxford: Oxford University Press.

4.22 ROLE OF COOLING IN ELECTRONICS RELIABILITY

Pradeep Lall

Introduction

The focus of this chapter is on cooling in electronic applications. The need and applicability of cooling to electronics has been discussed in detail with discussions on material limits, device performance parameters, and failure mechanisms. The fallacies with the current methodology for characterization of the functional relationship between temperature and electronic reliability have been discussed. It has been shown that material limits alone seldom justify the need for cooling. Temperature is more closely linked to degradation in performance. In many cases, just lowering temperature alone may or may not improve reliability—since a majority of the failure mechanisms are based on temperature cycling, spatial temperature gradient, and time-dependent temperature change. A methodology for evaluating the effects of temperature on reliability, determining the maximum operating temperature, and characterizing the functional relationship between temperature and electronic reliability has been discussed.

The cooling needs and constraints vary with the specific product–application requirement. The application areas have been broadly categorized into device level including PCB and IC packages and system level including space- and size-constrained and environment-constrained devices. The space- and size-constrained devices have been categorized into handheld communication devices such as smartphones and portable computing applications such as tablets and laptops. The environment-constrained devices include outdoor telecommunication enclosures and high-altitude electronics. For each electronic application, the temperature constraints and their basis have been discussed in detail. Typical material sets and heat transfer schemes and their relative effectiveness have also been discussed.

Application aspects of cooling in electronics have been discussed in detail. It is not the intent of this chapter to teach the reader basics of heat transfer—the reader is referred to previous chapters in the book or other basic texts on heat transfer for an in-depth treatment of the basics.

Understanding the Role of Temperature in Design

There has been a common belief that reliable electronics can be achieved by lowering temperature. Typically, statements such as “lower the temperature by 10°C and double reliability” are used to characterize the effect of temperature on reliability. The belief in the harmful effects of temperature has woven itself into today’s screening and thermal management processes. It is important to emphasize that cooling should be resorted to only when necessary. In this section, we will discuss some of the motivations for incorporating cooling in design and their validity.

The motivation for cooling electronics is based on three major premises: (1) material degrades or causes a fire hazard; (2) higher temperature activates failure mechanisms and decreases reliability; and (3) higher temperature degrades performance. Determination of requirements for cooling and design operation temperature is based on the reliability benefits, performance benefits, material limits, and auditability.

Material Limits

Materials typically used in electronics will withstand sustained operation at 125°C with the electronics operating in automotive underhood applications mounted on-engine or on-transmission expected to sustain exposure to high temperatures up to 175°C–200°C in the presence of vibration. The majority of the present-generation electronics products use lead-free solders. SnAgCu-based formulations including SAC105 and SAC305 and their variants are widely used, and have melting points in the neighborhood of 217°C–220°C. Typical peak reflow temperatures are in the

neighborhood of 230°C–240°C depending on the lead-free solder alloy formulation and the assembly. Based on historical precedence of products built using this profile over a number of product generations, it is fairly well known that the typical materials including molded-packages electronics, silicon circuits, discrete components, and printed-circuit boards (PCBs) do not undergo any permanent degradation when exposed to temperatures of up to 260°C. Table 4.22.1 shows the material limits for most of the common materials used in electronics. A brief review of the temperature limits indicates that material limits alone would seldom justify the need for cooling.

Impact on Performance

The parameters, which are sensitive to temperature, depend on device technology and are referred to as the temperature-sensitive electrical parameters (TSEPs). Table 4.22.2 shows the TSEP versus device technology.

The allowable junction temperature, T_J, is determined based on allowable values of the TSEP. The allowable variation in the values of the TSEP allows the device to function without significant degradation in performance. The reference temperature, T_R, is ordinarily measured with a thermocouple at the point of interest. The junction and reference temperature measured values are used to calculate the thermal resistance and the allowable power dissipation. The interested reader is referred to Lall et al. (1997) for a more detailed treatment of the Electrical Parameter Variations versus Temperature.

Impact on Reliability

The impact of temperature on reliability is typically the most misunderstood. Accelerated tests including burn-in are often run at high temperatures in hope that the high temperature will target the dominant failure mechanisms and reveal the suitability of a technology or design for a stress environment. It is often the case that the failure mechanisms precipitated in accelerated testing do not correspond to field life. Because of the nonlinearity of stress domain of the failure mechanisms—mechanisms accelerated by high-temperature accelerated stress may not exist in the equipment operating range of −55°C to 125°C. A stress domain is the set of values of stress magnitude over which the failure mechanism can be extrapolated in accordance with a single functional relationship. Linear stress domain signifies that the tendency of occurrence of any one failure mechanisms varies at a uniform rate with change in stress. Conversely, in the presence of a nonlinear domain, the accelerated test may not accurately predict how the device will fail during normal use. While high magnitudes of a stress may cause the device to fail instantly, low values of the same stress may not result in failure in any reasonably observable duration of time.

The naïve practitioner may often ignore the nonlinearity of the stress domain extrapolate the test results to the equipment operating temperature range. Further, the overall device, component, or assembly reliability is often incorrectly characterized by functional relationships such as the Arrhenius or Eyring models, and assigned an activation energy, and then used to extrapolate the test results to operating conditions. While individual failure mechanisms may be described using the Arrhenius or Eyring models, the device, component, and system's reliability dependence on temperature is more complex, often being determined by the dominant failure mechanisms and their dependence on temperature. One of the fairly obvious reasons for the continuous dependence on the use of a single activation energy to characterize the temperature dependence of a component is the imposing simplicity of extrapolation.

There are other circumstances in which the use of an activation energy to describe the temperature dependence of a device, component, or an assembly falls short of describing the true temperature dependence of electronic reliability. Several of the failure mechanisms may not be dependent on a steady-state temperature stress—but are instead driven by temperature cycling, spatial temperature gradient, and time-dependent temperature change. Such mechanisms have a minor or no dependence on the absolute value of temperature. Lowering temperature does not reduce the probability of failure by such mechanisms.

TABLE 4.22.1
Material Degradation Limits of Electronic Materials Based on Temperature Stress Alone

	Typical Material Sets	Material Degradation Limit due to Temperature Stress Only (°C)
	Integrated Circuits	
Mold compound	Epoxy	T_g = 155°C –175°C; T_{max} = 215°C –245°C
	Silicone Room-Temperature-Vulcanization (RTV), lead encapsulation	T_{max} = 260°C; T_{min} = –54°C
Wire bond	Gold	T_m = 1064.2°C
	Copper	T_m = 1084.7°C
	Aluminum	T_m = 660.31°C
Flip-chip	95Pb/5Sn	T_m = 308–312°C
	Electro-deposited Nickel	T_m = 1455°C
TAB	Polymide	T_g = 332°C –344°C; T_{max} = 344°C –360°C
	Copper metallization	T_m = 1084.7°C
Interposer substrate	BT-epoxy	T_g = 220–230°C; T_{max} = 150–170°C
	Elastomers	T_{max} = 300°C; T_{min} = –120°C
	Polymide	T_g = 332–344°C; T_{max} = 344–360°C
Package I/O	Alloy 42 (lead frame)	T_m = 1427°C
	Copper (lead frame)	T_m = 1084.7°C
	Sn1Ag0.5Cu solder	T_m = 216.84 ± 0.5°C
	Sn3Ag0.5Cu solder	T_m = 217–218°C
	63Sn/37Pb solder	T_m = 183°C
Metallization	Copper	T_m = 1084.7°C
	Aluminum	T_m = 660.31°C
Chip	Silicon	T_{max} = 300°C –350°C; T_m = 1414°C
	Gallium arsenide	T_m = 1283°C
	Discretes	
Spring contacts	Beryllium copper	T_m = 1587°C –1750°C
	Phosphor bronze	T_m = 900°C –1050°C
Capacitors, resistors, inductors	Ceramic (BeO)	T_m = 2527°C
	Ceramic (Al_2O_3)	T_{max} = 1000°C
	Copper	T_m = 1084.7°C
	Polymide	T_g = 332°C–344°C; T_{max} = 344°C–360°C
	Assembly Level	
Interconnections	Sn1Ag0.5Cu solder	T_m = 216.84 ± 0.5°C
	Sn3Ag0.5Cu solder	T_m = 217°C–218°C
	Solder joints (63Sn/37Pb)	T_m = 183°C
	Copper metallization	T_m = 1084.7°C
PCB	FR-4	T_g = 130°C–140°C
	High T_g laminate materials	T_g = 175°C–180°C
Plastic casings	ABS	T_{max} = 85°C
	PC/ABS	T_{max} = 90°C

T_g, glass transition temperature; T_{max}, maximum continuous service temperature; T_m, melting temperature.

TABLE 4.22.2
TSEP versus Device Technology

Device Technology	TSEP
MOSFET	V_{th}, $V_{DS(on)}$, $R_{DS(on)}$
IGBT	V_{th}, $V_{CE(s)}$
Bipolar	V_{BE}, $V_{CE(s)}$
Diode	V_F

Table 4.22.3 shows a summary of the variation in the temperature dependence of typical failure mechanisms in the temperature range of –55°C to 400°C (Lall et al. 1997). The interested reader is referred to Lall et al. (1997) for an in-depth treatment on the subject.

Thermal Characteristics of PCBs

Conventional PCB is typically a layered composite consisting of copper foil and a glass-reinforced polymer (FR-4). Figures 4.22.1 and 4.22.2 show the cross-section of four-layer PCB with metal layers and plated-through holes and blind vias.

Figure 4.22.3 shows the mathematical equivalent of the PCB cross section. Azar and Graebner (1996) and Graebner (1995) have shown that for many thermal calculations, it is convenient to treat the PCB as a homogeneous material, with two different effective thermal conductivities—one describing heat flow within the plane (in-plane) and another for heat flow through the thickness of the plane (through). The equations for calculating each of these conductivities, given the values for the thickness and thermal conductivity of each layer, are presented:

$$K_{\text{in-plane}} = \frac{\sum_{i=1}^{N} K_i t_i}{\sum_{i=1}^{N} t_i}, \tag{4.22.1}$$

$$K_{\text{through}} = \frac{\sum_{i=1}^{N} t_i}{\sum_{i=1}^{N} \frac{t_i}{K_i}}, \tag{4.22.2}$$

where K is the conductivity in (W/mK) and t is the thickness in (m). The subscripts in-plane, through, and I indicate the in-plane, through, and the individual-layer components, respectively. Equations 4.22.1 and 4.22.2 assume complete copper planes, which is typically not the case in most designs

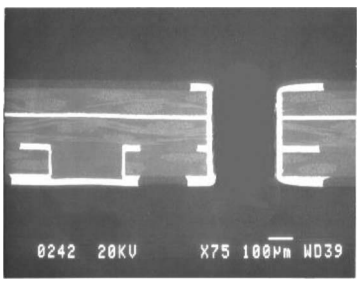


FIGURE 4.22.1 Cross-section of a four-layer PCB. (Courtesy of Motorola).

TABLE 4.22.3
Dominant Temperature Dependency over Steady-State Temperatures from –55°C to 500°C

		Temperature		
		–55°C < *T* <150°C	**150°C < *T* < 400°C**	***T* > 400°C**
Package Element	**Mechanism**	**Nature of Temperature Dependence**	**Nature of Temperature Dependence**	**Nature of Temperature Dependence**
		Assembly		
Solder joints	Thermal fatigue	ΔT	ΔT 150°C < *T* < 217°C None, for *T* > 217°C, SnAgCu solders melt 150°C < *T* < 183°C None, for *T* > 183°C, 62Sn/36Pb/2Ag solder melts	None For *T* > 217°C, SnAgCu solders melt For *T* > 183°C, 62Sn/36Pb/2Ag solder melts
	Creep–rupture	σ, *T* Dominant dependence on stress, σ	σ, *T* 150°C < *T* < 183°C (62Sn36Pb2Ag) 150°C < *T* < 217°C (SnAgCu) Dominant dependence on stress, σ None, for *T* > 217°C, SnAgCu solder melts For *T* > 183°C, 62Sn/36Pb/2Ag solder melts	None For *T* > 217°C, SnAgCu solders melt For *T* > 183°C, 62Sn/36Pb/2Ag solder melts
	Intermetallic growth	*T*, *t*	*T*, *t*	*T*, *t*
Underfill	Delamination	ΔT	ΔT	ΔT
PCB	Plated-through holes	ΔT	ΔT ΔT dependence accelerated for $T > T_g$	ΔT ΔT dependence accelerated for $T > T_g$
	High-density interconnect vias	ΔT	ΔT ΔT dependence accelerated for $T > T_g$	ΔT ΔT dependence accelerated for $T > T_g$
	Conductive filament formation	Moisture, voltage, *T*	Moisture, voltage, *T*	Moisture, voltage, *T*
		IC Package		
Wire	Flexure fatigue	ΔT	ΔT	ΔT
Wire bond	Shear fatigue	ΔT	ΔT	ΔT
	Kirkendall voiding	Temperature independent	*T* dependent for *T* > 150°C	*T* dependent

(*Continued*)

TABLE 4.22.3 (*Continued*)
Dominant Temperature Dependency over Steady-State Temperatures from –55°C to 500°C

		Temperature		
		–55°C < *T* <150°C	**150°C < *T* < 400°C**	***T* > 400°C**
Package Element	**Mechanism**	**Nature of Temperature Dependence**	**Nature of Temperature Dependence**	**Nature of Temperature Dependence**
		IC Package		
Beam lead (Elastomer-on-Flex Packages)	Flexure fatigue	ΔT	ΔT	ΔT
Underfill	Delamination at chip-underfill interface	ΔT	ΔT	ΔT
Die	Die fracture	ΔT	ΔT	ΔT
	Electrical overstress	Temperature independent	*T* dependent for *T* > 160°C	*T* dependent
Die adhesive	Die adhesive fatigue	ΔT	ΔT	ΔT
Encapsulant	Encapsulant reversion (plastic package only)	Temperature independent	*T* dependent for *T* > 300°C	*T* dependent
	Encapsulant cracking (plastic package only)	ΔT, $\partial T/\partial t$ (³25°C/s)	*T* dependent (for *T* > 215°C), D*T*, $\partial T/\partial t$, (³25°C/s)	*T* dependent ΔT $\partial T/\partial t$
Package	Stress corrosion	$\partial T/\partial t$	$\partial T/\partial t$	$\partial T/\partial t$
Die metallization	Corrosion	Mildly *T* dependent, $\partial T/\partial t$	$\partial T/\partial t$	$\partial T/\partial t$
	Electromigration	Structural nonuniformity dependent, ∇T	*T* dependent (for *T* > 150°C), ∇T dependent	*T* dependent (for *T* > 150°C), ΔT dependent
	Hillock formation	ΔT	ΔT	*T* dependent (for *T* > 400°C)
	Metallization migration	Temperature independent	Temperature independent	*T* dependent (for *T* > 500°C)
	Contact spiking	Temperature independent	Temperature independent	*T* dependent (for *T* > 400°C)
	Constraint cavitation	*T* dependent	T^{-1} dependent	T^{-1} dependent
	Slow trapping	Temperature independent	*T* dependent (for *T* > 175°C)	*T* dependent
	Electrostatic discharge	Temperature independent (in presence of protection circuits)	Temperature independent (in presence of protection circuits)	Temperature independent (in presence of protection circuits)
	Time-dependent dielectric breakdown	Voltage dependent, weak *T* dependence	Voltage dependent, weak *T* dependence	Voltage dependent, weak *T* dependence

(*Continued*)

TABLE 4.22.3 (*Continued*)
Dominant Temperature Dependency over Steady-State Temperatures from –55°C to 500°C

		Temperature		
		–55°C < *T* <150°C	**150°C < *T* < 400°C**	***T* > 400°C**
Package Element	**Mechanism**	**Nature of Temperature Dependence**	**Nature of Temperature Dependence**	**Nature of Temperature Dependence**
		IC Package		
Device	Ionic contamination	T^{-1} dependence (device not operational)	T^{-1} dependence (device not operational)	T^{-1} dependence (device not operational)
	Forward second breakdown	Temperature independent	*T* dependent (for *T* > 160°C)	*T* dependent
	Reverse second breakdown	Mild T^{-1} dependence	Mild T^{-1} dependence	Mild T^{-1} dependence
	Surface charge spreading	Temperature independent	*T* dependent (for *T* > 150°C)	*T* dependent (for *T* > 150°C)
Device/oxide interface	Hot electrons	T^{-1} dependent	Temperature independent	Temperature independent

Temperature dependency includes T (steady-state temperature), ΔT (temperature cycle), ΔT (temperature gradient), ∂T/∂t (time-dependent temperature change) (Lall et al. 1997).

(except for layers with ground planes). The nonground plane layers are accounted for calculating the fractional coverage of copper by

$$K_i = f_i K_{\mathrm{Cu}}, \tag{4.22.3}$$

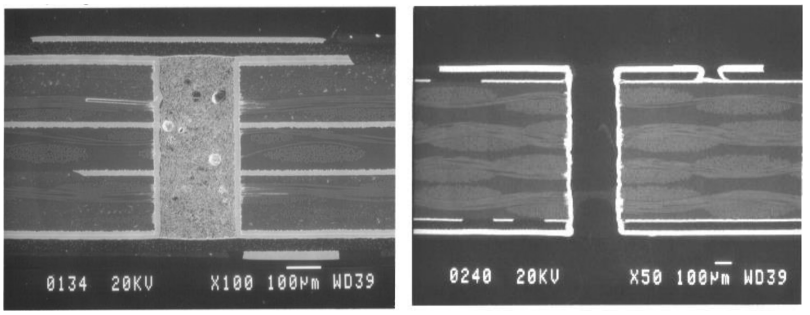


FIGURE 4.22.2 Cross-section of high-density interconnect PCBs with and without filled vias. (Courtesy of Motorola).

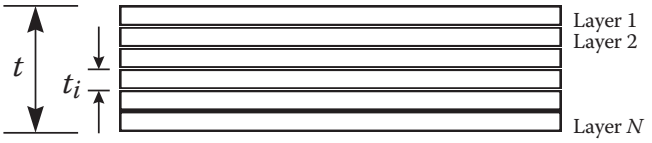


FIGURE 4.22.3 Mathematical equivalent of the PCB cross-section.

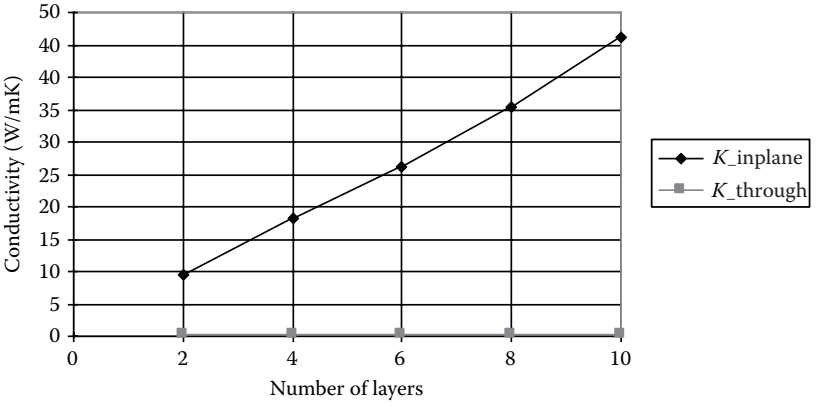


FIGURE 4.22.4 Conductivity versus number of layers for a 55 miles thick FR-4 PCB with 1/2 oz/ft^2 copper on the internal layers.

where f_i is the fractional coverage of copper for layer i. Figure 4.22.4 shows the in-plane and through conductivity for a 55 mils thick 2-layer, 4-layer, 6-layer, 8-layer, and 10-layer PCB with 1/2 oz/ft^2. Copper on the internal layers (assuming that the layers consist only of copper and FR-4, i.e., no vias) where the thermal conductivity for copper is 390 W/mK and thermal conductivity for FR4 is 0.25 W/mK. Figure 4.22.4 indicates that the $K_{\text{in-plane}}$ is two orders of magnitude higher than K_{through}. Heat flow through the copper layers dominates the heat transfer because of the large differential in the thermal conductivities of the copper and FR-4 layers. For the same reason, a typical PCB is not an efficient conductor of heat, leading to significant local variations in the PCB temperature in the vicinity of heat-generating components. Thermal conduction is not very efficient for heat flow in a direction lacking a continuous copper path.

Thermal Characteristics of Electronic Packages

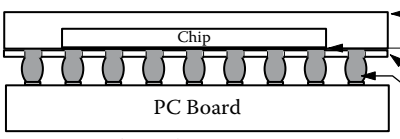
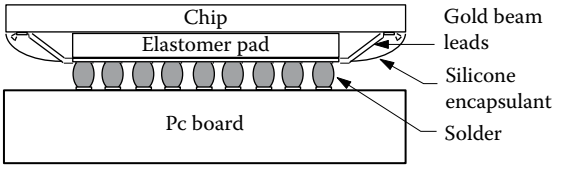
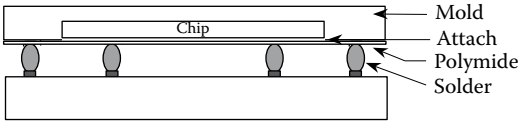
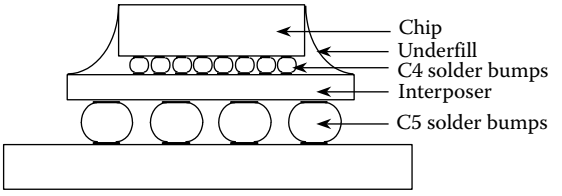
Package Architecture

Most electronic packages consist of a silicon chip bonded to an interposer or a lead frame through wire bonds or controlled collapse solder joints. The package I/O is either solder balls on the substrate or leads on the lead frame. Table 4.22.4 shows schematic of package cross section and package characteristics for common chip-scale package families including (1) Flex-Substrate BGA, (2) Elastomer-on-Flex Package, (3) Partial-Array Chip-Scale Package, and (4) Rigid or Flex Interposer Chip-Scale Package. Other conventional design includes Thin Small Outline Package (TSOP), Quad Flat Pack (QFP), Small Outline Integrated Circuit (SOIC), etc. The interested reader is referred to the IPC-SM-782 for a more detailed description of the standard package types.

Materials and Properties

Table 4.22.5 shows the typical material sets used in integrated circuit packages and their thermal conductivities. Die attach bonding is used to mechanically attach the die to the lead frame and also to provide a heat flow path to the lead frame. A typical adhesive is a thermally conductive epoxy composite with a thermal conductivity of 0.8 W/mK. The epoxy adhesive layer is typically 25 μm thick. Wire bonding and flip-chip are the commonly used techniques for the first-level interconnection in the plastic package interconnection technology. Traditionally, the gold wire has been the preferred material on account of its high oxidation resistance and low work hardening characteristics. Recent advances have resulted in the use of copper interconnects for wire bonding. The migration to copper interconnects has been triggered by the higher cost of gold wire in comparison with copper.

TABLE 4.22.4
Schematic of Package Cross-Section and Characteristics for the Common Chip-Sale Packages

Package Schematic	Characteristics
Package Schematic  Flex-Substrate BGA	*Flex-Substrate BGA* is a ball-grid array format chip-scale package. The chip is wire bonded to the single-sided flex-polymide substrate. Package I/Os are solder balls, which go-through holes in the substrate and connect to the solder pads on the chip side of the polymide substrate
 Elastomer-on-flex BGA	*Elastomer-on-Flex Package* involves a chip flipped on to an elastomeric substrate. All the chip I/Os are routed to a single-sided flexible circuit, which is bonded to the other side of the elastomeric substrate by beam leads (gold or copper 18 μm thick). The beam leads are extensions of the printed circuit on the flexible circuit, which are connected to solder balls
 Partial array BGA	*Partial-Array Chip-Scale Package* is a ball-grid array format chip-scale package. The chip is connected through wire bond interconnects to a multilayer organic substrate. Package I/Os are solder balls, which go-through holes in the substrate and connect to the solder pads on the chip side of the polymide substrate
 Rigid or flex interposer package	*Rigid or Flex-Interposer Chip-Scale Package* The rigid or flex interposer packages consists of a solder-bumped integrated circuit (IC), which is flip-chip bonded to an (rigid or flex) interposer substrate—approximately 8 mils thick—and then underfilled with an encapsulant. Chip I/Os are routed to package I/Os through the interposer substrate to solder bumps attached to the bottom side of the interposer substrate

The thermal conductivity of the gold wire is 319 W/mK while that of copper is in the neighborhood of 368–400 W/mK. The majority of electronic mold compounds (EMCs) in plastic packages use epoxies with fillers made of silica, alumina, aluminum–nitride particles. The fillers in the molding compound are intended to improve its mechanical strength, thermal conductivity, adhesion to package components, chemical resistance, electrical resistance, coefficient of expansion match, and thermal and moisture resistance in the operational temperature range. The thermal conductivity of a typical EMC is 0.67 W/mK with a glass transition temperature of 150°C. EMCs, in addition to providing protection for the silicon chip, are required to meet U-94, plastic flammability standards. In order to meet these conditions, flame retardants are added into molding compounds, which at high temperatures can work as catalyst and accelerate process of failure of bond wires. Commercially available PEMs are able to operate in the neighborhood of 175°C. Thermal degradation causes major changes in physical and chemical properties of the molding compound, which can severely affect the reliability of electronic devices. Prolonged exposure to very high-temperature results in degradation of the binding material in molding compound accompanied with oxidation producing a gradual loss of weight. Figures 4.22.5 and 4.22.6 show plots of percentage weight loss and aging duration, for mold compounds A, B, C, D for thermal aging at 200°C and 250°C.

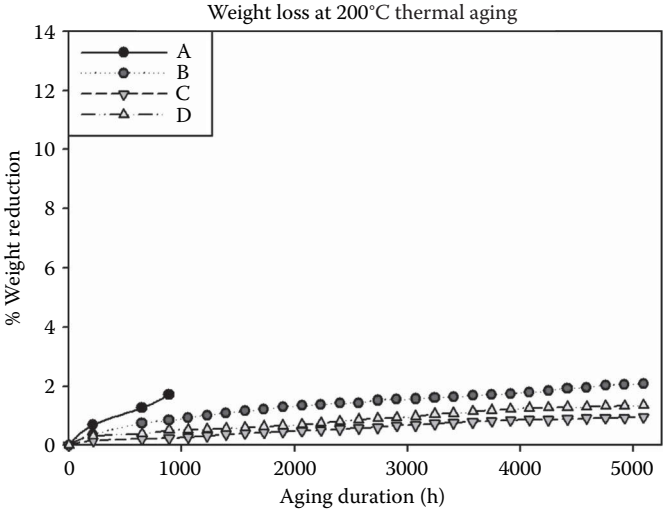


FIGURE 4.22.5 Weight loss of epoxy mold compounds A, B, C, and D at 200°C. (From Lall, P. et al., Degradation mechanisms in electronic mold compounds subjected to high temperature in neighborhood of 200°C, *Electronic Components and Technology Conference, 64th ECTC*, pp. 242–254, Orlando, FL, May 27–30, 2014.)

Mold compound-A is a high-temperature epoxy molding compound capable of low warpage and excellent reflow crack resistance. Mold compound-B is a higher temperature version of the mold compound-A. Mold compound-C is a multiaromatic epoxy with multiaromatic hardener and low-alpha filler with a filler content of 85%. Mold compound-D is a silicone encapsulation material capable of high electrical resistivity, high thermal and mechanical stability. Silicone-based encapsulants (EMC-D) exhibit the best high-temperature performance in the neighborhood of 200°C–250°C in comparison with high-temperature epoxies and multiaromatic epoxy with multiaromatic hardeners.

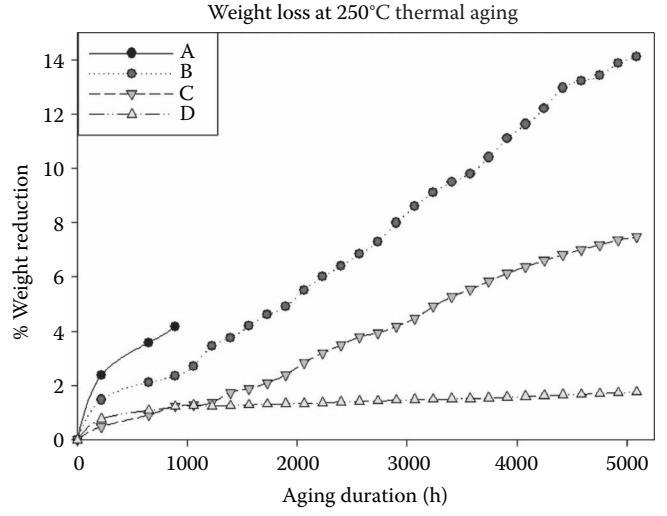


FIGURE 4.22.6 Weight loss of epoxy mold compounds A, B, C, and D at 250°C. (From Lall, P. et al., Degradation mechanisms in electronic mold compounds subjected to high temperature in neighborhood of 200°C, *Electronic Components and Technology Conference, 64th ECTC*, pp. 242–254, Orlando, FL, May 27–30, 2014.)

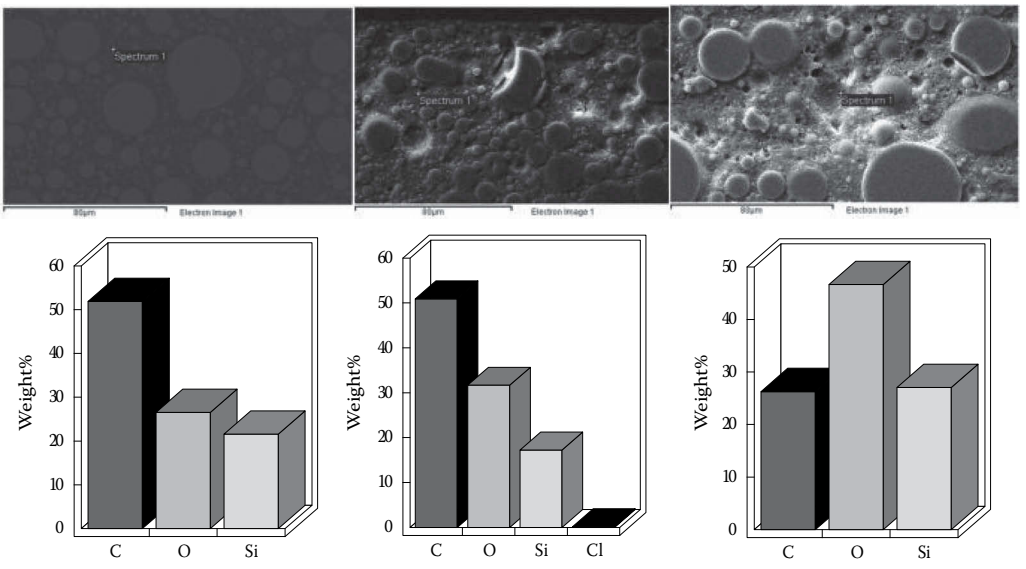


FIGURE 4.22.7 Molding compound "A", 200°C. (From Lall, P. et al., Degradation mechanisms in electronic mold compounds subjected to high temperature in neighborhood of 200°C, *Electronic Components and Technology Conference, 64th ECTC*, pp. 242–254, Orlando, FL, May 27–30, 2014.)

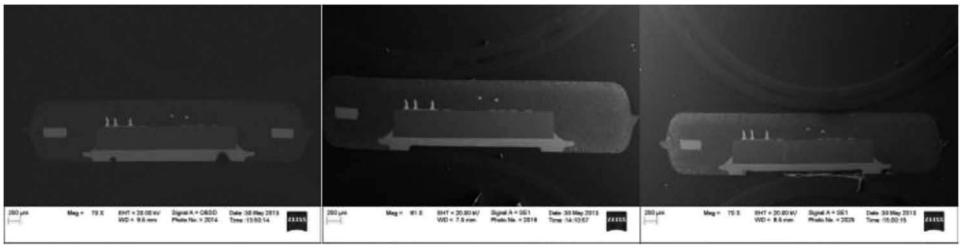


FIGURE 4.22.8 Molding compound "A", 250°C. (From Lall, P. et al., Degradation mechanisms in electronic mold compounds subjected to high temperature in neighborhood of 200°C, *Electronic Components and Technology Conference, 64th ECTC*, pp. 242–254, Orlando, FL, May 27–30, 2014.)

The high-temperature epoxy mold compounds (EMC-A) exhibit the poorest performance in comparison with multiaromatic epoxy with multiaromatic hardner and silicone-based encapsulants.

Lall et al. (2014) have studied the high-temperature survivability of a number of mold compound families at 200°C and 250°C of sustained operation. High-temperature epoxy molding compound-A exhibits rapid growth of oxidation layer was shown in Figures 4.22.7 and 4.22.8. Exposure to 200°C aging oxides is majority of the EMC-A volume (indicated by the light gray region), and exposure to 250°C results in complete oxidation of the EMC-A after 2000 h of aging. Mold compound-A exhibited the highest rate of weight loss and mold compound oxidation. EMC-A became very brittle after 2000 h, which its handling was a major issue. Mold compound cracked easily and broke during normal handling. EDX analysis was performed on white oxidation layer in the mold compound revealed higher oxygen content and lower carbon content in degraded layer as compared to time 0 samples. Loss of carbon is because of evaporation of products of decomposition of long polymer chains. Higher oxygen content shows that some of the products are oxidized. Results are shown in Figure 4.22.9.

Multiaromatic epoxy with multiaromatic hardener EMC-C exhibited an oxidation layer after 2000 h of aging at 200°C and 1000 h of aging at 250°C, shown in Figures 4.22.10 and 4.22.11. Oxidation layer is at along edges of the package, and from there it propagates toward the center of

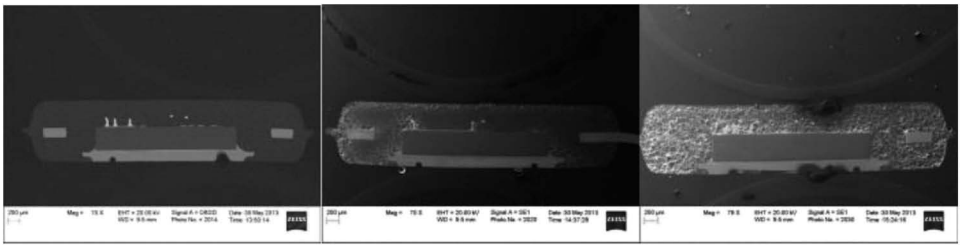


FIGURE 4.22.9 EDX analysis of EMC candidate-A—(a) time 0, (b) intermediate oxidation, and (c) complete oxidation. (From Lall, P. et al., Degradation mechanisms in electronic mold compounds subjected to high temperature in neighborhood of 200°C, *Electronic Components and Technology Conference, 64th ECTC*, pp. 242–254, Orlando, FL, May 27–30, 2014.)

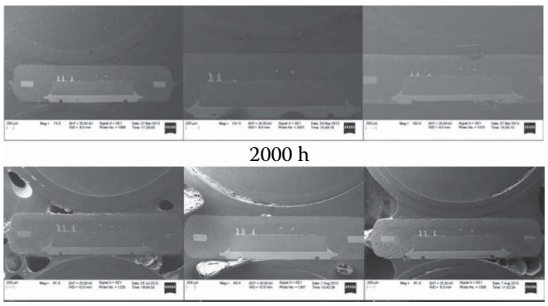


FIGURE 4.22.10 Molding compound "C", 200°C. (From Lall, P. et al., Degradation mechanisms in electronic mold compounds subjected to high temperature in neighborhood of 200°C, *Electronic Components and Technology Conference, 64th ECTC*, pp. 242–254, Orlando, FL, May 27–30, 2014.)

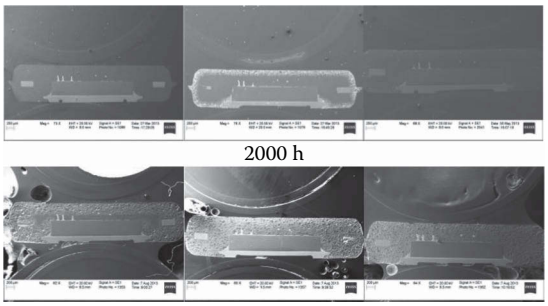


FIGURE 4.22.11 Molding compound "C", 250°C. (From Lall, P. et al., Degradation mechanisms in electronic mold compounds subjected to high temperature in neighborhood of 200°C, *Electronic Components and Technology Conference, 64th ECTC*, pp. 242–254, Orlando, FL, May 27–30, 2014.)

the package. For 200°C case, a large cross section of the mold compound was oxidized after 5000 h, and at 250°C, complete cross section area was oxidized. Silicone encapsulation material, EMC-D, results were very different and are shown in Figures 4.22.12 and 4.22.13. There was no oxidation area observed at edges of the package in this case; but there were oxidation spots spread throughout the cross-section in all cases. The density of the oxidation spots increased for the aged samples, but even after aging for 5000 h, however most of the area did not exhibit oxidation.

The predominant semiconductor material in the electronics industry today is silicon. The thermal conductivity of silicon is temperature dependent. Figure 4.22.14 shows the thermal conductivity of silicon versus temperature in the range of 75°C–150°C.

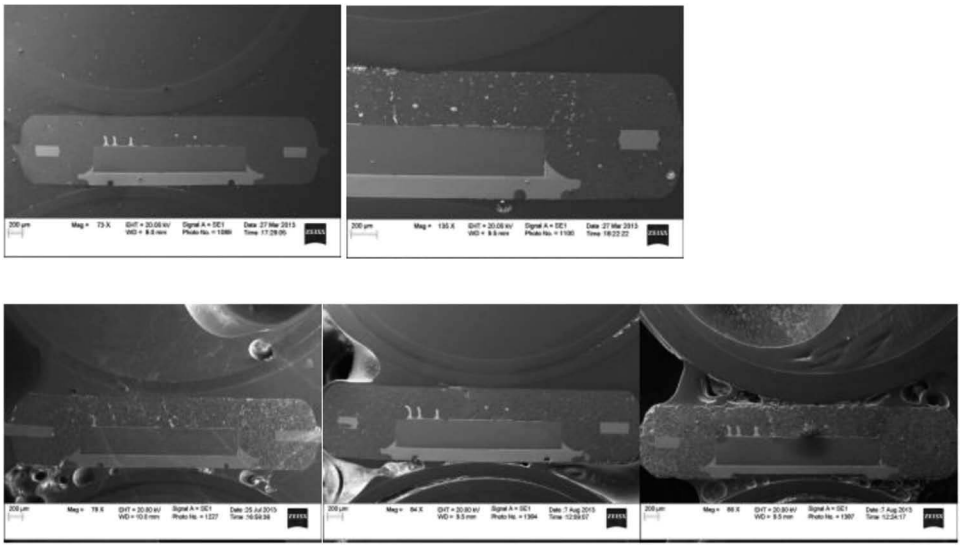


FIGURE 4.22.12 Molding compound "D," 200°C.

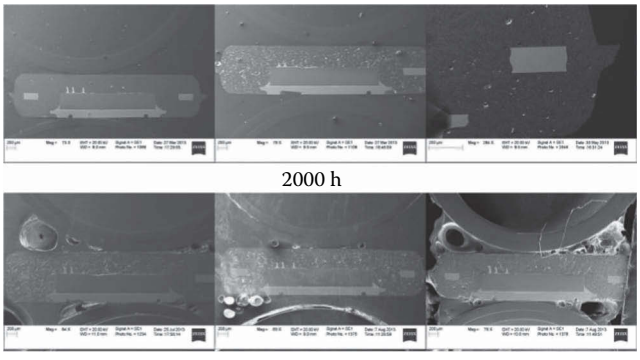


FIGURE 4.22.13 Molding compound "D," 250°C.

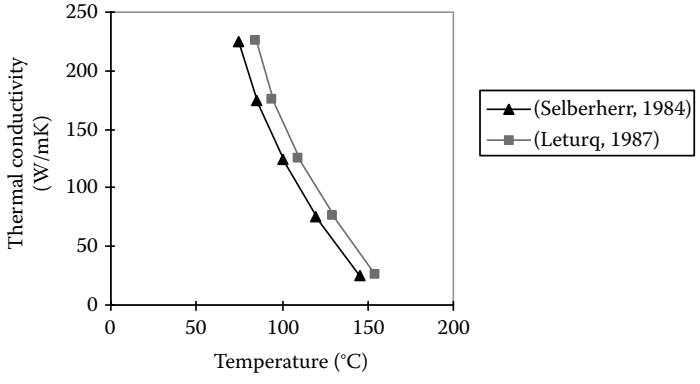


FIGURE 4.22.14 Thermal conductivity of silicon.

TABLE 4.22.5
Thermal Conductivity of Typical Packaging Materials

Package Element	Thermal Conductivity (W/mK)
Molding Compound	
Epoxies and silicones	0.36–2.09
Die Attach	
Epoxy	0.8
Solder Joints	
63Sn/37Pb	50
Lead Frame	
Alloy42	14.7
Cu–Zr	380
Cu–Fe	200–260
Cu–Mg	344
Wire bonds	
Gold	319
Copper	390

TABLE 4.22.6
Effect of Temperature on Elastic Modulus of SAC105 at Strain Rate of 10 s^{-1}

	SAC105, 10 s^{-1}, *E* (GPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	32.5	32	31.7	31.4	31.3
30	31.6	30.75	30.25	29.87	29.7
60	31.33	30.38	29.9	29.6	29.44

TABLE 4.22.7
Effect of Temperature on Elastic Modulus of SAC105 at Strain Rate of 35 s^{-1}

	SAC105, 35 s^{-1}, *E* (GPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	33.7	33.2	32.9	32.6	32.5
30	32.85	31.9	31.3	30.95	30.8
60	32.49	31.53	31	30.7	30.57

Effect of Aging on the Material Properties of Lead-Free Solders

Material properties for both SAC105 and SAC305 lead-free solder alloys are sensitive to the aging temperatures for a given period of time. Properties degrade nonlinearly under exposure to temperatures between 25°C and 125°C for times of 1 day to 61 days (Figures 4.22.15 through 4.22.30; Tables 4.22.6 through 4.22.21). The 1-day aged samples exhibit a higher elastic modulus and higher ultimate tensile strength compared to solder samples aged at the elevated temperature in each case.

TABLE 4.22.8
Effect of Temperature on Elastic Modulus of SAC105 at Strain Rate of 50 s^{-1}

	SAC105, 50 s^{-1}, *E* (GPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	34.2	33.7	33.4	33.1	33
30	33.27	32.5	31.95	31.55	31.35
60	32.97	32	31.5	31.15	31.01

TABLE 4.22.9
Effect of Temperature on Elastic Modulus of SAC105 at Strain Rate of 75 s^{-1}

	SAC105, 75 s^{-1}, *E* (GPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	36	35.5	35.2	34.9	34.8
30	35	34	33.5	33.2	32.95
60	34.7	33.7	33.2	32.9	32.7

TABLE 4.22.10
Effect of Temperature on Ultimate Tensile Strength (UTS) of SAC105 at Strain Rate of 10 s^{-1}

	SAC105, 10 s^{-1}, UTS (MPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	52	50	49	48.4	48
30	49.8	47.4	46.4	45.6	45.2
60	48.93	46.8	45.74	45.2	44.69

TABLE 4.22.11
Effect of Temperature on UTS of SAC105 at Strain Rate of 35 s^{-1}

	SAC105, 35 s^{-1}, UTS (MPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	65.2	63.2	62.2	61.6	61.2
30	62.5	60	58.6	58	57.5
60	61.35	59.16	58	57.4	56.98

Samples exposed to 60 days of isothermal aging exhibited the lowest elastic modulus and ultimate tensile strength of all the cases tested. Also, a special case where the rate of cooling effects (reflowed vs water quench) on material properties of SAC Alloys were determined for 1 day aging at 25°C (Lall et al. 2013).

Understanding Basic Semiconductor Thermal Parameters

Key parameters used to describe the thermal capability of a semiconductor device include (1) T_J, junction temperature; (2) T_C, case temperature; (3) T_A, ambient temperature; (4) TSP, temperature-sensitive

TABLE 4.22.12
Effect of Temperature on UTS of SAC105 at Strain Rate of 50 s^{-1}

	SAC105, 50 s^{-1}, UTS (MPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	71	69	68	67.4	67
30	66.75	65.3	64.4	63.8	63.5
60	65.1	63.74	63	62.6	62.38

TABLE 4.22.13
Effect of Temperature on UTS of SAC105 at Strain Rate of 75 s^{-1}

	SAC105, 75 s^{-1}, UTS (MPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	81.8	79.8	78.8	78.2	77.8
30	78.2	75.7	74.5	73.7	73.2
60	76.97	74.7	73.56	72.9	72.43

TABLE 4.22.14
Effect of Temperature on Elastic Modulus of SAC305 at Strain Rate of 10 s^{-1}

	SAC305, 10 s^{-1}, E (GPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	40.2	39.8	39.45	39	38.7
30	38.9	38	37.4	37.1	36.85
60	38.51	37.48	36.97	36.71	36.46

TABLE 4.22.15
Effect of Temperature on Elastic Modulus of SAC305 at Strain Rate of 35 s^{-1}

	SAC305, 35 s^{-1}, E (GPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	42.09	41.71	41.35	40.85	40.6
30	40.75	39.9	39.4	39.05	38.8
60	40.32	39.28	38.76	38.5	38.3

TABLE 4.22.16
Effect of Temperature on Elastic Modulus of SAC305 at Strain Rate of 50 s^{-1}

	SAC305, 50 s^{-1}, E (GPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	43.06	42.69	42.3	41.85	41.56
30	41.75	41	40.5	40.15	39.9
60	41.25	40.19	39.67	39.41	39.15

TABLE 4.22.17
Effect of Temperature on Elastic Modulus of SAC305 at Strain Rate of 75 s^{-1}

	SAC305, 75 s^{-1}, E (GPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	45	44.62	44.25	43.8	43.5
30	43.5	42.7	42.25	41.95	41.7
60	43.11	42.04	41.51	41.25	40.98

TABLE 4.22.18
Effect of Temperature on UTS of SAC305 at Strain Rate of 10 s^{-1}

	SAC305, 10 s^{-1}, UTS (MPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	55.9	54.6	53	51	50
30	53	50.75	49.2	48.1	47.4
60	52.2	49.35	47.93	47.21	46.5

TABLE 4.22.19
Effect of Temperature on UTS of SAC305 at Strain Rate of 35 s^{-1}

	SAC305, 35 s^{-1}, UTS (MPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	71	69.5	67.6	66	65
30	67.5	64.9	63.25	62.25	61.5
60	66.17	63.31	61.88	61.17	60.45

TABLE 4.22.20
Effect of Temperature on UTS of SAC305 at Strain Rate of 50 s^{-1}

	SAC305, 50 s^{-1}, UTS (MPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	78.3	76.9	75.5	74	72.8
30	74.2	72	70.85	69.9	69.25
60	72.98	70.34	69.02	68.36	67.7

TABLE 4.22.21
Effect of Temperature on UTS of SAC305 at Strain Rate of 75 s^{-1}

	SAC305, 75 s^{-1}, UTS (MPa), Aging Temperatures				
Aging (days)	**25°C**	**50°C**	**75°C**	**100°C**	**125°C**
1	91	89.5	87.8	85.9	85
30	86.25	84.25	82.7	81.55	80.75
60	84.81	81.93	80.49	79.77	79.05

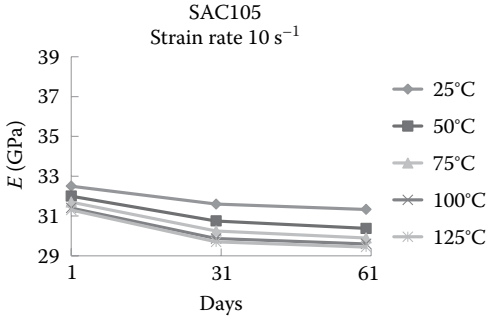


FIGURE 4.22.15 Effect of temperature on elastic modulus of SAC105 at strain rate of 10 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

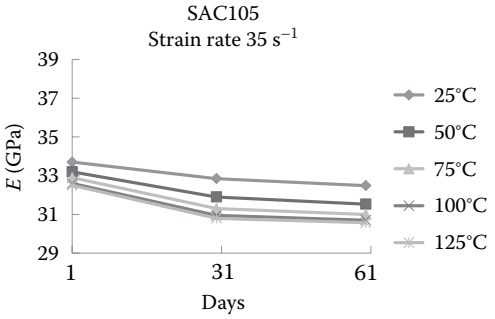


FIGURE 4.22.16 Effect of temperature on elastic modulus of SAC105 at strain rate of 35 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

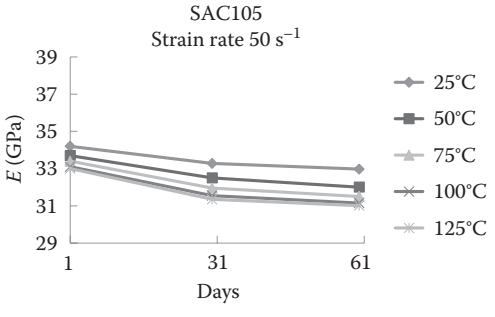


FIGURE 4.22.17 Effect of temperature on elastic modulus of SAC105 at strain rate of 50 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

parameter; (5) T_R, reference temperature (i.e., case or ambient); (6) $R_{\theta JR}$, junction-to-reference thermal resistance; (7) $R_{\theta JC}$, junction-to-case thermal resistance; (8) $R_{\theta JA}$, junction-to-ambient thermal resistance; (9) $R_{\theta JR(t),}$ junction-to-reference transient thermal resistance; (10) PD, power dissipation. The parameters describe the steady-state thermal capability of a power semiconductor device. The thermal behavior of a device is often represented by an electrical equivalent circuit. This circuit

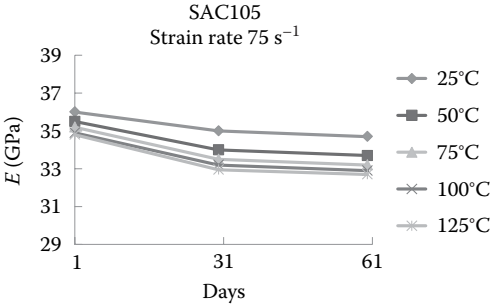


FIGURE 4.22.18 Effect of temperature on elastic modulus of SAC105 at strain rate of 75 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

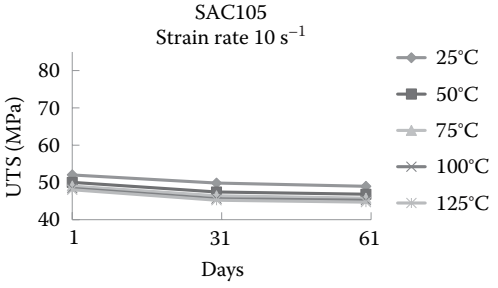


FIGURE 4.22.19 Effect of temperature on ultimate tensile strength of SAC105 at strain rate of 10 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

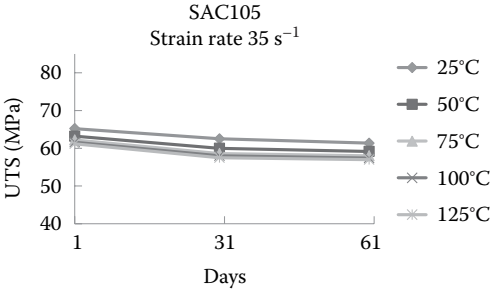


FIGURE 4.22.20 Effect of temperature on ultimate tensile strength of SAC105 at strain rate of 35 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

consists of a resistor–capacitor network as shown in Figure 4.22.31. Resistors R_1, R_2, and R_3 are all analogous to individual thermal resistances.

Resistor R_1 is the thermal resistance from the device's junction to its die bond. Resistor R_2 is the thermal resistance from the die bond to the device's case. Resistor R_3 is the thermal resistance from the device's case to ambient. The thermal resistance from the junction to some reference point is equal to the sum of the individual resistors between the two points. For instance, the thermal resistance $R\theta_{JC}$ from junction-to-case is equal to the sum of resistors R_1 and R_2. The thermal resistance

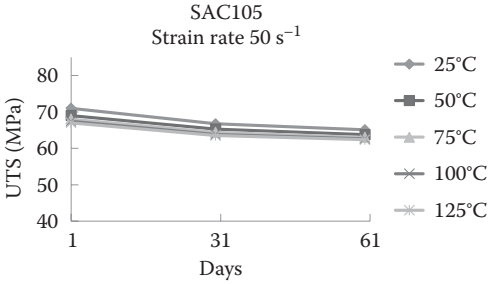


FIGURE 4.22.21 Effect of temperature on ultimate tensile strength of SAC105 at strain rate of 50 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

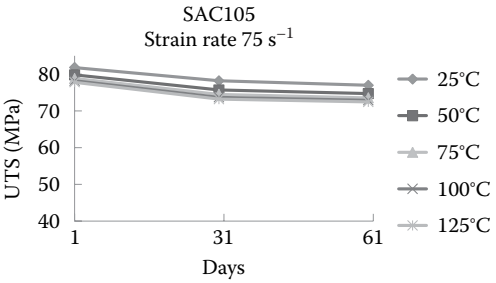


FIGURE 4.22.22 Effect of temperature on ultimate tensile strength of SAC105 at strain rate of 75 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

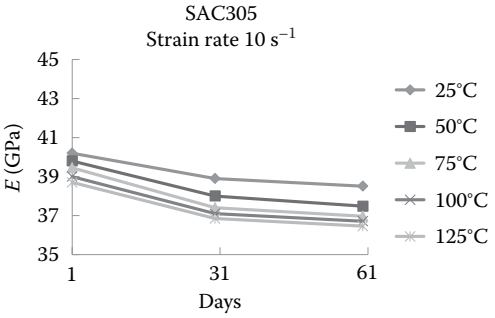


FIGURE 4.22.23 Effect of temperature on elastic modulus of SAC305 at strain rate of 10 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

$R\theta_{JA}$ from junction-to-ambient, therefore, is equal to the sum of resistors R_1, R_2, and R_3. The capacitors model the transient thermal response of the circuit. When heat is instantaneously applied and or generated, there is a charging effect that takes place. This response follows an RC time constant determined by the resistor–capacitor thermal network. Thermal resistance, at a given time, is called transient thermal resistance, $R\theta_{JR(t)}$.

$$R_{\theta JR} = \frac{T_J - T_R}{P_D}, \tag{4.22.4}$$

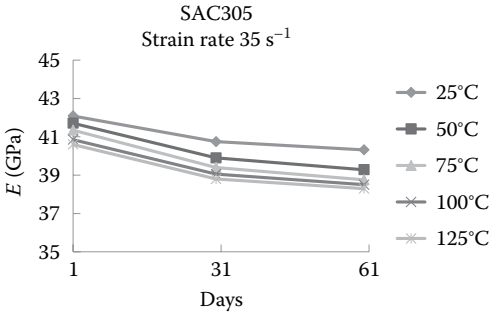


FIGURE 4.22.24 Effect of temperature on elastic modulus of SAC305 at strain rate of 35 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

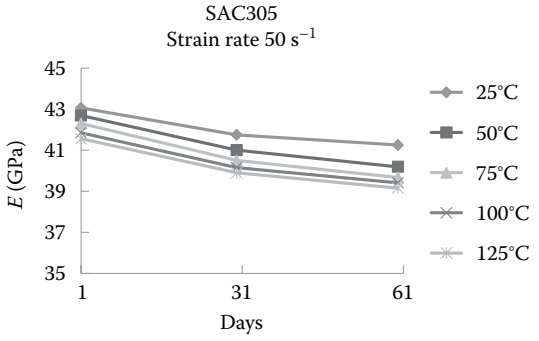


FIGURE 4.22.25 Effect of temperature on elastic modulus of SAC305 at strain rate of 50 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

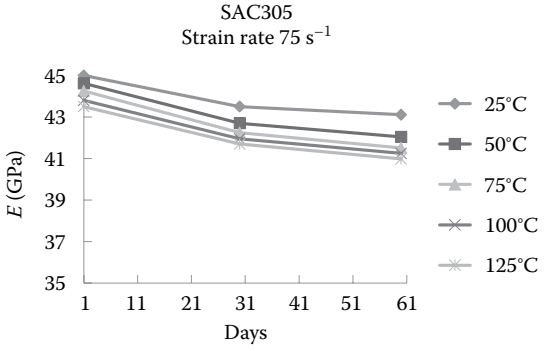


FIGURE 4.22.26 Effect of temperature on elastic modulus of SAC305 at strain rate of 75 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

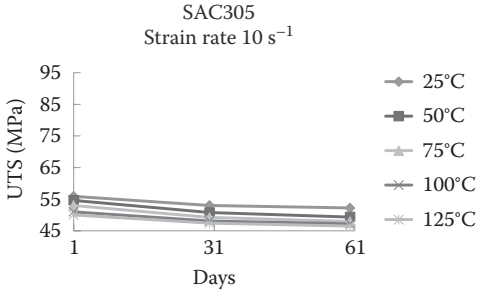


FIGURE 4.22.27 Effect of temperature on ultimate tensile strength of SAC305 at strain rate of 10 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

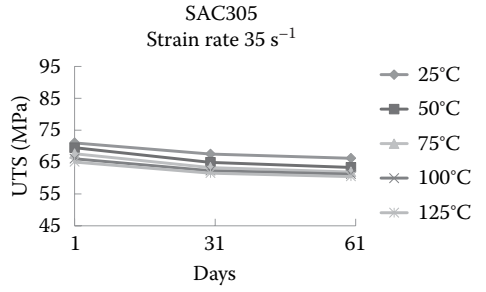


FIGURE 4.22.28 Effect of temperature on ultimate tensile strength of SAC305 at strain rate of 35 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

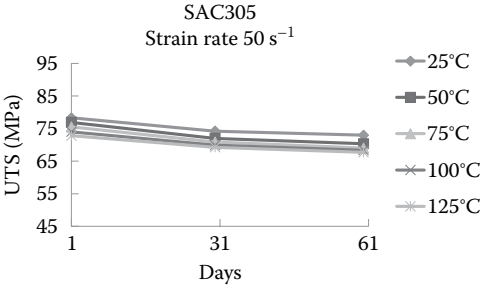


FIGURE 4.22.29 Effect of temperature on ultimate tensile strength of SAC305 at strain rate of 50 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

$$P_D = \frac{T_{max} - T_R}{R_{\theta JR}}, \tag{4.22.5}$$

$$T_J = P_D R_{\theta JR} + T_R. \tag{4.22.6}$$

Generally, T_J is measured utilizing an electrical characteristic of the device that is repeatable and an accurate function of temperature. A key factor in understanding thermal resistance measurements is that if the application configuration is different from that of the measured device, the thermal results will be different. Thermal data on manufacturers' data sheets are designed for comparison with other manufacturers' devices and to give the user a place to begin their thermal management solution.

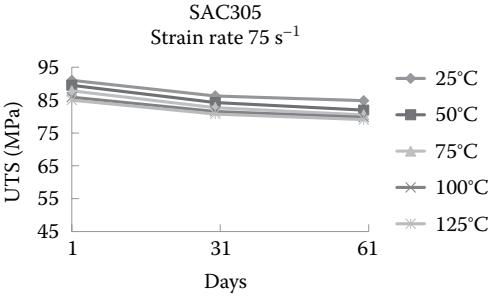


FIGURE 4.22.30 Effect of temperature on ultimate tensile strength of SAC305 at strain rate of 75 s^{-1}. (From Lall, P. et al., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high-temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, pp. 363–377, October 13–17, 2013.)

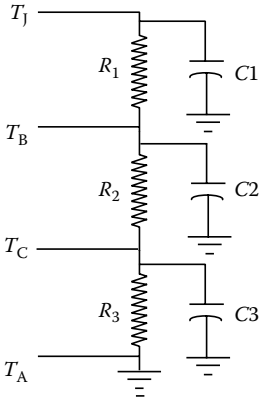


FIGURE 4.22.31 Thermal electrical equivalent circuit heat generated in a device's junction flows from a higher temperature region through each resistor–capacitor pair to a lower temperature region (From Mulgaonker, S. et al., *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A,* 17(4), 573–582, 1994; Mulgaonker, S. et al., *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A,* 18(4), 739–748, 1995. With permissions.) (Courtesy of Motorola).

The thermal performance of power-plastic ball grid array (PBGA) packages cannot be adequately quantified by the traditional $R\theta_{JA}$ measurements. Power-PBGA packages are typically designed thermal vias and "thermal balls," which couples the thermal performance of the package to that of the board. Mulgaonkar et al. (1994, 1995) have shown that the thermal performance of the package is dominated by the thermal performance of the PCB.

Figure 4.22.32 shows the comparison of thermal performance between a 100 lead 14 mm × 14 mm plastic QFP package and a 119 lead 14 mm × 22 mm PBGA package mounted on a single-layer and a four-layer board. There are 32 vias from the die paddle of PBGA to the array of 21 thermal balls. The thermal balls are soldered to an array of board pads. The board pads are connected to the ground plane in the PCB with 32 vias. The planes in the circuit board are solid 1 oz. copper. While the two packages have very similar thermal performance as measured on the standard single-layer PCB, there is a substantial difference in the performance on the four layer boards. PBGA packages are more closely coupled to the PCB than QFPs—in terms of their sensitivity to temperature and power dissipation of neighboring components. The higher thermal conductivity of the four-layer board with two solid 1 oz. planes causes more of the board to act as a heat sink. The effect is enhanced for the PBGA packages because there is metal conduction path from the die pad to the ground plane of the PCB. This path has much lower thermal resistance than the equivalent path for the QFP.

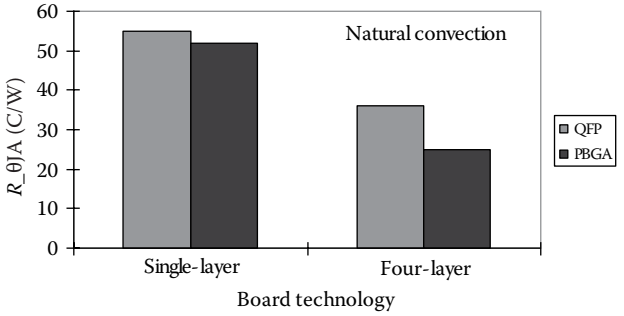


FIGURE 4.22.32 Comparison of thermal performance between a 100 lead 14 mm×14 mm plastic QFP package and a 119 lead 14 mm×22 mm PBGA package mounted on a single-layer and a four-layer board (From Mulgaonker, S. et al., *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A,* 17(4), 573–582, 1994; Mulgaonker, S. et al., *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A,* 18(4), 739–748, 1995. With permissions.) (Courtesy of Motorola).

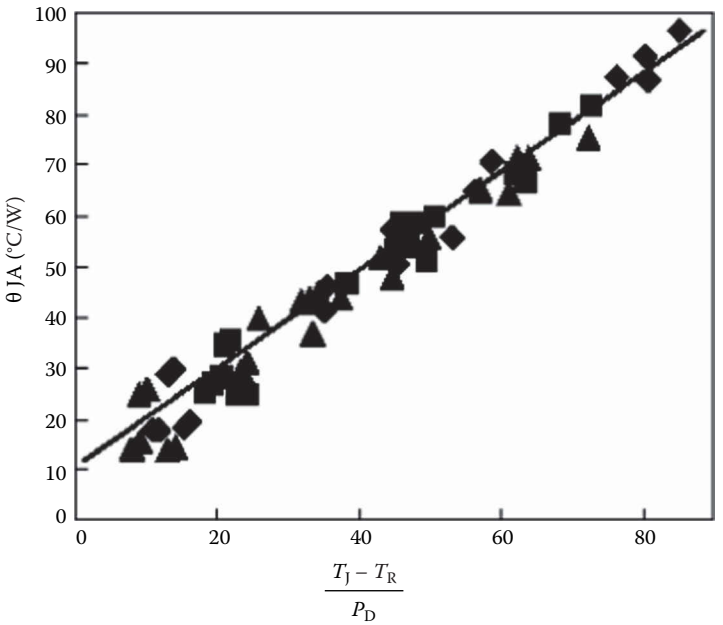


FIGURE 4.22.33 $R_{\theta JA}$ versus $(T_J - T_R)/P_D$ for a PBGA subjected to natural convection and various conditions of forced convection, 1 and 2 m/s (From Mulgaonker, S. et al., *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A,* 17(4), 573–582, 1994; Mulgaonker, S. et al., *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A,* 18(4), 739–748, 1995. With permissions.) (Courtesy of Motorola).

Figure 4.22.33 shows the $R\theta_{JA}$ versus $(T_J - T_R)/P_D$ for a PBGA on a populated board subjected to natural convection and various conditions of forced convection, 1 and 2 m/s. The thermal performance of PBGAs is linearly dependent on the board temperature rise above ambient (divided by package power dissipation). Mulgaonkar et al. (1994, 1995) have represented the thermal performance of electronic packages with equations of straight lines of the form "$y=mx+c$," where "m" is the slope and "c" is the intercept (Table 4.22.22).

A more traditional way to examine the data is to use a table of θ_{JA} determined by a variety of techniques (Table 4.22.23). The junction temperature depends on the environment including the

TABLE 4.22.22
Parameters for Thermal Performance of Electronic Packages (Mulgaonkar et al. 1994, 1995)

	Intercept	Slope
Natural convection	9.8	0.997
Forced 1 m/s	10.7	0.971
Forced 2 m/s	10.7	0.954

Source: Courtesy of Motorola.

TABLE 4.22.23
Effect of Board Configuration on the $R_{\theta JA}$ (Mulgaonkar et al. 1994, 1995)

	θ_{JA} (C/W)	
Board Type	**Natural Convection**	**1 m/s Forced Convection**
Single-layer board	52	41
Four-layer board	24	19
Eight components at 1 W	56–62	46–49
Eight components at 2 W		45–49
Sixteen components at 1 W	104	84

Source: Courtesy of Motorola.

conductivity of the board and the power dissipation of surrounding components. The single component on a multilayer board represents the lower extreme with the higher extreme represented by packages mounted closely together on both sides of the board. The values obtained from the single component on a multilayer board would predict a lower value of the junction temperatures for most applications than would be observed in the typical case with substantial power dissipation in other devices on the board. The effect of the board's power density on the junction temperature is shown by the doubling of the observed θ_{JA} when the packages are mounted on both sides of the board compared to the single sided board (Table 4.22.23). Mounting the packages on both sides of the board effectively halves the area available for power dissipation for each package (Mulgaonkar et al. 1994, 1995).

The traditional θ_{JA} is useful for comparing package performance and as a preliminary estimate to determine whether further analysis is needed. It gives no information to account for the range of thermal performance given as examples in Table 4.22.23. The board temperature is determined by a full board level thermal simulation. There are a number of commercial software codes 6–8 that perform a board level thermal solution with varying degrees of sophistication. These range in sophistication from the 2 1/2 dimensional finite difference or finite element codes to the computational fluid dynamics codes that simultaneously solve the conduction and the fluid flow convection. For all these simulation codes, a simplified thermal model for the package is required.

Thermal Interface Materials

Integrated circuits are kept within their operating temperature limits by transferring the generated heat to the ambient environment. This is often done by attaching a heat sink to the semiconductor package surface. When the electronic component and heat sink surfaces are brought together, typically <1% of the surfaces make physical contact. Much of the surfaces are separated by a layer of interstitial air. The reason is that all surfaces have a certain roughness due to microscopic hills and valleys. Superimposed on this surface roughness is a macroscopic nonplanarity in the form of a

concave, convex, or twisted shape. When the two such surfaces are brought together, only the hills of the surfaces come into physical contact. Some heat is conducted through the physical contact points, but much more has to transfer through the air gaps. Air is a poor conductor of heat and is replaced by a conductive thermal interface. Interface materials fill the voids and grooves created by imperfect surface finish of two mating surfaces, thus forcing air out and improving surface contact and the conduction of heat across the interface.

Typical thermal interface materials include greases, reactive compounds, elastomers, and pressure-sensitive adhesive films. All are designed to conform to surface irregularities. Some have secondary properties and functions, as well. Thermal greases and epoxies provide the lowest interface resistance, but they are pastes and require care in handling. Elastomers eliminate handling problems but they sometimes require high compressive loads even with well-prepared surfaces. Thermal tapes offer great convenience but their gap filling properties are limited.

Interpretation of Material Data

The thermal properties of interface materials are be measured by either ASTM test standard D5470-93 or military specification MIL149456A. Both methods are recognized throughout the industry as standards. High contact pressures (300 psi for ASTM and 500 psi for MIL SPEC) are used to reduce the effects of interface resistance generated by trapped air in the interface. The data gathered at high pressures are then used as the published thermal performance characteristics.

In typical microelectronics packaging applications, the contact pressure varies between 10 and 50 psi and data provided at 300 psi will be an underestimation of material thermal resistance. Any pressure exceeding 50 psi might damage the packaging itself. Typically, the manufacturers do not publish low-contact pressure data because of the case dependency of interface resistance. The thermal resistance of the interface material is a function of the contact pressure.

Types of Thermal Interface Materials

Thermal Greases

Thermal greases are made by dispersing thermally conductive ceramic fillers in silicone or hydrocarbon oils to form a paste. Table 4.22.24 shows some typical grease materials and their thermal conductivity. Sufficient grease is applied to one of the mating surfaces such that when pressed against the other surface, the grease flows into all voids to eliminate the interstitial air. Excess grease flows out past the edges and the thinnest possible thermal joint is formed as the two surfaces come into contact at their high points. Joint integrity has to be maintained with spring clips or mounting hardware. Thermal greases are notoriously "user unfriendly," but provide very low thermal resistance between reasonably flat surfaces. Grease does not provide electrical insulation between the two surfaces, and excess grease that flows from the joint should be cleaned up to prevent contamination problems. Greased joints can also dry out with time, resulting in increased thermal resistance.

Thermal Epoxies

Epoxies are useful in cases where a permanent thermal bond is needed, but soldering is not an option. There are three general types: (1) Silver-filled epoxy can be used from –55 to +125°C,

TABLE 4.22.24
Typical Thermal Greases and Their Thermal Conductivities

Thermal Grease	Thermal Conductivity (W/mK) @ 300 psi
Zinc oxide/silicone based	0.56
Al_2O_3/nonsilicone	2.78
AlN/nonsilicone	3.47

(2) Al_2O_3, and (3) Aluminum Nitride (AlN) (Table 4.22.25). Both Al_2O_3 and AlN types can operate up to 130°C for a limited amount of time. Thermally conductive compounds are an improvement on thermal grease as these compounds are converted to a cured rubber film after application at the thermal interface. Initially, these compounds flow as freely as grease to eliminate the air voids and reduce the thermal resistance of the interface. After the interface is formed, the compounds cure with heat to a rubbery state and also develop secondary properties such as adhesion. Formulations with adhesive properties do not require mechanical fasteners to maintain the integrity of the joint. Since the binder cures to a rubber, these compounds do not have the migration or the dry joint problems associated with thermal greases. Compounds can be used to fill large gaps where greases would bleed from the joint on account of their migratory nature. Clean-up is also simple as excess material is easily removed after it has been cured to a rubber. Cure times range from 24 h at 22°C or 2 h at 65°C for silver-filled epoxies to 1 h at 85°C for Al_2O_3- and AlN-filled epoxies.

Elastomer Pads

Thermally conductive elastomers are silicone elastomer pads filled with thermally conductive ceramic particles, often reinforced with woven glass fiber or dielectric film for added strength. These elastomers are available in thickness from about 0.1 to 5 mm and hardness from 5 to 85 Shore A. Unlike compounds and greases, elastomer pads provide electrical insulation and can be used between surfaces that are at different electrical potential. They are typically used under discrete power devices where electrical isolation is required.

Elastomers do not flow freely like the greases or compounds, but will deform if sufficient compressive load is applied to conform to surface irregularities. At low pressures, the elastomer cannot fill the voids between the surfaces and the thermal interface resistance is high. As pressure is increased, more of the microscopic voids are filled by the elastomer and the thermal resistance decreases. For most high durometer materials, mounting pressures around 300–500 psi eliminate the interstitial voids and reduce interface resistance to a minimum. Mounting pressure must be permanently maintained by using fasteners or springs to hold the two surfaces together.

Table 4.22.26 shows the thermal performance of some common conductive elastomers. Despite the 3–6 W/mK thermal conductivity, the thermal resistance at low pressures is in the neighborhood of 1.8°C/W, twice that of grease. Typical contact area at low pressures is in the neighborhood of 30%.

TABLE 4.22.25
Typical Thermal Epoxies and Their Thermal Conductivities

Thermal Epoxy	Cure Time	Thermal Conductivity (W/mK)
Silver filled	24 h @ 22°C or 2 h @ 65°C	1.39
Al_2O_3 filled	1 h @ 85°C	0.93
AlN filled	1 h @ 85°C	2.78

TABLE 4.22.26
Typical Elastomer Pad Materials and Their Thermal Conductivities

Elastomer Pads	Thermal Conductivity (W/mK)
Paraffin/Al	3.46–6.54
97% Graphite	3.38–5.62
Al_2O_3	1.63–4.33

Adhesive Tapes

Thermally conductive adhesive tapes are double-sided pressure-sensitive adhesive films filled with sufficient ceramic powder to balance their thermal and adhesive properties. The adhesive tape is usually supported either with an aluminum foil or a polyimide film for strength and ease of handling. Polyimide support also provides electrical insulation. Adhesive tapes perform much like the elastomeric films, in that they also require some initial mating pressure to conform to irregularities in the mating surfaces. They are also unable to fill large gaps between nonflat surfaces. However, once the joint is formed, the adhesive tapes require no mechanical support to maintain the mechanical or thermal integrity of the interface.

Adhesive tapes provide convenience in attaching a heat sink to a semiconductor package because, unlike liquid adhesives, no cure time is required. The film is applied to one of the surfaces, usually to the heat sink, and it is then forced into contact with the semiconductor package to complete the thermal joint. The application pressure is typically 10–50 psi for a few seconds duration. The bond thus formed can be considered permanent and the heat sink is reliably attached to the semiconductor.

However, this convenience comes at a price. The thermal resistance for adhesive tapes is only slightly better than a joint without any thermal interface material. This is because the thermal tapes do not fill gaps as well as liquids, and thermal joints made with tapes will normally include considerable interstitial air gaps. For the most part, the quality of the two joining surfaces will determine the amount of contact that can be achieved and the thermal performance that can be expected. The high shear strength of these thermal tapes means that reliable joints between heat sinks and semiconductors can be achieved, even with poor surfaces and no mechanical fasteners.

Computers

Portable computing applications bring with them the challenges of size, weight, and power consumption. Microprocessor (CPU) is the major source of heat dissipation in a portable computer. Cooling systems, in general, increase the size and weight of the portable computer.

Understanding Temperature Limits

Heat from a portable computer is passively dissipated by warming the outside surface (skin) of the plastic or the housing material over the ambient air and actively dissipated by using a fan and heat exchanger. Most fan designs can remove about 4–8 W of heat and the rest is passively dissipated by natural convection and radiation from the outside surface. The passive dissipation of the heat requires elevation of the notebook skin temperature over the air temperature. Typically, as a guideline the portable computer-skin temperature is kept lower than 50°C. Typical ambient operation temperatures are in the neighborhood of 25°C–35°C, an environment that applies to most air-conditioned places. Most personal computer chassis result in an internal 10°C temperature increase over ambient temperature. In such an environment, the portable computer-skin temperature will be 40°C–50°C (assuming 15°C above ambient). Since the human body temperature is 37.1°C, temperatures in the neighborhood of 40°C are perceived as normal and temperatures above 50°C are perceived as hot to touch. The temperature specifications for many keyboards are generally in the range 55°C–60°C and thus the keyboard operation will be at the specification limit. Similar logic can apply to other temperature-sensitive components (e.g., HDD, CD-ROM) inside the notebook (AMD 1998; Intel 1998b).

Figure 4.22.34 shows the maximum temperature versus component for a portable computer. Notebook computers are generally put together with standard components of varying maximum ambient temperatures. The failure mechanisms are device dependent. Floppy disk drive and compact disk drives start to degrade in the neighborhood of 50°C–55°C. The media in floppy disk drives warps and sticks to the casing at high temperatures. CD-ROM drives have laser diode aligned to the compact disc—large dimensional changes produce misalignment and malfunction.

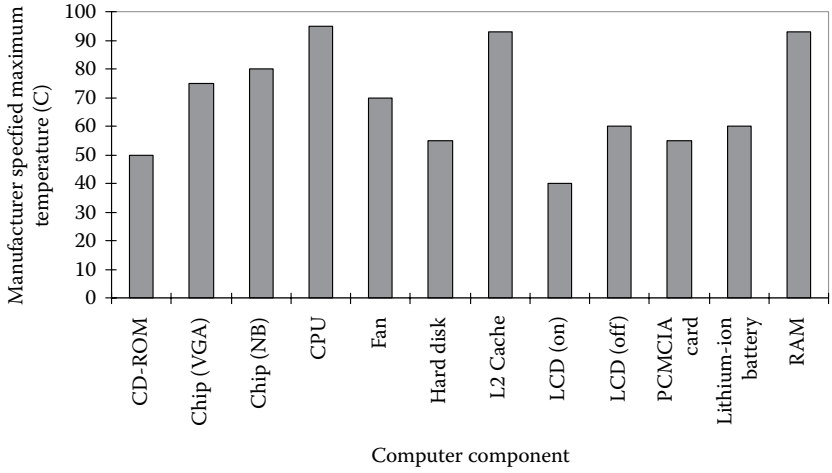


FIGURE 4.22.34 Manufacturer specified maximum temperature for personal computer components. (Based on data from Sterner, J., Thermal design challenges in a Pentium notebook computer, *Proceedings of the IMAPS International Systems Packaging Symposium*, pp. 107–112, December 1997. With permission.)

Temperature dependence of hard drive performance is often claimed by manufacturers of HDD cooling products. However, temperature (below a certain temperature limit, T_{limit}) does not directly affect the performance of the drive in any significant way. Seek times and data transfer rates generally stay the same with variation in temperature below T_{limit}. The rise in temperature, however, does require thermal recalibration, producing a very small performance degradation. Many SCSI drives, operating at 10,000 rpm, have a thermal recalibration loop. The rise in ambient temperature causes the drive to warm up—the platters expand, and their geometry changes, thus making recalibration necessary. During recalibration, the hard drive makes a short pause to recalibrate itself. Generally, hard drives operating in cooler environments recalibrate less frequently. However, new generation drives with embedded servo do not have a thermal recalibration loop, and thus their performance is totally unaffected by the operating temperature (up to a certain temperature limit). Hard disks are high precision machines with very tight tolerances. Large temperature variations ($T_{ambient} > T_{limit}$) cause dimensional changes, which cannot be compensated—causing malfunctions. Typical upper limit (T_{limit}) on hard drive temperature is in the neighborhood of 60°C. Cooling is generally recommended for all hard drives with rotation speeds higher than 7200 rpm (Steinbrecher 1998).

Figure 4.22.35 shows the power dissipation in typical ACTIVE and PASSIVE cooled portable computers. The ACTIVE and PASSIVE computers compared in Figure 4.22.35 have different equipment. Figure 4.22.36 shows the computer surface temperature versus power dissipation for a typical ACTIVE and typical PASSIVE cooled portable computer for an ambient temperature of 25°C. The bottom surface of the computer is insulated. The top (keyboard) and side walls are used as a radiator. The notebook casing is made of ABS/PC plastic or painted metal chassis (emissivity = 0.95) and has an isothermal casing temperature. The data presented in Figures 4.22.35 and 4.22.36 are for an ACTIVE cooled portable computer 10 in. × 12.3 in. × 1.0 in. base unit (excluding lid thickness) and a PASSIVE cooled portable computer with base size of 8.2 in. × 10.2 in. × 0.65 in.

In general, ACTIVE cooled portable computers can dissipate more power than PASSIVE cooled portable computers and still meet the temperature specification for the computer-casing temperature (Figure 4.22.35). While CPU is not the only component dissipating heat, it is certainly one of the major contributors to increase in computer surface temperature. Other major contributors are the graphics card, memory, hard drives, and power supplies. For operation at 15°C above ambient temperature (25°C), about 15–17 W of total platform power (entire notebook power minus the display power) can be managed passively in a typical notebook. The allowable power dissipation for

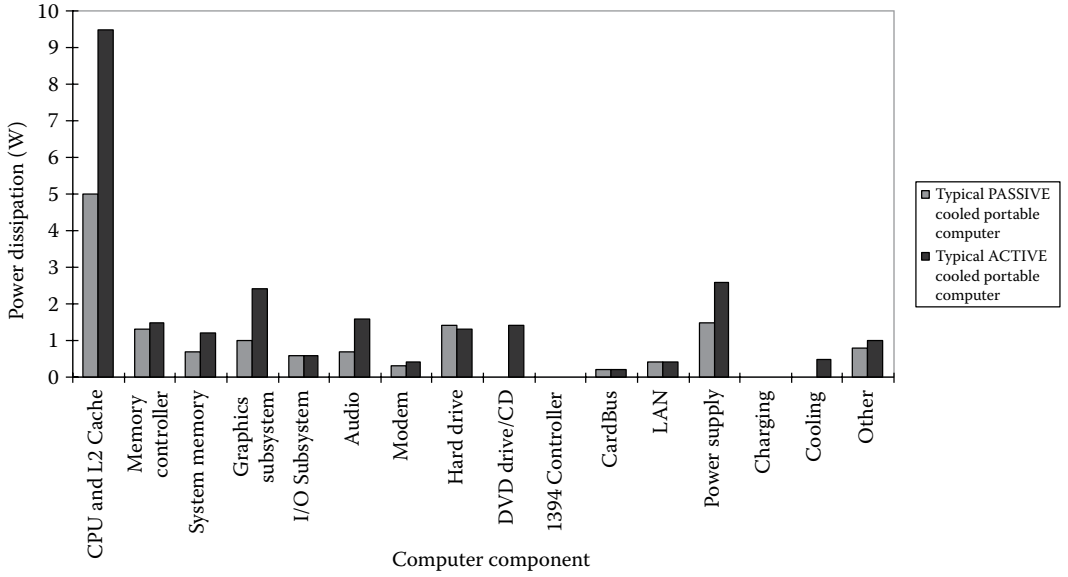


FIGURE 4.22.35 Power dissipation in typical ACTIVE and PASSIVE cooled portable computers. The ACTIVE and PASSIVE computers compared in the figure have different equipment. ACTIVE cooled portable computers can dissipate more power for any temperature specification of the computer-casing temperature. (Based on data from Intel Corporation, *Mobile Power Guidelines 2000*, Available from http://developer.intel.com/design/ mobile/intelpower/, Revision 0.8, September 16, 1998b.)

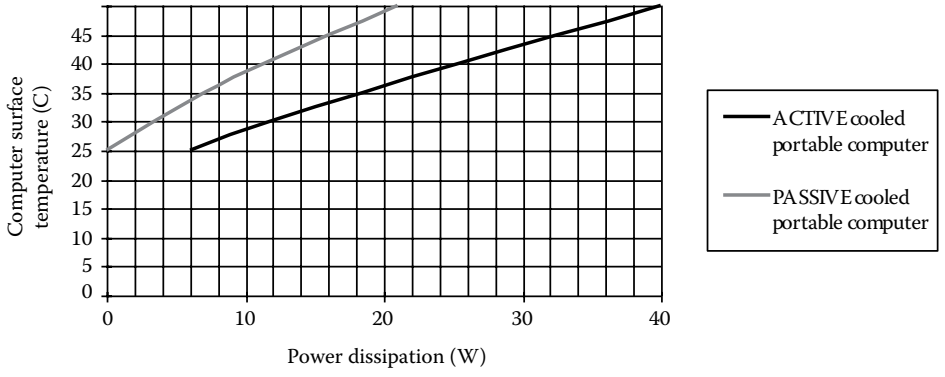


FIGURE 4.22.36 Computer surface temperature versus power dissipation for a typical ACTIVE and a typical PASSIVE cooled portable computers for an ambient temperature of 25°C. The ACTIVE and PASSIVE computers compared in the figure have different equipment. (Based on data fromIntel Corporation, *Mobile Power Guidelines 2000*, Available from http://developer.intel.com/design/ mobile/intelpower/, Revision 0.8, September 16, 1998b.)

an actively cooled portable computer is in the neighborhood of 24–25 W (Figure 4.22.36). [Based on data from Intel (1998b).]

Casing Materials

External skin temperature of the portable computer is designed to be safe to human touch. For most purposes, temperatures above 50°C are deemed unsafe to human touch. Computer casings are generally made of ABS (acrylnitrite–butadiene–styrene) resin, PC/ABS blend (polycarbonate/ABS), or

magnesium alloy. The maximum temperature for continuous operation for both ABS and PC/ABS casings is in the range of 85°C–90°C. Table 4.22.27 shows the properties for some typical ABS and PC/ABS blends used for computer housings.

The ABS resin has low-thermal conductivity and the molding process constraints limit the wall thickness of ABS enclosure to a minimum of 1.2–1.4 mm. Acrylonitrile–butadiene–styrene (ABS) thermoplastic resin is widely recognized as an engineering material offering a good balance of properties centering around toughness, hardness, and rigidity. Resin grades in this product family consist of a blend of an elastomeric component and an amorphous thermoplastic component. The elastomeric component is usually polybutadiene or a butadiene copolymer. The amorphous thermoplastic component is SAN, a copolymer of styrene and acrylonitrile. The flexibility offered by the use of the three monomer system (A, B, and S) allows the tailoring of the property profile.

Variation of the ABS material include blends of amorphous PC and ABS terpolymer, which combine the most desirable properties of both materials resulting in an optimum balance of performance, processability, and cost. By varying the ratio of PC and ABS, resins can be tailored to meet the specific property requirements of a wide range of high-performance applications from automotive body panels and computer housings, to instrument panels and cellular phones. CYCOLOY resins offer the processability of ABS together with the mechanical properties, impact and heat resistance of PC. These thermoplastic alloys are formulated for exceptional flow for filling thin-wall sections and complex parts.

In contrast, magnesium alloy can be injection molded to a wall thickness of 0.7 mm. Magnesium's best attributes are its high strength-to-weight ratio and light weight. As a structural material, it can be alloyed with several other elements including aluminum, manganese, rare-earth elements, silver, zinc, and zirconium. The various alloys have unusually high strength and are light weight. Magnesium is most commonly found alloyed with aluminum where its addition to aluminum increases hardness and improves corrosion resistance. Magnesium also possesses superior heat radiation properties, heat stability, electromagnetic interference shielding properties, and damping capabilities. Recent developments in thin-wall casting combined with magnesium's other characteristics are leading to increased applications in the electronics field for cellular telephones, video cameras, and personal computers (casings, battery shields, etc.). Table 4.22.28 shows the properties for typical magnesium alloys used for computer housings.

TABLE 4.22.27
Properties for Some Typical ABS and PC/ABS Blends Used for Computer Housings

	ABS	PC/ABS
Material Properties	**Acrylonitrile–Butadiene–Styrene**	**Polycarbonate/Acrylonitrile–Butadiene–Styrene**
Tensile strength—yield (psi)	5,300–6,500	8,200–9,100
Flexural strength—yield (psi)	9,400–11,000	12,500–14,800
Modulus of elasticity (psi)	3.40e+05 to 3.77e+05	3.20e+05 to 3.90e+05
Coefficient of Thermal Expansion (CTE) (ppm/°C)	100–140	80 normal-to flow, 76 parallel-to-flow
Deflection temperature at 0.46 Mpa (°C)	90	100
Deflection temperature at 1.8 Mpa (°C)	85	90
Max. continuous service temperature (°C)	85	90
Thermal conductivity (W/mK)	0.14–0.30	0.2–0.25
Specific heat (J/g °C)	1.2–1.4	1.2–1.4

TABLE 4.22.28
Properties for Typical Magnesium Alloys Used for Computer Housings

	Magnesium Alloy	
Material Properties	**AM60A-F**	**AZ91A-F**
Tensile strength—yield (ksi)	16.68	21.76
Poisson's ratio	0.35	0.35
Modulus of elasticity (psi)	6.52e+06	6.52e+06
CTE (ppm/°C)	25.6	26
Thermal conductivity (W/mK)	62	72
Specific heat (J/g °C)	1	1.05

Typical Heat Transfer Schemes

Figure 4.22.37 shows some of the typical active and passive heat transfer schemes used for portable computing applications. Figure 4.22.37a shows a passive heat transfer scheme with the use of a heat pipe along with a heat spreader to conduct the heat from the microprocessor. Figure 4.22.37b shows an active heat transfer scheme using a finned heat sink and fan. Figure 4.22.38 shows the representative Dual In-line Memory Module (DIMM) layout in a high-performance server. The buffered DIMM module shows a number of Dynamic Random Access Memory (DRAM) devices and a buffer chip. The basic layouts of other memory modules are similar. Figure 4.22.39 shows the temperature variation across a registered DIMM with a peak temperature of 120°C.

In desktop computers, fans are used to evacuate the hot air from the case. Figure 4.22.40 shows some of the common fan configurations in AT and ATX form-factor cases. Case fans come in many flavors, but usual sizes include 80 mm × 80 mm and 92 mm × 92 mm. The size and location of the fan depends on the computer case. Most desktop computers come in AT or ATX form-factor cases. In general, most designs use some combination of fans blowing air into and out of the computer case. The idea being that the CPU located directly in the airflow will get cooled. However, as the power supply gets hot, the fan blows warm air toward the CPU. There are two possibilities for installing the fan: Either above the power supply on the back (especially common with ATX cases) or on the lower part of the front side (common with AT cases). Care is exercised in determining the direction of air flow for each of the fans to make sure that (1) the fans do not create an air short circuit, which prevents circulation of air through the computer case and (2) the fans do not try to work against each other and try to create a vacuum inside the case.

Figure 4.22.41 shows the temperature variation across DRAM chips in a DIMM on a desktop machine. The increase in system performance with each new processor generation increases the

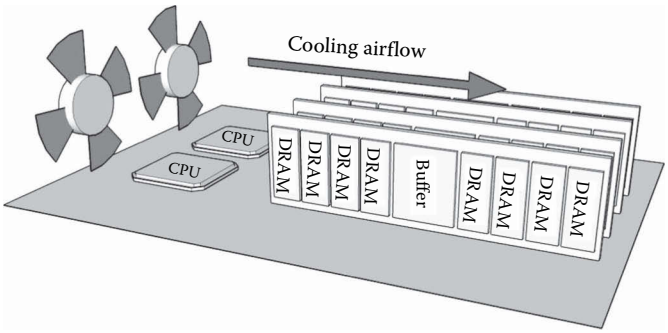


FIGURE 4.22.37 Some typical heat transfer schemes for portable computers. (From Sterner, J., Thermal Design Challenges in a Pentium Notebook Computer, *Proceedings of the IMAPS International Systems Packaging Symposium,* pp. 107–112, December 1997. With permission.)

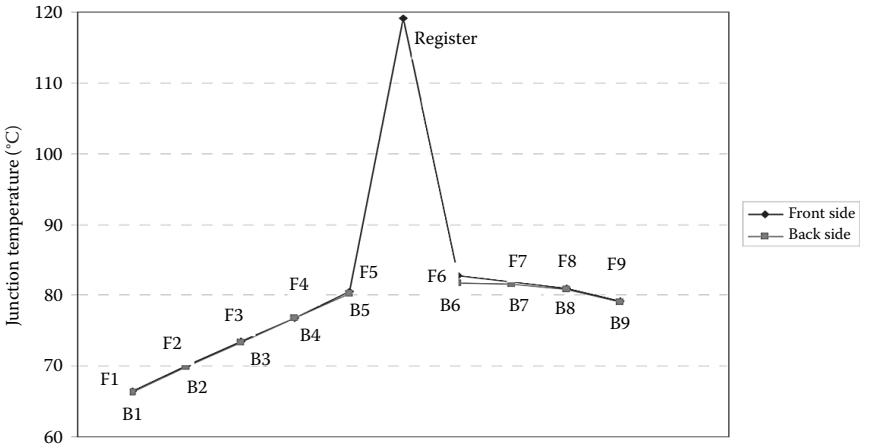


FIGURE 4.22.38 Some typical heat transfer schemes for portable computers. (From Liu, S. et al., Hardware/software techniques for DRAM thermal management, *Proceedings of the 17th IEEE International Symposium on High Performance Computer Architecture (HPCA)*, pp. 1–11, 2011.)

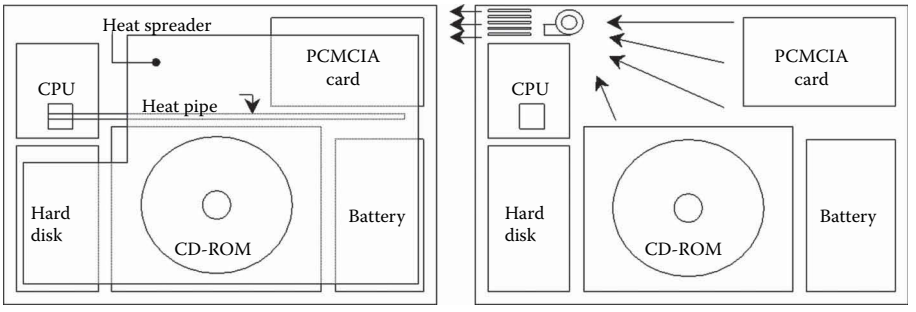


FIGURE 4.22.39 Temperature profile across an Registered Dual In-Line Memory Module (RDIMM). (From Zhu, Q. et al., Thermal management of high power memory module for server platforms, *Proceedings of the ITHERM Conference*, pp. 572–576, 2008.)

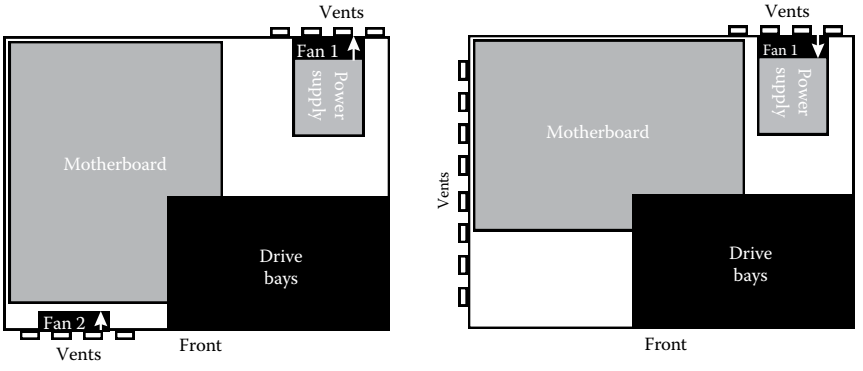


FIGURE 4.22.40 Some common fan configurations for AT and ATX form-factor desktop computer cases. (From Steinbrecher, T., *Heat-Sink Guide*, http://www.heatsink-guide.com/, 1998; AMD (Advanced Micro Devices), *AMD-K6 Processor Thermal Solution Design*, Application Note, Document Number 21085, Revision H, Amendment 0, November 1998.)

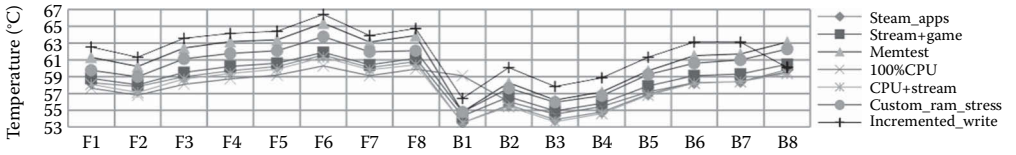


FIGURE 4.22.41 Temperature variation across DRAM chips in a DIMM on a desktop machine. (From Liu, S. et al., Hardware/software techniques for DRAM thermal management, *Proceedings of the 17th IEEE International Symposium on High Performance Computer Architecture (HPCA)*, pp. 1–11, 2011.)

system level thermal demands. In order to meet these thermal needs, the current systems require either multiple fans or heat sinks to maintain the necessary junction temperatures for the baseboard components. Need for multiple fans can be avoided by using FAN DUCTS. The fan duct distributes the air from a single fan to all parts of the system that require cooling—thus avoiding the need for multiple fans. A recommended path for a low-cost solution in these systems is to provide a fan duct that blows external air onto the processor core components (processor, chipset, graphics, graphics controller, and memory) (Intel 1998a).

Handheld Communication Devices

Mobile communication products such as cellular phones, pagers, two-way radios, personal digital assistants are being driven to increased functionality and smaller form factors by technological enhancements at device, package, and system levels.

Understanding Application Requirements

Typically, handheld products involve plastic cases, which have been snap-fitted or screwed together—space is at premium. The electronic packages are assembled in small, confined enclosures, and the stand-alone thermal performance of the packages does not guarantee similar levels of performance on the phone. The problem is further complicated by the fact that the plastic housings for handheld products have to remain cool enough to be safe to human touch.

Pagers typically are receivers and their active components dissipate low power (a few milliwatts) over a short operational time (several milliseconds). Cooling is generally not a concern. However, newer pagers including two-way paging involve transmitting signals with a power dissipation of approximately 1 W over an extended period of operation. Phones generate most of their heat in the power amplifier (PA) module (PA). The PA can dissipate power in the range of 1–3 W. Typically, cellular phones are two board designs including the main RF board and a keypad board. Typically, the keypad board has the front keypad popples for the phone and very few if any components. The PA module and other ICs are mounted on the main RF board. Two-way radios are similar to cellular phones except the power dissipation requirements are in general significantly larger. Handheld two-way radio PA may generate 1–10 W of power depending on the mode of operation. Two-way radios, like the cellular phones, are two-board designs including a main RF board and a keypad board. The PA in general is located on the main RF board.

Thermal Management: Some Test Cases

The choice of thermal management schemes is severely limited because of the space, size, weight, and cooling system power requirement constraints. The added weight of carrying a heavy cooling system in a handheld communication device is typically not an attractive option—let alone put up with the fan noise while talking on the phone or two-way radio. Cooling system power requirements further increase battery size and reduce standby and talk time. Passive cooling schemes in general are preferred—most of the heat generated in handheld portable products by the integrated circuits is

transmitted to the externals predominantly by conduction and to a lesser extent by natural convection and radiation.

Table 4.22.29 shows some of the common passive heat transfer schemes implemented by providing a solid conduction path between the PA module and the product casing. The solid conduction path between the PA and the system housing can be either a metal block or a thermally conductive pad (epoxy or silicone with a thermal conductivity of ~2 W/mK), a metal spring (~0.8 mm thick), or a thin metal foil (~0.04 mm thick). The material for metal block, spring, or foil could be either copper or aluminum. The metal block, conductive pad, or spring is simply inserted into the gap and the contact is maintained by the press fit. An adhesive tape may be required to attach the foil to the PA and housing. The metal foil can also be used as a heat spreader on the inside surface of the system case. If the power dissipated by the PA can be effectively transferred to the case wall (such as by a solid conduction path), then the heat spreader can quickly distribute the local heat flux over the larger area, reducing component temperatures.

Figure 4.22.42 shows the PA temperature versus power dissipation at the PA for various thermal enhancement techniques. Lee et al. (1998) showed that the least effective enhancement was made by positioning the thermal pad between the board (on the back side of the PA) and the case. The configuration with the thermal pad directly contacting the PA reduced PA temperature by 34%. The copper spring, due to its higher thermal conductivity and larger heat transfer cross-sectional area, provided the best solid conduction path. The copper foil provided an intermediate enhancement. In addition to the solid conduction path, adding the heat spreader to the case further reduced the PA temperature by an average of 26%. A combination of a thin metallic spring and a heat spreader on the enclosure provided the highest enhancement with a 68% reduction in PA temperature.

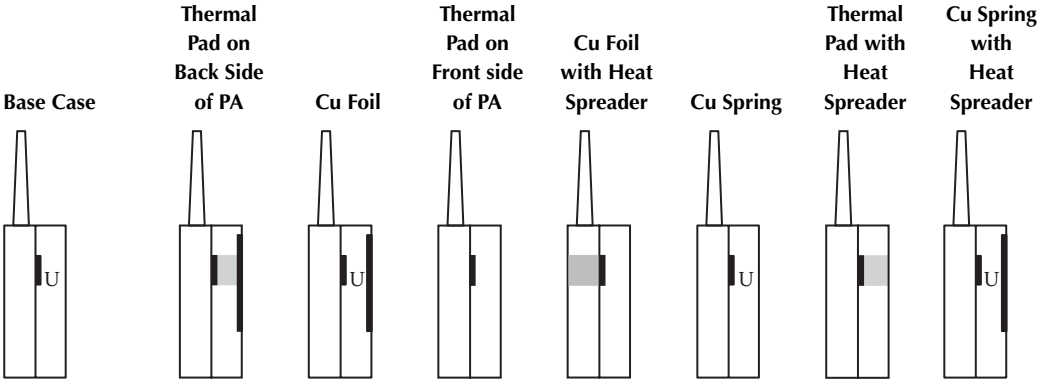
Casing Materials

Typically handheld portable product casings are made from ABS or PC/ABS blends. Materials are chosen based on various attributes including thin-wall molding, mold release, elastic modulus, and thermal conductivity. Table 4.22.30 shows the mechanical and thermal properties for typical handheld product casing materials.

Impact of Temperature on Handheld Product Reliability

In general, high temperature does not drive the reliability of handheld portable products. High-temperature bake is often used to cure moisture-related problems in PCBs and packages. A 24 h

TABLE 4.22.29
Thermal Enhancement Configurations (Lee et al. 1998)



Source: Courtesy of Motorola.

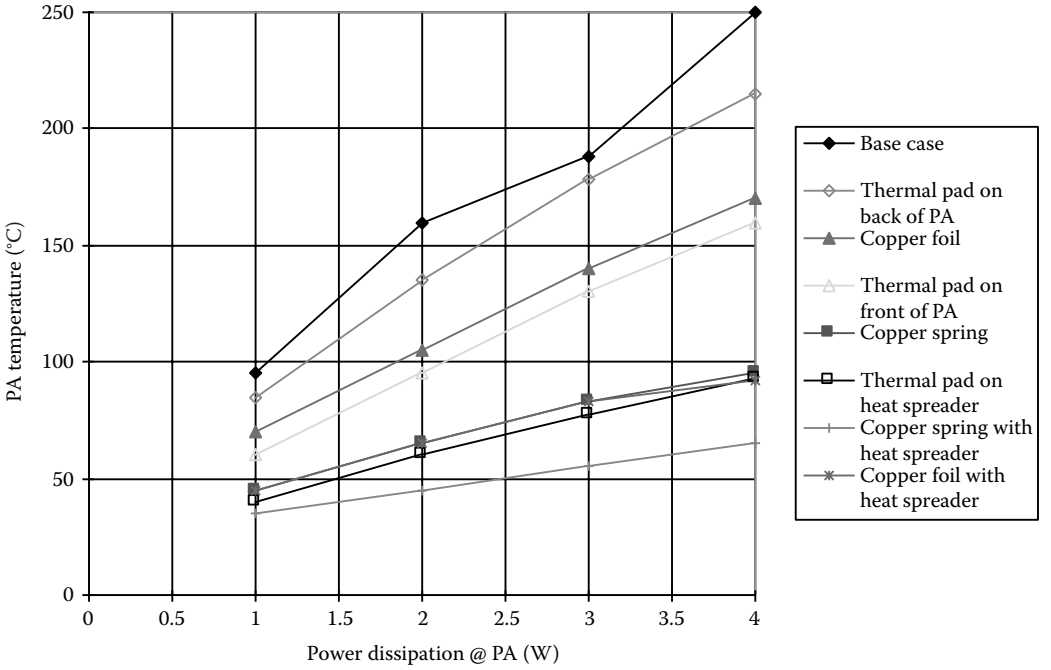


FIGURE 4.22.42 Comparison of various thermal enhancement techniques (the average experimental uncertainty is ±4%). (From Lee, T.T. et al., Thermal management of handheld telecommunication products, *Electronics Cooling*, Vol. 4, No. 2, May 1998.) (Courtesy of Motorola).

TABLE 4.22.30
Material Properties of the Housing for a Typical Handheld Portable Product

Material Properties	
Tensile strength—yield (ksi)	8.0–9.0
Flexural strength—yield (ksi)	11.5–14.0
Modulus of elasticity (psi)	3.40E+05
Thermal conductivity (W/m °C)	0.27
CTE (ppm/°F)	38
Specific heat (BTU/lb °F)	0.3
Density (lb/in.^{3})	0.043

bake at 150°C prevents popcorn or blistering of the packages or the PCB during reflow. However, high temperature during operation does impact the handheld product casing temperature.

Product reliability has a dominant dependence on the thermal cycling instead of high steady-state temperature. Thermal fatigue is a wear-out mechanism as a result of product usage (e.g., making phone calls, operating in home, office, car, leaving phone in the car, etc.) in different environments and product power cycling (turning ON and OFF). Figures 4.22.43 through 4.22.45 show the dominant chip-scale package failures caused by thermal fatigue in package form factors typically used in handheld communication products.

Figure 4.22.43 shows an cross-section of a Flex-Substrate BGA after thermal fatigue failure. The 48 I/O, 0.8 mm pitch Flex-Substrate BGA packages (SRAM) were soldered onto a 0.8 mm thick,

Immersion Ni/Au finish PC Board. The board assemblies were subjected to Liquid-Liquid Thermal Shock (LLTS), −55 to 125°C, 5.8 c/h. The primary crack propagation site is at the package-to-solder joint interface.

Partial-Array BGAs fail primarily by solder joint failure. Figure 4.22.44 shows a cross-section of a Partial-Array BGA Package (48 I/O, 0.5 mm pitch) after thermal fatigue failure. The 48 I/O, 0.5 mm pitch Partial-Array BGA was soldered to a 1.2 mm thick, OPC finish, PC Board and subjected to LLTS, −55 to 125°C, 5.8 c/h. The primary crack propagation site is at the package substrate-to-solder joint interface.

Elastomer-on-Flex Packages fail in thermal fatigue, primarily due to the fracture of the beam leads internal to the package. Figure 4.22.45 shows cross-section for Elastomer-on-Flex Packages (48 I/O, 0.65 mm and 0.75 mm pitch, FLASH) soldered onto 0.8 mm thick, Ni/Au finish PC boards. Beam lead failure typically occurs in the neighborhood of 400 (Vendor A) to 650 cycles (Vendor B) depending on the package Vendor. The board assemblies were subjected to LLTS −55 to 125°C, 5.8 c/h. The interested reader is referred to Lall et al. (1998) for a more detailed discussion of non-linear reliability modeling and failure mechanisms for the Elastomer-on-Flex Packages.

Figure 4.22.46 shows the thermal fatigue reliability for the Partial-Array BGA (0.5 mm pitch 48 I/O), Elastomer-on-Flex Package (0.65 mm pitch, 48 I/O), Flex-Substrate BGA (0.8 mm pitch, 48 I/O). Weibull distribution distinguishes between failure mechanisms by difference in the slopes of the distributions. The similarity in the slopes of the Flex-Substrate BGA and the Partial-Array

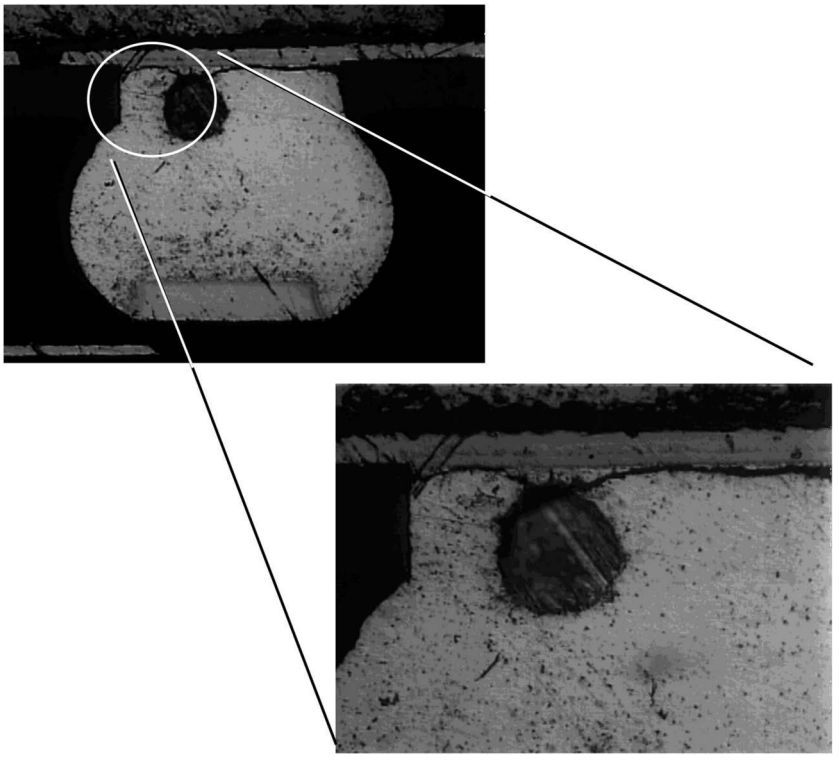


FIGURE 4.22.43 Cross-section of Flex-Substrate BGA thermal fatigue failure—the primary crack propagation site is at the solder joint to package interface. (From Lall, P. et al., *Influence of Temperature on Microelectronic and System Reliability*, Boca Raton, FL: CRC Press, 1997; Lall, P. and Banerji, K., Assembly-level reliability characterization of chip-scale packages, *Electronic Components and Technology Conference (ECTC)*, Seattle, Washington, DC, May 25–28, Session 13, Paper 2, 1998a; Lall, P. and Banerji, K., Assembly-level reliability of flex-substrate BGA, elastomer-on-flex packages, and 0.5 mm pitch partial array packages, *Surface Mount International Conference (SMI)*, San Jose, CA, August 23–27, 1998b. With permissions.)

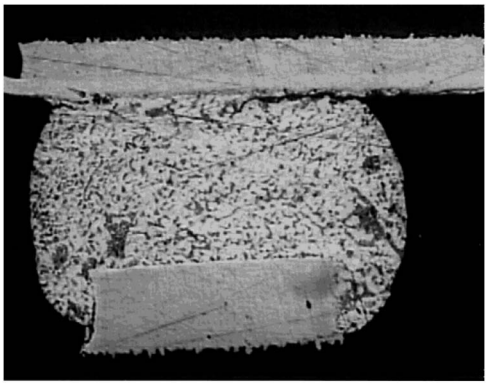


FIGURE 4.22.44 Cross-section of 0.5 mm pitch package after thermal fatigue failure. Notice the crack propagation at the package-to-solder interface. (From Lall, P. et al., *Influence of Temperature on Microelectronic and System Reliability*, Boca Raton, FL: CRC Press, 1997; Lall, P. and Banerji, K., Assembly-level reliability characterization of chip-scale packages, *Electronic Components and Technology Conference (ECTC)*, Seattle, Washington, DC, May 25–28, Session 13, Paper 2, 1998a; Lall, P. and Banerji, K., Assembly-level reliability of flex-substrate BGA, elastomer-on-flex packages, and 0.5 mm pitch partial array packages, *Surface Mount International Conference (SMI)*, San Jose, CA, August 23–27, 1998b. With permissions.)

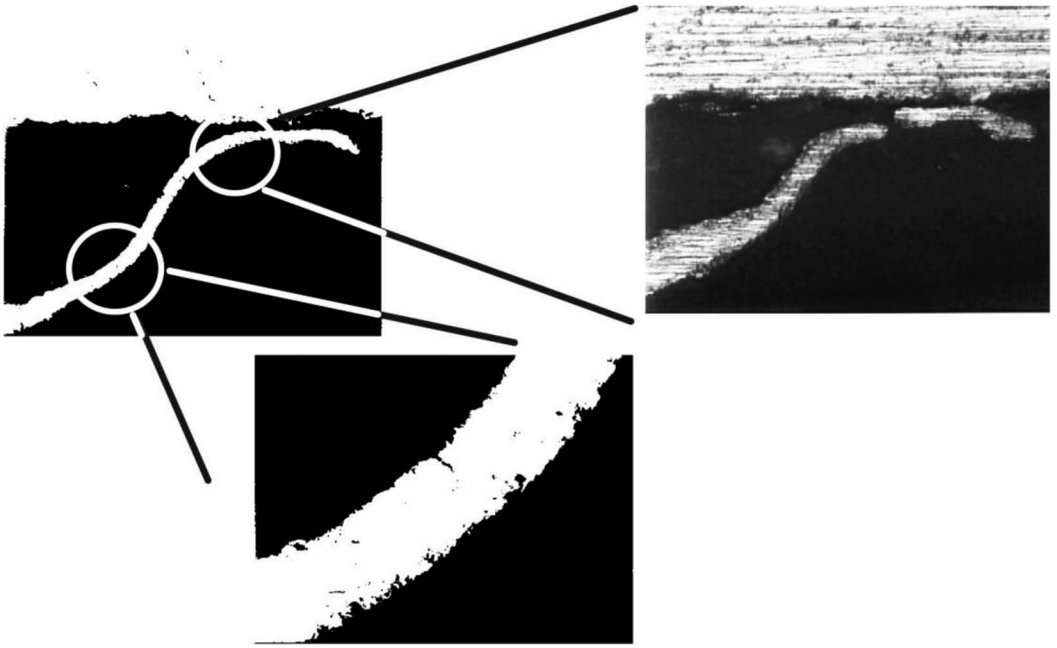


FIGURE 4.22.45 Cross-section of the Elastomer-on-Flex Package after thermal fatigue failure. The solder joints do not fail—the dominant failure mechanism is the thermal fatigue failure of the beam leads internal to the package. (Lall, P. et al., Reliability characterization of the SLICC package, *Electronic Components and Technology Conference (ECTC)*, Orlando, FL, May 28–31, pp. 1202–1210, 1996; Lall, P. et al., *Influence of Temperature on Microelectronic and System Reliability*, Boca Raton, FL: CRC Press, 1997; Lall, P. and Banerji, K., Assembly-level reliability characterization of chip-scale packages, *Electronic Components and Technology Conference (ECTC)*, Seattle, Washington, DC, May 25–28, Session 13, Paper 2, 1998a; Lall, P. and Banerji, K., Assembly-level reliability of flex-substrate BGA, elastomer-on-flex packages, and 0.5 mm pitch partial array packages, *Surface Mount International Conference (SMI)*, San Jose, CA, August 23–27, 1998b. With permissions.)

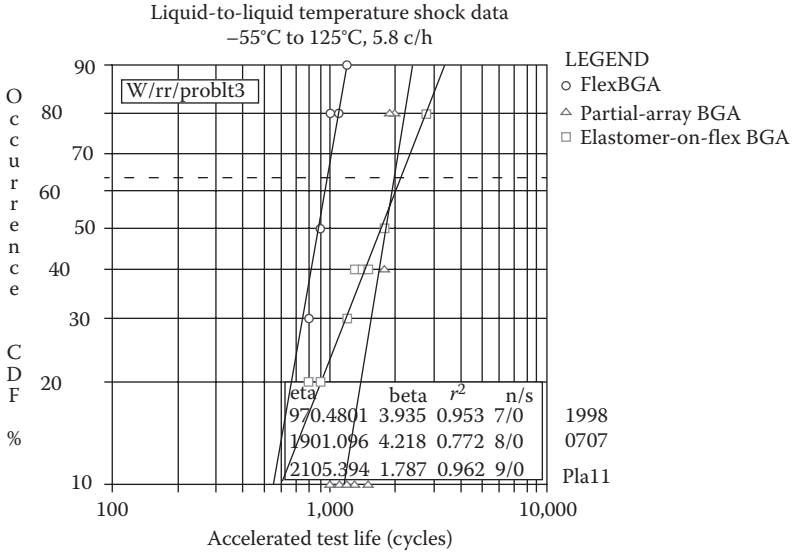


FIGURE 4.22.46 Weibull distributions for the Flex-Substrate BGA Package, Elastomer-on-Flex Package, and the Partial-Array BGA Package subjected to LLTS, –55 to 125°C, 5.8 c/h.

BGA Packages is because of similarity in the mechanisms of failure—solder joint fatigue. The difference in the slope of the Elastomer-on-Flex Package is because of a different failure mechanism—beam lead failure. Figure 4.22.47 shows the normalized accelerated test time-to-first failure of the packages (LLTS –55°C to 125°C, 5.8 c/h) versus package type.

Acceleration Transforms

Environmental usage profile has been developed based on phone usage profile, typical environmental extremes, and the measured thermal behavior of the radio. The profiles (Figure 4.22.48) have been developed for two distinct environmental conditions—extreme (hot, cold). Nonlinear finite element simulations have been run to evaluate the field life and correlate the LLTS cycles to years of field life.

Acceleration transforms have been derived based on nonlinear finite element model predictions. An acceleration transform quantifies the acceleration factor for an accelerated test environment compared to actual field use environments. The acceleration transform is in turn determined by several variables including the component and board—geometry, materials, architecture, and most

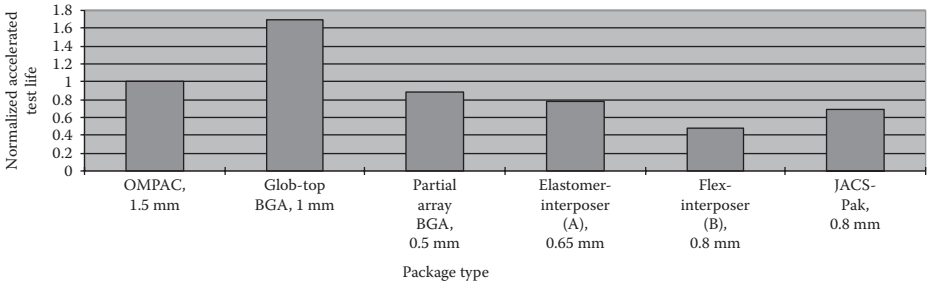


FIGURE 4.22.47 Comparison of the thermal fatigue reliability for the Partial-Array BGA (0.5 mm pitch 48 I/O), Elastomer-on-Flex Package (0.65 mm pitch, 48 I/O), Flex-Substrate BGA (0.8 mm pitch, 48 I/O), and Glob-top BGA (1 mm pitch, 196 I/O) versus OMPAC (1.5 mm pitch, 68 I/O).

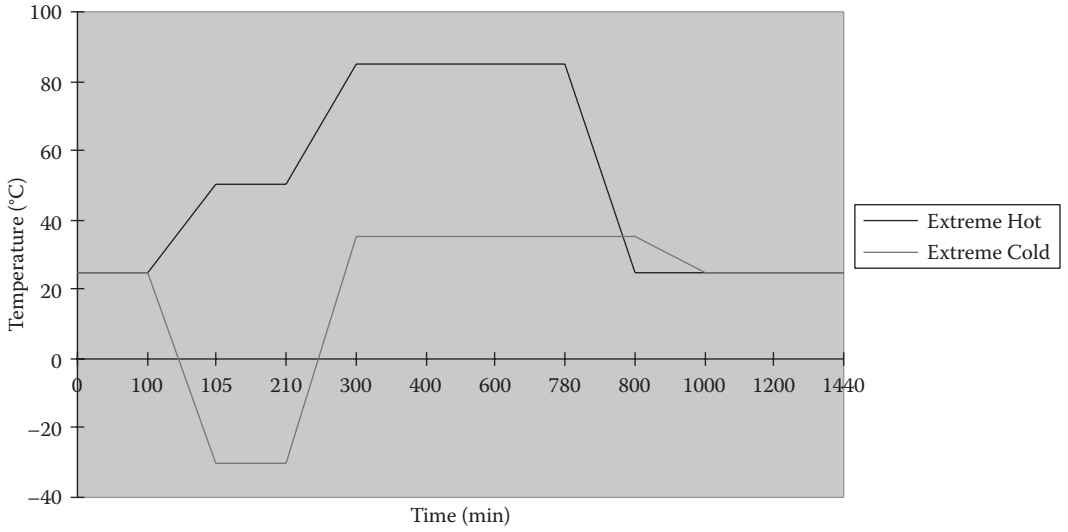


FIGURE 4.22.48 Schematic of temperature profile for extreme hot and extreme cold environments. (From Lall, P. and Banerji, K., Assembly-level reliability characterization of chip-scale packages, *Electronic Components and Technology Conference (ECTC)*, Seattle, Washington, DC, May 25–28, Session 13, Paper 2, 1998a; Lall, P. and Banerji, K., Assembly-level reliability of flex-substrate BGA, elastomer-on-flex packages, and 0.5 mm pitch partial array packages, *Surface Mount International Conference (SMI)*, San Jose, CA, August 23–27, 1998b. With permissions.)

importantly the field profile. Ideally, one would like to have one acceleration transform, which can be used to evaluate the field life once the life in accelerated test has been evaluated.

Unfortunately there is no universal acceleration transform—further, there can be no such universal relationship. The reason being that the variables that influence this acceleration transform vary from component to component and thus can be derived only for a class of components, e.g., plastic BGAs, ceramic BGAs, TSOPs, pendulum, etc., in specific product architectures. Care has to be exercised in using such a simplification of the damage mechanics as the "acceleration transform."

The acceleration transform lumps into itself several geometry and architecture variables integral to the component, solder joint geometry, and the assumed field profile. If any of the variables change, the acceleration transform has no option but to change. For this reason, acceleration transforms make sense for a particular set of conditions and for particular components. The field profile is the temperature–time plot of the field environment for the radio. The acceleration transform represents the relationship between the accelerated test life and field life for a particular accelerated test and particular field profile. Solder is viscoplastic in nature—which means that it is sensitive to absolute value of temperature, temperature ramp rate, and time spent at temperature. Changes in either the field profile or accelerated test conditions change the damage sustained in solder and thus the joint reliability. *For this reason, the acceleration transforms presented in this paper are valid for the accelerated test conditions and environmental profiles for which they have been derived.*

Figure 4.22.49 shows the acceleration transforms for the 0.5 mm pitch, 48 I/O Partial-Array Package and the Flex-Substrate BGA 0.8 mm pitch, 48 I/O package. The vertical bars indicate the time-to-first failure for the package in each environment. The acceleration transforms can be used to calculate the accelerated test requirement for each package or inversely predict field life in any field use environment or any combination thereof. It is known from Miner's superposition rule that

$$\sum_{i=1}^{m} \frac{n_i}{N_i} = 1, \tag{4.22.7}$$

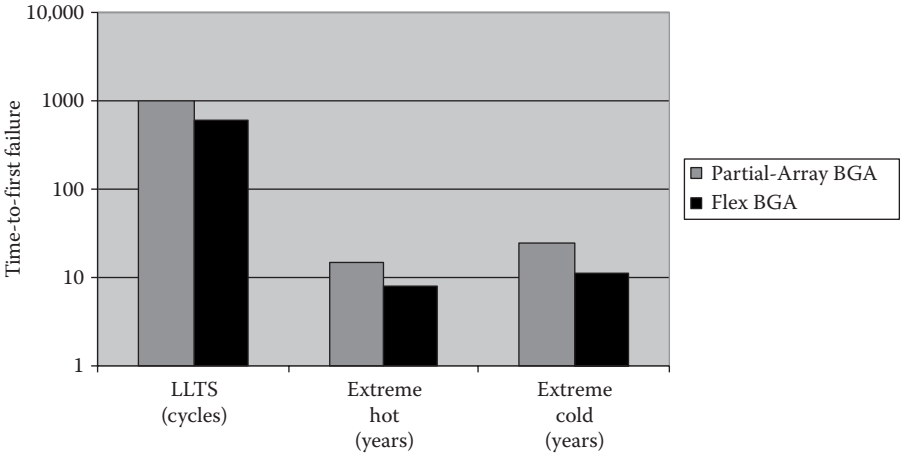


FIGURE 4.22.49 Acceleration transforms for Partial-Array BGA (0.5 mm pitch, 48 I/O Package) and Flex-Substrate BGA (0.8 mm pitch, 48 I/O Package) subjected to thermal fatigue. (From Lall, P. and Banerji, K., Assembly-level reliability characterization of chip-scale packages, *Electronic Components and Technology Conference (ECTC)*, Seattle, Washington, DC, May 25–28, Session 13, Paper 2, 1998a; Lall, P. and Banerji, K., Assembly-level reliability of flex-substrate BGA, elastomer-on-flex packages, and 0.5 mm pitch partial array packages, *Surface Mount International Conference (SMI)*, San Jose, CA, August 23–27, 1998b. With permissions.)

where n_i is the number of cycles in a particular environment, and N_i is the number of cycles to failure in the same environment, Miner's superposition rule states that the cumulative damage at end of life will be equal to 1. For example, the accelerated test requirement for a 10-year field life can be calculated, based on Figure 4.22.49 and Miner's superposition, as follows: $-n_i = 10$ years; $N_i = 15$ years for Extreme Hot Environment ($m = 1$); and $N_i = 1025$ for accelerated test environment, we want to find out $n_i = ?$

$$n_i = \frac{10}{15} \times (1025) = 684 \text{ cycles}. \tag{4.22.8}$$

Thus, 684 cycles of LLTS are required for the 0.5 mm pitch, 48 I/O Package to guarantee a 5-year life in Extreme Hot. The above method can also be used to calculate the accelerated test requirement for any combination of the above environments. For example, we want to calculate the accelerated test requirement for a product which will be in Extreme Hot for 6 years and in Extreme Cold for 4 years—$n_1 = 6$ years; $N_1 = 15$ years for Extreme Hot Environment; $n_2 = 4$ years; $N_2 = 25$ years for Extreme Cold Environment; and $N_i = 1025$ for accelerated test environment, we want to find out $n_i = ?$

$$n_i = \left(\frac{6}{15} + \frac{4}{25}\right) \times (1025) = 574 \text{ cycles}. \tag{4.22.9}$$

Thus, 574 cycles of LLTS are required for the 0.5 mm pitch, 48 I/O Package to guarantee a 10-year life (6 years in Extreme Hot and 4 years in Extreme Cold) (Lall et al. 1998).

Outdoor Telecommunication Electronics

Telecommunication electronics is extremely decentralized—located away from the controlled and protected environment of the central office—outside the plant (OSP).

Understanding the Environment

The OSP environment contains rain and humidity, dust and pollutants, significant daily and annual temperature swings, wide solar heat-load variations, and physical abuse. To protect electronics from this hostile environment, they are housed in a variety of OSP enclosures, which can range from small buildings (30 m^3) to small boxes (3×10^{-3} m3) that are mounted on the homes of customers. Heat densities within OSP enclosures can be large, often rivaling those of electronics designed for controlled indoor environments. Insolation is the amount of solar energy that is absorbed and transmitted into the enclosure.

The second source of heat in an outdoor enclosure is the solar energy that is absorbed and transmitted into the enclosure. The total insolation includes direct shortwave radiation from the sun, diffuse shortwave radiation that has been scattered by the atmosphere, and any shortwave radiation that is reflected from the ground or nearby surfaces. It is a complicated function of the orientation of the enclosure with respect to the earth–sun line; the amount of moisture, dust, and contaminants in the air; and the solar reflectivity and location of nearby surfaces (e.g., landscaping, fences, and buildings). Procedures for calculating the total shortwave radiation for an enclosure with a known location and orientation are well documented. [See References ASHRAE (1977) and HVAC (1987).]

Typically, outdoor enclosures are designed for the maximum solar load at the maximum ambient air temperature and the maximum internal heat load (Bellcore 1993). The maximum solar load is a function of the enclosure location and orientation. Figure 4.22.50 shows the total insolation incident on a 2 ft×4 ft×6 ft enclosure located in a low-reflecting open field environment (i.e., no surrounding reflecting surfaces). The shortwave absorptance of the enclosure surface affects the solar load absorbed into the enclosure. Table 4.22.31 shows typical values of the shortwave absorptance and longwave emittance for common enclosure surface treatments including polished aluminum and oil-based paints (white, light green, and light gray). Shortwave absorptance controls the fraction of

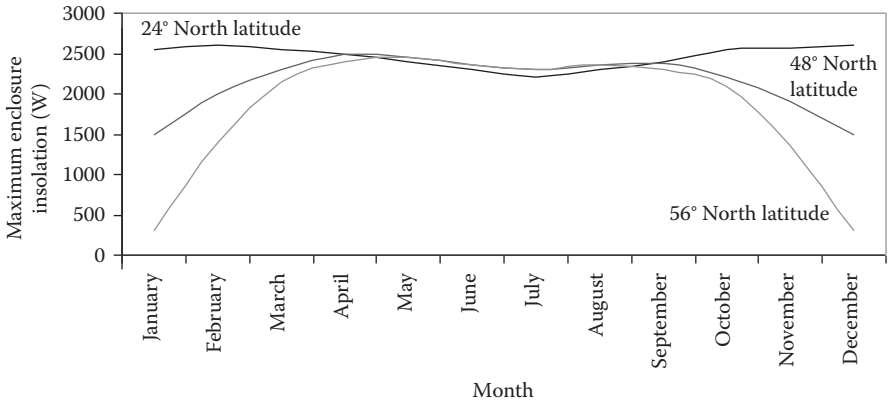


FIGURE 4.22.50 Maximum enclosure insolation versus calendar month and latitude.

TABLE 4.22.31
Typical Values of Shortwave Absorbtance for Common Enclosure Surface Treatments

Surfaces	Shortwave Absorptance	Longwave Emittance
Polished aluminum	0.03	0.05
White (oil-based paint)	0.2	0.9
Light green (oil-based paint)	0.5	0.9
Light gray (oil-based paint)	0.75	0.9

incident solar heat absorbed into the enclosure, and the longwave emittance controls the heat that is dissipated by radiation to the surroundings.

Hot surfaces radiate heat continuously in proportion to the fourth power of their absolute temperature (Stefan–Boltzmann law); the net heat radiated depends on the longwave emissivity of the surface and on the emissivities and temperatures of the surroundings.

The major heat sink for the heat dissipated from the enclosure is the local ambient air. The temperature of this sink varies widely with location, time of year, and time of day. Typical specification on the ambient air temperature range is −40°C to +46°C (−40°F to +115°F) (Bellcore 1993). Sky temperature is always lower than ambient air temperature, especially when the air contains little water vapor: This is the reason that frost occurs on terrestrial surfaces even when the air temperature is well above freezing. The atmosphere/remote sky is most effective as a radiant heat sink during the spring and fall, when dew point and dry-bulb temperatures are relatively low, and in dry dessert climates. It is not effective under overcast conditions or summer conditions of high dew point/dry-bulb temperatures.

Typical Heat Transfer Methods

Passive cooling methods are predominantly used in outdoor enclosures on account of their simplicity and lack of maintenance. The most common cooling method used today in the OSP is relatively simple. The enclosure is unventilated to protect the internal equipment from the rain, dust, and contaminants in the outside air. The internal heat is transferred primarily by convection to the inside surfaces of the enclosure, by conduction through the walls of the enclosure, and then by convection and radiation to the external heat sinks. Other passive methods include natural (free) convection in conjunction with phase change materials (PCMs) and solar reflectors. PCMs are substances that change phase, most often from solid to liquid, as they absorb heat. PCMs are selected for the temperature at which they change phase and for the latent heat associated with phase change. PCMs are sometimes used in conjunction with thermosiphons. Typical PCMs for high-temperature applications include waxes, salts, and paraffins. Water (ice) is used for low-temperature applications. The PCM is kept inside or attached to the enclosure in appropriately designed and sealed reservoirs (Figure 4.22.51). PCMs take advantage of thermal inertia and phase change effects. For example, an enclosure with PCMs during daylight hours will absorb heat through the cabinet walls and protect the electronics within the enclosure from overheating. The heat absorbed during the day will then be released to the outside world at night when it is cooler.

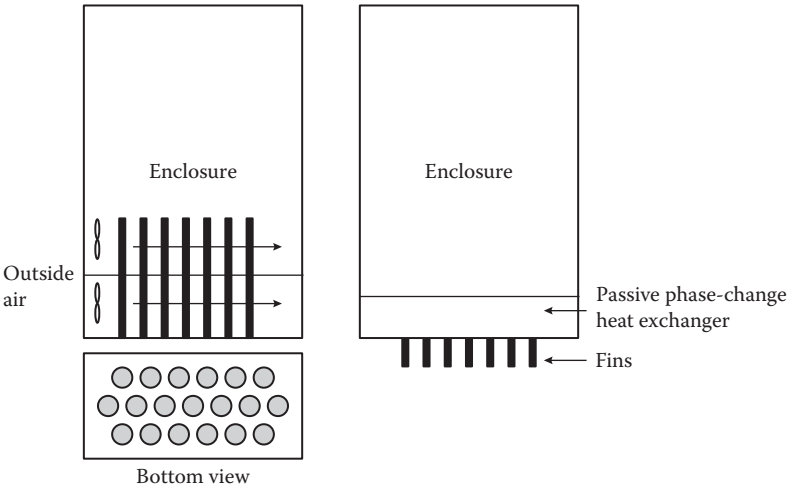


FIGURE 4.22.51 Heat transfer scheme for an outdoor enclosure using phase change materials. (From Marongiu, M.J. and Clarksean, R., Thermal management of outdoor enclosures using phase change materials, *Electronics Cooling*, Vol. 4, No. 1, January 1998.)

PCMs can also be incorporated into assisted systems for the cooling of enclosures. In order to enhance cooling, PCMs can be incorporated into a heat exchanger structure in which two fluids that are at different temperatures are separated by a PCM (possibly in encapsulated form). Hotter, internal air is first circulated through the heat exchanger and cooled by transferring its energy to the PCM, which slowly changes phase. This will occur during the hottest part of the day when external air temperatures prevent the use of outside air to cool the enclosure. Later in the day, when outside air temperatures drop, outside air is brought in to remove the heat stored in the PCM. Figure 4.22.51 shows a sketch for this simplified heat exchanger (Marongiu and Clarksean 1998).

High-Altitude Airborne Electronics

The heat transfer scheme is chosen based on the electronic module power dissipation surface density. Free convection is favored for power dissipation densities less than 500 W/m^2 (0.05 W/cm^2) on the surface of a fully populated PCB. Forced convection is used for power dissipation densities up to 4000 W/m^2. Heat pipes, liquid cooling, and boiling heat transfer is favored for power dissipation densities greater than 4000 W/m^2 (0.4 W/cm^2). This upper limit for forced air cooling is dictated by the fact that heat sinks cannot be used on the circuits due to the lack of space in avionics racks. The boards are designed for a minimum pitch.

Understanding the Environment

Figure 4.22.52 shows the decrease in density (kg/m^3) with increase in altitude (km). Air pressure at high altitudes greater than 21 km (70,000 ft) is so low that very little air is available for cooling. The volumetric flow (m^3/s) needed to cool the unit increases with increase in altitude. The static air temperature decreases with increase in altitude. In most hot day cases, the air temperature decreases with altitude. Typically, the reduction in temperature only occurs up to approximately 12 km (40,000 ft). On most cold days, the temperature actually increases with altitude in the first 3 km then the temperature begins to decrease. Cold days are not of great concern because the hot day defines the design point for worst case cooling conditions.

Impact of Equipment Location

Location of the equipment on aircraft affects electronic cooling designs based on the different qualities of air. Equipment located in the pressurized bay may be cooled by forced air cooling using

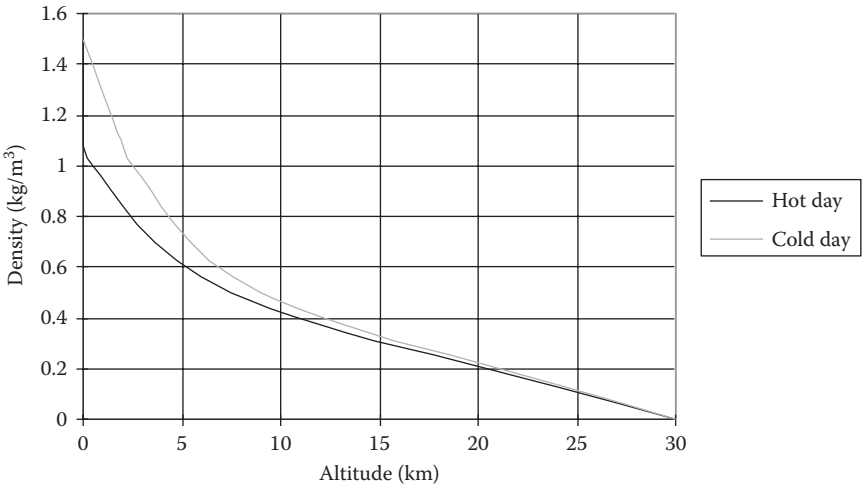


FIGURE 4.22.52 Variation in air density with altitude. (From Marthinuss Jr., J.E. and George, T.H., Cooling electronics at high altitudes made easy, *Electronics Cooling*, Vol. 4, No. 3, September 1998.)

a fan. The flow rate (Q) can be calculated from the overall power dissipation (P_d) and air temperature increase through the equipment by

$$Q = \frac{P_d}{(T_{out} - T_{in})\, m_p c_p}, \tag{4.22.10}$$

where T is the air temperature increase through the equipment, the subscripts out and in indicate the air temperature at the outlet and the inlet, m_p is the air specific mass, and c_p is the specific heat. The pressure drop can be calculated at the flow rate Q ($\nabla P = f(Q)$). A fan can be selected for the application based from a manufacturer's catalog using these two characteristics (Q, ∇P). Air velocity distribution obtained on each circuit should provide (1) a convective heat transfer coefficient sufficient to limit the temperature gradient between the components and the air and (2) provide uniformity of air flow distribution on each board. Local heat sinks can also be used on high power components, if space is available. Alternatively, a more powerful fan can be selected.

Electronics located in depressurized bays experiences a decrease in air density with increase in altitude. The air density decreases with decrease in atmospheric pressure. The air mass flow rate decreases with decrease in air density. This means that a fan with a higher volumetric flow is needed. Lower mass flow rate increases the outlet air temperature going through the equipment. This may or may not be compensated by the lower static air temperature at the inlet for higher altitudes.

Fighter aircraft typically have an environment control system—which can provide cool air to the electronic equipment. Flow rates and pressure drops obtained in such an environment is sufficient to cool the electronics efficiently. The main problem of forced convection cooling in such situation is due to air pollution. This air contains particles such as dust, sand, metals, and fluid droplets, such as oil, kerosene, and liquid water. Unfiltered air used for electronics cooling increases the risk of micro short circuits, abrasion, or chemical degradation of the components. Cooling with unfiltered air is obtained by avoiding contact between the components and the pollutants—air is passed through cold plates installed at the board level or at the rack level.

Figure 4.22.53 shows a compact-finned cold plate embedded in a double-sided circuit. This is an air flow-through module. The heat generated by the electronic components is transferred first by conduction through the board and then by forced air convection. The forced convection heat transfer coefficient is enhanced by the compact fins of the cold plate. Air is usually directed along the width of the circuit boards in order to limit the pressure drop.

Figure 4.22.54 shows a rack that has two cold plates cooled by the environmental control system air. The rack includes circuit boards, which are cooled by conduction enhanced by a metal (generally aluminum or copper) plate in contact with the wall cold plates through thermal clamps located on each side of the circuit.

Direct air forced convection on components may authorize power dissipation densities on a fully populated board, up to 3000 W/m^2, with an air velocity in the range of 5 m/s. This type of design requires nonpolluted air and a powerful fan on account of the high velocity of 5 m/s, which will generate high-pressure drops. An Air Flow-Through Module (AFT) may allow power densities up

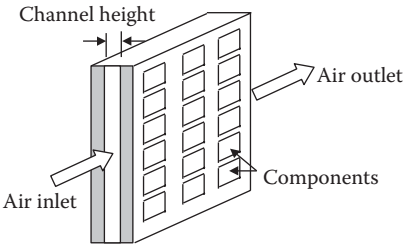


FIGURE 4.22.53 Schematic of a typical air flow-through module. (From Assouad, Y., Forced convection cooling of airborne electronics, *Electronics Cooling*, Vol. 3, No. 2, May 1997.)

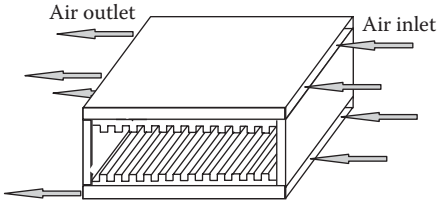


FIGURE 4.22.54 Schematic of a typical air-cooled rack for conduction-cooled electronic modules. (From Assouad, Y., Forced convection cooling of airborne electronics, *Electronics Cooling*, Vol. 3, No. 2, May 1997.)

to 3000 W/m^2 or more, but with a very large pressure drop penalty. For 3000 W/m^2, the pressure drop will be in the order of 5 mbar for a 100 mm wide module, which yields, for a 200 mm wide classical module, a total pressure drop of 10 mbar, not accounting for the distribution and collecting air chambers (Assouad 1997). Conduction-cooled modules in an air-cooled rack are limited to power densities up to 1500 W/m^2, due to the parabolic temperature gradient from conduction in the metal plate embedded in the PCB.

Fan Selection

Fan selection is based on attainable flow rate and pressure drop at the higher operating altitudes. The attainable flow rate is determined by the fan supply curve and the pressure drop is determined by the system pressure drop curve. The intersection of the two curves determines the operating point.

Conventional fans run at a constant speed. The operating point of a conventional fan varies with density since the pressure drop and flow rate curves vary with air density. The volumetric flow rate for a conventional fan is constant, while the pressure output is drastically different at different altitudes. The decrease in density produces a decrease in pressure drop and the volume flow rate.

$$\frac{Q_{h1}}{Q_{h2}} = \left(\frac{D_{h1}}{D_{h2}}\right)^3 \left(\frac{\omega_{h1}}{\omega_{h2}}\right), \tag{4.22.11}$$

where Q is the volume flow rate, D is the fan diameter, ω is the fan's rotational velocity (RPM), and the subscripts $h1$ and $h2$ indicate fan operation at heights $h1$ and $h2$.

$$\frac{\Delta P_{h1}}{\Delta P_{h2}} = \left(\frac{D_{h1}}{D_{h2}}\right)^3 \left(\frac{\omega_{h1}}{\omega_{h2}}\right)^3 \left(\frac{\rho_{h1}}{\rho_{h2}}\right), \tag{4.22.12}$$

where ΔP is the pressure drop, D is the fan diameter, ω is the fan's rotational velocity (RPM), ρ is the air density, and the subscripts $h1$ and $h2$ indicate fan operation at heights $h1$ and $h2$. Since the fan diameter and fan rotational velocity is constant, the equations can be reduced to

$$Q_{h1} = Q_{h2}, \tag{4.22.13}$$

$$\Delta P_{h1} = \Delta P_{h2} \left(\frac{\rho_{h1}}{\rho_{h2}}\right). \tag{4.22.14}$$

Equation 4.22.13 shows that the volumetric flow rate is constant with height for a conventional fan. Equation 4.22.14 shows that the pressure output of the fan is directly proportional to the density of the air. Thus, the mass flow rate will drop even though the volumetric flow rate is constant since the two are related as follows:

$$Q_h = \frac{\dot{m}_h}{\rho_h}, \tag{4.22.15}$$

where Q is the volume flow rate, $\dot{m}$ is the mass flow rate, ρ is the air density, and the subscript h denotes the altitude.

A slip fan is very different from a conventional fan. A slip fan allows the blades to slip or run at different speeds from the motor driving the fan, which allows the fan to work under many different altitudes and changing density conditions. The pressure flow rate curves for a slip fan at different altitudes are more consistent than a conventional or nonslip fan.

Summary

The need and applicability of cooling techniques in electronics has been discussed in terms of material limits, performance, and reliability. The use of acceleration transforms for determining the required operational temperature for temperature-dependent mechanisms has been discussed. The thermal management of electronic systems has been discussed based on application temperature constraints. Typical heat transfer schemes for each application environment have been discussed.

REFERENCES

AMD (Advanced Micro Devices), *AMD-K6 Processor Thermal Solution Design*, Application Note, Document Number 21085, Revision H, Amendment 0, November 1998.

ASHRAE, *Handbook and Product Directory—1977 Fundamentals*, Atlanta, GA: ASHRAE, Inc., 1977, Chapter 26.

Assouad, Y., Forced convection cooling of airborne electronics, *Electronics Cooling*, Vol. 3, No. 2, 1997.

Azar, K. and Graebner, K., Experimental determination of thermal conductivity of printed wiring boards, *Proceedings, SEMI-THERM XII Conference*, Austin, TX, March 1996, pp. 169–182.

Bellcore TA-NWT-000487, *Generic Requirements for Electronic Equipment Cabinets*, Issue 2, June 1993.

Graebner, J.E., Thermal conductivity of printing wiring boards, technical brief, *Electronics Cooling Magazine*, Vol. 1, No. 2, 1995, p. 27.

HVAC Systems and Applications – *1987 ASHRAE Handbook*, Atlanta, GA: ASHRAE, Inc., 1987, Chapter 47.

Intel Corporation, Low-profile fan duct board and chassis specification and low profile fan duct system ingredients specification, Available from http://developer.intel.com/ial/sdt/fanduct.htm, Revision 1.0, September 1998a.

Intel Corporation, Mobile power guidelines 2000, Available from http://developer.intel.com/design/ mobile/ intelpower/, Revision 0.8, September 16, 1998b.

Lall, P. and Banerji, K., Solder joint design for optimum process and reliability, *Surface Mount International (SMI) Conference*, San Jose, CA, September 9–11, pp. 97–108, 1997.

Lall, P. and Banerji, K., Assembly-level reliability characterization of chip-scale packages, *Electronic Components and Technology Conference (ECTC)*, Seattle, Washington, DC, May 25–28, Session 13, Paper 2, 1998a.

Lall, P. and Banerji, K., Assembly-level reliability of flex-substrate BGA, elastomer-on-flex packages, and 0.5 mm pitch partial array packages, *Surface Mount International Conference (SMI)*, San Jose, CA, August 23–27, 1998b.

Lall, P., Deshpande, S, Luo, Y., Bozack, M., Nguyen, L., and Murtuza, M., Degradation mechanisms in electronic mold compounds subjected to high temperature in neighborhood of 200°C, *Electronic Components and Technology Conference, 64th ECTC*, Orlando, FL, May 27–30, pp. 242–254, 2014.

Lall, P., Gold, G., Miles, B., Banerji, K., Thompson, P., Koehler, C., and Adhihetty, I., Reliability characterization of the SLICC package, *Electronic Components and Technology Conference (ECTC)*, Orlando, FL, May 28–31, pp. 1202–1210, 1996.

Lall, P., Pecht, M., and Hakim, E., *Influence of Temperature on Microelectronic and System Reliability*, Boca Raton, FL: CRC Press, 1997.

Lall, P., Shantaram, S., and Locker, D., High strain rate properties of SAC105 and SAC305 leadfree alloys after extended high temperature storage, *Proceedings of the SMTA*, Ft. Worth, TX, October 13–17, pp. 363–377, 2013.

Lee, T.T., Chambers, B., and Ramakrishna, K., Thermal management of handheld telecommunication products, *Electronics Cooling*, Vol. 4, No. 2, pp. 30–33, 1998.

Liu, S., Leung, B., Neckar, A., Memik, S.O., Memik, G., and Hardavellas, N., Hardware/software techniques for DRAM thermal management, *Proceedings of the 17th IEEE International Symposium on High Performance Computer Architecture (HPCA)*, San Antonio, TX, pp. 1–11, 2011.

Marongiu, M.J. and Clarksean, R., Thermal management of outdoor enclosures using phase change materials, *Electronics Cooling*, Vol. 4, No. 1, pp. 26–29, 1998.
Marthinuss Jr., J.E. and George, T.H., Cooling electronics at high altitudes made easy, *Electronics Cooling*, Vol. 4, No. 3, pp. 26–29, 1998.
Mulgaonker, S., Chambers, B., Mahalingam, M., Ganesan, G., Hause, V., Berg, H., Thermal performance limits of the QFP family, *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A*, Vol. 17, No. 4, pp. 573–582, 1994.
Mulgaonker, S., Chambers, B., Mahalingam, M., An Assessment of the Thermal Performance of the PBGA Family, *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A*, Vol. 18, No. 4, pp. 739–748, 1995.
Steinbrecher, T., *Heat-Sink Guide*, http://www.heatsink-guide.com/, 1998.
Sterner, J., Thermal Design Challenges in a Pentium Notebook Computer, *Proceedings of the IMAPS International Systems Packaging Symposium*, December 1997, pp. 107–112, 1997.
Zhu, Q., Li, X., and Wu, Y., Thermal management of high power memory module for server platforms, *Proceedings of the ITHERM Conference*, Orlando, FL, pp. 572–576, 2008.

Selected U.S. Electronic Thermal Management Patents

Active thermal isolation for temperature responsive sensors, Martinson, S.D., Hampton, VA; Gray, D.L., Newport News, VA; Carraway, D.L., Virginia Beach, VA; Reda, D.C., San Jose, CA. 5,311,772, May 17, 1994.
Active vaned passage casing treatment, Hobbs, D.E., South Windsor, CT. 5,431,533, July 11, 1995.
Air distribution system and manifold for cooling electronic components, Bartilson, B.W., Chippewa Falls, WI; Jirak, J.J., Jim Falls, WI. 5,321,581, June 14, 1994.
Apparatus for an air impingement heat sink using secondary flow generators, Lindquist, S.E., Boylston, MA; Bailey, D.A., Concord, MA. 5,304,845, April 19, 1994.
Appliance for attaching heat sink to pin grid array and socket, Perugini, M.N. Monroe, CT; Romatzick, Jr, D.H., Milford, CT. 5,396,402, March 7, 1995.
Cast casing treatment for compressor blades, Privett, J.D., North Palm Beach, FL; Byrne, W.P., Jupiter, FL; Nolcheff, N.A., Palm Beach Gardens, FL. 5,474,417, December 12, 1995.
Centrifugal compressor with a flow-stabilizing casting, Heinrich, K., Nussbaumen, Switzerland. 5,466,118, November 14, 1995.
Circuit with built-in heat sink, Mehta, A.M., Plantation, FL; Desai, V.D., Plantation, FL. 5,307,519, April 26, 1994.
Clib for clamping heat sink module to electronic module, Harmon, R.A., Hudson, MA; Urrata, Giovanni, W., MA. 5,381,305, January 10, 1995.
Clip-on heat sink, Villani, A., Shrewsbury, MA. 5,329,426, July 12, 1994.
Clip-on heat sink and method of cooling a computer chip package, Beijer, G., Moorpark, CA. 5,295,043, March 15, 1994.
Colored liquid crystal display having cooling, Hyatt, G.P., 7841 Jennifer Cir., La Palma, CA. 5,398,041, March 14, 1995.
Cooling device for a magnetron, Lim, J.H., 1429–18 Sankyeok 3ku, Buk-ku Daeku-si, Republic of Korea. 5,325,266, June 28, 1994.
Cooling system for integrated circuits, Hastings, R.J., Houston, TX; Davis, C.E., Houston, TX. 5,285,108, February 8, 1994.
Custom conformal heat sinking device for electronic circuit cards and methods of making the same. Santilli, M.A., 4840 Elkridge Rd., Lake Shastina, CA; Morgan, W.Z., 9714, Forest La. 6005 Dallas, TX. 5,403,973, April 4, 1995.
Disk augmented heat transfer system, Maling Jr., G.C., Schmidt, R.R. Both of Poughkeepsie, New York. 5,335,143, August 2, 1994.
Dual active surface, miniature, plug-type heat flux gauge, Liebert, C.H., Middleburg Heights, OH; Koch, Jr., J., Medina, OH. 5,314,247, May 24, 1994.
Electrical component assembly with heat sink, Cocconi, A.G., 725 S. Scottsdale Ave., Glendora, CA. 5,060,112, October 22, 1991.
Epoxy molding composition for surface mount applications, Gallagher, M.K., Lansdale, PA; Petti, M.A., North Wale, PA. 5,434,199, July 18, 1995.
Exposure control method for adjusting the temperature of a workpiece holding chuck attracting surface based on memorized data, Fujioka, H., Atsugi, Japan. 5,134,436, July 28, 1992.
Fin assembly for an integrated circuit, Chen, C.-C., Yi, L.H., China (Taiwan). 5,287,249, February 15, 1994.

Flexible thermal transfer apparatus for cooling electronic components. Tousignant, L.A., Shoreview, MN. 5,411,077, May 2, 1995.

Flow meter for measuring fluid flow with multiple temperature sensors, Ishikawa, K., Kyoto, Japan; Mihira, H., Kyoto, Japan; Kimura, N., Kyoto, Japan; Yamaguchi, M., Kyoto, Japan. 5,048,332, September 17, 1991.

Flow control method and means, Freeman, C., Farnsfield, England; Day, I.J., Cambridge, England. 5,275,528, January 4, 1994.

Flow control method and means, Freeman, C., Farnsfield, England; Day, I.J., Toft, England; Wright, W.B., Kegworth, England. 5,340,271, August 23, 1994.

Heat sink, Voorhes, D.W., Winchester, Massachusetts Goldman, R.D., Stoughton, MA; Lopez, R. R., Boxford, MA. 5,390,734, February 21, 1995.

Heat sink and cover for tab integrated circuits, Selna, E., Mountain View, CA; Ettehadish, E., Albany, CA; LaGassa, James, C., CA. 5,280,409, January 18, 1994.

Heat sinking apparatus for surface mountable power devices, Kent, K.L., Woodstock, IL; Glomski, J.D., Arlington Heights, IL. 5,365,399, November 15, 1994.

Heat sink assembly for solid state devices, Tata, P.D., Johnston; Rife, W.B., Greenville, RI. 5,397,919, March 14, 1995.

Heat sink attachment assembly, Jordan, W.D., Dallas, TX; Smithers, M.C., Lewisville, TX. 5,396,338, January 31, 1995.Heat sink device for electronics modules packaged in cylindrical castings, Sewell, M.W., Fredericksburg, VA. 5,060,115, October 22, 1991.

Heat sink for an electric circuit board, Andersson, B., Jonkoping, Sweden; Danielsson, H., Linkoping, Sweden; Johansson, H., Angstromm Anstroml, Sweden; Andersen, Hakan, V., Sweden. 5,077,638, December 31, 1991.

Heat sink method of manufacturing the same, and device of manufacturing the same, Nishiguchi, M., Yokohama, Japan. 5,339,215, August 16, 1994.

Heat sink mounting apparatus, Smithers, M.C., Lewisville, TX. 5,276,585, January 4, 1994.

Heart sink mounting system for semiconductor devices Goeschel, F.G., Mt. Clemens, MI, Lanting, M.L., Portage, MI, McConnell, A.M., Dearborn, MI. 5,283,467, February 1, 1994.

Heat sink plate for multiple semi-conductors, Villaume, H.F., Intervale, NH. 5,285,350, February 8, 1994.

High conduction cooling module having internal fins and complaint interfaces for VLSI chip technology, Horvath, J.L., Poughkeepsie, New York; Biskeborn, R.G., Pawling, New York; Harvilchuck, J.M., Billings, New York. 5,052,481, 1991.

High heat flux evaporative spray cooling, Tilton, D.E., Kent, Washington, DC; Tilton, C.L., Rent, Washington, DC. 5,220,804, June 22, 1993.

High performance centrifugal pump having an open-faced impeller, Lehe, J.-R., Vernon, France; Beaurain, Andre, C., France; Bosson, R., Vernon, France; Tiret, E., La Chapelle Reanville, France. 5,427,498, June 27, 1995.

Housing with recirculation control for use with banded axial-flow fans, Yapp, M.G., Needham, Massachusetts; Houten, R.V., Winchester, MA; Hickey, R.I., Concord, NH. 5,489,186, February 6, 1996.

Hybrid cooling system for electronic components, Fox, L., Boxboro, MA; Wade, P.C., Shirley, MA. 5,285,347, February 8, 1994.

Integral heat pipe, heat exchanger, and clamping plate, Davidson, H.L., San Carlos, CA; Ettenhadieh, E., Albany, CA. 5,253,702, October 19, 1993.

Integral heat pipe, heat exchanger and clamping plate, Ettehadieh, E., Albany, CA. 5,329,993, July 19, 1994.

Integral heat sink-terminal member structure of hybrid integrated circuit assembly and method of fabricating hybrid integrated circuit assembly using such structure, Matsumoto, H., Itami, Japan. 5,291,372, March 1, 1994.

Integrated circuit package having a cooling mechanism, Omeozawa, K., Tokyo, Japan. 5,329,426, July 12, 1994.

Liquid crystal display having conductive cooling, Hyatt, G.P., 7841 Jennifer Cir., La Palma, CA. 5,432,526, July 11, 1995.

Localized circuit card cooling device, Yiu, J.T., Chandler, AZ. 4,931,904, June 5, 1990.

Low profile integrated heat sink and fan assembly, Bailey, N.W., Sacramento, CA. 5,309,983, May 10, 1994.

Mechanical heat transfer device, Altoz, F.E., Catonsville, MD; Winn, W.H., Linthicum, MD. 4,273,183, June 16, 1981.

Metal heat sink baseplate for a semiconductor device, Della Bosca, P., Brugherio, Italy; Massironi, Angelo, C., Italy. 5,229,918, July 20, 1993.

Metal matrix composite heat transfer device and method, Monesano, M.J., Fairfax, VA. 5,287,248, February 15, 1994.

Method for repeatable temperature measurement using surface reflectivity, Thakur, R.P.S., Boise, ID; Sandhu, G.S., Boise, ID; Martin, A.L., Boise, ID. 5,350,236, September 27, 1994.
Method of making microchanneled heat exchangers utilizing sacrificial cores; Hoopman, T.L., River Falls, WI; Krinke, H.L., Marine, Washington, DC. 5,317,805, June 7, 1994.
Microprocessor heat dissipation apparatus for a printed circuit board, Koenen, D.J., Houston, TX. 5,272,599, December 21, 1993.
Mist supercooling of a heated surface, Lee, R.S.L., St. James, NY. 5,311,931, May 17, 1994.
Monolith heating element containing electrically conductive tin oxide containing coatings, Clough, T.J., Santa Monica, CA; Grosvenor, V.L., Topanga, CA; Pinsky, M., Thousand Oaks, CA. 5,304,783, April 19, 1994.
Organic coatings with ion reactive pigments especially for active metals, Mosser, M.F., Sellersville, PA; Harvey, III, W.A., MontClare, PA. 5,409,970, April 25, 1995.
Package clip on heat sink, Blomquist, M.L., 3474 Carl Ct. Newbury Park, CA. 5,367,433, November 22, 1994.
Performance characteristics stabilization in a radial compressor, Foerster, A., Worms, Germany; Engels, Berthold, Weisenheim am Sand, Germany; Hauck, P., Ludwigshafen, Germany. 5,333,990, August 2, 1994.
Pin-fin heat sink, Brady, K.J.M., NJ; Cohn, C., Wayne, NJ. 5,299,090, March 29, 1994.
Pivotal heat sink assembly, Frankeny, J.A., Taylor, Texas Frankeny, R.F., Austin, TX; Hermann, Karl, A., TX; Wustrau, Rolf, A., TX. 5,161,087, November 3, 1992.
Plasma momentum meter for momentum flux measurements. Zonca, F., Rome, New Jersey, Italy; Cohen, S.A., Hopewell, NJ; Bennett, T., Princeton, NJ; Timberlake, J.R., Allentown, NJ. 5,239,563, August 24, 1993.
Powder epoxy resin coating composition, Matsuzaki, K., Yono, Japan; Ono, K., Tokyo, Japan; Iwamoto, S., Soka, Japan; Osa, M., Saitama, Japan; Watanabe, T., Soka, Japan. 5,418,265, May 23, 1995.
Printed circuit board with cooling monitoring system, Bartilson, B.W., Chippewa Falls, WI; Schlimme, E.F., Chippewa Falls, WI. 5,281,026, January 25, 1994.
Radiation detector suitable for tympanic temperature measurement, Pompei, F., Wellesley Hills, MA; Looney, Jr., J.M., Clearwater, FL. 5,271,407, December 21, 1993.
Rapid method for determining critical vapor pressure, Wood, R.W., Elkhorn, WI; Hansen, L.D., Orem, Utah; Crawford, J.W., Salem, Utah. 5,266,492, November 30, 1993.
Remote high temperature insulatorless heat-flux gauge, Noel, B.W., Espanola, NM. 5,273,359, December 28, 1993.
Remote temperature and/or temperature difference measuring device, Barral, J.-P., Seyssins, France; Million, B., Saint-MartinD'Hres, France; Colomb, Francois, Grenoble, France. 5,352,039, October 4, 1994.
Rotor case treatment, Koff, S.G., Cambridge, Massachusetts; Nikkanen, J.P., West Hartford, CT; Mazzawy, R.S., South Glastonbury, CT. 5,308,225, May 3, 1994.
Self clamping heat sink assembly, Brauer, E.A., Indianapolis, IN. 5,309,979, May 10, 1994.
Single flow turbopump with integrated boosting, Girault, J.-P.N., Tourny, France. 5,232,333, August 3, 1993.
Solid-state imaging device having temperature sensor; Miyaguchi, K., Hamamatsu, Japan; Muraki, T., Hamamatsu, Japan. 5,508,740, April 16, 1996.
Stacking heatpipe for three dimensional electronic packaging, Davidson, H.L., San carlos, CA; Nishtala, S., Santa Clara, CA. 5,181,167, January 19, 1993.
Surface-to-air heat exchanged for electronic devices, Pitasi, M.J., Newbury, MA. 5,293,930, March 1, 1994.
System for repeatable temperature measurement using surface reflectivity, Thakur, R.P.S. Boise, Idaho; Sandhu, G.S., Boise, ID; Martin, A.L., Boise, ID. 5,364,187, November 15, 1994.
Temperature controller for semiconductor device, Watanabe, H., Kanagawa, Japan. 4,689,659, August 25, 1987.
Thermal joint, Gruber, P.A., Mohegan Lake, NY; Zingher, A.R., White Plains, NY. 5,291,371, March 1, 1994.
Thermal transfer plate and integrated circuit chip or other electrical component assemblies including such plate, Hopfer, A.N., Park Ridge, IL. 5,282,111, January 25, 1994.
Thermopile differential thermal analysis sensor. Schaefer, J.W., Wilmington, DE; Danley, R.L., Collingswood, NJ. 5,288,147, February 22, 1994.Toughened fiber-reinforced composites, Qureshi, S.P., Alpharetta, Georgia; Hoffman, R.E., Cumming, Georgia; Newman-Evans, R.H., Alpharetta, Georgia. 5,268,223, December 7, 1993.
Translating wedge heat sink, Sewell, M.W. Dahlgren, VA. 5,218,517, Jun. 8, 1993.
Ultrathin film thermocouples and method of manufacture, Rimai, D.S., Webster, NY; Anne, R.E., Caledonia, NY; Bowen, R.C., Rochester, NY. 5,411,600, May 2, 1995.
Vaned shroud for centrifugal compressor, Palmer, D.L., Cave Creek, AZ. 5,277,541, January 11, 1994.

Bibliography

Ajiki, T., Sugumoto, M., Higuchi, H., and Kumada, S., A new cyclic biased T.H.B. Test for power dissipating IC's. *International Reliability Physics Symposium*, San Diego, CA, (1979), 118–126.

Alexander, D.R., An electrical overstress failure model for bipolar semiconductor components. *IEEE Transactions on Components, Hybrids and Manufacturing Technology*, 1 (1978), 345–353.

Arrhenius, S., Ueber die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren, Zeitschrift für Physikalische Chemie, 4, (1889), 226.

Assouad, Y., Caplot, M., and Gautier, T., *Advanced Cooling Techniques Comparison for Airborne Electronic Circuits*, Austin, TX: Semitherm, (1997).

Banerji, K. and Darveaux, R., Effect of aging on the strength of and ductility of controlled collapse solder joints, *First International Conference on Microstructural & Mechanical Properties of Aging Materials*, TMS, Warrendale, PA, (1993).

Beasley, K., New standards for old. *Quality and Reliability International*, 6 (1990), 289–294.

Berry, R.S., Rice, S.A., and Ross, J., *Physical Chemistry*, John Wiley & Sons, New York, (1980).

Blanks, Henry J., Arrhenius and the temperature dependence of non-constant failure rate, *Quality and Reliability Engineering International*, 6, (1990), 259–265.

Blech, I.A., Meieran, E.S., Direct transmission electron microscope observations of electrotransport in aluminum films. *Applied Physics Letters*, 11 (1967), 263.

Bowles, John, A survey of reliability prediction procedures for microelectronic devices, *IEEE Transactions on Reliability*, 41(1), (1992), 2–12.

Boyko, K.C. and Gelarch, D.L., Time dependent dielectric breakdown of 210A oxides. *27th Annual Reliability Physics Symposium,* Phoenix, AZ, *IEEE*, (1989), 1–8.

Bradley, E., Lall, P., and Banerji, K., Effect of thermal aging on the microstructure and reliability of ball grid array (BGA) solder joints, *Proceedings of Surface Mount International 1996*, San Jose, CA, September 8–12, (1996), 95–106.

Canali, C, Fatini, F., Gaviraghi, S., and Senin, A., Reliability problems in TTL-LS devices. *Microelectronics and Reliability*, 21 (1981), 637–651.

Clark, I.D., Harrison, L.G., Kondratiev, V.N., Szabo, Z.G., and Wayne, R.P., *The Theory of Kinetics*, Edited by C.H. Bamford and C.F.H. Tipper, Elsevier Scientific, Amsterdam, (1979).

CNET, *Recueil de Donnees de Fiabilitie du CNET (Collection of Reliability DATA from CNET)*, Centre National d'Etudes des Telecommunications (National Center for Telecommunication Studies) (1983).

Cowan, J.P., *Handbook of Environmental Acoustics*, New York: Van Nostrand Reinhold, (1994).

Coyne, J.C., An approximate thermal model for outdoor electronics cabinets, *The Bell System Technical Journal*, 61, (1982), 247–256.

Crook, D.A., Techniques for evaluating long term oxide reliability at wafer level. *IEDM Technical Digest*, (1978), 444–448.

Crook, D.A., Method of determining reliability screens for time dependent dielectric breakdown. *17th Annual Reliability Physics Symposium*, (1979), 1–7.

Darveaux, R., Banerji, K., Mawer, A., and Dody, G., *Reliability of Plastic Ball Grid Array Assembly,* Chapter 13, pp. 380–442, Ball grid array technology, edited by John Lau, McGraw Hill, Inc., New York, (1994).

DiStefano, T.H. and Shatzkes, M., Dielectric stability and breakdown in wide bandgap insulators. *Journal Vacuum Science Technology*, 12 (1975), 37.

Evans, M.G. and Polanyi, M., Inertia and driving force of chemical reactions. *Faraday Society (London) Transactions*, 34, (1938), 11.

Eyring, H., Lin, S.H., and Lin, S.M., *Basic Chemical Kinetics*, John Wiley & Sons, New York, 1980.

Fukuzawa, I., Ishiguro, S., and Nanbu, S., Moisture resistance degradation of plastic LSIs by reflow soldering. *Proceedings 23rd Annual International Reliability Physics Symposium*, Orlando, FL, (1985), 192–197.

Gaffeny, J., Internal lead fatigue through thermal expansion in semiconductor devices. *IEEE Transactions, Electronic Devices*, ED-15 (1968), 617.

Hakim, E.B., Reliability prediction: Is Arrhenius erroneous? *Solid State Technology*, (1990), 57.

Handbook of Reliability Data for Components Used in Telecommunication Systems, Issue 4 (HRD4), British Telecom, January (1987).

Harari, E., Dielectric breakdown in electrically stressed thin films of thermal SiO_2. *Journal of Applied Physics*, 49 (1978), 2478.

Harman, G.G., Metallurgical failure modes of wire bonds. *12th International Reliability Physics Symposium*, Las Vegas, NV, (1974), 131–141.

Hart, A., Teng, T.T., and McKenna, A., Reliability influences from electrical overstress on LSI devices. *18th Annual Proceedings International Reliability Physics Symposium*, Las Vegas, NV, (1980), 190–196.

Holland, S., Chen, I.C., Ma, T.P., and Hu, C., Physical models for gate oxide breakdown. *IEEE Electron Device Letters*, 5 (1984), 302.

Honda, K., Ohsawa, A., and Toyokura, N., Breakdown in silicon oxides: correlation with Cu precipitates. *Applied Physics Letters*, 45 (1984), 270.

Hsu, F.C. and Chiu, K.Y., Temperature dependence of hot electron induced degradation in MOSFET's. *IEEE Electron Devices Letters*, EDL-5 (1984), pp.148–150.

Hu, C., Hot electron effects in MOSFET'S. *IEDM Technical Digest*, (1983), 176.

Hu, J., Pecht, M., and Dasgupta, A., A probabilistic approach for predicting thermal fatigue life of wirebonding in microelectronics, *ASME Journal of Electronic Packaging*, 113 (1991), 275.IITRI/Honeywell, *VHSIC/ VHSIC-Like Reliability Prediction Modeling*, Contract No F30602–86-C-0261, 15 January (1988).

Klinger, D.J., On the notion of activation energy in reliability: Arrhenius, Eyring, and thermodynamics, *1991 Proceedings of Annual Reliability and Maintainability Symposium*, Las Vegas, NV, (1991), pp. 295–300.

Ko, P.K., *Hot Electron Effects in MOSFETS*, Ph.D Dissertation, EECS Dept., U.C. Berkeley (1982).

Kraus, A.D. and Bar-Cohen, A., *Thermal Analysis and Control of Electronic Equipment*, New York: Hemisphere Publishing, (1983).

Kuo, W. and Kuo, Y., Facing the headache of early failures: state of the art review of burn-in decisions. *Proceedings IEEE*, 71 (1983), 1257–1266.

LaCombe, D.J., Dening, D.C., and Christou, A., A new failure mechanism in thin gold films at elevated temperature. *20th Proceedings of International Reliability Physics Symposium*, San Diego, NV, (1982), 81–87.

Lall, P., Pecht, M., and Hakim, E., *Estimating the Role of Temperature on Microelectronic Device Reliability*, John Wiley & Sons, New York.

Lee, J., Chen, I.C., and Hu, C., Modeling and characterization of gate oxide reliability. *IEEE Transactions on Electron Devices*, 35 (1988a), 2268.

Lee, J., Chen, I.C., and Hu, C., Statistical modeling of silicon dioxide reliability. *26th Annual International Reliability Physics Symposium*, Monterey, CA, (1988b), 131–138.

Leturcq, P., Dorkel, J.-M., Napieralski, A., and Lachiver, E., A new approach to thermal analysis of power devices, *IEEE Transactions on Electron Devices*, ED24, (1987), 1147–1156.

Liehr, M., Bronner, G.B., and Lewis, J.E., Stacking-fault-induced defect creation in SiO_2 on Si(100). *Applied Physics Letters*, 52 (1988), 1892.

Lin, P.S.D., Marcus, R.B., and Sheng, T.T., Leakage and breakdown in thin oxide capacitors: correlation with decorated stacking faults. *Journal of Electrochemical Society*, 130 (1983), 1878.

Lloyd, J.R., Shatzkes, M., and Challaner, D.C., Kinetic study of electromigration failure in Cr/Al-Cu thin film conductors covered with polyimide and the problem of the stress dependent activation energy. *IEEE International Reliability Physics Symposium*, Monterey, CA, (1988), pp. 216–225.

LSI Logic, *Reliability Manual and Data Summary*, Milpitas, CA, (1990), 12–18.

Machiels, F., Lijbers, G., Allaire, R., and Poiblaud, G., Derating of results according to various models and prediction of lifetimes of plastic SMD's under humidity stress conditions. *5th International Conference Quality in Electronic Components, Failure Prevention, Detection and Analysis*, Bordeaux, France (1991), pp. 869–874.

Mahidhara, R., Solberg, V., DiStefano, T., and Greathouse, S., Solder joint integrity in Tessera's μBGA package, *Design and Reliability of Solders and Solder Interconnections*, TMS, Warrendale, PA, (1997), 403–412.

Matsumoto, H., Sawada, K., Asai, S., Hirayama, M., and Nagasawa, K., Effect of long term stress on IGFET degradations due to hot electron trapping. *IEEE Transactions on Electron Devices*, ED-28 (1981), 923–928.

McAteer, O.J., *Electrostatic Discharge Control*, McGraw-Hill Publishing Company, New York, (1989).

McLinn, J.A., Constant failure rate: A paradigm transition. *Quality and Reliability International*, 6 (1990), 237–241.

McPherson, J.W. and Baglee, D.A., Acceleration factors for thin gate oxide stressing. *23rd Annual Reliability Physics Symposium, IEEE*, Orlando, FL, (1985), 1–5.

MIL-HDBK 217F, *Reliability Prediction of Electronic Equipment,* MIL-HDBK 217F, U.S. Department of Defense, Washington DC (1991).

Moss, R.Y., Caution: Electrostatic discharge at work, *IEEE Transactions Component Hybrids and Manufacturing Technology*, 5, (1982), 512–515.

Ning, T.H., Cook, P.W., Dennard, R.H., Osburn, C.M., Schuster, S.E., and Yu, H.W. 1 mm MOSFET VLSI technology Part IV: hot electrons design constraints. *IEEE Transactions on Electron Devices*, ED-26 (1979), 346–353.

Nippon Telegraph and Telephone Corporation (NTT), *Standard Reliability Table for Semiconductor Devices*, March (1985).

O'Connor, P.D.T., Reliability prediction: help or hoax? *Solid State Technology*, (1990), 59–61.

Pancholy, R.K. and Jhnoki, T., CMOS/SOS gate protection networks. *IEEE Transactions on Electronic Devices*, ED-25 (1978), 917–925.

Pecht, M., U.S. *Japanese Technology Evaluation Center Visit*, Private Communications, 1993.
Pecht, M. and Ko, W., A corrosion rate equation for microelectronic die metallization, *International. Journal for Hybrid Microelectronics*, 13 (1990), 41–52.
Pecht, M., Lall, P., and Dasgupta, A., A failure prediction model for wire bonds. *Proceedings of the 1989 International Symposium on Hybrids Microelectronics*, Baltimore, MD, (1989), 607–613.
Pecht, M., Lall, P., Hakim, E.B., The influence of temperature on integrated circuit failure mechanisms. *Quality and Reliability Engineering International Journal*, 8 (1992), 167–175.
Pecht, M. and Ramappan, V., Are components still the major problem: a review of electronic system and device field failure returns. *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, 15(6), December (1992), 1–5.
Phillips, W.E., *Microelectronic Ultrasonic Bonding*, G.G. Harman, ed., National Bureau of Standards (U.S.), Spec. Publ. 400–2, (1974), 80–86.
Ravi, K.V. and Philosky, E.M., Reliability improvement of wire bonds subjected to fatigue stresses. *10th Annual Proceedings IEEE Reliability Physics, Symposium*, Las Vegas, NV, (1972), 143–149.
Reif, F., *Fundamentals of Statistical and Thermal Physics*, McGraw-Hill, New York, 1965.
Ricco, B., Azbel, M. Ya, and Brodsky, M.H., Novel mechanism for tunneling and breakdown of thin SiO_2 films. *Physics Review Letters*, 51 (1983), 1795.
Rome Air Development Center, RAC (Reliability Analysis Center), *IC Quality Grades: Impact on System Reliability and Life Cycle Cost, SOAR-3*, (1985).
Runayan, W.R., *Silicon Semiconductor Technology*, McGraw-Hill Company, New York, (1965).
Schnable, G.L., Failure mechanisms in microelectronic devices. *Microelectronics and Reliability, Artech House Inc.*, 1 (1988), 25–87.
Schwarzenberger, A.P., Ross, C.A., Evetts, J.E., and Greer, A.L., Electromigration in presence of a temperature gradient: Experimental study and modeling. *Journal of Electronic Materials*, 17 (1988), 473–478.
Selberherr, S., *Analysis and Simulation of Semiconductor Devices*, New York: Springer-Verlag, (1984), 119.
Setliff, J.E., A review of commercial microcircuit qualification and reliability methodology, *Proceedings of the 1991 Advanced Microelectronics Technology, Qualification, Reliability and Logistics Workshop*, August 13–15, Seattle, WA, (1991), 325–335.
Shirley, G.C. and Hong, C.E., Optimal acceleration of cyclic THB tests for plastic packages devices. *29th International Reliability Physics Symposium*, Las Vegas, NV, (1991), 12–21.
Siemens Standard, SN29500, *Reliability and Quality Specification Failure Rates of Components*, (1986).
Smith, J.S., El overstress failure analysis in microcircuits. *16th Annual Proceedings International Reliability Physics Symposium*, San Diego, CA, (1978), 41–46.
Stojadinovic, N.D., Failure physics of integrated circuits: A review. *Microelectronics and Reliability*, 23 (1983), 609–707.
Swartz, G.A., Gate oxide integrity of NMOS transistor arrays. *IEEE Transactions on Electron Devices*, ED-33, (1986), 1826–1829.
Sze, S.M., *Physics of Semiconductor Devices*, 2nd edition, John Wiley & Sons, New York, (1981), 30–47.
Takeda, E., Nakagome, Y., Kume, H., Suzuki, N., and Asai, S., Comparison of characteristic of N-channel and P-channel MOSFETs for VLSI. *IEEE Transactions on Electron. Devices*, ED-30 (1983), 675–680.
Thomas, R.W. and Calabrese, D.W., Phenomenological observations on electromigration. *21st Proceedings of International Reliability Physics Symposium*, Phoenix, AZ, (1983), 1–9.
Villela, F. and Nowakowski, M. F., Investigation of fatigue problems in 1-mil diameter thermocompression and ultrasonic bonding of aluminum wire. NASA Technical Memorandum, NASA TM-X-64566 (1970).
Wendt, H., Cerva, H., Lehmann, V., and Pamler, W., Impact of copper contamination on the quality of silicon oxides. *Journal of Applied Physics*, 65 (1989), 2402.
Wigner, E.P., The transition state method, *Faraday Society (London) Transactions*, 34, (1938), 29.
Witzmann, S. and Giroux, Y., Mechanical integrity of the IC device package: A key factor in achieving failure free product performance, *Transactions of the First International High Temperature Electronics Conference*, Albuquerque, (June 1991), p. 137.
Wong, K.L., The bathtub does not hold water any more. *Quality and Reliability International*, 6 (1990a), 279–282.
Wong, K.L., What is wrong with the existing reliability prediction methods? *Quality and Reliability Engineering International*, 6, (1990b), 251–257.
Woods, M.H. and Rossenburg, S., EPROM reliability; Parts I and II: Electronics, *Microelectronics and Reliability*, 53 (1980), 133–141.
Wunsch, D.C. and Bell, R.R., Determination of threshold failure levels of semiconductor diodes and transistors due to pulse voltages. *IEEE Transactions on Electron Devices*, NS-15 (1968), 244–259.

4.23 DIRECT CONTACT HEAT TRANSFER

Harold R. Jacobs

Introduction

Direct contact heat transfer refers to the transport of thermal energy between two or more mass streams without separation by intervening walls. Generally, the interaction is between only two streams and there is generally a degree of mass transfer as well. In the above definition, the words mass streams are used; however, the term direct contact can also refer to batch processing such as in a stirred reactor. Nonetheless, the study of continuous processing is more likely. In process heat transfer, direct contactors are used in absorption refrigeration, distillation, electric power generation (as open feedwater heaters and in conjunction with dry cooling towers in the so-called Heller cycle), food processing, drying, emergency cooling of nuclear reactors, quenching of hot gases, and in wet cooling towers. Other process equipment includes condensers such as barometric condensers, and spray condensers. The process of liquid spray combustion (Law, 1984) in a variety of applications also is a direct contact heat transfer process through the fact that the fuel is injected as a spray of drops which must be preheated to the point of evaporation. In general, the vaporization process, normally required as the combustion takes place, fits the definition of direct contact heat transfer. Many environmental processes such as the formation of ice, snow, or rain utilize the same physics as in other direct contact processes. Thus, direct contact heat transfer is ubiquitous.

In process-type direct contact heat and mass transfer, we normally deal with only two mass streams, although three or more can occur. For example, in some pipeline problems and in the extraction of heavy oils in the petroleum industry, water, oil, and air are sometimes encountered. This is often the case in off-shore production.

In process heat transfer equipment, the mass streams can be cocurrent, countercurrent, or even cross flow. The streams may be composed of, for example, systems that are liquid–liquid, liquid–vapor, liquid–gas, liquid–solid, gas–solid, or even solid–solid. The streams can be composed of the same chemicals, as in the case of steam-water, or can be partially miscible or completely so. One or both streams can be dispersed (made up of defined masses, particles, drops, bubbles) or continuous.

As noted previously, heat transfer may be accompanied by mass transfer such as in the case of dryers, distillation columns, wet cooling towers, condensers, and in combustion systems. The possibilities are legion. Thus, in this section we must limit ourselves to a few fundamental processes. The reader is referred to the literature or specific review articles such as Sideman (1966), Sideman and Moalem-Maron (1982), Jacobs (1988a), articles by Jacobs (1988a,b), Taitel and Tamir (1969).

The fundamental processes to be included here are sensible heat transfer between parallel fluid streams in co-or counterflow; sensible heat transfer to a disperse flow of particles, drops, or bubbles; flow with condensation, and flow with evaporation with similar geometries to that for sensible heat transfer.

Heat Transfer between Continuous Parallel Streams

Falling Films

In cooling towers, wetted wall towers, packed beds, etc., a liquid is injected into the device in a number of ways, resulting in a thin film of liquid running over solid vertical or inclined surfaces in the form of a coherent thin film or in rivulets under the influence of gravity. Such flows are generally classified under the name of falling films. The solid substructure can take any variety of forms including a bed of spheres, raschig rings, berl saddles, or a structured array of metal chevrons, for

example. The packing can be in an organized array or random and can include particles of different sizes and be of a variety of materials, generally wettable by one of the streams.

The initial work on falling films can be traced to Nusselt (1916). Since that time, many studies have been performed. Hasson et al. (1964a,b) investigated a vertically falling film while neglecting the effects of surface tension, interfacial shear, and pressure gradients. He obtained a universal nondimensional curve for the developing film thickness. Murty and Sastri (1974) studied the problem of a film exiting a slot flowing down an inclined plane of defined height. The surrounding gas was assumed to be quiescent, and pressure variation and surface tension were assumed to be negligible. Assuming a smooth interface and continuity of velocity and shear at the interface, they determined the thickness of the film as a function of the length along the plate. Murty and Sarma (1976) presented an expression for film thickness for either laminar or turbulent flow down an inclined wall with a cocurrent gas stream and incorporated interfacial shear and gravity effects in the nondimensional film thickness. Experimental nondimensional velocity profiles were used to present a universal, nondimensional film thickness. Tekic et al. (1984) repeated the model of Murty and Sastri (1974) but incorporated normal and tangential stresses on the film interface and neglected the interfacial shear of the still air. Their results presented a family of curves for liquids with different physical properties, but underpredicted the film entrance length.

While cocurrent flows are interesting, they result in minimum heat transfer when compared to countercurrent flows. Countercurrent flows, however, can strip the film from the surface, depending on the properties of the fluid external to the film. Since the external fluids are typically a gas or a vapor, the velocities when stripping occurs are normally quite high. Nevertheless, this is a common problem in cooling towers and other devices, and in the case of cooling towers can lead to the need for excessive make-up water and a highly visible plume. Usually the falling film develops a wavy interface, followed by the formation of larger waves on its surface prior to entrainment of liquid in the gas stream. The waves can be initiated by either Tollmein–Schlichting or Kelvin–Helmholtz instabilities. Ostrach and Koestel (1965) discussed these and other instabilities associated with two-phase flows. The Tollmein–Schlichting waves are associated with transition from laminar to turbulent flows while the Kelvin–Helmholtz instabilities are interfacial phenomena resulting from the shear at the interface due to relative velocity differences. The latter are applicable to otherwise laminar flows.

Ishii and Grolmes (1975) discuss entrainment of liquid into the gas stream. For the case of vertical falling films in countercurrent gas flow, possible causes of entrainment are shearing droplets from the top of roll waves and the formation of large-amplitude bulges in the liquid. Unstable films are particularly troublesome in wetted wall columns (which are simply vertical pipes with a gas flowing up the center and an annular film flowing down the surface). Such devices have been reviewed by Bharathan (1988). Tien and Liu (1979) outline an overview of the theory and experiment. Subsequently, McQuillan et al. (1985) and Bankoff and Lee (1986) reviewed these apparatuses and established a data bank of experimental flooding points. Stefan and Mayinger (1990) studied such systems at high gas pressure.

Heat and mass transport to falling films were reviewed by Seban (1978) and others (Faghri, 1983; Faghri and Saban, 1985). He discussed heating at the wall, evaporation, absorption, and the effects of film thickness for laminar and turbulent flows, but omitted interfacial shear effects. Experimental work since Seban's review include that of Segev et al. (1981) and Lee and Bankoff (1983). Theoretical studies were conducted by Jacobs and Fletcher (1988) and Jacobs and Dolan (1994) of the counterflow falling film problem for laminar flows with heat transfer from the gas to a film flowing over an adiabatic wall. Thus, their models apply to packed beds and other devices where the supporting surfaces are only to provide for the formation of the film. (In fact, the packing continuously reforms the film in lengths sufficiently small so as to maximize the local heat transfer coefficient on an individual packing element.) The first paper used integral methods and was applicable for only very small and large values of the gas velocity. The asymptotic solutions indicated that for a given countercurrent velocity a minimum would occur in the heat transfer coefficient. The second paper utilized the boundary layer equations for the film and the Navier–Stokes equations for

the external gas stream, and was able to define the extent of the phenomenon. The nature of the flow is shown in Figures 4.23.1a and b. Comparison of the results are shown in Figure 4.23.2. It is clear that both analyses predict a region of minimum heat transfer. Figure 4.23.3 shows a comparison of the mean interfacial friction coefficient, and with Figure 4.23.2, indicates an overprediction of drag and an under-prediction of heat transfer by the integral model. The results indicate that a gas velocity of at least 10 times that of the mean film should be used, but that the velocity should be less than that necessary to cause entrainment. The heat transfer in the gas liquid film system is controlled by the gas-side heat transfer for the case of sensible heat transfer.

No solution currently exists for the wavy film countercurrent flow situation. Subject to the lack of experimental data, it is recommended that one assume that the surface waves should only influence the heat transfer internal to the film. For situations where the thermal penetration is less than the film thickness, the average Nusselt number for the internal resistance is

$$\mathrm{Nu} = \sqrt{\frac{6}{\pi}}\mathrm{Re}^{1/2}\mathrm{Pr}^{1/2} \tag{4.23.1}$$

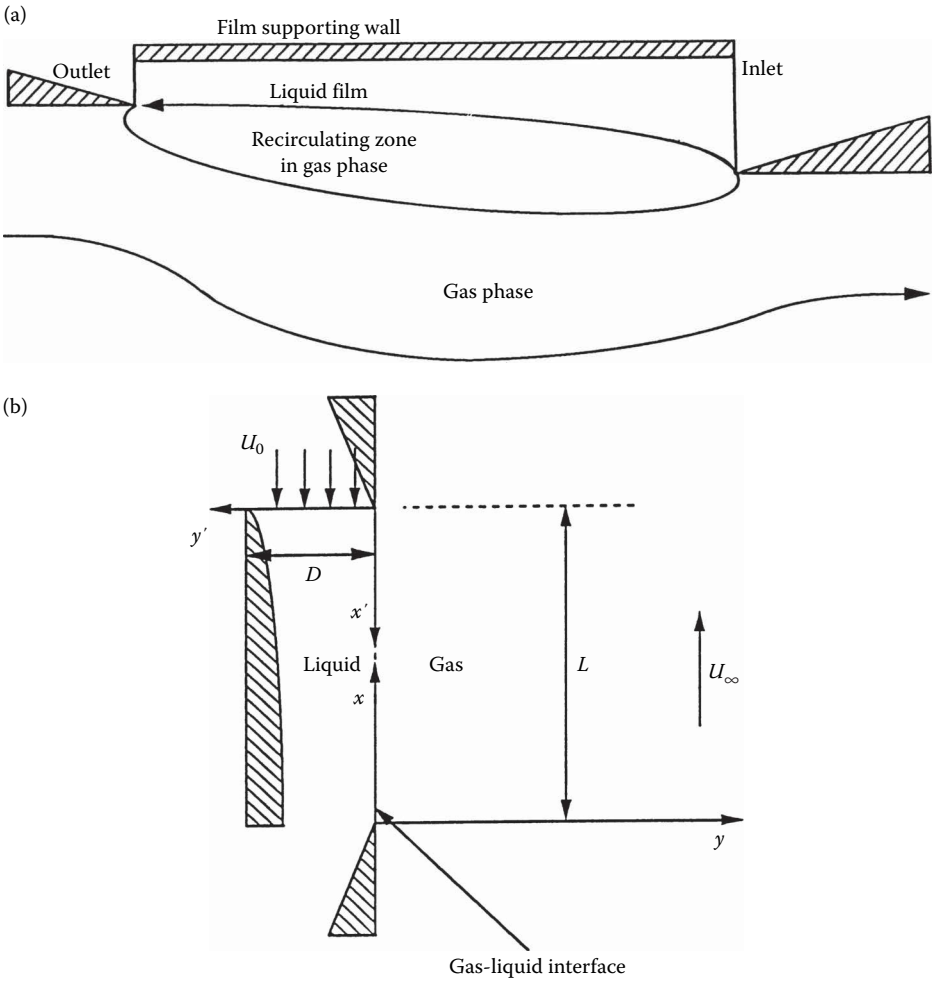


FIGURE 4.23.1 (a) Flow of gas over falling liquid film with decreasing thickness. (b) Geometry used in mathematical model for falling film and gas regions. (From Jacobs, H.R. and Dolan, S.M., *J. Energy, Heat Mass Transfer*, 16, 77, 1994. With permission.)

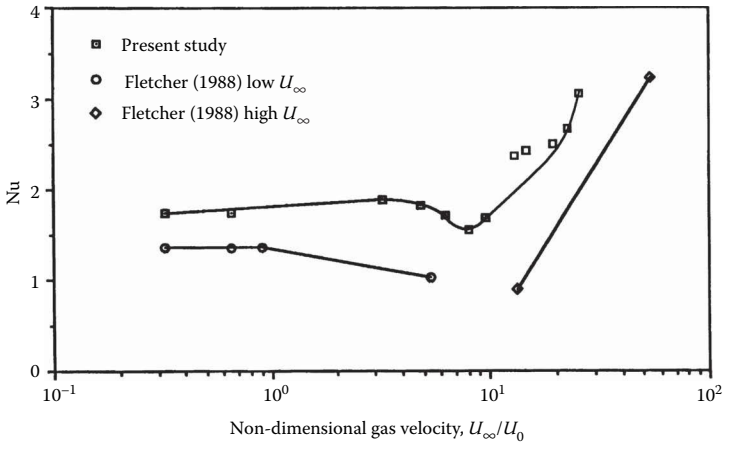


FIGURE 4.23.2 Relationship between nondimensional gas velocity and average Nusselt number, U_∞ = gas velocity, U_0 = liquid velocity at inlet. (From Jacobs, H.R. and Dolan, S.M., *J. Energy, Heat Mass Transfer*, 16, 77, 1994. With permission.)

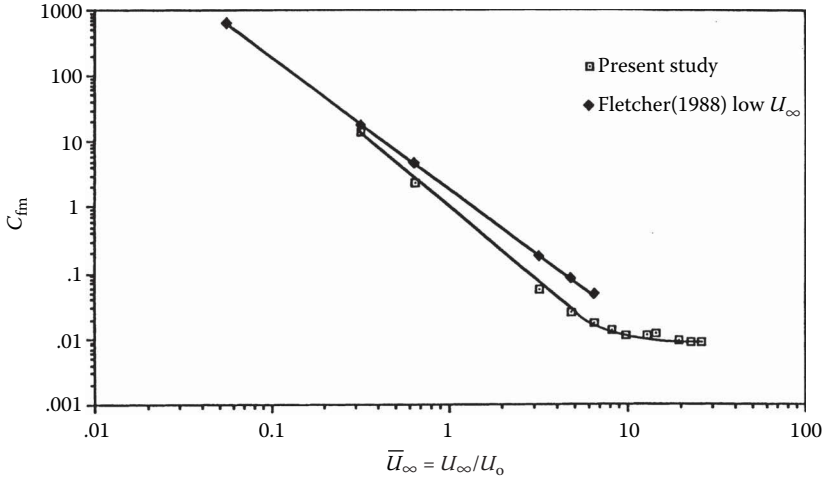


FIGURE 4.23.3 Effect of U_∞ changes on frictional drag.

where $\mathrm{Re} = uD/v$ is the film thickness and u is the mean film velocity. For a wavy film, the heat transfer coefficient will be increased by 15% (Levich, 1962) based on an assumed similarity to gas absorption. Similarly for turbulent films, mass transfer analogs should offer some assistance although there exist few controlled experiments (Levich, 1962).

Heat Transfer to Jets and Sheets

In many types of process equipment (Figure 4.23.4), the apparatus is simply a vessel containing a few partial trays which are solid or perforated. The trays have a lip so as to hold a quantity of liquid sufficient to provide the needed pressure head to form a series of jets or a sheet of fluid falling under gravity. Such devices are used in humidifiers and direct contact condensers or as mass transfer devices in dealing with gases or vapors (List, 1982).

In such equipment the heat transfer process is described as heat transfer to jets and sheets rather than films, as the fluid is not supported where the primary heat transfer takes place (on the jets and

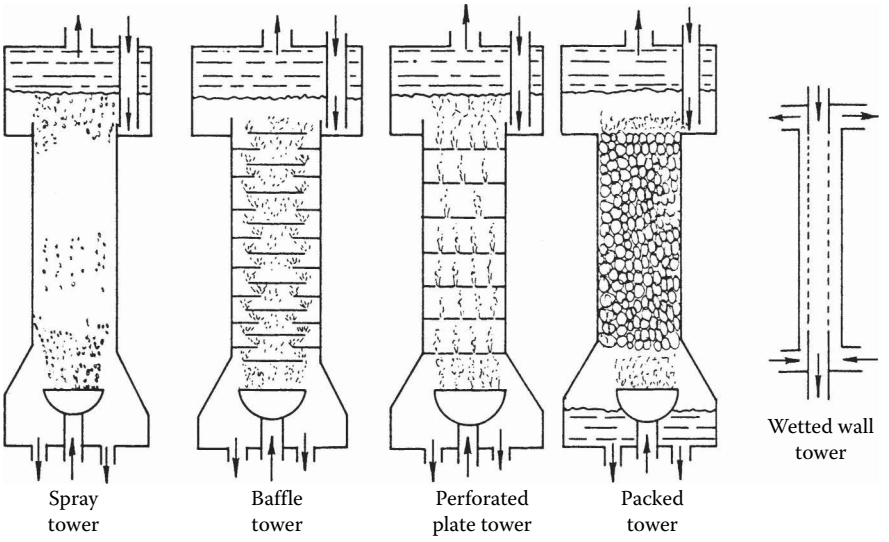


FIGURE 4.23.4 Schematic of various types of direct contact counterflow devices.

sheets). The series of trays act as mixers. The liquid is caught on a tray and leaves it at its mean temperature. The reforming of jets or sheets provides for a higher heat transfer coefficient than would occur for a stream that fell the entire length of the column and, thus, requires a shorter column to accomplish the required energy transfer. The major disadvantage is a greater pressure loss in the device than for an empty vessel (which is called a spray column).

Typically, the hydrodynamics of a jet or sheet can be assumed to be either a slug flow or fully developed. Thus, the analysis of the heat transfer via the energy equation follows that of a Graetz problem where only the axial component of the advection term and the conduction term across the jet or sheet remain. External to the jet or sheet the heat transfer coefficient must be determined. In most cases, values for flow over a solid surface can be used. The flow over a jet or sheet in most column designs is crossflow and, thus, the external heat transfer coefficient is independent of axial location. This allows one to adapt the Graetz-type solutions to solving for the mean axial temperature of the jet or sheet, as was done by Jacobs and Eden (1986). For liquid jets or sheets exposed to a gas or vapor, especially for sensible heat transfer, the external resistance to heat flow may be controlling, and it may be possible to assume negligible gradient across the jet or sheet. Otherwise, the fully developed internal thermal resistance can be used for axial lengths greater than the thermal developing region; however, to minimize the internal resistance requires the use of shorter jets or sheets and requires the distance between trays to be less than the thermal entrance length, as predicted by the Graetz solution. For turbulent flows in the liquid, the maximum heat transfer would neglect the internal resistance altogether. Where the external thermal resistance is not controlling and the stream is turbulent, it is advisable to use a Graetz-type solution with an effective thermal conductivity given by

$$k_{\in \mathrm{H}} = k + \frac{\rho c \in_{\mathrm{t}}}{\mathrm{Pr}_{\mathrm{l}}} \tag{4.23.2}$$

The above discussion applies to smooth, laminar surfaces. As with thin films, waves will eventually occur on the surface of these streams, leading to a better mixing in the jets or sheets. The wavy regime typically decreases the internal resistance by as much as 15%. Eventually, the jets and/or sheets will break down into drops or a spray in a manner similar to that discussed in Section 4.18. While it is possible to have a jet break down into a fine spray, if the jet Reynolds number is sufficiently

high, a freely falling jet will typically break down into drops that are twice the diameter of the jet or twice the thickness of a sheet. In many instances this does not provide a sufficient increase in surface area to result in higher heat transfer rates than would occur in a continuous sheet, although internal mixing may increase. Thus, an analysis for a continuous jet or sheet may be applicable even after breakup.

Sensible Heat Transfer to Dispersed Media: Drops, Particles, Bubbles

Sensible heat transfer to disperse media from a continuous phase is dependent upon the proximity of other particles, drops, or bubbles and to their own geometry, as well as their velocity through the continuous surrounding media and its thermal properties. For solid particles, the heat transfer is generally described in terms of an equivalent sphere. An equivalent diameter is obtained from

$$D_{\text{equiv}} = \frac{6V}{A_s} \tag{4.23.3}$$

For drops and bubbles, the problem is more difficult than for solid particles, as they are easily distorted, subject to both breakup and coalescence, and may have significant internal circulation depending on how they are formed (Rose and Kintner, 1966; Sideman and Isenberg, 1967; Sideman and Taitel, 1964; Sirignano, 1983; Sundararajan and Ayyaswamy, 1984, 1985a,b). Close proximity to other drops or bubbles, as associated with volume fraction or holdup equaling 5% or greater, will make drops that would otherwise distort tend to remain spherical. Further, the presence of impurities tends to stabilize the interfaces of bubbles with the surrounding media no matter whether the impurities are in the dispersed or continuous phase. Thus, models of drops within swarms, based on transport in single bubbles typically overpredict heat and mass transfer since they overpredict drop or bubble distortion and internal circulation. Drop dynamics based on drag for single spheres will be in error, as the drag will vary for drops in a swarm due to differences in the surrounding flow. In the midst of the array the spheres have a greater drag, but at the front the drag can be less than for a single sphere and the drag at the rear can be higher than that within the array for staggered or random arrays. For inline arrays, the drag can be greatest at the front of a column, leading to a bunching up of particles.

For holdups less than 5%, assuming rigid spheres, the expression of Ahmad and Yovanovich (1994) for external to the sphere heat transfer is recommended:

$$\mathrm{Nu_D} = 2.0 + \frac{0.775\mathrm{Re_D^{1/2}Pr^{1/2}}\Big/\sqrt{2\gamma+1}}{\left[1+\left(\dfrac{1}{(2\gamma+1)^3\,\mathrm{P_r}}\right)^{1/6}\right]} \tag{4.23.4}$$

where:

$$\gamma = \frac{1.0}{\mathrm{Re_D^{1/4}}} \quad \text{for} \quad \gamma < 1 \tag{4.23.5}$$

and $\gamma = 1.0$ for $\gamma > 1$.

For holdups greater than 5%, Wilson and Jacobs (1993) recommended for isothermal spheres in the midst of a swarm:

$$\mathrm{Nu_D} = \left\{1 - 0.186\left(\phi - 0.42\right)\right\}\left(0.877\mathrm{Re_D^{1/2}} + 0.152\mathrm{Re_D^{2/3}}\right)\mathrm{Pr}^{1/3} \tag{4.23.6}$$

where the velocity is the superficial velocity through the bed. This expression compares well with the correlation of Whitaker (1972) for a densely packed bed as well as for single spheres.

Similarly, Evans (1994) studied the heat transfer to particles at the leading and trailing edge of a hexagonal array of spheres. Using a partial spherical shell with slip at the surface to define the effect of the surrounding particles, he showed the effect mentioned earlier for random arrays. Sirignano (1993) has studied inline arrays for up to three particles deep and showed the effects mentioned earlier for such particles. While inline effects may apply to a single column of bubbles, the hydrodynamic instability of such a flow makes the staggered array of Evans (1994) more practical in attempting to model the flows. Analogy between heat and mass transfer in packed beds has been explored by Roberson and Jacobs (1990) and Thomson and Jacobs (1984).

An experimental model for the behavior of drops or particles in swarms was presented by Wagner and Stichlmair (1997). They provide the following expressions for swarm velocity relative to the continuous phase as a function of the Archimedes number, Ar, void fraction, and so-called fluid number, K_f:

$$\frac{\mathrm{Re}_{ss}}{K_f^{0.086}} = \frac{4.55\,\mathrm{Ar}\left(1-\epsilon_d\right)^{4.65}\Big/K_f^{0.1}}{\left[15.6+\mathrm{Ar}\left(1-\epsilon_d\right)^{4.65}\Big/K_f^{0.1}\right]^{1.5}} \tag{4.23.7}$$

where $\mathrm{Re}_{ss} = \dfrac{W_{ss} D_p \rho_c}{\mu_c}$, $\quad K_f = \sigma^3 \rho_c^3 \Big/ \left(\mu_c^4 g \Delta\rho\right)$, $\quad \mathrm{Ar} = d_p^3 \rho_c \Delta\rho g / \mu_c^2$

and W_{ss} is the superficial velocity of the multiphase system.

Figure 4.23.5 presents a comparison of experimental data with the correlation of Equation 4.23.7, and Figure 4.23.6 shows the flooding limit obtained.

The internal to the drop or particle heat transfer is based on the one-dimensional conduction problem in a sphere. This, of course, must be modified for internal circulation and geometrical distortion. Grace (1983) provides a mapping of drop distortion for single drops of a pure fluid in a

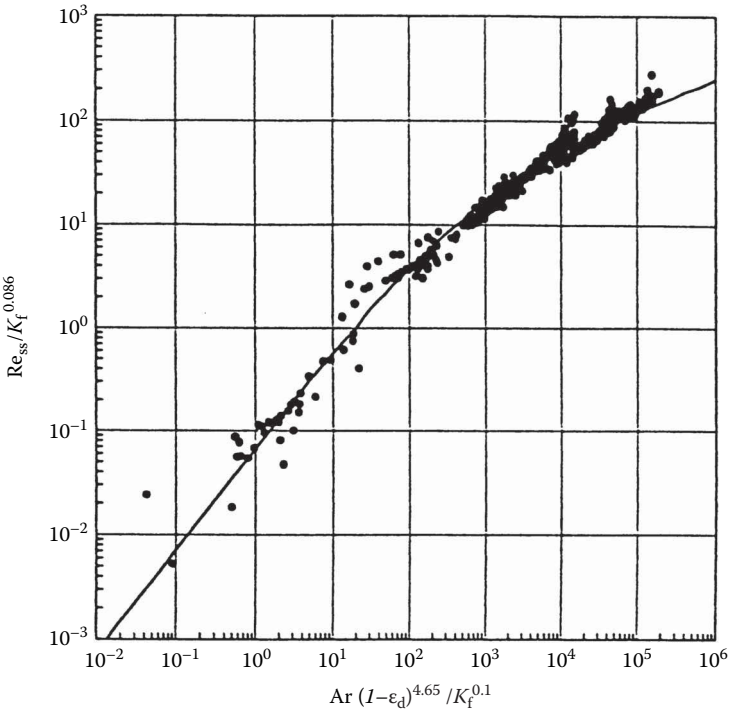


FIGURE 4.23.5 Comparison of Equation 4.23.7 with experimental data.

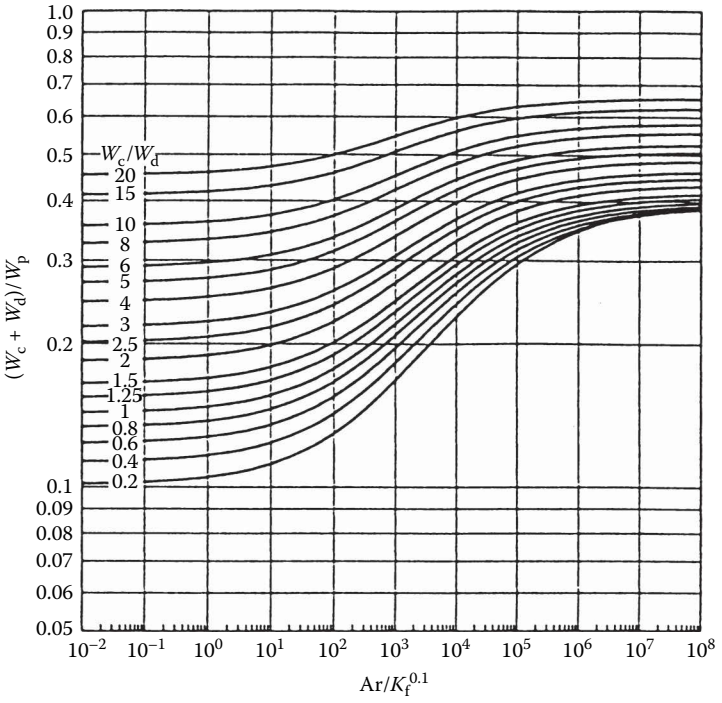


FIGURE 4.23.6 Flooding rates.

second pure immiscible fluid, as shown in Figure 4.23.7. Impurities, however, tend to immobilize the interface and the proximity of other drops will reduce any recirculation behind the drops and return them to sphericity. Jacobs and Eden (1986) provide for a correction to the molecular diffusivity which can account for both internal circulation and distortion in drops in a swarm or dense array as shown in Figure 4.23.8. For a single column of bubbles, Calderbank and Korchinski (1956) also provide an effective thermal diffusivity based on studies of bromobenzene in glycerol. Jacobs and Eden's studies were for isobutane in a geothermal brine.

For rapidly internally circulating drops, a number of investigators have used the Hadamard (1911) model for flow inside a sphere. Of these models, Sideman (1966) recommends

$$\mathrm{Nu}_{\mathrm{di}} = 2 + 1.13\left(\mathrm{Pe}K_{\mathrm{v}}\right)^{0.5} \tag{4.23.8}$$

where

$$K_{\mathrm{v}} = 1 - \frac{1.45\left(2 + 3\mu_{\mathrm{d}}/\mu_{\mathrm{c}}\right)}{1 + \dfrac{\rho_{\mathrm{d}}\mu_{\mathrm{d}}^{1/2}}{\rho_{\mathrm{c}}\mu_{\mathrm{c}}^{1/2}}}\mathrm{Re}_{\mathrm{cD}}^{1/2} \tag{4.23.9}$$

and subscripts c and d refer to the continuous and drop fluid properties, respectively.

The drop temperature is predicted from

$$\frac{\rho_{\mathrm{d}} C_{\mathrm{pd}} D}{6}\frac{\mathrm{d}T_{\mathrm{d}}}{\mathrm{d}t} = U_0\left(T_{\mathrm{d}} - T_{\mathrm{c}}\right) \tag{4.23.10}$$

where U_0 is the overall surface-based heat transfer coefficient. When the drop is not strongly circulating the spherical conduction problem must be solved. The temperature distribution in

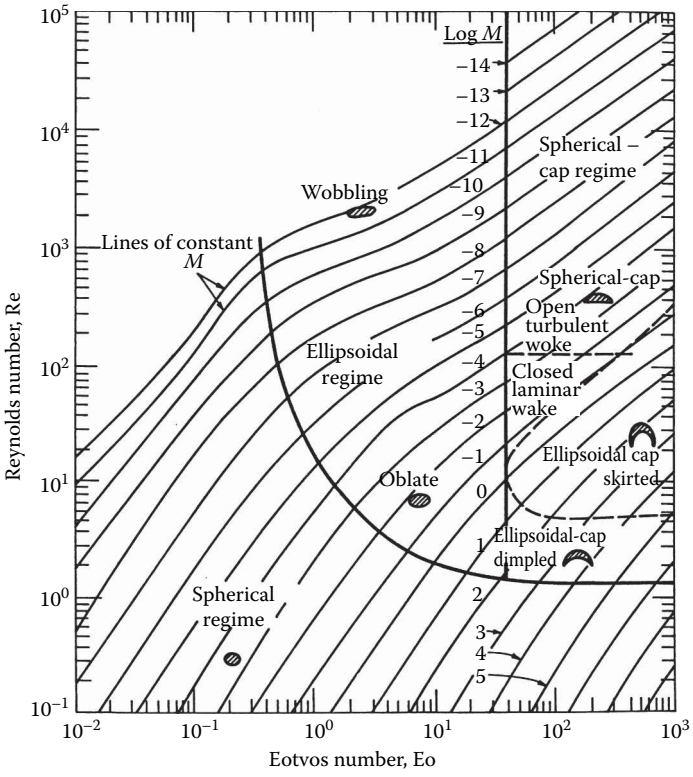


FIGURE 4.23.7 Distortion map as a function of Reynolds and Eötvos numbers. (From Grace, J.R., *Handbook of Fluids in Motion*, Butterworths Group, Ann Arbor, MI, 1983. With permission.)

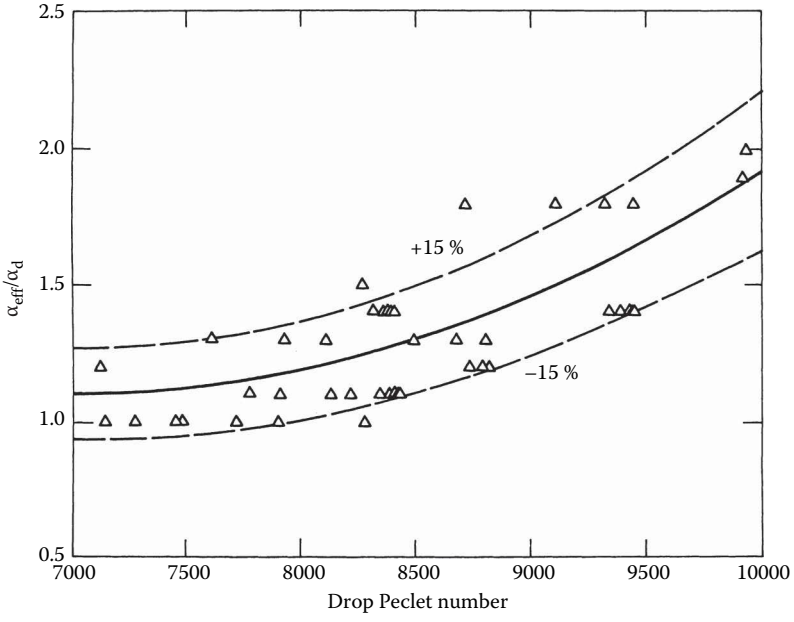


FIGURE 4.23.8 Ratio of effective thermal diffusivity as a function of drop Peclet number. (From Jacobs, H.R. and Eden, T.J., *Proceedings of 8th International Heat Conference*, Vol. 6, 3013, 1986. With permission.)

a sphere subject to a constant heat transfer coefficient and fixed continuous phase temperature is given by

$$\frac{T-T_\infty}{T_0-T_\infty}=2\sum_{n=1}^{\infty}\frac{\mathrm{Bi}\sin\mu_n e^{-\mu_n F_0}}{\left(\mu_n-\sin\mu_n\cos\mu_n\right)}\frac{\sin\left(\mu_n r/R\right)}{\mu_n r/R} \tag{4.23.11}$$

where T_0 is the initial sphere temperature, Bi is the Biot modulus, F_0 is the Fourier modulus, and $\mu_n=\lambda_n R$. Here μ_n are the zeros of $\mu_n \cos \mu_n=(1–\mathrm{Bi}) \sin \mu_n$. Solution of Equation 4.23.11 can be obtained from the so-called Heisler charts, or calculated.

In the time interval 0 to t, the ratio between the total heat transfer q from or to the sphere, and its initial internal energy $q_i=4/3R\ c(T_0-T_\infty)$ is

$$\frac{q}{q_i}=6\sum_{n=1}^{\infty}\frac{\sin\mu_n-\mu_n\cos\mu_n}{\mu_n^3\left(\mu_n-\mu_n\cos\mu_n\right)}\left(1-e^{-\mu_n^2F_0}\right) \tag{4.23.12}$$

Use of Equation 4.23.11 or 4.23.12 with values of α_{eff} and an appropriate heat transfer coefficient allows for the prediction of sensible heat transfer to individual drops or particles for low holdup. Equation 4.23.11 can be used as the Kernel in a Duhamel's superposition integral to solve for the temperature of a drop moving through a continuous phase, the temperature of which is variable with the trajectory of the drop or particle.

Direct Contact Heat Transfer with Change of Phase

Evaporation and Boiling

In Section 3 the topics of phase change from a vapor to a liquid and the reverse are introduced. These topics apply to direct contact heat transfer as well; however, the effects of solid surfaces do not come into play. The interface where the heat transfer, and thus the mass transfer to a different phase, takes place is between two independent mass streams.

At the beginning of Section 3.4.1, Carey introduces the concept of metastable liquids and the role of surface tension. Consider a gas bubble in a liquid. Carey shows that if the gas is a saturated vapor and the liquid is the same fluid, that the liquid must be superheated or in a metastable state if the pressure balance, including surface tension effects, is to be in equilibrium.

For pure liquids, where there exists a continuous phase with temperature sufficiently greater than the saturation temperature of a second fluid immiscible with the first, it is possible to demonstrate the existence of metastable states. If several subcooled drops of a fluid are injected into a second, the drops will have a pressure greater than the surroundings. The drops will gradually heat and will rise to a temperature equal to or nearly equal to the continuous phase. If the temperature is below the Spinodal limit (maximum superheat limit for a liquid) of the drops, nothing may happen. For example if the drops are pentane and the fluid is pure water, the drops may rise to the surface after a period of time and then evaporate. If the water temperature is near the Spinodal lmit, the drops will rise, but before reaching the surface may instantaneously vaporize in an explosive way. The presence of impurities in either phase will cause the drops to vaporize and form two-phase bubbles below the surface. The impurities act as nucleation sites for the phase change just as impurities in a metal or surface flaws serve as nucleation sites for boiling from solid surfaces. Some early experiments to illustrate boiling nucleation were with mercury films on a surface with an oil as the vaporizing medium. They were used to illustrate the need for nucleation sites. Thus, although the thermodynamics of direct contact boiling is related to conventional boiling, the effect of the interfaces between liquids makes them different and metastable states perhaps play a more important role. In most industrial situations, impurites are present and alter the situation for direct contact boiling.

Four basic problems are encountered in direct contact boiling/evaporation. These are (1) boiling at the interface between two parallel streams, (2) evaporation of drops of a spray from an immiscible

liquid, (3) evaporation of drops within a liquid or gas/vapor mixture, and (4) the evaporation of the continuous phase around the dispersed phase. The first problem is related directly to film boiling and follows its analyses. The second problem is of interest due to a number of things. First, if the interfacial properties are such that the liquid drops wet the liquid below them, there will exist a metastable state with the liquid in the drops having their phase change only from their upper surfaces. Such a situation was addressed by Kodres et al. (1980, 1981) who showed photographically the difference between this situation and that of Liedenfrost boiling, where the drop is held above the surface by the vaporizing film. (See *Numerical Heat Transfer*, 3:21–34, 1998.) Development of a model for the evaporation rate must be accomplished numerically and depends on the fluid properties (Prakash and Sirignano, 1980; Rangel and Sirignano, 1987).

The evaporation of a drop in a gas stream is dependent upon the nature of the gas phase. For the evaporation of a drop into a gas made up primarily of different substances; for example, a fuel drop in air combustion products requires the solution of the mass diffusion problem coupled with the heat transfer. The same is true of water drops in an air stream as related to the design of cooling towers discussed in Section 4.13. However, the evaporation of drops into a superheated stream of its own vapor is considerably easier to analyze, depending on the drop size. Carey and Hawks (1995) report for the evaporation of large drops in their own vapor

$$\mathrm{Nu}_\mathrm{D} = 2\frac{\ln\left(\mathrm{Ja}_\mathrm{v}-1\right)}{\mathrm{Ja}_\mathrm{v}} \tag{4.23.13}$$

where $J_{\mathrm{av}} = e_{\mathrm{pv}}\left(T_\infty - T_\mathrm{R}\right)/h_{\mathrm{fg}}$ and for microdrops

$$\mathrm{Nu}_\mathrm{D} = 2\frac{\ln\left(\mathrm{Ja}_\mathrm{v}-1\right)}{\mathrm{Ja}_\mathrm{v}}\left[1+\frac{0.4\,\mathrm{Ja}_\mathrm{v}}{1+10\,\mathrm{Ja}_\mathrm{v}^2}\left(\frac{K_{\mathrm{np}}}{0.001+K_{\mathrm{nD}}}\right)\cdot\sigma_\mathrm{T}\left(1+\gamma\right)\left(9\gamma+5\right)K_{\mathrm{nD}}+\sigma_\mathrm{T}\left(1+\gamma\right)\right] \tag{4.23.14}$$

where K_n is the mean free path of a molecule divided by the drop diameter, γ is the ratio of specific heat, and ϕ_r is the accommodation coefficient. The reader is referred to Sanders and Dwyer (1987) for further review of individual droplet processes. They present a comprehensive numerical model of single-droplet dynamics with and without combustion. Their studies include mass transfer with accelerating and decelerating drops. They solve the complete Navier–Stokes equations, energy equations, mass conservation equations, and equations of state.

The last evaporation/boiling model deals with droplets vaporizing in an immiscible liquid. Such a system was investigated by Coban and Boehm (1986) and by numerous others for application to desalination, geothermal power production, and production of electricity using solar ponds, etc. A volatile fluid is injected into a hot immiscible continuous stream. Unlike evaporation within a gas stream, here the vapor stays with the drop. The problem has been studied by Klipstein (1963), Simpson et al. (1974), Sideman (1966), and Raina and co-workers (1982, 1984a,b, 1985), and Battya et al. (1983). The models are all single-drop/bubble studies and are assumed to strongly depend on the external heat transfer and surface tension effects. In most of the systems studied, a drop of hydrocarbon was placed in water, the temperature of which was well above the boiling point of the hydrocarbon. For such systems, the surface of the two-phase bubble is wetted by the hydrocarbon liquid. This suggests that the hydrocarbon must reach a degree of superheat before evaporation occurs, but not the explosive condition close to the Spinodal line. Various models assume that the liquid settles to the bottom of the drop, as in the early models of Sideman, or to totally wet the walls with a uniform film. The former models do not reproduce well the experimental results while the latter model, first proposed by Simpson, does. It results in

$$h_0 = \frac{2.57\left(\dfrac{D(t)}{D_0}\right)^{1/6}}{1+0.206\left(\dfrac{D_t}{D_0}\right)^{5/12}}\frac{K_\mathrm{w}}{\mathrm{m^2K}} \tag{4.23.15}$$

This equation predicts an increasing heat transfer in the early stages of evaporation and a decrease in the latter stages.

The Sideman (1966) model depended only on the external heat transfer coefficient and fraction of the bubble surface wetted by the remaining liquid. It predicts the heat transfer decreases with time as it ignores heat transfer through the vapor. Coban and Boehm (1986) account for heat transfer through the vapor as well as directly to the liquid, resulting in

$$\mathrm{Nu}_{\mathrm{lD}} = 2\mathrm{Pe}_{\mathrm{c}}^{1/3}\left\{0.466\left(\pi - B + \sin\frac{\mathrm{d}B}{\mathrm{d}t}\right)^{2/3} + \frac{5K_{\mathrm{dv}}}{R}(1-0.466)\left(\pi - B + \sin\frac{\mathrm{d}B}{\mathrm{d}t}\right)^{2/3}\right\} \quad (4.23.16)$$

where c indicates the continuous phase, d the drop, and l and v are properties of the liquid and vapor, respectively. The above model assumes no sloshing of the liquid and B is the segment angle defining the surface of the liquid in the bottom of the drop. The expression is for the overall heat transfer.

The last evaporation phenomenon addressed is for a hot liquid drop or particle in a subcooled liquid with subsequent vaporization of the coolant. Kazimi et al. (1974) dealt with this as a transient film boiling problem but neglected the effects of both gravity and continuous phase motion. Dhir and Purohit (1977) solved the problem of subcooled film boiling from spheres including buoyancy effects. They predict

$$\mathrm{Nu}_{\mathrm{Dv}} = \mathrm{Nu}_0 + \mathrm{Nu}_{\mathrm{nc}}\frac{\mathrm{PrSc}}{\mathrm{Pr_l Sh}\mu} + \frac{\mathrm{Pr_v}\sigma}{\mathrm{Sh}}\frac{\left(T_{\mathrm{w}}^4 - T_{\mathrm{SAT}}^4\right)D}{h_{\mathrm{fg}}\mu_{\mathrm{v}}} \quad (4.23.17)$$

where

$$\mathrm{Nu}_0 = 0.8\left(\frac{g\rho_{\mathrm{v}}\left(\rho_2 - \rho_{\mathrm{v}}\right)h_{\mathrm{fg}}D^3}{M_{\mathrm{v}}k_{\mathrm{v}}\left(T_{\mathrm{w}} - T_{\infty}\right)}\right)^{1/4} \quad (4.23.18)$$

and

$$\mathrm{Nu}_{\mathrm{nc}} = 0.9\left(\frac{g\rho_{\mathrm{L}}c_{\mathrm{pL}}B_0\left(T_{\mathrm{SAT}} - T_{\infty}\right)D^3}{\mu_{\mathrm{L}}K_{\mathrm{L}}}\right)^{1/2} \quad (4.23.19)$$

for the average Nusselt number for a sphere in a quiescent pool.

For forced flow over a sphere or a falling sphere

$$\mathrm{Nu} = \mathrm{Nu}_0 = 0.8\mathrm{Re}_{\mathrm{cD}}^{1/2}\left[1 + \frac{\mathrm{ScPr}v}{\mathrm{ShPr_L}\mu_{\mathrm{c}}}\right] \quad (4.23.20)$$

where $\mathrm{Sc} = \dfrac{C_{\mathrm{pL}}\left(T_{\mathrm{SAT}} - T_{\infty}\right)}{h_{\mathrm{fg}}}$ and $\mathrm{Sh} = \dfrac{C_{\mathrm{pr}}\left(T_{\mathrm{w}} - T_{\infty}\right)}{h_{\mathrm{fg}}}$.

For boiling in cryogenic fluids, Hendricks and Baumeister (1969) investigated boiling of nitrogen and helium from a sphere. They show a dependence on the Bond number, $\mathrm{Bo} = (P_{\mathrm{L}} - P_{\mathrm{r}})g\,R^2/(\sigma gc)$, where σ is the surface tension. The results are shown in Figure 4.23.9, where Nu′ and Nu − 2 and $Rd' = \dfrac{h_{\mathrm{fg}} + 0.5C_{\mathrm{p}}\Delta T}{C_{\mathrm{p}}\Delta T}$.

Direct Contact Condensation

Direct contact condensation occurs in many industrial processes including reflux columns and in separators as well as in open feedwater heaters, etc. Jacobs (1988a) provides a significant discussion in his review article on *Direct Contact Heat Transfer for Process Technologies*. The first industrial

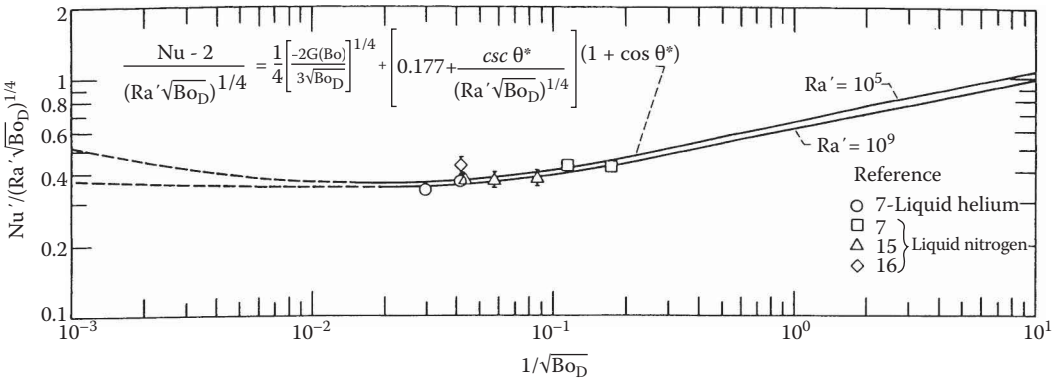


FIGURE 4.23.9 Experimental and predicted results for film boiling off submerged spheres. (From Hendricks, R.C. and Baumeister, K.J., Film boiling from submerged spheres, NASA, Washington, DC, 1969.)

use of direct contact condensers was in steam engines and dates back prior to the turn of the century. The first design article was written by Hausbrand and was published in English in 1933. In a review of the literature, Jacobs and Fannar (1977) categorized direct contact condensers as being "drop type," "jet and sheet type," and "film type."

Drop type refers to condensation on a spray of coolant that is injected into a chamber filled with vapor or a gas-vapor mixer with the intent to condense the vapor. The coolant can be the same fluid as the vapor or could be immiscible with the condensate. Kutateladze (1952) was the first to recommend the drops be assumed spherical and that the heat transfer is dominated by the transient conduction within the drops when dealing with a pure vapor. This assumption is reasonable in most situations as injectors used in direct contact condensers produce drops in the size range of 0.25–2.5 mm in diameter. Such drops fall at sufficiently slow velocity that they remain spherical with little or no internal circulation. The Kutateladze model neglects the added mass. While such an analysis is appropriate for a Jakob number tending to infinity or time to zero, it overpredicts the heat transfer. Jacobs and Cook (1978) developed a theoretical model that accounted for the added mass and predicts the experiments of Ford and Lekic (1973) exactly.

The model development shows that the final to initial drop size is given by

$$R_f / R_1 = \left(1 + \frac{1}{\mathrm{Ja}}\right)^{1/3} \tag{4.23.21}$$

where

$$\mathrm{Ja} = \frac{\rho_L h_{fg}}{\rho_d C_{pd} \left(T_{SAT} - T_0\right)} \tag{4.23.22}$$

in which the subscripts L and d refer to the condensate liquid and drop, respectively. Thus, one can deal with a drop being of a different fluid than the condensate. Their solution solves the problem of conduction in a sphere subject to the boundary condition

$$q'' = \frac{k_c \left(T_{SAT} - T_{RI}\right)}{R(t) - \mathrm{Ri}} \tag{4.23.23}$$

The nondimensional radius at time t is given as

$$\frac{T_{\text{mean}}(t)}{T_{\text{SAT}}-T_i}=1+\frac{6}{\pi}\sum_{n=1}^{\infty}\frac{1}{n^2}\exp\left(-n^2\pi^2\,\frac{4at}{D_0^2}\right) \tag{2.23.24a}$$

This is a solution of a spherical Stefan-type problem assuming that the condensate film is thin. It is appropriate for large values of Ja. In the limit as Ja tends to infinity

$$\frac{R(t)}{R_1}=1+\frac{3}{\text{Ja}}\int_0^{R_i}\left(\frac{r}{R_1}\right)^2\frac{\left(T_{(r,t)}-T_i\right)}{\left(T_{\text{SAT}}-T_i\right)}\frac{\mathrm{d}r}{\text{Ri}} \tag{4.23.24b}$$

For conditions where there are noncondensibles present, it is necessary to solve the flow field external to the drops. Such studies have been conducted by Chung et al. (1981a,b; 1984a,b) and by Huang and Ayyaswamy (1987a,b) who also included the effects of internal circulation. They found the latter to have negligible effects, except at extremely high velocities or low interfacial tension. Celata et al. (1990) attempted to form uniform drops by application of a high-frequency acoustic field. At Re as low as 100, but for Pe of 1000 to 5000, they found heat transfer coefficients higher than predicted for infinite Pe. The best fit to their data was obtained using Equation 4.23.24; however, even it underpredicted the early temperature rise, indicating even stronger mixing than predicted by the Hadamard model for flow inside a sphere.

The influence of noncondensibles on condensation rates is not unlike that predicted for boundary layer type problems, with as little as 5% noncondensibles reducing the heat transfer by as much as 50% to 100%. More studies of this influence are needed.

In addition to accurately predicting the performance of a spray type (drop type condenser), it is required to have data on drop size distribution for a given spray nozzle. Brown (1951) and Isachenko and Kushnyrev (1974) present some work, but more is needed. For example, in order to design a barometric condenser the vapor velocity must not exceed the terminal velocity of the smallest drop; yet to fully utilize the cooling capacity of the spray, the size of the largest drops must be known. From a practical view, this means that the diameter of the condenser for a simple spray column is set by the smallest drops and its height by the largest.

Collapsing Bubble Condensation

Direct contact condensation can be accomplished by injecting a jet of bubbles into a continuous liquid coolant. Depending on the jet velocity, the jet can break down quickly or remain coherent for some time. If the jet is injected at close to sonic velocities, the latter is true. Of course, the vapor can bubble through in a series of bubble trains as well. Thus, single vapor bubbles as well as jets have been studied. Jet flow is impacted by its orientation with gravity, with vertical jets foreshortened or lengthened depending on whether gravity is opposing or favorable. Horizontal jets are typically curved due to buoyancy effects. The first definitve study was by Weimer et al. (1973). Chen and Faeth (1982, 1983) and Sun (1985) used sophisticated two-fluid numerical models to analyze jet collapse, yet they do not adequately predict the heat transfer between the phases (Eden, 1995). Thus, dependence on experimental studies is still required.

The study of bubble collapse is not only of interest as it applies to condensers, but also for other applications. Sideman and Moalem-Maron (1982) reported a wide range of experimental data. Floreschuetz and Chao (1965) modeled a drop collapsing in its own vapor. Jacobs et al. (1978) were the first to account for the internal resistance of the condensate in the collapse of a saturated vapor bubble in an immiscible liquid. They show that the collapse rate for a pure vapor rising at a constant velocity is

$$\frac{D(t)}{D_0}=\left[1-0.901\left(\text{Re}_L\text{P}_r\right)^{1/2}\phi\,\text{Ja}^*\,\frac{\rho_f}{\rho_v}F_{0L}\right] \tag{4.23.25}$$

where

$$\phi = \sqrt{\left(C_1^2 + C_1\right)} - C_1 \tag{4.23.26}$$

and

$$C_1 = 0.75 \frac{\left(\rho k C_p\right)_f}{\left(\rho k C_p\right)_1} \tag{4.23.27}$$

Jacobs and Major (1982) dealt with bubbles containing a mixture of vapor and a noncondensible gas. They showed that for small bubbles less than or equal to 3.5 mm in radius, the heat transfer was controlled by diffusion, but larger bubbles have good internal circulation due to distortion and oscillation and their analysis can be accomplished by assuming perfect mixing during the collapse.

The effect of adjacent bubbles in a collapsing train was dealt with by Moalem-Maron et al. (1972). Golafshani (1983) improved on the model, using the more sophisticated model of Jacobs and Major (1982). For pure vapor bubbles they found little influence based on the models, indicating that the external thermal resistance dominated the heat transfer.

Despite the relative acceptance of the theories, Johnson et al. (1983) and Sudhoff (1982) found that experiments with bubble trains are plagued by a so-called persistent bubble. Johnson et al. carried out experiments in a very deep heat exchanger. The visual observations indicated that the persistent bubble finally disappears and they conclude that this is due to absorption of the non-condensible gases present by the surrounding liquid. So far, no other investigators have offered hypotheses to explain this phenomenon; however, if the arguments of Johnson et al. are valid, it may be possible that a heat exchanger could be designed with a depth such that condensation may be accomplished and noncondensible gases readily separated.

Condensation on Jets and Sheets

Barometric condensers are designed in the form of spray columns, baffle tray columns, or sieve tray columns. When designed as spray columns, a nozzle forms a coarse spray as discussed in the preceding section. When designed as a baffle tray column, the coolant flows into a tray, and then over its lip in a sheet to the tray below, and subsequently to the next tray and so forth until it and the condensate form a pool at the bottom of the condenser.

In a sieve tray column, the coolant leaves the tray through perforations as a series of jets. Both sheets and jets can break down into lesser streams and/or drops as discussed by Bogy (1979). The jets are laminar but instabilities can make them wavy or even turbulent prior to breaking into drops. If the jets or sheets remain laminar or just wavy, they can be analyzed as a Graetz-type problem similar to the analysis for sensible heat transfer. Hasson et al. (1964a,b) applied such analyses to condensing steam on sheets and fan jets with reasonable success. Jacobs and Nadig (1987) solved the problems for an axisymmetric jet or a sheet including both the resistance of the added condensate mass and the effects of noncondensibles for the vapor and coolant being the same or immiscible fluids. They presented curves of coolant capacity utilization rather than heat transfer coefficients for design purposes. Results for steam-water are presented in Figures 4.23.10a and b. Knowing the noncondensible concentration, and the degree of utilization, the Graetz number and, thus, the length of the sheet can be determined. Extension of this work to wavy or turbulent flow is desirable; however, it is likely that waves will not increase the heat transfer by more than 15% based on the sensible heating information reported earlier for jets and falling films.

When as little as 2% noncondensibles are present, the heat transfer is controlled by the external diffusion of vapor to the interface, thus turbulence in the liquid will also have small effects.

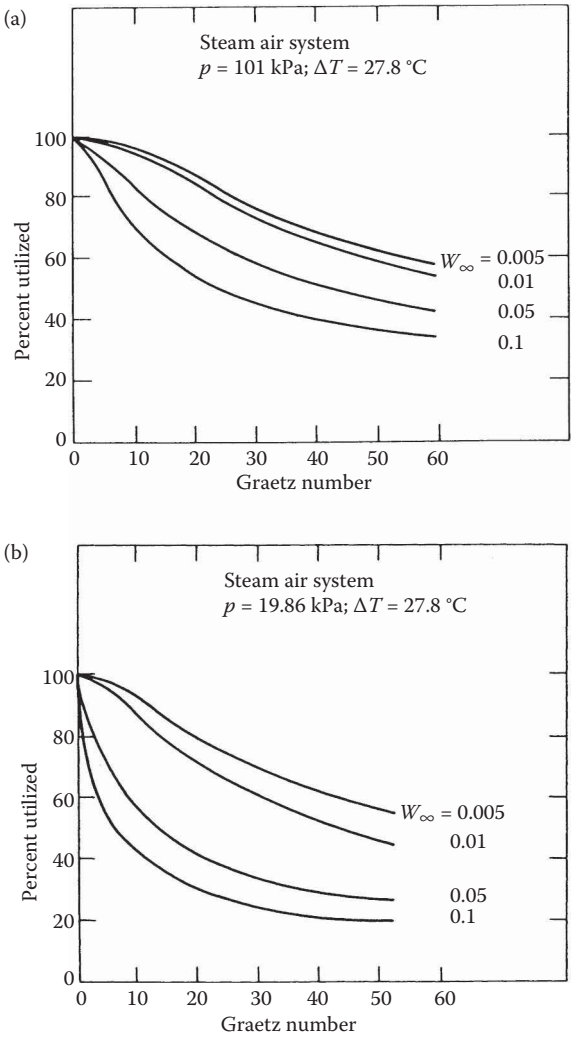


FIGURE 4.23.10 (a) Effect of W_∞ on percent coolant utilized for condensation on a sheet, $P=10$ kPa and (b) effect of W_∞ on percent coolant utilized for condensation on a sheet, $P=19.86$ kPa. (From Jacobs, H.R. and Nadig, R., *ASME J. Heat Transfer*, 109, 1013, 1987. With permission.)

Liu et al. (1989) carried out experiments with round jets, which ultimately broke down and formed drops whose diameter were twice that of the jet. Their results, shown in Figures 4.23.11a and b indicate that the heat transfer for condensation was little different than if the jet had not fragmented. This was because although the surface area was greater, the rate of diffusion was smaller, thus negating the former. Kim and Mills (1989a,b) studied condensation on turbulent jets and noted that the external to the jet heat transfer resistance governed, as noted above for sheets.

Direct Contact Condensation Falling Films

In addition to the barometric condensers mentioned in the preceding subsections, some condensers have been designed as packed beds. The packing can be varied from structured to nonstructured types and be made of a variety of materials. The materials are chosen so that the coolant

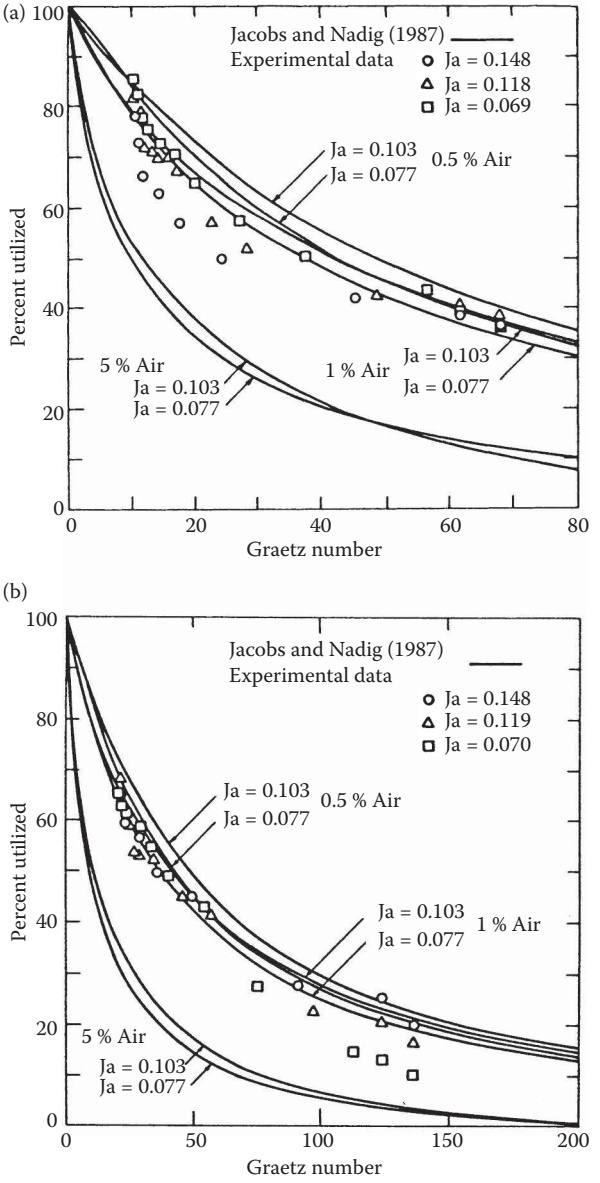


FIGURE 4.23.11 (a) Percent coolant utilized for 12.7 mm head of water and (b) percent coolant utilized for 50.8 mm head of water. (From Liu, T.L., Jacobs, H.R., and Chen, K., *ASME J. Heat Transfer*, 111, 585, 1989. With permission.)

wets the packing and flows across it as a thin film which is reformed as it drips from one level of packing to the next in a manner analogous to condensate dripping from one tube to the next in a surface condenser.

The modeling of falling films of coolant with subsequent condensation are similar to those for sensible heat transfer mentioned previously, but considerably more complicated, especially with noncondensibles present. Seven equations must be solved simultaneously. Even for a pure vapor the problem is difficult, as shown by Jacobs and Bogart (1980), Rao and Sarma (1984), etc. Jacobs and Nadig (1983) deal with noncondensibles being present as well as the coolant and condensate being immiscible.

Jacobs et al. (1982) studied film condensation on a thin film flowing over a sphere. Using their results, they present a correlation for the mean fluid temperature leaving the sphere as

$$\frac{T_m - T_{SAT}}{T_0 - T_{SAT}} = 0.03\ \mathrm{Ja}^{-0.4}\,\overline{R}^{-1/2}\mathrm{Re}^{1/3}\mathrm{Pr}^{0.36} \tag{4.23.28}$$

for the coolant and condensate being the same. In the above, $\overline{R}$ is $R\Big/\left(\frac{v^2}{g}\right)^{1/3}$ and R is the radius of the sphere.

The results showed very good agreement when compared to the experimental data of Finklestein and Tamir (1976).

Repeated layers of spheres with reforming films which flow onto the succeeding layer with a new T_0 equal to T_m at the base of the previous sphere provide a model for a packed bed of spheres.

Jacobs et al. (1979) and Bharathan and Althof (1984) carried out experiments with a variety of packings and relative flow rates of vapor and coolants that account for a nonstagnant vapor. Celata et al. (1989) studied steam condensation on a turbulent film of water. They indicate that laminar models can be used when modifying the liquid thermal conductivity as

$$k_{eff} = k_f\left(0.21\mathrm{Re}^{0.34}/\mathrm{Pr}^{0.24}\right) \tag{4.23.29}$$

Direct Contact Heat Transfer with Solidification

Direct contact with solidification is found in several applications including the manufacture of small beads where a molten spray is injected into a coolant. Typical of this is the formation of lead and steel shot for shotguns. It is also used in freezing of food, freeze-desalting of brine, and freeze-coating of thin films on wires. Johnson et al. (1973) investigated freeze-desalting by injecting a cold refrigerant into brine. Ice formed at the interface and was recovered with the refrigerant. Cheung and Cha (1987) and Cheung (1990) studied freeze-coating of a moving wire with the intent of forming a uniform coating. These problems are all similar to condensation problems assuming, for example, condensation on a drop or condensation on a jet or sheet with the velocity in the stream being slug flow. The difference in the analysis is that the Jakob number is replaced by the Stefan number. Freezing of liquid drops follows the Stefan problem in conduction, depending on the circulation within the drops and external shear. For metal drops, the high surface tension eliminates the circulation and the solution is in fact a spherical Stefan problem. Of course, no analytical solution exists and numerical or integral methods must be used to solve the nonlinear problem.

Summary

The length of this section on direct contact heat transfer does not provide for a complete development of the subject. However, the references cited in the text will provide further insight to the multitude of complex problems that are direct contact in nature. The reader has been referred to other sections within the text as well. Of particular importance are the development of models for multiphase flow. So-called two-fluid models are often used to determine the heat transfer in direct contact systems. These models are typically one-dimensional Eulerian models. However, when dealing with one or more dispersed phases in a continuum, a so-called Lagrangian–Eulerian model also is used. The author and others have developed computer codes to deal with a range of such problems; nonetheless, continued work is needed, especially for problems where the holdup of the dispersed phase is large.

REFERENCES

Ahmad, G.R. and Yovanovich, M.M., Approximate analytical solution of forced convection heat transfer from isothermal spheres for all Prandtl numbers, *ASME J. Heat Transfer*, 116, 838–843, 1994.

Bankoff, S.G. and Lee, S.C., A critical review of the flooding literature, *Multiphase Sci. Technol*., 2, 95–180, 1986.

Battya, P., Raghaven, V.R., and Seetharamu, K.N., A theoretical analysis of direct contact evaporation of a moving drop in an immiscible liquid, *Proceedings of 3rd International Conference Numerical Methods in Thermal Problems*, Swansea, UK, 1983, pp. 1083–1093.

Bharathan, D., Direct contact evaporation, *Direct Contact Heat Transfer*, Kreith, F. and Boehm, R.F., Eds., Hemisphere, New York, 1988, Chapter 11, 203–221.

Bharathan, D. and Althof, J., An experimental study of steam condensation on water in countercurrent flow in the presence of inert gases, ASME Paper No. 84-WA-Sol-25, American Society of Mechanical Engineers, New York, 1984.

Bogy, D.B., Drop formation in a circular jet, *Annu. Rev. Fluid Mech*., 11, 207–228, 1979.

Brown, G., Heat transmission during condensation of steam on a spray of water drops, Institute of Mechanical Engineers, General Discussion on Heat Transfer, 1951, pp. 49–51.

Calderbank, P.H. and Korchinski, I.J.O., Circulation in liquid drops: (A heat-transfer study), *Chem. Eng. Sci*., 6, 65–78, 1956.

Carey, V.P. and Hawks, N.E., Stochastic modeling of molecular transport to an evaporating microdroplet in a superheated gas, *ASME J. Heat Transfer*, 117, 432–439, 1995.

Celata, G.P., Cumo, M., D'Annibale, F., and Farello, G.E., Direct contact condensation of steam on droplets, *Int. J. Multiphase Flow*, 117(2), 191–211, 1990.

Celata, G.P., Cumo, M., Farello, G.E., and Focandi, G., A Theoretical and experimental study of direct contact condenstation on water in turbulent flow, *Exp. Heat Transfer*, 2(2), 153–172, 1989.

Chen, L.D. and Faeth, G.M., Condensation of submerged vapor jets in subcooled liquids, *ASME J. Heat Transfer*, 104, 774–780, 1982.

Chen, L.D. and Faeth, G.M., Structure of turbulent reacting gas jets submerged in liquid metals, *Combust. Sci. Technol*., 31, 277–296, 1983.

Cheung, F.B., Solidification on a chilled continuous surface moving in a parallel free stream, *ASME J. Heat Transfer,* 112, 521–523, 1990.

Cheung, F.B. and Cha, S.W., Flow-freezing interaction during freeze coating on a non-isothermal moving cylinder, ASME Paper No. 87-Ht-4, American Society of Mechanical Engineers, New York, 1987.

Chung, J.N. and Ayyaswamy, P.S., Material removal associated with condensation on a droplet in motion, *Int. J. Multiphase Flow*, 7, 329–342, 1981a.

Chung, J.N. and Ayyaswamy, P.S., Laminar condensation heat and mass transfer of a moving drop, *AlChE J*., 27, 327–347, 1981b.

Chung, J.N., Ayyaswamy, P.S., and Sabhal, S.S., Laminar condensation on a moving drop, I. Singular perturbation technique, *J. Fluid Mech*., 139, 105–130, 1984a.

Chung, J.N., Ayyaswamy, P.S., and Sadhal, S.S., Laminar condensation on a moving drop. II. Numerical solutions, *J. Fluid Mech*., 139, 131–144, 1984b.

Coban, R. and Boehm, R.F., Numerical and experimental modelings of three phase direct contact heat exchangers, *Proceedings of 8th International Heat Transfer Conference*, Toronto, 1986, 3019–3024.

Dhir, V.K. and Purohit, G.P., Subcooled film boiling heat transfer from spheres, ASME Paper No. 77-HT-78, American Society of Mechanical Engineers, New York, 1977.

Eden, T.J., Buoyant condensing vapor jets, Ph.D. dissertation, Penn State University, University Park, PA, 1995.

Evans, M.S., Numerical solution of axisymmetric flow and heat transfer from spherical particles at the leading and trailing edges of an array, MSME thesis, Pennsylvania State University, 1994.

Faghri, A., Prediction of heat and mass transfer for absorption of a gas to a liquid film, *Numer. Methods Therm. Prob*., III, 981–990, 1983.

Faghri, A. and Saban, R.A., Heat transfer in wavy liquid films, *Int. J. Heat Mass Transfer*, 28(3), 506–509, 1985.

Faghri, A. and Seban, R.A., Heat and mass transfer to a turbulent liquid film, *Int. J. Heat Mass Transfer*, 31(4), 891–894, 1998.

Finkelstein, Y. and Tamir, A., Interfacial heat transfer coefficients of various vapors in direct contact condensation, *Chem. Eng. J*., 12, 199–209, 1976.

Floreschuetz, L.W. and Chao, B.T., On the mechanics of vapor bubble collapse, *ASME J. Heat Transfer*, 87, 209–220, 1965.

Ford, J.D. and Lekic, A., Rate of growth of drops during condensation, *Int. J. Heat Mass Transfer*, 16, 61–66, 1973.

Golafshani, M., Bubble type direct contact condensers, MS thesis, Department of Mechanical and Industrial Engineering, University of Utah, 1983.

Grace, J.R., Hydrodynamics of liquid drops in immiscible liquids, *Handbook of Fluids in Motion*, Cheremisinoff, N.P. and Gupta, R., Eds., Butterfield Group, Ann Arbor, MI, 1983, 1003–1025.

Hadamard, J., Mouvement permanent lent d'une sphere liquide et visquesuse dans une liquide visqueux, *C. R. Acad. Sci. Paris*, 152, 1735–1738, 1911.

Hasson, D., Luss, D., and Peck, R., Theoretical analysis of vapor condensation on liquid jets, *Int. J. Heat Mass Transfer*, 7, 969–981, 1964a.

Hasson, D., Luss, D., and Naron, V., An experimental study of steam condensing on a laminar water sheet, *Int. J. Heat Mass Transfer*, 7, 983–1001, 1964b.

Hausbrand, E., *Condensing and Cooling Apparatus*, 5th ed., Van Nostrand Reinhold, New York, 1933.

Hendricks, R.C. and Baumeister, K.J., Film boiling from submerged spheres, NASA TN D-5124, National Aeronautics and Space Administration, Washington, DC, 1969.

Huang, L.J. and Ayyaswamy, P.S., Drag coefficients associated with a moving drop experiencing condensation, *ASME J. Heat Transfer*, 109, 1003–1006, 1987a.

Huang, L.J. and Ayyaswamy, P.S., Heat and mass transfer associated with a spray drop experiencing condensation: A fully transient analysis, *Int. J. Heat Mass Transfer*, 30, 881–891, 1987b.

Isachenko, V.P. and Kushnyrev, V.J., Condensation heat transfer in dispersed liquid spray, *Proceedings of 5th International Heat Transfer Conference*, Tokyo, Japan, Vol. III, pp. 217–220, 1974.

Ishii, M., *Thermo-Fluid Dynamics of Two-Phase Flow*, Eyrolles, Scientific and Medical Publication of France, Paris, 1975.

Ishii, M. and Grolmes, M.A., Inception criteria for droplet entertainment in two-phase concurrent film flow, *AIChE J.*, 21, 308–318, 1975.

Jacobs, H.R., Direct contact condensation, *Direct Contact Heat Transfer,* Kreith, F. and Boehm, R.F., Eds., Hemisphere, New York, 1988a, 223–234.

Jacobs, H.R., Thermal and hydraulic design of direct contact spray columns for use in extracting heat from geothermal brines, *Direct Contact Heat Transfer*, Kreith, F. and Boehm, R.F., Eds., Hemisphere, New York, 1988b, 343–370.

Jacobs, H.R. and Bogart, J.A., Condensation on an immiscible thin film, ASME Paper No. 80-HT-110, American Society of Mechanical Engineers, New York, 1980.

Jacobs, H.R., Bogart, J.A., and Pensel, R.W., Condensation on a thin film flowing over an adiabatic sphere, *Proceedings of 7th International Heat Transfer Conference*, Vol. 5, pp. 89–94, 1982.

Jacbos, H.R. and Cook, D.S., Direct contact condensation on a non-circulating drop, *Proceedings of 6th International Heat Transfer Conference*, Toronto, ON, Vol. 3, pp. 389–393, 1978.

Jacobs, H.R. and Dolan, S.M., The Transfer of heat and momentum between a laminar falling film and a countercurrent gas stream, *J. Energy, Heat Mass Transfer*, 16, 77–85, 1994.

Jacobs, H.R. and Eden, T.J., Direct contact heat transfer in a sieve tray column, *Proceedings of 8th International Heat Transfer Conference*, San Francisco, CA, Vol. 6, pp. 3013–3018, 1986.

Jacobs, H.R. and Fannir, H., Direct contact condensers: A literative survey, USDOE Report, DOE/1523-3, Department of Energy, Washington, DC, 1977.

Jacobs, H.R., Fannar, H., and Beggs, G.C., Collapse of a bubble vapor in an immiscible liquid, *Proceedings of 6th International Heat Transfer Conference*, Vol. 2, 383–388, 1978.

Jacobs, H.R. and Fletcher, M.J., Heat transfer from a falling film to a countercurrent gas stream, *Proceedings of ASME National Heat Transfer Conference*, Houston, TX, HTD Vol. 96, No. 3, pp. 281–289, 1988.

Jacobs, H.R. and Major, B.H., The effects of noncondensible gases on bubble condensation in a immiscible liquid, *ASME J. Heat Transfer*, 104, 487–492, 1982.

Jacobs, H.R. and Nadig, R., Condensation on an immiscible falling film in the presence of a non-condesible gas, *Heat Exchangers for Two Phase Applications*, ASME HTD-Vol. 27, American Society of Mechanical Engineers, New York, 1983, 99–106.

Jacobs, H.R. and Nadig, R., Condensation on coolant jets and sheets including the effects of noncondensible gases, *ASME J. Heat Transfer*, 109, 1013–1020, 1987.

Jacobs, H.R., Thomas, K.D., and Boehm, R.F., Direct contact condensation of immiscible fluids in tacked beds, *Condensation Heat Transfer*, American Society of Mechanical Engineers, New York, 1979, 103–110.

Johnson, K.M., Jacobs, H.R., and Boehm, R.F., Collapse height for condensing vapor bubbles in an immiscible liquid, *Proceedings of Joint ASME/JSME Heat Transfer Conference*, Vol. 2, pp. 155–163, 1983.

Johnson, W., Pallone, A., and Probstein, R., Freeze desalting—A new approach, *Proceeding of 4th International Symposium on Fresh Water from the Sea*, Heidelberg, Germany, September 9–14, 3, 371–382, 1973.

Kazami, M.S., Todreas, N.E., Lanning, D.D., and Rohsenow, W.M., A theoretical study of the dynamics growth of a vapor film around a hot sphere in a coolant, *Proceedings of 5th International Heat Transfer Conference*, Tokyo, Japan, Vol. IV, pp. 91–95, 1974.

Kim, S. and Mills, A.F., Condensation on coherent liquid jets. I. Experimental studies, *ASME J. Heat Transfer*, III, 1068–1074, 1989a.

Kim, S. and Mills, A.F., Condensation on coherent liquid jets. II. A theoretical study, *ASME J. Heat Transfer*, III, 1075–1082, 1989b.

Klipstein, D.H., Heat transfer to a vaporizing drop, Doctor of Science thesis, Massachusetts Institute of Technology, 1963.

Kodres, C.A., Jacobs, H.R., and Boehm, R.F., A Numerical method for determining direct contact heat transfer takes to a superheated, evaporating, floating droplet, *J. Numer. Heat Transfer*, 3, 21–34, 1980.

Kodres, C.A, Jacobs, H.R. and Boehm, R.F., Design and evaluation of evaporating lens direct contact heat exchangers, *ASME J. Energy Resour. Technol.*, 103, 128–132, 1981.

Kutateladze, S.S., Heat transfer in boiling and condensation, Moscow, English Transl. AEC TR-3770, U.S. Atomic Energy Commission, Washington, DC, 1952.

Law, C.K., Heat and mass transfer in combustion: Fundamental concepts and analytical techniques, *Prog. Energy Combust. Sci*., 10, 295–318, 1984.

Lee, S.C. and Bankoff, S.G., Stability of steam-water countercurrent flow in an inclined channel: Flooding, *ASME J. Heat Transfer*, 105, 713–718, 1983.

Levich, V.G., *Physiochemical Hydrodynamics*, Prentice Hall, Englewood Cliffs, NJ, 1962.

List, E.J., *Turbulent Jets and Plumes*, *Annual Review of Fluid Mechanics*, Academic Press, New York, 1982, 189–212.

Liu, T.L., Jacobs, H.R., and Chen, K., An experimental study of direct contact condensation on a fragmenting circular jet, *ASME J. Heat Transfer*, 111, 585–588, 1989.

McQuillan, K.W., Whalley, P.B., and Hewitt, G.F., Flooding in vertical two phase flow, *Int. J. Multiphase Flow*, 11(6), 741–760, 1985.

Moalem-Maron, D., Sideman, S., Hetsroni, G., and Orell, A., Condensation of bubble trains: An approximate solution, *Prog. Heat Mass Transfer*, 6, 155–177, 1972.

Murty, N.S. and Sarma, V.M.K., Direct contact heating of laminar fall liquid jet, *Int. J. Heat Mass Transfer*, 19, 115–117, 1976.

Murty, N.S. and Sastri, V.K.M., Condensation on a falling laminar liquid film, *Proceedings of 5th International Heat Transfer Conference*, Tokyo, Japan, Vol. III, 231–235, 1974.

Nusselt, W., Die Oberflachenkondensation des Wasserdamfes, *Z. Ver. Dt. Ing*., 60, 541–569, 1916.

Ostrach, S. and Koestel, A., Film instabilities in two-phase flow, *AlChE J*., II, 294–303, 1965.

Prakash, S. and Sirignano, W.A., Theory of convection droplet vaporization with unsteady heat transfer in the circulating phase, *Int. J. Heat Mass Transfer*, 23, 253–268, 1980.

Raina, G. and Grover, P., Direct contact heat transfer with change of phase: Theoretical model, *AlChE J*., 28, 515–517, 1982.

Raina, G. and Grover, P., Direct contact heat transfer with change of phase: Theoretical model incorporating sloshing effects, *AlChE J*., 31, 507–510, 1985.

Raina, G. and Wanchoo, R., Direct contact heat transfer with phase change: Theoretical expression for instantaneous velocity of a two-phase bubble, *Int. Commun. Heat Mass Transfer*, 11, 227–237, 1984a.

Raina, G., Wanchoo, R., and Grover, R, Direct heat transfer with phase change: Motion of evaporating droplets, *AlChE J*., 30, 835–837, 1984b.

Rangel, R.H. and Sirignano, W.A., Vaporization ignition and combustion of two parallel fuel droplet streams, *Proceedings of ASME/JSME Thermal Engineering Joint Conference*, Honolulu, Hawaii, Vol. 6, pp. 27–34, 1987.

Rao, V.D. and Sarma, P.K., Condensation heat transfer on laminar liquid film, *ASME J. Heat Transfer*, 106, 518–523, 1984.

Roberson, M.E. and Jacobs, H.R., An experimental study of mass transfer in packed beds as an analog to convective heat transfer, *Proceedings of 9th International Heat Transfer Conference*, Jerusalem, Israel, Vol. 5, pp. 189–194, 1990.

Rose, P.M. and Kintner, R.C., Mass transfer from large oscillating drops, *AlChE J*., 12(3), 530–534, 1966.

Sanders, B.R. and Dwyer, H.H., Modeling unsteady droplet combustion processes, *Proceedings of ASME/JSME Thermal Engineering Joint Conference*, Honolulu, Hawaii, Vol. 1, pp. 3–10, 1987.

Seban, R.A., Transport to falling films, *Proceedings of 6th International Heat Transfer Conference*, Toronto, ON, Vol. 6, pp. 417–428, 1978.

Segev, A., Flanigan, L.J., Kurth, R.E., and Collier, R.P., Experimental study of countercurrent steam condensation, *ASME J. Heat Transfer*, 103, 307–311, 1981.

Sideman, S., *Direct Contact Heat Transfer between Immisicible Liquids*, *Advances in Chemical Engineering*, Academic Press, New York, 1966, 207–286.

Sideman, S. and Isenberg, J., Direct contact heat transfer with change of phase: Bubble growth in three-phase systems, *Desalin.*, 2, 207–214, 1967.

Sideman, S. and Moalem-Maron, D., *Direct Contact Condensation*, *Advances in Heat Transfer*, Academic Press, New York, 1982, 228–276.

Sideman, S. and Taitel, Y., Direct contact heat transfer with change of phase: Evaporation of drops in an immiscible liquid medium, *Int. J. Heat Mass Transfer*, 7, 1273–1289, 1964.

Simpson, H.C., Beggs, G.C., and Nazir, M., Evaporation of butane drops in brine, *Desalin.*, 15, 11–23, 1974.

Sirignano, W.A., Fuel droplet vaporization and spray combustion theory, *Prog. Energy Combust. Sci.*, 9, 291–322, 1983.

Stefan, M. and Mayinger, F., Countercurrent flow limitation in vertical ducts at high system pressure, *Proceedings of International Heat Transfer Conference*, Jerusalem, Israel, Vol. 6, pp. 47–52, 1990.

Sudhoff, B., Director Warmeubergang bei der Kondensation in Blap Finsaulen, Ph.D. dissertation, Universitat Dortmund, Federal Republic of Germany, Dortmund, Germany, 1982.

Sun, T.Y., A theoretical and experimental study of noncondensible turbulent bubbly jets, Ph.D. dissertation, Department of Mechanical Engineering. Pennsylvania State University, 1985.

Sundararajan, T. and Ayyaswamy, P.S., Hydrodynamics and heat transfer associated with condensation on a moving drop: solutions for intermediate Reynolds numbers, *J. Fluid Mech.*, 149, 33–58, 1984.

Sundararajan, T. and Ayyaswamy, P.S., Heat and mass transfer associated with condensation on a moving drop: solution for intermediate Reynolds numbers by a boundary layer formulation, *ASME J. Heat Transfer*, 107, 409–416, 1985a.

Sundararajan, T. and Ayyaswamy, P.S., Numerical evaluation of heat and mass transfer to a moving drop experiencing condensation, *Numerical Heat Transfer*, 8, 689–706, 1985b.

Taitel, Y. and Tamir, A., Condensation in the presence of a noncondensible gas in direct contact, *Int. J. Heat and Mass Transfer*, 12, 1157–1169, 1969.

Tekic, M.N., Prosarac, D., and Petrovic, D., Entrance region lengths of laminar falling films, *Chem. Eng. Sci.*, 39, 165–167, 1984.

Thompson, R.J. and Jacobs, H.R., Mass and simultaneous convective heat transfer for flow in a packed bed of spherical particles, *AlChE Symp. Ser.*, 80(236), 240–248, 1984.

Tien, C.L. and Liu, C.P. Survey on vertical two-phase countercurrent flooding, EPRI NP-984, Electrical Power Research Institute, Palo Alto, CA, 1979.

Wagner, I. and Stichlmair, J., Motion of swarms of droplets, *Proceedings of International Symposium Liquid-Liquid Two Phase Flow Transport Phenomena*, International Center Heat and Mass Transfer, Antalya, Turkey, 1997, 87–96.

Weimer, J.C., Faeth, G.M., and Olson, D.R., Penetration of vapor jets in subcooled liquids, *AlChE J.*, 19(3), 552–558, 1973.

Whitaker, S., Forced convection heat transfer correlations for flow in pipes, past flat plates, single cylinders, single spheres and for flow in packed beds and tube bundles, *AlChE J.*, 18, 361–371, 1972.

Wilson, M.R. and Jacobs, H.R., A numerical solution of the heat transfer from a dense array of spherical particles, *ANS Proceedings of National Heat Transfer Conference*, HTC, Atlanta, GA, Vol. 7, pp. 225–233, 1993.

4.24 HEAT TRANSFER IN THE PRESENCE OF SYNTHETIC JETS

Mangesh Chaudhari and Amit Agrawal

Introduction

A synthetic jet is synthesized directly from the fluid of the system in which it is embedded. It is commonly formed from a small closed cavity having a diaphragm at one end and a small opening at the other. In the process of jet formation, the fluid is alternately sucked into and ejected from the cavity due to vibration of the diaphragm (Figure 4.24.1). Owing to zero net mass addition to the system, the synthetic jet is also known as a zero-net-mass-flux (ZNMF) device. At low actuating frequency (less than the jet formation frequency) of vibration of the diaphragm, the fluid enters the cavity during the suction stroke and ejected out from the cavity during the ejection stroke, without any jet formation. However, an increase in the excitation frequency increases the rate of suction and ejection of fluid from the cavity, and beyond a certain frequency (the jet formation frequency) the fluid vortices detach from the cavity surface and travel into the surrounding medium. Further increase in the excitation frequency increases the formation rate of vortices; and these vortices coalesce with each other in the medium to form a synthetic jet. However, there is an upper limit on frequency until which the synthetic jet forms.

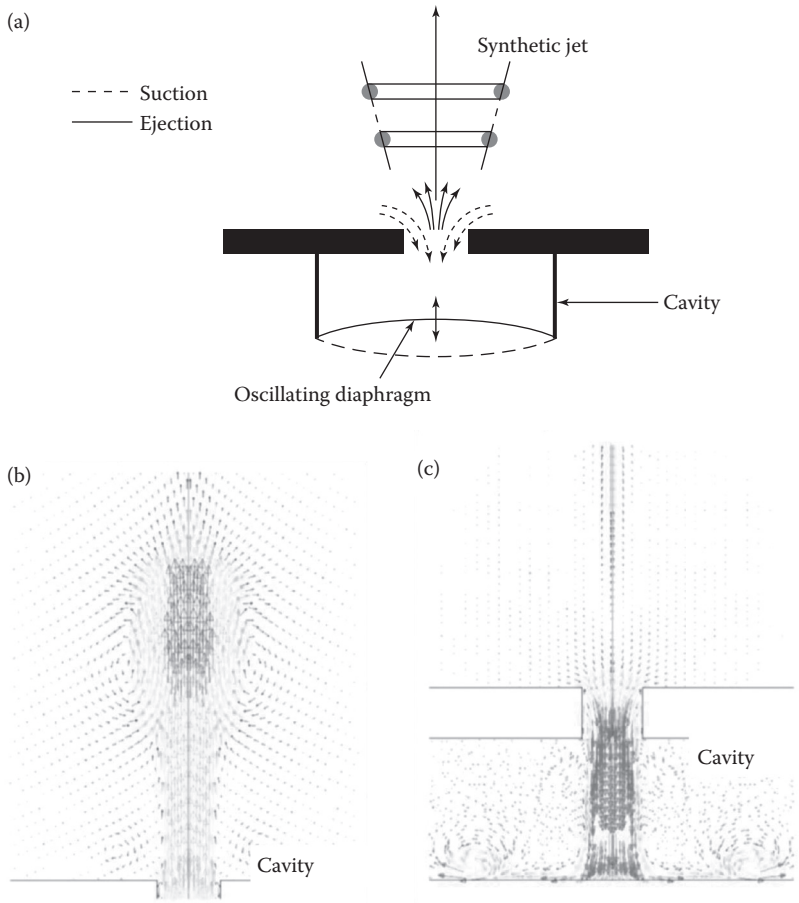


FIGURE 4.24.1 (a) Schematic of synthetic jet operation showing the suction and ejection strokes, respectively. (b) and (c) Velocity vectors obtained from computational fluid dynamics calculations by Jain et al. (2011), showing vortex formation during the suction and ejection strokes, respectively. (From Jain, M. et al., *Sens. Actuators, A,* 165, 351–366, 2011. With permission.)

The aforementioned feature obviates the need for input piping or complex fluidic packaging and makes synthetic jets ideally suited for low-cost batch fabrication using micromachining techniques. Owing to the pulsating nature of the flow, the entrainment of fresh ambient fluid into the jet is high. This leads to a relatively large mass flux downstream of the orifice and a relatively large heat transfer coefficient. The use of "synthetic jet" to drive the air is a relatively new approach for electronics cooling, and its usefulness has been explored in previous studies (Mahalingam and Glezer, 2005; Pavlova and Amitay, 2006; Arik, 2007; Chaudhari et al., 2010a).

Owing to the vortical nature of the flow during the suction and ejection processes, no potential core exists in the synthetic jet, which makes this flow different as compared to a normal jet. It has been further observed that there is substantial mixing of the fluid near the orifice of a synthetic jet; and that the turbulence intensities are about four times higher than that of a normal jet. These features make the use of synthetic jet for cooling of a hot surface attractive; their use for cooling of compact devices has also been explored.

Electronic cooling is an area where the use of synthetic jet has been actively explored. Besides cooling, synthetic jets have a number of other potential engineering applications, such as boundary-layer separation control, jet vectoring, better mixing of fuel in the engine combustion chamber, creation of local turbulence, and vehicle propulsion. A synthetic jet is therefore a promising novel technique having some unique characteristics (better mixing, zero net mass flux, etc.), which makes them particularly useful for several engineering applications.

Flow Characteristics of Synthetic Jet

The synthetic jet is formed in a stagnation fluid due to vibration of a diaphragm in a closed cavity having one or more openings at some other end of the cavity. Owing to the oscillatory nature of the flow, calculation of Reynolds number is not straightforward. The Reynolds number of the synthetic jet is calculated as (Smith and Glezer, 1998):

$$\mathrm{Re} = \frac{U_o d}{\upsilon} \tag{4.24.1}$$

where d is the orifice diameter, υ is the kinematic viscosity, and U_o is the average orifice velocity during the ejection part of the cycle. The last parameter is calculated from measurement (using, for example, a hotwire) of the instantaneous streamwise velocity ($u(t)$) at the exit and centerline of the orifice, as

$$L_o = \int_0^{\tau/2} u(t)\,\mathrm{d}t \tag{4.24.2}$$

$$U_o = L_o f \tag{4.24.3}$$

where L_o is the stroke length, f is the excitation frequency (or inverse of time period τ), and t is time. Various researchers have explored the heat transfer and flow characteristics of synthetic jets both experimentally and numerically. Initial studies were geared toward understanding the flow physics of synthetic jet, and later their heat transfer characteristics were explored. Schlieren imaging, particle image velocimetry, hot wire anemometry, thermal imaging, liquid crystals, and interferometery are some of the measurement techniques used for the study of synthetic jet (Smith and Glezer, 1998; Zhang and Tan, 2007; Chaudhari et al., 2009; Qayoum et al., 2010; Bhapkar et al., 2014). The study has been carried out for both planar and axisymmetric synthetic jets using different actuators (such as piezoelectric and electromagnetic). Lumped element modeling (LEM) (Gallas et al., 2003) and similarity analysis (Agrawal and Verma, 2008) of a synthetic jet have been carried out in the

literature. It was noticed that the synthetic jet formation is governed by the jet Strouhal number (or Stokes number) and found to be different for planar and axisymmetric jets (Holman et al., 2005).

Broadly, the following set of events takes place from the time of starting the actuator and as the excitation frequency is increased. The smoke required in the experiments is generated using an incense stick (*agarbatti*) and images are captured using a digital camera.

1. When the actuator is off, there is no flow from the orifice and the smoke from the incense stick rises vertically upward (not shown). When the actuator is excited at a very low frequency (around 1 Hz), initially there is a continuous suction and ejection of the smoke in the form of vortex rings from the orifice, due to the movement of the diaphragm. However, the vortex rings generated during the ejection stroke are sucked back into the cavity during the suction stroke; so effectively, the vortex rings do not come out of the cavity. The effect of the pressure waves generated by the oscillatory diaphragm can however be seen as a disturbance to the otherwise undisturbed smoke line flowing in the vertical direction (Figure 4.24.2a). Although clearly visible to the naked eye, the sucking back of the vortices is rather difficult to photograph. The generation and suction of vortices are similar to those shown in the numerical simulations (Figure 4.24.1c).
2. As the excitation frequency is increased, the frequency of vortex formation also increases, but still the vortices that are formed are eventually sucked back into the cavity. This occurs until a frequency of approximately 6–8 Hz is reached. Interestingly, the frequency of excitation seems to be reflected in the smoke line as a corresponding frequency of disturbance. Eventually, a critical frequency (at about 9 Hz), termed as the starting point for the synthetic jet to form, is reached. At this point, for the first time, the vortex ring crosses the smoke line and propagates in the horizontal direction (not shown). Due to increase in the momentum of the vortex ring, it is able to overcome the suction created during the backward movement of the diaphragm.
3. On a further increase of frequency (to about 20 Hz), the vortex rings are expelled out of the orifice (Figure 4.24.2b). The vortex-shedding frequency increases with the excitation frequency, along with the momentum of the rings. The rings now travel some distance in the horizontal direction beyond which they become indistinguishable from the ambient fluid. Initially, the distance between the rings is large but it reduces with an increase in the excitation frequency. The vortex rings lose their momentum in the far field, and eventually coalesce with the rings ejected in the previous forcing cycle.

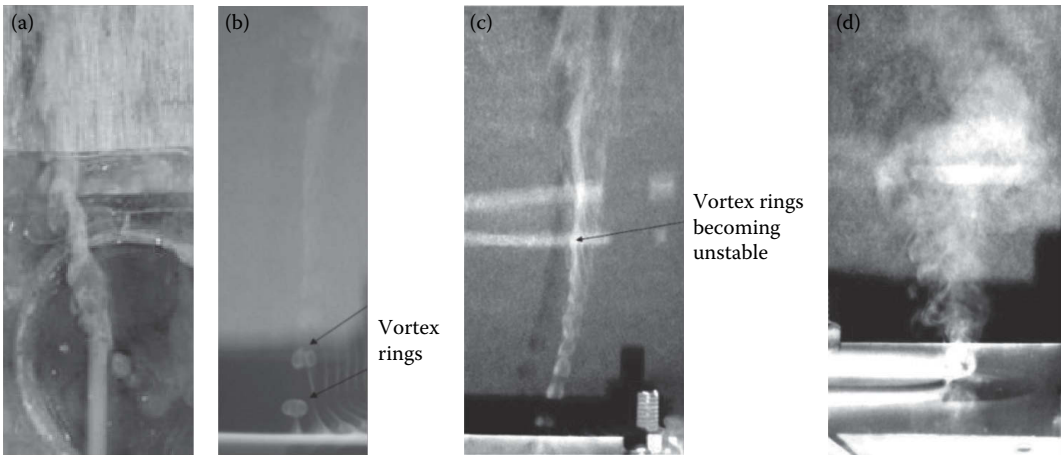


FIGURE 4.24.2 Flow visualization images showing the various stages of synthetic jet. (a)–(d) correspond to a gradual increase in the excitation frequency. See text for details.

4. As the frequency is further increased (to 20–30 Hz), coalescence starts taking place much closer to the orifice and slowly it takes the form of a continuous jet that appears laminar in appearance even in the far field. At a frequency of about 40 Hz, instabilities are seen to set up in the far field of the jet (Figure 4.24.2c). Due to these instabilities, the flow seems to become helical intermittently, and the laminar structure of the jet breaks down. The absence of potential core is also evident from the image.
5. Increasing the frequency further to about 100 Hz increases the instabilities and eventually a turbulent jet is seen to form (Figure 4.24.2d). The turbulence intensity of this jet increases with a corresponding enhancement in the dissipation of smoke in the flow field. Different length scales corresponding to eddies of different sizes can be discerned.

Smith and Glezer (1998) and Mallinson et al. (2001) observed that there is large entrainment in the near field of both planar and axisymmetric synthetic jets, due to the oscillating nature of the flow. The synthetic jet is found to be wider, slower, and having more momentum as compared to a continuous jet in the near field, while in the far field the characteristics of synthetic and continuous jets are similar in several aspects (Smith and Swift, 2003; Agrawal and Verma, 2008). The synthetic jet flow characteristics have been studied for different flow and geometrical parameters of the cavity (Chaudhari et al., 2009) using an electromagnetic actuator. The geometrical parameters of the cavity include the depth of cavity, orifice diameter, shape of orifice, and number of orifices. Hotwire anemometer is used for measurement of suction and ejection velocity. The ejection velocities are found to be larger as compared to the suction velocities. The ratio of ejection to suction average velocities decreases monotonically with frequency—from about 2.3 to 1, for 3 mm orifice diameter (Chaudhari et al., 2009). Note that this ratio is dependent on the orifice diameter, as the orifice diameter increases the ratio decreases suggesting a relative decrease in the ejection velocity. A lower bound and an upper bound on the formation frequency of a synthetic jet are also observed from Figure 4.24.3. Also, two resonant frequencies have been identified out of which the first resonance frequency (in the range of 250–300 Hz) seems almost independent of both the cavity depth and orifice diameter, whereas the second resonance frequency was found to be dependent on the cavity height and it decreased with an increase in cavity height. Also, the second resonance peak dampens out with an increase in the orifice diameter. The first resonance frequency is independent of all the configurations of the actuator known as diaphragm resonance frequency, and the second resonance frequency depends on the cavity volume termed as Helmholtz resonance frequency. These resonance frequencies can be varied as per the requirements, by proper selection of the actuator as well as the cavity volume. Turbulence intensities were found to be larger (ranging around 3–4) for different orifice diameter and almost independent of the cavity volume and frequencies for the same actuator (Chaudhari et al., 2009). As compared to the normal jet, the synthetic jet turbulence intensities are very high near the orifice, i.e., the fluid mixing takes place at a faster rate, which leads to the use of a synthetic jet for cooling of surfaces in compact devices.

Chaudhari et al. (2009) observed that the input power requirement for a synthetic jet is about 1.4–1.6 W. Interestingly, the power requirement at the resonance frequency is slightly lower (by about 10%) as compared to that at the other frequencies. The issue of noise generation due to the operation of the synthetic jet has been addressed by Bhapkar et al. (2013, 2014). Noise levels between 48–68 dB have been measured by them, which are well within the comfort limit. Operating the synthetic jet cavity at the resonance frequency is found to be beneficial from several aspects: high ejection velocity, low noise generation, low input power requirement, and high heat transfer coefficient. The reason for this observation stems from the fact that a part of the input power supplied to the actuator is consumed in moving the diaphragm, while the remaining amount is dissipated through noise and other losses. A larger amount of noise at frequencies other than the resonance frequency and at higher excitation voltages can therefore be expected.

It was noticed that the synthetic jet spreads rapidly along the direction of the minor axis of the rectangular orifice, while along the major axis the synthetic jet initially contracted before spreading

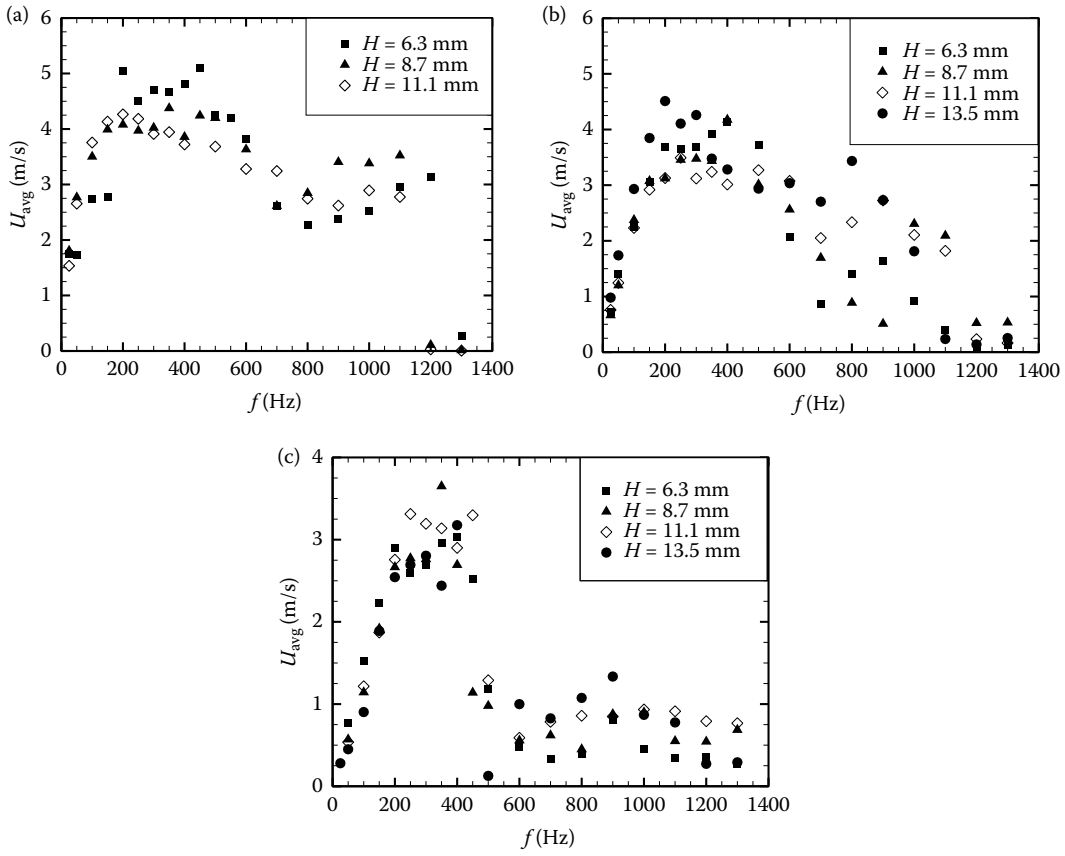


FIGURE 4.24.3 Variation of average velocity with excitation frequency for different cavity depths and orifice diameters. The orifice diameters are: (a) 3 mm, (b) 5 mm, and (c) 8 mm. (From Chaudhari, M. et al., *Exp. Therm. Fluid Sci*., 33, 439–448, 2009. With permission.)

slowly. The cooling region was observed to be wider with a synthetic jet than with a continuous jet, as deduced from the local temperature distribution (Zhang and Tan, 2007).

Heat Transfer Characteristics

The flow behavior of synthetic jet suggests that it can be employed for enhancing the heat transfer rate from a surface. The hot surfaces can be cooled by employing a synthetic jet parallel to the surface (such as in a duct or microchannel) or by impinging on a surface. The study was carried out to measure the performance of impinging synthetic jet for cooling of electronic components. The experiments were conducted for different excitation frequencies at different orifices to impinging surface distances. Synthetic jet gives optimum cooling performance at orifice to a surface distance of six times the orifice diameter (Figure 4.24.4), for all Reynolds numbers investigated in Pavlova and Amitay (2006) and Chaudhari et al. (2010a).

The vortices of the synthetic jet impinging on the surface individually are responsible for enhancing the heat transfer from the surface; after impingement these vortices break down into smaller secondary vortices to increase the rate of heat transfer. The maximum heat transfer coefficient with synthetic jet is found to be 4–11 times that of the natural convection heat transfer coefficient (Arik, 2007; Chaudhari et al., 2010a). The exponent of Reynolds number for the maximum Nusselt number is found to vary between 0.6 and 1.4 (Figure 4.24.4), which is larger as compared to a continuous jet (~0.5).

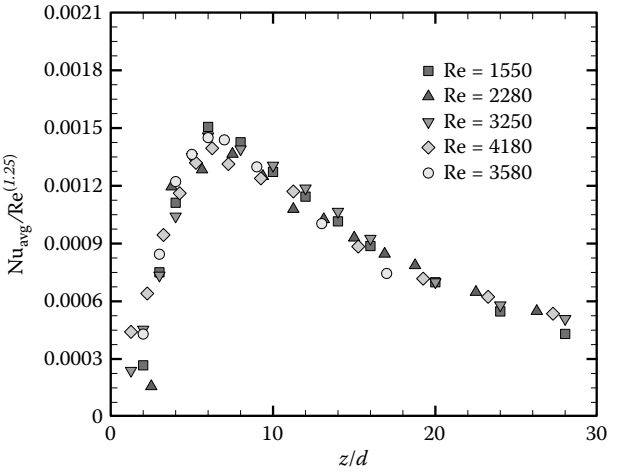


FIGURE 4.24.4 $Nu_{avg}/Re_{1.25}$ vs z/d showing dependence of Nusselt number on Reynolds number. (From Chaudhari, M. et al., *Int. J. Heat Mass Transfer,* 53, 1057–1069, 2010a. With permission.)

Note that the data for different Reynolds numbers collapse well on to a single curve with a maximum scatter of 3% around the mean value. Due to the high exponent of Reynolds number, synthetic jets are expected to perform substantially better at high Reynolds number as compared to a continuous jet.

The heat transfer performance of the continuous jet is found to be better at the lower axial distances as compared to the synthetic jet for the same Reynolds number. The fresh fluid input is provided by the continuous jet at all axial distances but for the synthetic jet at lower axial distances the same fluid is sucked and ejected, which degrades the cooling performance of the synthetic jet (Chaudhari et al., 2010a). Based on the experimental data for the circular orifice synthetic jet, the following correlation has been proposed to calculate the average Nusselt number:

$$\frac{\mathrm{Nu_{avg}}}{\mathrm{Pr}^{0.333}} = 7.624(\mathrm{Re})^{0.792}\left(\frac{L}{d}\right)^{-2.186}\left(\frac{R}{d}\right)^{2.258}\left(\frac{z}{d}\right)^{-0.632} \tag{4.24.4}$$

In the above correlation, d is orifice diameter, L is enclosure size, ($2R$) is heater size, and z is axial distance between orifice and surface. The above correlation is valid for: Re = 1150–4180, L/d = 7.86–22, R/d = 1.5–4, and $z/d > 3$ (Chaudhari et al., 2010a). A general correlation for the stagnation point heat transfer performance of an impinging synthetic jet is obtained for a wide parameter range, incorporating the effect of Reynolds number ($500 < \mathrm{Re} < 1500$), jet-to-surface spacing ($2 < z/d < 16$), and stroke length ($2 < L_0/D < 40$) (Persoons et al., 2011),

$$\mathrm{Nu_0} = 0.1676\mathrm{Re}^{0.686}\mathrm{Pr}^{0.4} f(H) \tag{4.24.5}$$

where, $f(z) = \dfrac{1+1.108\left(\dfrac{z/d}{5.21}\right)^2}{1+\left(\dfrac{z/d}{5.21}\right)^{2.487}}$.

The cooling performance of the impinging synthetic jet has been studied for different sizes and shapes of orifice (viz., square, rectangular, and elliptical with different aspect ratios, oval and diamond). It has been observed that a square orifice performs better as compared to the circular and rectangular

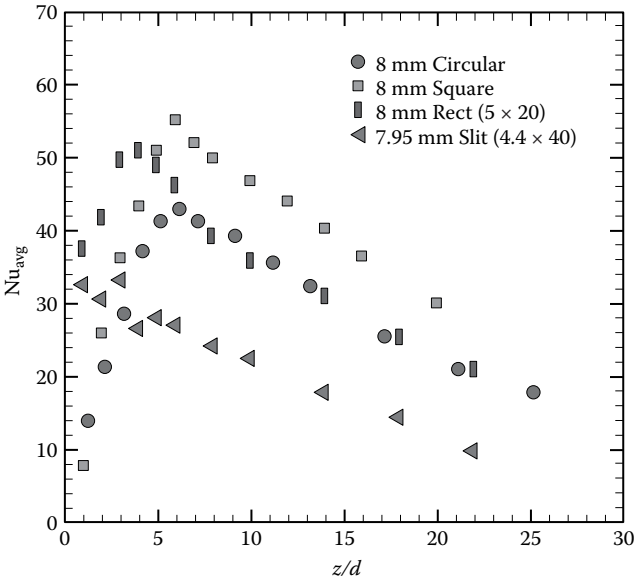


FIGURE 4.24.5 Variation of average Nusselt number with the normalized axial distance for different shapes of orifice having same hydraulic diameter with excitation frequency of 200 Hz. (From Chaudhari, M. et al., *Exp. Therm. Fluid Sci.*, 34, 246–256, 2010b. With permission.)

orifices for the same hydraulic diameter due to the larger mass flux (Chaudhari et al., 2010b) as shown in Figure 4.24.5. Further, a rectangular orifice gives better cooling performance as compared to the other shapes of orifice with the same hydraulic diameter at lower axial distances. For rectangular orifices, at lower axial distances, there is no recirculation and the shape of the orifice is retained as the flow emerges from the orifice; therefore, covering larger surface area for cooling. The following correlation from the experimental data for all orifice shapes is proposed by Chaudhari et al. (2010b):

$$\frac{\mathrm{Nu}_{\mathrm{avg}}}{\mathrm{Pr}^{0.333}} = 0.014(\mathrm{Re})^{0.841}\left(\frac{L}{d}\right)^{1.514}\left(\frac{R}{d}\right)^{-1.725}\left(\frac{z}{d}\right)^{-0.607} \tag{4.24.6}$$

The above correlation is valid for: $L/d = 7.86–22$, $R/d = 1.5–4$, $z/d = 3–25$, and $\mathrm{Re} = 600–4200$. The diamond and oval shape orifices give 17% and 7%, respectively, increase in heat transfer as compared to a circular orifice of the same hydraulic diameter (Bhapkar et al., 2014; Mangate and Chaudhari, 2015).

Multiple-orifice, single-cavity impinging synthetic jet cooling performance has also been explored for different locations and number of satellite orifices with respect to a center orifice. Further results are available for different Reynolds numbers and normalized axial distance (Chaudhari et al., 2011). The maximum heat transfer coefficient with multiple-orifice, single-cavity is found to be 30% higher as compared to a single orifice synthetic jet. Further, two peaks of the Nusselt number are observed with respect to normalized axial distance (z/d) for multiple-orifice single-cavity impinging synthetic jet, as shown in Figure 4.24.6. The first peak is located at $z/d = 2$, and the second peak is at $z/d = 9$. The appearance of the first peak is particularly significant as it can be utilized in space-constrained applications. The two peaks are observed for 1, 2, and 4 number of satellite orifices along with the center orifice due to the enhanced circulation of the fluid. However, for 8 number of satellite orifices, a ring is formed around the center orifice, which restricts entrainment of the fresh fluid and only a single peak is observed. The first peak shows a larger increase in Nusselt number with respect to the baseline value of no satellite orifice as compared to the second

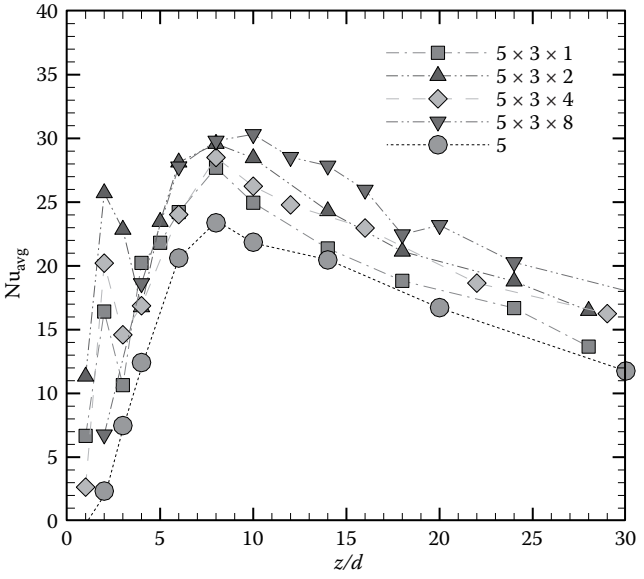


FIGURE 4.24.6 Comparison of average Nusselt number for 5 mm center orifice and different satellite holes around center of same pitch along the normalized axial distance for Re=2600. (From Chaudhari, M. et al., *Int. J. Heat Mass Transfer*, 54, 2056–2065, 2011. With permission.)

peak. There are a sharp increase and a sharp decrease of the average Nusselt number at around $z/d=2$ for the satellite orifices of 1, 2, and 4 number. The two satellite orifices along with the center orifice give maximum cooling performance at $z/d=2$, which is required in the cooling of compact devices. The number of satellite orifices with the center orifice plays an important role to get the two peaks of Nusselt number. Multiple satellite orifices without the center orifice cover the maximum surface area, which gives better performance as compared to a single center orifice. According to requirements, a proper configuration can thus be selected with different number of satellite holes and with/without the center orifice.

Besides improvement in the value of heat transfer coefficient and appearance of an additional peak, the benefit with multi-orifice, single-cavity synthetic jet is a slight reduction in the input power requirement as compared to a single center orifice synthetic jet. The power consumption is reduced due to a reduction in the resistance to flow, owing to a larger number of openings as compared to a single center orifice. The reduction in power consumption in conjunction with the larger heat transfer coefficient with multi-orifice synthetic jet suggests that they can be beneficially employed in cooling systems.

Other than employing synthetic jets in impinging mode, studies related to flow in microchannels and server ducts have been carried out, where the synthetic jet flow is parallel to the hot surface. Lee et al. (2012) numerically studied the flow and heat transfer characteristics in a microchannel using the synthetic jet. Cross flow in addition to the synthetic jet flow removes the hot spots from the electronic components. A synthetic jet can thus be effectively used for the local cooling of electronic components of a blade server (Chaudhari et al., 2012). The heat transfer characteristics of an inclined impinging synthetic jet have also been carried out for different angle of inclinations ($\theta=40°$–$90°$). The normal impingements ($\theta=90°$) resulted in the maximum value of the heat transfer coefficient for all orifice to surface spacings (Bhapkar et al., 2013).

From an application point of view, the utility of the synthetic jet has been explored by studying the thermal performance of heat sinks. Also, use of the synthetic jet in the duct (as a blade server) has been explored. It was observed that approximately 40% more heat dissipation occurred with the synthetic jet-based heat sinks as compared to that with a steady flow from a ducted fan blowing air

through the heat sink. The average heat transfer coefficient in the channel flow between the fins was 2.5 times that of the steady flow in the duct at the same Reynolds number. The thermal effectiveness of the synthetic jet heat sinks is more than twice of the fan-heat sink combination at the same power (Mahalingam and Glezer, 2005). The shape of the orifice with the same hydraulic diameter of the synthetic jet has a negligible effect on the thermal performance of the heat sink. A commercial fan that covers the entire surface of heat sink performs better than the single orifice synthetic jet (Chaudhari et al., 2012).

REFERENCES

Agrawal, A., and Verma, G. 2008. Similarity analysis of planar and axisymmetric turbulent synthetic jets, *Int. J. Heat Mass Transfer*, 51, 6194–6198.

Arik, M. 2007. An investigation into feasibility of impingement heat transfer and acoustic abatement of meso scale synthetic jets, *Appl. Therm. Eng.*, 27, 1483–1494.

Bhapkar, U., Srivastava, A., and Agrawal, A. 2013. Acoustic and heat transfer aspects of an inclined impinging synthetic jet, *Int. J. Therm. Sci.*, 74, 145–155.

Bhapkar, U., Srivastava, A. and Agrawal, A. 2014. Acoustic and heat transfer characteristics of an impinging elliptical synthetic jet generated by acoustic actuator, *Int. J. Heat Mass Transfer*, 79, 12–23.

Chaudhari, M., Verma, G., Puranik, B., and Agrawal, A. 2009. Frequency response of a synthetic jet cavity, *Exp. Therm. Fluid Sci.*, 33, 439–448.

Chaudhari, M., Puranik, B., and Agrawal, A. 2010a. Heat transfer characteristics of synthetic jet impingement cooling, *Int. J. Heat Mass Transfer*, 53, 1057–1069.

Chaudhari, M., Puranik, B., and Agrawal, A. 2010b. Effect of orifice shape in synthetic jet based impingement cooling, *Exp. Therm. Fluid Sci.*, 34, 246–256.

Chaudhari, M., Puranik, B., and Agrawal, A. 2011. Multiple orifice synthetic jet for improvement in impingement heat transfer, *Int. J. Heat Mass Transfer*, 54, 2056–2065.

Chaudhari, M., Puranik, B., and Agrawal, A. 2012. Heat transfer characteristics of a heat sink in presence of a synthetic jet, *IEEE Trans. Compon., Packag., Manuf. Technol.*, 2 (3), 457–463.

Gallas, Q., Holman, R., Nishida, T., Carroll, B., Sheplak, M., and Cattafesta, L. 2003. Lumped element modeling of piezoelectric-driven synthetic jet actuators, *AIAA J.*, 41, 240–247.

Holman, R., Utturkar, Y., Mittal, R., Smith, L. B., and Cattafesta, L. 2005. Formation criterion for synthetic jets, *AIAA J.*, 43, 2110–2116.

Jain, M., Puranik, B., and Agrawal, A. 2011. A numerical investigation of effects of cavity and orifice parameters on the characteristics of a synthetic jet flow, *Sens. Actuators, A*, 165, 351–366.

Lee, A., Yeoh, G. H., Timchenko, V., and Reizes, J. A. 2012 Heat transfer enhancement in micro-channel with multiple synthetic jets, *Appl. Therm. Eng.*, 48, 275–288.

Mahalingam, R., and Glezer, A. 2005. Design and thermal characteristics of a synthetic jet ejector heat sink, *J. Electron. Packag.*, 127, 172–177.

Mallinson, S., Reizes, J., and Hong, G. 2001. An experimental and numerical study of synthetic jet flow, *Aeronaut. J.*, 105, 41–49.

Mangate, L. D., and Chaudhari, M. 2015. Heat transfer and acoustic study of impinging synthetic jet using diamond and oval shape orifice, *Int. J. Therm. Sci.*, 89, 100–109.

Pavlova, A., and Amitay, M. 2006. Electronic cooling using synthetic jet impingement, *J. Heat Transfer*, 128, 897–907.

Persoons, T., McGuinn, A., and Murray, D. B. 2011. A general correlation for the stagnation point Nusselt number of an axisymmetric impinging synthetic jet, *Int. J. Heat Mass Transfer*, 54, 3900–3908.

Qayoum, A., Gupta, V., Panigrahi, P., and Muralidhar, K. 2010. Perturbation of a laminar boundary layer by a synthetic jet for heat transfer enhancement, *Int. J. Heat Mass Transfer*, 53, 5032–5057.

Smith, B., and Glezer, A. 1998. The formation and evolution of synthetic jets, *Phys. Fluids*, 10, 2281–2297.

Smith, B., and Swift, G. 2003. A comparison between synthetic jets and continuous jets, *Exp. Fluids*, 34, 467–472.

Zhang, J., and Tan, X. 2007. Experimental study on flow and heat transfer characteristics of synthetic jet driven by piezoelectric actuator, *Sci. China, Ser. E: Technol. Sci.*, 50, 221–229.

4.25 TEMPERATURE AND HEAT TRANSFER MEASUREMENTS

Robert J. Moffat and Tadhg O'Donovan

There are two different kinds of materials to consider with respect to experimental methods: the unit operations of measurement (transducers and their environmental errors) and the strategy of experimentation. This section deals only with the unit operations: transducers, their calibrations, and corrections for environmental errors.

Temperature Measurement

An International Practical Temperature Scale (IPTS) has been defined in terms of a set of fixed points (melting points of pure substances) along with a method for interpolating between the fixed points. The IPTS agrees with the thermodynamic temperature scale within a few Kelvin over most of its range. The IPTS is the basis for all commerce and science, and all calibrations are made with respect to the IPTS temperature. The scale is revised periodically.

Accurate calibrations are not enough to ensure accurate data, however. If a sensor has been installed to measure a gas temperature or a surface temperature, any difference between the sensor temperature and the measurement objective due to heat transfer with the environment of the sensor is an "error." In most temperature-measuring applications, the environmental errors are far larger than the calibration tolerance on the sensor and must be dealt with just as carefully as the calibration.

Thermocouples

Any pair of thermoelectrically dissimilar materials can be used as a thermocouple. The pair need only be joined together at one end and connected to a voltage-measuring instrument at the other to form a usable system. A thermocouple develops its signal in response to the temperature difference from one end of the pair to the other. The temperature at one end, known as the reference junction end, must be known accurately before the temperature at the other end can be deduced from the voltage.

Thermocouples are the most commonly used electrical output sensors for temperature measurement. With different materials for different ranges, thermocouples have been used from cryogenic temperatures (a few Kelvin) to over 3000 K. In the moderate-temperature range, ambient to 1200°C, manufacturer's quoted calibration accuracy can be as good as ±3/8% of reading (referred to 0°C) for precision-grade base metal thermocouples. Broader tolerances apply at very high temperature and very low temperatures. Thermocouple signals are DC voltages in the range from a few microvolts to a few tens of microvolts per degree Centigrade. Because of their low signal levels, thermocouple circuits must be protected from ground loops, galvanic effects, and from pickup due to electrostatic or electromagnetic interactions with their surroundings. Thermocouples are low-impedance devices. Multiple channels of thermocouples can be fed to a single voltage reader using low-noise-level scanners or preamplifiers and electronic multiplexers.

The alloys most frequently used for temperature measurement are listed in Table 4.25.1. These alloys have been developed, over the years, for the linearity, stability, and reproducibility of their EMF vs. temperature characteristics and for their high-temperature capability.

Calibration data for thermocouples are periodically reviewed by the National Institutes of Science and Technology based on the then-current IPTS. Values in Table 4.25.1 illustrate the approximate levels, which can be expected, and are from the National Bureau of Standards Monograph 125. Maximum temperatures listed in this table are estimates consistent with a reasonable service lifetime. Allowable atmosphere refers to the composition in contact with the thermoelements themselves. Accuracy estimates are provided for two levels of precision: standard grade and precision grade, where these data are available.

TABLE 4.25.1
Application Characteristics of Some Common Thermocouple Alloys

Max *T*, °F	Max *T*, °C	Allowable Atmos. (Hot)	Material Names	ANSI Type[a]	Color Code	Output, mV/100°F	Accuracy, % Standard[a]	Accuracy, % Precision[a]
5072	2800	Inert, H_2, vacuum	Tungsten/tungsten 26% rhenium	—	—	0.86	—	—
5000	2760	Inert, H_2, vacuum	Tungsten 5% rhenium/tungsten 26% rhenium	—	—	0.76	—	—
4000	2210	Inert, H_2	Tungsten 3% rhenium/tungsten 35% rhenium	—	—	0.74	—	—
3720	1800	Oxidizing[b]	Platinum 30% rhodium/platinum 6% rhodium	B	—	0.43	1/2	1/4
2900	1600	Oxidizing[b]	Platinum 13% rhodium/platinum	R	—	0.64	1/4	1/4
2800	1540	Oxidizing[b]	Platinum 10% rhodium/platinum	S	—	0.57	1/4	1/4
2372	1300	Oxidizing[b,c]	Platinel II (5355)/ Platinel II (7674)	—	—	2.20	5/8	—
2300	1260	Oxidizing	Chromel/Alumel,[d] Tophel/Nial,[e] Advance T1/T2,[f] Thermo-Kanathal P/N[g]	K	Yellow red	2.20	4°F, or 3/4%	2°F, or 3/8%
1800	980	Reducing[a]	Chromel/constantan	E	Purple red	4.20	1/2	3/8
1600	875	Reducing	Iron/constantan	J	White red	3.00	4°F, or 3/4%	2°F, or 3/8%
750	400	Reducing	Copper/constantan	T	Blue red	2.50	3/4	3/8

[a] Per ANSI C96.1 Standard.
[b] Avoid contact with carbon, hydrogen, metallic vapors, silica, reducing atmosphere.
[c] @ Engelhard Corp.
[d] @ Hoskins Mfg. Co.
[e] Wilber B. Driver Co.
[f] Driver-Harris Co.
[g] The Kanthal Corp.

Noble metal and refractory metal thermocouples are often used with substitute lead wires, as a cost-saving measure. These lead wires, described in Table 4.25.2, are cheaper and easier to handle than the high-temperature thermocouples. They have the same temperature–EMF characteristics as their primary thermoelements, but only over the range of temperatures the lead wires will usually encounter (up to a few hundred degrees Centigrade). Except for the substitute alloys, thermocouple extension wires have the same composition as thermocouple wires, differing only in the type of insulation and the accuracy of calibration, which is not held as closely for extension wire as for thermocouple-grade wire.

Any instrument capable of reading low DC voltages (on the order of millivolts) with 5–10 μV resolution will suffice for temperature measurements. Galvanometric measuring instruments can be used, but, since they draw current, the voltage available at the terminals of the instrument depends not only on the voltage output of the thermocouple loop but also on the resistance of the instrument

TABLE 4.25.2
Substitute Material Extension Wires for Thermocouples

Thermocouple Material	Thermocouple Type[a]	Extension Wire, Type[a]	Color for (+) Wire	Color for (–) Wire	Overall Color
Tungsten/tungsten 26% rhenium	—	Alloys 200/226[b]	—	—	—
Tungsten 5% rhenium/tungsten 26% rhenium	—	Alloys (405/426)[b]	White	Red	Red[b]
Tungsten 3% rhenium/tungsten 25% rhenium	—	Alloys (203/225)[b]	White/yellow	White/red	Yellow/red[b]
Platinum/platinum rhodium	S, R	SX, SR	Black	Red	Green
Platinel II-5355/Platinel II-7674	—	P2X[c]	Yellow	Red	Black[c]
Chromel/Alumel, Tophel/Nial, Advance, Thermokanthal[d]	K	KX	Yellow	Red	Yellow
Chromel/constantan	E	EX	Purple	Red	Purple
Iron/constantan	J	JX	White	Red	Black
Copper/constantan	T	TX	Blue	Red	Blue

[a] ANSI, except where noted otherwise.
[b] Designations affixed by Hoskins Mfg. Co.
[c] Engelhard Mfg. Co.
[d] Registered trade mark names.

and the loop together. Such instruments are normally marked to indicate the external resistance for which they have been calibrated. Potentiometric instruments, either manually balanced or automatically balanced, draw no current when in balance, hence can be used with thermocouple loops of any arbitrary resistance without error. High-input impedance voltmeters draw very low currents and, except for very high resistance circuits, are not affected by the loop resistance.

Thermocouple Theory

Equation 4.25.1 is the general form describing the EMF generated in a two-wire thermocouple (Moffat, 1962). The same form can be derived from either the free-electron theory of metals or from thermodynamic arguments alone: the output of a thermocouple can be described as the sum of a set of terms, one arising in each wire in the circuit.

The junctions do not generate the EMF: they are merely electrical connections between the wires. For a two-wire circuit,

$$\mathrm{EMF} = \int_0^L \varepsilon_1 \frac{dT}{dx}\,dx + \int_L^0 \varepsilon_2 \frac{dT}{dx}\,dx \tag{4.25.1}$$

where

ε_1 and ε_2 = the total thermoelectric power of materials 1 and 2, respectively, mV/°C. The value of ε is equal to the sum of the Thomson coefficient and the temperature derivative of the Peltier coefficient for the material.
T = temperature, °C
x = distance along the wire, m
L = length of the wire, m

This form for expressing the output of a two-wire circuit applies regardless of whether the wires are uniform in composition or not. If a circuit contained four wires (two thermocouple wires and two extension wires), then Equation 4.25.1 would be written with four terms, one for each length of wire.

When the wire is uniform in composition and both wires begin at ($\boldsymbol{T}_0$) and both end at ($\boldsymbol{T}_\mathbf{L}$), the two terms can be collected into one integral:

$$\text{EMF} = \int_{T_0}^{T_\text{L}} (\varepsilon_1 - \varepsilon_2)\,\text{d}T \tag{4.25.2}$$

The EMF–temperature (*E–T*) tables produced by NIST and others are "solutions" to Equation 4.25.2 and can be used only when the following three conditions are met:

1. The thermoelectric power, ε, is not a function of position, i.e., the wires are homogeneous;
2. There are only two wires in the circuit;
3. Each wire begins at $\boldsymbol{T}_\mathbf{0}$ and ends at $\boldsymbol{T}_\mathbf{L}$.

When the circuit consists entirely of pairs of materials, Equation 4.25.2 can be used directly as the basis for understanding the source of the EMF. As an example, consider the three-pair system shown in Figure 4.25.1. For that circuit, Equation 4.25.2 would have three terms: one for each pair. The total EMF generated by the circuit would be the sum of the EMFs generated in the thermocouple pair and in the extension wire pair. The pair of copper wires would not contribute to the net EMF, assuming the two copper wires were perfectly matched. The EMF contributed by each pair would be proportional to the temperature difference from end to end of that pair, as shown in Equations 4.25.3 and 4.25.4.

$$\text{EMF} = \int_{T_1}^{T_2} (\varepsilon_{\text{cn}} - \varepsilon_{\text{cu}})\,\text{d}T + \int_{T_2}^{T_3} (\varepsilon_+ - \varepsilon_-)_{\text{LEADS}}\,\text{d}T + \int_{T_3}^{T_4} (\varepsilon_+ - \varepsilon_-)_{\text{TC}}\,\text{d}T \tag{4.25.3}$$

$$\text{EMF} = 0 + (T_3 - T_2)\,(\varepsilon_+ - \varepsilon_-)_{\text{LEADS}} + (T_4 - T_3)\,(\varepsilon_+ - \varepsilon_-)_{\text{TC}} \tag{4.25.4}$$

Most thermocouple circuits consist only of pairs of wires and can be understood in terms of these two equations, but some require a more detailed treatment. A graphical method of analysis is available, based on Equation 4.25.1.

The temperature–EMF calibrations of the more common materials are shown in Figure 4.25.2 derived from NBS Monograph 125 and other sources. This figure provides the input data for a simple graphical technique for describing the EMF generation in a circuit. Each curve in Figure 4.25.2 represents the output that would be derived from a thermocouple made of material X used with platinum when the cold end is held at 0°C and the hot end is held at $\boldsymbol{T}$.

Those elements commonly used as "first names" for thermocouple pairs, i.e., Chromel (Chromel–Alumel), iron (–constantan), copper (–constantan), have positive slopes in Figure 4.25.2.

The simplest thermocouple circuit for temperature measurement consists of two wires joined together at one end (forming the "measuring junction") with their other ends connected directly to a measuring instrument, as shown in the upper portion of Figure 4.25.3. The EMF generation

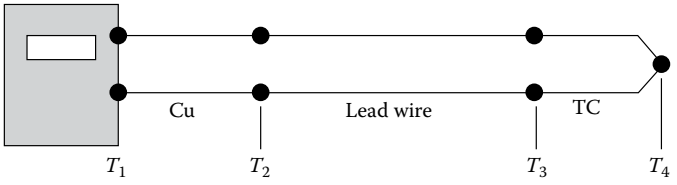


FIGURE 4.25.1 A three-pair circuit.

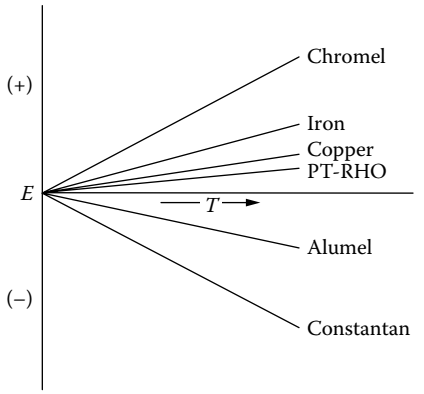


FIGURE 4.25.2 *E–T* calibrations for several common thermocouple materials.

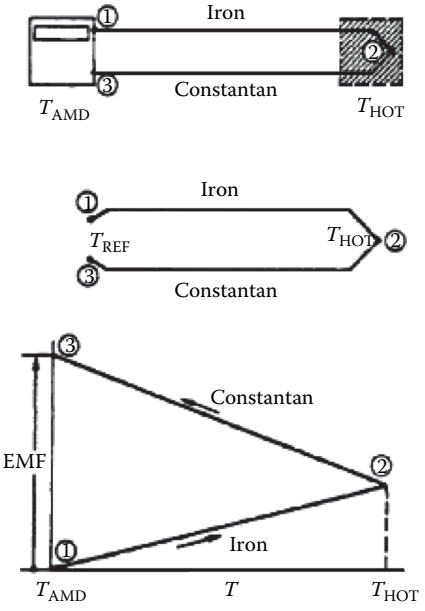


FIGURE 4.25.3 *E–T* diagram of a thermocouple using an ambient reference.

in this circuit is graphically represented in the lower portion, an ***E–T*** diagram, using the data in Figure 4.25.2. The ***E–T*** diagram is used for examining the EMF generated in the circuit, to be certain that it arises only from the desired thermocouple materials, and that all segments of the circuit are properly connected. The ***E–T*** diagram is not used for evaluating the output—that is done using the tables after the circuit has been shown to be correctly wired.

To construct an ***E–T*** diagram, first sketch the physical system considered and assign a number to each "point of interest" on that sketch and a temperature. On ***E–T*** coordinates, locate point 1 at 0/mV and at its assigned temperature. Then start a line from point 1, moving toward the temperature of point 2, and copying the shape of the calibration curve of the iron wire (see Figure 4.25.2). From 2 to 3, use the constantan calibration curve. The difference in elevation of points 1 and 3 describes the net EMF generated in the circuit between points 1 and 3, and describes the polarity. When point 3 lies physically above point 1 in the ***E–T*** diagram, it is, by convention, electrically negative with respect to point 1.

The simple triangular shape shown in Figure 4.25.3 identifies a proper circuit. Any thermocouple circuit whose ***E–T*** diagram is equivalent to that is appropriate for temperature measurement and its EMF may be interpreted using the conventional tables. Any circuit whose ***E–T*** diagram is not equivalent to the pattern circuit should be rewired.

Thermocouples generate their signal in response to the temperature difference between the two ends of the loop. For accurate measurements, the temperature of the "reference junction" must be known. Laboratory users often use an ice bath made from a good-quality Dewar flask or vacuum-insulated bottle of at least 1 pt capacity, as shown in Figure 4.25.4. The flask should be filled with finely crushed ice and then flooded with water to fill the interstices between the ice particles. The reference thermocouple is inserted into a thin-walled glass tube containing a small amount of silicone oil and submerged six or eight diameters into the ice pack. The oil assures good thermal contact between the thermocouple junction and the ice–water mixture. The tube should be sealed at the top to prevent atmospheric moisture from condensing inside it, which would cause corrosion when using iron–constantan thermocouples. Figure 4.25.5 shows in iron–constantan thermocouple circuit with an ice bath. The individual thermocouple wires are connected to copper wires in the ice bath,

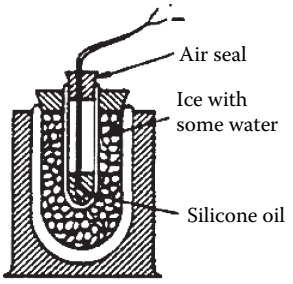


FIGURE 4.25.4 Characteristics of a good ice bath.

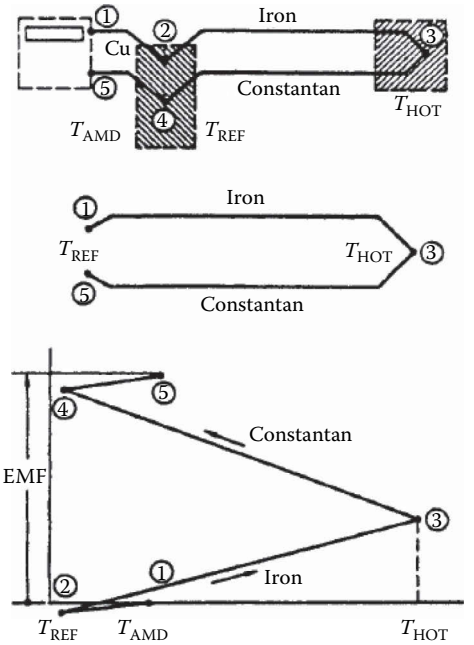


FIGURE 4.25.5 A thermocouple circuit using an ice bath reference and its *E–T* diagram.

and the two copper wires taken to the voltmeter. The lower portion of this figure shows the ***E–T*** diagram for this circuit, and proves that the output of this circuit is entirely due to the temperature difference from end to end of the iron–constantan loop: the two copper wires do not contribute to the output.

Calibration

Thermocouple calibrations are provided by the wire manufacturers to tolerances agreed upon industry-wide, as summarized in Table 4.25.1. These tolerances include two components: the uncertainty in the average slope of the calibration curve of the wire, and the effects of local inhomogeneities in the wire. It is difficult to improve the accuracy of a thermocouple by calibrating it. For a truly significant calibration, the thermocouple would have to be exposed to the same temperature during calibration, at every point along it, as it would encounter in service. In an oven calibration, most of the signal is generated in the material at the mouth of the oven, as could be recognized by considering the temperature gradient distribution along the wire. The material inside the oven contributes little or nothing to the signal.

Thermistors

Thermistors are electrical resistance temperature transducers, made from a semiconducting material or a doped polycrystalline ceramic; their electrical resistance varies inversely, and exponentially, with temperature. The resistance of a 5000 Ω thermistor may go down by 200 Ω for each degree Centigrade increase in temperature in the vicinity of the initial temperature. Interrogated by a 1.0 mA current source, this yields a signal of 200 mV/°C. As a consequence of this large signal, thermistors are frequently used in systems where high sensitivity is required. It is not uncommon to find thermistor data logged to the nearest 0.001°C. This does not mean that the data are accurate to 0.001°C, simply that the data are readable to that precision. A thermistor probe is sensitive to the same environmental errors that afflict any immersion sensor: its accuracy depends on its environment.

Thermistor probes can be used between –183°C (the oxygen point) and +327°C (the lead point), but most applications are between –80°C and +150°C. The sensitivity of a thermistor (i.e., the percent change in resistance per degree Centigrade change in thermistor temperature) varies markedly with temperature, being highest at cryogenic temperatures.

Thermistor probes range in size from 0.25-mm spherical beads (glass covered) to 6-mm-diameter steel-jacketed cylinders. Lead wires are proportionately sized. Disks and pad-mounted sensors are available in a wide range of shapes, usually representing a custom design "gone commercial." Aside from the unmounted spherical probes and the cylindrical probes, there is nothing standard about the probe shapes.

Calibration

Thermistor probes vary in resistance from a few hundred ohms to megaohms. Probe resistance is frequently quoted at 25°C, with no power dissipation in the thermistor. The commercial range is from about 2000 to 30,000 Ω. Representative values of the sensitivity coefficient (% change in resistance per degree Centigrade) are given in Table 4.25.3 and resistance values themselves, in Table 4.25.4.

Proprietary probes are available that "linearize" thermistors by placing them in combination with other resistors to form a circuit whose overall resistance varies linearly with temperature over some range. These compound probes can be summed, differenced, and averaged as can any linear sensor. Modern manufacturing practices allow matched sets to be made, interchangeable within ±0.1°C.

Thermal Characteristics

Thermistor probes are generally interrogated using a low current, either AC or DC. A level of about 10 μA would be typical. With a probe resistance of 10 K Ω, 0.01 W must be dissipated into

TABLE 4.25.3
Thermistor Temperature Coefficient Variations with Temperature

Temperature, °C	Condition	$\Delta R/R$, %
–183	Liquid oxygen	–61.8
–80	Dry ice	–13.4
–40	Frozen mercury	–9.2
0	Ice point	–6.7
25	Room temperature	–5.2
100	Boiling water	–3.6
327	Melting lead	–1.4

TABLE 4.25.4
Thermistor Resistance Variation with Temperature

Temperature, °C	Resistance, Ω	Temperature, °C	Resistance, Ω
–80	1.66 M	0	7355
–40	75.79 K	25	2252
–30	39.86 K	100	152.8
–20	21.87 K	120	87.7
–10	12.46 K	150	41.9

TABLE 4.25.5
Representative Thermal Dissipation Constants for Two Thermistor Probe Designs

Environment	1.0-cm Disk	5.0-cm Cylinder
Still air	8 mW/°C	1 mW/°C
Still oil	55	—
Still water	—	3.5
Oil at 1 m/s	250	—

its surrounding material. This current results in the probe running slightly above the temperature of the medium into which it is installed: the "self-heating" effect. Since thermistors are often used where very small changes in temperature are important, even small amounts of self-heating may be significant.

The self-heating response is discussed in terms of the "dissipation constant" of the probe, in milliwatts per degree Centigrade. The dissipation constant depends on the thermal resistance between the thermistor and its surroundings. For fluid-sensing probes, the self-heating varies with velocity and thermal conductivity, while for solid immersion probes, it varies with the method of attachment and type of substrate.

Dissipation constants for representative probes are given in Table 4.25.5. The self-heating effect must be considered in calibration as well as in use.

The transient response of a thermistor is more complex than that of a thermocouple and, size for size, they are not as well suited to transient measurements.

Thermistor probes are sold with calibration tables of resistance vs. temperature at some specified accuracy, on the order of ±0.1 or 0.2 K, depending on the grade of probe purchased. These tables are

typically in increments of 1 K. For computer interpretation, they should be fit to the Steinhart–Hart form[2] and the coefficients determined for least error.

$$\frac{1}{T} = A_0 + A_1 \ln(R) + A_3 \ln(R^3) \tag{4.25.5}$$

Resistance Temperature Detectors

The terms resistance temperature detector (RTD) and resistance thermometer are used interchangeably to describe temperature sensors containing either a fine wire or a thin film metallic element whose resistance increases with temperature. In use, a small current (AC or DC) is passed through the element, and its resistance measured. The temperature of the element is then deduced from the measured resistance using a calibration equation or table lookup.

RTDs are used both for standards and calibration laboratories and for field service. Field-service probes are generally encased in stainless steel protective tubes with either wire or film elements bonded to sturdy support structures. They are made to take considerable physical abuse. Laboratory standard-grade probes are often enclosed in quartz tubes, with the resistance wire mounted in a strain-free manner on a delicate mandrel.

High-quality resistance thermometers have been used as defining instruments over part of the range of the IPTS. Because of this association with high-precision thermometry, resistance thermometers, in general, have acquired a reputation for high precision and stability. Commercial probes, however, are far different in design from the standards-grade probes, and their stability and precision depend on their design and manufacture.

RTDs are often recommended for single-point measurements in steady-state service at temperatures below 1000°C, where longtime stability and traceable accuracy are required and where reasonably good heat transfer conditions exist between the probe and its environment.

They are not recommended for use in still air, or in low-conductivity environments. RTDs self-heat, which causes an error when the probes are used in a situation with poor heat transfer. They are not recommended for transient service or dynamic temperature measurements unless specifically designed for such service. The probes tend to have complex transient characteristics and are not amenable to simple time-constant compensation.

Physical Characteristics

The physical characteristics of any given resistance thermometer represent a compromise between two opposing sets of requirements. For accuracy, repeatability, and speed of response, a delicate, low-mass sensing element is desired, supported in a strain-free manner in good thermal contact with its surroundings. For durability, a rugged sensor is indicated, mounted firmly to a sturdy structure inside a robust, sealed protection tube.

Both the short-term calibration (resistance vs. specimen temperature) and the long-term stability (drift) are directly affected by the mechanical configuration of the probe. The electrical resistance of the sensing element is a function of its temperature and state of mechanical strain (Figure 4.25.6).

The sensing elements used in field-service RTD probes range from thin metallic films deposited on rectangular ceramic wafers ($0.5 \times 1.0 \times 2.0$ mm) with pigtail leads (0.25 mm diameter and 2.5 cm long) to glass-encapsulated, wire-wound mandrels (4 mm in diameter and 2.0 cm long), again with pigtail leads. Bonding the sensor to its support provides good mechanical protection to the element, but subjects the element to strain due to thermal expansion. As long as this process is repeatable, the calibration is stable.

Electrical Characteristics

RTDs are available commercially with resistances from 20 to 20,000 Ω with 100 Ω being common. Bifilar windings are frequently used in wire-wound elements, to reduce the electrical noise pickup.

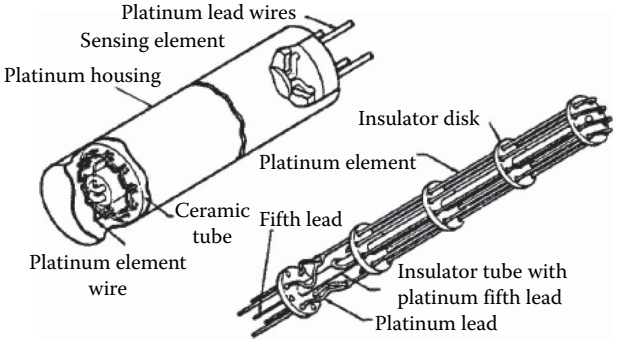


FIGURE 4.25.6 Slack-wire platinum resistance thermometer.

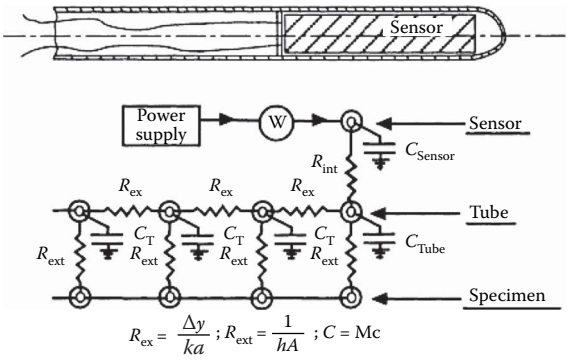


FIGURE 4.25.7 Thermal circuit representation of a typical resistance thermometer.

This is more important in the quartz-jacketed probes than in those with stainless steel protection tubes. Twisted pair lead wires are recommended.

Thermal Characteristics

Figure 4.25.7 shows a simplified cross section of a typical resistance thermometer and a thermal circuit, which can be used to discuss its behavior. In forming such a thermal circuit model, each element of the probe is described by its resistive and capacitive attributes following conventional heat transfer practice. The principal components are

- The external thermal resistance per unit length
- The thermal capacitance of the protective tube per unit length, $\boldsymbol{C_T}$
- The radial internal thermal resistance between the sensor and the protective tube, $\boldsymbol{R_{int}}$
- The capacitance of the sensor element and its support, $\boldsymbol{C_{sensor}}$
- The axial internal thermal resistance of the stem, per unit length, $\boldsymbol{R_T}$

This circuit can be used to predict the temperature distribution within the probe both at steady state and during transients and can be refined, if needed, by subdividing the resistance and capacitance entities.

Steady-State Self-Heating

Interrogating an RTD by passing a current through it dissipates power in the element, shown in Figure 4.25.7 as $\boldsymbol{W}$, which goes off as heat transfer through the internal and external resistances. This self-heating causes the sensing element to stabilize at a temperature higher than its surroundings

and constitutes an "error" if the intent is to measure the surrounding temperature. The amount of the self-heating error depends on three factors:

- The amount of power dissipated in the element
- The internal thermal resistance of the probe, as a consequence of its design
- The external thermal resistance between the surface of the probe and the surrounding material

The self-heating temperature rise is given by Equation 4.25.6:

$$T_{sens} - T_{surr} = W(R_{int} + R_{ext}) \tag{4.25.6}$$

The internal thermal resistance of a probe, $\boldsymbol{R_{int}}$, measured in degree Centigrade per Watt, describes the temperature rise of the sensing element above the surface temperature of the probe, per unit of power dissipated. The internal thermal resistance can be deduced from measurements of the sensor temperature at several different current levels when the probe is maintained in a well-stirred ice bath, where the external thermal resistance is very low. The slope of the apparent temperature vs. power dissipated line, °C/W, is the internal thermal resistance. When an RTD is used in a gas or liquid, the external resistance between the probe and its surroundings must be estimated from standard heat transfer data. The external resistance is $1/(\boldsymbol{hA})$, °C/W.

A typical cylindrical probe exposed to still air will display self-heating errors on the order of 0.1°C–1.0°C per mW (commercial probes of 1.5–5 mm in diameter). At 1 m/s air velocity, the self-heating error is reduced to between 0.03°C and 0.3°C. In water at 1 m/s velocity, the self-heating effect would be reduced by a factor of four or five compared to the values in moving air, depending on the relative importance of the internal and the external thermal resistances.

Calibration and Drift

The relationship between resistance and temperature must be determined for each probe or acquired from the manufacturer. Generally speaking, the reported values will require interpolation.

The resistance–temperature characteristic of a probe may drift (i.e., change with time), while the probe is in service. Manufacturers of laboratory-grade probes will specify the expected drift rate, usually in terms of the expected error in temperature over an interval of time. Two sample specifications are "0.01°C per 100 h" for a low-resistance, high-temperature probe (0.22 Ω at 0°C, 1100°C maximum service temperature) and "0.01°C per year" for a moderate-resistance, moderate-temperature probe (25.5 Ω at 0°C, 250°C maximum service temperature). The drift of the resistance–temperature relationship takes place more rapidly at high temperatures.

Infrared Thermography

Advances in infrared thermal imaging technology, of all the temperature measurement techniques, have been the most significant in recent years. There is now even an add-on device for smart phones to convert the camera to a thermography device. The spatial and temporal resolution of this non-invasive technique has increased to 15 μm and in excess of 500 Hz respectively.

Surface temperatures and gas temperatures can be deduced from radiation measurements. Surface-temperature measurements are based on the emitted infrared energy, while gas-temperature measurements use specific emission lines from the gas itself or from a tracer inserted into the gas. Radiation devices, such as pyrometers for surface-temperature measurement systems (single-point), are commercially available, at low cost, which can measure temperature to ±1.5% of reading or 1.5°C (whichever is greater), above 0°C, if the emissivity of the surface is known. The device referenced requires a spot size of 1.9 cm diameter, viewed from 95 cm. Spectroscopic gas-temperature measurements can be accurate to ±3 or 4% of reading, but require a significant investment in effort as well as equipment (on the order of 1– 2 years and $100,000–$200,000). Several techniques based on Raman scattering have been used in combustion systems. Planar-laser-induced fluorescence has shown considerable promise as one of the newer methods.

Infrared emission from a surface is described by two laws: the Stefan–Boltzmann law describing the total emitted radiation as a function of temperature, and Planck's law describing its distribution as a function of temperature. These laws form the basis for all radiation-based surface-temperature detectors.

Early radiometers focused the total infrared energy on a thermopile bolometer and used the temperature rise across its calibrated heat loss path to measure the incident energy flux. Solid-state photon detectors have replaced thermopile bolometers as the detector of choice. Such a detector will respond to any photon having energy above a certain level (specific to the detector). Since the energy of a photon is inversely proportional to its wavelength, detectors respond to all wavelengths below some value. Modern detectors use band-pass filters to limit the wavelength band of photons admitted to the detector and rely on Planck's law to infer the temperature from the energy flux:

$$E_{b,\lambda} = \frac{C_1 \lambda^{-5}}{e^{C_2/\lambda T} - 1} \tag{4.25.7}$$

where

$E_{b,\lambda}$ = radiated power at the wavelength λ, W/m²
T = temperature, K
$C_1 = 3.743 \times 108$, W μm⁴/m²
$C_2 = 1.4387 \times 104$, μm K

Commercial radiation temperature detectors use different wave bands for different temperature ranges, with different detectors for each band. The emissivity of the surface must be known, as a function of temperature, in the wavelength band used by the detector.

Radiation detectors are vulnerable to interference from four sources: low signal-to-noise ratio at low temperatures (below a few hundred degrees Centigrade); radiation from the surroundings reflecting into the detector (also usually more important at low temperatures); low spatial resolution (also more evident at low temperatures); uncertainty in the emissivity of the surface (at all temperatures); and absorption of radiation into water vapor and CO_2 in the line of sight (at any temperature).

A fiber-optic blackbody temperature detector system is offered by the Luxtron Corporation for standards room and field service above 300°C. The unit consists of a blackbody capsule fiber-optically coupled to a filtered, band-limited photon detector. The accuracy of 0.01°C–0.1°C is claimed, depending on temperature level.

A fluoroptic temperature-measuring system is also offered by the same company, for use only at lower temperatures (–200°C to +450°C). This system uses an ultraviolet-stimulated phosphor on the end of an optical fiber as its sensor. The fluorescent signal from the phosphor decays with time, and its "time constant" is a function of temperature. The accuracy of ±0.5°C is claimed for measurements within ±50°C of a calibration point, or ±1°C within 100°C.

Temperature-Sensitive Paints, Crayons, and Badges

Temperature-sensitive paints, crayons, and badges are available from several suppliers (Omega Engineering, Inc., Stamford, CT, and others in Germany and Japan). Each undergoes an irreversible change (e.g., a change in color or a change from solid to liquid) at one specified temperature. With a range of paints, temperatures from ambient to about 1500°C can be covered. The accuracy generally quoted is about ±1% of level, although melting standards are available to ±0.5°C.

The phase-change materials melt at well-defined temperatures, yielding easily discernible evidence that their event temperature has been exceeded. When more than one phase-change paint is applied to the same specimen, there can be interference if the melt from the low-melting paint touches the high-melting material. Color change materials do not interfere, but are more difficult to interpret. The calibration of high-temperature paints (both phase change and color change) may shift when they are used on heavily oxidized materials, due to alloying of the oxide with the paint. Recommended practice is to calibrate the paints on specimens of the application material. The event

temperature that will cause transformation depends on the time at temperature: short exposure to a high temperature often has the same effect as long exposure to a lower temperature.

The paints and crayons are nonmetallic and, therefore, tend to have higher emissivities for thermal radiation than metals. They should be used only over small areas of metallic surfaces, compared with the metal thickness, or else their different emissivities may lead to a shift in the operating temperature of the parts.

The principal disadvantages of the paints and crayons are that they require visual interpretation, which can be highly subjective, and they are one-shot, irreversible indicators that respond to the highest temperature encountered during the test cycle. They cannot record whether the peak was reached during normal operation or during soak-back.

Liquid crystals can be divided into three groups, depending on their molecular arrangements: (1) smectic, (2) nematic, and (3) cholesteric. Most of the temperature-sensitive liquid crystals now in use are cholesteric: made from esters of cholesterol. Their molecules are arranged in planar layers of molecules with their long axes parallel and in the plane of the layer. The molecules in each layer are rotated with respect to those in its neighboring layers by about 15 min of arc in a continuous, helical pattern along an axis normal to the layers.

The colors reflected from cholesteric liquid crystals are thought to be due to Bragg diffraction from the aligned layers. The "wrap angle" between adjacent layers increases with temperature; hence, the color of the liquid crystal shifts toward short wavelengths (toward blue) as the temperature is raised. The color can also be affected by electric fields, magnetic fields, pressure, shear stress, and some chemical vapors.

Warm cholesterics are colorless liquids and they pass through a series of bright colors as they are heated through their "color-play" temperature band. The first color to appear is a deep red, followed by yellow, green, blue, and violet. Further heating yields a colorless liquid again. This cycle is reversible and repeatable, and the color–temperature relationship can be calibrated.

Liquid crystals selectively reflect only a small fraction of the incident light; hence, to enhance the brightness of the color image, they must be backed up with black paint or a nonreflecting surface.

A typical calibration is shown in Figure 4.25.8 for liquid crystals painted over black paint on an aluminum calibration strip. The upper part of Figure 4.25.8 describes the color variation, while the lower part shows the imposed linear temperature distribution. The hot end is blue, the cold end is red. Color-play intervals range from 0.5°C to 10.0°C. Liquid crystals whose color-play interval is on the order of 0.5°C–2.0°C are often referred to as narrow-band materials, while those whose interval extends to 5.0°C–10°C are called wide band. Narrow-band images are easy to interpret by eye. Wide-band images show only subtle variations of color for small changes in temperature, and accurate work requires digital image handling or multiple images taken with different filters.

Several different narrow-band liquid crystals can be mixed together to make a single, multi-event paint covering a wide range of temperatures, provided their color-play intervals do not overlap. Such a mixture yields a set of color-play bands, one for each component.

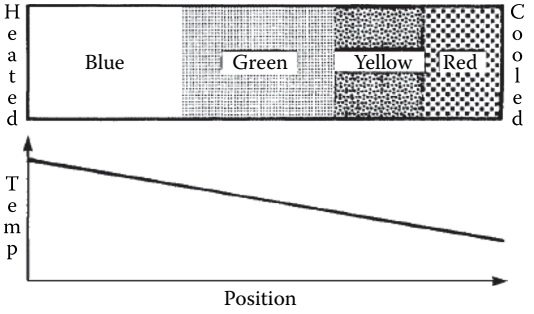


FIGURE 4.25.8 Schematic representation of a calibration strip.

Calibration

Liquid crystals are sold by event temperature and color-play bandwidth, with a nominal accuracy of ±1°C on the event temperature. In many applications, especially if the image is to be visually interpreted, no further calibration is needed.

The accuracy attainable with a liquid crystal is related to the width of the color-play interval. With narrow-band material (a color-play interval of about 1.0°C), visual interpretation can be done with an uncertainty of 0.25°C–0.5°C. With digital image interpretation, spectrally controlled lighting and appropriate corrections for reflected light interference, the uncertainty can be held below 0.25°C.

Early users reported that the perceived color of a liquid crystal depended on both the lighting angle and the viewing angle. This dependence can be eliminated by using a light source along the line of sight (coaxial viewing and illumination).

Multiple-Event Paints

Several narrow-band paints can be mixed together to make a single paint with all the characteristics of each component, if their color-play intervals do not overlap. Each component retains its original calibration and acts independently of the other components.

Figure 4.25.9 shows the image from a five-event paint used to map the adiabatic wall temperature isotherms around a heated block in mixed convection. The outermost isotherm is 30°C, and the events are spaced apart at 5°C intervals. Determination of the temperatures from a multiple-event image requires that the temperature be known at one point in the image.

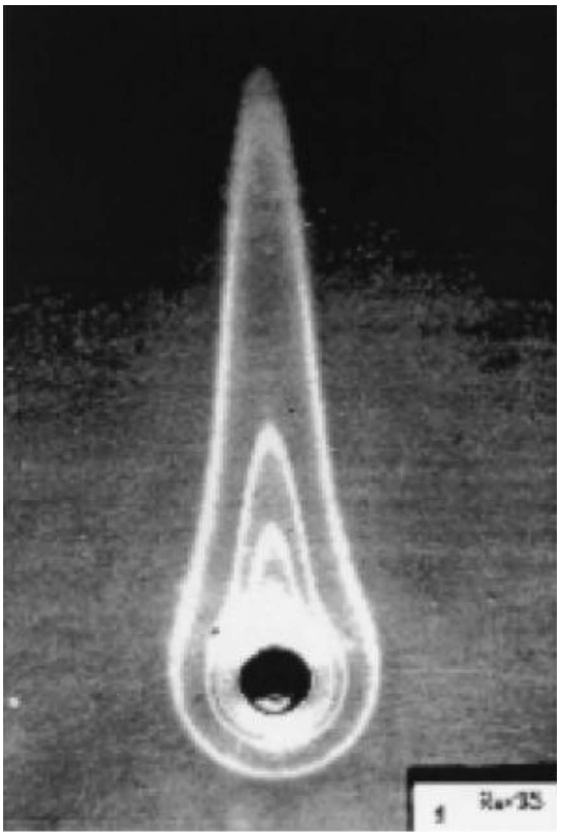


FIGURE 4.25.9 Multi-event liquid crystal used to visualize the isotherm pattern above a heated spot in mixed convection.

Liquid Crystals in Water

Liquid crystals can be used to mark the temperature distribution in water and some other liquids by adding a small quantity of encapsulated liquid crystal material to the liquid and photographing the color distribution using planar lighting. Velocity and temperature distributions can be determined by photographing the liquid crystal particles using a known exposure time. The temperature is deduced from the particle color, and the velocity by the length of the streak its image forms. Figure 4.25.10 shows the velocity and temperature distributions in a shear-driven, water-filled cavity 30 s after the impulsive start of belt motion. In this view, the belt is at the top of the image, and moved from left to right. The water was stably stratified initially, with the top being 4°C hotter than the bottom. This technique was demonstrated by Rhee et al. (1984) and has been used by several workers.

Image Processing

Several schemes have been proposed to remove the subjectivity from the interpretation of liquid crystal images. Akino et al. (1989), and others have processed RGB video images of narrow-band images using multiple filters to extract images of specified isochromes, related to temperatures through a calibration. Hollingsworth et al. (1989) processed RGB images of wide-band images using chromaticity coordinates (hue, saturation, and intensity) and extracted temperature at each pixel, rather than along isochromes.

Heat Flux Measurement

Heat flux to or from a surface can be measured directly, using heat flux meters, or inferred from an overall energy balance, or inferred from temperature–time measurements at the surface or within the body. There are no primary standards for heat flux measurement.

Three general classes of heat flux meters are in common use: slug calorimeters, planar heat flux gauges (sometimes called Schmidt–Boelter gauges), and circular foil gauges (sometimes called Gardon gauges). Sensitivities range from microvolts per kW/m^2 to millivolts per W/m^2. Planar gauges can be used for radiant or convective heat loads. Circular foil gauges should be used only for radiant loads.

Thus far both temperature and surface heat flux measurement techniques described have largely been limited to steady-state or low-speed transient measurements. To advance our understanding

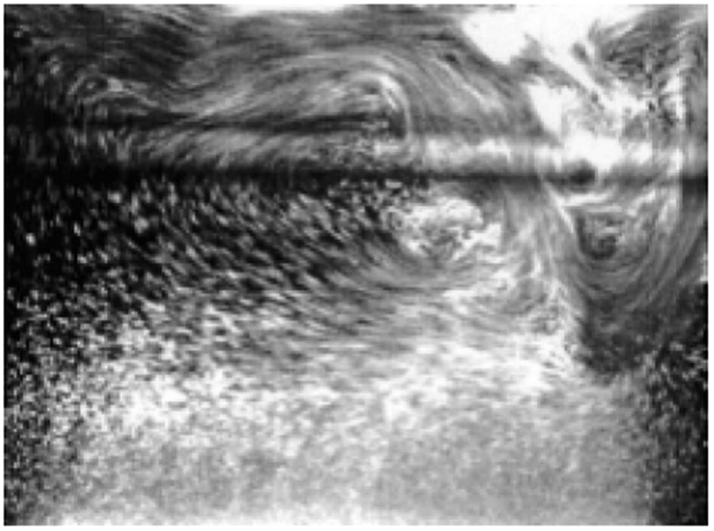


FIGURE 4.25.10 Liquid crystal visualization of the velocity and temperature distribution in a water-filled tank.

of convective heat transfer in many applications, it is important to shed new light on the complex coupling of surface heat flux and local fluid velocity. While several fluid flow measurement techniques, such as laser Doppler velocimetry (LDV), particle image velocimetry (PIV), and hot-wire anemometry (HWA) can resolve velocities at rates of up to 100 kHz, temperature and heat flux measurements are more limited; fine-wire thermocouples have reported maximum frequency response in the region of 330 Hz. This section will also describe the current state-of-the-art and potential of techniques such as infrared thermal imaging and hot-film anemometry to improve the temporal resolution of temperature and surface heat flux measurement techniques.

Slug Calorimeter

The slug calorimeter is an energy balance transducer consisting of a known mass of material instrumented so that its temperature can be measured. A simple version is shown in Figure 4.25.11. If losses are negligibly small and the mass and the specific heat are constant, the instantaneous heat flux is deduced from

$$\dot{q}''_{\text{in}} A = Mc \frac{\partial T}{\partial t} \tag{4.25.8}$$

where

T = average temperature of the slug, °C
M = mass of the slug, kg
c = specific heat, J/kg-°C
A = face area, m^2
τ = time

The variation of slug temperature with time is used to infer net heat transfer rate to the gauge. Slug calorimeters are used mainly when the heat flux, or the heat transfer coefficient, is expected to be relatively constant. They are of less value when the input flux changes arbitrarily because of the inaccuracies inherent in differentiating the signals.

Planar Heat Flux Gauge

Planar heat flux gauges use Fourier's law to deduce the heat flux from a steady-state measurement of the temperature difference across a thin sheet of thermally insulating material. The planar gauge geometry is shown in Figure 4.25.12. The working equation for a planar gauge is

$$\text{EMF} = N\varepsilon\Delta T = \frac{N\varepsilon t}{k}\dot{q}'' \tag{4.25.9}$$

where

N = number of junction pairs
ε = thermoelectric power of the thermoelement, mV/°C
t = thickness of the insulator, m
k = conductivity of the insulator, W/m·°C
$\dot{q}''$ = heat flux through the gauge, W/m^2

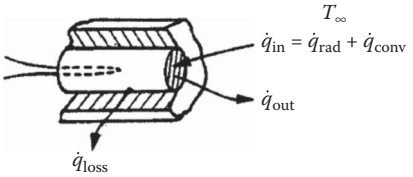


FIGURE 4.25.11 A simple slug calorimeter.

FIGURE 4.25.12 A typical planar heat flux gauge.

The figure shows one thermocouple junction on the top and one on the bottom surface of the insulator. Most gauges use multiple junctions (thermopile). The thermoelements may be wire (down to 0.025 mm diameter) or thin films deposited on the insulator (10–20 Å). The assembly is usually sandwiched between two sheets of protective material to form an integral unit. Up to 150°C application temperature, these units are often made of Kapton, and provided with a contact adhesive. They may be as thin as 0.15 mm overall.

Gauges should not be removed and reinstalled without recalibration, as the act of removing them from the surface may delaminate the gauge, changing its thermal resistance, and therefore its calibration.

Hot-Film Anemometry

While this technique is not currently commercially available, it has been used by researchers since at least the 1980s, by adapting set-ups similar to those employed for hot wire anemometry fluid velocity measurement. A hot-film sensor typically consists of a nickel sensor element on a thin polyimide film and is should be flush-mounted on a heated surface. The sensor element forms one arm of a Wheatstone bridge and as the electrical resistance of the film increases with temperature, a decade resistor is used to set and maintain the film at, or above the temperature of the surrounding heated surface. The power or heat dissipated from the surface can then be measured by monitoring the temperature-dependent resistance of the sensor and the voltage across the sensor:

$$q_{\text{dissipated}} = \frac{R_{\text{film}}}{(R_1 + R_{\text{probe}})^2} \times E^2 \tag{4.25.10}$$

where

R_1 = bridge resistor in line with probe
$R_{\text{probe}} = R_{\text{film}} + R_{\text{sensor\_leads}}$
E = voltage

The setup can respond at frequencies up to 100 kHz but the accuracy of the technique is limited by a number of factors. Firstly, to achieve a measureable signal, it is important for the sensor to operate at a temperature in excess of the surrounding surface. This introduces an error to the technique, which is proportional to the "overheat"; typical "overheat" values range from 1°C to 20°C. There are two proposed correction techniques for this error. The first relies on a calculation of the conduction losses across the substrate and adhesive layer but accuracy is compromised by the ability to measure the temperature locally and to estimate the distance across which the temperature difference occurs. The second considers the sensor signal under adiabatic test conditions, but as these tests are not simultaneous, it is thereafter difficult to ascertain the time-resolved surface heat flux.

A further aspect to be considered is the "effective surface area" of the film element. Lateral conduction from the sensor element heats the substrate locally, increasing the heated area above that of the element. The effective surface area of the sensor therefore is the equivalent area of the sensor at the elevated overheat temperature; this can range up to a factor of 10 times the geometric surface area, depending on the overheat and convective heat flux range. Studies by Scholten and Murrary (1996), and O'Donovan et al. (2011) have shown that these combined effects can lead to a significant uncertainty in the time-average heat flux measurement (up to 25%) but with some optimization of sensor overheat, uncertainties of less than 6% have been reported.

Liu and Sullivan (1996), Scholten and Murray (1998), and O'Donovan and Murray (2007) have successfully demonstrated that the technique can resolve surface heat flux measurements at high

temporal resolution. These studies have reported on the time-varying effects of vortices shedding from cylinders in cross-flow and impinging from air jets on the boundary layer. The correction techniques presented are based on time-averaged and steady-state measurements and the influence on the magnitude of the local heat flux measurements is, as yet, unresolved. Further experimental and numerical modeling strategies are required to perfect this technique and develop it for commercial readiness.

Calibration

Calibration of the hot-film sensor for heat flux applications is done by reference to a known/established point, such a front stagnation of a heated cylinder in crossflow. Calibration requires both a heated and adiabatic test for each overheat considered to ascertain the effective surface area of the sensor.

Circular Foil Gauges

A circular foil gauge consists of a thin circular disk of metal supported by its edge from a structure of constant and uniform temperature. The circular foil gauge is often called a Gardon gauge. A constantan foil is often used, with a copper support structure. Two copper wires complete the circuit: one attached to the center of the foil disk and one to the support structure. The copper–constantan thermocouple thus formed produces an EMF determined by the temperature difference from the center of the foil disk to its rim. That temperature difference is directly proportional to the average heat flux on the disk. A cross-sectional view of a circular foil gauge is shown in Figure 4.25.13.

The working equation for a circular foil gauge is

$$\text{EMF} = \varepsilon \frac{R^2}{4kt} \dot{q}'' \tag{4.25.11}$$

where

ε = thermoelectric power, mV/°C
R = radius of the disk, m
k = thermal conductivity of the disk, W/m·°C
t = thickness of the disk, m
q'' = heat flux absorbed by the disk, W/m^2 (must be uniform)

The output signal is thus directly proportional to the heat flux on the disk. Cooling passages are frequently built into the support structure to maintain the edge of the disk (the heat sink for the foil disk) at a constant temperature.

Calibration

Calibration of the Gardon-type heat flux meters is most easily done by comparison, using a radiation calibrator.

Planar gauges can be calibrated either by conduction or radiation, but the results will depend on the calibration method for some gauges.

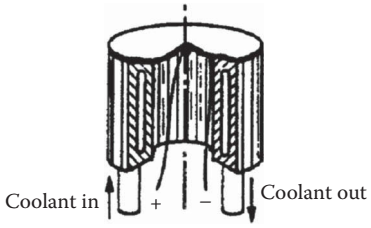


FIGURE 4.25.13 A water-cooled circular foil gauge (Gardon gauge).

Sensor Environmental Errors

Temperature sensors generate signals in response to their own temperatures, but are usually installed to measure the temperature of some fluid or solid. There is heat transfer between the sensor and all of its surroundings, with the result that the sensor usually equilibrates at some temperature different from the fluid or solid it is installed in. This difference is considered an error in the measurement.

Similarly, heat flux gauges are generally installed so one can infer the heat flux, which would have been there had the gauge not altered the system behavior. But heat flux gauges do disturb the system, and the heat flux at the gauge location, when the gauge is there, may be significantly different from that which would have been there without the gauge. This system disturbance effect must also be considered an error.

Steady-State Errors in Gas-Temperature Measurement

All immersion-type temperature sensors (thermocouples, resistance detectors, and thermistors) are subject to the same environmental errors, which are frequently larger than the calibration errors of the sensors. Large probes are usually affected more than small ones; hence, RTDs and thermistors (selected by investigators who wish to claim high accuracy for their data) are more vulnerable to environmental errors (due to their larger size and their self-heating errors). This aspect of accuracy is sometimes overlooked.

Sensor installations for gas-temperature measurements should be checked for all three of the usual steady-state environmental errors: velocity error, radiation error, and conduction error. The same equations apply to all sensors, with appropriate dimensions and constants.

$$\text{Velocity error:} \quad E_v = (1-\alpha)\frac{V^2}{2g_c J c_p} \tag{4.25.12}$$

$$\text{Radiation error:} \quad E_r = \frac{\sigma\varepsilon}{h}(T_{sens}^4 - T_{surr}^4) \tag{4.25.13}$$

$$\text{Conduction error:} \quad E_c = \frac{T_{gas} - T_{mount}}{\cosh\left[L\sqrt{\dfrac{hA_c}{kA_k}}\right]} \tag{4.25.14}$$

where

E_v = velocity error, °
α = recovery factor, —
V = velocity, ft/s
g_c = universal gravitational constant
J = Joules constant, Btu/ft²
c_p = specific heat, Btu/lbm, °F
and E_r = radiation error, °R
σ = Stefan–Boltzmann constant
ε = emissivity
h = heat transfer coefficient, Btu/s ft², °F
T_{sens} = indicated temperature, °R
T_{surr} = surrounding temperature, °R
E_c = conduction error, °R
T_{gas} = gas temperature, °R
T_{mount} = mount temperature, °R
L = length of exposed junction, ft
h = heat transfer coefficient, Btu/s ft², °F
A_c = heat transfer area, ft²

k = thermal conductivity, Btu/s ft
A_k = conduction area, ft^2

Velocity error depends upon the recovery factor, which varies with the Prandtl number of the fluid. The Prandtl numbers of most liquids are greater than 1; hence, the recovery factor α is greater than 1 and probes tend to read higher than the stagnation temperature in high-speed liquid flows. With thermistors and RTDs in liquids, the self-heating effect and the velocity error both tend to cause high readings. In gases, where the Prandtl number is less than 1, the two effects are of opposite sign and may partly cancel each other.

Radiation and conduction errors vary inversely with the heat transfer coefficient. Both tend to be larger for larger-diameter probes since, all other factors remaining the same, the heat transfer coefficient will be lower for a large-diameter probe. This results in larger radiation and conduction errors. In liquids, radiation error is not a problem, but velocity error and conduction error may both be significant. Conduction error becomes a problem in liquid-temperature measurements when thermowells are used. The depth of immersion of the well is frequently too short to eliminate conduction error.

Steady-State Errors in Solid and Surface-Temperature Measurements

When probes are used to measure solid temperature by inserting them into a hole in the specimen, they are subject to conduction errors proportional to their size and conductivity. A general rule of thumb is to keep the insertion depth at least 20 times the diameter (or wall thickness) of the probe. This assumes a close-fitting hole, backfilled with a material with higher thermal conductivity than air. For more-exact advice regarding a specific installation, a careful thermal circuit analysis of the installation should be developed, and its results used to guide the selection of diametrical clearance, backfill materials, and penetration depth.

A thermocouple attached to a hot surface surrounded by cooler fluid will exchange heat with the fluid by convection and with the surrounding solids by radiation. Heat lost from the thermocouple must be made up from the surface by conduction, which will result in a cold spot at the point of attachment.

Figure 4.25.14 shows the system disturbance error caused by a surface-attached thermocouple, as a fraction of the maximum possible error for the installation. If the surface is irradiated (e.g., by heating lamps), the irradiation will raise the surface temperature, but will also affect the system disturbance error. The effect on the system disturbance error caused by turning on the irradiation is similar to that of raising the temperature of the surrounding fluid to a new value, T_∞°,

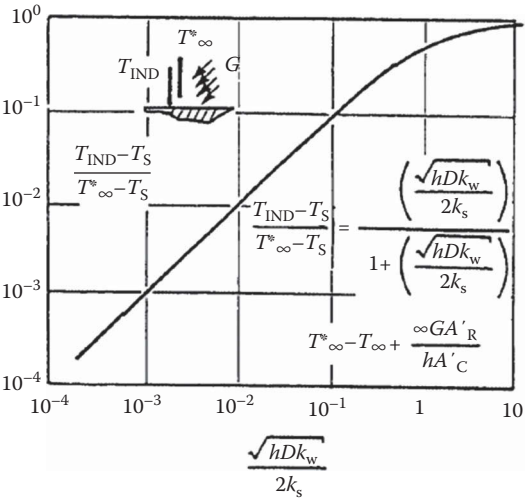


FIGURE 4.25.14 System disturbance errors caused by an attached thermocouple (worst case).

where

T_{ind} = indicated temperature from an otherwise error-free thermocouple, °C
T_s = undisturbed substrate temperature, °C
T_∞° = effective fluid temperature, °C
h = heat transfer coefficient, TC to fluid, W/m, °C
D = outside diameter of TC, m
k_w = effective thermal conductivity of TC, W/m, °C
k_s = thermal conductivity of the substrate, W/m, °C

The effective gas temperature is defined in terms of the incident irradiation and the heat transfer coefficient as

$$T_\infty^\circ = T_\infty + \frac{\alpha G A_R}{h A_C} T_\infty^* \tag{4.25.15}$$

where

T_∞ = actual gas temperature, °C
α = absorptivity of the TC for thermal radiation
G = incident thermal radiation flux, W/m^2
A_R/A_c = ratio of irradiated surface to convective surface
h = heat transfer coefficient between the TC and the gas, W/m °C

Steady-State Errors in Heat Flux Gauges for Convective Heat Transfer

If the gauge is not flush with the surface, it may disturb the flow, and if it is not at the same temperature as the surface, it will disturb the heat transfer. Thus, the gauge may properly report the heat flux, which is present when the gauge is present, but that may be significantly different from the heat flux that would have been there if the gauge had not been there.

For planar gauges, both effects are usually small. The thermal resistance of such a gauge is generally small, and they are thin enough to avoid disturbing most flows. Circular foil gauges pose a more serious problem, since they are often cooled significantly below the temperature of the surrounding surface. Dropping the wall temperature at the gauge location can significantly increase the local heat load in two ways: one due to the fact that, for a given value of h, a cold spot receives a higher heat load from the gas stream. The second effect arises because the value of the heat transfer coefficient itself depends on the local wall temperature distribution: a local cold spot under a hot gas flow will experience a higher heat transfer coefficient than would have existed had the surface been of uniform temperature.

Evaluating the Heat Transfer Coefficient

The heat transfer coefficient is a defined quantity, given by

$$h = \frac{\dot{q}''_{conv}}{(T_o - T_{ref})} \tag{4.25.16}$$

where

h = heat transfer coefficient, W/m^2, °C
$\dot{q}''_{conv}$ = convective heat flux, W/m^2
T_o = temperature of the considered surface, °C
T_{ref} = temperature used as reference for this definition, °C

Different reference temperatures are conventionally used for different situations:

- T_∞: The free-stream temperature. Used for isolated objects of uniform temperature in a uniform free stream, where an average value of h is desired, which describes the overall heat transfer between the object and the flow. Also used in boundary layer heat transfer analyses where local values of h are needed to deal with locally varying conditions.

- T_m: The mixed mean fluid temperature. Used for internal flows where the intent of the calculation is to describe the changes in mixed mean fluid temperature (e.g., heat exchangers).
- $T_{adiabatic}$: The adiabatic surface temperature. Used for isolated objects or small regions of uniform temperature in either internal or external flows, where the overall thermal boundary conditions are neither uniform heat flux nor uniform temperature.

For a given data set, the value of the heat transfer coefficient will depend on the reference temperature chosen, and h should be subscripted to inform later users which reference was used: e.g., h_∞, h_m, or $h_{adiabatic}$.

Direct Methods

The two most commonly used methods for measuring the heat transfer coefficient are both derived from the same energy balance equation:

$$hA(T_o - T_{ref}) = \dot{e}_{in} + \dot{q}_{cond,in} + \dot{q}_{rad,in} - Mc\frac{dT}{d\tau} \tag{4.25.17}$$

where

h = the heat transfer coefficient, W/m, °C
A = the area available for convective transport, m^2
T_{ref} = the reference temperature used in defining h, °C
T_o = the average surface temperature over the area A, °C
$\dot{e}_{in}$ = externally provided input, W
$\dot{q}_{cond,in}$ = net energy conducted in, W
$\dot{q}_{rad,in}$ = net energy radiated in, W
$Mc\ dT/d\tau$ = rate of increase of thermal energy stored within the system, W

Steady State

In the steady-state method, the transient term is zero (or nearly so), and h is determined by measuring the input power and the operating temperature, and correcting for losses. Equation 4.25.17 can be applied to differentially small elements or to whole specimens. The considered region must be reasonably uniform in temperature, so the energy storage term and the convective heat transfer term use the same value.

For tests of isolated objects, or embedded calorimeter sections, steady-state tests usually use high-conductivity specimens (e.g., copper or aluminum) with embedded electric heaters. The resulting value of h is the average over the area of the specimen. The Biot number, hL/k, for the specimen should be low (on the order of 0.01 or less) if only one temperature sensor is used in the specimen, so the surface temperature can be determined from the embedded sensor.

If a single heated element is used within an array of unheated elements, the resulting heat transfer coefficient is implicitly defined as $h_{adiabatic}$ and should be identified as such. Heat transfer coefficients measured with single-active-element tests cannot be used with the mixed mean fluid temperature.

When the variation of h over a surface is required, one common steady-state technique is to stretch a thin foil (stainless steel, or carbon impregnated paper, or gold deposited on polycarbonate) over an insulating substrate, and electrically heat the foil surface. Liquid crystals or infrared techniques can be used to map the surface temperature, from which the heat transfer coefficient distribution can be determined. The "heated foil with liquid crystal" approach was used by Cooper et al. in 1975 to measure heat transfer coefficients, and has since been used by many others. Hippensteele et al. (1985) have made extensive use of the foil technique in studies of gas turbine heat transfer. An example of their work on the end wall of a turbine cascade is shown in Figure 4.25.15. Hollingsworth et al. (1989) used a stainless steel foil heater for a study in air for an electronics cooling application, illustrated in Figure 4.25.16.

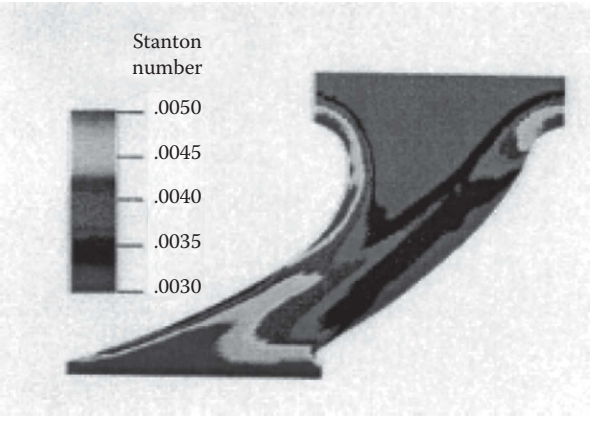


FIGURE 4.25.15 Heat transfer coefficient distribution on the end wall of a turbine cascade. (From Hippensteele, S.A. et al., NASA Technical Memorandum 86900, March, 1985. With permission.)

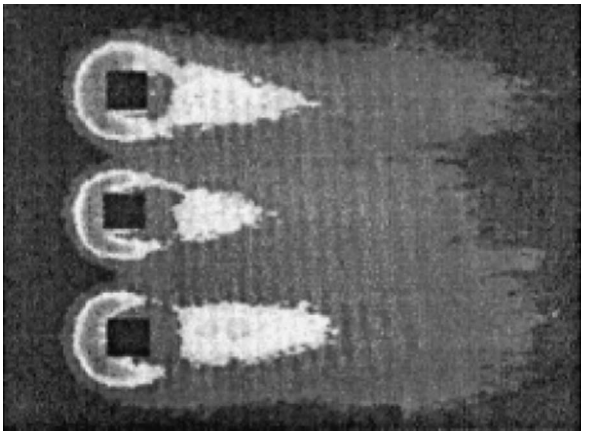


FIGURE 4.25.16 Visualization of the heat transfer coefficient distribution on a heated plate around three unheated cubes.

Advances in imaging and thermal imaging technology now enable measurements in excess of 500 Hz. The limiting factor for the response time of the experimental setup therefore is more likely to be the thickness and material properties (thermal conductivity and capacitance) of the foil. In order to resolve the time-varying heat transfer coefficient, it is explained by Golobic et al. (2009) that the time constant ($\delta^2 \rho c_p / k$) for the foil (including paint/coating) should be much less than the required measurement timescale. A finite difference that considers an energy balance across each pixel (foil volume) can resolve the time-varying heat transfer coefficient at high temporal resolution.

Another steady-state technique that reveals the distribution of h on the surface was introduced by den Ouden and Hoogendoorn (1974). It used a uniform and constant-temperature substrate (a tank of warm water or a copper block) covered with a layer of known thermal resistance (a plate of glass or a thin layer of epoxy). The surface was painted with liquid crystals (or visualized using infrared imaging) and the surface-temperature distribution determined. The inner (uniform) and outer (measured) temperature distributions are then used as boundary conditions to a three-dimensional conduction solver that calculates the total heat flux at each point on the surface. The total heat flux is corrected for radiation to yield the net convective transport at each point, from which h can be determined.

This method appears to have an advantage in accuracy over the heated foil technique because of the more accurate handling of substrate conduction.

Transient Lumped Parameter Systems

In the lumped parameter transient method, the specimen is assumed to be uniform in temperature at every instant through the transient. The power, $\dot{e}_{in}$, in Equation 4.25.17 is usually zero, although that is not necessary (one could simply change the power level at time zero to initiate the transient). At time zero, a transient is initiated, and the response curve recorded.

The data can be interpreted, and the validity of the first-order assumption tested at the same time by plotting $(T-T_{final})/(T_{initial}-T_{final})$ on the log scale of semilog coordinates, with time on the algebraic scale. If the line is straight, then the system is first order and the characteristic time can be determined from any two points on the line by

$$\tau = \frac{(t_2 - t_1)}{\ln\left\{\dfrac{T_{fin} - T_1}{T_{fin} - T_2}\right\}} \tag{4.25.18}$$

where

τ = characteristic time, Mc/hA, s
t_1 = time at the first instant
t_2 = time at the second instant
T_1 = specimen temperature at the first instant, °C
T_2 = specimen temperature at the second instant, °C
T_{fin} = specimen temperature after a long time (fluid temperature), °C

The heat transfer coefficient is extracted from the time-constant definition.

Indirect Methods

An increasingly popular method is the extraction of h from surface-temperature variations after a step in flow temperature using an inverse calculation method. The simplest inverse method assumes one-dimensional conduction into an infinitely thick plate of constant material properties. Even highly irregular geometries can be studied with this technique, if the streamwise extent of the specimen is small and the testing time is short. A short time interval is necessary so the penetration of the thermal wave is limited to a thin layer near the surface. The short streamwise extent is necessary so the temperature response of the surface upstream does not alter the thermal step applied to the downstream surface. This technique has been used to determine the heat transfer coefficient distribution on the inside walls of passages of irregular shape, by making the passage in a transparent material.

Naphthalene Sublimation

The equations for mass diffusion are similar to those for heat transfer, except for replacing the Prandtl number in the heat transfer equation by the Schmidt number in the diffusion equation. Thus, one could expect that the distribution of the mass transfer coefficients on a surface would mimic the distribution of the heat transfer coefficients.

The most commonly used analog technique is naphthalene sublimation. As early as 1940, the mass transfer/heat transfer similarity was used to estimate the heat transfer coefficient distribution. Naphthalene is a solid material that sublimes at a reasonable rate in air at ambient temperature. Specimens can be cast in naphthalene with good precision, and the recession of the surface mapped as a function of position and time using automated or semiautomated measuring equipment. The surface recession over a known interval of time is a measure of the mass transfer rate, from which the mass transfer coefficient can be deduced.

Naphthalene experiments are generally done at a uniform temperature; hence, a uniform vapor pressure exists at the surface. This corresponds to the heat transfer situation of heat transfer from a uniform temperature surface. No counterpart of the uniform heat flux situation has been produced using naphthalene, nor have there been experiments corresponding to variable wall temperature.

Naphthalene sublimation experiments do not suffer from any counterpart of the conduction heat transfer in the substrate. Conduction makes it difficult to work near discontinuities in wall temperature in a heat transfer experiment. Details of the fine structure of mass transfer near obstructions and discontinuities can be resolved in naphthalene experiments, but those details might not exist in a heat transfer process. The Prandtl number of air is much lower than the Schmidt number of naphthalene diffusing in the air; hence, thermal conduction would tend to blur out sharp gradients in the temperature field more than diffusion would blur out gradients in naphthalene concentration.

The Schmidt number of naphthalene in air is about 2.5, far different than the Prandtl number of air (0.71); hence, the mass transfer coefficient deduced from a naphthalene experiment is not numerically equal to the heat transfer coefficient which would have existed at those conditions. The usual recommendation is to adjust for the Prandtl number of Schmidt number using a relation of the form:

$$\mathrm{St}\,\mathrm{Pr}^{2/3} = f\{\mathrm{Re}\} = \mathrm{Sh}_j \mathrm{Sc}_j^{2/3} \tag{4.25.19}$$

based on laminar results. That recommendation has not been seriously tested by experiments in turbulent and separated flows. By using nominal values of the Schmidt number and Prandtl number, the heat transfer Stanton number would be 2.3 times higher than the measured Sherwood number and an uncertainty of 10% in that ratio would alter the inferred heat transfer coefficient by 23%.

System Performance Matching. Sometimes the "effective average heat transfer coefficient" for a system is inferred from the overall behavior of the system, e.g., estimating h from the effectiveness of a heat exchanger. Values deduced by this means cannot be expected to agree well with direct measurements unless a very sophisticated system description model is used.

REFERENCES

Akino, N., Kunugi, T., Ichimiya, K., Mitsushiro, K., and Ueda, M., Improved liquid-crystal thermometry excluding human color sensation, *ASME J. Heat Transfer*, 111, 558–565, 1989.

Cooper, T. E., Field, R. J., and Meyer, J. F., Liquid crystal thermography and its application to the study of convective heat transfer, *J. Heat Transfer*, 97, 442–450, 1975.

den Ouden, C. and Hoogendoorn, C. J., Local convective heat transfer coefficients for jets impinging on a plate: Experiments using a liquid crystal technique, in *Proceedings of the 5th International Heat Transfer Conference*, September 3–7, Tokyo, Japan, Vol. 5, A75-14226 03-34, Society of Heat Transfer of Japan, pp. 293–297, 1974.

Golobic, I., Petkovsek, J., Baselj, M., Papez, A., and Kenning, D., Experimental determination of transient wall temperature distributions close to growing vapor bubbles, *Heat Mass Transfer*, 45, 857–866, 2009.

Hippensteele, S. A., Russell, L. M., and Torres, F. J., Local heat transfer measurements on a large scale model turbine blade airfoil using a composite of a heater element and liquid crystals, NASA Technical Memorandum 86900, March 1985.

Hollingsworth, K., Boehman, A. L., Smith, E. G., and Moffat, R. J., Measurement of temperature and heat transfer coefficient distributions in a complex flow using liquid crystal thermography and true-color image processing, in *Collected Papers in Heat Transfer*, ASME HTD, Vol. 123, pp. 35–42, Winter Annual Meeting, San Francisco, CA, 1989.

Liu, T. and Sullivan, J. P. Heat transfer and flow structures in an excited circular impinging jet, *Int. J. Heat Mass Transfer*, 39, 3695–3706, 1996.

Moffat, R. J., The gradient approach to thermocouple circuitry. In Herzfeld, C. M. (Ed.), *Temperature, Its Measurement and Control in Science and Industry*, Rienhold, New York, 1962.

O'Donovan, T. S. and Murray, D. B., Jet impingement heat transfer—Part II: A temporal investigation of heat transfer and local fluid velocities, *Int. J. Heat Mass Transfer*, 50, 3302–3314, 2007.
O'Donovan, T. S., Persoons, T., and Murray, D. B., High-resolution hot-film measurement of surface heat flux to an impinging jet, *Meas. Sci. Technol.*, 22, 105402, 2011.
Rhee, H. S., Koseff, J. R., and Street, R. L., Flow visualization of a recirculating flow by rheoscopic liquid and liquid crystal techniques, *Exp. Fluids*, 2, 57–64, 1984.
Scholten, J. W. and Murray, D. B., Measurement of Convective heat transfer using hot film sensors: Correction for sensor overheat, *ASME J. Heat Transfer*, 118, 982–984, 1996.
Scholten, J. W. and Murray, D. B., Unsteady heat transfer and velocity of a cylinder in cross flow - 1. Low freestream turbulence, *Int. J. Heat Mass Transfer*, 41, 1139–1148, 1998.
Steinhart, J. S. and Hart, S. R., Calibration curves for thermistors, *Deep-Sea Res.*, 15, 497, 1968.

4.26 FLOW MEASUREMENT

Jungho Kim, S.A. Sherif, and Alan T. McDonald

Introduction

Flow measurements are perhaps the most common type of measurements made. Many types of flow measuring devices are available. The choice of meter type depends on the required accuracy, range, cost, ease of reading or data reduction, and service life. The simplest and cheapest device that gives the required accuracy should be chosen.

Flow measurement can be classified into four broad types: (1) direct methods, (2) flow restriction methods, (3) linear flow meters, and (4) traversing methods. Each of these is discussed in the following subsections.

Direct Methods

Tanks can be used to determine the flow rate for steady liquid flows by measuring the volume or mass of liquid collected during a known time interval. If the time interval is long enough, flow rates may be determined precisely by using tanks. Compressibility must be considered in gas volume measurements. It is not practical to measure the mass of gas, but a volume sample can be collected by placing an inverted "bell" over water and holding the pressure constant by counterweights. No calibration is required when volume measurements are set up carefully; this is a great advantage of direct methods.

Restriction Flow Meters for Flow in Ducts

Most restriction flow meters for internal flow (except the laminar flow element) are based on acceleration of a fluid stream through some form of nozzle, shown schematically in Figure 4.26.1. Flow separating from the sharp edge of the nozzle throat forms a recirculation zone shown by the dashed lines downstream from the nozzle. The main flow stream continues to accelerate from the nozzle throat to form a vena contracta at location 2 in Figure 4.26.1 and then decelerates again to fill the duct. At the vena contracta, the flow area is a minimum, the flow streamlines are essentially straight, and the pressure is uniform across the channel section. The theoretical flow rate is

$$\dot{m}_{\text{theoretical}} = \frac{A_2}{\sqrt{1-\left(A_2/A_2\right)^2}}\sqrt{2\rho\left(p_1 - p_2\right)}. \tag{4.26.1}$$

Equation 4.26.1 shows the general relationship for a restriction flow meter: Mass flow rate is proportional to the square root of the pressure differential across the meter taps. This relationship limits the flow rates that can be measured accurately to approximately a 4:1 range.

Several factors limit the utility of Equation 4.26.1 for calculating the actual mass flow rate through a meter. The actual flow area at Section 2 is unknown when the vena contracta is pronounced (e.g., for orifice plates when D_t is a small fraction of D_1). The velocity profiles approach uniform flow only at large Reynolds numbers. Frictional effects can become important (especially downstream from the meter) when the meter contours are abrupt. Finally, the location of the pressure taps influences the differential pressure reading, $p_1 - p_2$.

The actual mass flow rate is given by

$$\dot{m}_{\text{actual}} = \frac{CA_t}{\sqrt{1-\left(A_2/A_2\right)^2}}\sqrt{2\rho\left(p_1 - p_2\right)}, \tag{4.26.2}$$

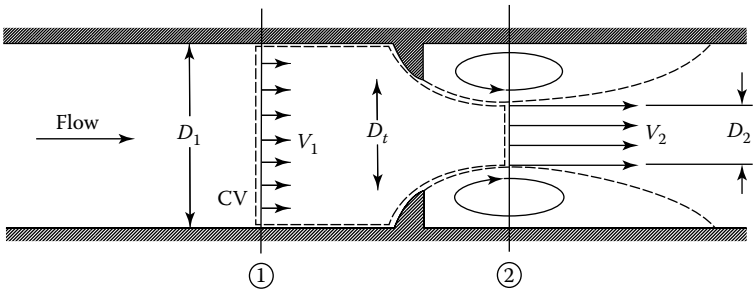


FIGURE 4.26.1 Internal flow through a generalized nozzle, showing control volume used for analysis.

where C is an empirical discharge coefficient.

If $\beta = D_t/D_1$, then $(A_t/A_1)^2 = (D_t/D_1)^4 = \beta^4$, and

$$\dot{m}_{\text{actual}} = \frac{CA_t}{\sqrt{1-\beta^4}}\sqrt{2\rho(p_1 - p_2)}, \tag{4.26.3}$$

where $1/(1-\beta^4)^{1/2}$ is the velocity correction factor. Combining the discharge coefficient and velocity correction factor into a single *flow coefficient*,

$$K \equiv \frac{C}{\sqrt{1-\beta^4}} \tag{4.26.4}$$

yields the mass flow rate in the form:

$$\dot{m}_{\text{actual}} = KA_t\sqrt{2\rho(p_1 - p_2)}. \tag{4.26.5}$$

Test data can be used to develop empirical equations to predict flow coefficients vs. pipe diameter and Reynolds numbers for standard metering systems. The accuracy of the equations (within specified ranges) is often adequate enough to use the meter without calibration. Otherwise, the coefficients must be measured experimentally.

For the turbulent flow regime ($Re_D > 4000$), the flow coefficient may be expressed by an equation of the form:

$$K = K_\infty + \frac{1}{\sqrt{1-\beta^4}}\frac{b}{\text{Re}_{D_1}^n}, \tag{4.26.6}$$

where subscript ∞ denotes the flow coefficient at infinite Reynolds numbers and constants b and n allow for scaling to finite Reynolds numbers. Correlating equations and curves of flow coefficients vs. Reynolds number are given for specific metering elements in the next three subsections following the general comparison of the characteristics of the orifice plate, flow nozzle, and venturi meters in Table 4.26.1 (see Fox and McDonald, 1992).

Flow meter coefficients reported in the literature have been measured with fully developed turbulent velocity distributions at the meter inlet. When a flow meter is installed downstream from a valve, elbow, or other disturbance, a straight section of pipe must be placed in front of the meter. Approximately 10 diameters of straight pipe upstream are required for venturi meters and up to 40 diameters for orifice plate or flow nozzle meters. Some design data for incompressible flow are given in the following subsections. The same basic methods can be extended to compressible flows.

TABLE 4.26.1
Characteristics of Orifice, Flow Nozzle, and Venturi Flow Meters

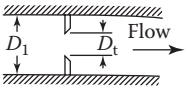
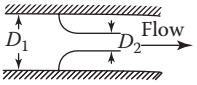
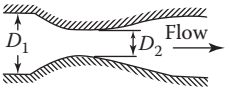
Flow Meter Type	Diagram	Head Loss	Cost
Orifice		High	Low
Flow nozzle		Intermediate	Intermediate
Venturi		Low	High

FIGURE 4.26.2 Orifice geometry and pressure tap locations. (From Miller, R.W., *Flow Measurement Engineering Handbook*, McGraw-Hill, New York, 1996. With permission.)

Orifice Plates

The orifice plate (Figure 4.26.2) may be clamped between pipe flanges. Since its geometry is simple, it is low in cost and easy to install or replace. The sharp edge of the orifice will not affect the scale or suspended matter. However, suspended matter can build up at the inlet side of a concentric orifice in a horizontal pipe; an eccentric orifice may be placed flush with the bottom of the pipe to avoid this difficulty. The primary disadvantages of the orifice are its limited capacity and the high permanent head loss caused by uncontrolled expansion downstream from the metering element.

Pressure taps for the orifices may be placed in several locations as shown in Figure 4.26.2 (see Miller, 1996 for additional details). Since the location of the pressure taps influences the empirically determined flow coefficient, one must select handbook values of K consistent with the pressure tap locations.

The correlating equation recommended for a concentric orifice with corner taps is

$$C = 0.5959 + 0.0312\beta^{2.1} - 0.184\beta^{8} + \frac{91.71\beta^{2.5}}{\mathrm{Re}_{D_1}^{0.75}}. \tag{4.26.7}$$

Equation 4.26.7 predicts orifice discharge coefficients within ±0.6% for $0.2 < \beta < 0.75$ and for $10^4 < \mathrm{Re}_{D_1} < 10^7$. Some flow coefficients calculated from Equation 4.26.7 are presented in Figure 4.26.3. Flow coefficients are relatively insensitive to Reynolds number for $\mathrm{Re}_{D_1} > 10^5$ when $\beta > 0.5$.

Flow Nozzles

Flow nozzles may be used as metering elements in either plenums or ducts, as shown in Figure 4.26.4; the nozzle section is approximately a quarter ellipse. Design details and recommended locations for pressure taps are given in *ASHRAE Handbook: Fundamentals* (1981).

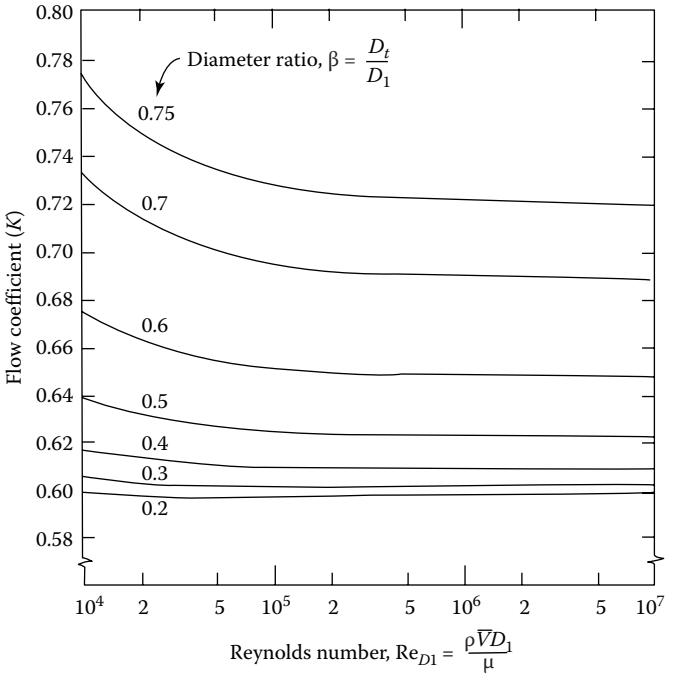


FIGURE 4.26.3 Flow coefficients for concentric orifices with corner taps.

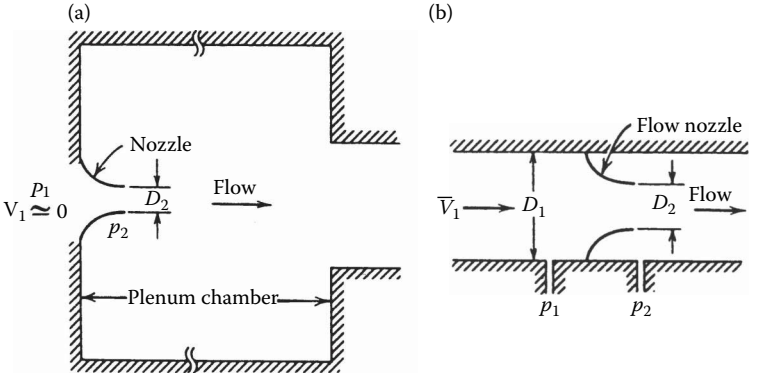


FIGURE 4.26.4 Typical installations of nozzle flow meters: (a) plenum and (b) in duct.

A similar correlating equation is available for orifice plates with D and $D/2$ taps. Flange taps require a different correlation for every line size. Pipe taps, located at $2 - 1/2\ D$ and $8D$, no longer are recommended. The correlating equation recommended for ASME long-radius flow nozzles is

$$C = 0.9975 - \frac{6.53\ \beta^{0.5}}{\mathrm{Re}_{D_1}^{0.5}}. \tag{4.26.8}$$

Equation 4.26.8 predicts discharge coefficients for flow nozzles within ±2.0% for $0.25 < \beta < 0.75$ for $10^4 < \mathrm{Re}_{D_1} < 10^7$. Some flow coefficients calculated from Equation 4.26.8 are presented in Figure 4.26.5. (K can be >1 when the velocity correction factor exceeds 1.) For plenum installation, nozzles may be fabricated from spun aluminum, molded fiber glass, or other inexpensive materials.

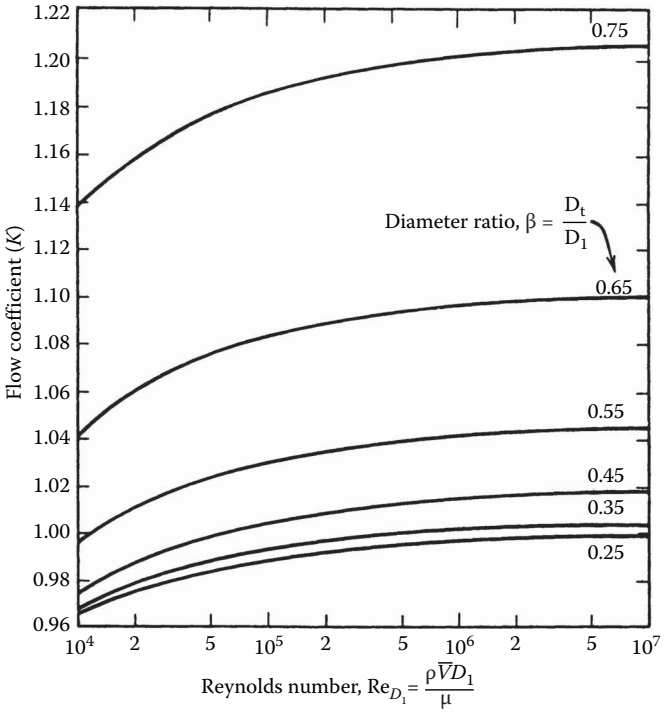


FIGURE 4.26.5 Flow coefficients for ASME long-radius flow nozzles.

Typical flow coefficients are in the range $0.95 < K < 0.99$; the larger values apply at high Reynolds numbers. Thus, the mass flow rate can be computed within approximately ±2% using $K=0.97$.

Venturi Meters

Venturi meters are generally made from castings machined to close tolerances to duplicate the performance of the standard design, so they are heavy, bulky, and expensive. The conical diffuser section downstream from the throat gives excellent pressure recovery; the overall head loss is low. Venturi meters are self-cleaning because of their smooth internal contours.

Experimentally measured discharge coefficients for venturi meters range from 0.980 to 0.995 at high Reynolds numbers ($Re_{D_1} > 2\times10^5$). Thus, $C=0.99$ can be used to calculate mass flow rate within ~±1% at high Reynolds numbers. For specific information at Reynolds numbers below 10^5, consult manufacturers' literature.

Orifice plates, flow nozzles, and venturis all produce pressure drops proportional to the flow rate squared, according to Equation 4.26.4. In practice, a meter must be sized to accommodate the largest flow rate expected. Because the pressure drop vs. flow rate relationship is nonlinear, a limited range of flow rates can be measured accurately. Flow meters with single throats are usually considered for flow rates over a 4:1 range.

Unrecoverable head loss across a metering element may be expressed as a fraction of the differential pressure across the element. These losses are shown in Figure 4.26.6.

Linear Flow Meters

The output of flow meters is typically calibrated to read a volume flow rate of a specified liquid or gas under certain conditions. For example, the volume flow rate of a liquid could be expressed in terms of liters per minute (LPM) of water, while that of a gas could be expressed in terms of standard cubic feet per minute (SCFM) of air. When using flow meters to determine flow rate at

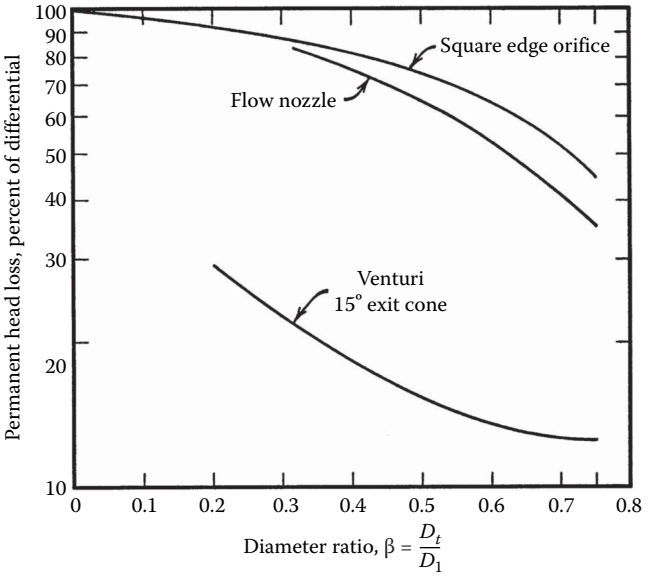


FIGURE 4.26.6 Permanent head loss produced by various flow metering elements. (From Miller, R.W., *Flow Measurement Engineering Handbook*, McGraw-Hill, New York, 1996. With permission.)

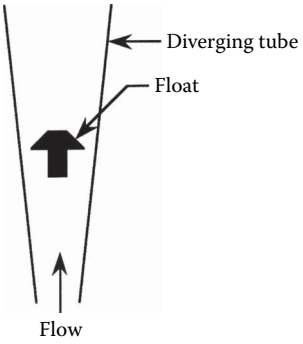


FIGURE 4.26.7 Schematic of rotameter operation.

conditions different from the calibration conditions, corrections to the output must be made to obtain the true flow rate.

Rotameters

A typical rotameter consists of a float contained within a diverging tube (Figure 4.26.7). The float is raised by a combination of drag and buoyancy. The height to which the float rises in the tube can be made to vary by varying the shape of the diverging tube—tubes with linear or logarithmic variations can easily be fabricated.

Glass tube rotameters can be used to temperatures as high as 200°C and pressures up to 20 atm, depending on the diameter of the tube. Plastic tube rotameters are also available for use at lower temperatures and pressures. Flow rates at very high temperatures and pressures can be measured using metal tubes—detectors mounted on the outside of the tube are used to determine the float position.

Rotameters usually come with calibrations based on water or air at STP (standard temperature and pressure). The output is generally in SCFM of air or LPM of water. For measuring flow at other conditions or for other fluids, corrections must be made to determine the flow rate. The response of

a rotameter can be determined by performing a force balance on the float. The height of the float is determined by a balance between the drag force on the float and the effective weight of the float:

$$F_D = \left(\rho_{float} - \rho_{fluid}\right) g V_{float}, \tag{4.26.9}$$

where the drag force is given by

$$F_D = C_D A_{float} \frac{\rho_{fluid} u_{fluid}^2}{2}. \tag{4.26.10}$$

Combining Equations 4.26.9 and 4.26.10 and solving for the fluid velocity u_{fluid} yields

$$u_{fluid} = \sqrt{\frac{2g\left(\rho_{float} - \rho_{fluid}\right)V_{float}}{C_D A_p \rho_{fluid}}}. \tag{4.26.11}$$

The mass flow rate is given by

$$\dot{m} = \rho_{fluid} A_{meter} u_{fluid}, \tag{4.26.12}$$

where A_{meter} is the area between the float and the tube. Combining Equations 4.26.11 and 4.26.12 yields

$$\dot{m} = C\rho_{fluid} A_{meter} \sqrt{\frac{\rho_{float}}{\rho_{fluid}} - 1}, \tag{4.26.13}$$

where it has been assumed that the drag coefficient, C_D, is constant. Many floats are designed for constant C_D over a wide flow range.

Consider now the case where a given flow meter indicates the same flow rate for two different fluid flows, i.e., ammeter is constant for both cases. Then,

$$\frac{\dot{m}_1}{\dot{m}_2} = \frac{\rho_{fluid,1}}{\rho_{fluid,2}} \frac{\sqrt{\rho_{float}/\rho_{fluid,1} - 1}}{\rho_{float}/\rho_{fluid,2} - 1}. \tag{4.26.14}$$

Since $\dot{m} = \rho_{fluid}\dot{V}$,

$$\frac{\dot{V}_1}{\dot{V}_2} = \frac{\sqrt{\rho_{float}/\rho_{fluid,1} - 1}}{\rho_{float}/\rho_{fluid,2} - 1}. \tag{4.26.15}$$

For cases where $\rho_{float} \gg \rho_{fluid}$ (gas flows), Equation 4.26.15 reduces to

$$\frac{\dot{V}_1}{\dot{V}_2} = \frac{\sqrt{\rho_{fluid,2}}}{\rho_{fluid,1}}. \tag{4.26.16}$$

Example: Assume we have an oil with a specific gravity of 0.8 flowing through a rotameter with a float made of steel. The steel has a specific gravity of 8.0. If the meter was calibrated for water, what is the correction that must be made to the indicated flow rate in order to obtain the true flow rate of oil?

Solution: Because the fluid density is not negligible compared to the float density, Equation 4.26.15 must be used:

$$\frac{\dot{V}_{oil}}{\dot{V}_{water}} = \frac{\sqrt{\rho_{float}/\rho_{oil} - 1}}{\rho_{float}/\rho_{water} - 1} = \frac{\sqrt{8.0/0.8 - 1}}{\sqrt{8.0/1.0 - 1}} = 1.13.$$

The true flow rate of oil, therefore, would be 1.13 times that indicated.

Target Flow Meters

These meters measure the drag force on a disk or target placed in the flow. The meter may consist simply of a hinged disk in the flow with a dial indicator showing the amount of deflection. More sophisticated meters use a strain gauge or a bellows to measure the force. Target meters can be used in dirty flows and also to sense the flow direction. They are subject to the same corrections as for rotameters.

Turbine Flow Meters

Turbine flow meters consist of a free-running, vaned impeller mounted in a cylindrical section of a tube. A cutaway of some turbine flow meters is shown in Figure 4.26.8. The rate of angular rotation is determined by counting the number of blade passages per unit time magnetically, mechanically, optically, or by some other method. No penetrations or seals in the duct are usually required, enabling turbine flow meters to measure flow rates in corrosive or toxic liquids. They are generally used for liquid flows, and are characterized by high accuracy and wide dynamic range. Reading accuracies of 0.5% over a 30:1 range are common. If units with ball bearings are used, they should be inspected occasionally to check for bearing wear. The life of the unit can be dramatically increased by operating the unit well below the maximum rated speed, however. The response of turbine flow meters can be affected by viscosity. A field calibration should be performed when the working fluid has a viscosity significantly higher than that for water.

Paddlewheel flow meters are an inexpensive alternative to turbine flow meters when extreme accuracies are not required. The rate of rotation of a paddlewheel partially immersed in the flow is used as a measure of the flow rate. Unlike turbine flow meters, paddlewheels can be used in dirty flows.

Positive Displacement Meters

Positive displacement meters measure flow rate by separating the flow into discrete increments that are moved from one side of the flow meter to the other. An example of a positive displacement meter is shown in Figure 4.26.9. Positive displacement meters can be used to measure the flow rate of high-viscosity fluids.

Vortex Shedding Meters

When a bluff body is placed in a flow, vortices can be shed downstream of the body (Figure 4.26.10). Vortex shedding meters measure the frequency of the asymmetric pressure variations that occur on the body as the vortices are shed and relate this to the flow rate. The vortex shedding frequency is governed by the Strouhal number (St)

$$\mathrm{St} = \frac{f_s d}{V}, \tag{4.26.17}$$

where

f_s = vortex shedding frequency
d = characteristic dimension of the body
V = velocity of the flow

Experiments have shown that St is constant within 1% for Reynolds numbers between 104 and 106. The frequency is directly proportional to the velocity (and thus the flow rate) around the body in this range. These flow meters are insensitive to dirty flows and can be used under a wide variety of conditions because they have no moving parts.

Thermal Mass Flow Meters

These determine flow rate of clean gases by passing the fluid through a precision tube that is heated in one section; then, the temperature of the tube wall upstream and downstream of the heated section is measured (Figure 4.26.11). The difference in temperature between the two sensors is a measure of the mass flow rate. Calibration is usually performed using nitrogen and a correction factor is applied for other gases.

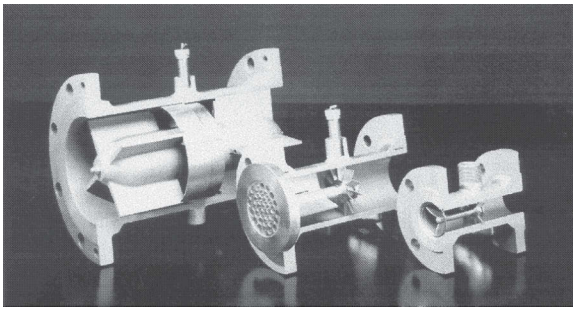


FIGURE 4.26.8 Some turbine flow meters. (Courtesy of Smith Meter, Erie, PA.)

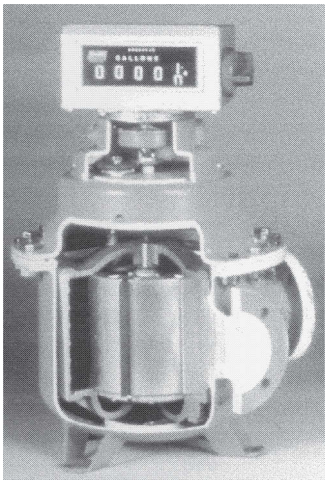


FIGURE 4.26.9 A cutaway of a positive displacement meter. (Courtesy of Smith Meter, Erie, PA.)

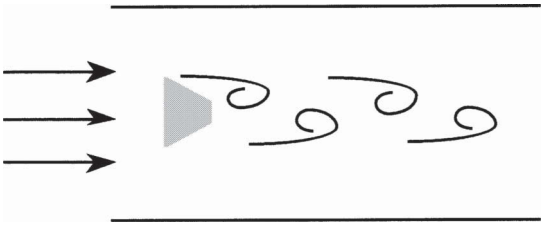


FIGURE 4.26.10 Vortices shedding off a body placed in a moving stream.

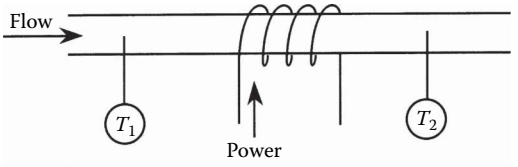


FIGURE 4.26.11 Schematic of thermal mass flow meter operation.

An alternate scheme measures the temperature of the gas; power is supplied to the downstream sensor such that its temperature is kept at a fixed amount above the upstream sensor. The feedback circuit needed to do this is similar to that used in hot-wire anemometry. The amount of power supplied to the downstream sensor is a measure of the gas flow rate. The advantages of these meters are that they give a direct indication of the flow rate without requiring corrections for temperature or pressure variations, they have no moving parts, and are noninvasive.

Ultrasonic Flow Meters

These meters determine flow rate by measuring the Doppler shift of an ultrasonic signal due to the movement of particles in the flow (Figure 4.26.12). Particles or bubbles in the fluid are usually required. Recently, meters have been developed that can measure the flow of clean fluids by measuring the velocity of turbulent swirls in the liquid. The meter is noninvasive, and can measure flow rate of dirty fluids or slurries.

The transit time method relies on measuring the difference in the time it takes for a signal to traverse the flow in one direction compared to a signal traveling in the opposite direction. Because the difference in transit times are so small, the time difference is usually measured as a phase difference between the two sound signals. When the particles move at flow speed, the frequency shift is proportional to flow speed. For a suitably chosen path, output is proportional to volume flow rate. Ultrasonic meters may require calibration in place. One or two transducers may be used.

Laminar Flow Elements

A laminar flow element (LFE) operates by passing a gas through a bundle of tubes or a honeycomb (Figure 4.26.13). The small diameter of these flow passages in the matrix laminarize the flow, and the pressure drop between the inlet and exit of the tubes is measured. If the flow passages were circular, the pressure drop across the element would be

$$\Delta P = \frac{128\mu l \dot{V}}{\pi D^4}, \tag{4.26.18}$$

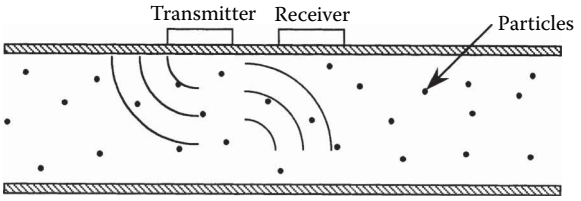


FIGURE 4.26.12 Schematic of an ultrasonic flow meter.

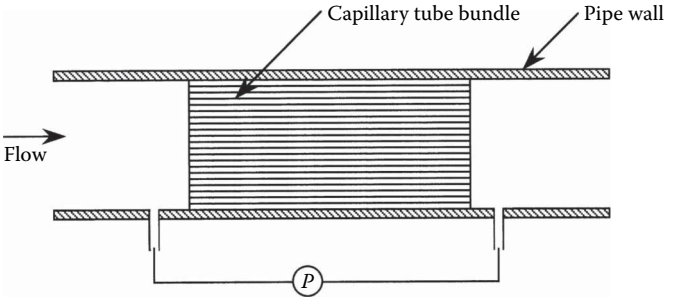


FIGURE 4.26.13 Schematic of a laminar flow element.

which is linear with the volumetric flow rate. Because the flow passages are not necessarily circular, calibration of the flow meter is usually required. Corrections are also needed for changes in fluid viscosity and temperature. The gas must be clean. The LFE may be used with reasonable accuracy over a 10:1 flow rate range. An LFE costs approximately as much as a venturi, but is much lighter and smaller. Thus, the LFE is widely used in applications where compactness and extended range are important.

Coriolis Meters

These meters are true mass flow meters. The concept behind Coriolis meters is shown in Figure 4.26.14. Consider the case where the flow moves through a U-tube that is vibrated at its resonant frequency using an electromagnetic coil. If fluid flows into the tube when the tube is moving upward, the fluid needs to be accelerated and will cause the tube to bend down. The fluid that has traveled around the bend at this time already has an upward momentum and needs to be decelerated, causing this part of the tube to bend upward. The tube, therefore, will twist in the opposite direction during the downward motion. The difference in velocity of the two legs of the vibrating tube is a measure of the mass flow rate. The density of the fluid flowing in the tube can be measured by measuring the change in resonant frequency of the tube. The instrument measures mass flow rate directly and, thus, is ideal for two-phase or liquid–solid flow measurements. Pressure drop across the coriolis meter may be high, but its useful flow rate range is 100:1.

Electromagnetic Flow Meters

Electromagnetic flow meters create a magnetic field across a pipe. When a conductive fluid passes through the field, a voltage is generated at right angles to the field and velocity vectors. Electrodes placed on a pipe detect the resulting signal voltage, which is proportional to the average axial velocity when the profile is axisymmetric. The minimum flow speed should be above ~0.3 m/s, but there are no restrictions on the Reynolds number. The flow rate range normally quoted is 10:1.7

Traversing Methods

In situations such as in air handling or refrigeration equipment, it may be impractical or impossible to install a fixed flow meter; but it may be possible to measure flow rate using a traversing technique.

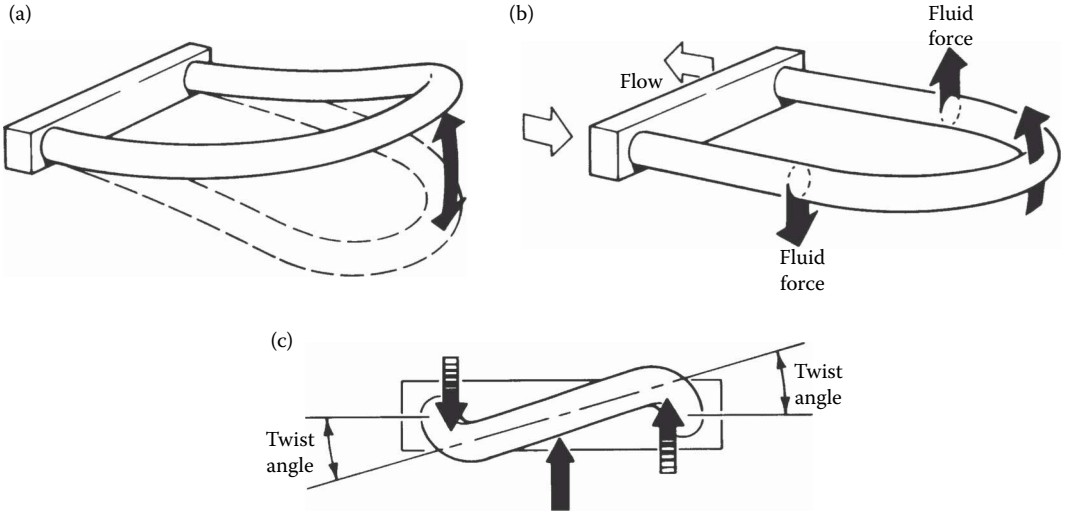


FIGURE 4.26.14 Principle of operation of a Coriolis meter: (a) vibrating flow tube, (b) fluid forces reacting to vibration of flow tube, and (c) end view of flow tube showing twist. (Courtesy of Micro Motion Inc., Boulder, CO.)

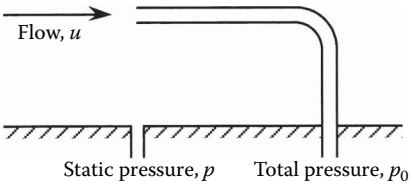


FIGURE 4.26.15 Basic total pressure probe setup.

To measure flow rate by the traversing method, the duct cross section is subdivided into segments of equal area. The fluid velocity is measured at the center of each area segment, and the volume flow rate for each segment is approximated by the product of the measured velocity and segment area. Flow rate through the entire duct is the sum of these segmental flow rates. Methods for measuring the flow velocity are discussed in the following subsections.

Pressure Probes

Perhaps the simplest and least expensive method of measuring velocity is through the use of Pitot probes. Consider the stagnating flow shown in Figure 4.26.15. Applying Bernoulli's equation to the flow between a point far upstream of the probe and the probe tip, assuming an incompressible flow, yields

$$p_0 = p + \frac{1}{2}\rho u^2. \tag{4.26.19}$$

The total pressure sensed by the tube (p_0) is greater than the free stream pressure (p) by the dynamic head $1/2\rho u^2$. The free stream velocity (u), thus, can be determined by measuring the static and total pressures. Equation 4.26.19 can be applied to flows whose Mach number (Ma=u/a) is <0.2 with <0.5% uncertainty. The speed of sound is given by

$$a = \sqrt{\gamma R_g T}\,, \tag{4.26.20}$$

where T is the static absolute temperature of the flow, R_g is the real gas constant, and γ is the ratio of specific heats. When the flow Mach number exceeds 0.2, the density of the fluid along the stagnation streamline changes and compressibility effects must be taken into account. For $0.2 < \mathrm{Ma} < 1.0$, the relation between velocity and pressure difference is given by

$$\mathrm{Ma}^2 = \frac{2}{\gamma - 1}\left[\left(\frac{p_0}{p}\right)^{\frac{\gamma-1}{\gamma}} - 1\right]. \tag{4.26.21}$$

When the flow is supersonic, the flow adjusts to the stagnation condition through a bow shock off the tip of the Pitot probe (Figure 4.26.16). The relationship between the static pressure ahead of the shock (measured by a static pressure tap) and the total pressure behind the shock (indicated by the probe) is given by

$$\frac{p_1}{p_{2,0}} = \frac{\left(\frac{2\gamma}{\gamma+1}\mathrm{Ma}_1^2 - \frac{\gamma-1}{\gamma+1}\right)^{\frac{1}{\gamma-1}}}{\left(\frac{\gamma+1}{2}\mathrm{Ma}_1^2\right)^{\frac{\gamma}{\gamma-1}}}. \tag{4.26.22}$$

Equation (4.26.22) can be solved to determine the Mach number and, hence, the flow velocity ahead of the shock.

Although the measurement of velocity using pressure probes is relatively straightforward, there are many situations where the ideal situation indicated in Figure 4.26.15 does not occur. Some factors that influence the measurement are

- Alignment of the probe with the flow
- Wall effects
- Turbulence effects
- Pitot probe design
- Static pressure tap geometry
- Streamline curvature

Each of these effects is briefly discussed below.

Alignment Effects

The change in stagnation pressure due to flow misalignment for various probes is shown in Figure 4.26.17. It is seen that an open-ended tube can be misaligned by up to ±10° without error in the pressure reading, while a tube with a chamfered opening can tolerate misalignments up to ±30°.

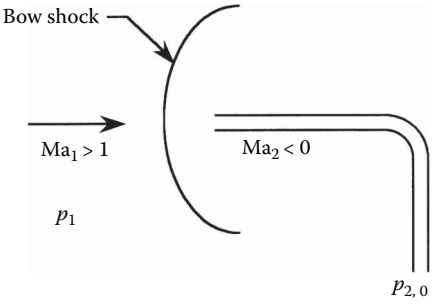


FIGURE 4.26.16 Supersonic flow around a total pressure probe.

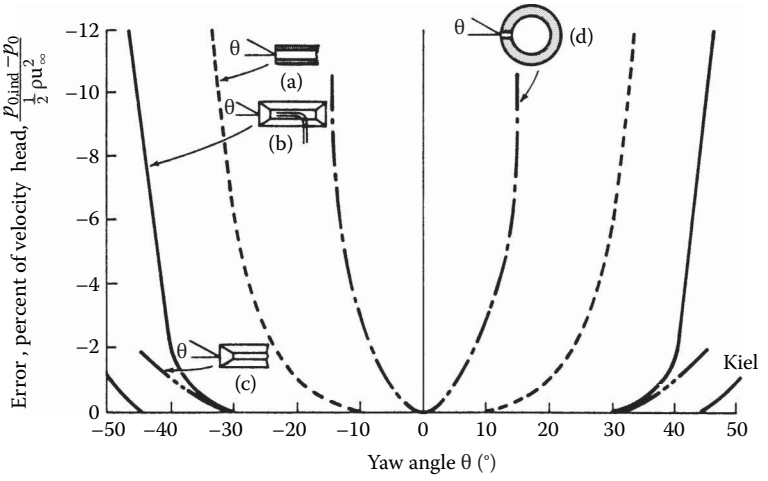


FIGURE 4.26.17 Error in stagnation pressure with changes in yaw angle for (a) open-ended tube, (b) channel tube, (c) chamfered tube, and (d) tube with orifice in the side. p_0 and $p_{0,ind}$ are the true and indicated pressures, respectively. (From Krause, L.N. and Gettleman, C.C., *ISA Proc.*, 7, 134, 1952. With permission.)

Wall Effects

The presence of a wall changes the flow around the probe, and this can result in pressure readings that are lower than when a wall is not present. For a round probe touching the wall, the velocity indicated can be up to 1.5% lower than the actual value.

Turbulence Effects

The effects of turbulence on velocity measurements are dependent on many factors such as turbulence scale and structure, fluctuation frequency, and internal geometry of the probe. The few studies that have been performed were inconclusive, with some studies indicating a decrease in the measured velocity with increasing turbulence and others indicating the opposite trend. A first-order analysis indicates that 20% turbulence intensities can result in a 2% increase in the measured velocity.

Static Pressure Taps

Static pressure taps should be drilled perpendicular to the wall; no burrs or chamfers should be present. The holes should be made as small as possible without causing excessively slow response times at the transducer.

Streamline Curvature

If the streamlines in a flow are curved (perhaps due to obstacles in the flow), then velocity and pressure gradients can occur across the channel and the static pressure obtained from a wall pressure tap may not be representative of the total pressure being measured. The local static pressure at the measuring location must either be measured or computed.

Design of Pitot-Static Probes

Pitot-static probes combine the measurement of total and static pressure in one unit. A commercially available Pitot-static probe is shown in Figure 4.26.18. The location of the static pressure

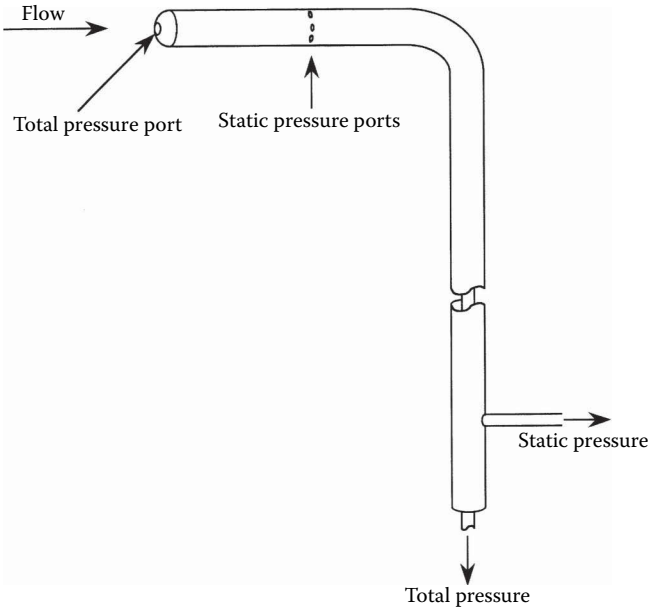


FIGURE 4.26.18 A pitot–static probe.

taps is important. The static pressure taps should be located sufficiently far downstream of the tip so that flow acceleration effects can be ignored, and sufficiently far upstream of the support so that blockage effects are small. If the static pressure taps are located eight tube diameters downstream of the tip and eight tube diameters upstream of the support, then the effects of both acceleration and blockage will be below 1%. Flow direction can be obtained using pitot probes containing multiple sensing holes. Prism probes can also be used.

Hot-Wire Anemometry

Hot-wire anemometry is used when one wishes to measure rapidly varying velocities with good spatial resolution. Velocity fluctuations up to 50 kHz can easily be obtained. Typically, a heated, thin wire is held in the flow, and the cooling effect on the wire due to the velocity of the flow is measured. The disadvantages compared to pressure probes are that they are relatively expensive, they use frail wires, and they need to be calibrated in a known flow environment. Corrections also need to be made to account for free stream temperature variations and for flow along the wire for slanted wires. There are two types of hot-wire anemometry—constant temperature and constant current.

Constant Current Anemometry

This method consists of sending a constant current through a thin wire and measuring the change in resistance of the wire with velocity. Although much simpler than the constant temperature anemometry, constant current anemometry has several disadvantages that has resulted in almost exclusive use of constant temperature anemometry today. These disadvantages include relatively low frequency response and the possibility of wire burnout at large currents and low velocities. However, they can be used when many channels of data need to be acquired, when turbulence information is not needed, or when the flow temperature needs to be measured. Numerous schemes for treating the resulting data have been published. Digital processing techniques, including fast Fourier transforms, can be used to obtain mean values and moments, and to analyze signal frequency content and correlations.

Constant Temperature Anemometry Fundamentals

The basic circuit is shown on Figure 4.26.19. The wire is represented by resistance R_w in a Wheatstone bridge. Resistors R_1, R_2, and R_c are very stable with temperature. The imbalance in

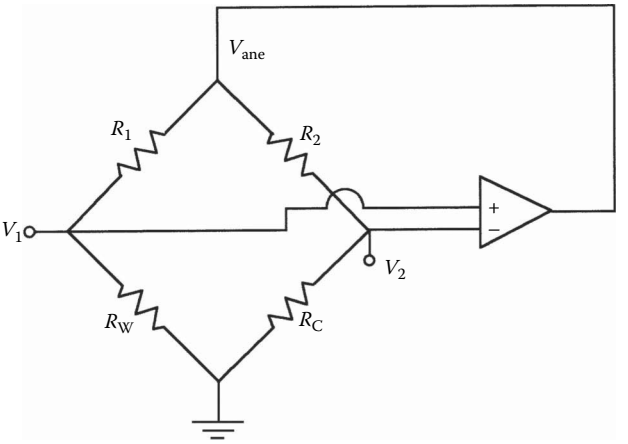


FIGURE 4.26.19 Schematic of a hot-wire circuit.

the bridge is measured using an op-amp (operational amplifier), and the op-amp outputs whatever voltage is necessary to bring the bridge back into balance. For example, consider the case where the bridge is balanced at a given flow velocity, i.e., $R_w/R_1 = R_c/R_2$ and $V_1 = V_2$. If the velocity over the wire increases, the wire temperature (and therefore its resistance) will decrease, causing the bridge to become unbalanced ($R_w/R_1 < R_c/R_2$ and $V_1 < V_2$). The op-amp senses this imbalance in the bridge and increases its output voltage causing more current to flow through the wire, bringing the bridge back into balance. The operating temperature of the wire is controlled by changing the value of R_c. The output voltage of the op-amp is a measure of the amount of cooling the wire experiences. Because the temperature of the wire does not change, the frequency response is not limited by the thermal mass of the wires and high frequency responses are possible.

Equations Governing Hot-Wire Anemometry

The cooling law is assumed to be of the form

$$\mathrm{Nu} = A_0 + B_0 \mathrm{Re}^{05}. \tag{4.26.23}$$

Multicomponent Probes

More than one component of velocity can be measured if multiple wires are present. Consider the case where a wire is inclined relative to the axis of the flow, as shown in Figure 4.26.20. The single wire probe shown in Figure 4.26.21a is often used to measure the axial component of velocity. Cross wires (Figure 4.26.21b) inclined at θ_1 and θ_2 from the normal can be used to obtain the instantaneous U and V components of velocity. The cross-correlation between u' and v' ($u'V'$), also known as the turbulent shear stress, can be obtained if the instantaneous velocities U and V are known. The wires will be sensitive to the velocities according to

$$U_{\mathrm{eff},1} = U\cos(\theta_1) + V\sin(\theta_1), \tag{4.26.24}$$

$$U_{\mathrm{eff},2} = U\cos(\theta_2) - V\sin(\theta_2). \tag{4.26.25}$$

If the effective velocities over the two wires are known along with the probe geometry, then the velocities U and V can be determined. The instantaneous velocity can be broken down into the mean and fluctuating components:

$$U = \overline{U} + u', \tag{4.26.26}$$

$$V = \overline{V} + v'. \tag{4.26.27}$$

Since

$$\overline{uv} = \overline{\left(\overline{U} + u'\right)\left(\overline{V} + v'\right)} = \overline{UV} + \overline{u'v'}, \tag{4.26.28}$$

the turbulent shear stress can be written as

$$\overline{u'v'} = \overline{uv} - \overline{UV}. \tag{4.26.29}$$

These results must be corrected to take into account the directional response of the wire, as discussed in the next subsection. Three-wire probes to measure three components of velocity are also available (Figure 4.26.21c).

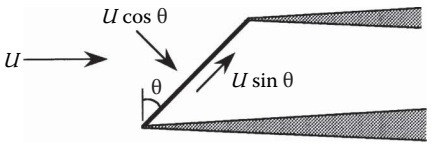


FIGURE 4.26.20 Normal and tangential components of velocity for a wire inclined to the flow.

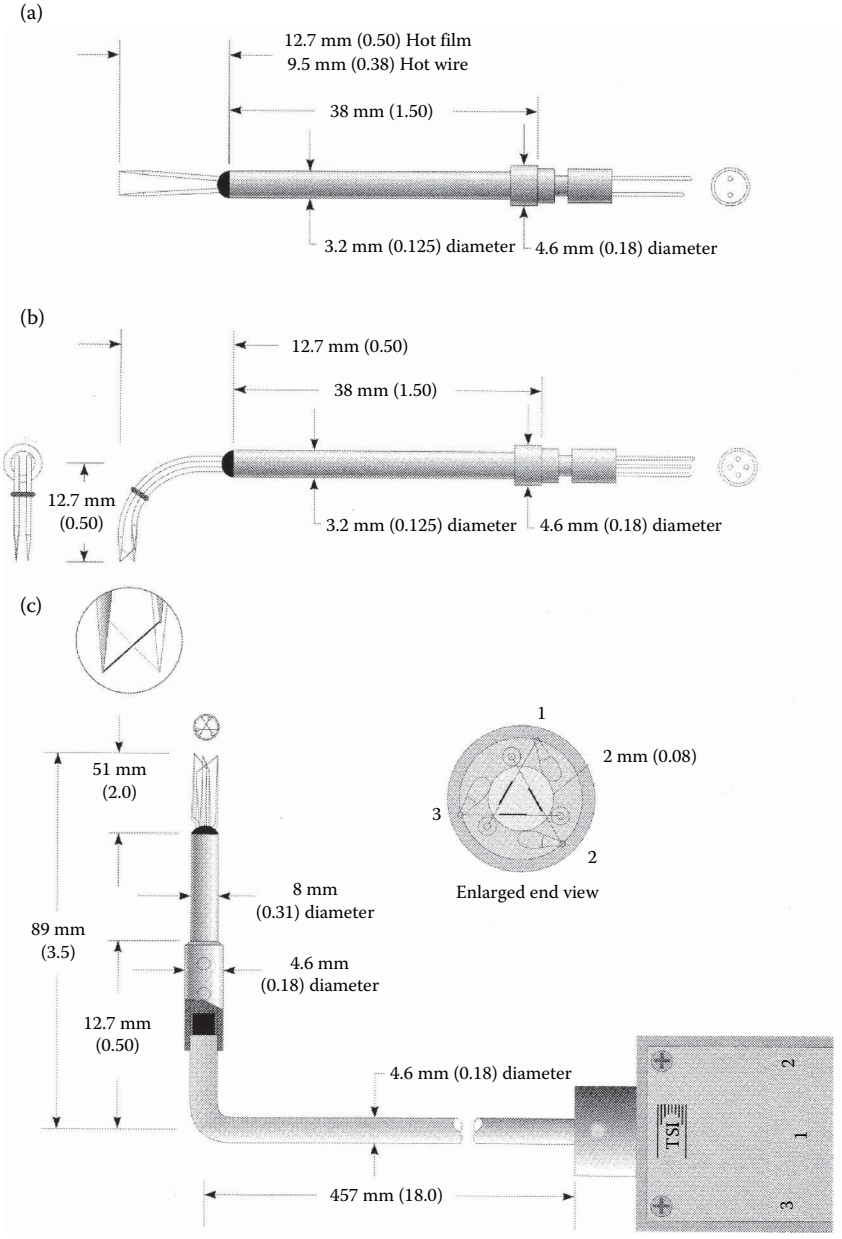


FIGURE 4.26.21 Some hot-wire probe configurations: (a) single-wire probe, (b) boundary layer cross-wire probe, and (c) triple sensor cross-flow probe. (Courtesy of TSI Inc., St. Paul, MN.)

Directional Response

For small inclinations θ to the flow, the hot wire will only respond to the velocity component normal to the wire or $U\cos\theta$. For large inclinations, a significant axial flow develops along the wire resulting in additional cooling as a result of the finite length of the wire. The equation usually used to describe this effect is based on the work by Champagne et al. (1967) and is given by

$$U_{\text{eff}}^2 = U^2\left(\cos^2\theta + k^2\sin^2\theta\right), \tag{4.26.30}$$

where k is a constant that varies linearly from 0.2 for a wire with $l/d=200$ to zero for $l/d=600$. Champagne et al. (1967) applied these results to estimate the overall error for a 45° cross wire operated at constant temperature. They found that

$$\left(\frac{\overline{u'^2}}{U^2}\right)_{\text{m}} = \left(\frac{\overline{u'^2}}{U^2}\right)_{\text{i}}, \tag{4.26.31}$$

$$\left(\frac{u'v'}{U^2}\right)_{\text{m}} = \frac{1+k^2}{1-k^2}\left(\frac{\overline{u'v'}}{U^2}\right)_{\text{i}}, \tag{4.26.32}$$

$$\left(\frac{\overline{v'^2}}{U^2}\right)_{\text{m}} = \frac{1+k^2}{1-3k^2+4k^4}\left(\frac{\overline{v'^2}}{U^2}\right)_{\text{i}}, \tag{4.26.33}$$

where the subscript m is the measured value and the subscript i is the ideal value for a wire of infinite length.

Corrections due to Temperature

Probably the most significant error when using a properly calibrated hot wire occurs as a result of drifts in the flow temperature from the calibration temperature. If the flow temperature is higher than the calibration temperature, less energy is required to keep the wire at a constant temperature for a given velocity, and the indicated flow velocity will be lower than the true flow velocity. However, a first-order correction can be made to account for the drift in flow temperature is described below. The energy dissipated by the wire is proportional to the wire-to-flow temperature difference and is also proportional to the voltage across the wire:

$$q_{\text{w}} \sim \Delta T \sim V^2. \tag{4.26.34}$$

The following equation can be used to correct for small changes in temperature:

$$V_{\text{corr}} = V_{\text{m}}\sqrt{\frac{T_{\text{w}} - T_{\text{ref}}}{T_{\text{w}} - T_{\infty}}}, \tag{4.26.35}$$

where

V_{corr} = corrected voltage
V_{m} = measured voltage
T_{w} = wire temperature
T_{ref} = temperature of the calibration flow
T_{∞} = temperature of the flow

It should be noted that the sensitivity to ambient temperature changes decreases with higher wire temperatures—the wire should be operated at the highest possible temperature.

Laser Doppler Anemometry

Laser Doppler velocimetry (LDV) uses light to measure the velocity of particles in a flow through a small volume. If the particles faithfully follow the flow, then the velocity of the particle can be assumed to be the velocity of the flow itself. LDV has numerous advantages over other methods of measuring velocity. First, it is nonintrusive so it can be used in environments that are hostile to probes or where probes excessively disturb the flow. Second, it does not depend on temperature or pressure changes in the flow. Third, it can detect flow reversals. Fourth, it can measure very high velocities if sufficiently fast electronics are available. Fifth, velocity is measured directly and is unaffected by changes in temperature, density, or composition in the flow field. LDV has been used in flows ranging from natural convection to flows in internal combustion engines and gas turbines. The disadvantages include the need for particles (seeded or naturally occurring) that follow the fluid motion, high cost, the need for optical access to the test section, and the need for careful alignment.

Principle of Operation

The basic setup for a popular type of LDV system is shown in Figure 4.26.22. A single, polarized, continuous wave laser is split into two parallel beams using a beam splitter. The beams pass through

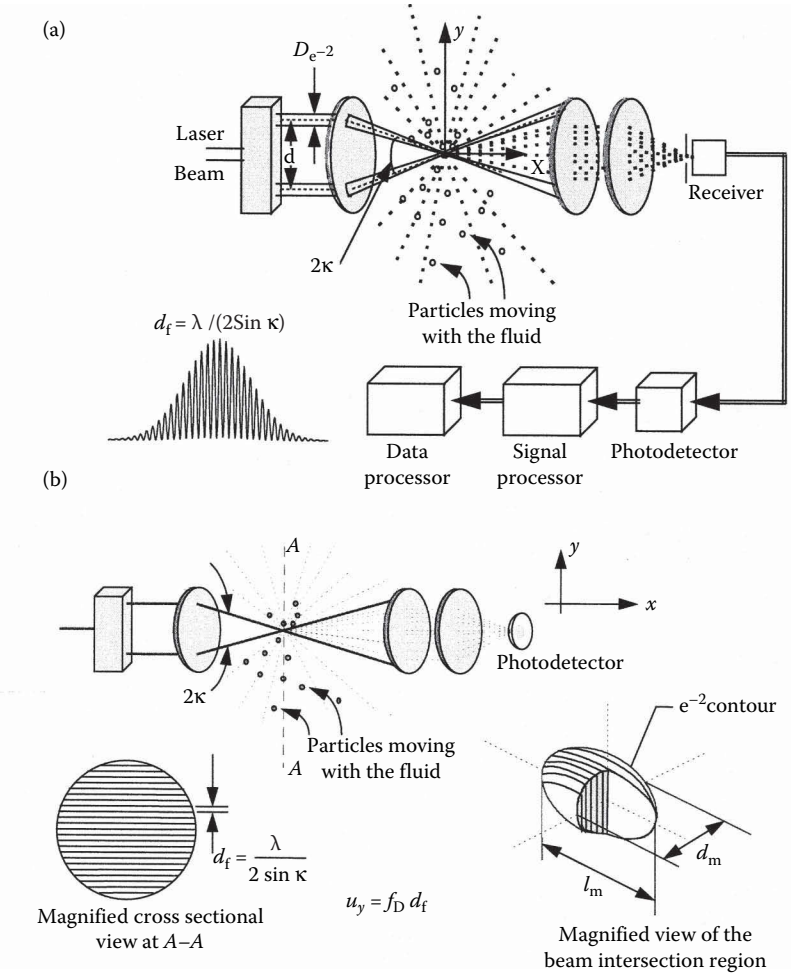


FIGURE 4.26.22 Basic setup for a two-component LDV: (a) schematic of a dual beam system and (b) details of the beam crossing. (Courtesy of TSI Inc., St. Paul, MN.)

a focusing lens that causes the beams to intersect in the focal plane of the lens, simultaneously reducing the beam diameter. The light beams approximate cylinders in the intersection region. The two beams constructively and destructively interfere with each other in the intersection regions, setting up alternating bright and dark fringes that are parallel to the axis of the laser beams. A particle moving through the measuring volume will scatter light with the intensity of the scattered light depending on whether or not the particle is in a bright or dark fringe. Only the component of velocity perpendicular to the fringes and to the axis of the laser beams will be measured. The collecting lens collects some of the scattered light and focuses it onto a photodetector that outputs a signal dependent on the light intensity. Since the spacing between the fringes can be calculated or measured, the velocity of the particle traversing the measuring volume can be determined by measuring the frequency of the signal from the photomultiplier.

Measuring Volume Dimensions

The dimensions of the measuring volume can be computed from the following relations:

$$d_{\mathrm{m}} = \frac{4 f \lambda}{\pi D_{\mathrm{e}}^{-2}}, \tag{4.26.36}$$

$$l_{\mathrm{m}} = \frac{d_{\mathrm{m}}}{\tan \kappa}, \tag{4.26.37}$$

$$d_{\mathrm{f}} = \frac{\lambda}{2 \sin \kappa}, \tag{4.26.38}$$

where

d_{m} = measuring volume diameter
l_{m} = measuring volume length
d_{f} = fringe spacing
f = focal length of the transmitting lens
D_{e}^{-2} = diameter of the laser beam at the e^{-2} intensity point entering the transmitting lens
λ = wavelength of the laser light
κ = half angle between the beams

He–Ne or argon-ion lasers are typically used. Note that in order to get a small measuring volume diameter, the focal length needs to be short, or the beam diameter needs to be large. Similar effects are seen for the measurement volume length. Ideally, one would like to have a measuring volume that is small compared to the scales of interest, with at least 10 fringes in the measuring volume.

Backscatter Mode

A configuration that is often used is shown in Figure 4.26.23. The backscatter configuration has the advantage that the optics can be made very compact, and because all the optics can be put in a single

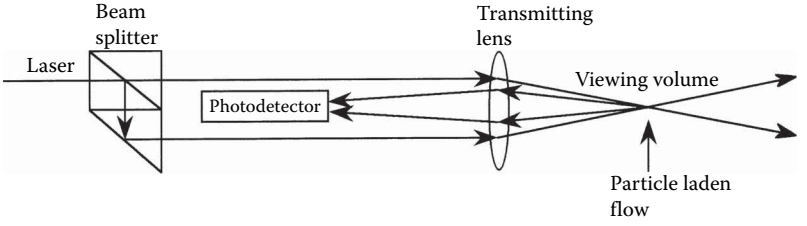


FIGURE 4.26.23 LDV configuration for backscatter.

housing, the optics need to be aligned only once. Much less light is scattered in the backward direction than in the forward direction; however, so much more sensitive photodetectors are required to detect the weaker signal.

Photodetectors

Two types of photodetectors are generally used to convert the scattered light collected by the collecting lens to an electronic signal. In the photomultiplier, photons strike a coating of photoemissive material, causing electrons to be emitted. These electrons are accelerated by a high-voltage field and impact onto a dynode that emits even more electrons. This amplification process is repeated, resulting in final gains of 10^3–10^7. Frequency responses up to 200 MHz can be obtained using photomultipliers.

Photodiodes are light-sensitive semiconductor junctions. An amplification process similar to that in photomultiplier tubes occurs in photodiodes, but gains are limited to ~10^3. The frequency response tends to be lower as well. However, photodiodes are much less expensive than photomultipliers.

Signal Processing

A typical output signal from a photodetector is shown in Figure 4.26.24. It is seen to consist of a pedestal along with the Doppler signal. Most commercial LDV systems currently on the market use fast Fourier transform or correlation methods to determine the frequency.

Particle Tracking

It is important that the particles used in the LDV system be able to follow the flow. Consider a spherical particle placed in a uniform flow. The drag force on the particle is due to the difference in speed between the particle and the carrier fluid, i.e.,

$$\text{Force} = ma = m\frac{\mathrm{d}U_\infty}{\mathrm{d}t} = \rho_\mathrm{P}\frac{\pi D_\mathrm{P}^3}{6}\frac{\mathrm{d}U_\infty}{\mathrm{d}t} = \text{Drag}. \tag{4.26.39}$$

The drag on a spherical particle is given by

$$\text{Drag} = C_\mathrm{d}\frac{1}{2}\rho_\infty\frac{\pi D_\mathrm{P}^2}{4}\left(U_\infty - U_\mathrm{P}\right)^2, \tag{4.26.40}$$

where

$$C_\mathrm{d} = \frac{24}{\mathrm{Re}_\mathrm{P}} = \frac{24\nu}{D_\mathrm{P}\left(U_\infty - U_\mathrm{P}\right)} \tag{4.26.41}$$

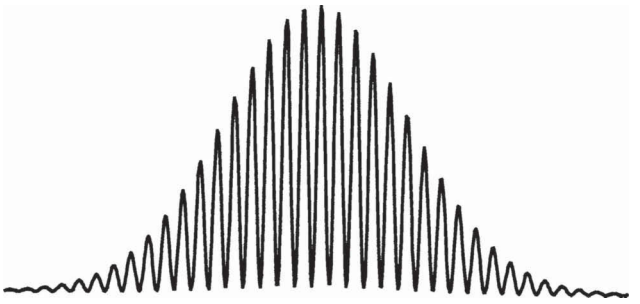


FIGURE 4.26.24 Typical photodetector signal due to scattering from a single particle. (Courtesy of TSI Inc., St. Paul. MN.)

for $0 < \mathrm{Re}_p < 100$. Combining Equations 4.26.39 through 4.26.41 and integrating yields

$$\frac{(U_\infty - U_P)}{U_\infty} = \exp(-t/K), \tag{4.26.42}$$

where

$$K = \frac{D_P^2 \rho_P}{18 \nu \rho_\infty}. \tag{4.26.43}$$

Here, K is the time constant, and represents the time it takes for the initial velocity difference between the particle and the fluid to decay by $1/e$. Particles with K small compared to the time scales of interest will faithfully follow the flow.

Flow Reversal Measurements

If the fringes are stationary, as was assumed until now, one cannot tell the direction from which the particle passes the fringes, i.e., a particle moving with a velocity U will produce the same signal as a particle moving with velocity $-U$. While the direction of the flow is unambiguous in many cases, there are other flows (recirculating flows, flows with high turbulence intensity, impinging jets, shear flows) where the flow direction is not known. This problem can be easily rectified by shifting the frequency of one of the laser beams. If one of the beams is shifted in frequency, this will cause the fringes to move as well. If the fringes move at a speed faster than the maximum absolute flow velocity, then the indicated velocity will be the flow velocity plus the fringe velocity. Zero flow velocity would correspond to the fringe velocity; negative flow velocities would correspond to the indicated velocities lower than the fringe velocity; and positive flow velocities would correspond to the indicated velocities higher than the fringe velocity.

Frequency shifting is usually accomplished by inserting an acoustic–optic Bragg cell in one of the split beams. Typical shifts are 40 MHz, i.e., a 40 MHz signal would be seen at the output of the photodetector at zero flow velocity.

Multicomponent Systems

Measurement of more than one component of velocity is more difficult than simply having two one-component systems because the signals from the two beams interfere with each other. It is necessary to be able to distinguish which pair of beams the signal came from. One can either use two pairs of beams of different color or different polarities. In the two-color method, four beams consisting of two different colors are usually used. An argon-ion laser produces both blue and green lines. This light is split and converged to form two mutually orthogonal dual-beam systems. The scattered light is collected by a pair of receiving optics/photodetector systems, one of which accepts only blue light and one of which only accepts green light. The scattered light is separated using a dichroic mirror that transmits the blue light and reflects the green light.

Particle Image Velocimetry

A relatively new method of measuring velocity at many points in the flow is by particle image velocimetry (PIV). This method enables simultaneous measurement of velocity at many points in the flow by measuring the distance that particles carried by the flow travel over a known time interval. Accuracy and spatial resolution are comparable to LDV.

Typical applications of PIV are to obtain the turbulent flow structure of complex steady flow fields where mapping is too difficult and time-consuming, and periodic flow fields where the velocity field can be obtained at different phases. An example of flow field velocity vectors obtained using PIV is shown in Figure 4.26.25.

The fundamental principle of PIV by optical imaging is the measurement of displacements δx, δy of image markers, which travel with the fluid. In PIV, the time between the marks is fixed, while

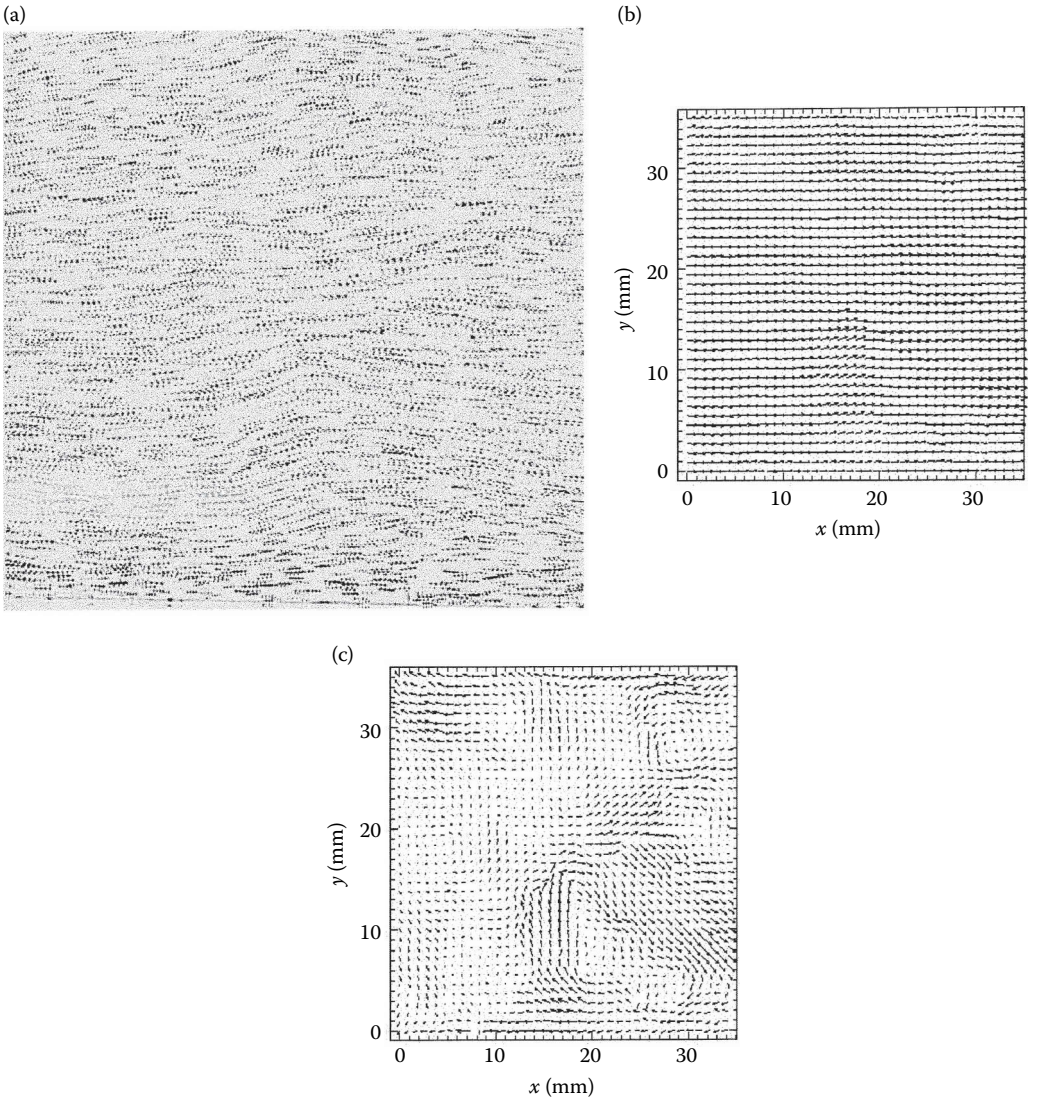


FIGURE 4.26.25 Examples of velocity vectors obtained using PIV: (a) double exposure of particles in flow field, (b) velocity vectors, and (c) velocity vectors with mean velocity subtracted out. (Courtesy of Ken Kiger, University of Maryland, College Park, MD.)

the distance between fringes is fixed in LDV. The time interval between marks in PIV must be small compared to the time scales of interest in the flow. Typical markers that have been used are H_2 bubbles, fluorescent particles, and naturally occurring small particles in liquids. The two main components of a PIV system are the imaging system that produces a double-exposed photographic image of the particles in the flow field, and the interrogation system that extracts and presents the velocity field information contained in the photograph.

Imaging System

Typically, the beam from a doubly pulsed laser is converted into a sheet of laser light using a cylindrical lens, which is then used to illuminate the markers in the flow. A camera located perpendicular to the flow records the doubly exposed image onto film. Typical light sources are pulsed lasers or shuttered continuous wave (CW) lasers. Because pulsed lasers have larger energy output than CW

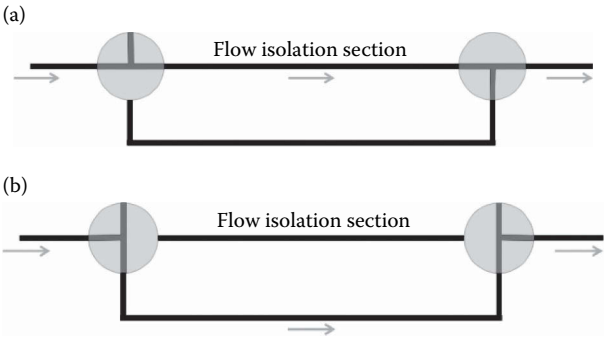


FIGURE 4.26.26 Example of (a) a double-exposed image of a particle shifted from one another, (b) the corresponding spatial autocorrelation. (Courtesy of Ken Kiger, University of Maryland, College Park, MD.)

lasers, they are generally used to illuminate small particles in the flow. Single CW laser beams swept across the flow using a rotating, faceted mirror produce a pseudo light sheet and can also be used to illuminate particles.

Film is usually necessary to resolve the markers if reasonable spatial resolution is desired. Larger markers in conjunction with high-resolution video cameras can be used, but care must be taken to ensure that the markers faithfully follow the flow. High-resolution cameras do have the advantage that near online processing of the data can be performed, however, with the appropriate digitizing and computing equipment.

The result of the imaging system is to produce two sets of markers that are displaced in time. The two sets can be on a single sheet of film, or can be recorded separately on different frames of a video camera. An example of two sets of markers is shown on Figure 4.26.26a.

Interrogation System

To obtain the velocity vectors, the doubly exposed image is first divided up into small subareas over which the average velocity vector will be computed. The size of the subareas should be as small as possible to obtain the highest number of velocity vectors; but they should not be so small that an insufficient number of particle pairs result. A good rule of thumb is to choose a subarea size that contains ~10 particle pairs. The position of each of the particles is then determined, and a spatial autocorrelation performed (Figure 4.26.26b). The location of the peaks in the spatial correlation plane yields the average particle displacement over that subarea. The largest peak results from a self-correlation and is not of interest. Two other peaks, one in the positive direction and the other in the negative direction, indicate the possible average particle displacement over that subarea.

Flow Direction

It is not possible to determine the *direction* of the velocity vector from a doubly exposed photograph alone. This is not a problem if the main flow direction is known *a priori*, but it does present problems in highly turbulent flows or if flow reversals occur. Image shifting of the second set of particles relative to the first can be used to resolve this directional ambiguity. Suppose the second set of images is shifted by a positive amount. A stationary marker will be displaced by this amount. A moving marker will have an additional displacement as a result of its movement. A shift greater than that expected at the largest negative flow velocity then eliminates ambiguity regarding flow direction since the second image is always on the positive side of the first image. This technique is analogous to using frequency shifting in LDV measurements to determine the direction of the particle movement.

Rotating mirrors and crystals have been used to produce this shift in images. More recently, cameras have been developed specifically for PIV that electronically shift the first set of images by

a certain number of pixels before acquiring the second set, eliminating any additional mechanical complexity.

Viscosity Measurements

Viscometry is the technique of measuring the viscosity of a fluid. Viscometers are classified as rotational, capillary, or miscellaneous, depending on the technique employed. Rotational viscometers use the principle that a rotating body immersed in a liquid experiences a viscous drag, which is a function of the viscosity of the liquid, the shape and size of the body, and the speed of its rotation. They are widely used because measurements can be carried out for extended periods of time.

Capillary viscometry uses the principle that when a liquid passes in laminar flow through a tube, the viscosity of the liquid can be determined from measurements of the volume flow rate, the applied pressure, and the tube dimensions. Viscometers that cannot be classified either as rotational or capillary include the falling ball viscometer. Its method of operation is based on Stokes' law, which relates the viscosity of a Newtonian fluid to the velocity of a sphere falling through it. Falling-ball viscometers are often employed for reasonably viscous fluids. Rising bubble viscometers utilize the principle that the rise of an air bubble through a liquid medium gives a visual measurement of liquid viscosity. Because of their simplicity, rising bubble viscometers are commonly used to estimate the viscosity of varnish, lacquer, and other similar media.

Void Fraction Measurement

The measurement of the void fraction (α) in a two-phase flow of liquid and gas (or vapor)—defined as the fraction of the tube cross-sectional area occupied by the gas—is important for two-phase flow closure models. Some methods that have been developed to measure α are described below.

Fast closing valve. A simple and accurate method to measure α is to use two valves to quickly isolate a section of pipe containing the two-phase flow. The void fraction is determined by dividing the volume of trapped liquid by the total trapped volume. Ball valves are generally used since their internal flow area can be made to match the tube's inner diameter to minimize flow disturbance. The valves can be rigidly linked together so they close simultaneously; fast closing solenoid-actuated valves can be used. Two three-way valves can be arranged as shown in Figure 4.26.27 such that the flow can bypass the measurement section during measurement, thereby avoiding the need for flow shutdown.

Capacitance sensor. This technique exploits the difference in the dielectric constant between liquid and gas to determine α through the change in capacitance between two conductors. The

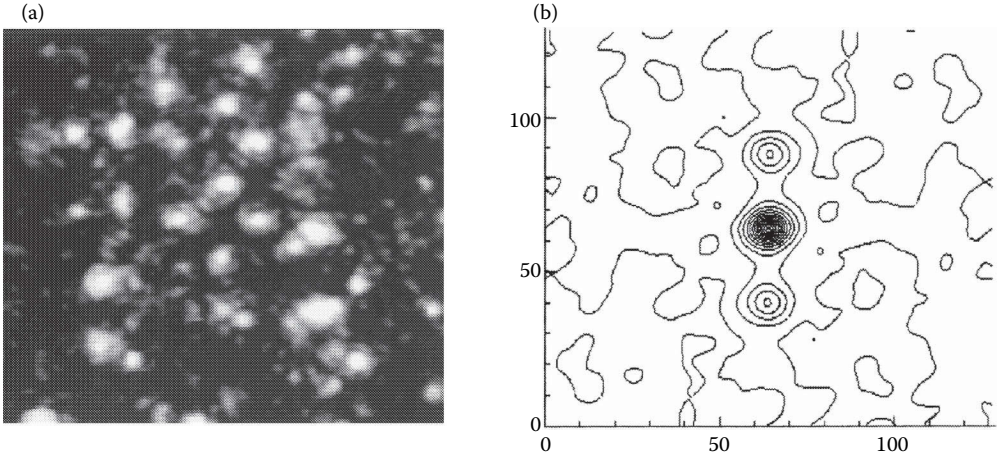


FIGURE 4.26.27 Three-way ball valve arrangement that allows the flow to be uninterrupted during measurement. (a) Straight through flow during normal operation. (b) Bypassed flow during void fraction measurement.

conductors can be deposited around the two halves of the tube or as helical strips wound around the tube. The measured capacitance has been found to be relatively insensitive to the flow pattern. Calibration of the sensor vs. void fraction is usually performed with the results expressed as the capacitance change normalized by the maximum capacitance change:

$$\alpha = f\left(\frac{C - C_1}{C_2 - C_1}\right), \tag{4.26.44}$$

where C is the measured capacitance, C_1 is the capacitance with the tube filled with Fluid 1, and C_2 is the capacitance with the tube filled with Fluid 2. Depending on the size of the conductors and the speed of the measurement, flow patterns can be identified based on time resolved measurements.

Tomography. If multiple electrodes are arranged around the circumference of the tube, the capacitance (for insulating fluids), resistance (for conducting fluid), or impedance (uses both resistance and capacitance) between any pair of electrodes can be measured using tomography. Techniques based on X-rays, gamma rays, and neutrons have also been proposed. The resolution is limited by the number of elements in the sensor array. An example of the void fraction measured using tomography is shown in Figure 4.26.28.

Conductivity wire probe. The conductivity probe exploits the difference in electrical conductivity between the liquid and gas to obtain the local void fraction in conducting liquids (Neal and Bankoff, 1963). A wire is covered with insulation except for its tip and placed in the flow. The impedance between the tip and the liquid is measured, and high and low impedance are seen when gas and liquid occupy the tip, respectively. The void fraction is given by the fraction of time the tip is occupied by gas.

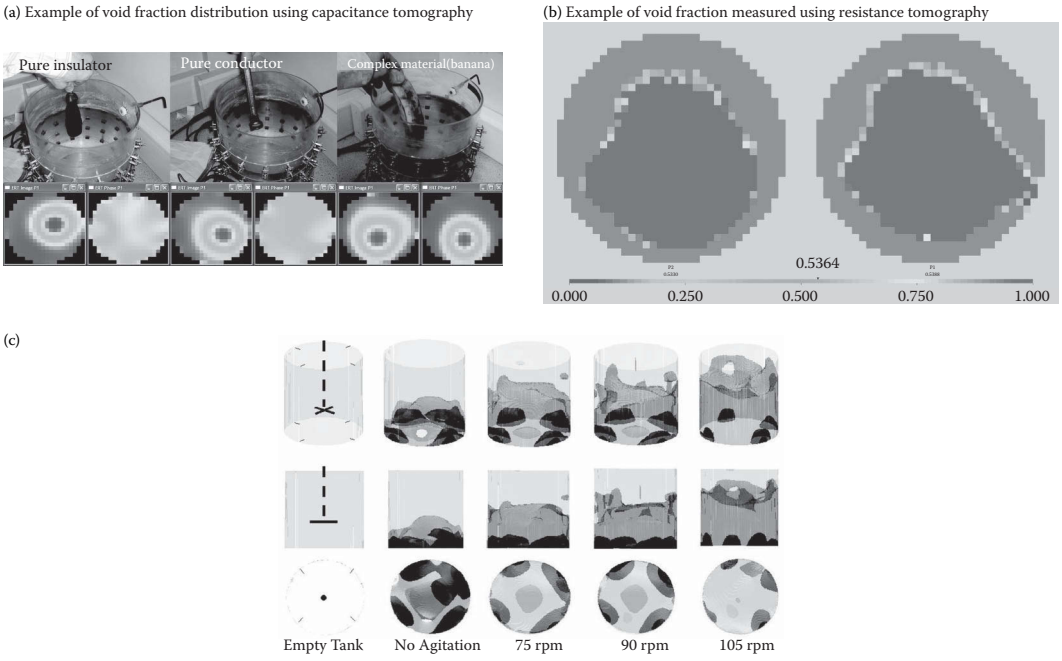


FIGURE 4.26.28 Sample results of void distribution measured using electrical capacitance tomography (top left), electrical resistance tomography (top right), and electrical impedance tomography (bottom). (Courtesy of Industrial Tomography Systems.)

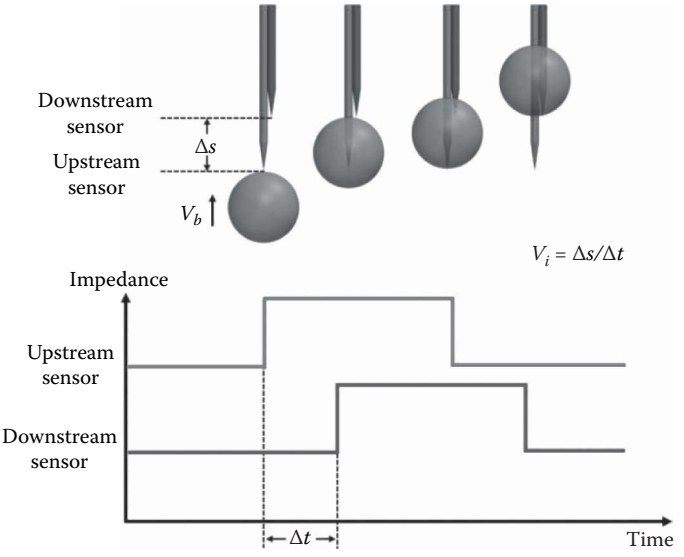


FIGURE 4.26.29 Conductivity probe measurement principle. (Courtesy of Prof. Seungjin Kim, Pennsylvania State University.)

Multiple sensors offset from each other axially and laterally can be used to obtain information about the interfacial velocity and area. For example, the interfacial velocity can be measured using the double sensor configuration shown in Figure 4.26.29 by dividing the distance between the sensors (Δs) and the measured time delay (Δt). Double and four sensor probes have been used to measure the interfacial area as well as the Sauter mean diameter and bubble frequency. Details of the measurement theory and signal processing scheme for multisensor conductivity probes including recent improvements made on the probe can be found in the literature (Kataoka et al., 1986; Kim et al., 2000; Bernard et al., 2015).

Optical probe. Optical probes are similar to conductivity probes, but depend on the index of refraction change between liquid and gas to measure the local void fraction (Abuaf et al., 1978). Typically, the tip of an optical fiber or light guide is shaped into a cone such that light traveling through the fiber is reflected back to the source by total internal reflection when gas is present. When the tip is covered by liquid, however, light is refracted into the liquid with little being reflected to the source since the probe has similar index of refraction as the liquid. The presence of bubbles and droplets as well as their arrival times and frequency can be measured in this way. If two probes are used in close proximity to each other, the velocity can also be measured.

Wire mesh sensors. The technique developed by Prasser et al. (1998, 2001) is used to measure the void fraction distribution at many points in a two-phase flow of a conducting liquid. In their sensor, two orthogonal electrode grids (a transmitter grid and a receiver grid) containing 16 sensors each are separated axially by a distance of 1.5 mm. The electrical conductivity measured using a high-speed multiplexer circuit between each of the wires in the transmitter grid and those in the receiver grid is used to determine if liquid or gas is present at each crossing point. This technique can provide detailed information regarding the local flow structure. Although the measurement is intrusive, the flow structure can be captured before the bubbles are fragmented by the wires. Development of the technique over the years has led to sensors as large as 128×128 crossing points that can be read at frame rates up to 1280 Hz. Three-layer sensors have been developed for the measurement of velocity, interfacial area density, and bubble size distribution. The reader is referred to Valasco Pena and Rodriguez (2015) for an overview of recent developments.

Defining Terms

Flow meter: Device used to measure mass flow rate or volume flow rate of fluid flowing in a duct.

Restriction flow meter: Flow meter that causes flowing fluid to accelerate in a nozzle, creating a pressure change that can be measured and related to flow rate.

Thermal anemometer: Heated sensor used to infer local fluid velocity by sensing changes in heat transfer from a small electrically heated surface exposed to the fluid flow.

Traverse: Systematic procedure used to traverse a probe across a duct cross section to measure flow rate through the duct.

REFERENCES

Abuaf, N., Jones, O.C., and Zimmer, G.A., 1978. Optical probe for local void fraction and interface velocity measurements, *Review of Scientific Instruments*, 49, 1090.

ASHRAE, 1981. *Handbook Fundamentals*, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

Bernard, M., Worosz, T., Kim, S., and Hoxie, C., 2015. Considerations in the practical application of the multisensor conductivity probe for two-phase flow, *Nuclear Technology*, 190(3), 225–235.

Champagne, F.H., Sleicher, C.A., and Wehrman, O.H. 1967. Turbulence measurements with inclined hotwires, I and II., *Journal of Fluid Mechanics*, 28, 153–182.

Fox, R.W. and McDonald, A.T., 1992. *Introduction to Fluid Mechanics*, 4th ed., John Wiley & Sons, New York.

Kataoka, I., Ishii, M., and Serizawa, A., 1986. Local formulation and measurements of interfacial area concentration in two-phase flow, *International Journal of Multiphase Flow*, 12(4), 505–529.

Kim, S., Fu, X. Y., Wang, X., and Ishii, M., 2000. Development of the miniaturized four-sensor conductivity probe and the signal processing scheme, *International Journal of Heat and Mass Transfer*, 43(22), 4101–4118.

Krause, L.N. and Gettleman, C.C., 1952. Considerations entering into the selection of probes for pressure measurement in jet engines, *ISA Proceedings*, 7, 134.

Miller, R.W., 1996. *Flow Measurement Engineering Handbook*, 3rd ed., McGraw-Hill, New York.

Neal, L.G. and Bankoff, S.G., 1963. A high resolution resistivity probe for determination of local void properties in gas-liquid flow, *AIChE Journal*, 9, 490–494.

Prasser, H.M., Bottger, A., and Zschau, J., 1998, A new electrode-mesh tomograph for gas-liquid flows, *Flow Measurement and Instrumentation*, 9(2), 111–119.

Prasser, H.M., Scholz, D., and Zippe, C., 2001. Bubble size measurement using wire mesh sensors, *Flow Measurement and Instrumentation*, 12(4), 299–312.

Velasco Pena, H.F. and Rodriguez, O.M.H., 2015. Applications of wire-mesh sensors in multiphase flows, *Flow Measurement and Instrumentation*, 45, 255–273.

Additional Reading

This section presents only an overview of flow measurement methods. The following references contain a wealth of further information. Baker[1] surveys the field and discusses precision, calibration, probe, and tracer methods, and likely developments. Spitzer[4] presents an excellent practical discussion of flow measurement.

Baker, R.C., 1989. *An Introductory Guide to Flow Measurement*, Institution of Mechanical Engineers, London.

Bruun, H.H., 1995. *Hot-Wire Anemometry: Principles and Signal Analysis*, Oxford University Press, New York.

ISO 7145, 1982. Determination of flowrate of fluids in closed conduits or circular cross sections method of velocity determination at one point in the cross section, ISO UDC 532.57.082.25:532.542. International Standards Organization, Geneva.

Spitzer, R.W., Ed., 1991. *Flow Measurement: A Practical Guide for Measurement and Control*, Instrument Society of America, Research Triangle Park, NC.

Van Wazer, J.R., Lyons, J.W., Kim, K.Y., and Colwell, R.E., 1963. *Viscosity and Flow Measurement: A Laboratory Handbook of Rheology*, Interscience, New York.

White, F.M., 1994. *Fluid Mechanics*, 3rd ed., McGraw-Hill, New York.

4.27 APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS AND GENETIC METHODS IN THERMAL ENGINEERING

Arturo Pacheco-Vega, Gerardo Diaz, Mihir Sen, and K.T. Yang

Introduction

Recent decades have seen enormous strides in technologies that can be loosely grouped together under the general term *artificial intelligence* (AI). Since, at the same time, relatively powerful desktop computers have become commonly available, applications of these methods are quickly spreading to many areas of science and engineering outside the field of computer science, especially for complex systems where more traditional methods have failed to be useful. The techniques of artificial neural networks (ANNs) and genetic methods have been shown to be particularly useful in the thermal sciences. This section is an overview of the state-of-the-art technology in this area. It is fundamentally tutorial in nature so that the reader unfamiliar with these techniques can use it as a point of departure. We will describe the methodology and scope of these methods, and then concentrate on specific applications that illustrate their use. An extensive bibliography is also provided for readers who are interested in pursuing the subject further.

Thermal systems are natural or engineered systems involving temperature differences. The engineering discipline of thermal sciences, particularly heat transfer, specifically deals with the analysis, design, and control of such systems, and has a long history of development in response to needs in a great variety of applications. Much of the advancement in the field has been the result of first-principle analyses for simpler phenomena, supplemented by experimental correlations for the more complex ones for which the first principles are either not known or the problem is too complex for their use. Unfortunately, the existing methods of analyses are no longer adequate to deal with many current applications, even though analyses are still critically needed for a variety of reasons, such as, for instance, energy efficiency, environmental concern, optimum design and control, cost, and competitiveness in the market place. While the thermal sciences as a basic engineering discipline must and should continue to develop and advance our quantitative understanding of the fundamental phenomena involved, new methods of analysis must be found in the interim, which can be used as a surrogate for such understanding. This is particularly true in the analysis and control of dynamic thermal systems. Obvious examples in this area are dynamic performance of complex heat exchangers and storage devices, network dynamics and hydronic control for heating, ventilating, and air-conditioning (HVAC) applications, dynamic thermal sensor development, thermal packaging of microelectronic devices, and thermal processing of conventional and advanced materials, just to mention a few.

AI and Its Scope

AI, though easy to recognize, is difficult to define: the word artificial is used to signify something made by humans, but intelligence itself is a more elusive concept. In fact, one can debate whether machines can actually think. Working definitions would be to say that it is "the study of how to make computers do things which, at the moment, people do better" (Rich and Knight, 1991); the "study that encompasses computational techniques for tasks that apparently require intelligence when performed by humans" (Tanimoto, 1990); "the discipline that aims to understand the nature of human intelligence through the construction of computer programs that imitate intelligent behavior" (Bonnet, 1985); or "the study of mental faculties through the use of computational models" (Charniak and McDermott, 1985). Though every definition has its drawbacks, it appears that some elements of a knowledge base, search techniques, and logical reasoning are usually involved in AI and that the techniques usually attempt a simulation of a specific aspect of human or animal behavior. There are also related and overlapping disciplines that go by other names. *Soft computing*, for

example, is a collection of techniques including fuzzy logic, neural networks, genetic algorithms, and probabilistic reasoning that are based on the idea that "intelligent systems should exploit, whenever possible, the tolerance for imprecision and uncertainty" (Bouchon-Meunier et al., 1995; Jang et al., 1997). *Knowledge engineering* and *intelligent information systems* are terms that are used when the emphasis is on the database. *Virtual reality, artificial life, machine intelligence, data mining, soft computing*, and *deep learning* are also related ideas that are being widely used. There is obviously some overlap between all these terminologies that we will not attempt to sort out here. The point is that these techniques are able to deal with complex problems that would be difficult to handle otherwise.

AI has a wide variety of applications in the natural and engineering sciences, among which are pattern recognition, decision making, system control, information processing, natural languages, symbolic mathematics, computer-aided instruction, speech recognition, vision, and robotics. These applications have spawned a large number of usable algorithms, such as optimization and search techniques, that can be used in other fields. Out of the broad set of techniques available, we will, for reasons of space, be concerned with two that have been shown to be useful in the thermal sciences: ANNs and genetic methods. We will exclude any discussion of topics such as knowledge-based techniques, fuzzy logic, and other AI methodologies.

Why AI?

At the outset, one must ask why AI techniques have any role in the thermal sciences. The answer to this lies at the core of the subject, i.e., in heat transfer. Heat transfer, like its sister discipline fluid mechanics, is based on mechanics and thermodynamics and has long been considered as one of the basic engineering sciences. At a fundamental level, it is deterministic and well understood; the equations for conduction, convection, and radiation can be easily written down, and the material properties involved can also be measured. So it is proper to discuss why AI and allied techniques are needed at all.

A major application of heat transfer is in devices such as heat exchangers, which are one of the basic components of thermal systems. A wide variety of heat exchangers exists in the market and for each it is necessary, for design and selection purposes, to know the heat transfer rate under prescribed operating conditions. In a given device exchanging heat between two fluids, the heat transfer rate depends on the flow rates and the inlet temperatures of each fluid. From the heat transfer rate, secondary quantities such as the fluid outlet temperatures can be determined. Though the governing equations can be written down, it is extremely difficult to solve them numerically under actual operating conditions. Among the reasons for this are as follows: complicated heat and fluid flow geometries, turbulence in the flow, existence of hydrodynamic and thermal entrance regions, nonuniform local heat transfer rates and fluid temperatures, secondary flows in the tube bends, vortices in the neighborhood of the tube–fin junctions, airside flow development in fin passages, heat conduction along tube walls, natural convection within the tubes and between the fins, and temperature dependence of fluid properties.

The computation must take into account the large ratio between the largest and the smallest length scales at which computations must be carried out. Accurate numerical solutions, thus, cannot be obtained with reasonable effort. As a consequence of this, we are usually forced to abandon the first-principles approach in favor of empirical information. It must be emphasized that each one of the complicating factors can be handled alone and can even be said to be understood, but their occurring together is what makes the problem difficult.

We can, therefore, for purposes of this discussion, define a *complex* system to be one that is made up of a large number of simple subsystems that interact with each other. Empirical information in the form of correlations is commonly used to help in the analysis of thermal systems that are complex. These correlations usually have little physical basis and are proposed to have the simplest form that will give the best accuracy. The vagueness in this procedure gives rise to a variety of correlations for even the most straightforward calculation such as, for example, that of turbulent heat

transfer in a duct. Furthermore, as our ability to carry out complicated calculations for the purpose of design increases, the definition of the "simplest" form changes. At this point in time, it may be said that simplicity in the correlation may no longer be of concern, but accuracy in the predictions is. In this regard, AI has much to offer. It is typically used in instances where the ultimate purpose is to detect and make use of patterns, rather than to uncover new physics.

Heat Exchangers

Since much of the following discussion will center on compact heat exchangers, we will explain this problem a little further. For the interested reader, Section 4.1 provides a detailed description of heat exchangers. Empirical information about a given heat exchanger is experimentally determined by the manufacturer and is presented to the user, i.e., the thermal system designer, in the form of the heat transfer rate as a function of the four control variables: the two flow rates and the two inlet temperatures. Being a four-variable function, it is difficult to represent completely, though partial information can be provided in the form of charts or graphs. In principle, this function depends on the geometry of the heat exchanger, the materials with which it is made, the surface conditions, the fluids used, etc. It completely characterizes the heat exchanger, and is the information that must be transferred in some form from the manufacturer to the design engineer. Since the raw experimental data would require a large amount of space for data storage, it is necessary to compress the information in some fashion for transfer from manufacturer to user, and from which it can later be accurately recovered. There are different ways in which this can be done.

In the simplest procedure, the information is compressed as two heat transfer correlations, one for the inside and another for the outside, thus characterizing the heat exchanger performance by only these two quantities. For each side, a characteristic temperature difference is assumed. However, in reality, this is found to be inaccurate; the heat transfer coefficients are known to vary considerably with flow rates and fluid properties. Furthermore, without accurate tube wall temperature measurements, which are not easy to make, it is difficult to separate the experimentally determined overall thermal resistance into its internal and external components to find each transfer coefficient separately. It is possible to assume separate correlations for the heat transfer on either side and to use a regression analysis to determine the constants. Summing the resistances, however, produces a correlation that is not linearizable and, hence, could give more than one set of constants. This is discussed in further detail later.

We can resort to dimensionless analysis to reduce the number of variables involved in the correlation. The internal and external Nusselt (or Stanton) numbers should thus depend on the corresponding Reynolds and Prandtl numbers. A correlation using nondimensional variables has the advantage that it can be used for a wide range of length scales, flow rates, and fluids. This is the most common procedure. However, loss of accuracy is the price to be paid for this generality since not all the physical phenomena involved in the heat transfer process, like the previously cited examples show, scale exactly with the limited set of nondimensional variables used. Property variations are one example; there will be a significant difference in the heat rate predictions if the manufacturer and the user assume different values of the viscosity of the liquid (the viscosity of water, for example, at 40°C is 65% of what it is at 20°C). Letting a characteristic temperature difference scale a temperature profile assumes that the profile will be similar under all conditions, which is not the case. Moreover, the user of the information, the thermal system designer, is usually interested only in the heat transfer rate and not in intermediate variables like the internal and external heat transfer coefficients. It is preferable to sacrifice generality of application and correlate the heat rate directly with respect to the other variables.

The errors that are usually present in the predictions of a heat transfer correlation are due less to errors in measurement than to data compression represented by the correlation. Accuracies of a few percent would be achievable if the predictions had the same error level as the measurements. The two approaches that we will consider in detail to resolve this problem are outlined as follows.

1. The ANN technique offers an alternative approach to the problem of information compression. It is a procedure that is usually used to predict the response of a complex physical system that cannot be easily modeled mathematically. The network is first trained by experimentally obtained input–output sets of data, after which it can be used for prediction. The manufacturer can train a network using the experimental data; the constants or parameters of the trained network can then be transferred to the user who can calculate the performance of the heat exchanger under any other flow rate or inlet temperature condition.
2. The search for the correlation that best fits the data can be carried out using genetic methods. Given a specific form of the correlation, the constants therein can be obtained through the use of a genetic algorithm (GA). Extending the idea farther, the best correlation function itself can be selected among a restricted class by genetic programming (GP). The objective is not that the form of the correlation be simple—just that its predictions be accurate.

Artificial Neural Networks

In this subsection, we will discuss the ANN technique, which is generally considered to be a subclass of AI, and its application to the analysis of complex thermal systems (Yang, 2008). Applications of ANNs have been found in such diverse fields as philosophy, psychology, business and economics, sociology, science, as well as in engineering. The common denominator is the complexity of the field.

The technique is rooted in and inspired by the biological network of neurons in the human brain that learns from external experience, handles imprecise information, stores the essential characteristics of the external input, and generalizes previous experience (Eeckman, 1992). In the biological network of interconnecting neurons, each receives many input signals from other neurons and gives only one output signal, which is sent to other neurons as part of their inputs. If the sum of the inputs to a given neuron exceeds a set threshold, normally determined by the electric potential of the receiver neuron that may be modified under different circumstances, the neuron fires and sends a signal to all the connected receiver neurons. If not, the signal is not transmitted. The firing decision represents the key to the learning and memory ability of the neural network.

The ANN attempts to mimic the biological neural network: the processing unit is the artificial neuron; it has synapses or interneuron connections characterized by synaptic weights; an operator performs a summation of the input signals weighted by the respective synapses; an activation function limits the permissible amplitude range of the output signal. It is also important to realize the essential difference between a biologic neural network and an ANN. Biological neurons function much slower than the computer calculations associated with an artificial neuron in an ANN. On the other hand, the delivery of information across the biological neural network is much faster. The biological one compensates for the relatively slow chemical reactions in a neuron by having an enormous number of interconnected neurons doing massively parallel processing, while the number of artificial neurons must necessarily be limited by the available hardware.

In this subsection, we will briefly discuss the basic principles and characteristics of the multilayer ANN, along with the details of the computations made in the feedforward mode and the associated backpropagation algorithm that is used for training. Issues related to the actual implementation of the algorithm will also be noted and discussed. Specific examples on the performance of two different compact heat exchangers analyzed by the ANN approach will then be shown, followed by a discussion on how the technique can also be applied to the dynamic performance of heat exchangers as well as to their control in real thermal systems. Finally, the potential of applying similar ANN techniques to other thermal system problems and their specific advantages will be delineated.

Methodology of ANNs

The interested reader is referred to the text by Haykin (1994) for an account of the history of ANN and its mathematical background. Many different definitions of ANNs are available; the one proposed by Schalkoff (1997) is that an ANN is a network composed of a number of artificial neurons. Each neuron has an input/output characteristic and implements a local computation or function. The output of any neuron is determined by this function, its interconnection with other neurons, and external inputs. The network usually develops an overall functionality through one or more forms of training; this is the learning process. Many different network structures and configurations have been proposed, along with their own methodologies of training (Warwick et al., 1992).

Feedforward Network

There are many different types of ANNs, but one of the most appropriate for engineering applications is the supervised, fully connected multilayer configuration (Zeng, 1998) in which learning is accomplished by comparing the output of the network with the data used for training. The *feedforward* or *multilayer perceptron* is the only configuration that will be described in some detail here. Figure 4.27.1 shows such an ANN consisting of a series of layers, each with a number of nodes. The first and last layers are for input and output, respectively, while the others are the hidden layers. The network is said to be fully connected when any node in a given layer is connected to all the nodes in the adjacent layers.

We introduce the following notation: (i, j) is the jth node in the ith layer. The line connecting a node (i, j) to another node in the next layer $i+1$ represents the synapse between the two nodes. $x_{i,j}$ is the input of the node (i, j), $y_{i,j}$ is its output, $\theta_{i,j}$ is its bias, and $w^{i,j}_{i-1,k}$ is the synaptic weight between nodes $(i-1, k)$ and (i, j). The total number of layers, including those for input and output, is I, and the number of nodes in the ith layer is J_i. The input information is propagated forward through the network; J_1 values enter the network and J_I leave. The flow of information through the layers is a function of the computational processing occurring at every internal node in the network. The relation between the output of node $(i-1, k)$ in one layer and the input of node (i, j) in the following layer is

$$x_{i,j} = \theta_{i,j} + \sum_{k=1}^{J_{i-1}} w^{i,j}_{i-1,k} y_{i-1,k} \tag{4.27.1}$$

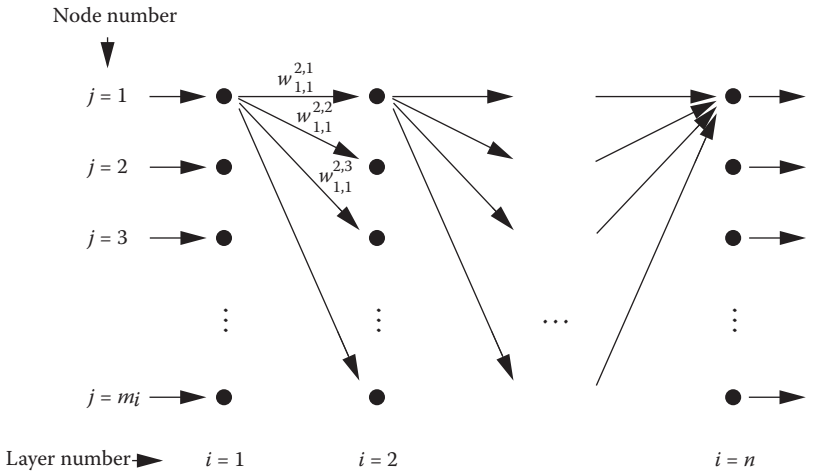


FIGURE 4.27.1 Schematic of a fully connected multilayer ANN.

Thus, the input $x_{i,j}$ of node (i, j) consists of a sum of all the outputs from the previous nodes modified by the respective internode synaptic weights $w_{i-1,k}^{i,j}$ and a bias $\theta_{i,j}$. The weights are characteristic of the connection between the nodes and the bias of the node itself. The bias represents the propensity for the combined incoming input to trigger a response from the node and presents a degree of freedom, which gives additional flexibility in the training process. Similarly, the synaptic weights are the weighting functions that determine the relative importance of the signals originated from the previous nodes.

The input and output of the node (i, j) are related by

$$y_{i,j} = \phi_{i,j}(x_{i,j}) \tag{4.27.2}$$

where $\phi_{i,j}(x)$, called the *activation* or *threshold* function, plays the role of the biological neuron determining whether it should fire or not on the basis of the input to that neuron. A schematic of the nodal operation is shown in Figure 4.27.2. It is obvious that the activation function plays a central role in the processing of information through the ANN. Keeping in mind the analogy with the biological neuron, when the input signal is small, the neuron suppresses the signal altogether, resulting in a vanishing output, and when the input exceeds a certain threshold, the neuron fires and sends a signal to all the neurons in the next layer. This behavior is determined by the activation function. Several appropriate activation functions have been studied (Haykin, 1994; Schalkoff, 1997). For instance, a simple step function can be used, but the presence of noncontinuous derivatives causes computing difficulties. The most popular one is the logistic sigmoid function

$$\phi_{i,j}(\xi) = \frac{1}{1+e^{-\xi/c}}$$

for $i > 1$, where c determines the steepness of the function. For $i = 1$, $\phi_{i,j}(\xi) = \xi$ is used instead. The sigmoid function is an approximation of the step function, but with continuous derivatives. The nonlinear nature of the sigmoid function is particularly beneficial in the simulation of practical problems. For any input $x_{i,j}$, the output of a node $y_{i,j}$ always lies between 0 and 1. Thus, from a computational point of view, it is desirable to normalize all the input and output data with the largest and smallest values of each of the data sets.

Training

For a given network, the weights and biases must be adjusted for known input–output values through a process known as training. The backpropagation method is a widely used deterministic training algorithm for this type of ANN (Rumelhart et al., 1986). The central idea of this method is to minimize an error function by the method of steepest descent to add small changes in the direction of minimization. This algorithm may be found in many texts on ANN (for instance, Rzempoluck, 1998), and only a brief outline will be given here.

In usual complex thermal system applications where no physical models are available, appropriate training data come from experiments. The first step in the training algorithm is to assign initial values to the synaptic weights and biases in the network based on the chosen ANN configuration. The values may be either positive or negative and, in general, are taken to be less than unity in absolute value. The second step is to initiate the feedforward of information starting from the input layer. In this manner, successive input and output of each node in each layer can all be computed. When

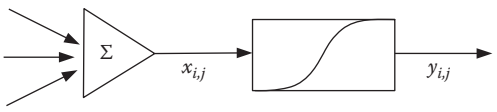


FIGURE 4.27.2 Nodal operation in an ANN.

finally $i=I$, the value of $y_{I,j}$ will be the output of the network. Training of the network consists of modifying the synaptic weights and biases until the output values differ little from the experimental data, which are the targets. This is done by means of the backpropagation method. First, an error $\delta_{I,j}$ is quantified by

$$\delta_{I,j} = (t_{I,j} - y_{I,j}) y_{I,j} (1 - y_{I,j})$$

where $t_{I,j}$ is the target output for the j node of the last layer. This equation is simply a finite difference approximation of the derivative of the sigmoid function. After calculating all the $\delta_{I,j}$, the computation then moves back to the layer $I-1$. Since the target outputs for this layer do not exist, a surrogate error defined as

$$\delta_{I-1,k} = y_{I-1,k}(1 - y_{I-1,k}) \sum_{j=1}^{J_I} \delta_{I,j} w_{I-1,k}^{I,j}$$

is used instead.

A similar error $\delta_{i,j}$ is used for the rest of the inner layers. These calculations are then continued layer by layer backward until layer 2. It is seen that the nodes of layer 1 have neither δ nor θ values assigned, since the input values are all known and invariant. After all the errors $\delta_{i,j}$ are known, the changes in the synaptic weights and biases can then be calculated by the generalized delta rule (Rumelhart et al., 1986):

$$\Delta w_{i-1,k}^{i,j} = \lambda \delta_{i,j} y_{i-1,k} \tag{4.27.3}$$

$$\Delta \theta_{i,j} = \lambda \delta_{i,j} \tag{4.27.4}$$

for $i < I$, from which all the new weights and biases can be determined. The quantity λ is known as the learning rate that is used to scale down the degree of change made to the nodes and connections. The larger the training rate, the faster the network will learn, but the chances of the ANN to reach the desired outcome may become smaller as a result of possible oscillating error behaviors. Small training rates would normally imply the need for longer training to achieve the same accuracy. Its value, usually around 0.4, is determined by numerical experimentation for any given problem.

A cycle of training consists of computing a new set of synaptic weights and biases successively for all the experimental runs in the training data. The calculations are then repeated over many cycles while recording an error quantity E for a given run within each cycle, where

$$E = \frac{1}{2} \sum_{j=1}^{J_I} (t_{I,j} - y_{I,j})^2 \tag{4.27.5}$$

The output error of the ANN at the end of each cycle can be based on either a maximum or averaged value for a given cycle. Note that the weights and biases are continuously updated throughout the training runs and cycles. The training is terminated when the error of the last cycle, barring the existence of local minima, falls below a prescribed threshold. The final set of weights and biases can then be used for prediction purposes, and the corresponding ANN becomes a model of the input–output relation of the thermal system problem.

Implementation Issues

In the implementation of a supervised, fully connected multilayered ANN, the user is faced with several uncertain choices that include the number of hidden layers, the number of nodes in each layer, the initial assignment of weights and biases, the training rate, the minimum number of training data sets and runs, the learning rate, and the range within which the input–output data are

normalized. Such choices are by no means trivial, and are rather important in achieving good ANN results. Since there is no general sound theoretical basis for specific choices, past experience and numerical experimentation are still the best guides, despite the fact that much research is now going on to provide a rational basis (Zeng, 1998).

On the issue of number of hidden layers, there is a sufficient, but certainly not necessary, theoretical basis known as the Kolmogorov mapping neural network existence theorem as presented by Hecht-Nielsen (1987), which essentially stipulates that only one hidden layer of artificial neurons is sufficient to model the input–output relations as long as the hidden layer has $2J_1 + 1$ nodes. Since in realistic problems involving a large set of input parameters, the nodes in the hidden layer would be excessive to satisfy this requirement, the general practice is to use two hidden layers as a starting point, and then add more layers as the need arises, while keeping a reasonable number of nodes in each layer (Flood and Kartam, 1994).

A slightly better situation is in the choice of the number of nodes in each layer and in the entire network. Increasing the number of internal nodes provides a greater capacity to fit the training data. In practice, however, too many nodes suffer the same fate as the polynomial curve-fitting routine by collocation at specific data points in which the interpolations between data points may lead to large errors. In addition, a large number of internal nodes slows down the ANN both in training and in prediction. One interesting suggestion given by Rogers (1994) and Jenkins (1995) is that

$$N_t = 1 + N_n \frac{J_1 + J_I + 1}{J_I}$$

where N_t is the number of training data sets, and N_n is the total number of internal nodes in the network. If N_t and J_I are known in a given problem, this equation determines the suggested minimum number of internal nodes. Also, if N_n, J_1, and J_I are known, it gives the minimum value of N_t. The number of data sets used should be larger than that given by this equation to ensure the adequate determination of the weights and biases in the training process. Other suggested procedures for choosing the parameters of the network include the one proposed by Karmin (1990) that involve first training a relatively large network that is then reduced in size by removing nodes which do not significantly affect the results, and the so-called radial Gaussian system, which adds hidden neurons to the network in an automatic sequential and systematic way during the training process (Gagarin et al., 1994). Also available is the use of evolutionary programming approaches to optimize ANN configurations (Angeline et al., 1994). Some authors (see, for example, Thibault and Grandjean, 1991) present studies of the effect of varying these parameters.

The issue of assigning the initial synaptic weights and biases is less certain. Despite the fact that better initial guesses would require less training efforts, or even less training data, such initial guesses are generally unavailable in applying the ANN analysis to a new problem. The initial assignment then normally comes from a random number generator of bounded numbers. Unfortunately, this does not guarantee that the training will converge to the final weights and biases for which the error is a global minimum. Also, the ANN may take a large number of training cycles to reach the desired level of error. Wessels and Barnard (1992), Drago and Ridella (1992), and Lehtokangas et al. (1995) suggested other methods for determining the initial assignment so that the network converges faster and avoids local minima. On the other hand, when the ANN needs upgrading with additional or new experimental data sets, the initial weights and biases are simply the existing ones.

During the training process, the weights and biases continuously change as training proceeds in accordance with Equations 4.27.3 and 4.27.4, which are the simplest correction formulas to use. Other possibilities, however, are also available (Kamarthi et al., 1992). The choice of the training rate λ is largely by trials. It should be selected to be as large as possible, but not too large to lead to nonconvergent oscillatory error behaviors. Finally, since the sigmoid function has the asymptotic limits of [0, 1] and may thus cause computational problems in these limits, it is desirable to normalize all physical variables into a more restricted range such as [0.15, 0.85]. The choice is somewhat

arbitrary. However, pushing the limits closer to [0, 1] does commonly produce more accurate training results at the expense of larger computational efforts.

Application of ANNs to Compact Heat Exchangers

In this subsection, the ANN analysis will be applied to the prediction of the performance of two different types of compact heat exchangers, one being a single-row, fin–tube heat exchanger (called heat exchanger 1), and the other a much more complicated multirow, multicolumn fin–tube heat exchanger (heat exchanger 2). In both cases, air is either heated or cooled on the fin side by water flowing inside the serpentine tubes. Except at the tube ends, the air is in a cross-flow configuration. Details of the analyses are available in the literature (Diaz et al., 1996, 1998b, 1999b; Pacheco-Vega et al., 1999, 2001a). For either heat exchanger, the normal practice is to predict the heat transfer rates by using separate dimensionless correlations for the air- and waterside coefficients of heat transfer based on the experimental data and definitions of specific temperature differences.

Heat Exchanger 1

The simpler single-row heat exchanger, a typical example being shown in Figure 4.27.3, is treated first. It is a nominal 18 in. × 24 in. plate fin–tube type exchanger manufactured by the Trane Company, with a single circuit of 12 tubes connected by bends. The experimental data were obtained in a variable-speed, open wind-tunnel facility shown schematically in Figure 4.27.4. A PID (proportional-integral-derivative)-controlled electrical resistance heater provides hot water and its flow rate is measured by a turbine flow meter. All temperatures are measured by type T thermocouples. Additional experimental details can be found in the thesis by Zhao (1995). A total of $N = 259$ test runs were made, of which only the data for $N_t = 197$ runs were used for training, while the rest were used for testing the predictions. It is advisable to include the extreme cases in the training data sets so that the predictions will be within the same range. For the ANN analysis, there are four input nodes, each corresponding to the normalized quantities: air flow rate $\dot{m}_a$, water flow rate $\dot{m}_w$, inlet air temperature T_a^{in}, and inlet water temperature T_w^{in}. There is a single output node for the normalized heat transfer rate $\dot{Q}$. Normalization of the variables was done by limiting them within the range [0.15, 0.85]. Coefficients of heat transfer have not been used since that would imply making some assumptions about the similarity of the temperature fields.

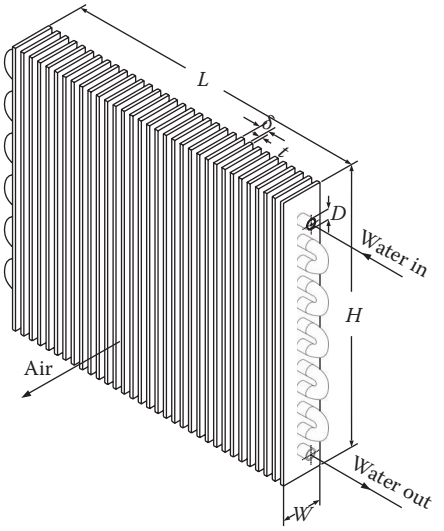


FIGURE 4.27.3 Schematic of compact heat exchanger 1.

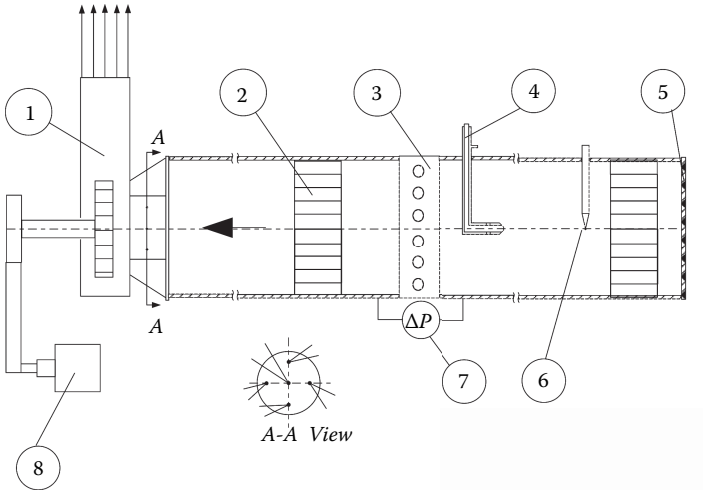


FIGURE 4.27.4 Schematic of test facility: (1) centrifugal fan, (2) flow straightener, (3) heat exchanger, (4) Pitot static tube, (5) screen, (6) thermocouple, (7) differential pressure gauge, (8) motor. View A–A shows the placement of five thermocouples.

Fourteen different ANN configurations were studied, as shown in Table 4.27.1. As an example, the training results of the 4-5-2-1-1 configuration, with three hidden layers with 5, 2, and 1 nodes, respectively, are considered in detail. The input and output layers have 4 nodes and 1 node, respectively, corresponding to the four input variables and a single output. Training was carried out to 200,000 cycles to show how the errors change along the way. The average and maximum values of the errors for all the runs can be found, where the error for each run is defined in Equation 4.27.5. These errors are shown in Figure 4.27.5. It is seen that the maximum error asymptotes at about 150,000 cycles, while the corresponding level of the average error is reached at about 100,000. In either case, the error levels are sufficiently small.

After training, the ANNs were used to predict $N_p = 62$ testing data, which were not used in the training process; the mean and standard deviations of the error for each configuration, R and σ, respectively, are shown in Table 4.27.1. R and σ are defined as

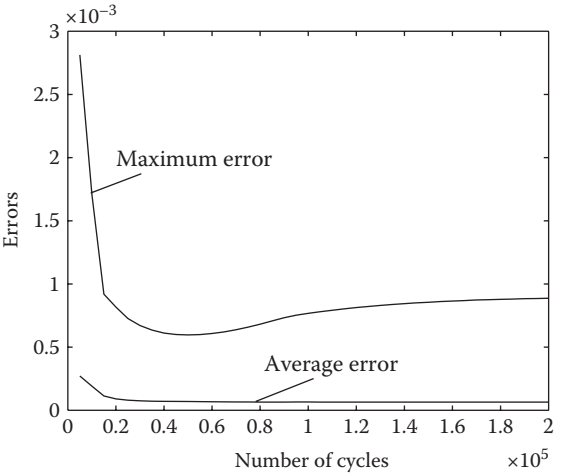


FIGURE 4.27.5 Training error results for configuration 4-5-2-1-1 ANN.

TABLE 4.27.1
Comparison of Heat Transfer Rates Predicted by Different ANN Configurations for Heat Exchanger 1

Configuration	R	σ
4-1-1	1.02373	0.266
4-2-1	0.98732	0.084
4-5-1	0.99796	0.018
4-1-1-1	1.00065	0.265
4-2-1-1	0.96579	0.089
4-5-1-1	1.00075	0.035
4-5-2-1	1.00400	0.018
4-5-5-1	1.00288	0.015
4-1-1-1-1	0.95743	0.258
4-5-1-1-1	0.99481	0.032
4-5-2-1-1	1.00212	0.018
4-5-5-1-1	1.00214	0.016
4-5-5-2-1	1.00397	0.019
4-5-5-5-1	1.00147	0.022

$$R = \frac{1}{N_p}\sum_{r=1}^{N_p} R_r \tag{4.27.6}$$

$$\sigma = \sqrt{\sum_{r=1}^{N_p}\frac{(R_r - R)^2}{N_p}} \tag{4.27.7}$$

where R_r is the ratio $\dot{Q}^e / \dot{Q}^p_{ANN}$ for run number r, $\dot{Q}^e$ is the experimental heat transfer rate, and $\dot{Q}^p_{ANN}$ is the corresponding prediction of the ANN. R is an indication of the average accuracy of the prediction, while σ is that of the scatter, both quantities being important for an assessment of the relative success of the ANN analysis. The network configuration with R closest to unity is 4-1-1-1, while 4-5-5-1 is the one with the smallest σ. If both factors are taken into account, it seems that 4-5-1-1 would be the best, even though the exact criterion is of the user's choice. It is also of interest to note that adding more hidden layers may not improve the ANN results. Comparisons of the values of R_r for all test cases are shown in Figure 4.27.6 for two configurations. It is seen, that although the 4-5-1-1 configuration is the second best in R, there are still several points at which the predictions differ from the experiments by more than 14%. The 4-5-5-1 network, on the other hand, has errors confined to 3.7%.

The effect of the normalization range for the physical variables was also studied. Additional trainings were carried out for the 4-5-5-1 network using the different normalization range of [0.05, 0.95]. For 100,000 training cycles, the results show that $R = 1.00063$ and $\sigma = 0.016$. Thus, in this case, more accurate averaged results can be obtained with the range closer to [0, 1].

We also compare the heat transfer rates obtained by the ANN analysis based on the 4-5-5-1 configuration, $\dot{Q}^p_{ANN}$, and those determined from the dimensionless correlations of the coefficients of heat transfer, $\dot{Q}^p_{cor}$. For the experimental data used, the least-square correlation equations have been given by Zhao (1995) and Zhao et al. (1995) to be

$$\varepsilon Nu_a = 0.1368 Re_a^{0.585} Pr_a^{1/3} \tag{4.27.8}$$

$$Nu_w = 0.01854 Re_w^{0.752} Pr_w^{0.3} \tag{4.27.9}$$

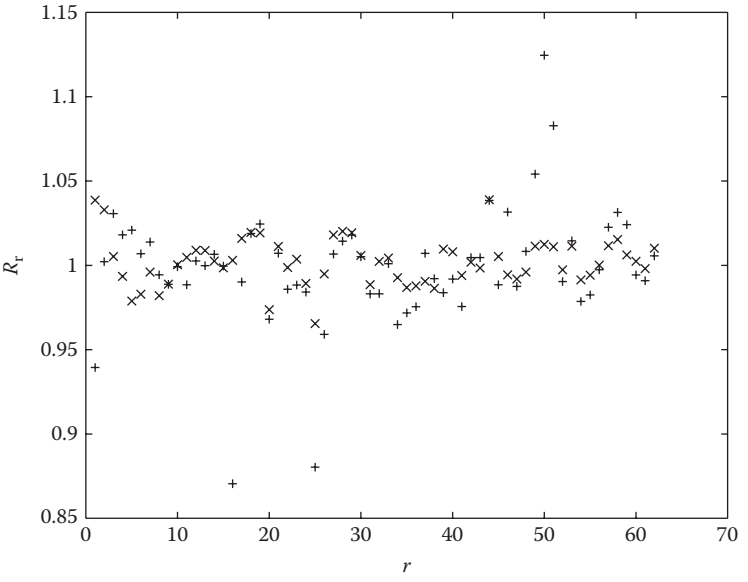


FIGURE 4.27.6 Ratio of heat transfer rates R_r for all testing runs (× 4-5-5-1; + 4-5-1-1) for heat exchanger 1.

applicable for $200 < \mathrm{Re_a} < 700$ and $800 < \mathrm{Re_w} < 4.5 \times 10^4$, where ε is the fin effectiveness. The Reynolds, Nusselt, and Prandtl numbers are defined as follows,

$$\mathrm{Re_a} = \frac{V_a \delta}{\nu_a}; \mathrm{Nu_a} = \frac{h_a \delta}{k_a}; \mathrm{Pr} = \frac{\nu_a}{\alpha_a}$$

$$\mathrm{Re_w} = \frac{V_w D}{\nu_w}; \mathrm{Nu_w} = \frac{h_w D}{k_w}; \mathrm{Pr} = \frac{\nu_w}{\alpha_w}$$

where the subscripts a and w refer to the air- and waterside, respectively, V is the average flow velocity, δ is the fin spacing, D is the inside diameter of the tube, and ν and k are the kinematic viscosity and thermal conductivity of the fluids, respectively. The correlations are based on the maximum temperature differences between the two fluids. The results are shown in Figure 4.27.7, where the superscript e is used for the experimental values and p for the predicted. For most of the data, the ANN error is within 0.7%, while the predictions of the correlation are of the order of ±10%. The superiority of the ANN is evident.

These results suggest that the ANNs have the ability of recognizing all the consistent patterns in the training data including the relevant physics as well as random and biased measurement errors. It can perhaps be said that it catches the underlying physics much better than the correlations do, since the error level is consistent with the uncertainty in the experimental data (Zhao, 1995). However, the ANN does not know and does not have to know what the physics is. It completely bypasses simplifying assumptions such as the use of coefficients of heat transfer. On the other hand, any unintended and biased errors in the training data set are also picked up by the ANN. The trained ANN, therefore, is not better than the training data, but not worse either.

Heat Exchanger 2

A second demonstration of the ANN analysis is on a multirow multicolumn compact heat exchanger shown schematically in Figure 4.27.8. Detailed heat and mass transfer measurements for this heat exchanger with different fin spacings and a wide range of operating conditions were carried out and the data reported by McQuiston (1978b). The air- and waterside thermal resistances were decoupled by using a high Reynolds number turbulent flow on the water side such that the waterside

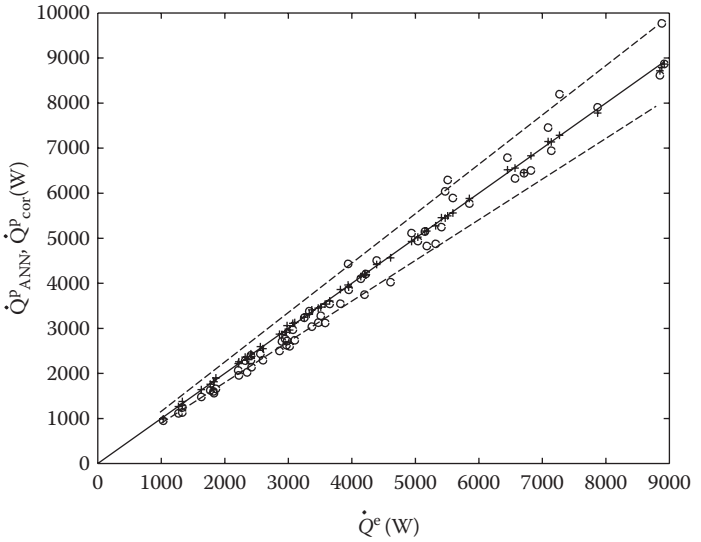


FIGURE 4.27.7 Comparison of 4-5-5-1 ANN (+) and correlation (o) predictions for heat exchanger 1.

coefficients of heat transfer were much higher than those on the air side; the waterside resistance could be estimated and subtracted from the overall thermal resistance to yield the resistance of the air side only. Details of the experimental facility and instrumentation can be found in McQuiston (1978b). The types of heat exchangers tested were for air-conditioning applications and, therefore, under certain operating conditions, condensation would occur. The heat transfer data were analyzed depending on whether the airside surface was dry, covered with droplets of water in dropwise condensation, or with water films in filmwise condensation. All physical dimensions are shown in Figure 4.27.8; the only geometrical quantity that was allowed to vary was the fin spacing, as it is the most important physical variable in such heat exchangers. On the other hand, the experimental data covered a wide range of operating conditions on the air side. McQuiston (1978a) correlated his data in terms of three airside heat transfer coefficients, nondimensionalized as Colburn j factors. These are the sensible heat j_s, the latent heat j_m, and the total heat j_t defined by

$$j_s = \frac{h_a}{G_c c_{p,a}} \mathrm{Pr}_a^{2/3}$$

$$j_m = \frac{h_{m,a}}{G_c} \mathrm{Sc}_a^{2/3}$$

$$j_t = \frac{h_{d,a}}{G_c} \mathrm{Sc}_a^{2/3}$$

where h_a is the heat transfer coefficient, and $h_{m,a}$ and $h_{d,a}$ are the mass transfer coefficients associated with the j factors, G_c is the air flow mass flux, $c_{p,a}$ is the specific heat of air, and Sc_a is the Schmidt number of air defined by $\mathrm{Sc}_a = \nu_a / D_a$, where ν_a is the kinematic viscosity of air and D_a is the mass diffusivity of water vapor in air. The j factors were correlated as functions of Reynolds numbers Re_D and $Re\delta$ and the geometrical factors $(4/\pi)(x_a/D_h)(x_b/D)\sigma_f$ and $\delta/(\delta - t)$, where

$$\mathrm{Re}_D = \frac{G_c D}{\mu_a}$$

$$\mathrm{Re}_\delta = \frac{G_c \delta}{\mu_a}$$

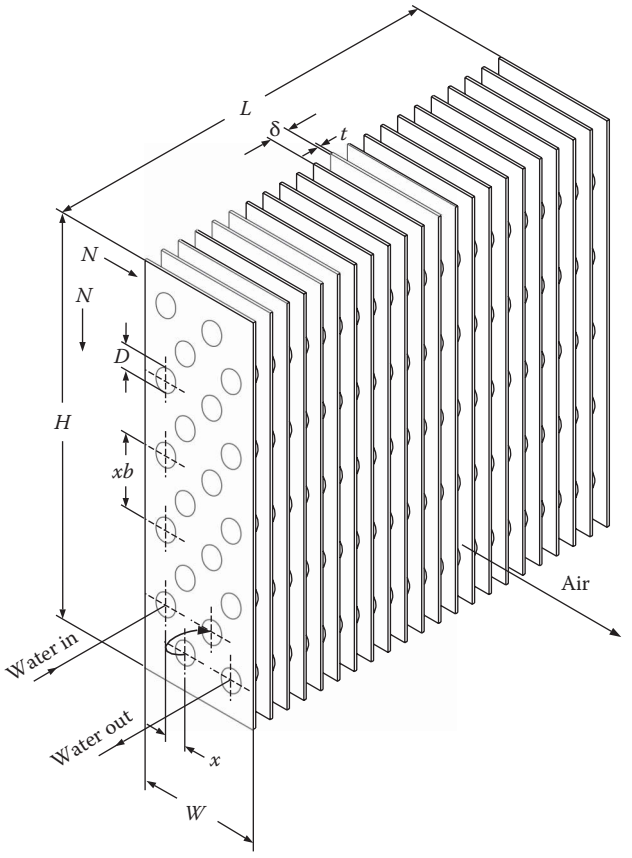


FIGURE 4.27.8 Schematic of compact heat exchanger 2. Dimensions are $L=12$ in., $D=0.392$ in., $t=0.006$ in., $x_a=0.866$ in., $x_b=1$ in., $N_r=4$, $N_c=5$, 0.071 in. $< \delta <$ 0.0250 in.

D is the outside diameter of the tube, δ is the fin spacing, and μ_a is the dynamic viscosity of air. Also x_a and x_b are the two tube-center distances shown in Figure 4.27.8, D_h is the hydraulic diameter, t is the fin thickness, and σ_f is the ratio of the free-flow area of air to the frontal area.

The McQuiston (1978a) correlations are

$$j_s = 0.0014 + 0.2618\mathrm{Re}_D^{-0.4}\left(\frac{4}{\pi}\frac{x_a}{D_h}\frac{x_b}{D}\right)^{-0.15} f_s(\delta) \tag{4.27.10}$$

$$j_t = 0.0014 + 0.2618\mathrm{Re}_D^{-0.4}\left(\frac{4}{\pi}\frac{x_a}{D_h}\frac{x_b}{D}\right)^{-0.15} f_t(\delta) \tag{4.27.11}$$

where

$$f_s(\delta) = \begin{cases} (0.9 + 4.3\times10^{-5}\,\mathrm{Re}_\delta^{1.25})\left(\dfrac{\delta}{\delta - t}\right)^{-1} & \text{dropwise condensation} \\ 1.0 & \text{dry surface} \\ 0.84 + 4.0\times10^{-5}\,\mathrm{Re}_\delta^{1.25} & \text{film condensation} \end{cases} \tag{4.27.12}$$

and

$$f_t(\delta)=\begin{cases}(0.80+4.0\times10^{-5}\,\mathrm{Re}_\delta^{1.25})\left(\dfrac{\delta}{\delta-t}\right)^4 & \text{dropwise condensation}\\ 1.0 & \text{dry surface}\\ (0.95+4.0\times10^{-5}\,\mathrm{Re}_\delta^{1.25})\left(\dfrac{\delta}{\delta-t}\right)^2 & \text{film condensation}\end{cases} \tag{4.27.13}$$

The ranges of validity are number of rows $N_r=4$, $100 < \mathrm{Re}_D < 4000$, 3/8 in. $< D <$ 5/8 in., 1 in. $< x_b <$ 2 in., 4 fins/in. $< l/\delta <$ 14 fins/in., and 200 ft/min $< V <$ 800 ft/min. Gray and Webb (1986) used these and other data to find

$$j_s = 0.14\mathrm{Re}_D^{-0.328}\left(\frac{x_b}{x_a}\right)^{-0.502}\left(\frac{\delta-t}{D}\right)^{0.0312} \tag{4.27.14}$$

for the dry-surface conditions in the ranges $500 < \mathrm{Re}_D < 24{,}700$, $1.97 < x_b/D < 2.55$, $1.97 < x_a/D < 2.58$, $0.08 < (\delta-t)/D < 0.64$, and $1 < N_r < 8$. Either correlation, i.e., Equations 4.27.10 and 4.27.11 or Equation 4.27.14, can be used to determine the total heat transfer rates under the dry-surface conditions, while those in Equations 4.27.10 and 4.27.11 are for the sensible and total heat transfer rates under wet-surface conditions.

Several ANN configurations were tried, and the best results were given by the fully connected 5-5-3-3 ANN configuration shown in Figure 4.27.9. The five input nodes correspond to the variables: inlet air dry bulb temperature $T_{a,db}^{in}$, inlet air wet bulb temperature $T_{a,wb}^{in}$, inlet water temperature T_w^{in}, the airflow Reynolds number Re_D, and the fin spacing δ. The three output nodes correspond to j_s, j_t, and the total heat transfer rate $\dot{Q}_t$. The j factors were not necessary for the heat rate predictions and were computed merely for the purpose of comparison. All variables were normalized in the same way as in the previous example. The two j factor predictions can be directly compared with the computed data given by McQuiston (1978a). A total of 327 experimental runs were reported by McQuiston (1978b), of which 117 runs were associated with observed dropwise condensation, 91 were related to dry-surface conditions, and the remaining 119 were with filmwise condensation. The data were analyzed in two different ways. First, each of the three surface conditions was analyzed separately to train three ANNs using the respective data sets. Second, the entire 327-run data sets were combined to train an ANN. The purpose here was to determine whether the network based on training data separated by some physical condition would perform better than another trained with the combined data set.

The errors obtained by these four ANNs are shown in Table 4.27.2. The errors indicated are root-mean-square values of the percentage difference between the predicted and experimental data for 800,000 training cycles. For dropwise condensation, the error is of the order of 3.3% in j_s, 3.9% in j_t, and only 1.4% in the total heat transfer rates. From the discussions in the previous example, this result is somewhat expected, since both j_s and j_t are not expected to include all the essential physics, while $\dot{Q}_t$, being a physical quantity itself, the ANN better recognizes the true relationship between $\dot{Q}_t$ and the physical input data. It is seen that this is also true for the dry surface and the surface with film condensation. It is also of interest to note that the ANN gives much better prediction for dry surfaces than for wet. This is perhaps also expected, since the physical phenomena with condensation are more complex. On the other hand, when the ANN is trained by the entire set of experimental data without separating them according to surface conditions, the errors all increase. This indicates that the ANN has more difficulty in differentiating between the different physics involved. However, it must be added that an error of 5% or less can still be counted as a relatively good prediction. It is even more remarkable to note that the predictions for $\dot{Q}_t$, which is what the user of the information is ultimately interested in, have errors only of the order of 2.7%.

Table 4.27.3 shows comparisons between the predictions of ANNs and the dimensionless correlations given earlier for j_s. Table 4.27.4 is a similar comparison for j_t. It is seen that despite the fact that

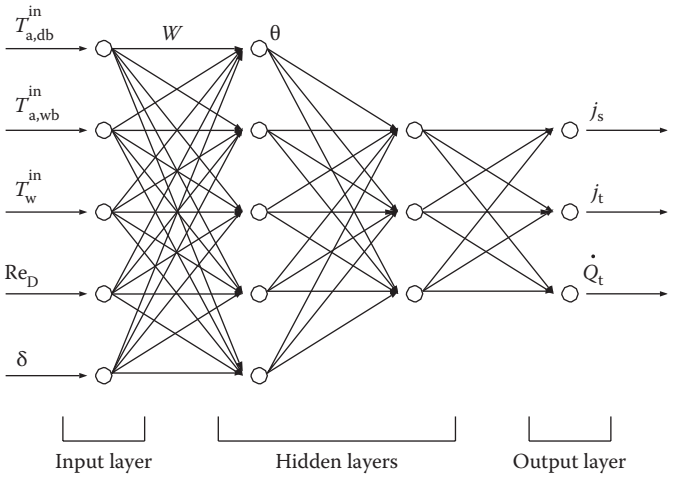


FIGURE 4.27.9 Schematic of a 5-5-3-3 ANN for heat exchanger 2.

TABLE 4.27.2
Percentage Errors in j_s, j_t, and $\dot{Q}_t$ for Heat Exchanger 2

Surface	j_s	j_t	$\dot{Q}_t$
Dropwise condensation	3.32	3.87	1.446
Dry surface	1.002	1.002	0.928
Filmwise condensation	2.58	3.15	1.960
Combined	4.58	5.05	2.69

the Gray and Webb (1986) correlation indeed performs better than those of McQuiston (1978a) under the dry-surface conditions, all the correlation predictions are much worse than those from the ANN. Actual comparisons of the j_t factors under both wet-surface conditions between the McQuiston correlations and the ANN results are shown in Figures 4.27.10 and 4.27.11, where the superscripts p and e on the j factors refer, respectively, to ANN predictions and the experimental data. This example shows that the ANN is able to predict the total heat transfer well, even for the fairly complex case of a heat exchanger in which there may be condensation phase change on the air side.

Separation of Data for Training and Testing

One of the main issues during the implementation of the ANN model of a complex system is the choice of the minimum number of training data sets and runs, since the goal is to build the best possible model over the widest parameter range of operating conditions. The data can be separated in different ways into training and testing; the usual approach is to use a fraction of the total available data during each stage of the process. This has two disadvantages: first, the data set available for training is smaller than the total amount of information available so that the predictions are not the best possible, and second, if the training data do not include the extreme values, the predictions fall in the extrapolated range and are hence less reliable. The issue of extrapolation and its impact on the reliability of ANNs, as modeling tools of complex systems, has been studied by Pacheco-Vega et al. (2001b) for the case of limited experimental data in heat exchangers.

The case of separating the data for training and testing is analyzed next by considering the data reported by McQuiston (1978b) for the heat exchanger 2 discussed in the previous section. Additional details are in Pacheco-Vega et al. (2001a). The ANN chosen is a 5-5-3-1 configuration,

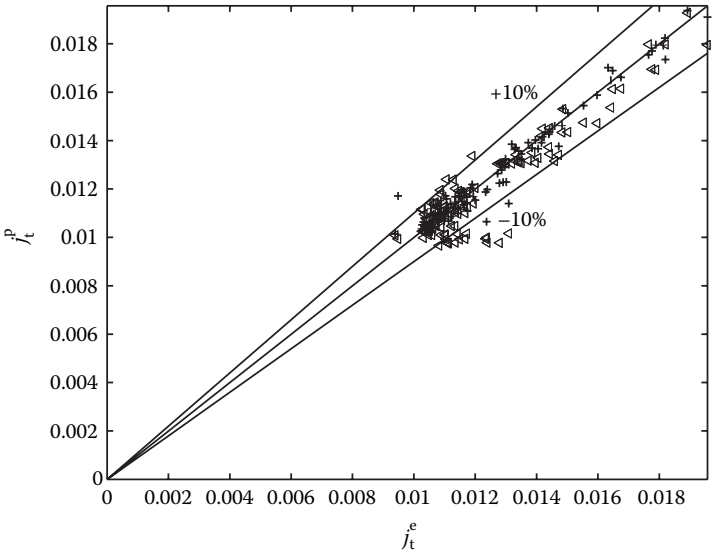


FIGURE 4.27.10 Experimental vs. predicted j_t for heat exchanger 2 with dropwise condensation. ANN: (+); McQuiston: (◁).

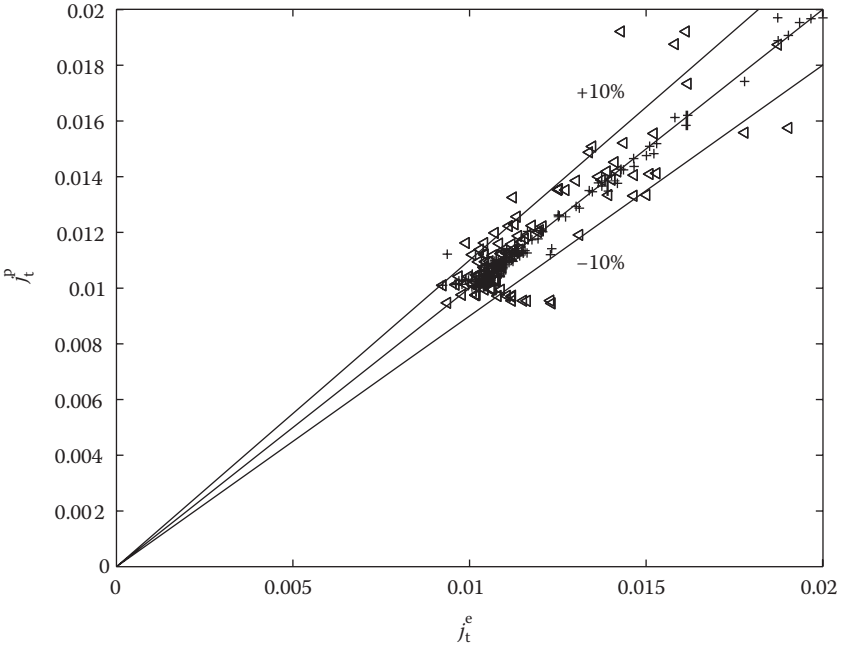


FIGURE 4.27.11 Experimental vs. predicted j_t for heat exchanger 2 with film condensation. ANN: (+); McQuiston: (◁).

similar to the one schematically shown in Figure 4.27.9—the only difference being that the fourth layer has only one output, the heat rate. The procedure is outlined as follows:

- The $M = 327$ available sets of experimental data are first randomly sorted to avoid introducing any bias in the selection process and their order is then fixed.

TABLE 4.27.3
Percentage Errors in j_s Predictions for Heat Exchanger 2

Surface	ANN	McQuiston	Gray and Webb
Dropwise condensation	3.32	8.50	—
Dry surface	1.002	14.57	11.62
Filmwise condensation	2.58	9.01	—
Combined	4.58	—	—

TABLE 4.27.4
Percentage Errors in j_t Predictions for Heat Exchanger 2

Surface	ANN	McQuiston	Gray and Webb
Dropwise condensation	3.87	7.55	—
Dry surface	1.002	14.57	11.62
Filmwise condensation	3.15	14.98	—
Combined	5.05	—	—

- From the set, only the first M_a are chosen for training and the rest $M_b = M - M_a$ kept aside. The fraction used for training is defined as $P_s = M_a/M$.
- For the M_a data sets, the rms values of the relative output errors

$$S_{\dot{Q}} = \left[\frac{1}{M_a} \sum \left(\frac{\dot{Q}^p - \dot{Q}^e}{\dot{Q}^e} \right)^2 \right]^{1/2}$$

  are calculated at each cycle of the training process in order to evaluate the performance of the network and to update the weights. The number of cycles providing a reasonably low level of error, $S_{\dot{Q}}$, in the training process is fixed for the entire process.
- After the training is finished, three different data sets are then tested: (1) the same M_a data that were used for training are tested, (2) the M_b data left out of the training process are tested, and (3) the complete M data sets are tested. In each case, the percentage error between the predicted and experimental values is calculated, being E_a, E_b, and E, respectively.
- Finally, without reordering the M data sets, the procedure described here is repeated for different values of the percentage of splitting, i.e., $P_s = 10\%$, 20%,…, 90%, 95%, and 99%.

Note that the exact shape of the error vs. P_s curve depends on the initial order of the data sets, but some general features can be identified. To remove the influence of the initial random ordering of the data sets, and get the overall characteristics of the error, the curves are calculated ten times and the results averaged.

Figure 4.27.12 shows the average error in prediction calculated in the three different ways described earlier as a function of the training fraction P_s. The error bars illustrate the standard deviations, σ_a, σ_b, and σ corresponding to E_a, E_b, and E, respectively, that resulted from the 10 different curves. From this figure, it can be seen that as P_s increases, on average, the prediction errors for all three cases asymptote to approximately the same values. For any P_s, E_a is always small since the training and testing data are identical. At small values of P_s, E_b and E are both very large indicating that an insufficient fraction of the data has been used for training. As M_a is increased, better predictions are obtained. Beyond $P_s = 60\%$ approximately, the differences in the prediction errors for all

data sets are small. The same trend is observed for σ_s, which become smaller as $P_s \rightarrow 100\%$, indicating that the prediction is somewhat insensitive to the initial random ordering of the data. Near the end, however, at $P_s = 99\%$ for example, σ_b becomes large again because now M_b is very small and the error depends greatly on the data set that is picked for testing. It is thus possible to deduce that, in this case at least, if more than 60% of the available data is used for training, the results will be essentially the same. However, it is always better to use all of the available data for training, and use the same for testing. This gives us the best prediction over the widest parameter range.

Thermal System Dynamics and Control Using ANNs

In many practical thermal systems, the operations are time dependent. A dynamic situation occurs either when a disturbance is introduced into the system or when the conditions change with time; in either case, the thermal system will respond in a time-dependent manner. In many instances, there is a need to quantitatively know what these responses may be. Ideally, this calls for dynamic system models that relate the outputs to the inputs. Unfortunately, again because of the complexity of practical thermal systems, only very few attempts at developing such models have been made and only for simple cases (Cohen and Johnson, 1956; Mozley, 1956; Thal-Larsen, 1960; Gartner and Harrison, 1965; Yamashita et al., 1978; Schmidt and Willmott, 1981; Spiga and Spiga, 1992). This is understandable since modeling is difficult even in static cases. On the other hand, in many applications, such as HVAC systems and industrial thermal processes, the control strategy is often based on the dynamic system or process models to achieve robustness and optimality. The usual PID controllers are simply not sufficient, as they often require manual supervision. Another difficulty with the dynamic thermal system models, even in those cases where they do exist, is that they are incapable of accommodating the degradation of system components over time and, hence, they may become less adequate as time goes on. One good example is the effect of fouling in heat exchanger pipes. Interestingly enough, ANNs can offer solutions to all these problems. A broad discussion on ANN-based dynamic system modeling techniques has been given by Ayoubi (1997).

Dynamic Model

The central issue in developing ANN-based dynamic models for thermal systems is the addition of time as a variable for both training and predictions.

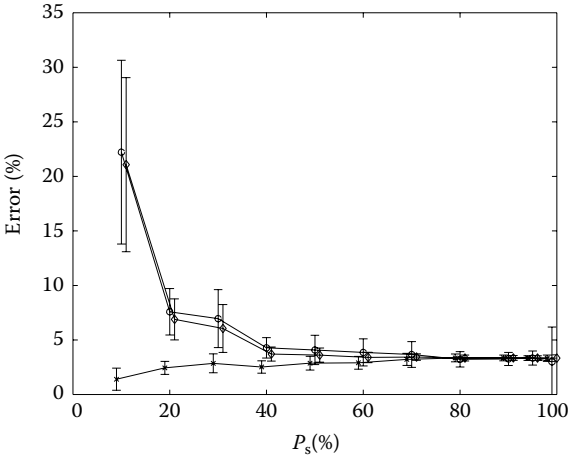


FIGURE 4.27.12 ANN prediction errors vs. percentage of data used for training; –×– error E_a using training data; –o– error E_b using data not used for training; –◇– error E using complete data. Error bars indicate standard deviations. The E_a and E curves have been shifted horizontally by –1% and 1%, respectively, for clarity.

This can be done in two different ways as shown in Figures 4.27.13. In the first methodology, depicted in Figure 4.27.13a, we train the ANN in the same way that we train it for steady state problems, where $f(t)$ corresponds to a time-dependent forcing function and t is time, which is included as an additional input variable. The main problem with using this method to predict the dynamic behavior of a system is that time needs to remain within the range used for training the ANN. Extrapolation with respect to the training data lowers the accuracy of the prediction due to activation-function saturation. The other method, shown in Figure 4.27.13b, differs with respect to the first one in the fact that no explicit information about time is provided to the network. In this methodology, used in Diaz et al. (1998a,b, 1999a, 2000a,b, 2001b,c), values of the variables at time t are provided as inputs to the network and the values of the same variables at time $t+\Delta t$ are used as the outputs. The time step Δt does not need to stay constant over time, though it should be chosen so that incremental values of all the variables remain small. The trainings are successively carried out over as many Δts as needed to cover the entire time horizon. Once trained, the ANN can predict the values of the input variables for an instant Δt later. For higher order dynamical systems, the past information at several previous instants in time might be required in order to obtain an accurate prediction with the ANN.

For a demonstration of this method, the heat exchanger 1 situation discussed earlier is again considered. The ANN to be trained had a 3-5-2 configuration. Dynamic data were obtained by varying the water inlet temperatures by changing the heater settings while keeping the other variables, the mass flow rates of both fluids and the air inlet temperature, all constant. The input to the ANN were $T_w^{in}(t)$, $T_a^{out}(t)$, $T_w^{out}(t)$ and the output were $T_a^{out}(t+\Delta t)$ and $T_w^{out}(t+\Delta t)$. Training data were obtained from experiments in which the water inlet temperature was varied in small increments of 5.56°C from 32.2°C up to 65.6°C. For testing the predictions of the ANN, three additional sets of experiments were performed.

1. The system was first brought up to a temperature of 60°C and the heater was then set at 37.8°C. The resulting sudden decrease of the air and water outlet temperatures is shown in Figure 4.27.14. The ANN predictions are excellent. The slight oscillations in the measured water outlet temperatures were likely due to local turbulence in the flow.
2. In the second set of experiments, the water inlet temperature was ramped up manually. The corresponding results are in Figure 4.27.15, which shows the same excellent prediction by the ANN.
3. The final set of experiments was designed to observe if the ANN could predict the system behavior when an input variable different from the one in the training data is used. The changes in the air and water outlet temperatures were measured when the mass flow rate of air was first increased to a value greater than the one in the training data, and then decreased to a lower value. The results are shown in Figure 4.27.16 where the decrease is

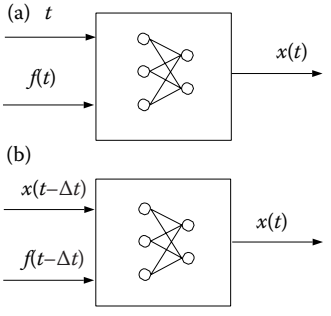


FIGURE 4.27.13 Two training methods for dynamic problems.

seen around $t = 10$ s, and the increase at $t = 120$ s. The predictions, as expected, are not very good. It is remarkable, though, that the ANN approach still predicts the correct overall trends and that the errors are small.

Dynamic Control

In the usual dynamic control systems for industrial processing and HVAC systems, the common practice is to use a feedback control involving a system model, sometimes known as the plant model, in conjunction with a standard PID controller. Such a control system, unfortunately, requires constant attention and manual supervision to achieve optimality. Although there are automatic tuning algorithms available (Zhuang and Atherton, 1993), these controllers still require a certain level of human supervision. As previously discussed, there are many advantages in

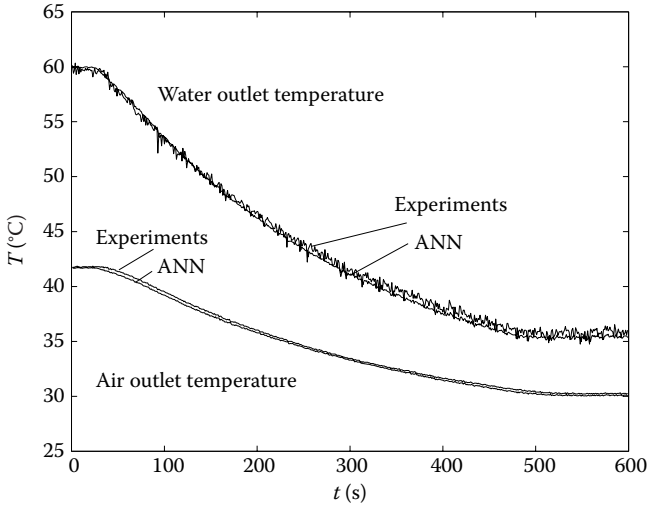


FIGURE 4.27.14 ANN dynamic prediction for sudden cooling.

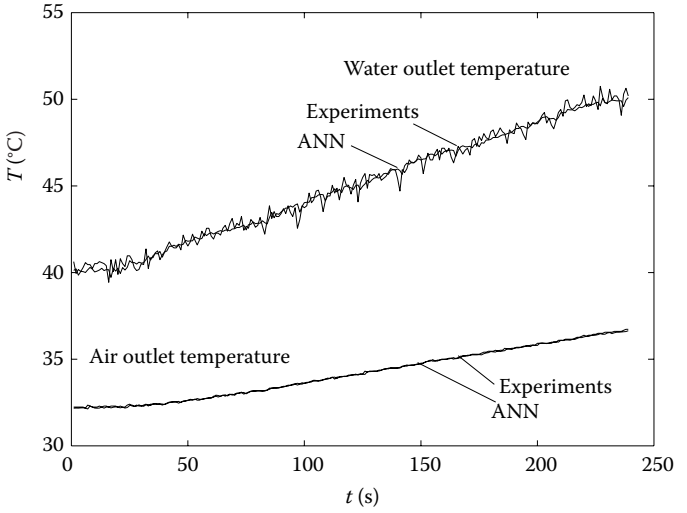


FIGURE 4.27.15 ANN dynamic prediction for ramp heating.

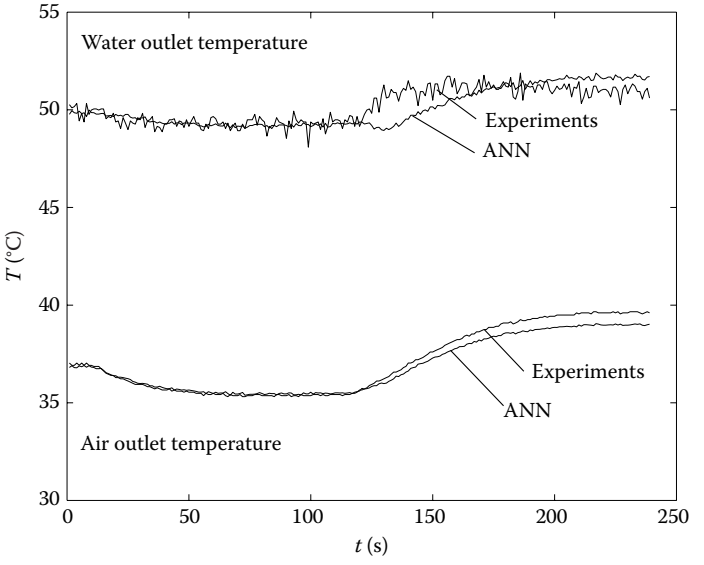


FIGURE 4.27.16 ANN dynamic model prediction for changes in the air mass flow rates.

using a plant model based on ANNs; control strategies have been developed along these lines. There is the technique of direct inverse control that utilizes an inverse system model (Hunt et al., 1992), which is then cascaded with the controlled system so that the composed system results in an identity mapping between the desired response and the controlled system output. Thus, the network acts directly as the controller in such a configuration. An extension to this strategy that incorporated an orthogonal weight estimator routine was given by Pican et al. (1998). Then there is the technique of internal model control (IMC) (Psichogios and Ungar, 1991; Hunt et al., 1992) where a forward system and inverse model are used directly as elements within the feedback loop. The system model is placed in parallel with the real plant, and the difference between the system and model outputs is used as a feedback signal, which is then processed by a controller subsystem in the forward path. In the IMC strategy, a separate linear filter is often introduced to provide desirable robustness and tracking response to the closed-loop system. The IMC has been widely applied in process-control systems.

A third dynamic-control strategy is known as predictive control in which an ANN model predicts the future plant response over the specific time horizon (Hunt et al., 1992; Ayoubi, 1997; Jeannette et al., 1998). The predictions supplied by the network are passed to a numerical optimization routine, which attempts to minimize a specified performance criterion in the calculation of a suitable control signal. One further possibility is to train another ANN to mimic the action of the optimization routine. The advantage here is that the outer loop consisting of the plant model and the optimization routine is no longer needed when training is complete, thus greatly simplifying the control system operation. Finally, mention should be made of the adaptive control strategies (Brown and Harris, 1994; Narendra and Mukhopadhyay, 1996), which have been under very active development in the past decades, and have specific important applications in the area of thermal systems.

ANNs with IMC

IMC consists of having a model of a plant M in parallel with the real system P, as shown in Figure 4.27.17. The difference between the outputs of P and M is used as the feedback for a controller C that is located in the forward path of the control scheme. The training procedure of such a control system using ANNs has two steps.

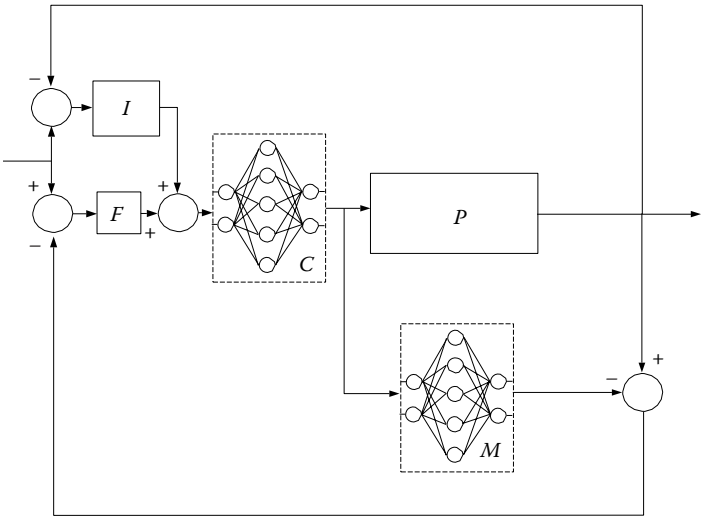


FIGURE 4.27.17 General IMC structure plus integral control.

- We first train an ANN to learn the dynamics of the process by providing known input and output data sets. This is *M*.
- Then, another ANN is trained to learn the inverse dynamics of the process and to function as a nonlinear controller *C*. It is trained to invert the model *M* instead of trying to learn the inverse dynamics of the actual process. By training in this way, we make sure that we invert the steady state gain of the model so that the offset can be eliminated.

For our experiments, we consider the heat exchanger 1 configuration described earlier, where we trained the plant model *M* with information related to $T_{\text{out}}^{\text{a}}$ and $\dot{m}_a$. These data were obtained by taking measurements of the system subject to small increments in the set point temperature. The controller *C* is obtained by using a synthetic signal, which is the desired value of the air speed. This signal is supplied to *M* to give a certain value of $T_{\text{out}}^{\text{a}}$, which is then supplied as the input to the controller. The training algorithm adjusts the weights of *C* to reduce the error between the synthetic signal and the controller output.

Since the ANNs only provide an approximation to the behavior of the actual plant, we used a one parameter filter *F*, following the suggestion of Nahas et al. (1992), preceding the controller in the forward path to account for plant–model mismatch. An integral control path *I* was also added in parallel with *F* to help obtain an offset-free controlling action. There are two constants that have to be chosen by trial and error, the first for the integral controller and the other for the filter. As a large percentage of the controllers that are currently being used correspond to proportional-integral (PI) and PID schemes, standard PI and PID controllers were used to compare the performance with the ANN controller.

Comparison with PID: Step Change in Set Point

The first test was designed to observe the performance of the controller subject to a step change in the value of the set point temperature $T_{\text{out}}^{\text{a}}$. The system was taken up to a point in which the outlet air temperature was near 32°C. The controller was turned on for 40 s until the temperature remained within a band of ±0.1°C. The set point was then increased to 36°C. Both controllers performed well and behaved in a similar way when controlling the system at large values of air speed. However, on approaching the lower end of air speeds, the system became very hard to control for two reasons. One is the effect of the delay involved, and the other is the high sensitivity of $T_{\text{out}}^{\text{a}}$ to $\dot{m}_{\text{a}}$ at low air speeds. This test brings the system from a very easy-to-control point at 32°C to a hard-to-control

state at 36°C. The results are shown in Figure 4.27.18. It is seen that although the ANN controller has a slightly larger overshoot, it presents less oscillations and it is able to bring the system to a stable condition. On the other hand, both PI and PID controllers oscillate significantly more and are not able to bring the system to a steady state, but keep T_{out}^{a} within 36 ± 0.1°C by constantly adjusting the air speed. Thus, the ANN controller uses less energy and is more stable by keeping the system steady instead of generating an oscillatory controlling action.

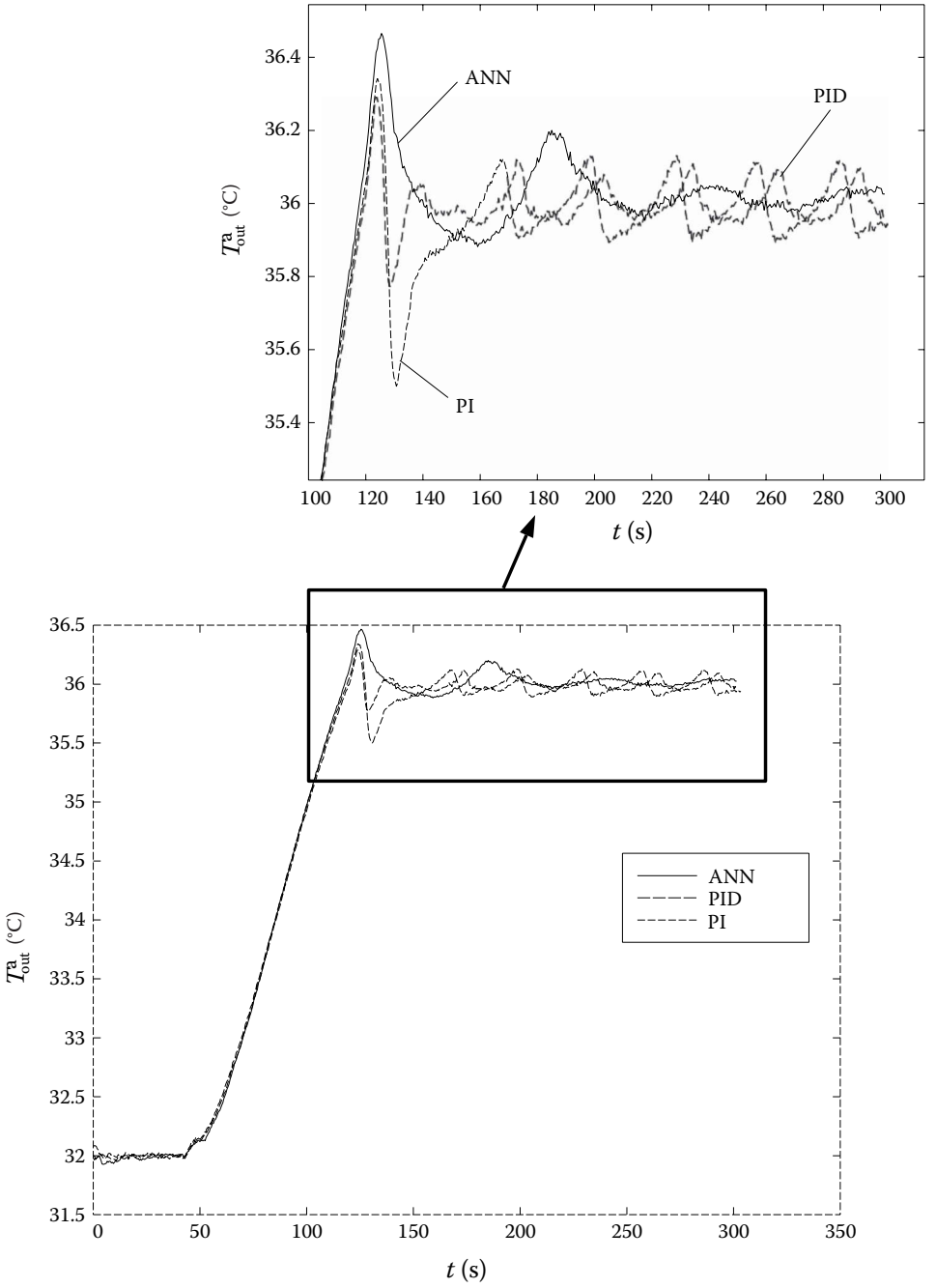


FIGURE 4.27.18 Change in the setpoint temperature. —, ANN; – –, PID; ---, PI.

Comparison with PID: Disturbance Rejection

We now analyze the disturbance rejection capabilities of the control system. In this test, a disturbance is applied to the plant in the form of a pulse in the following way. Once the system is at steady state operation, one of the valves on the water side was completely shut down for a short time. Once again, the controllers were tested at a state that was hard to control, i.e., with $T^{a}_{out} = 36°C$ and a low air speed. The PI controller showed the worst performance and is left out of the comparison shown in Figure 4.27.19. Figure 4.27.19a shows the change in the water flow rate, which is the disturbance itself; the water flow is shut down between $t=40$ s and $t=70$ s. After the disturbance pulse, the controller brings the system back to steady state. Figure 4.27.19b,c shows the change in T^{a}_{out} and $\dot{m}_a$, respectively. Once again it can be seen that the PID is not able to bring the system to a steady state condition while the oscillations of the ANN controller are quickly damped out. It can be seen in Figure 4.27.19c that the PID controller, in trying to control the temperature, generates an oscillatory air speed.

Neurocontroller as an Iterated Map

For a static ANN, the input and output variables are time-independent. In a digitally controlled time-dependent system, however, the variables are functions of time, which are sampled periodically with a constant time interval Δt. Here we are interested in a control methodology that brings the single-input–single-output thermal system to a given steady state. We are thus interested in a single variable to be controlled, $y(t)$, and a single control variable, $x(t)$, where t is time. A trained ANN can be used to predict $y^{i+1}=y((i+1)\Delta t)$ knowing $y^{i}=y(i\Delta t)$ if the system satisfies an unknown

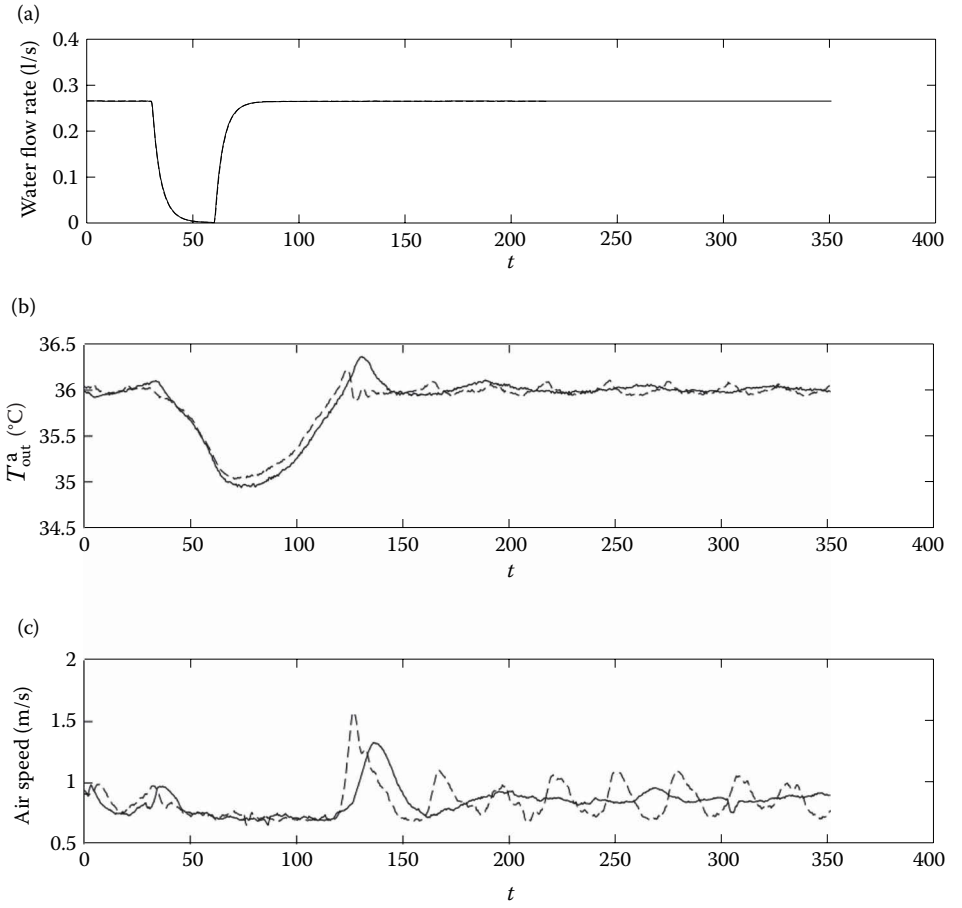


FIGURE 4.27.19 Disturbance rejection. —, ANN; – –, PID.

differential equation of first order. If the equation is of higher order, one can, in principle, include sufficient time derivatives of a vector version of $y(t)$ to reduce all scalar equations in its components to first order. In practice, this is difficult since analog signals corresponding to derivatives may not be readily available. However, a finite difference approximation of the derivatives shows that, provided Δt is small enough, this is equivalent to adding information about y at previous instants in time, i.e., y^{i-1}, y^{i-2}, etc., as many as needed. Previous values of $x^i = x(t)$ may also be needed for systems of higher order. A memory unit, which is part of the computerized control system, stores y^i and x^i to provide information on previous instants to the ANN.

A vector $\mathbf{u}$ of dimension n will be defined such that on the input side of the ANN $\mathbf{u}^i = (y^{i-n+1}, y^{i-n+2}, \ldots, y^{i-1}, y^i)$, and on the output side $\mathbf{u}^{i+1} = (y^{i-n+2}, y^{i-n+3}, \ldots, y^i, y^{i+1})$. The overall effect of the control system is that of a nonlinear map

$$\mathbf{u}^{i+1} = \mathbf{F}(\mathbf{u}^i) \tag{4.27.15}$$

For closed-loop operation, a controller generates the control variable $x(t)$. The controller is modeled by

$$x^i = g(y^i, y^{i-1}, \ldots, x^{i-1}, x^{i-2}, \ldots, y^*) \tag{4.27.16}$$

and the ANN by

$$y^{i+1} = f(y^i, y^{i-1}, \ldots, x^i, x^{i-1}, x^{i-2}, \ldots) \tag{4.27.17}$$

where y^* is a reference value of the controlled variable. Once again the map represented by Equation 4.27.15 applies where $\mathbf{u}^i = (y^{i-n+1}, y^{i-n+2}, \ldots, y^{i-1}, y^i, x^{i-m+1}, x^{i-m+2}, \ldots, x^{i-1}, x^i)$.

Consider the nonlinear map Equation 4.27.15 for open-loop control, or a similar map obtained by using Equations 4.27.16 and 4.27.17 in closed-loop control. The fixed points $\bar{\mathbf{u}}$ map to themselves and are hence solutions of $\bar{\mathbf{u}} = \mathbf{F}(\bar{\mathbf{u}})$. A nonlinear map can have more than one fixed point with different stability characteristics. To study the stability of a given fixed point, the map is linearized around it to get

$$\mathbf{u}^{i+1} - \bar{\mathbf{u}} = \mathbf{J}(\mathbf{u}^i - \bar{\mathbf{u}})$$

where $\mathbf{J}$ is the Jacobian of $\mathbf{F}$ evaluated at the fixed point. The spectral radius of $\mathbf{J}$, denoted by r, is the largest of the absolute values of its eigenvalues. If $r < 1$, the images of the map converge to the fixed point and it is stable; otherwise it is unstable (Hale and Kocak, 1991).

Closed-Loop Control

The control strategy used for closed-loop operation is based on internal model control (Hunt et al., 1992). It consists of a model M of a plant in parallel with the plant P itself, as shown in Figure 4.27.17. The difference between the outputs of P and M is used as the feedback for a controller C that is located in the forward path of the control scheme. The training procedure of such a control system using ANNs has two steps. An ANN is first trained to learn the dynamics of the process by providing known input and output data sets; this is M. Then, another network C is trained to learn the inverse dynamics of M to function as a controller. Further details of the control scheme are in Diaz et al. (2001b).

In the experiments, M was trained with information related to $y(t) = T_{\text{out}}^{\text{air}}(t)$ and $x(t) = U(t)$. The data were obtained by taking measurements of the system subject to small increments in the set point temperature. The controller C was obtained by using a synthetic signal, which is the desired U. This was supplied to M to give a y, which was the input to the controller. The training algorithm adjusted the weights of C to reduce the error between the synthetic signal and the controller output.

Since the ANNs only provide an approximation to the behavior of the actual plant, a filter F (Diaz et al., 1999a, 2001a, 2004) was used preceding the controller in the forward path to account for plant–model mismatch. An integral control path I was also added in parallel with F to help obtain an offset-free controlling action.

The controller obeys the relationship

$$x^i = g(x^{i-2}, x^{i-1}, y^{i-2}, y^{i-1}, y^i, T_{\text{ref}})$$

and the model

$$y^{i+1} = f(x^{i-2}, x^{i-1}, x^i, y^{i-2}, y^{i-1}, y^i)$$

where $x(t) = U(t)$. T_{ref} is the desired set point temperature of $T_{\text{out}}^{\text{air}}$. Thus, the map is

$$\begin{aligned} u_1^{i+1} &= u_2^i \\ u_2^{i+1} &= g(u_1^i, u_2^i, u_3^i, u_4^i, u_5^i, T_{\text{ref}}) \\ u_3^{i+1} &= u_4^i \\ u_4^{i+1} &= u_5^i \end{aligned} \tag{4.27.18}$$

$$u_5^{i+1} = f(u_1^i, u_2^i, g(u_1^i, u_2^i, u_3^i, u_4^i, u_5^i, T_{\text{ref}}), u_3^i, u_4^i, u_5^i) \tag{4.27.19}$$

where $\mathbf{u}^i = (x^{i-2}, x^{i-1}, y^{i-2}, y^{i-1}, y^i)$. Equation 4.27.18 must be computed before Equation 4.27.19. The Jacobian of this map is

$$\mathbf{J} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \mu_1 & \mu_2 & \mu_3 & \mu_4 & \mu_5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \end{bmatrix}$$

where $\mu_k = \overline{\partial g / u_k^i}$.

To illustrate the difference between stable and unstable behaviors of the system, two different controllers, C_1 and C_2 with $r < 1$ and $r > 1$ respectively, were trained. The weights for the stable controller were found using the stabilization algorithm proposed here. The reference temperature was chosen to be $T_{\text{ref}} = 34.0°\text{C}$. The flow rates and temperatures in the test facility were first adjusted to get $T_{\text{out}}^{\text{air}}$ as close to T_{ref} as possible, and then the controller was turned on. Figure 4.27.20 shows the response of the two controllers. For the stable controller C_1, $T_{\text{out}}^{\text{air}}$ oscillates a few times but eventually goes to T_{ref}. On the other hand, for otherwise identical conditions, controller C_2 takes the system to the maximum air speed, $U_{\max}$, where it remains without accomplishing the control task.

Additional Applications of ANNs in Thermal Engineering

In the past, applications of ANN to engineering problems have been attempted in structural engineering and engineering mechanics (Zeng, 1998). Tentative studies of applying ANN to problems in thermal systems have a relatively short history. With the exception of neural network systems and control systems applied to HVAC systems, the studies have been somewhat sporadic and in only some distinct areas of application. For heat transfer data analysis and correlation, an ANN-based

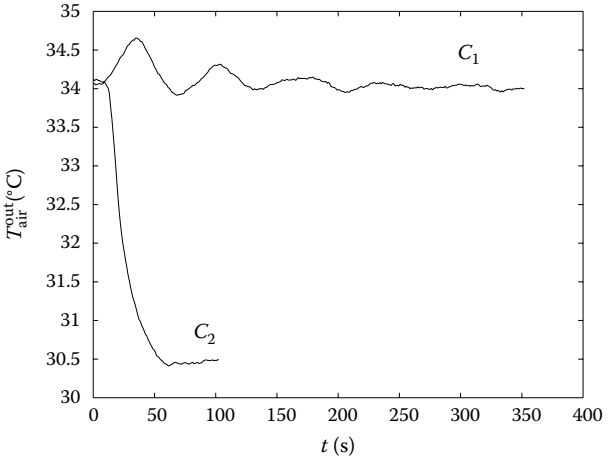


FIGURE 4.27.20 Performance of controllers C_1 and C_2, closed-loop control.

methodology was proposed by Thibault and Grandjean (1991) and a similar methodology was introduced by Jambunathan et al. (1996) to predict coefficients of heat transfer in convective-flow systems using liquid crystal thermography. Both steady and unsteady heat conduction problems were treated by ANN in studies by Gobovic and Zaghloul (1993), Yentis Jr. and Zaghloul (1994), and Kuroe and Kimura (1995). Kaminski et al. (1996) gave an interesting description of the thermal deterioration process based on a combined ANN and GA analysis. An ANN was also used to identify the location and strength of unknown heat sources by sparse temperature measurements (Momose et al., 1993). In addition, an ANN-based attempt was made to predict measured intrinsic thermodynamic properties (Normandin et al., 1993), including those of new refrigerants (Sozen et al., 2010).

ANN analyses have also been utilized to study and predict the performance of specific thermal devices and systems. Good estimates of the thermal storage loads and the dynamic system operation of typical thermal storage systems were obtained by Ferrano and Wong (1990) and Ito et al. (1995). Heat exchanger performance and control were studied by Diaz et al. (1996, 1998b, 1999b), Lavric et al. (1994, 1995), Bittanti and Piroddi (1997), Pacheco-Vega et al. (2001a,b) and, more recently, Ruiz-Mercado et al. (2009). Several industrial applications of ANN were also demonstrated, in a fluidized-bed dryer (Zbicinski et al., 1996), in a liquid-sodium reflux-pool boiler solar receiver (Fowler et al., 1997), in a steel annealing furnace (Pican et al., 1998), and in the design of a chemical injection-system retrofit fuzzy control system in a thermal power plant (Moon and Cho, 1996). Other specific applications include manufacturing and materials processing involving microelectronic manufacturing (Mahajan and Wang, 1993), a coordinate grinder (Yang et al., 1995), rapid thermal processing control (Fortuna et al., 1996), sensors and sensor analysis involving thermal image processing (Naka et al., 1993), and blast furnace probe temperatures (Bulsari and Saxin, 1995).

On the other hand, there is a coherent body of published literature dealing with the analysis and control of HVAC systems. Such systems are generally quite large and rather complex, and robust optimal control is essential in assuring their satisfactory operation. In current practice, many HVAC control processes utilize the standard PID control strategy, which, however, is not always autonomous and often requires constant manual supervision in order to achieve optimal control. Interest in the application of ANNs to the control of HVAC systems is due to its high potential of achieving an improvement. Since the appropriate theoretical dynamic model of the HVAC system is difficult to obtain, the role of ANN is to provide such a model from existing experimental data on the system operation. Several studies are available dealing with the modeling of HVAC systems, including ANN-based nonlinear HVAC model computations (Mistry and Nair, 1993), models for building energy use and HVAC system identification (Kreider et al., 1992, 1995; Chow and Teeter, 1997),

and system models for district heating (Waldemark et al., 1992). The first proof of the possibility of utilizing an ANN model as a part of the control system was given by Curtiss et al. (1993), and subsequent studies showed how the control systems including the ANN models could be implemented (Alessandri et al., 1994; Carriere et al., 1994). It is interesting to note in an experimental study how well one such ANN-based control system performed in a simple hydronic HVAC system, compared to the performance of the standard PID control (Jeannette et al., 1998).

Adams et al. (1990) optimized the protection of feedwater heaters. Leitch and Quek (1992) worked on an architecture for integrated process supervision. Kos (1993) took an ANN approach to thermal placement in power electronics. Terry and Himmelblau (1993) did data rectification and gross error detection in a steady-state process via ANNs. Huang and Nelson (1994) also applied this technique to determine the delay time for a HVAC plant to respond to control actions. Ding and Wong (1990) controlled a simulated hydronic system using an ANN. Dhar et al. (1995) used an ANN with wavelet basis functions to predict hourly heating and cooling energy use in commercial buildings with the temperature as the only weather variable. Inoue et al. (1995) studied thermal comfort and climate control. Kawashima et al. (1995) did a 24 h thermal load prediction. Li et al. (1996) developed a fault diagnosis method for heating systems using neural networks. Matsumoto et al. (1997) studied the effect of pin fin arrangement on endwall heat transfer. Takemori et al. (1991) worked on a neural network air-conditioning system for individual comfort. Dounis et al. (1992) implemented an AI technique in thermal comfort control for passive solar buildings.

All the studies mentioned here represent tentative attempts to apply the ANN analysis to thermal system problems. Since good results have been obtained so far, there is no reason why the ANN approach cannot be applied to many other thermal problems with equal success, particularly in the analysis of dynamic systems and their control.

General Discussion of ANNs

It must be emphasized that the implementation of ANN predictions in an industrial setting is really very simple and straightforward. For instance, once the network is trained for a given device by the manufacturer, it can simply be transmitted to the intended customer by giving him or her the sets of weights and biases and the corresponding software for the appropriate network configuration. The customer can then independently carry out the neural computing. An example of the weights and biases for heat exchanger 1 for the 4-2-1 ANN configuration is given in Table 4.27.5. It should be mentioned that the training process of an ANN is somewhat computing intensive, even though not excessive in terms of the computing resources available today, and all computations can be readily handled by personal workstations. However, once the training is completed, the predictions using the finalized sets of synaptic weights and biases can be carried out on any personal computer.

There are several inherent and excellent attributes of the ANN approach, which the traditional analysis, including the dimensionless correlations of the experimental data, simply does not have. The neural network, by its ability for pattern recognition, is capable of capturing all the relevant physics, but some of this physics is often not accounted for in the common data correlations by dimensionless parameters. It should be realized, too, in this regard, that since the effects of properties, which represent part of the physics, are expected to show up in the input–output relations in the training data for the ANN, there is no need to explicitly use such properties in the ANN analysis, while such use is necessary in the traditional analysis. Unfortunately, in many applications such properties, which may be either intrinsic or apparent, may not be known or not known accurately, and thus may often lead to difficulties. Furthermore, the complexity in the physics is not an issue in the ANN approach, while such complexity is often the critical barrier in the traditional analysis. Also, it is a relatively simple matter, by continuing training with new data, to adapt to changing conditions, a situation which often exists in real-world applications. The traditional analysis is not capable of doing so, since the physical mechanisms responsible for such changing conditions are not known. In addition, dynamic thermal-system models, which are needed for developing optimal

TABLE 4.27.5
Synaptic Weights and Biases of a 4-2-1 ANN Configuration for Heat Exchanger 1

i	j	k	l	$w_{i,k}^{j,l}$
1	1	2	1	–8.744
1	1	2	2	0.401
1	2	2	1	1.321
1	2	2	2	1.120
1	3	2	1	0.772
1	3	2	2	1.356
1	4	2	1	–0.303
1	4	2	2	–0.223
2	1	3	1	–7.741
2	2	3	1	8.576

i	j	$\theta_{i,j}$
2	1	–1.574
2	2	–2.474
3	1	–1.848

control strategies, are difficult to develop by the traditional analysis even for simple systems; they are readily accessible by the ANN approach. On the other hand, there are several shortcomings relative to ANN implementations, but these can be remedied by adequate numerical experimentation, past experience in similar applications, or continuing research in the field of ANN.

The attractiveness of the ANN approach has been demonstrated in the preceding subsections by considering the static and dynamic performances of typical compact heat exchangers. Such heat exchangers are already in common use today, and the traditional analysis has not yet produced satisfactory theories to predict their behavior. Even though the applications of the ANNs have been demonstrated here by considering them as examples, the potential applications in other thermal systems are essentially unlimited—thermal systems still have not been analyzed by the ANN approach to any large extent. In order to promote its application to thermal science and engineering problems, it would be desirable to look at application areas in which the ANN approach would make significant impacts. In this regard, the most significant areas lie where the lack of physical models, system complexities, and unavailability of properties or constitutive relations represent critical barriers to the advancement of the field in terms of product development, process development and control, manufacturing, sensitivity and risk analyses, and other similar areas which involve simulation, design, and control. This is particularly imperative in those areas where there are plenty of experimental data that can be used to train the networks. To give some indication to this vast potential, several examples of critical technology areas can be mentioned for the applications of the ANN approach.

Thermal Networks

The focus so far has been on components, especially on heat exchangers. However, even if the behavior of a single heat exchanger could be accurately predicted, it would be difficult to do so for a large network of these devices as are commonly used in practice. Thermal networks, as used in heating and cooling systems in a single building or in a collection of buildings, may consist of chilled water for cooling purposes, steam for heating, air for ventilation or, more commonly, a combination of these. Current design is heavily based on overspecification so that it is not unusual to find simultaneous heating and cooling to obtain the desired comfort level, especially during periods

of change in weather. A hydraulic network may consist of a large number of interconnected pipes with valves, heat exchangers, pumps, and other components online. The simplest network problem consists in determining the flow in each branch of the network. Methods based on the solution of the nonlinear pressure drop and flow rate network equations have been developed to solve this problem. The control of the flow in a given branch by change of pump or valve parameters is a more difficult problem. Though one can in principle resolve the network for the new conditions, this could be time-consuming for a large network and, thus, difficult to use in real time. In addition the exact, current state of the network and its physical properties may not be known for computational purposes. Thus, for example, the heat transfer coefficients at a heat exchanger will be affected by fouling, which will in turn change its ability to transfer heat. It is possible that the response of the network be simulated by an ANN. This ANN can be trained by using experimental data so that the response of changing any parameter can be found. One of the advantages is that this can be done online so that changes in system performance are reflected in the simulation.

Thermal Processing of Materials

In almost all materials processing operations involving thermal effects, the most critical information is the direct relationship between the desired characteristics of the end product, such as, for instance, material microstructure defect elimination and the processing parameters. Physical models that provide such relationship are either not available or, if available, involve either apparent properties or constitutive relations, which are not available. In the latter case, the lack of such information essentially nullifies the usage of such models. The ANN analysis simply bypasses all these uncertainties and leads immediately to the direct quantitative relationship between the desired end-product characteristics and the processing parameters. This advantage becomes particularly significant in dealing with new engineering materials, foods, and pharmaceutical and biological materials.

Environmental and Safety Studies

There are many complex thermal systems dealing with environmental and safety control in buildings, factories, mines, and the like. HVAC and indoor pollution systems and fire protection systems are good examples. Despite the recent advances made in the control of HVAC systems by means of the ANN approach, much more needs to be done, especially in the area of simulating realistic dynamic conditions involving, for instance, time delays. Performance-oriented risk analysis of fire hazards is drawing much attention, but very little is known concerning the realistic assessment of likely dynamic fire-and smoke-spread scenarios, which are needed in risk analysis and also in the design of fire mitigation systems for enclosed habitat spaces such as high rises and mines. The capability of the ANN approach needs to be explored and exploited to deal with such complex dynamic systems.

Tribological Systems

In all systems of any size involving relative motions of components in contact, friction inevitably occurs and produces heat, wear, or both. To minimize or control friction, theories or models are needed for the physical, chemical, and other phenomena occurring at the surfaces of contact or interfaces. Such models would necessarily depend on the physical structure and properties at these interfaces. Such theories or models are either not available or difficult to apply because of the unknown characterization of these interfaces, even though the methods of preparing such interfaces and the lubricant used are known. ANNs, in general, are capable of catching all the relevant physics and recognizing the unique surface characterization, and then incorporating such characterizations into the ANN-based input–output relation.

Reactive Systems

One major barrier to our knowledge of reacting systems from molecular–biological systems in single cells to large scale reactors to combustion phenomena is the lack of reaction kinetics and the corresponding energy transport information. Models are impossible without such information. For the ANN

analysis, such information can again be bypassed and, therefore, is not needed in the analysis. The effects of the reaction kinetics are reflected in the output data used in the training of the ANN and, therefore, they are accounted for in the trained ANN, which then becomes the model of the reactive system.

Multiphase Phenomena and Systems

In systems involving components of multiple phases such as mixtures of liquids and gases, liquids and solids, solids and gases, binary gas mixtures (Diaz and Campo, 2009) and all three phases in turbulent pipe flows with heat transfer and multicomponent heterogeneous mixtures in food processing, the extreme complexity prevents the development of reasonable models for use in simulation, optimization, and control. Attempts to correlate the data would not be particularly fruitful either because of the large data scatter due to unknown parameters. This type of complex system is a natural for ANN-based analysis, modeling, and dynamic control, as recently demonstrated by the application of ANNs to predict the convective heat transfer in evaporative processes and small-scale units (Romero-Mendez et al., 2014, 2016).

Inverse Problems and Parameter Estimation

The importance of solving inverse problems and problems in parameter estimation has long been recognized in many engineering applications, and this importance has also received ever-increasing attention in the field of thermal systems. The application of ANN analysis to such problems is not new, but should be emphasized. As is already obvious, the inputs and outputs of an ANN, especially in dynamic systems, can be switched as in the case of an inverse ANN model.

Genetic Algorithms (GAs)

Evolutionary programming, of which genetic algorithms and programming are examples, allow programs to change or evolve as they compute. GAs, specifically, are based on the principle of Darwinian selection. One of their most important applications in the thermal sciences is in the area of optimization of various kinds.

Optimization by itself is fundamental to many applications. In engineering, for example, it is important to the design of systems; analysis permits the prediction of the behavior of a given system, but optimization is the technique that searches among all possible designs of the system to find the one that is the best for the application. The importance of this problem has given rise to a wide variety of techniques that help search for the optimum solution. There are searches that are gradient-based and those that are not. In the former, the search for the optimum solution, as for example, the maximum of a function of many variables, starts from some point and directs itself in an incremental fashion towards the optimum; at each stage, the gradient of the function surface determines the direction of the search. Local optima can be found in this way, the search for global optimum being more difficult. Again, if one visualizes a multivariable function, it can have many peaks, any one of which can be approached by a hill-climbing algorithm. To find the highest of these peaks, the entire domain has to be searched; the narrower this peak the finer the searching "comb" must be. For many applications, this brute force approach is too expensive in terms of computational time. Alternatives, like simulated annealing, are techniques that have been proposed, and the GA is one of them.

In what follows, we will provide an overview of genetic algorithm and programming. A numerical example will be explained in some detail. The methodology will be applied to one of the heat exchangers discussed earlier. There will be a discussion on other applications in thermal engineering and comments will be made on potential uses in the future.

Methodology of GAs

GAs are discussed in detail by Holland (1975, 1992), Mitchell (1997), Goldberg (1989), Michalewicz (1992), and Chipperfield (1997). One of the principal advantages of this method is its ability to pick

out a global extremum in a problem with multiple local extrema. For example, we can discuss finding the maximum of a function $f(x)$ in a given domain $a < x < b$. In outline, the steps of the procedure are the following.

- First, an initial population of n members, $x_1, x_2, \ldots, x_n \in [a, b]$ is randomly generated.
- Then, for each x, a fitness is evaluated. The fitness, or effectiveness, is the parameter that determines how good the current x is in terms of being close to an optimum. Clearly, in this case, the fitness is the function $f(x)$ itself, since the higher the value of $f(x)$, the closer we are to the maximum.
- The probability distribution for the next generation is found based on the fitness values of each member of the population. Pairs of parents are then selected on the basis of this distribution.
- The offsprings of these parents are found by crossover and mutation. In crossover, two numbers in binary representation, for example, produce two others by interchanging part of their bits. After this, and based on a preselected probability, some bits are randomly changed from 0 to 1 or vice versa. Crossover and mutation create a new generation with a population that is more likely to be fitter than the previous generation.
- The process is continued as long as desired or until the largest fitness in a generation does not change much any more.

The procedure can be easily generalized to a function of many variables.

Let us consider a numerical example that is shown in detail in Table 4.27.6. Suppose that one has to find the x at which $f(x)=x(1-x)$ is globally a maximum between 0 and 1. We have taken $n=6$, meaning that each generation will have six numbers. Thus, for a start, six random numbers are selected between 0 and 1. Now we choose n_b,which is the number of bits used to represent a number in binary form. Taking $n_b=5$, we can write the numbers in binary form normalized between 0 and the largest number possible for n_b bits, which is $2^{n_b-1} = 31$. In one run, the numbers chosen and written down in the first column of the table labeled $G=0$, are 25, 30, 28, 19, 3, and 1, respectively. The fitnesses of each one of the numbers, i.e., $f(x)$, are computed and shown in the second column. These values are normalized by their sum and shown in the third column as $s(x)$. The normalized fitnesses are drawn on a roulette wheel in Figure 4.27.21. The probability of crossover is taken to be 100%, meaning that crossover will always occur. Pairs of numbers are chosen by spinning the wheel, the numbers having a bigger piece of the wheel having a larger probability of being selected. This produces column four marked $G=1/4$, and shuffling to produce random pairing gives column five, marked $G=1/2$. The numbers are now split up in pairs, and crossover applied to each pair. The

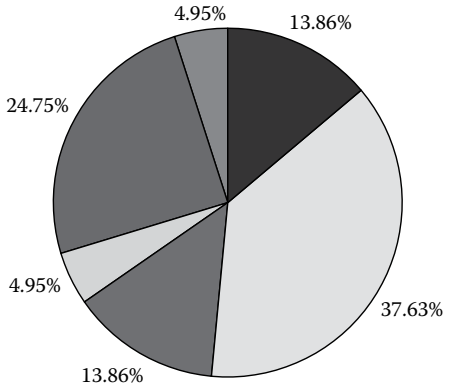


FIGURE 4.27.21 Distribution of fitnesses.

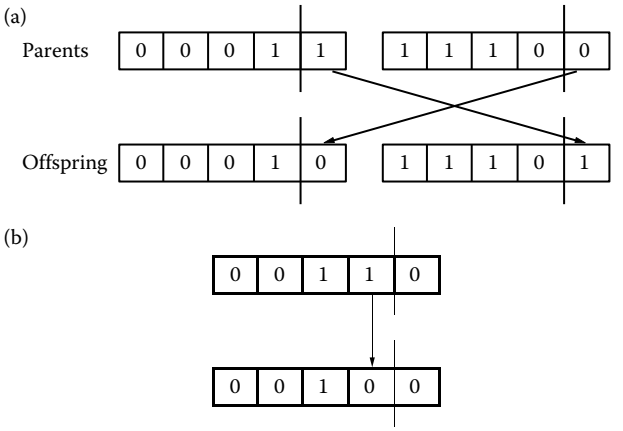


FIGURE 4.27.22 (a) Crossover and (b) mutation in a genetic algorithm.

TABLE 4.27.6
Example Using the Genetic Algorithm

G=0	*f*(*x*)	*s*(*x*)	*G*=1/4	*G*=1/2	*G*=3/4	*G*=1
11001	0.1561	0.2475	00011	00011	00010	00010
11110	0.0312	0.0496	00011	11100	11101	11101
11100	0.0874	0.1386	11110	00011	10011	10011
10011	0.2373	0.3762	10011	10011	00011	00011
00011	0.0874	0.1386	00011	11110	11011	11011
00001	0.0312	0.0495	11100	00011	00110	00100

first pair [0 0 0 1 1] and [1 1 1 0 0] produces [0 0 0 1 0] and [1 1 1 0 1]. This is illustrated in Figure 4.27.22a, where the crossover position is between the fourth and fifth bit; the bits to the right of this line are interchanged. Crossover positions in the other pairs are randomly selected. Crossover produces column six, marked as $G=3/4$. Finally, one of the numbers, in this case the last number in the list [0 0 1 1 0], is mutated to [0 0 1 0 0] by changing one randomly selected bit from 1 to 0, as shown in Figure 4.27.22b. From the numbers in generation $G=0$, these steps have now produced a new generation $G=1$. The process is repeated until the largest fitness in each generation increases no more. In this particular case, values within 3.22% of the exact value of x for maximum $f(x)$, which is the best that can be done using 5 bits, were usually obtained within 10 generations.

Applications of GAs to Compact Heat Exchangers

The following analysis is based on the data collected on a single-row heat exchanger referred to earlier as heat exchanger 1. In the following, a set of $N=214$ experimental runs provided the database. The heat rate is determined by

$$\dot{Q} = \dot{m}_a c_{p,a} \left(T_a^{out} - T_a^{in} \right) \tag{4.27.20}$$

$$= \dot{m}_w c_w \left(T_w^{in} - T_w^{out} \right) \tag{4.27.21}$$

For prediction purposes, we will use functions of this type

$$\dot{Q} = \dot{Q}(T_w^{in}, T_a^{in}, \dot{m}_a, \dot{m}_w)$$

The conventional way of correlating data is to determine correlations for inner and outer heat transfer coefficients. For example, power laws of the following form

$$\varepsilon \mathrm{Nu}_\mathrm{a} = a\mathrm{Re}_\mathrm{a}^\mathrm{m}\mathrm{Pr}_\mathrm{a}^{1/3} \tag{4.27.22}$$

$$\mathrm{Nu}_\mathrm{w} = b\mathrm{Re}_\mathrm{w}^\mathrm{n}\mathrm{Pr}_\mathrm{w}^{0.3} \tag{4.27.23}$$

are common. The two Nusselt numbers provide the heat transfer coefficients on each side and the overall heat transfer coefficient, U, is related to h_a and h_w by

$$\frac{1}{UA_\mathrm{a}} = \frac{1}{h_\mathrm{w}A_\mathrm{w}} + \frac{1}{\varepsilon h_\mathrm{a}A_\mathrm{a}}$$

To find the constants a, b, m, n, the mean square error

$$S_U = \frac{1}{N}\sum\left(\frac{1}{U^\mathrm{p}} - \frac{1}{U^\mathrm{e}}\right)^2 \tag{4.27.24}$$

must be minimized, where N is the number of experimental data sets, U^p is the prediction made by the power-law correlation, and U^e is the experimental value for that run. The sum is over all N runs.

This procedure was carried out for the data collected. It was found that S_U had local minima for many different sets of the constants, the following two being examples.

Correlation	*a*	*b*	*m*	*n*
A	0.1018	0.0299	0.591	0.787
B	0.0910	0.0916	0.626	0.631

Figure 4.27.23 shows a section of the S_U surface that passes though the two minima, *A* and *B*. The coordinate z is a linear combination of the constants a, b, m, and n such that it is zero and unity at the two minima. Though the values of S_U for the two correlations are very similar and the heat rate

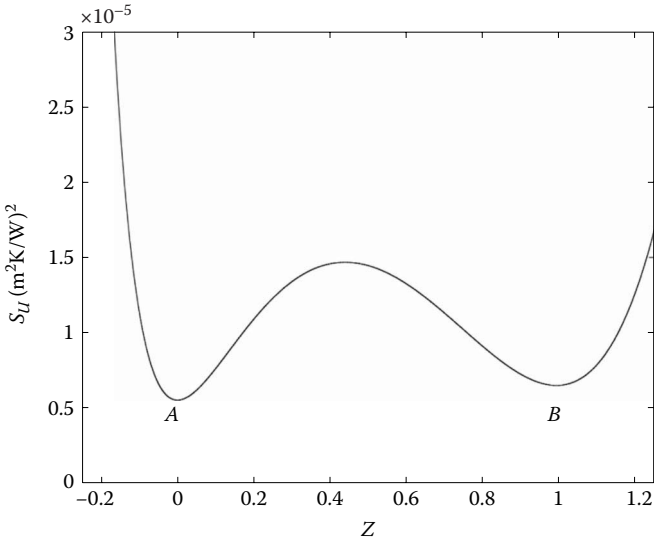


FIGURE 4.27.23 Section of $S_\mathrm{U}(a, b, m, n)$ surface.

predictions for the two correlations are also almost equally accurate, the predictions on the thermal resistances on either side are different. Figure 4.27.24 shows the ratio of the predicted air- and waterside Nusselt numbers using these two correlations. R_a is the ratio of the Nusselt number of the airside predicted by Correlation *A* divided by that predicted by Correlation *B*. R_w is the same value for the waterside. The predictions, particularly the one on the water side, are very different.

There are several reasons for this multiplicity of minima of S_U. Experimentally, it is very difficult to measure the temperature at the wall separating the two fluids, or even to specify where it should be measured, and mathematically, it is due to the nonlinearity of the function to be minimized. This raises the question as to which of the local minima is the "correct" one. A possible conclusion is that the one which gives the smallest value of the function should be used. This leads to the search for the global minimum, which can be done using the GA.

For these data, Pacheco-Vega et al. (1998) conducted a global search among a proposed set of heat transfer correlations using the GA. The experimentally determined heat rate of the heat exchanger was correlated with the flow rates and input temperatures, with all values being normalized. To reduce the number of possibilities, the total thermal resistance was correlated with the mass flow rates in the form

$$\frac{T_w^{in} - T_a^{in}}{\dot{Q}} = f(\dot{m}_a, \dot{m}_w) \tag{4.27.25}$$

The functions that were used are indicated in Table 4.27.7. The GA was used to seek the values of the constants associated with each correlation, the objective being to minimize the variance.

$$S_Q = \frac{1}{N}\sum\left(\dot{Q}^p - \dot{Q}^e\right)^2 \tag{4.27.26}$$

where the sum is over all *N* runs between the predictions of a correlation, $\dot{Q}^p$, and the actual experimental values, $\dot{Q}^e$. Since the unknowns are the set of constants, *a*, *b*, *c*, and sometimes *d*, a single binary string represents them; the first part of the string is *a*, the next is *b*, and so on. The rest of the GA is as in the numerical example given before. The results obtained for each correlation are also summarized in the table in descending order of S_Q. The last column shows the mean square error σ

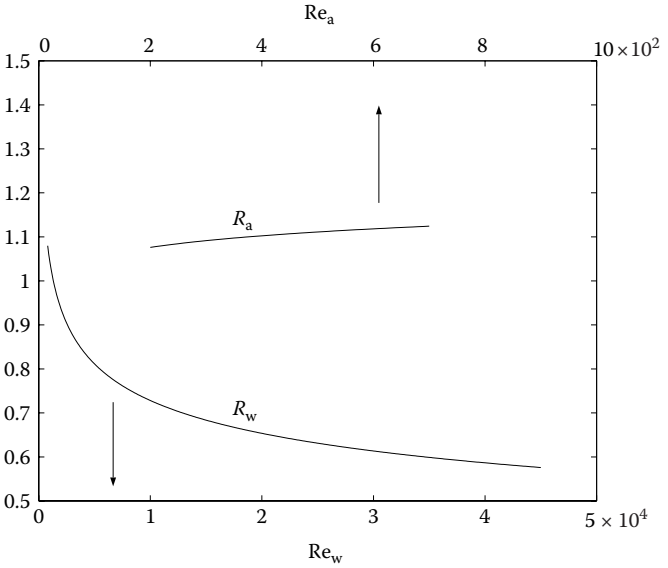


FIGURE 4.27.24 Ratio of the predicted air- and water-side Nusselt numbers.

TABLE 4.27.7
Comparison of Best Fits for Different Correlations

Correlation	*f*	*a*	*b*	*c*	*d*	σ
Power law	$a\dot{m}_w^{-b} + c\dot{m}_a^{-d}$	0.1875	0.9997	0.5722	0.5847	0.0252
Inverse linear	$(a+b\dot{m}_w)^{-1} + (c+d\dot{m}_a)^{-1}$	–0.0171	5.3946	0.4414	1.3666	0.0326
Inverse exponential	$(a+e^{b\dot{m}_w})^{-1} + (c+e^{d\dot{m}_a})^{-1}$	–0.9276	3.8522	–0.4476	0.6097	0.0575
Exponential	$ae^{-b\dot{m}_w} + ce^{-d\dot{m}_a}$	3.4367	6.8201	1.7347	0.8398	0.0894
Inverse quadratic	$(a+b\dot{m}_w^2)^{-1} + (c+d\dot{m}_a^2)^{-1}$	0.2891	20.3781	0.7159	0.7578	0.0859
Inverse logarithmic	$(a+b\ln\dot{m}_w)^{-1} + (c+d\ln\dot{m}_a)^{-1}$	0.4050	0.0625	–0.5603	0.2048	0.1165
Logarithmic	$a - b\ln\dot{m}_w - c\ln\dot{m}_a$	0.6875	0.4714	0.4902	—	0.1664
Linear	$a - b\dot{m}_w - c\dot{m}_a$	2.3087	0.8533	0.8218	—	0.2118
Quadratic	$a - b\dot{m}_w^2 - c\dot{m}_a^2$	1.8229	0.6156	0.5937	—	0.2468

defined in a manner similar to Equations 4.27.6 and 4.27.7. The parameters used for the computations are population size 20, number of generations 1000, bits for each variable 30, probability of crossover 1, and probability of mutation 0.03.

Some correlations are clearly seen to be superior to others. However, the difference in S_Q between the first- and second-place correlations, the power law, and inverse logarithmic, which have mean errors of 2.5% and 3.3%, respectively, is only about 8%, indicating that either could do just as well in predictions even though their functional forms are very different. In fact, the mean error in many of the correlations is quite acceptable. Figure 4.27.25 shows the predictions of the power-law correlation vs. the experimental values, all in normalized variables. The prediction is seen to be very good. The quadratic correlation on the other hand, is the worst in the set of correlations considered, and Figure 4.27.26 shows its predictions. It must also be remarked that, because of the random

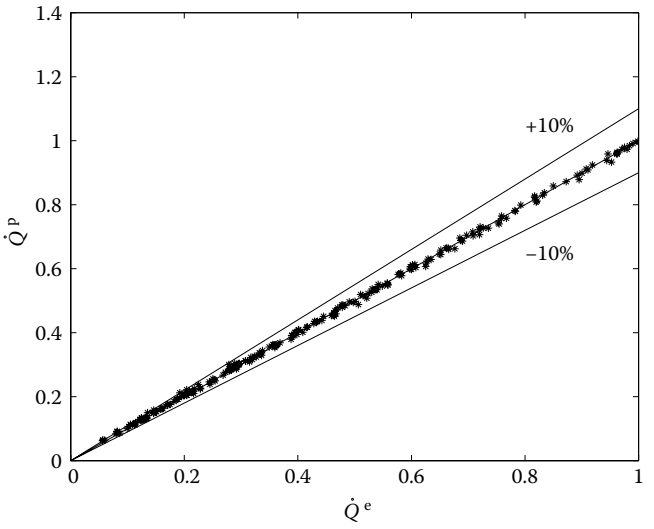


FIGURE 4.27.25 Experimental vs. predicted normalized heat flow rates for a power-law correlation. The straight line is the line of equality between prediction and experiment, and the broken lines are ±10%.

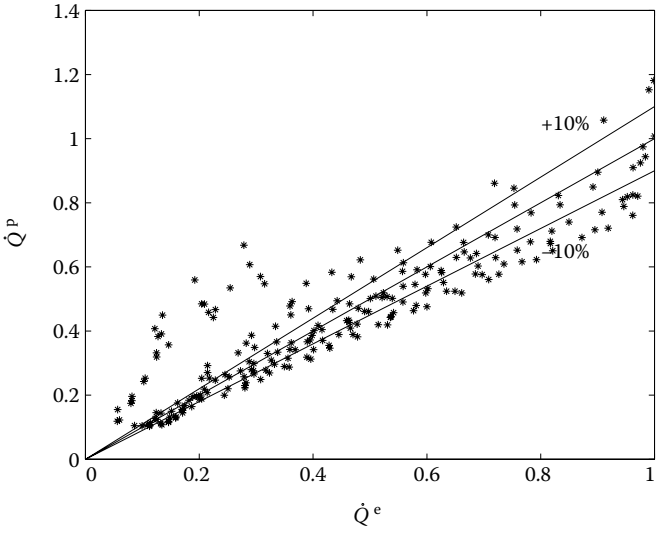


FIGURE 4.27.26 Experimental vs. predicted normalized heat flow rates for a quadratic correlation. The straight line is the line of equality between prediction and experiment, and the broken lines are ±10%.

numbers used in the procedure, the computer program gives slightly different results each time it is run, changing the lineup of the less appropriate correlations somewhat.

Additional Applications of GAs in Thermal Engineering

Though the GA is a relatively new technique in relation to its application to thermal engineering, there are a number of different applications that have already been successful. Davalos and Rubinsky (1996) adopted an evolutionary genetic approach for numerical heat-transfer computations. Shape optimization is another area that has been developed. Fabbri (1997) used a GA to determine the optimum shape of a fin. The two-dimensional temperature distribution for a given fin shape was found using a finite-element method. The fin shape was proposed as a polynomial, the coefficients of which have to be calculated. The fin was optimized for polynomials of degree 1 through 5. von Wolfersdorf et al. (1997) did shape optimization of cooling channels using GAs. The design procedure is inherently an optimization process. Androulakis and Venkatasubramanian (1991) developed a methodology for design and optimization that was applied to heat exchanger networks; the proposed algorithm was able to find local solutions where gradient-based methods failed. Abdel-Magid and Dawoud (1995) optimized the parameters of an integral and a proportional-plus-integral controller of a reheat thermal system with GAs. The fact that the GAs can be used to optimize in the presence of variables that take on discrete values was put to advantage by Schmitt et al. (1996) who used it for the design of a compact high-intensity cooler.

The placing of electronic components as heat sources is a problem that has become very important recently from the point of view of computers. Queipo et al. (1994) applied GAs to the optimized cooling of electronic components. Tang and Carothers (1996) showed that the GA worked better than some other methods for the optimum placement of chips. Queipo and Gil (1997) worked on the multiobjective optimization of component placement and presented a solution methodology for the collocation of convectively and conductively air-cooled electronic components on planar printed wiring boards. Meysenc et al. (1997) studied the optimization of microchannels for the cooling of high-power transistors. Inverse problems may also involve the optimization of the solution. Allred and Kelly (1992) modified the GA for extracting thermal profiles from infrared image data that can be useful for the detection of malfunctioning electronic components. Jones et al. (1995)

used thermal tomographic methods for the detection of inhomogeneities in materials by finding local variations in the thermal conductivity. Raudensky et al. (1995) used the GA in the solution of inverse heat conduction problems. Okamoto et al. (1996) reconstructed a three-dimensional density distribution from limited projection images with the GA. Wood (1996) studied an inverse thermal field problem based on noisy measurements and compared a GA and the sequential function specification method. Li and Yang (1997) used a GA for inverse radiation problems. Castrogiovanni and Sforza (1996, 1997) studied high-heat flux-flow boiling systems using a numerical method in which the boiling-induced turbulent eddy diffusivity term was used with an adaptive GA closure scheme to predict the partial nucleate boiling regime.

Genetic Programming

The GP technique (Koza, 1992, 1994) is an extension of the GA procedure in which computer codes take the place of numbers. Developed originally to automatically generate computer programs, in the context of thermal engineering applications, it can be used, for example, in symbolic regression to search within a set of functions for the one which best fits experimental data. The procedure is similar to that for the GA, except for the crossover operation. If each function is represented in tree form, though not necessarily of the same length, crossover can be achieved by cutting and grafting. A typical representation of the function $5x\cos(5x+1)$ is illustrated, for example, in Figure 4.27.27a, while Figure 4.27.27b, shows the result of the crossover operation on the two functions $5x(x+1)$ and $x(5x+1)$ to give $5x(5x+1)$ and $x(x+1)$. The crossover points may be different for each parent. In general, as shown in the figures, branch nodes may be mathematical operators (such as sin, cos, exp, log, +, −, *, /, ∧), or may be Boolean (such as AND, OR, and NOT) or conditional (IF-THEN-ELSE, etc.) operators. Leaf or terminal nodes x_j, $j=1, \ldots, N_v$, on the other hand, are the variables in a particular problem, or constants $\mathbf{C} \in \mathbb{R}$ to be determined as part of the solution.

Computer Representation

A key element in GP is the representation of functional forms conserving a correct syntax, since both the crossover and mutation operations directly lean on it. Such representation depends on the programming language being used for its coding. For instance, due to the natural way of portraying these tree structures, GP was ordinarily coded in LISP (Koza, 1992; Sette and Boullart, 2001). However, to increase speed in the computations, other object-oriented programming languages like C or C++ have also been used. Recently, Cai et al. (2006) introduced a GP algorithm coded in MATLAB by taking advantage of the fact that MATLAB enables handling different data types in a

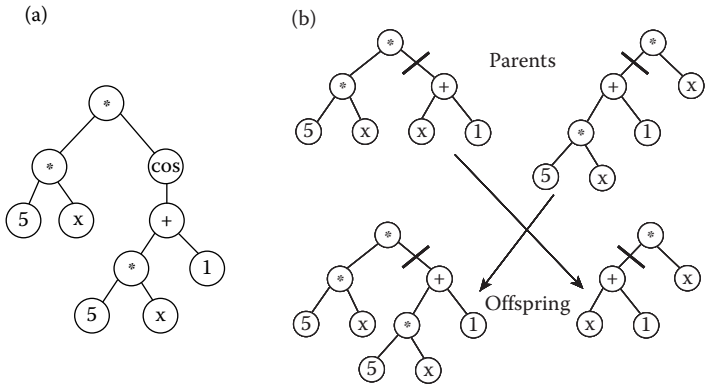


FIGURE 4.27.27 Function representation as (a) parse tree and (b) crossover in GP. (a) Function is $5x\cos(5x+1)$. (b) Parents are $5x(x+1)$ and $x(5x+1)$, and offspring are $5x(5x+1)$ and $x(x+1)$.

straightforward manner, such that the syntactically-correct crossover operation can be achieved. In their code, trees representing the functions are stored as rectangular arrays, an example of which, for the crossover of functions $5x(x+1)$ and $x(5x+1)$ considered before, is illustrated in Figure 4.27.28. An important advantage of using arrays is that the "size" L_f, of the function can be easily computed as the product of the number of rows and the number of columns of the smallest rectangular layout, which, for symbolic regression in a variety of applications is key as an objective may be to find the smallest heat transfer correlation that provides the highest accuracy. As an example, the size of the functions mentioned earlier is, respectively, $L_f=4\times3=12$ and $L_f=4\times4=16$.

Fitness and Penalty

If GP is used to perform symbolic regression, then the objective is to minimize the variance of the error between predictions and the data. Thus, the fitness is naturally defined as the reciprocal of the variance

$$F_f = \left(\frac{1}{N} \sum_{i=1}^{N} \left[f^t(x_j)_i - f^p(x_j, \mathbf{C})_i \right]^2 \right)^{-1}$$

between the target data $f^t(x_j)$ for $i=1, \ldots, N$, and the values predicted by the candidate functions $f^p(x_j)$ for $i=1, \ldots, N$. In heat transfer applications, f is either a transfer conductance or a dimensionless form of the heat transfer coefficients. A potential problem with using the fitness directly as defined here, however, is that it may generate functions of complicated forms with many terms. To prevent large functions and favor more compact ones, the fitness may be penalized according to their corresponding size. Although this can be done in many ways, McKay et al. (1997) has proposed to use a logistic sigmoid penalty function $g(L_f)$, with prescribed parameters a_1 and a_2, to define a penalized fitness Q_f, as

$$Q_f = F_f g(L_f) \tag{4.27.27}$$

$$g(L_f) = \frac{1}{1+\exp[a_1(L_f - a_2)]} \tag{4.27.28}$$

The effect that a_1 and a_2 have on the results from GP is illustrated in a later section.

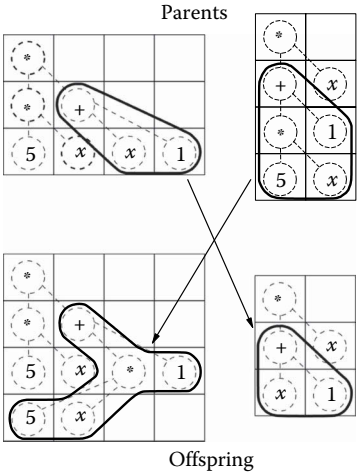


FIGURE 4.27.28 Crossover of functions represented as arrays: parents are $5x(x+1)$ and $x(5x+1)$, and offspring are $5x(5x+1)$ and $x(x+1)$.

Methodology of GP

The steps of the procedure are the following.

- First, an initial set of M correlations coded as parse trees is randomly generated from the chosen sets of functional operators and terminals. For the subsequent generations, the old population is used.
- Then, for each f, the value of the fitness is calculated.
- The probability distribution for the next generation, on the basis of which parents are chosen for replacement, is calculated from the fitness values.
- As in the GA procedure, once the parents are selected, crossover and mutation are applied according to preselected probabilities p_c and p_m, respectively.
- Once the optimum functional form is determined, the final step of each iteration in GP is the application of the GA to seek for the optimum values constants in it.
- Once crossover and mutation have been applied to the complete population, and the optimum set of constants are determined, a new population that keeps the fittest member of the previous generation is created.
- The process is repeated until some criterion based on convergence or maximum number of generations, G_{max}, is achieved.

Since GP is a probabilistic technique, every run gives a slightly different answer. To understand this inherent variation, it is useful to perform multiple runs along with a statistical analysis of the results.

Applications of GP to Compact Heat Exchangers

The procedure is illustrated by its application to experimental data obtained from tests on the heat exchanger 2 described in an earlier section. Though the data included operating conditions for which either dry surface, dropwise or film condensation would occur, only the dry-surface data were considered by Cai et al. (2006) for the analysis. From the measurements, the correlation originally proposed by McQuiston (1978a) is given in Equation 4.27.10, with $f_s(\delta)=1$ in Equation 4.27.12; i.e.,

$$j_s = 0.0014 + 0.2618\mathrm{Re}_D^{-0.4}\left(\frac{4}{\pi}\frac{x_a}{D_h}\frac{x_b}{D}\right)^{-0.15} \tag{4.27.29}$$

where j is the sensible Colburn j factor, defined earlier as $j_s = h_a \mathrm{Pr}_a^{2/3} / G_c c_{p,a}$, Re_D is Reynolds number, and the term in parenthesis is a nondimensional geometrical parameter representing an airside area ratio.

To perform symbolic regression for these data, the fitness function was defined as

$$F_j = \left(S_j\right)^{-1} = \left(\frac{1}{N}\sum_{i=1}^{N}\left(j_{s,i}^{e} - j_{s,i}^{p}\right)^2\right)^{-1}$$

where $j_{s,i}^{e}$, $i=1, \ldots, N$, are the measurements and $j_{s,i}^{p}$, $i=1, \ldots, N$, are the predicted values from each of the M correlations in the population. The algorithm seeks a correlation function and the corresponding constants that maximize F_j (or minimize S_j), with $M=100$, $G_{max}=800$, $p_c=0.8$, and $p_m=0.2$; $x_1=\mathrm{Re}_D$, and $x_2 = \frac{4}{\pi}\frac{x_a}{D_h}\frac{x_b}{D}$ are the variables in the terminal set. To build Q_j as in Equation 4.27.27, the values used in Equation 4.27.28 are $a_1=0.2$ and $a_2=30$.

A typical evolution of the algorithm is shown in Figure 4.27.29, where the two curves correspond to the values of unpenalized and penalized fitnesses, F_j and Q_j respectively, obtained from the best correlation found in each generation. The figure shows that because Q_j rejects good candidate

correlations that are large in size, the values of F_j are larger than those of Q_j. In addition, though both curves follow similar paths, only that of Q_j keeps increasing as the best correlation from the previous generation is based on Q_j, rather than F_j; this means that a correlation that has a large F_j but is also large in size, may not be preserved for the subsequent iteration.

The following are two examples of the correlations that result from the procedure:

$$j_s = \frac{1.82}{103.81 + 0.0299\ \mathrm{Re_D} + \frac{4}{\pi}\frac{x_a}{D_h}\frac{x_b}{D}}$$

$$j_s = \frac{66.39 - \frac{1.7824}{\pi}\frac{x_a}{D_h}\frac{x_b}{D}}{3881.61 + \mathrm{Re_D}}$$

Though these correlations are different in form, and their sizes, L_j, are 10 and 16 respectively, their RMS errors in predictions of the j factor are less than 0.1% (6.32% and 6.24%), indicating that either could be used. This multiplicity of solutions in functional space, similar to that in the parameter space of an assumed form (Pacheco-Vega et al., 2001a), was also noticed by Pacheco-Vega et al. (2003a) using artificial data. After several runs of the procedure, a slightly more complex correlation, with L_j=24, that provides better predictions is

$$j_s = \frac{2205.32}{1.39\times10^5 + 24.16\ \mathrm{Re_D} + \left(\frac{4}{\pi}\frac{x_a}{D_h}\frac{x_b}{D}\right)\mathrm{Re_D}} \tag{4.27.30}$$

A comparison of the RMS percentage error in j_s from the correlations obtained with GP against that developed by Pacheco-Vega et al. (2003b) using global regression with the same functional form as that of Equation 4.27.29, and against Equation 4.27.29 itself, is shown in Table 4.27.8 in descending order of magnitude. From the results, it is possible to observe that regardless of differences in accuracy, all the correlations found by the GP method give a smaller error than that of Equation 4.27.29. On the other hand, the table illustrates that even though the global-regression-based correlation is the best possible that can be obtained from the assumed functional form, Equation 4.27.30 is seen to give a slightly smaller error.

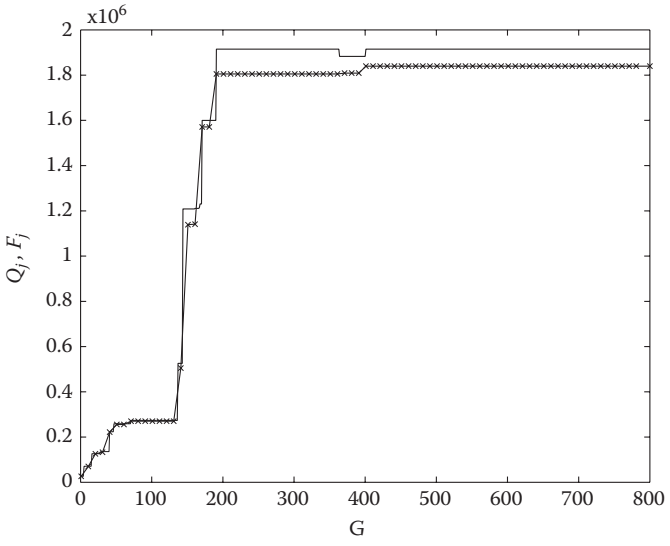


FIGURE 4.27.29 Evolution of fitness and penalized fitness with generation number. (–) fitness function F_j, (–×–) penalized fitness Q_j.

TABLE 4.27.8
Comparison of RMS Errors in j_s for Heat Exchanger 2

Prediction Method	Error (%)
Equation 4.27.29 McQuiston	14.74
Pacheco-Vega et al. (2003b)	6.21
Equation 4.27.30	6.18

A comparison between the experimentally-determined j factor and that predicted from Equation 4.27.30 is illustrated in Figure 4.27.30. The predictions from Equation 4.27.29 are also included as a reference. As expected, the scatter in the predictions from the GP-based correlation is much smaller.

Effect of Penalty Parameters

In limiting the size of the correlations for practical thermal engineering applications, it is possible to choose different penalty functions, $g(L_f)$, in Equation 4.27.27. Using the sigmoidal form as in Equation 4.27.28, however, prevents the fitness in Equation 4.27.28 from becoming either unbounded or singular thus avoiding computational problems, but the choice of a_1 and a_2 has an effect on the results. This effect was analyzed and reported by Cai et al. (2006), with different values of a_1 and a_2, maintaining one variable fixed while changing the other. Three runs were made for each a_1, a_2 value and the results are presented next.

Figure 4.27.31 presents the results for a_1 constant and varying a_2, with Figure 4.27.31a showing the penalty function vs. correlation size, and Figure 4.27.31b illustrating the corresponding penalized fitness values. The result of each run is also marked on them. As can be seen from the figures and Equation 4.27.28, the parameter a_2 provides a nominal size of the correlation by simply shifting the penalty function horizontally. A larger a_2 enables a larger correlation size with a smaller error to be selected. The choice of a_2 is subjective in that a smaller error is to be preferred, but at the expense of functional complexity.

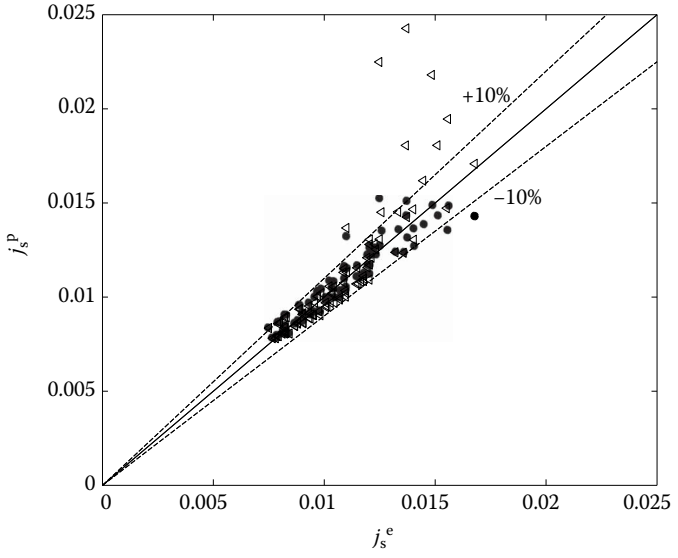


FIGURE 4.27.30 Experimental vs. predicted j factors for a compact heat exchanger from Equation 4.27.30: (•). Also shown are predictions of Equation 4.27.29 by McQuiston: (◁). Straight line is the perfect prediction.

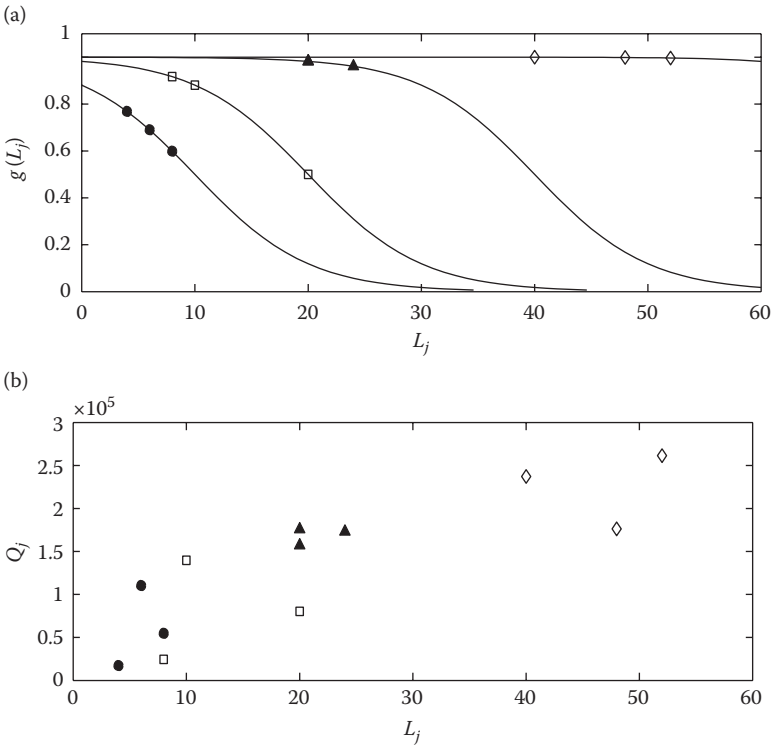


FIGURE 4.27.31 (a) Penalty and (b) penalized fitness functions vs. correlation size for different a_2: (•) 10, (□) 20, (▲) 40, (◊) 80.

The results for variable a_1 and constant a_2 are shown in Figure 4.27.32, with Figure 4.27.32a illustrating the penalty function and Figure 4.27.32b the penalized fitness function. In this case, a_1 is a measure of the slope of the penalty function and determines how strongly a large functional form of the correlation is penalized. From the figures, it is clear that for small values of a_1, the penalty function changes smoothly providing a gradually increasing rejection of large correlations, while for the large values, the fitness curve shows a sudden onset of rejection. Though the choice of a_1 is subjective, an intermediate value for a_1 is appropriate for finding relatively compact well-fitted correlations.

Additional Applications of GPs in Thermal Engineering

Applications involving GP are rarer than those of the GA. Lee et al. (1997) studied the problem of correlating the critical heat flux (CHF) for upward water flow in vertical round tubes under low-pressure and low-flow conditions. Correlations were developed for two types of conditions, namely, inlet and local. In these, the data included the tube diameter, fluid pressure, and mass flux, along with either the heated length and the subcooling enthalpy for the earlier or the critical quality for the latter. Genetic programming was used as a symbolic regression tool to find the functional form and the constants in it that provided the best accuracy. The final correlations that were found had predictions better than those in the literature. In a different study, the modeling of solid oxide fuel cell systems with semi-empirical equations was performed by Chakraborty (2009). Given the data for fixed and dynamic loads, the GP technique was used to find the correlation functions that could accurately describe the electrochemical and thermal properties of solid-oxide fuel-cell stacks. The results from the GP-based models for the static current–voltage behavior, and the transient response

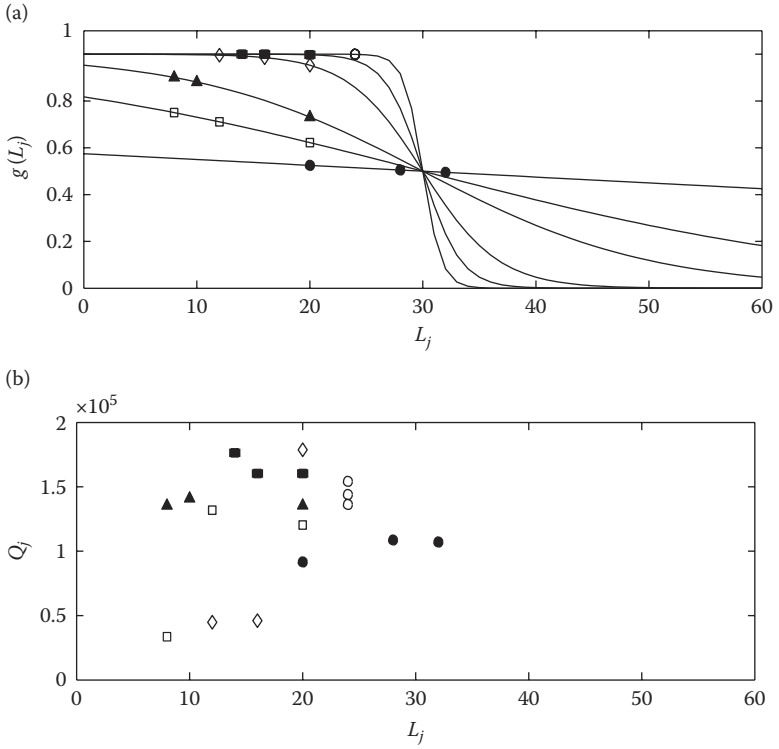


FIGURE 4.27.32 (a) Penalty and (b) penalized fitness functions vs. correlation size for different a_1: (•) 0.01, (□) 0.05, (▲) 0.1, (◊) 0.3, (■) 0.6, (o) 1.2.

under changing load conditions, were more accurate that the state-of-the-art radial basis function neural network. On the other hand, Lee and Tong (2011) conducted a study on the development of reliable models towards forecasting energy consumption from experimental data. The investigation focused on combining the residual modification information with GP sign estimation to improve a "grey" forecasting model. The energy consumption in China was considered to demonstrate the effectiveness of the proposed GP-based forecasting model.

An experimental investigation of helically finned tubes was reported by Zdaniuk et al. (2011) with the idea of investigating the performance of two correlation techniques to be used for system optimization. From the data collected, empirical correlations for the friction factor f and the Colburn j factor were developed by symbolic regression and used to find the geometric parameters and operating conditions, which would maximize heat transfer enhancement efficiency index. The results from the GP-based models were better suited for optimization purposes than those of an ANN approach. A similar study that introduced a new correlation based on GP to describe choke performance in both critical and subcritical flow conditions in pipeline flow regimes was conducted by Kaydani et al. (2014). The GP correlation equation was developed using field data collected from various Iranian oil well pipelines, and its accuracy compared to several other empirical models. A statistical analysis of the results from the symbolic-regression-based model show that for either flow condition, it can be applied effectively yielding high accuracy for choke performance pipeline forecasting.

In a recent paper, Liu et al. (2015b) followed the idea first proposed by Cai et al. (2006) to integrate the GP technique with the GA for deriving more accurate heat transfer correlations. Three different experimental data sets, corresponding to a shell-and-tube heat exchanger with continuous helical baffles, a single row heat exchanger with helically finned tubes and a finned oval-tube heat

exchanger with double rows of tubes, respectively, were considered for their study. In each case, the final correlation functions were found to be more accurate in predicting their thermal behavior than currently used models. Based on the aforementioned work, Liu et al. (2015a) proposed a method for synthesizing both economic cost and energy benefit aspects of the sinter cooling bed. The predictions were based on correlation equations derived by symbolic regression for the annual energy gains as functions of the parameters of operation. A bi-objective optimization of the sinter cooling bed was finally performed to achieve the optimal conditions in reference to both waste heat utilization and economic cost. The advantage of the GP method in seeking an optimum functional form was exploited in these applications.

General Discussion of GAs and GPs

The evolutionary programming method has the advantage that, unlike the ANN, a functional form of the relationship is obtained. Genetic algorithms, GP, and symbolic regression are relatively new techniques from the perspective of thermal engineering, and we can only expect the applications to grow. There are a number of areas in prediction, control, and design where these techniques can be effectively used. One of these, in which progress can be expected, is in thermal-hydronic networks. Networks are complex systems built up from a large number of simple components; though the behavior of each component may be well understood, the behavior of the network requires massive computations that may not be practical. Optimization of networks is an important issue from the perspective of design, since it is not obvious what the most energy-efficient network, given certain constraints, should be. The constraints are usually in the form of the locations that must be served and the range of thermal loads that are needed at each position. A search methodology based on the calculation of every possible network configuration would be very expensive in terms of computational time. An alternative based on evolutionary techniques would be much more practical. Under this procedure, a set of networks that satisfy the constraints would be proposed as candidates for the optimum. From this set, a new and more fit generation would evolve and the process repeated until the design does not change much. The definition of fitness, for this purpose, would be based on the energy requirements of the network.

Concluding Remarks

AI is a broad term and many of its techniques have not been reviewed here. Reviews like those of Culp et al. (1990) on the impact of AI within the HVAC industry, and Cahill (1994) and Paoletti and Sciubba (1997) on the use of a knowledge-based expert system for the design of complex thermal systems like power plants, should be consulted. Other examples on the use of AI in thermal engineering are the following. Rasmussen et al. (1990) identified the distillation process dynamics comparing process knowledge and black box-based approaches. Gao et al. (1990) worked on the synthesis of optimized heat exchanger networks. Fujii and Lutzenhiser (1992) used intelligent systems in residential air-conditioning and natural cooling systems. Loveday and Virk (1992) applied the technique to buildings. Dregalin and Nazyrova (1993) introduced AI elements to heat power engineering. Sandberg et al. (1995) studied an intelligent distributed heater used for industrial applications. Yang et al. (1996) developed an analysis of intelligent control strategies of a thermal energy storage air-conditioning system. These papers indicate some of the wide variety of approaches that can be considered in thermal engineering.

Despite the tremendous strides made in advancing our knowledge base in the field of thermal science and engineering in the last 100 years or so, researchers in the field have often felt apprehensive that their hands were tied by the relatively limited though popular approaches in the traditional analysis of thermal phenomena: solving the governing first principles or modeling equations and finding the necessary correlations from experimental data. Even with the ever-increasing computing resources, which have become available, difficulties still exist since the systems are much

more complex, and the need is to respond more quickly to the applications and challenges of the fast-changing world of technology. The AI community provides some possible solutions, and it is surprising that there has not been much more concerted effort in the thermal-science research community to seek the application of these new approaches, which have the potential for addressing these needs. This section has discussed two of these techniques, the ANN and the genetic algorithmic approaches, that are beginning to take root now. The purpose here has been to introduce the reader to these methodologies, and it is up to him or her to apply them and similar procedures to a broader set of problems. AI techniques are well suited to complex problems, and thermal engineering presents many examples of this. It is not a substitute, but rather a complement, for incisive physical insight and first-principle analyses.

ACKNOWLEDGMENTS

We are grateful for the contributions our collaborators X. Zhao and R.L. McClain have made to our understanding of the subject. We are also thankful for the financial support of the late D.K. Dorini to the Hydronics Laboratory where most of our work has been carried out, and that of the Mexican Council of Science and Technology (CONACyT) and the Organization of American States for support.

REFERENCES

Abdel-Magid, Y. and Dawoud, M., 1995. Tuning of AGC of interconnected reheat thermal systems with genetic algorithms. In *Proceedings of the IEEE International Conference System Man Cybernetics,* Vancouver, BC, volume 3, pp. 2622–2627.

Adams, L., Hoeinghaus, R., and Lantz, J., 1990. Optimization and protection of feedwater heaters. In *Proceedings of the 33rd Power Instrumentation Symposium*, Toronto, ON, volume 33, pp. 47–52.

Alessandri, A., Verona, F., Parisini, T., and Torrini, A., 1994. Neural approximations for the optimal control of heating systems. In *Proceedings of the IEEE Conference Control Applications*, Glasgow, UK, volume 3, pp. 1613–1618.

Allred, L. and Kelly, G., 1992. Modified genetic algorithm for extracting thermal profiles from infrared image data. In *Proceedings of the SPIE–International Society for Optical Engineering. Neural Stochastic Methods Image and Signal Processing*, San Diego, CA, volume 1766, pp. 77–81.

Androulakis, I. and Venkatasubramanian, V., 1991. Genetic algorithmic framework for process design and optimization. *Comput. Chem. Eng.*, 15(4):217–228.

Angeline, P., Saunder, G., and Pollack, J., 1994. Complete induction of recurrent neural networks. In Sebald, A. and Fogel, L., editors, *Proceedings of the Third Annual Conference Evolutionary Programming*, pp. 1–8. World Scientific, Singapore.

Ayoubi, M., 1997. Dynamic multi-layer perception: application to the nonlinear identification and predictive control of a heat exchanger. In Kalkkuhl, J., Hunt, K., Zbikowski, R., and Dzielinski, A., editors, *Applications of Neural Adaptive Control Technology*, pp. 205–230. World Scientific, Singapore.

Bittanti, S. and Piroddi, L., 1997. Nonlinear identification and control of a heat exchanger: A neural network approach. *J. Franklin Inst.*, 334B(1):135–153.

Bonnet, A., 1985. *Artificial Intelligence, Promise and Performance*. Prentice-Hall, Englewood Cliffs, NJ.

Bouchon-Meunier, B., Yager, R., and Zadeh, L., 1995. *Fuzzy Logic and Soft Computing*. World Scientific, Singapore.

Brown, M. and Harris, C., 1994. *Neurofuzzy Adaptive Modelling and Control*. Prentice-Hall, Englewood Cliffs, NJ.

Bulsari, A. and Saxin, H., 1995. Classification of blast furnace probe temperatures using neural networks. *Steel Res.*, 66:231–236.

Cahill, E., 1994. Knowledge-based algorithm construction for real-world engineering PDEs. *Math. Comput. Simul.*, 36(4–6):389–400.

Cai, W., Pacheco-Vega, A., Sen, M., and Yang, K., 2006. Heat transfer correlations by symbolic regression. *Int. J. Heat Mass Transfer*, 49(23–24):4352–4359.

Carriere, A., Cela, A., and Hamam, Y., 1994. Neural network-based adaptive control of a non-linear system: Application to a thermal process. In *Proceedings of the IEEE International Conference Systems, Man Cybernetics*, San Antonio, TX, volume 2, pp. 1133–1138.

Castrogiovanni, A. and Sforza, P., 1996. Scaling of high heat flux flow boiling flow systems. In *Proceedings of the ASME Heat Transfer Division, International Mechanical Engineering Congress Exposition*, Atlanta, GA, volume 334 of 3, pp. 39–46.

Castrogiovanni, A. and Sforza, P., 1997. A genetic algorithm model for high heat flux flow boiling. *Exp. Thermal Fluid Sci.*, 15(3):193–201.

Chakraborty, U., 2009. Static and dynamic modeling of solid oxide fuel cell using genetic programming. *Energy*, 34(6):740–751.

Charniak, E. and McDermott, D., 1985. *Introduction to Artificial Intelligence*. Addison-Wesley, Reading, MA.

Chipperfield, A., 1997. Introduction to genetic algorithms. In Zalzala, A. and Fleming, P., editors, *Genetic Algorithms in Engineering Systems*. Institute of Electrical and Electronics Engineers, New York.

Chow, M. and Teeter, J., 1997. Reduced-order functional link neural network for HVAC thermal system identification and modeling. In *Proceedings of the IEEE International Conference Neural Networks*, Houston, TX, volume 1, pp. 5–9.

Cohen, W. and Johnson, E., 1956. Dynamic characteristics of double-pipe heat exchangers. *Ind. Eng. Chem.*, 48(6):1031–1034.

Culp, C., Haberl, J., Norford, L., Brothers, P., and Hall, J., 1990. Impact of AI technology within the HVAC industry. *ASHRAE J.*, 32(12):12–22.

Curtiss, P., Kreider, J., and Brandemuehl, M., 1993. Artificial neural networks proof of concept for local and global control of commercial building HVAC systems. In *Proceedings of the ASME International Solar Energy Conference*, Washington, DC, volume 1, pp. 429–443.

Davalos, R. and Rubinsky, B., 1996. An evolutionary-genetic approach to heat transfer analysis. *ASME J. Heat Transfer*, 118(3):528–531.

Dhar, A., Claridge, D., and Ruch, D., 1995. Artificial neural network with wavelet basis functions to predict hourly heating and cooling energy use in commercial buildings with temperature as the only weather variable. In *Proceedings of the ASME-JSME-JSES International Solar Energy Conference*, Maui, Hawaii, volume 1, pp. 145–156.

Diaz, G. and Campo, A., 2009. Artificial neural networks to correlate in-tube turbulent forced convection of binary gas mixtures. *Int. J. Therm. Sci.*, 48:1392–1397.

Diaz, G., Sen, M., and McClain, R., 2000a. Adaptive neurocontrol and minimization of energy consumption of a heat exchanger test facility. In *Proceedings of the ASME IMECE, HTD-Vol.366-3*, Orlando, FL, *Heat Transfer Division-2000*, pp. 111–116.

Diaz, G., Sen, M., Yang, K., and McClain, R., 1996. Analysis of data from single-row heat exchanger experiments using an artificial neural network. In *FED. Proceedings of the ASME International Mechanical Engineering Congress Exposition*, Atlanta, GA, volume 242, pp. 45–52.

Diaz, G., Sen, M., Yang, K., and McClain, R., 1998a. Artificial neural network simulation of the dynamic behavior of heat exchangers. In *Proceedings of the Seventh Latin American Heat and Mass Transfer Conference*, Salta, Argentina, volume 1, pp. 23–28.

Diaz, G., Sen, M., Yang, K., and McClain, R., 1998b. Neurocontrol of heat exchangers (in Spanish). In *Proceedings of the 4th International Meeting in Mechanical Engineering*, pp. 123–128, Mexico.

Diaz, G., Sen, M., Yang, K., and McClain, R., 1999a. Artificial neural network control of an experimental heat exchanger facility. In *Proceedings of the ASME Heat Transfer Division-1999, 1999 ASME IMECE*, Nashville, TN, volume 364-4, pp. 325–330.

Diaz, G., Sen, M., Yang, K., and McClain, R., 1999b. Simulation of heat exchanger performance by artificial neural networks. *Int. J. HVAC&R Res.*, 5(3):195–208.

Diaz, G., Sen, M., Yang, K., and McClain, R., 2000b. Use of artificial neural networks for temperature control in heat exchangers. In *Proceedings of the Fourth ISHMT-ASME Heat and Mass Transfer Conference and Fifteenth National Heat and Mass Transfer Conference*, pp. 963–968, Pune, India.

Diaz, G., Sen, M., Yang, K., and McClain, R., 2001a. Adaptive neurocontrol of heat exchangers. *ASME J. Heat Transfer*, 123(3):556–562.

Diaz, G., Sen, M., Yang, K., and McClain, R., 2001b. Dynamic prediction and control of heat exchangers using artificial neural networks. *Int. J. Heat Mass Transfer*, 44(9):1671–1679.

Diaz, G., Sen, M., Yang, K., and McClain, R., 2001c. On-line training of artificial neural networks for control of a heat exchanger test facility. In *Proceedings of the ASME NHTC'01*, volume 10, pp. 359–365, Anaheim, CA.

Diaz, G., Sen, M., Yang, K., and McClain, R., 2004. Stabilization of thermal neurocontrollers. *Appl. Artif. Intell.*, 18(5):447–466.

Ding, Y. and Wong, K., 1990. Control of a simulated dual-temperature hydronic system using neural network approach. *ASHRAE Trans.*, 96(2):727–732.

Dounis, A., Santamouris, M., and Lefas, C., 1992. Implementation of artificial intelligence techniques in thermal comfort control for passive solar buildings. *Energy Convers. Manage.*, 33(3):175–182.

Drago, G. and Ridella, S., 1992. Statistically controlled activation weight initialization. *IEEE Trans. Neural Networks*, 3(4):627–631.

Dregalin, A. and Nazyrova, R., 1993. On introduction of artificial intelligence elements to heat power engineering. *Izv. VUZ: Aviatsionnaya Tekhnika*, 4:90–94.

Eeckman, F., 1992. *Analysis and Modeling of Neural Systems*. Kluwer Academic, Boston, MA.

Fabbri, G., 1997. A genetic algorithm for fin profile optimization. *Int. J. Heat Mass Transfer*, 40(9):2165–2172.

Ferrano, F. and Wong, K.-F., 1990. Prediction of thermal storage loads using a neural networks. *ASHRAE Trans.*, 96(2):723–726.

Flood, I. and Kartam, N., 1994. Neural networks in civil engineering. 1. Principles and understanding. *ASCE J. Comput. Civil Eng.*, 8(2):131–148.

Fortuna, L., Muscato, G., Nunnari, G., and Papaleo, R., 1996. Neural network approach to control the temperature on rapid thermal processing. In *Proceedings of the 8th Mediterranean Electrotechnical Conference MELECON Industrial Applied Power Systems, Computer Science and Telecommunications*, Bari, Italy, volume 2 of Part 2, pp. 649–652.

Fowler, M., Klett, D., Moreno, J., and Heermann, P., 1997. Using artificial neural networks to predict the performance of a liquid sodium reflux pool boiler solar receiver. In *Proceedings of the International Solar Energy Conference*, Washington, DC, pp. 93–104.

Fujii, H. and Lutzenhiser, L., 1992. Japanese residential air-conditioning: natural cooling and intelligent systems. *Energy Build.*, 18(3–4):221–233.

Gagarin, N., Flood, I., and Albrecht, P., 1994. Computing track attributes with artificial neural networks. *ASCE J. Comput. Civil Eng.*, 8(2):179–200.

Gao, W., Yu, W., and Han, F., 1990. An artificial intelligence approach to the synthesis of optimized heat exchanger networks. *J. Chem. Ind. Eng. (China)*, 41(3):353.

Gartner, J. and Harrison, H., 1965. Dynamic characteristics of water-to-air cross flow heat exchangers. *ASHRAE Trans.*, 71:212–223.

Gobovic, D. and Zaghloul, M., 1993. Design of locally connected CMOS neural cells to solve the steady-state heat flow problem. In *Proceedings of the 36th Midwest Symposium Circuits Systems*, Detroit, MI, volume 1, pp. 755–757.

Goldberg, D., 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA.

Gray, D. and Webb, R., 1986. Heat transfer and friction correlations for plate finned-tube heat exchangers having plain fins. In *Proceedings of the 8th International Heat Transfer Conference*, San Francisco, CA, volume 6, pp. 2745–2750.

Hale, J. and Kocak, H., 1991. *Dynamics and Bifurcations*. Springer-Verlag, New York.

Haykin, S., 1994. *Neural Networks: A Comprehensive Foundation*. Macmillan, New York.

Hecht-Nielsen, R., 1987. Kolmogorov's mapping neural network existence theorem. In *IEEE First International Conference in Neural Networks*, San Diego, CA, volume III, pp. 11–13.

Holland, J., 1975. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

Holland, J., 1992. Genetic algorithms. *Sci. Am.*, 267(1):66–72.

Huang, S.-H. and Nelson, R., 1994. Delay time determination using an artificial neural network. *ASHRAE Trans.*, 100(1):831–840.

Hunt, K., Sbarbaro, D., Zbikowski, R., and Gawthrop, P., 1992. Neural networks for control systems-a survey. *Automatica*, 28(6):1083–1112.

Inoue, Y., Sugiura, S., Murakami, K., and Kotera, N., 1995. Thermal comfort and climate control. *ASHRAE J.*, 37(10):28–32.

Ito, K., Shiba, T., and Yokoyama, R., 1995. Operational advisor for a heat pump/thermal storage system (development and evaluation of a method based on a layered neural network). *Trans. Jpn. Soc. Mech. Eng. C*, 61(590):3909–3915.

Jambunathan, K., Hartle, S., Ashforth-Frost, S., and Fontama, V., 1996. Evaluating convective heat transfer coefficients using neural networks. *Int. J. Heat Mass Transfer*, 39(11):2329–2332.

Jang, J.-S., Sun, C.-T., and Mizutani, E., 1997. *Neuro-Fuzzy and Soft Computing*. Prentice-Hall, Upper Saddle River, NJ.

Jeannette, E., K. Assawamartbunlue, P. C., and Kreider, J., 1998. Experimental results of a predictive neural network HVAC controller. *ASHRAE Trans.*, 104(2):192–197.

Jenkins, W., 1995. Neural network-based approximations for structural analysis. In Topping, B., editor, *Neural Networks and Evolutionary Computing for Civil and Structural Engineering*, pp. 25–35. Civil-Computing Press, Edinburgh, UK.

Jones, M., Tezuka, A., and Yamada, Y., 1995. Thermal tomographic detection of inhomogeneities. *ASME J. Heat Transfer*, 117(4):969–975.

Kamarthi, S., Sanvido, V., and Kumara, R., 1992. Neuroform-neural network system for vertical form-work selection. *ASCE J. Comput. Civil Eng.*, 6(2):178–199.

Kaminski, W., Strumillo, P., and Tomczak, E., 1996. Genetic algorithms and artificial neural networks for description of thermal deterioration processes. *Drying Technol.*, 14(9):2117–2133.

Karmin, E., 1990. Simple procedure for pruning back propagation-trained neural networks. *IEEE Trans. Neural Networks*, 1(2):239–242.

Kawashima, M., Dorgan, C., and Mitchell, J., 1995. Hourly thermal load prediction for the next 24 hours by Arima, Ewma, LR, and an artificial neural network. *ASHRAE Trans.*, 1(2):186–200.

Kaydani, H., Najafzadeh, M., and Mohebbi, A., 2014. Wellhead choke performance in oil well pipeline systems based on genetic programming. *J. Pipeline Syst. Eng. Pract.*, 5(3):06014001.

Kos, A., 1993. Approach to thermal placement in power electronics using neural networks. In *Proceedings of the IEEE International Symposium Circuits Systems*, Chicago, IL, volume 4, pp. 2427–2430.

Koza, J., 1992. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA.

Koza, J., 1994. *Genetic Programming II: Automatic Discovery of Reusable Programs*. MIT Press, Cambridge, MA.

Kreider, J., Claridge, D., Curtiss, P., Dodier, R., Haberl, J., and Krarti, M., 1995. Building energy use prediction and system identification using recurrent neural networks. *ASME J. Solar Energy Eng.*, 117(3):161–166.

Kreider, J., Wang, X., Anderson, D., and Dow, J., 1992. Expert systems, neural networks and artificial intelligence applications in commercial building HVAC operations. *Automat. Constr.*, 1(3):225–238.

Kuroe, Y. and Kimura, I., 1995. Modeling of unsteady heat conduction field by using composite recurrent neural networks. In *Proceedings of the IEEE International Conference in Neural Networks*, Perth, WA, volume 1 of Part 1, pp. 323–328.

Lavric, D., Lavric, V., Muntean, O., and Danciu, E., 1994. Auto-organizing algorithm for design of fin heat exchangers. *Rev. Roum. Chim.*, 39(11):1241–1256.

Lavric, D., Lavric, V., and Woinaroschy, A., 1995. Designing fin heat exchanger with a neural network. *Rev. Roum. Chim.*, 10(2):561–565.

Lee, D.-G., Kim, H.-G., Baek, W.-P., and Chang, S., 1997. Critical heat flux prediction using genetic programming for water flow in vertical round tubes. *Int. Commun. Heat Mass Transfer*, 24(7):919–929.

Lee, Y. and Tong, L., 2011. Forecasting energy consumption using a grey model improved by incorporating genetic programming. *Energy Convers. Manage.*, 52(1):147–152.

Lehtokangas, M., Saarinen, J., and Kaski, K., 1995. Initializing weights of a multilayer perceptron network by using the orthogonal least square algorithm. *Neural Comput.*, 7:982–999.

Leitch, R. and Quek, C., 1992. Architecture for integrated process supervision. In *IEEE Proceedings of the Part D: Control Theory and Applications*, volume 139 of 3, pp. 317–327.

Li, X., Vaezi-Nejad, H., and Visier, J.-C., 1996. Development of a fault diagnosis method for heating systems using neural networks. *ASHRAE Trans.*, 102(1):607–614.

Li, H. and Yang, C., 1997. A genetic algorithm for inverse radiation problems. *Int. J. Heat Mass Transfer*, 40(7):1545–1549.

Liu, Y., Yang, J., Wang, J., Ding, X., Cheng, Z., and Wang, Q., 2015a. Prediction, parametric analysis and bi-objective optimization of waste heat utilization in sinter cooling bed using evolutionary algorithm. *Energy*, 90(1):24–45.

Liu, Y., Yang, J., Xu, J., Cheng, Z., and Wang, Q., 2015b. Integration of genetic programing with genetic algorithm for correlating heat transfer problems. *ASME J. Heat Transfer*, 137(6):061012.

Loveday, D. and Virk, G., 1992. Artificial intelligence for buildings. *Appl. Energy*, 41(3):201–221.

Mahajan, R. and Wang, X., 1993. Neural network models for thermally based microelectronic manufacturing processes. *J. Electrochem. Soc.*, 140(8):2287–2293.

Matsumoto, R., Kikkawa, S., and Senda, M., 1997. Effect of pin fin arrangement on endwall heat transfer. *JSME Int. J. B: Fluids Thermal Eng.*, 40(1):142–151.

McKay, B., Willis, M., and Barton, G., 1997. Steady-state modelling of chemical process systems using genetic programming. *Comput. Chem. Eng.*, 21(9):981–996.

McQuiston, F., 1978a. Correlation of heat, mass and momentum transport coefficients for plate fin tube heat transfer surfaces with staggered tubes. *ASHRAE Trans.*, 84(1):294–309.

McQuiston, F., 1978b. Heat, mass and momentum transfer data for five plate-fin-tube heat transfer surfaces. *ASHRAE Trans.*, 84(1):266–293.

Meysenc, L., Saludjian, L., Bricard, A., Rael, S., and Schaeffer, C., 1997. High heat flux IGBT micro exchanger setup. *IEEE Trans. Components Packag. Manuf. Technol. A*, 20(3):334–341.

Michalewicz, Z., 1992. *Genetic Algorithms+Data Structures=Evolution Programs.* 2nd ed., Springer-Verlag, Berlin.

Mistry, S. and Nair, S., 1993. Nonlinear HVAC computations using neural networks. *ASHRAE Trans.*, 99(1):775–784.

Mitchell, M., 1997. *An Introduction to Genetic Algorithms.* MIT Press, Cambridge, MA.

Momose, K., Murai, T., Asami, T., and Hosokawa, Y., 1993. Identification of heat-source using neural network. *Trans. Jpn. Soc. Mech. Eng. C*, 59(567):3431–3436.

Moon, C. and Cho, B., 1996. Design of fuzzy control system for chemical injection system retrofit using neural network model in thermal power plant. In *Proceedings of the 1st International Workshop Neural Networks Identification, Control, Robotics, and Signal/Image Process.*, Venice, Italy, pp. 263–269.

Mozley, J., 1956. Predicting dynamics of concentric pipe heat exchangers. *Ind. Eng. Chem.*, 48(6):1035–1041.

Nahas, E., Henson, M., and Seborg, D., 1992. Nonlinear internal model control strategy for neural network models. *Comput. Chem. Eng.*, 16(12):1039–1057.

Naka, M., Imai, T., Shida, T., Sato, M., Ito, R., and Akamine, I., 1993. Thermal image processing using neural network. In *Proceedings of the International Joint Conference Neural Networks*, volume 3 of Part 3, pp. 2065–2068, Nagoya, Japan.

Narendra, K. and Mukhopadhyay, S., 1996. Intelligent control using neural networks. In Gupta, M. and Sinha, N., editors, *Intelligent Control Systems, Theory and Applications*, pp. 151–186. Institute of Electrical and Electronics Engineers, New York.

Normandin, A., Grandjean, B., and Thibault, J., 1993. PVT data analysis using neural network models. *Indust. Engrg. Chem. Research*, 32(5):970–975.

Okamoto, K., Tsuru, D., and Fumizawa, M., 1996. Reconstruction of three-dimensional density distribution from the limited projection images with genetic algorithm. *J. Flow Visual. Image Process.*, 3(4):253–263.

Pacheco-Vega, A., Cai, W., Sen, M., and Yang, K., 2003a. Heat transfer correlations in an air-water fin-tube compact heat exchanger by symbolic regression. In *Proceedings of the 2003 ASME International Mechanical Engineering Congress and Exposition*, Washington, DC. IMECE2003/HTD-41977.

Pacheco-Vega, A., Diaz, G., Sen, M., Yang, K., and McClain, R., 2001a. Heat rate predictions in humid air-water heat exchangers using correlations and neural networks. *ASME J. Heat Transfer*, 123(2):348–354.

Pacheco-Vega, A., Sen, M., and Yang, K., 2003b. Simultaneous determination of in-and over-tube heat transfer correlations in heat exchangers by global regression. *Int. J. Heat Mass Transfer*, 46(6):1029–1040.

Pacheco-Vega, A., Sen, M., Yang, K., and McClain, R., 1998. Genetic-algorithm-based-predictions of fin-tube heat exchanger performance. In Lee, J., editor, *Proceedings of the 11th International Heat Transfer Conference*, volume 6, pp. 137–142. Taylor & Francis, New York.

Pacheco-Vega, A., Sen, M., Yang, K., and McClain, R., 1999. Prediction of humid air heat exchanger performance using artificial neural networks. In *Proceedings of the 1999 ASME International Mechanical Engineering Congress and Exposition*. ASME, New York.

Pacheco-Vega, A., Sen, M., Yang, K., and McClain, R., 2001b. Neural network analysis of fin-tube refrigerating heat exchanger with limited experimental data. *Int. J. Heat Mass Transfer*, 44(4):763–770.

Paoletti, B. and Sciubba, E., 1997. Artificial intelligence in thermal systems design: concepts and applications. In Boehm, R., editor, *Developments in the Design of Thermal Systems.* Cambridge University Press, New York.

Pican, N., Alexandre, F., Haton, J.-P., and Bresson, P., 1998. Neurocontrol of temperatures with OWE neural network architecture in continuous steel annealing furnace. In Soulie, F. and Gallinari, P., editors, *Industrial Applications of Neural Networks*, pp. 9–15. World Scientific, Singapore.

Psichogios, D. and Ungar, L., 1991. Direct and indirect model based control using artificial neural networks. *Indust. Engrg. Chem. Res.*, 30:2564–2573.

Queipo, N., Devarakonda, R., and Humphrey, J., 1994. Genetic algorithms for thermosciences research: application to the optimized cooling of electronic components. *Int. J. Heat Mass Transfer*, 37(6):893–908.

Queipo, N. and Gil, G., 1997. Multiobjective optimization of component placement on planar printed wiring boards. In *Proceedings of the IEEE 13th Annual Semiconductor Thermal Measurement and Management Symposium*, Austin, TX, pp. 92–105.

Rasmussen, K., Nielsen, C., and Jorgensen, S., 1990. Identification of distillation process dynamics comparing process knowledge and black box based approaches. In *Proceedings of the American Control Conference*, San Diego, CA, pp. 3116–3121.

Raudensky, M., Woodbury, K., Kral, J., and Brezina, T., 1995. Genetic algorithm in solution of inverse heat conduction problems. *Numer. Heat Transfer, Part B*, 28(3):293–306.

Rich, E. and Knight, K., 1991. *Artificial Intelligence*. 2nd ed., McGraw-Hill, New York.

Rogers, J., 1994. Simulating structure analysis with neural networks. *ASCE J. Comput. Civil Eng.*, 8(2):252–265.

Romero-Mendez, R., Hidalgo-Lopez, J., Duran-Garcia, H., and Pacheco-Vega, A., 2014. Use of artificial neural networks for prediction of convective heat transfer in evaporative units. *Ingenieria Investigacion y Tecnologia*, XV(1):93–101.

Romero-Mendez, R., Vazquez, P. L., Oviedo-Tolentino, F., Duran-Garcia, H., Perez-Gutierrez, F., and Pacheco-Vega, A., 2016. Use of artificial neural networks for prediction of the convective heat transfer in evaporative mini-tubes. *Ing., Invest. Tecnol*., XVII(1):23–34.

Ruiz-Mercado, C., Pacheco-Vega, A., and Torres-Chavez, G., 2009. A Takagi-Sugeno fuzzy dynamic model of a concentric-tubes heat exchanger. *Chem. Prod. Process Model.*, 4(2):1–22.

Rumelhart, D., Hinton, G., and Williams, R., 1986. Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Miscrostructure of Cognition*, volume 1, pp. 8.318–8.362. MIT Press, Cambridge, MA.

Rzempoluck, E., 1998. *Neural Network Data and Analysis Using Simulnet*. Springer-Verlag, New York.

Sandberg, C., Whitney, W., Nassar, A., and Kuse, G., 1995. Intelligent distributed heater used for industrial applications. In *Proceedings of the IEEE International Conference Systems, Man Cybernetics*, Vancouver, BC*, Part 4*, volume 4, pp. 3346–3351.

Schalkoff, R., 1997. *Artificial Neural Networks*. McGraw-Hill, New York.

Schmidt, F. and Willmott, A., 1981. *Thermal Energy Storage and Regeneration*. McGraw-Hill, New York.

Schmitt, T., Dhingra, A., Landis, F., and Kojasoy, G., 1996. Genetic algorithm optimization technique for compact high intensity cooler design. *J. Enhanc. Heat Transfer*, 3(4):281–290.

Sette, S. and Boullart, L., 2001. Genetic programming: principles and applications. *Eng. Appl. Artif. Intell.*, 14(1):727–736.

Sozen, A., Arcaklioglu, E., and Menlik, T., 2010. Derivation of empirical equations for thermodynamic properties of a ozone safe refrigerant (R404a) using artificial neural network. *Expert Syst. Appl.*, 37(2):1158–1168.

Spiga, M. and Spiga, G., 1992. Step response of the cross-flow heat-exchanger with finite wall capacitance. *Int. J. Heat Mass Transfer*, 35(2):559–565.

Takemori, T., Miyasaka, N., and Hirose, S., 1991. Neural network air-conditioning system for individual comfort. In *Proceedings of the SPIE International Society of Optical Engineering Applications Artificial Neural Networks II*, Orlando, FL, volume 1469, pp. 157–165.

Tang, M. and Carothers, J., 1996. Multichip module placement with heat consideration. In *Proceedings of the 9th Annual IEEE International ASIC Conference Exhibition*, Rochester, New York, pp. 175–178.

Tanimoto, S., 1990. *The Elements of Artificial Intelligence Using Common LISP*. Computer Science Press, New York.

Terry, P. and Himmelblau, D., 1993. Data rectification and gross error detection in a steady-state process via artificial neural networks. *Indust. Engrg. Chem. Res.*, 32(12):3020–3028.

Thal-Larsen, H., 1960. Dynamics of heat exchangers and their models. *ASME J. Basic Eng.*, 82:489–504.

Thibault, J. and Grandjean, B., 1991. Neural network methodology for heat transfer data analysis. *Int. J. Heat Mass Transfer*, 34(8):2063–2070.

von Wolfersdorf, J., Achermann, E., and Weigand, B., 1997. Shape optimization of cooling channels using genetic algorithms. *ASME J. Heat Transfer*, 119(2):380–388.

Waldemark, J., Wiklund, H., Andersson, S., and Sandberg, O., 1992. Neural network modelling of the heat load in district heating systems. *Fernwärme Int.*, 21(9):424–437.

Warwick, K., Irwin, G., and Hunt, K., 1992. *Neural Networks for Control and Systems*. Short Run Press, Exeter, UK.

Wessels, L. and Barnard, E., 1992. Avoiding fake local minima by proper initialization of connections. *IEEE Trans. Neural Networks*, 3(6):899–905.

Wood, R., 1996. Inverse thermal field problem based on noisy measurements: comparison of a genetic algorithm and the sequential function specification method. *Eng. Comput.*, 13(6):43–59.

Yamashita, H., Zzumi, R., and Yamaguchi, S., 1978. Analysis of the dynamic characteristics of cross-flow heat exchangers with both fluids mixed. *Bull. JSME*, 21(153):479–485.

Yang, K., 2008. Artificial neural networks (ANNs): A new paradigm for thermal science and engineering. *ASME J. Heat Transfer*, 130(9):093001.

Yang, Q.-H., Ma, H., Liu, X., and Shi, H., 1995. Thermal deformation compensation for a coordinate grinder by neural networks. In *Proceedings of the SPIE International Society of Optical Engineering International Conference Intelligent Manufacture*, Wuhan, China, pp. 499–502.

Yang, K., Su, C., and Hwang, R., 1996. The analysis on intelligent control strategies of a thermal energy storage air-conditioning system. *Energy*, 21(4):319–324.

Yentis Jr., R. and Zaghloul, M., 1994. CMOS implementation of locally connected neural cells to solve the steady-state heat flow problem. In *Proceedings of the 37th Midwest Symposium Circuits Systems, Part 1*, volume 1, pp. 503–506, Lafayette, LA.

Zbicinski, I., Strumillo, P., and Kaminski, W., 1996. Hybrid neural model of thermal drying in a fluidized bed. *Comput. Chem. Eng.*, 20(Suppl. A):S695–S700.

Zdaniuk, G., Walters, D., Luck, R., and Chamra, L., 2011. A comparison of artificial neural networks and symbolic-regression-based correlations for optimization of helically finned tubes in heat exchangers. *J. Enhanced Heat Transfer*, 18(2):115–125.

Zeng, P., 1998. Neural computing in mechanics. *Appl. Mech. Rev.*, 51(2):173–197.

Zhao, X., 1995. Performance of a single-row heat exchanger at low in-tube flow rates. MS Thesis, University of Notre Dame, Notre Dame, IN.

Zhao, X., Sen, M., Yang, K., and McClain, R., 1995. An artificial neural network model of a heat exchanger. In *Symposium Thermal Science Engineering in Honor of Chancellor Chang-Lin Tien*, Berkeley, CA, pp. 83–88.

Zhuang, M. and Atherton, D., 1993. Automatic tuning of optimum PID controllers. *Control Theory and Applications, IEEE Proc. D*, 140(3):216–224.

4.28 THERMAL ASPECTS OF PAPER MAKING

Martine Rueff and Evelyne Mauret

Introduction

Paper making is a very ancient process that aims at producing fibrous materials, used in the everyday life, and characterized by an incredibly diversity of applications. However, papers are largely composed of cellulosic fibers issued from plant biomass only (mainly wood but also perennial and annual plants) even if some other fibers can be used for imparting very specific properties (synthetic, mineral, and animal fibers). So, paper and board are bio-based and recyclable materials, which make them perfectly suited to meet the requirements of a sustainable world.

Industrialization of paper making began in the nineteenth century with the invention of the internal sizing based on the use of rosin added to the fiber suspension in acidic conditions, in 1807. This treatment aimed at decreasing the surface energy of the fibers, allowed the elimination of a gluing operation previously performed on dry paper sheets, this operation being manual to a large extent. The first paper machines including a drying section equipped with heated cylinders thus appeared around 1820. Of course, from the nineteenth century, important modifications of the paper making unit operations (disintegration, refining, forming, pressing, drying, etc., operations) were conducted to improve the productivity of the paper machine and the quality of paper. The use of wood as the main source of cellulosic fibers instead of rags was also an important step and radically changed the way paper was manufactured. During the twentieth century, the development of new treatments of the fiber suspension in neutral or alkaline conditions allowed the use of calcium carbonate as mineral fillers, among other things. This was also a drastic change in the monitoring of the paper machines. From 1970, the progressive closing of the water circuits and the increasing use of recycled fibers were other important features in the evolution of the paper making practices. Today, changes are made in a more incremental way and the optimization of the process is made to limit the impact on our environment and the production costs; the main aim is to decrease the amounts of water, raw materials, and energy used for manufacturing paper. In that sense, the drying section of a paper machine is of utmost importance because it contributes largely to intensive energy consumption.

In the first section of this chapter, we will briefly describe the paper making process and the specific role of drying. Then, the kinds of cellulosic wood fibers used in paper making will be presented, as well as how these fibers interact with water. Finally, the main properties of the paper are described with a special focus on the effect of drying on some properties.

The second section is devoted to the description of the drying from a more formal point of view. A brief description of drying is made, including the driving forces and the physical phenomena taking place. This section will focus then on the multicylinder drying section, its organization, and the main equipment, which play a role on heat transfer and the economy of the process. Finally, physical properties of the paper constituents, in relation with drying, will be presented.

The Paper Making Process

From Pulp Disintegration to Drying and Reeling

This section is only a brief introduction to paper making. This process is described on a very broad basis and only the Fourdrinier paper machine, which is the most common paper machine, is considered. Interested readers may refer to more developed introduction (Scott et al., 1995; Gavelin and Smook, 1998; Rojas and Hubbe, 2004; Lundell et al., 2011) or to specialized literature (Jokio, 1999; Paulapuro, 1999; Ek et al., 2009; Karlsson, 2009; Holik, 2013).

Paper is produced from an aqueous suspension of lignocellulosic fibers mainly obtained from wood (softwood and hardwood species) by a process that combines batch and continuous unit

operations. In most cases, fibers are received as pulp bales. These commercial bales are disintegrated in pulpers in order to form a fibrous suspension as homogeneous as possible. Pulpers are tanks equipped with bottom-mounted turbines: They provide high shearing effects to the suspension and allow breaking bonds between fibers thus disrupting flakes of pulp. Pulpers work at about 40–50 g/L consistency, preferably in batch conditions so that the disintegration duration can be adapted to the properties of the pulp. Nevertheless, because of increasing production rates, some installations switch to continuous mode. This has a negative impact on the quality of the fiber suspension and leads to a poor dispersion. Whatever the operating mode considered, it should be kept in mind that the fibers used in paper making exhibit high aspect ratios and, for this reason, have a high propensity to flocculate (see Table 4.28.1). Obtaining a perfectly dispersed suspension is thus not possible apart from very low fiber concentrations (below 1 g/L), which are not representative of the conditions encountered in a paper machine.

The stored pulp suspension then undergoes a mechanical treatment, called refining, whose objective is to modify the properties of the fibers in order to meet the end-use requirements regarding the considered grade of paper. Refining induces external and internal fibrillation of the fibers resulting from the compression and shear forces acting during the operation. This leads to an increase of the swelling and the conformability/flexibility of the fibers, their specific surface area, and the production of cellulosic fines. Cellulosic fines are small-sized elements whose largest dimension is less than approximately 100 μm (the upper limit ranges between 75 and 200 μm according to the definition). Refining also induces a decrease of the fiber length. So, owing to the physical modifications undergone by the fibers, refining significantly influences the physical and mechanical properties of paper, much more than the other operations in the paper making process. Generally, refining is a continuous process involving disk refiners with grooved plates (see Figure 4.28.1). The flow of the fiber suspension is forced in between the plates—one rotating (the rotor plate) and one stationary

TABLE 4.28.1
Morphological Properties of Softwood and Hardwood Fibers

	Length (mm)	Width (μm)	Aspect Ratio (Length/Width)
Softwood	2–4	20–40	50–200
Hardwood	1–2	14–40	30–90

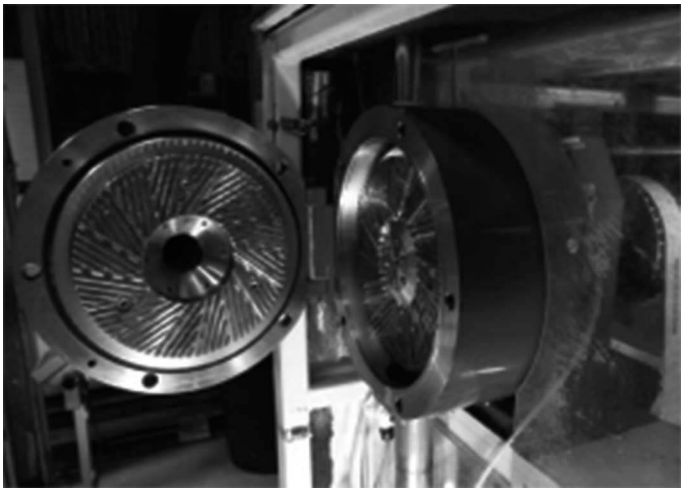


FIGURE 4.28.1 A disk refiner(©CHEZIERE).

(the stator plate). During the operation, the two plates are nearly in contact with gap clearance ranging from about 100 to 200 μm. The design of the disks (bar pattern, groove depth, bar spacing, etc.) and the refining conditions (pressure, flow rate, concentration, gap clearance between the disks, rotational frequency, etc.) are very important and directly impact the compression and shear forces acting on the fiber flocs as well as the energy consumption.

It is worth noting that many paper grades are made from blends of pulps: virgin or recycled pulps, mechanical or chemical pulps, hardwood or softwood pulps, etc. Whenever possible, and because the fibers issued from different sources have different properties, the pulps are disintegrated and refined separately.

After these two steps of disintegration and refining, the pulps are blended in the desired proportions and diluted to low concentrations (ranging from 5 to 10 g/L in most cases) by using water recovered from the sheet-forming section. The diluted pulp is then subjected to several stages of cleaning (centrifugal cleaning and pressure screening) and enters the headbox. The headbox is a fluid disperser device designed to create high levels of turbulence and shear and ensuring two main actions: the rupture of the fiber flocs and the homogeneous and regular delivery of the fiber suspension onto an endless moving wire that constitutes the dewatering fabric. On this fabric, a wet web is formed by filtration of the deposited suspension (see Figure 4.28.2). This web is drained by gravity, thanks to dynamic devices (foils, and possibly rolls at low-speed paper machines) and wet suction boxes until the wet line. Then, a diphasic flow of water and air takes place through the formed sheet, and the drainage is made by dry suction boxes. During the sheet formation, huge quantities of water (named white water) are removed and collected in specific trays from where it is pumped and partly reused for dilution purposes (this circuit constitutes the so-called short loop of the paper machine). The dryness of the wet web increases continuously during the sheet formation and reaches approximately 20% at the end of the moving wire before entering the press section.

It is worth noting here that many papers contain inorganic components in significant quantities. Chemical additives are also added into the fiber suspension from the disintegration stage to the headbox (Eklund and Lindström, 1991; Roberts, 1996; Scott, 1996; Neimo, 1999). The added amount of mineral fillers (ground calcium carbonate, clay, talc, etc.) depends on the paper grade and varies from 0 (blotting papers, tissue, etc.) to 20% approximately (printing and writing papers, for instance). In some specific cases such as decorative papers, higher quantities of inorganic compounds (up to 50%) are added and, quite often, pigments are used for these grades (titanium dioxide,

FIGURE 4.28.2 The forming section of a pilot paper machine (the wet line is visible).

precipitated calcium carbonate). Using the words pigment or filler is a pure convention for papermakers: pigments, which are more refined than fillers in terms of purity, specific surface area, optical properties, are by far more expensive and only used for high added-value papers when introduced into the fibrous suspension. Fillers are generally less expensive than virgin pulps and their introduction allows decreasing the production costs and improving certain properties of the paper (optical properties and printability, for instance). The chemical additives play an important role by improving the productivity of the paper machine (processing chemicals) and the end-use properties of the paper (performance additives). Retention and drainage agents are by far the most important processing chemicals. The combined introduction of cationic and anionic polyelectrolytes, which constitute the most common system used, allows monitoring the retention in small-sized elements (cellulosic fines, mineral fillers, and pigments) during the sheet formation on the forming fabric. Indeed, at this stage, the drainage of large amounts of water combined with the pressure pulsations induced by the drainage elements (rolls, foils, suction boxes) leads to a loss in fine elements. A large fraction of these elements are recycled with the white water in the short loop of the paper machine. Nevertheless, this loss contributes to lengthen the nonsteady periods when changes in the production operating conditions occur and then to decrease the productivity; it also causes an asymmetrical distribution of components in the thickness of the paper sheet, to cite a few of the numerous effects of low retention levels. The increase in the flocculation level of the fiber suspension allows for an entrapment of the finest elements in the formed flocs, which can be easily retained on the fabric. It also has a positive effect on the drainage rate of the suspension that may result in an increase of the production rate. On the other hand, the use of such additives negatively impacts the quality of the sheet formation (mass distribution) and that of the produced paper. In the case of overdosage, the presence of large flocs, from which water cannot be easily removed, may lower the dryness of the wet web entering the press section and such heterogeneities lead to the creation of defects in the paper sheet during pressing. Consequently, it appears that the optimal level of flocculation is a subtle balance influenced by the properties of the fibers, the conditions of turbulence in the headbox, the intensity of the pressure pulsations during the sheet formation, as well as the properties of the chemicals and their dosage in a given application (Eklund and Lindström, 1991; Scott, 1996; Neimo, 1999; Rojas and Hubbe, 2004; Hubbe et al., 2009).

After forming, the web is conveyed from the formation fabric to the press section where it passes through a series of nips formed by two rolls running under pressure to remove a significant fraction of the remaining water. Porous belts, called felts, support the sheet in the nips and contribute to dewatering (see Figure 4.28.3). The dryness increases again and, depending on the properties

FIGURE 4.28.3 Example of a press section of a pilot paper machine (the rolls, the nip, the belt, and the paper sheet are visible).

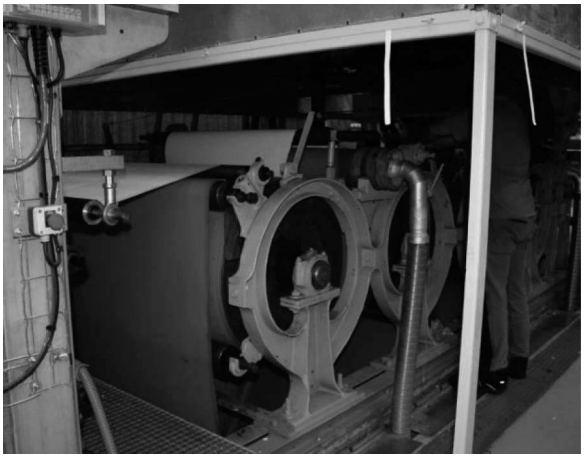


FIGURE 4.28.4 Example of a drying section of a pilot paper machine.

of the wet web and the technology used, is comprised between 40% and 50% at the end of the press section.

The final section corresponds to the drying stage. In most cases, rotating steam-heated cylinders fed with low pressure steam are used for drying (see Figure 4.28.4). Steam pressure is generally below 6 bar and so the surface temperature of the dryers does not exceed 130°C. Dryer fabrics are used in order to improve the contact area between the wet sheet and the drying cylinders and, as will be explained later, the drying fabrics may be used to carry the web in the paths between two cylinders. Paper drying is then predominantly based on conduction heat transfer but convection also plays a role, particularly during the travel of the sheet between two cylinders. There is indeed a relative movement between the web and the surrounding air and the drying section is ventilated to extract the water vapor evaporated from the sheet. Combining forced convection (impingement drying) to contact drying on drying cylinders is also a method to improve the capacity of a multi-cylinder drying section. Different technical solutions are available on the market. Though radiation heat transfer (infrared drying) may also be considered in order to improve the drying capacity of a paper machine, and thus its performance, it is in fact rarely implemented in the conventional drying section. It is more often used to even out the moisture profile before online surface treatment or at the end of the drying section as an alternative to remoisturizing the web, for instance. In this case, systems equipped with low power density infrared, acting over 150 mm width, may be used. The main application of infrared in paper making is to perform contactless drying after surface treatment. The dryness of the sheet after drying is generally comprised between 92% and 95%, each grade having specific targets. The drying operation will be described in detail in the third section of this chapter.

Finally, the paper enters the reel section where it is rolled up. The obtained parent roll is removed from the reeling machine and transferred to other places where the next operations can be conducted. It is worth noting that some finishing operations aiming at improving paper surface properties may also be performed online. In order to summarize the description of the paper making process, which transforms a fibrous suspension into a dry paper, a very simplified drawing of the paper machine is shown in Figure 4.28.5.

To conclude, and before describing the finishing and converting operations, it is important to point out here that the characteristics of the paper machines mainly depend on the grade of the paper produced and particularly on its basis weight. Paper machines can run at velocities ranging from 300 m/min (for carton boards, for instance) to 1500 m/min (for printing papers) or even at higher velocities. Their size may also vary considerably in terms of width and length. Paper machine may be from 50 to 150 m long, the drying section being the longest, up to 100 m long in certain cases.

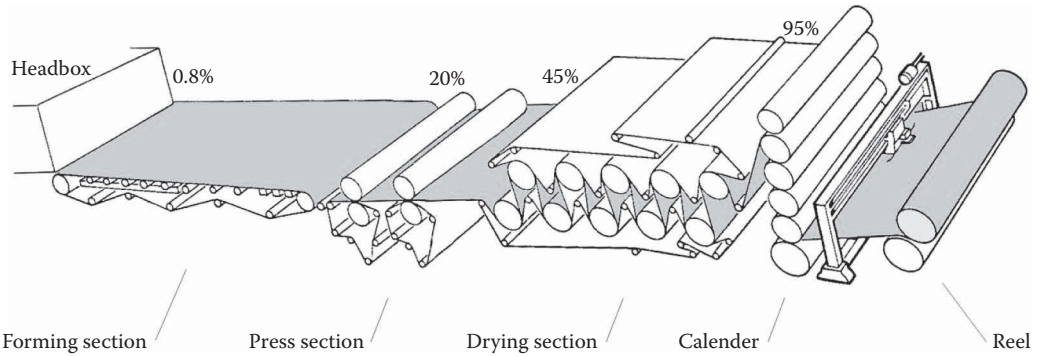


FIGURE 4.28.5 Very simplified drawing of a paper machine.

Paper sheets from 1 to 10 m wide can be produced. On modern paper machines, production rates can thus reach 40 tons/h. The paper making process involves the use of huge quantities of water: working at a low concentration is required in order to limit flocculation of the suspension resulting from the relatively high aspect ratio of the lignocellulosic fibers. Paper making can thus be viewed as a process allowing the removal of water from a fiber suspension. As a first approximation, 200 kg of water per kg of fibers are drained in the forming section, whereas the amount of removed water is close to 2 kg/kg and 1 kg/kg in the press section and the drying section, respectively (Ghosh, 2011; Karlsson, 2009, Chap. 2, p. 40). In most cases, the water is recycled in the process and, today, most paper machines allow producing paper with reduced consumption of water, generally $<5\ m^3$ of water per ton of paper produced. Certain paper machines work with fully closed circuits and the water consumption is reduced to nearly zero. Table 4.28.2 summarizes the amounts of water present at different locations of the paper machine.

The energy consumption of a paper machine is high and generally varies between 200 and 600 kWh/ton of paper produced for electricity and between 1300 and 1800 kWh/ton for thermal energy (McIvor et al., 1999; Molin et al., 2010). The largest consumption of thermal energy is in the drying section that represents about 40% of the total capital costs of a paper machine (Ghosh, 2011).

Finishing and Converting Operations

Sizing, coating, and calendering are considered as finishing operations. Generally, sizing consists of applying to the paper surface a thin aqueous layer of starch that may contain mineral pigments. To this purpose, the dried sheet passes through two soft rolls of the so-called size press and the size solution is metered in the nip. The amount deposited is generally comprised between 2 and 10 g of dry matter per square meter of paper. The objective is to improve the properties of paper

TABLE 4.28.2
Examples of Involved Quantities of Water and Dryness at Different Locations of the Paper Machine

	Dryness (%)	Moisture Content (kg/kg Dry Matter)	Cumulated Fractions of Eliminated Water (%)
Headbox	0.8	124	0
Wet line	2	49	60.5
Dry suction boxes	14	6	95.2
Couch roll	20	4	96.8
Press section	40	1.5	98.8
Drying section	95	0.05	>99.9

by increasing its strength and limiting dusting problems. This operation also decreases the surface porosity and, consequently, slows down the penetration rate of liquids into the sheet. Specific chemicals can also be added to decrease the surface energy of paper. A great number of paper grades undergo such a treatment from printing papers to liners or fluting media. Compared to sizing, coating aims at modifying the surface properties of the sheet in a more drastic way. The formulation of the deposited dispersion is complex: it is made of water, mineral pigments, binders (polymer emulsions, water-soluble polymers), and additives (thickeners, dyes, optical brightening agents, etc.). Coated papers are printing grades (magazines, paper boards, etc.), mainly because the printing quality is significantly improved after coating and calendering resulting in very good properties in terms of surface smoothness, brightness, opacity, surface energy, among others. Coaters are based on several techniques (roll coater, air-knife coater, blade coater, etc.) depending on the required properties, the targeted quantity deposited during the coating, etc. The basis weight of the dry coated layer ranges from 10 to 30 g/m² and, when necessary, several coating layers can be deposited so that the final basis weight can reach more than 100 g/m². Both sizing and coating operations require the presence of a devoted drying section combining contact-free drying devices to preserve the wet coating layer, which could adhere on the dryer shell (infrared and air dryers) and conventional drying cylinders (Hung et al., 2014).

Calendering, which is based on a mechanical treatment, also aims at improving the quality of the paper surface and leveling the thickness profile. This operation is carried out on steel rolls disposed in a stack in which paper is subjected to friction and compression forces. The main operating parameters impacting the effect of calendering are the nip pressure, the number of nips, and the retention time. Supercalendering is a more intensive operation generally used for coated papers. In this case, paper passes between soft (polymer, compressed paper, or cotton) and hard (steel) rolls disposed alternately in a stack. The operation involves high pressure and the effect on paper is important: very glossy papers can be obtained by supercalendering. Calendering is generally performed on the paper sheet just before it is rolled up, whereas supercalendering is performed offline. Several other techniques of calendering exist, but they are not detailed here: softcalendering, extended-nip calendering, etc. The effect of calendering is more pronounced if the moisture content in the paper sheet is increased and if the operation is conducted at a high temperature. This is generally done by spraying water on the paper just before the nip and by heating the rolls to temperatures ranging from 80°C to 200°C, depending on the desired intensity of the operation. At such temperatures, water acts as a plasticizer and facilitates the softening of the wood polymers as well as that of the coated polymers, thus increasing the efficiency of the calendering. Most often, the calendering operation aims at modifying the surface only and not the bulk of the paper: supercalendering must be conducted so that the increase of moisture content in the paper and of the temperature remains localized at the surface of the sheet.

Depending on their use, paper or boards can be subjected to various converting treatments and, among these treatments, one operation is particularly important because it refers to a grade of paper used in everyday life. This is thermoforming on which the production of corrugated medium is based. Such an operation is performed on a corrugator fed with several sheets of paper. Some of the sheets (the fluting medium) are subjected to pressure by being pressed in between heated corrugated metal rolls, their surface temperature being around 180°C. A corrugated paper is then obtained and, once a suspension of starch granules has been deposited on the tips of the flutes, the fluting paper is joined to the liner to form a complex by gluing. This operation is repeated on the other side of the corrugated sheet to form a corrugated board.

Calendering and thermoforming of the sheet are thus ensured by the fact that paper is a thermoplastic material. The softening of paper results from the combined action of heat and water. In the dry state, the wood polymers have very close softening properties. The glass transition temperatures of amorphous cellulose, hemicelluloses, and lignin are about 200°C. When the water content increases, the glass transition temperature of the polysaccharides may reach ambient or subambient temperatures, whereas that of lignin is about 100°C (Back and Salmen, 1982; Szczesniak et al.,

2008; Paes et al., 2010; Horwath et al., 2011). For these finishing and converting operations, heat transfer through the paper thickness, moisture content, and thermal properties of the sheet, therefore, play a major role. Nevertheless, the operating temperature must be adapted to avoid thermal degradation of the fibers that begins at about 220°C (Yao et al., 2008).

The Cellulosic Fibers and Their Behavior during the Paper Making Process

As already mentioned, cellulose, hemicelluloses, and lignin are the main constituents of all wood species. The cellulose, which is a linear polymer of cellobiose, constitutes the strength-bearing component of the fiber (see Figure 4.28.6). In its pristine natural state, its degree of polymerization may be very high (10,000 or more) depending on the considered vegetal species. Intra- and intermolecule hydrogen bonds impart to the cellulose a high level of organization and crystallinity. Cellulose molecules are thus strongly linked to form supramolecular assemblies such as microfibrils in which crystalline and amorphous regions alternate. Amorphous zones result from defects occurring during the growth of the trees and contribute to a large extent to decrease the strength of such materials.

Cellulose microfibrils are embedded in an amorphous matrix of hemicelluloses whose chemical composition and structure can significantly vary from one vegetal species to another. Hemicelluloses are by far more accessible to chemicals than cellulose and consequently more reactive. They are made of branched polysaccharides of hexoses or pentoses. Hemicelluloses play a crucial role in paper making because they may bear functional groups (carboxylic groups) that can be ionized in the conventional conditions of pH used for the production of paper (that is between 5.5 and 8 in most cases), thus conferring an anionic charge to the fibers. Lignin is a very complex phenolic macromolecule whose composition also differs by species. In wood, lignin can be viewed as a binder between the fibers. It imparts rigidity to the material, but it also confers resistance to wood against microbiological attacks. The wood fibers are, therefore, formed by the association of cellulose microfibrils and matrices of hemicelluloses and lignin. The concentration of these constituents varies in the fiber cell walls, which comprise three main layers: the middle lamella at the outside of the fiber, the primary and the secondary walls (see Figure 4.28.7). The middle lamella is rich in lignin and ensures the cohesiveness between the fibers. The primary wall also contains lignin in significant quantity together with cellulose and hemicelluloses. In the primary wall, cellulose exhibits lower degrees of crystallinity and polymerization than in the secondary wall, which constitutes the thickest region of the fiber. The secondary wall is divided in three sublayers (S_1, S_2, and S_3) with different compositions and thicknesses. The sublayers are very rich in cellulose and hemicelluloses. The orientation of the cellulose fibrils varies from sublayer to sublayer and this organization largely contributes to the strength of the cellulosic fiber. The lumen corresponds to the central opening of the fiber.

The proportions of the three main components in wood are on average 40%, 35%, and 25% for cellulose, hemicelluloses, and lignin, respectively (%, w/w), as illustrated in Table 4.28.3.

Fibers used in paper making can be obtained from wood by chemical or mechanical treatment. Chemical pulping of wood consists of dissolving lignin by removing the middle lamella: it is the most efficient way to get well-separated individual fibers exhibiting good properties. Kraft pulping uses a mixture of sodium hydroxide and sodium sulfide. It is the preponderant process with more than 90% of the chemical pulps produced in the world because it is based on a very efficient chemical recovery system. The other process is the acid sulfite pulping, which is used for producing

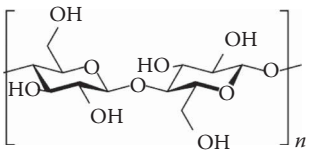


FIGURE 4.28.6 The chemical structure of cellulose.

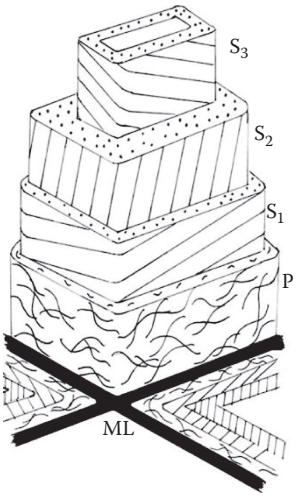


FIGURE 4.28.7 Ultrastructure of a wood fiber (*ML*, middle lamella; *P*, primary wall; S_1, S_2, S_3: sublayers of the secondary wall).

TABLE 4.28.3
Examples of Chemical Composition of Wood Species

	Spruce	Pine	Birch
Cellulose	41	40	40
Hemicelluloses	30	29	37
Glucomannan	18	17	3
Xylan	8	7	30
Others	4	5	4
Lignin	27	27	20
Extractive substances	2	4	3
Total	100	100	100

Source: Eklund, D., and T. Lindström, *Paper Chemistry. An Introduction*, Grankulla: DT PaperScience Publications, 1991.
Note: Weight fraction in %.

pulps with high content in cellulose, hemicelluloses being quickly degraded in acidic conditions (see Table 4.28.4).

In the mechanical processes, fiber separation is caused by shear forces exerted on pieces of wood, which can be pretreated (by hot water, steam, or chemicals) in order to soften the lignin and facilitate the grinding process. Mechanical defibering process does not allow a good separation of the fibers: the pulp contains shives or bundles of fibers and also a high quantity of fines whose production results from the very intense shear forces occurring during the operation. The properties of the obtained pulps are thus different from those of the chemical pulps. The fibers have less affinity to water, are less flexible, and form paper with relatively poor mechanical properties. The yields of the mechanical pulping processes are high (more than 90%) when compared to those encountered in chemical pulping (about 50%) where most of the lignin and simultaneously a fraction of the hemicelluloses are dissolved.

The fibers obtained after pulping (mechanical or chemical) are more or less brown, depending on the process used. This color is mainly due to the presence of lignin and may even be more

TABLE 4.28.4
Examples of Chemical Composition of Chemical Pulps

	Pine Kraft	Birch Kraft	Birch Sulfite
Cellulose	35	34	40
Hemicelluloses	9	17	6
Glucomannan	4	1	1
Xylan	5	16	5
Others	—	—	—
Lignin	3	2	2
Extractive substances	0.5	0.5	1
Total	47	53	49

Source: Eklund, D., and T. Lindström, *Paper Chemistry. An Introduction*, Grankulla: DT PaperScience Publications, 1991.
Note: Weight fraction in %.

pronounced after a kraft chemical pulping, for instance. Removing the remaining lignin or converting the chromophore groups to colorless ones is thus important when white grades of pulps are required. This is done by using oxidants (ClO_2, H_2O_2, O_2, Na_2O_2, etc.) during the bleaching operation. After the pulping and bleaching operations, the obtained pulps are pressed, dried, and packaged in the form of bales. The combined effect of the chemical pulping and bleaching decreases the degree of polymerization of the cellulose polymers to about 1000–1500.

Besides the virgin pulps, recycled pulps play a major role today. They represent more than 50% of the sources of fibers for the paper making industry: Today, this proportion exceeds 60% in Europe. Recycled papers from various sources (magazines for the low-quality grades or kraft linerboards for high-quality grades, for instance) are disintegrated in the paper mill, in pulpers working at a high consistency (around 10%, w/w). Compared to the production of paper from virgin pulps, the use of recycled fibers requires intense cleaning stages, also operating at a high consistency (between 3% and 5%, w/w). Finally, although refining is a key unit operation for virgin pulps, it is performed under very moderate conditions for recycled fibers that are characterized by their weakness. It is worth noting that recycled fibers can be composed of mechanical or chemical fibers or a mixture of these two types of fibers and also that their properties and behavior are neither easily predictable nor reproducible.

The affinity of wood fibers (and nonwood fibers as well) with water is a key parameter of the paper making process. Once immersed in water, the cellulosic fibers swell almost instantly and this affinity helps their dispersion, which is also favored by the presence of anionic charges on their surface (Eklund and Lindström, 1991; Roberts, 1996; Hubbe et al., 2007). The behavior of fibers in water depends on several factors, the most important one being their chemical composition and the accessibility of the fiber wall to water. Thus, thanks to the presence of OH groups onto cellulose and hemicelluloses, hydrogen bonds can be formed between the water molecules and the main constituents of the fiber, both at its surface and in its wall. Lignin that is more hydrophobic limits the sorption of water. Moreover, during chemical pulping, the lignin present in the primary and secondary walls is removed to a great extent. This elimination, more or less easy depending on how the lignin and the carbohydrates are bound, leads to the creation of voids or pores inside the fiber walls. These pores also favor the penetration of water. Consequently, the tendency of chemical fibers to swell in water is enhanced when compared to that of mechanical fibers, which do not undergo the same phenomena during pulping and which contain high amounts of lignin. Ionized chemical groups also promote interactions of the fibers with water molecules. Hardwood fibers are more charged than softwood fibers due to the presence of hemicelluloses containing a greater amount of carboxylic functions and they therefore exhibit a greater tendency to swelling compared to softwood. Sulfonic

acidic groups, ionized whatever the pH, must also be taken into account for certain high yield pulps; they result from the fact that wood chips can be impregnated by sodium sulfite in order to soften them and facilitate the subsequent operation of defibering. It is also important to mention here that the refining stage plays a major role: by promoting external fibrillation (thus increasing the specific surface area as illustrated in Figure 4.28.8) and internal fibrillation (which corresponds to delamination phenomena of the ultrastructure of the fibers), refining has a considerable effect on fiber swelling. Finally, increasing the pH of the suspension (which leads to an increase in the amount of ionized carboxylic groups) or decreasing the ionic strength (which limits the screening effects of the electrical charges born by the fibers) also favor the affinity of the fibers with water but to a lesser extent than refining.

Water present in the wet web is commonly referred as free and bound water even if a continuum does exist between the totally free water (capillary condensation) and the strongly bound water (adsorbed as a monolayer) as illustrated on Figure 4.28.9. In physisorption (or physical adsorption), intermolecular forces (e.g., van der Waals forces or hydrogen bonding) are responsible for adsorption; the surface is not altered chemically (chemisorption). Multilayer adsorption is frequently observed. The adsorbate uptake increases usually with increased partial pressure of the fluid and decreasing temperature. The water present far from the cellulosic surface, even in the pores of the fiber wall, must be considered as free water. On the contrary, bound water strongly interacts with cellulose and hemicelluloses at very short distance: it is present onto the surface of the fiber (outside and lumen surface) and in the fiber wall.

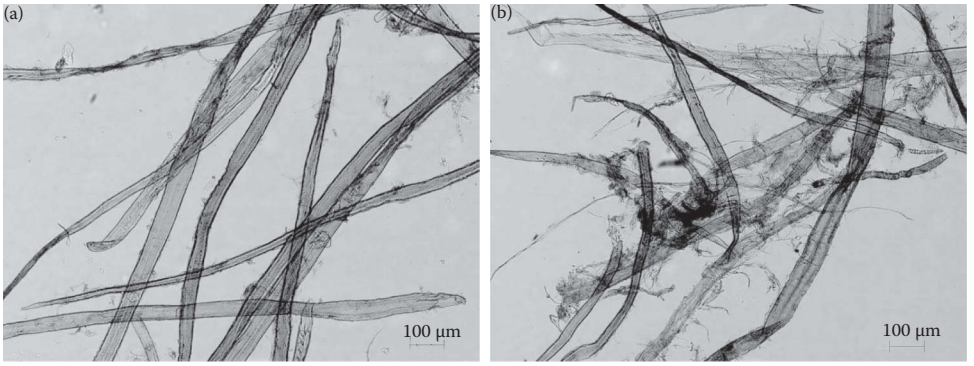


FIGURE 4.28.8 Optical photomicrographs of softwood chemical kraft fibers: slightly refined (a) and highly refined (b).

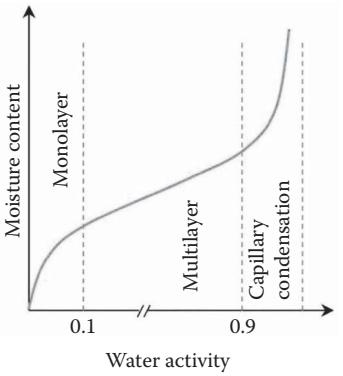


FIGURE 4.28.9 Typical shape of a sorption isotherm showing the different states of water. Activity of water is equal to the relative humidity of the ambient air (see Equation 4.28.2).

Bound and free water show different properties. Viscosity, density, and boiling point of bound water are, for instance, higher than those of free water. In contrast, the dielectric constant is lower for bound water. The free water present between the fibers is relatively easily removed from the wet web and a significant part is eliminated during the forming and pressing operations. However, the water present inside the fiber requires more energy to be removed. The swelling of the fibers may be evaluated by measuring the water retention value (WRV). It corresponds to the amount of water remaining in a small sample of pulp after centrifugation (generally at 3000 g for 10–20 min):

$$\mathrm{WRV} = \frac{\text{Weight of the centrifugated sample}}{\text{Weight of the dried sample}} - 1. \tag{4.28.1}$$

The WRV is then expressed in grams of water per gram of dried fibers. This value lies between 0.8 g/g for unrefined unbleached softwood kraft pulps and 2.0 g/g for refined bleached hardwood kraft pulps. When small-sized cellulosic elements are present in significant amounts in the pulp, the WRV increases because it also includes capillary water retained onto the surface of these elements. Thus, for values greater than 2 g/g, the WRV does not reflect the swelling of the fiber only and it must be considered with caution.

The elimination of water from the fiber wall has several consequences (Niskanen, 1999a; Hubbe et al., 2007): pressing and drying affect the cross-sectional shape of the fibers. Indeed, when the water is removed, the fibers shrink and the lateral shrinkage of the fiber is by far more important (about 15%–30%) than the longitudinal shrinkage (about 2%). Finally, removing the water from the cell wall leads to the creation of new hydrogen bonds between the cellulose chains. These bonds are often not easily broken during subsequent rehydration of the fibers. It implies that a never-dried fiber swells to a greater extent than fibers that have undergone a drying stage. This phenomenon that limits the accessibility of the wall to water is often called hornification. It also modifies the mechanical properties of the fibers. So, when a never-dried single fiber is dried under restraint, its tensile strength increases and its elongation at break decreases, which results from the rearrangement of the polymers in the cell wall (Niskanen, 1999b).

In summary, it is worth noting here that numerous parameters such as the chemical composition of the fibers, the amount of fine elements, the level of refining, the pH of the suspension, and its ionic strength have a direct impact on the swelling of the fibers and, consequently, on the specific energy consumption during its drying.

Paper: Composition, Grades, and Main Properties

A sheet of paper is a complex porous structure made of cellulosic fibers and fines and, in many cases, of mineral fillers or pigments and additives. The paper composition depends on its end-use and three main grades of paper can be considered according to Paulapuro's classification (Paulapuro, 2000): printing and writing papers (copy papers, newsprint papers, magazines, etc.), packaging papers (paper boards, corrugated board, liquid packaging boards, etc.), and specialty papers (tissue, security papers, labels, etc.). Figure 4.28.10 shows scanning electron microscope (SEM) pictures of a cardboard and a coated paper.

Paper is characterized by its basis weight and its apparent density. Typically, basis weight varies from 20 g/m² (for tissue paper, for instance) to more than 400 g/m² (for boards). The apparent density of paper is typically in the range of 0.5–1 g/cm^3, whereas the density of pure cellulose is close to 1.5 g/cm^3. Consequently, the sheet porosity, which is the ratio of the pore volume to the total volume of the paper, ranges from 0.3 to 0.7 in most cases. Papermakers often used the specific volume or bulk, which ranges from 1.2 to 2 cm^3/g. In a paper sheet, the porous network has interfiber pore sizes of the order of 0.5–10 μm (Corte, 1957).

The properties of paper are largely governed by its structure. In a paper sheet produced on an industrial paper machine, the fibers exhibit specific orientations. Due to the forming process, the

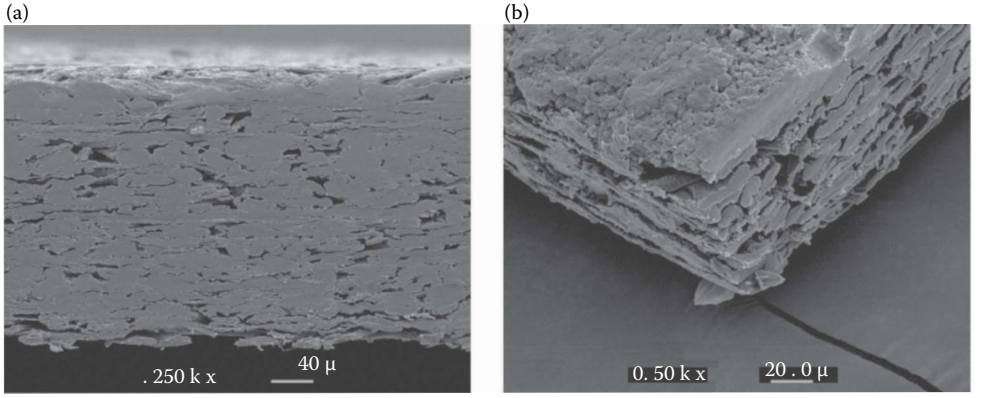


FIGURE 4.28.10 SEM pictures of cardboard (a: view of a cross section) and a high-gloss ink jet paper (b: view of the coated layer and the cross sections).

fibers are thus preferentially oriented with their main axis perpendicular to the thickness of the sheet: the paper can thus be viewed as a layered structure. Moreover, if we consider the plane of the sheet, the fibers also have a preferred orientation in the direction of the production named the machine direction. This orientation can be adjusted; it results from the difference between the velocity of the suspension leaving the headbox and that of the forming fabric. Increasing this difference leads to a more marked orientation. Fiber orientation influences the properties of paper and, depending on the tested direction, paper shows different values in terms of tensile strength, Young's modulus, tearing strength, etc. Moreover, all along its production, the wet paper sheet undergoes strains in the machine direction that results from different velocities between press section rolls, dryer rolls, etc. This strain positively impacts the main mechanical properties of paper in the machine direction. On the other hand, in the cross direction, particularly in the regions close to the edges, and in the transverse direction (or *z*-direction), the sheet can move more freely. The elimination of the water present between the fibers induces capillary forces, which contribute to hold the fibers together. The fibers come into contact as water is removed and the network thus gains strength to withstand external forces. This phenomenon, combined with the shrinkage of the individual fibers, causes a global shrinkage of the fibrous network during drying. The in-plane shrinkage may be noticeable and reach values generally between 2% and 10% in the cross direction. The magnitude of shrinkage in the *z*-direction is much greater. For instance, Wahlström (1999) measured a global shrinkage in the range from 2% to 6% in the cross direction, whereas it was between 40% and 60% in the *z*-direction. Shrinkage depends on the initial swelling of the fibers and on their specific surface area, these properties being largely influenced by the refining of the suspension. The lateral shrinkage of the fibers during drying creates shear stresses and compressive stresses (microcompressions) in the regions of bonding between fibers. Together with the fiber orientation, shrinkage thus induces differences in terms of properties between the cross direction and the machine direction. For instance, the elongation at failure obtained from a tensile test is greater in the cross direction than in the machine direction because of this anisotropic shrinkage. The existence of a shrinkage profile, as already mentioned, also explains why properties can significantly vary along the cross direction. Consequently, elongation at failure in the cross direction is greater for paper strips sampled near the edge of the sheet than that of paper strips sampled in the middle of the sheet. It should be noted that it is possible, in a paper machine, to limit, at least in part, the cross-directional shrinkage but not that in the *z*-direction.

Paper produced on a classical Fourdrinier paper machine also shows an unequal distribution of fine elements in the thickness. This originates from a nonsymmetrical forming; the use of retention agents helps to improve the distribution profile. Besides these considerations, it is important to note that paper does not consist of individual fibers homogeneously distributed in the sheet, but is made of flocs containing fibers and fine elements. Indeed, depending on the formation conditions

(intensity of the turbulence in the headbox and of pressure pulsations in the forming section, dosage of the retention, and drainage agents), the flocculation of the fibers is more or less pronounced, which affects the floc size and the quality of the sheet formation through nonuniform mass distribution. All these phenomena are detailed in Niskanen (1999b).

The mechanical strength of a paper sheet depends on numerous parameters such as the fiber strength, its aspect ratio, the number and nature of bonds between fibers as well as their distribution, which itself is impacted by the sheet formation, the stress distribution, etc. The involved bonds result from different mechanisms and, until now, their relative importance has not been totally elucidated (Lindström et al., 2005; Hubbe, 2006). Hydrogen bonding is generally considered as the most important mechanism but attractive van der Waals forces, attractive and repulsive electrostatic forces acting between charged groups, among others, must also be considered. These forces take place mainly at a molecular level and very short distances, which seems, in a first approach, incompatible with the roughness of the fibers. Nevertheless, several phenomena may justify the creation of bonds, which occur during the progressive elimination of water resulting from the drying of paper. Capillary forces thus play an important role. Their increase during drying is a driving force that pulls the fibers together thus allowing close contact and bonding. The diffusion mechanism corresponds to the extension in water of segments of macromolecules or fibrils out of the surface of the fibers. When two surfaces are brought closer together, macromolecule or fibril segments may come into contact so that bonds are created. This explains why swollen macromolecules such as hemicelluloses are crucial for obtaining paper with high strength. Mechanical entanglement of fibrils and fibers also helps to improve the strength. This discussion emphasizes the importance of the conformability and flexibility of the fibers and of the presence of fibrils at the fiber surface for an adequate formation of contact surfaces. As discussed earlier, these properties are influenced by the refining stage, which explains why this operation is fundamental in order to increase paper strength.

Because paper is a hygroscopic material, its moisture content depends on the relative humidity of the surrounding atmosphere. Relative humidity is defined as the ratio between the ambient partial pressure of water vapor, e', and the partial pressure of water vapor at saturation at the same total pressure and temperature, e'_w:

$$RH\% = \frac{e'(p,T,r)}{e'_w(p,T)} \times 100, \tag{4.28.2}$$

where e'_w depends on the total pressure of the system, p, and temperature, T. It is very close to the saturation pressure of pure water at low temperature. The partial pressure of water vapor in the air, e', depends also on the mixing ratio of the air (r is the ratio of the mass of water vapor in the air to the mass of dry air).

At equilibrium, the relative humidity of the surrounding air and the activity of water in the material are equal. Working at constant pressure (atmospheric pressure), the total pressure of the system is omitted in the variables. The activity of water in the material, a_w, which depends then on the temperature and moisture content, X, of the material, is

$$a_w(T,X) = \frac{e'(T,X)}{p_{sat}(T)}, \tag{4.28.3}$$

where e' is the partial pressure of water vapor in the material (considering the comment on pressure, this partial pressure is a function of the temperature and moisture content) and p_{sat} is the saturation pressure of pure water.

Sorption isotherms provide the relationship between the equilibrium moisture content in the paper sheet and the relative humidity of the ambient air (or water activity in the material) for a given temperature (Figure 4.28.11). The general shape of paper sorption/desorption isotherms looks like type II of IUPAC (International Union of Pure and Applied Chemistry) classification. Sorption

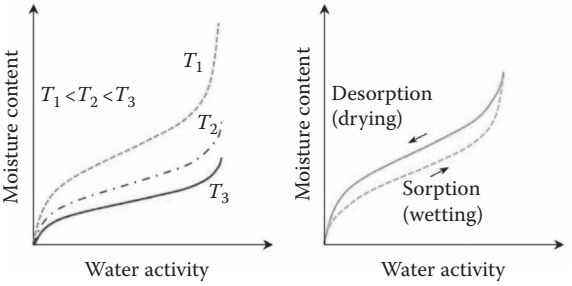


FIGURE 4.28.11 Typical shape of sorption isotherms at different temperatures (left) and hysteresis effect (right). Activity of water is equal to the relative humidity of the ambient air (see Equation 4.28.2).

isotherms may be different for water sorption and desorption, particularly during the first expositions of paper to various atmospheres (Figure 4.28.11). This hysteresis is due to the fact that a total rewetting of paper after a first drying is not possible, because of the irreversible formation of bonds between cellulose chains in the walls of the fibers. This phenomenon was already described when hornification was defined. Other aspects of the hygroscopic behavior of paper are discussed later.

The properties of paper are influenced by its moisture content and decreasing the amount of water sorbed by a paper is a way to increase most of its mechanical properties. Nevertheless, when paper becomes too dry, such as for relative humidities below 15% and moisture contents close to zero, it becomes very brittle and the mechanical properties may be significantly reduced. This phenomenon emphasizes the fact that water acts as a plasticizer in the sheet structure. Consequently, overdrying can be detrimental and this is why paper dryness above 95% should be avoided.

An increase in the moisture content of paper also induces hygroexpansion, which corresponds to the strain measured for a given change in relative humidity. This phenomenon depends on the ability of the paper to sorb water and on the degree of bonding of the fibrous network. Consequently, the composition of the paper needs to be considered: mineral fillers or addition of chemicals that limit the water sorption improve the dimensional stability of paper. The composition of the fibers themselves is also important: papers made from chemical pulps exhibit poor dimensional stability compared to those made from mechanical pulps. Increasing the refining level contributes to decrease in the dimensional stability because the degree of bonding is greater. This hygroexpansion is also largely governed by the fiber orientation in the fibrous network and its shrinkage during drying. The dimensional stability of freely dried sheets is thus drastically lower than that observed with a paper dried under restraint. Consequently, hygroexpansion is much higher in the cross direction than in the machine direction. The difference in composition of the two sides of the paper sheet leads to differences of local moisture contents and here again the two sides may have different hygroexpansion. The fact that hygroexpansion locally varies in the sheet leads to specific behaviors such as paper curl, cockles, etc., and also alter the runnability of the subsequent operations undergone by the material. Hygroexpansion is a major concern for paper producers, printers, and converters. Therefore, many investigations have focused on the sorption/desorption isotherms of paper and their link with some physical properties. One can find in the literature a comprehensive list of publications on this subject, most data concerning results at a low temperature, that is to say room temperature.

In summary, drying plays a major role in paper making. It is an energy-consuming operation and its relative share in the production costs of a paper machine ranges between 70% and 80% (Ghosh, 2011). The drying operation also impacts to a great extent the properties of the paper mainly through the anisotropic shrinkage of the sheet occurring when water is eliminated. Mechanical properties and dimensional stability are particularly influenced. Table 4.28.5 shows properties affected by drying for different grades of paper. For all these reasons, a good control of this operation is necessary to limit the increase in the production costs and to control the paper properties for its intended end use.

TABLE 4.28.5
Properties of Various Grades of Papers Affected by Drying

Newsprint	Woodfree Uncoated	Woodfree Coated	SC and LWC
Strength properties	Strength properties	Strength properties	Porosity
Porosity	Curling tendency	Stiffness	Moisture profile
Curling tendency	Cockling tendency	Moisture profile	Strength properties
Moisture profile	Moisture profile	Oil absorption	Stretch at break
Stretch at break	Porosity	Porosity	Shrinkage
Shrinkage	Oil absorption	Dimensional stability	
	Shrinkage		

Source: Karlsson, M., *Papermaking Science and Technology*, Helsinki, FI, Fapet Oy, 2009, p. 85.
SC, supercalendered; LWC, lightweight coated.

Drying of Paper: General Aspects

As already mentioned, drying is the final dewatering step in the paper making process, which is usually implemented when the mechanical water removal techniques have reached their limits; the moisture content of the web entering the drying section is generally between 1.0 and 1.4 kg water per kg dry matter depending on the fibrous composition, refining level of the pulp and technology used. However, in the case when high bulk is desired, moderate pressing is applied and thus drying starts at a relatively high moisture of paper. Drying is an important stage of the manufacturing process because it determines the final structure of the fibrous network and the distribution profile of the small particles. It also plays a significant role on the properties of the end product (mechanical properties, optical properties, etc.) owing to the creation of bonds between fibers and the development of stresses within the fiber walls during consolidation of the material as already mentioned in a previous section of this chapter.

Drying involves simultaneous heat and mass transfer:

- Heat transfer to the material to be dried, allowing water phase change
- Mass transfer from the bulk to the exterior of the material

Contact drying, in combination with convection drying, is the most common method in the paper making industry as shown in Table 4.28.6 (Karlsson, 2009; Polat and Mujumdar, 2014). Conventional drying is performed in a multicylinder drying section that is surrounded by a hood. Figure 4.28.12 illustrates two different organizations of dryers (double felted and single felted) that are generally seen in the drying section of a paper making plant. The reasons for these configurations will be

TABLE 4.28.6
Share of the Drying Technologies by Application

Technology	Share in Industry (%)	Evaporation Rate ($kg/m^2 \cdot h$)	Energy Consumption (MJ/kg evap)	Main Application
Cylinder	85–90	15–35	2.8–3.5	Base paper and boards
Yankee	4–5	30–200	2.8–5.0	Board and tissue
Infrared	3–4	90–120	15.0–8.0	Coatings
Impingement	2–3	40–130	2.8–5.0	Papers and coatings
Through air	1–2	150–550	3.4.4–4.5	Tissue

Source: Karlsson, M., *Papermaking Science and Technology*, Helsinki, FI, Fapet Oy, 2009, p. 21.

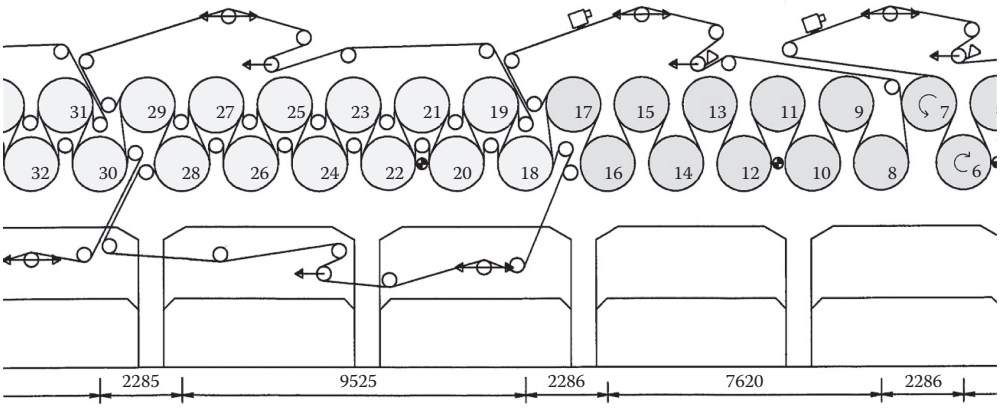


FIGURE 4.28.12 Two typical arrangements of the dryers at a paper machine.

discussed later in this section. Figure 4.28.12 also shows the travel of the paper from one dryer to the next one and the presence of fabrics, the role of which will also be described later.

Other technologies using infrared radiation or convection are used for specific purpose or grades, as shown in Table 4.28.6. The order of magnitude of evaporation rates attained in each type of dryer is reported in the same table. It should be noted that the evaporation rate can be defined by different manners. For instance, for a cylinder dryer, the surface area can be either the surface of paper in contact with the dryer or the lateral area of the dryer. The former is used in Table 4.28.6. It is noticeable that the evaporation rates obtained with all the other technologies are significantly higher than with a conventional steam-heated dryer. However, the reader should be aware that these systems are used for a specific purpose.

In addition, it should be noted that the drying section is equipped with runnability devices, ventilation systems, heat recovery exchangers, etc., and that the dryers are heated with saturated (or slightly overheated) steam; the drying section needs a steam–condensate system. None of these systems are included in Figure 4.28.12. Depending on the design of the hood, the ventilation systems, and also the operating conditions (steam temperature, air temperature, etc.), the performance of a drying section can vary to a certain extent for the same grade of paper.

This chapter is focused on the case of the multicylinder drying section only. After a brief description of the phenomena taking place during paper drying, the role of the different elements and the effect of operating parameters are explained in more detail.

The Different Stages of the Drying Process

A drying curve for contact drying of a capillary porous medium such a paper, performed under laboratory-controlled conditions, can be divided into three different periods:

- Temperature increase: The sheet temperature increases and a temperature gradient develops within the thickness of the sheet. Moisture content only slightly decreases.
- Constant evaporation rate: The wet sheet behaves like a water surface. This period lasts as long as the surface is fed with water from the bulk of the material. The main mechanisms, described by Dreshfield and Han (1956), are the following. Water evaporates at the free surface but vapor is also produced close to the hot surface. While migrating toward the free surface, this vapor is subjected to evaporation–condensation cycles. Liquid water moves toward both surfaces under the effect of capillary pressure. Sheet permeability is therefore an important parameter. Simultaneously, structural modifications take place resulting from substantial shrinkage and increase in air content.

- Decreasing evaporation rate: At a certain moment, the superficial layers in contact with the hot surface are no more saturated with water and the evaporation rate therefore decreases. The evaporation front recedes within the bulk of the material. The surface area in contact with the hot surface decreases and also the thermal conductivity of the layers close to the hot surface, increasing the resistance to heat transfer. As a consequence, the temperature of the sheet decreases until an equilibrium is reached and then it reincreases (Deshfield and Han, 1956; Jones, 1969). According to Dreshfield and Han (1956), the external conditions are less important here than that in the constant rate period. The value of the moisture content when the evaporation rate starts to decrease is called the critical moisture content. Its value depends on various parameters such as the basis weight of the web (the higher the basis weight, the higher the value of the critical moisture content; this is due to the gradients that develop within the sheet thickness), drying rate (the faster the drying, the higher the value of the critical moisture content; this is also due to the gradients within the sheet thickness), heterogeneities of the sheet, refining degree of the fibers or moisture content of the web at the commencement of drying among others. For example, laboratory experiments give values ranging between 0.3 and 0.6 kg water/kg dry matter for papers having a basis weight below 100 g/m^2 and as high as 1–1.2 kg water/kg dry matter for very heavy papers (200–800 g/m^2).
- In the case of convection drying, the same three drying regimes are encountered. During the constant rate period, the sheet is at a temperature close to the wet bulb temperature of the air.

Among the different mechanisms that may contribute to moisture movement during the drying of porous materials, that is to say liquid migration due to gravity, capillary migration, diffusion due to a concentration gradient, diffusion under the effect of a partial pressure gradient and thermomigration, capillary migration of liquid water, and diffusion of water vapor are the most important. It should be noted that, in the case of paper drying, the Fickian diffusion only is considered and not the Knudsen diffusion that occurs when the scale length of a system is comparable to or smaller than the mean free path of the particles involved. However, when the surface of paper has been coated, the Knudsen diffusion cannot be neglected in the coat because of the small pore sizes in this layer (<0.5 μm).

During the constant rate period, mass transfer is limited by the boundary layer at the free surface. The mass flux is given by

$$J_m = \frac{D_{va} \cdot M_v}{R \cdot T_p \cdot \delta_m} \cdot p \cdot \ln \frac{p - e'_a}{p - e'_p}, \tag{4.28.4}$$

where p is the pressure in the air surrounding the dryer, e'_a is the partial pressure of water vapor in the air, e'_p is the partial pressure of water vapor at the sheet surface, D_{va} is the coefficient of diffusion of water in air, T_p is the temperature of paper at the surface, and δ_m is the thickness of the boundary layer. This expression can also be written as

$$J_m = \frac{k_m \cdot M_v}{R \cdot T_p} \cdot p \cdot \ln \frac{p - e'_a}{p - e'_p}, \tag{4.28.5}$$

where k_m is the mass transfer coefficient.

The mass transfer coefficient is generally assessed using the Chilton–Colburn analogy (Bergman et al., 2011). It is expressed as

$$k_m = \frac{h^*_{pa}}{\rho_a \cdot c_{pa} \cdot Le^{2/3}}, \tag{4.28.6}$$

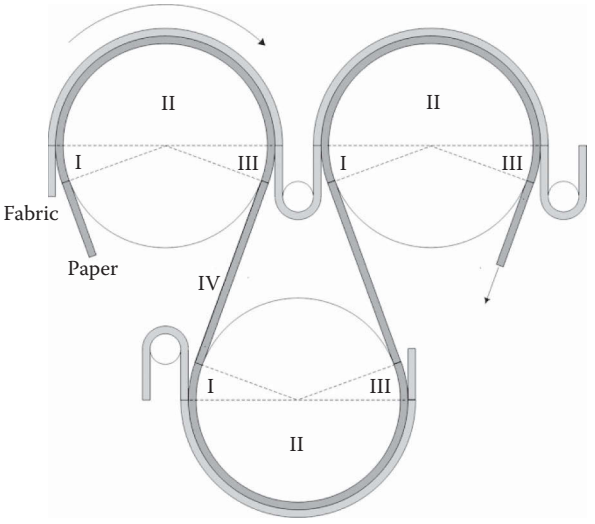


FIGURE 4.28.13 The four drying phases of paper in a drying section, as defined by Nissan (1960, 1961).

where h_{pa}^{*} is the corrective heat transfer coefficient and *Le* is the Lewis number. It should be noted that when the temperature difference between the air and the surface is small, the influence of mass transfer on heat transfer is negligible. When the temperature difference is large, for example, in impingement drying, the corrective heat transfer coefficient in the Chilton–Colburn equation is derived from the heat transfer coefficient modified thanks to Ackermann correction in order to include the effect of evaporation (Karlsson, 2009, Chap. 2, p. 58). However, most computational drying simulations nowadays include the correction factor whatever the evaporation rate (Zvolinschi et al., 2006; Roonprasang, 2008; Ghodbanan et al, 2015).

In a paper machine, the problem is more complex than in laboratory experiments. Figure 4.28.13 shows the different stages of drying as described by Nissan and coworkers (1960, 1961). It should be noted that the wrapping angles may be different in a real paper machine and that the drawing is not to scale. As shown in this figure, the drying process can be divided into four stages:

- I: The paper comes in contact with the dryer face (hot surface).
- II: The fabric covers the paper.
- III: Due to the wrapping angle of the fabric, the paper is no more in contact with this fabric.
- IV: The paper leaves the dryer and travels in the open draw.

This process is repeated as many times as there are dryers. However, the conditions may be different at each location (evolution of the temperature of the sheet and of its moisture content, temperature of the dryer, etc.). Globally, the sheet is warmed up during the phases I–III and cooled down in phase IV because of the evaporation process. Depending on the properties of the fabric, evaporation takes place when the sheet is in contact with the dryer or only in the free draw. So external factors play a role on the kinetics of drying as well as internal factors.

The heat transfer coefficient between air and paper is generally calculated from the correlation expressions for a flat plate (e.g., see Section 3.2.2 in Chapter 3 of this handbook):

- For laminar flow over a flat plate,

$$\text{when } \mathrm{Re} < 5 \cdot 10^5 : \overline{\mathrm{Nu_L}} = 0.664 \cdot \mathrm{Re_L}^{1/2} \cdot \mathrm{Pr}^{1/3}. \tag{4.28.7}$$

- For turbulent flow over a flat plate,

$$\text{when } \mathrm{Re} \geq 5 \cdot 10^5 : \overline{\mathrm{Nu_L}} = \left(0.037 \cdot \mathrm{Re_L}^{4/5} - 871\right) \cdot \mathrm{Pr}^{1/3}, \tag{4.28.8}$$

where $\overline{Nu_L}$ is the average Nusselt number, the values of the Reynolds and Prandtl numbers being evaluated at the mean film temperature. There are also other correlation expressions in the literature based on experimental work for paper machines (Karlsson, 2009) and one should be aware that simulation results should be validated experimentally.

In the drying section of a paper machine, the length of the sheet over a dryer (more than 2 m) and the relative speed between air and paper (at least 4 m/s) are generally sufficient for reaching the turbulent flow conditions. The values of the Prandtl number depend on temperature and air humidity in the surrounding air. For conditions in a conventional drying section, it ranges between 0.7 and 0.8 and the order of magnitude of the heat transfer coefficient is 20–40 W/m^2·K, depending on the machine speed and the state of air.

Other factors such as the presence of the fabric also affect the value of the mass transfer coefficient. The actual coefficient at the surrounding air interface is reduced. In drying simulations, a corrective factor called fabric reduction factor is applied.

In the paths between two cylinders, the correlations for heat transfer over a flat plate are also used. The exchange surface area is increased because the two sides of the web can be exposed to air. In recent publications (Zvolinschi et al., 2006; Ghodbanan et al., 2015), authors also use the Churchill correlation for a bank of cylinders (Bergman et al., 2011):

$$\overline{Nu_D} = 0.3 + \frac{0.62 \cdot Re_D^{1/2} \cdot Pr^{1/3}}{\left[1+\left(0.4/Pr\right)^{2/3}\right]^{1/4}} \cdot \left[1+\left(\frac{Re_D \cdot Pr^{1/3}}{282000}\right)^{5/8}\right]^{4/5}. \tag{4.28.9}$$

Global Heat Transfer Coefficient on the Drying Cylinder

The rate of heat transfer from the hot steam inside the cylinder to the paper on the outside depends on the overall temperature gradient and on the various thermal resistances present in the system:

$$Q = U \cdot A \cdot \left(T_0 - T_7\right). \tag{4.28.10}$$

As shown in Figure 4.28.14, there are a number of resistances to heat transfer in series due to

- The presence of noncondensable gases
- A film of condensate

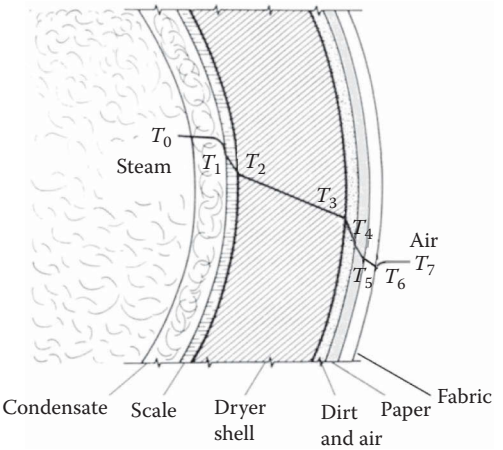


FIGURE 4.28.14 Resistances to heat transfer in paper drying on a steam-heated dryer cylinder.

- Rust and scale on the internal parts of the dryer shell
- Dryer shell
- Sheet-to-dryer contact including
  - Dirt and air between the outside surface of the dryer and the wet paper
  - Paper
  - Dryer fabric
  - Boundary layer of air

Neglecting the presence of noncondensable gases, the condensate is thus the first barrier to heat transfer. The thickness of the condensate layer is greatly affected by dryer speed; as the cylinder rotational speed increases, the contribution of the condensate layer to the overall resistance also increases.

The rust and scale deposited on the dryer shell is responsible for the second thermal resistance. Its value depends both on the nature of dryer shell and the thickness of the fouling layer. Dryers are generally made of cast iron. The thickness of the shell wall being around 25 mm, the thermal resistance due to the shell is approximately $5.3 \cdot 10^{-4}$ or 5.3×10^{-4} K m^2/W.

The third primary resistance to heat transfer is due to the contact between the outside surface of the dryer and the wet paper. Dirt on the outside of the dryer, the thin film of air that separates the dryer from the paper, and the outer surface conditions of the paper are the primary contributors to sheet-to-dryer contact resistance.

The relative contribution of each of these three resistances (condensate layer, dryer shell, and sheet-to-dryer contact) to the overall thermal resistance is greatly dependent on cylinder speed, as depicted in Figure 4.28.15. Pulkowski and Wedel (1988), for example, showed that the condensate layer is the main obstacle to increasing cylinder speeds and thus productivity.

Concerning the cylinder to paper coefficient, experimental studies performed both at the laboratory and field scales show that this heat transfer coefficient is dependent on the sheet moisture content, ranging, for instance, from about 200 W/m^2·K for a dry paper to 1000 W/m^2·K for a wet paper at 1 kg/kg moisture content. The variations are almost linear in this moisture range. However, the numerical values of the coefficients vary among authors (Wilhelmsson, 1990) and also depend on the grade of paper, etc.

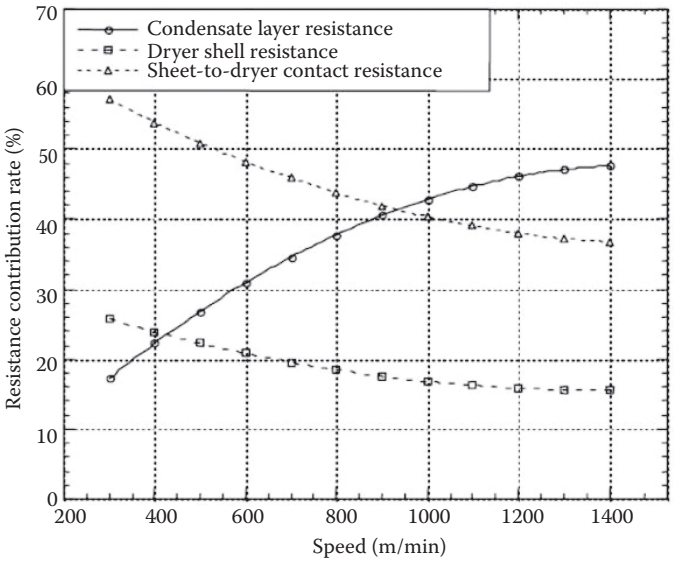


FIGURE 4.28.15 Share of the three main contributors to heat transfer resistance as a function of machine speed.

Organization of the Cylinder Dryers

In the conventional, the so-called double-tier drying section shown in Figure 4.28.12 on the left-hand side (cylinders 18–29), the steam-heated dryers are organized in two rows. The web travels alternately from one cylinder row to the other, for instance, from a cylinder in the bottom row to the next one in the top row and then to the following one in the bottom row and so on until the end of the drying section. Dryer fabrics are used to hold the sheet in close contact with the cylinders through the dryer section. Separate fabrics are used for the top and bottom rows of a dryer group, as shown in Figure 4.28.12.

In the second arrangement, displayed on the right-hand side of the same figure (cylinders 8–17), the rolls are also organized in two rows but the same fabric is used in a mechanical group of dryers. In this configuration, the web is sandwiched in between the dryer and the fabric in the top row and it is the fabric that is in contact with the dryer in the bottom row, the paper web being outside. The first "Unorun" sections were developed and commercialized around 1975 in order to allow high paper machine speeds. Sheet flutter and billowing in the early part of the drying section were indeed one of the major obstacles to achieving higher speeds with the double-tier configuration. As shown in the figure, the web is unsupported in the free draws between two drying cylinders in a double-tier group and, being relatively fragile because of its moisture content, would break under the effect of flutter. Using one fabric only was a solution to carry the web. In addition, there are more unfavorable conditions for heat transfer in the "Unorun" configuration because the fabric thermal resistance between the dryer and the paper decreases the temperature of the sheet, so the bottom row of dryers was no more heated with steam. This configuration is used at the beginning of the drying section in order to reduce production loss, before using two-tier groups when the sheet is dry enough to support the open draws.

The design of the drying section has been significantly improved over the 40 last years. In the single-tier configuration (Figure 4.28.16), there is only one row of dryers; the second one has been replaced by vacuum rolls, with smaller diameters than the cylinder dryers. This change was made for mechanical reasons to counteract the effect of the centrifugal force, which tends to pull the web apart from the dryer; this effect is significant at speed above 1200 m/min. In addition, exposing the free side of the sheet to air is favorable to mass transfer as well as the suction applied in the roll (around 1 kPa).

Modern and rapid machines are equipped with single-tier groups only. The heated groups may be in top or bottom position.

Cylinder dryers were usually 1500 mm in diameter (or less at small paper machines). With the development of single-tier drying groups, dryer diameters have been increased to 1800–2200 mm varying from one manufacturer to another. The suction rolls are smaller in diameter than the drying cylinders, with diameters depending on the manufacturers. This configuration allows increasing the wrapping angle of the fabric on the steam-heated dryers, which is beneficial for the drying process.

Initially, such new configurations were considered to be detrimental to the quality of the end paper because the same side of the sheet is in contact with the dryer surface in a given group and so

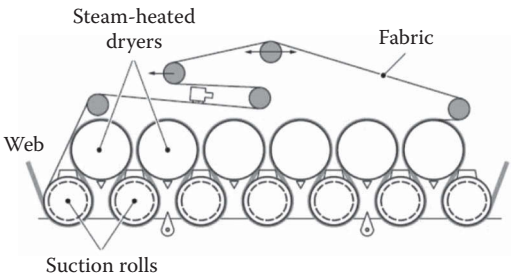


FIGURE 4.28.16 Configuration of a single-tier drying section.

gradients (moisture, temperature, etc.) could develop within the thickness of the sheet. In comparison, in the double-tier configuration, the side in contact with the lateral surface of the dryers alternates at each dryer, which reduces the gradients. Nevertheless, drying the paper on one side only for a certain period of time should be a problem for the high basis weight grades but is not so for the low basis-weight papers, which are concerned by the high production speeds. It has been found that the number of breaks in the drying section was significantly reduced (breaks decrease paper machine productivity) and that some physical properties of the sheet such as dimensional stability, tensile strength, and Young's modulus could be improved.

It should be noted that the single-felted section has been designed to enable increasing paper machine speed. As speed increases, the effect of the positive and negative nips combined with centrifugal forces caused the sheet to drop off the bottom dryer. This causes stretching and wrinkling of the sheet. Web stabilizers have specially been designed to hold the paper sheet in contact with the fabric in the draws between the dryers. The operating principle of modern devices is often based on the Coanda effect. Interested readers are referred to Rouhiainen (2008), Karlsson (2009), and Ghosh (2011) for further details.

Dryer Fabrics

This subsection is only a brief description of the role of the dryer fabrics, which are essential in different ways in the industrial process and some of their influencing properties. Interested readers could refer to, for instance, Adamur (1997) and Lang (2009a,b).The primary function of a dryer fabric is to maintain and/or improve the sheet-to-dryer contact in order to promote heat transfer. The second one is to carry the sheet and to transmit power to undriven rolls. Besides, dryer fabrics limit the cross-directional shrinkage of the paper (shrinkage was discussed in Section 4.28.2). Regarding the first function, there is always a layer of air between the sheet and the dryer that is carried by the moving sheet and dryer. The dryer fabric presses the sheet onto the dryer reducing the thickness of the air layer (as in the lubrication theory) and also minimizing any sheet lifting due to heating the sheet. Dryer felts have been used since 1820, when the paper machines were equipped with dryers. Since the 1960s, these felts have been replaced with fabrics. Dryer fabrics are nowadays woven or spiral materials constructed of synthetic monofilaments (Figure 4.28.17). Using monofilament yarns is extremely important because it allows uniform tension of the fabric in the cross direction of the paper machine owing to the good dimensional stability of monofilaments, increase life time of fabrics, cleanability of the fabric, among others. It should be noted that dryer fabrics are supplied as long strips and joining together of the two ends of the strip to form an endless belt is achieved once it is has been placed at the dryer group. Special seams are developed for this purpose; they must be resistant and should not damage the paper.

Drying fabrics must be heat resistant, rugged, and designed for both drying efficiency and runnability. In order to hold the sheet tightly in contact with the cylinders through the dryer section, adequate tension must be maintained on the fabric. Like a belt, the fabric tension exerts pressure on the paper, whose average value is proportional to the effective tension and inversely proportional to the radius of the cylinder (Karlsson, 2009, Chap. 3, p. 111; Lang, 2009a,b). However, paper is

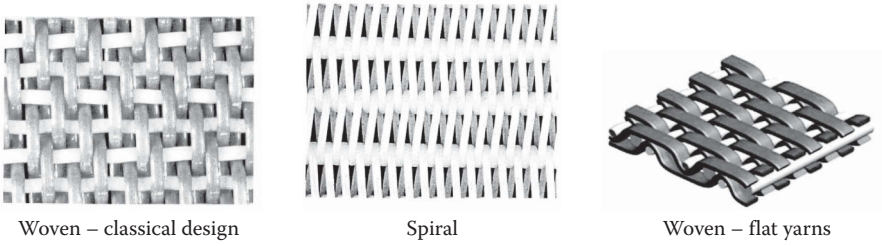


FIGURE 4.28.17 Three examples of dryer fabric patterns.

locally subjected to much higher pressure because the contact area between paper and the fabric is only a very small fraction of the area covered by the fabric (a few percent). The shape of the fabric yarns and the number of contact points are thus important parameters with regard to paper quality. The tension of a dryer fabric used to be in the 1.5 kN/m–2.2 kN/m range but higher fabric tensions (3.5 kN/m and up to 4.5 kN/m) are being utilized on high-speed paper machines to perform well. Appropriate tension values also depend on paper-related characteristics such as basis weight, permeability, and may be different from the wet end to the dry end of the drying section.

In the drying section, air must pass through the fabric. The permeability of the fabric is, therefore, an important factor. It is expressed as the flow rate of air that passes through per unit area of fabric at a given pressure drop (100 Pa which is the SI standard—permeability is expressed in $m^3/m^2 \cdot h$, or 125 Pa generally when permeability is expressed in cubic feet per square foot and per minute [cfm]). Fabric permeability is determined by the structure of the fabric and the sizes and types of yarns used. Dryer fabrics carry air into the dryer pockets due to the effect of entrained boundary layers. For the case of machines with no pocket ventilation system or poorly operating systems, the pumping of entrained boundary layer air into the pocket may reduce moisture levels and consequently improve evaporation rates. The air pumping effect is of no benefit on machines with well-designed ventilation systems and may lead to undesirable effects such as sheet flutter and breaks. The amount of boundary air carried is directly proportional to the speed of the fabric and is strongly dependent on fabric permeability. The aerodynamics of drying fabrics has been little studied (e.g., see Fagerholm, 1990; Widlund et al., 1997; Angrilli et al., 2000). The air-carrying capacity of a drying fabric is characterized by a drag coefficient. It is the ratio of the amounts of air carried, respectively, by the fabric and the reference surface (a glass plate). Typically, dryer fabrics have air drag coefficient values in the range from 2.0 to 14.7 (Adamur, 1997). The air-carrying ability of a common fabric is 2.0 on the paper side and 2.4 on the back side, and that of a flat ribbon fabric with symmetrical weave design is 1.5 (it is the same on both sides). Normally, machines run as high as fabric permeability and air-carrying characteristics as possible, but below the point that they cause sheet handling problems and also because the pocket ventilation systems are designed for a certain range of permeabilities. There is no benefit to use higher permeability fabrics.

Fabric permeability slightly affects heat transfer and thus evaporation when the sheet is in contact with the dryer but in most cases, this is not a major factor, especially with the recent paper machine designs.

Sheet surface quality can also be a limiting factor in the fabric permeability. In most conventional weave patterns, the higher the permeability, the coarser the fabric surface. A coarse fabric surface can damage the sheet surface.

Steam and Condensate Systems

The purpose of the steam and condensate system is to provide a sufficient amount of steam to the dryers and to handle the condensed steam. Dryers are organized in mechanical groups, with one fabric per group, but also in thermal groups fed at the different steam pressures. There are in general four to five thermal groups in a drying section (but the number of groups can be larger, for instance, for heavy grades or high-speed paper machines). Indeed, it is necessary to heat up the web sheet progressively and so the first dryers of the drying section run at relatively low temperatures. In addition, when the paper is wet, there is a risk of picking at the surface of the dryer if the contact temperature is too high. Once the sheet has been warmed up, water evaporates relatively easily (the drying rate depends on internal and external factors). Thus, the next dryer groups constitute the main groups of the drying section and are operated at a slightly higher temperature. Once the critical moisture content has been reached, water is more difficult to remove and it is necessary to increase the drying temperature to maintain a reasonable evaporation rate. This final stage of the drying is made in the last groups of the drying section. In practice, the dryers groups are fed with steam at different pressure, increasing from the wet end of the paper machine to the dry end.

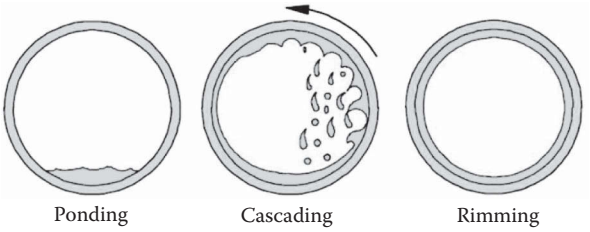


FIGURE 4.28.18 The three main condensate regimes in a steam-heated dryer.

For most printing and copying grades, the steam pressure does not exceed 0.65–1 bar in the first dryer group (Karlsson, 2009, Chap. 3, p. 116). Higher values (3–4 bar) are encountered for linerboard and corrugating medium. The steam pressure in the final groups does not generally exceed 4–6 bar, except for linerboard and corrugating medium (11 bar).

Saturated steam (or slightly superheated steam) is fed into the dryers. The steam condenses on the inner surface of the dryer and forms a layer of condensate. Depending on the speed of the paper machine, the diameter of the cylinder, and also the thickness of the condensate layer, the condensate layer behaves differently (Gavelin, 1964; Karlsson, 2009; Ghosh, 2011) as shown in Figure 4.28.18. This is explained by the balance of the forces acting: gravitational forces, centrifugal force, and frictional forces. At low speed, the condensates accumulate at the bottom of the dryer and form a pond. This is called the ponding or puddle regime. As the dryer speed increases, this pond moves in the direction of dryer rotation and widens. As the speed is further increased, the trailing edge of the pond extends over the horizontal centerline and the condensate tumbles back to the bottom of the dryer cylinder because the gravity force acting on the condensate layer when it arrives at the top of the rotating cylinder is greater than centrifugal force, the film falls down into the puddle. This is called the cascading regime. At higher speeds, the centrifugal force is sufficient and the layer can cover the inner face of the cylinder. This regime is called the rimming regime. Rimming condensate presents a resistance to the flow of heat from steam inside the dryer cylinder to the inner surface of the dryer cylinder.

As a first approximation, it can be demonstrated, from a simple force balance, that the rimming regime is attained when the speed of the dryer is equal to $\sqrt{5 \cdot g \cdot (D/2)}$, where g is the gravitational constant and D is the inner diameter of the cylinder. In practice, other factors such as the thickness of the layer of condensates must also be taken into account. The actual speed is greater. At low speeds, the rimming condensate film does not move at the same speed as the cylinder neither does its velocity constant. As the cylinder surface rises, gravity decelerates the condensate film. As the cylinder surface descends, gravity accelerates the condensate film. An oscillation is created by this alternating acceleration and deceleration. This natural oscillation produces turbulence in the rimming layer, which reduces the heat transfer resistance. The velocity of this oscillation depends on both the speed of the dryer and on the thickness of the condensate film. At high speeds, the oscillation of the condensate film is greatly reduced. As a result, the condensate layer produces a very significant reduction in the transfer of heat from the dryer cylinder, particularly in high-speed operations with thick condensate films: the condensate behaves like a stagnant film. This is why the inner surface of the dryer shells is equipped with turbulence bars (spoiler bars) in order to create microcascades in the condensate film to improve heat transfer eventually, resulting in higher drying rates.

Turbulence bars are not useful for machine speeds below 400 m/min, whereas they improve the drying rate by 5%–12% at 600 m/min and 12%–20% at 900 m/min (Karlsson, 2009, Chap. 3, p. 116).

The spacing of the bars and their cross sections are important parameters. Typically, 18 rows of bars are used in a 1.5 m diameter dryer and 21 rows of bars are used in a 1.8 m diameter dryer. High-performance bars today are 19 mm wide and 12 mm high.

The steam is fed and the condensate removed from the dryers thanks to rotary joints. There is one device per dryer fulfilling both functions except at high-speed paper machines where steam is fed at one side of the paper machine and the condensate removed at the other one. Using multiport systems could help improve the steam distribution inside the dryer and minimize the layer of condensates (Choi et al., 2001; Tim et al., 2010; Shin et al., 2011) but this development remains at the pilot stage.

Above a certain speed, the condensate cannot be removed from the dryer without creating a two-phase flow. So steam must be allowed to pass through the syphon pipe to remove the condensate (and also air and other noncondensable gases). It requires creating a pressure differential between the cylinder and the condensate header. The rate of blow through steam depends on the pressure differential and the frictional resistance of the syphon. The condensate removal tube, equipped with a special shoe, constitutes the syphon. There are two types of syphons:

- The rotary syphon that rotates with the dryer. It is convenient for relatively low speeds, e.g., 300–1000 m/min. However, as the centrifugal force that is opposed to the removal of condensates increases with the rotation speed, this type of syphon requires increasing the differential pressure as the paper machine speed, resulting in an increase of the blow-through steam. At high differential pressures, the amount of blow-through steam may be as high as 50%, or more, of the incoming steam. Therefore, rotary syphons are not suited to paper machines running at velocities more than 600 m/min.
- The stationary syphon that does not move with the dryer. In this case, the weight of the water column in the syphon pipe is the force that opposes to the removal of the condensates. Therefore, the required differential pressure is much lower than that for the rotary syphon. The value only varies little with the rotational speed of the dryer. However, the syphon shoe is more difficult to position.

Typical values of the pressure differential are 0.5–1 bar for rotary syphons and 0.15–0.25 bar for the stationary ones.

The blow-through steam must be reutilized before the condensate is returned to the boiler. The two-phase flow from a group of dryers is collected in a separator maintained at constant level where the blow-through steam is separated from the liquid. In addition, a little quantity of flash steam is produced. The condensate is pumped from the separator, and finally returns to the boiler, and the vapor is reused. In the cascade design, this vapor is reused in a group operated at a lower pressure, except of course for the first thermal group (the vapor from the first separator feeds an air/vapor heat exchanger). The vapor is mixed with some fresh steam (low-pressure steam) in a header in order to feed this group with steam. The cascade system is the predominant configuration. The disadvantage of the interconnections from the cascade configuration is that the pressures and differential pressures are not independent. This drawback can be resolved by a thermocompressor unit. A thermocompressor is an ejector that uses high-pressure steam (motive steam) from the steam header to compress low-pressure steam (here, the vapor from a separator tank) to a desired pressure. In this way, the blow-through steam can be recirculated to the same steam group, making the different steam groups independent. In this configuration makeup steam can also be used, if necessary. Adequate sizing of the thermocompressor is essential and the presence of noncondensable gases is a problem. Another disadvantage is the need for high-pressure steam and, in case of breaks or grades with low drying demand, the blow-through steam is dumped to the condenser.

The steam and condensate system is equipped with a condenser, which condenses the residual vapor before it returns to the boiler. The amount of condensed vapor is an indicator of the performance of the steam–condensate system. It should not be greater than 3% of the total steam consumption.

Hood, Ventilation, and Heat Recovery

Apart from the mechanical problems such as sheet flutter, etc., which would cause sheet breaks, the drying process is not an efficient process without combining appropriate ventilation and air exhaust

system with the dryers. Figures 4.28.12 and 4.28.13, for instance, show that many dead zones, the so-called pockets, exist in the drying section because of the sheet and/or fabric paths. A large part of evaporation from the web takes place in these zones. In order to be able to dry the paper and maintain high-evaporation rates, air must be kept reasonably dry. For this purpose, the dryers are enclosed in a hood, equipped with air systems, which have two main goals: remove the water vapor evaporated from the sheet in the dryer section and control the temperature, humidity, and air flow in the dryer pockets across the full machine width. In addition, the presence of the hood permits to maintain good working conditions in the machine room in terms of heat, humidity, and noise; protect the building and machinery from deterioration because of the humidity; and reduce emissions and mist to the outside of the mill (Gavelin, 1964; Karlsson, 2009; Ghosh, 2011).

A properly designed and operated hood air system will enable to remove vapor uniformly both in the machine and the cross directions. There are three main types of hoods: the semiopen hood, the closed hood, and the high-efficiency hood. The closed hood and high-efficiency hood completely surround the drying section and also the basement. Average values of the ventilation requirements and effect on exhaust air state are given in Table 4.28.7.

In the relatively low-speed paper machines (old machines), warm air is blown at the bottom part of the hood, below the cylinders thanks to blow rods (holes or slots). The temperature of the supply air is generally comprised between 70°C and 90°C and air velocity is slow because the objective is only to ventilate the drying section and avoid condensation in the hood.

In the recent designs of paper machines, the blow rolls or blow boxes are installed outside the pockets to blow air through the fabric. The air flow rate can reach 800 m^3/h per meter width of the paper machine and air is generally blown at a temperature comprised between 90°C and 125°C in the case of blow rolls and up to 145°C for blow boxes. This explains the differences in the properties of the exhaust air and air ratios shown in Table 4.28.7. Of course, the operating conditions of the closed and high-efficiency hoods improve the energy performance of the paper machine. At comparable evaporation rate (kg of water evaporated per time unit), the data in Table 4.28.7 show that the air exhaust flow rate will be lower with closed hoods compared with a semiopen hood, and thus the number of exhaust ventilators and the electrical power consumption will be lower. The amount of supply air is also lower for closed hoods. However, the blow boxes need high-pressure ventilators, so electrical consumption may not be different. Finally, it should be noted that the difference between the amounts of dry air in the exhaust air and supply air comes from the leakage air that is taken from the machine room (infiltration air). The machine room is conditioned and must be renewed regularly. As the amount of leakage air is lower in the case of closed hoods, another advantage of closed hoods can be pointed out.

The supply air is taken from either the outside of the building or from the machine room. It can be heated by different ways:

- In a heat exchanger, air–exhaust air, air–condensates, air–vapor (from a separator), air–steam, or
- By mixing the air stream with hot air (for instance, from a turbo blower used for the production of vacuum in the wet end) or flue gas (for instance, flue gas from infrared dryers) from another part of the mill.

TABLE 4.28.7
Ventilation Requirements at a Paper Machine Drying Section

	Semiopen	Closed	High Efficiency
Exhaust air temperature, °C	50–60	65–85	80–90
Exhaust air mixing ratio	0.04–0.06	0.12–0.14	0.16–0.18
Exhaust air per kg water evap., kg/kg	15–30	8–13	6–8
Supply air/exhaust air ratio, %	20–30	60–80	70–80

Several heat exchangers are generally installed in series, for example, air–exhaust air followed by air–vapor and then air–steam. Heat recovery from the exhaust air is possible only if the temperature is high enough. About 6% to 10%–15% of the heat at exhaust can be recovered for heating supply air (Smook, 1992; Karlsson, 2009). There are still possibilities to use the waste heat before it is released in the atmosphere.

Physical Properties of Paper Drying

Global Changes in Paper Structure during Drying

Transport mechanisms during the thermal drying of deformable porous media, including paper, are influenced by the structural changes that take place and also depend on the moisture content of the material.

The wet paper is composed of the fibrous material, liquid water that can be present inside the fibers and in the voids between the fibers and a gaseous phase (air and water vapor). Assuming that the volume of the fiber walls is strictly equal to the sum of the volumes of the dry solids and liquid water, and neglecting in-plane shrinkage, the fractions of volume occupied by the solid, liquid, and gaseous phases, expressed in percent, are given as

$$\varepsilon_s = 100 \cdot \frac{W}{\delta} \cdot \frac{1}{\rho_s}, \varepsilon_\ell = 100 \cdot \frac{W}{\delta} \cdot \frac{X}{\rho_\ell}, \varepsilon_g = 100 - 100 \cdot \frac{W}{\delta} \cdot \left(\frac{1}{\rho_s} + \frac{X}{\rho_\ell} \right), \tag{4.28.11}$$

where W is the bone-dry basis weight of the sheet, X is the moisture content, and δ is the thickness of the sheet at moisture X. It should be noted that the thickness of the wet sheet can be measured only at the laboratory scale. At the moisture content where a sheet enters the drying section, the web is no more saturated with water. The volume fraction occupied by the solids increases during drying because of the transverse shrinkage of the sheet, the gaseous volume fraction also increases, and the liquid volume fraction decreases. A typical example is shown in Figure 4.28.19. However, numerical values depend on numerous parameters, type of fibers, refining level of the fibers, pressing intensity, drying conditions, and among others. These variations affect the transport properties of the web drying, effective diffusivity of vapor, capillary flow, etc., as already discussed in the previous section, which are important for paper quality parameters. Shrinkage is largely governed by the interactions between fibers but it may be limited by the design of the dryer (Stenström, 2004).

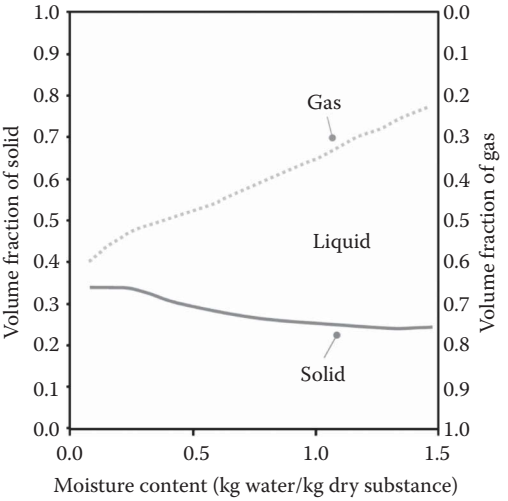


FIGURE 4.28.19 Example of variation of the fractions of volumes of the solid and gaseous phases in a paper during drying.

Specific Heat Capacity

In general, the specific heat capacity c_p of dry (air dry) paper made of pulp only is quite close to that of natural wood pulps or various vegetable fibers, wood, and cellulose, that is, about 1300–1400 kJ/kg·K. Values only slightly vary depending upon the chemical composition of fibers and also, similar to wood, depend on temperature. This is, however, generally neglected at the drying temperatures of the conventional drying sections.

As already mentioned, paper generally contains fillers, such as calcium carbonate or clay (kaolin), or, for special grades, titanium dioxide. These inorganic materials have a lower thermal capacity than the fibers, ranging from 750 to 900 J/kg·K depending on the substance and temperature. For example, the value of the specific thermal capacity of calcium carbonate may be assessed by the following expression $c_{\mathrm{p}} = 828 + 0.5T - 1.3 \cdot 10^7 / T^2$ and that of titanium dioxide by $c_p = 626 + 0.4T - 2.2 \cdot 10^6 / T^2$ (Green and Perry, 2008). The specific heat capacity of clay is 920 J/kg·K. The specific heat capacity of filler-containing paper is, therefore, a little lower than that of the fibers.

Wet paper has a greater specific thermal capacity than dry paper because of the influence of water.

The specific thermal capacity of paper is commonly calculated as the weighed sum of the specific thermal capacities of each component:

$$c_{\mathrm{p\,paper}} = x_{\mathrm{fibers}} \cdot c_{\mathrm{p\,fibers}} + x_{\mathrm{fillers}} \cdot c_{\mathrm{p\,fillers}} + x_{\mathrm{water}} \cdot c_{\mathrm{p\,water}}, \tag{4.28.12}$$

where the symbol x denotes the mass fraction of each component. The interested reader can find values of the specific thermal capacity of various wet cellulosic materials (pulps) in Shanbhag (1981).

Heat of Desorption

Paper is a hygroscopic material as already explained and depicted in Figure 4.28.11. The hygroscopic nature of the fibers also plays a role in the drying of the material. When the dryness is greater than 75%–85% (moisture content $X < 0.3 - 0.25$ kg/kg dry substance), the free water has been evaporated and the remaining water is hygroscopically bound water (physisorption), which requires more energy for its removal. Additional energy is necessary to overcome these bonding effects. For example, at 80°C and 95% dryness, 2800 kJ/kg water is needed to evaporate water from pulp compared to the latent heat of water of about 2300 kJ/kg at the same temperature. This additional energy is called the heat of (de)sorption.

The differential heat of (de)sorption, h_{s}, is the additional heat required to remove water at a given equilibrium moisture content, X:

$$h_{\mathrm{evap}}(T, X) = h_{\mathrm{s}}(T, X) + h_{\mathrm{c}}(T). \tag{4.28.13}$$

The integral heat of (de)sorption is the cumulative additional energy needed to reach a given equilibrium moisture content starting from a very large moisture content value per unit mass of water removed.

Owing to the complex and inhomogeneous structure of pulp fibers, the sorption (desorption) of water cannot be expressed by simple mathematical relations. Values are determined experimentally, either from calorimetric measurements or, more commonly, from (de)sorption isotherms, because this type of measurement is easier to perform. For several cellulosic materials, a saturation of the isosteric heat of sorption has been reported for very low equilibrium moisture contents, that is, to say below 0.04, e.g., see Leuk et al. (2016), from isotherm measurements and also from direct measurements using calorimetry. The common interpretation of these findings is that at some point below monolayer moisture, the average distance between the individual water molecules adsorbed to the surface is large enough so that there is no significant interaction between these molecules. Below this point, the adhesion interaction takes place only between the material surface and the

water molecules; therefore, the adhesion energy saturates. This is a somewhat simplistic explanation because in complex, mesoporous materials like pulp fibers there is no sharp threshold value for monolayer moisture. Nevertheless, the basic idea should still hold.

In drying calculations, even if it is not entirely correct, it is commonly assumed that the equilibrium is described by a (de)sorption isotherm, instantly attained and valid everywhere inside the web. The isotherms are used to compute the local partial pressure of water as a function of the moisture content and temperature and also the desorption energy.

Applying the Clausius–Clapeyron equation, the heat of sorption, which is related to the work needed to compress the vapor from the partial pressure e' to the saturation pressure p_{sat}, is equal to

$$h_s(T,X) = -\frac{R}{M_v}\frac{\partial \ln a_w(T,z)}{\partial(1/T)}\bigg|_{z=X}. \tag{4.28.14}$$

This method requires modeling the (de)sorption isotherms as a function of both temperature and moisture content and then calculate the first partial derivative of the model with respect to $1/T$ at constant moisture content. It should be noted that a high value of goodness of fit is not sufficient to obtain correct values of the (de)sorption heat at a low moisture content, that is to say below 0.08 kg water per kg dry substance. The most commonly used model in recent publications related to paper drying section simulation is the so-called Heikkilä formula (Heikkilä, 1993):

$$a_w(T,X) = 1 - \exp\left[-A \cdot X^B - C \cdot (T - 273.15) \cdot X^D\right],$$

where A, B, C, and D are empirical coefficients.

Other empirical models give similar quality of fit of experimental desorption data (Leuk et al., 2016) such as Anderson and McCarty formula (Anderson and McCarty, 1966)

$$a_w(T,X) = 1 - \exp\left[-(A + B \cdot T) \cdot \exp\left[-(C \cdot T + D) \cdot X\right]\right].$$

It is noteworthy that there are only few data available in the literature for the temperature range covered by the paper industrial drying process (Nelson, 1982; Karlsson, 2009; Rhim and Lee, 2009; Leuk et al., 2016).

Figure 4.28.20 displays an example of desorption energy graph as a function of moisture content for various values of the temperature. These curves clearly show that the heat of desorption significantly increases when the moisture content decreases and it decreases with temperature. From these values, the integrated desorption heat is 55 kJ/kg water at 80°C and 220 kJ/kg water at 25°C when the moisture content is decreased from 0.6 to 0.04 kg/kg dry substance. Thus, the additional heat

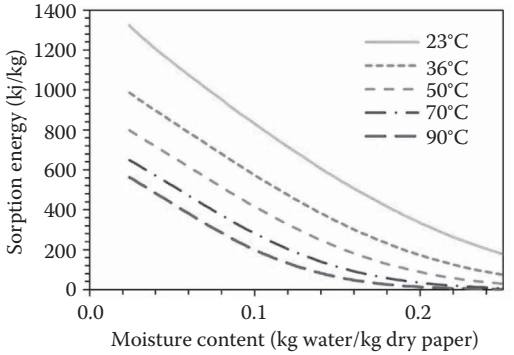


FIGURE 4.28.20 Influence of temperature on the heat of desorption of paper.

needed to desorb water accounts only for a few percent of the energy to evaporate water in industrial conditions; it is useful to recall here that the latent heat of evaporation of water is 2310 kJ/kg at 80°C and 2440 kJ/kg at 25°C. It is, however, not totally negligible and overdrying should be avoided for economic reasons in addition to its effects on the physical properties of paper.

In the case of filler-containing papers, the heat of (de)sorption should be calculated taking into account the mass fractions of the fibers and the fillers and the heat of (de)sorption of each component. The fillers used in paper making are generally not hygroscopic.

Experimental studies show that refined pulp behaves like the same untreated pulp (Seaborg and Stamm, 1931; Leuk et al., 2016), whereas bleaching or predrying decreases sorption heat (Leuk et al., 2016). However, these results do not imply that a paper made of a refined pulp behaves exactly as the paper produced from the same unrefined fibers while drying: there will be more water to remove after pressing and shrinkage will be more pronounced. In addition, it is noteworthy that, in the industrial drying process, the point when water is more difficult to remove may be reached earlier than the value obtained in quasistatic tests because of the gradients that develop in the transverse direction of the sheet (thickness) and nonuniformities in mass distribution.

Thermal Conductivity

Regarding thermal aspects, paper is an orthotropic material because of its anisotropic structure (machine, cross, and transverse directions). The thermal conductivity in the transverse direction is the most studied property, and thermal conductivity values, in the absence of other precision, are related to the property in this direction, i.e., thickness direction. This property is important for many applications where paper is relatively dry, such as copy or calendering. Dry paper is an insulating medium; its thermal conductivity ranges between 0.04 and 0.20 W/m·K. As an example, typical values collected from the literature (Kerekes and Crotogino, 1980; Burnside and Crotogino, 1984; Kartovaara et al., 1985; Lopatkiewicz et al., 2011; Lavrykov and Ramarao, 2012) are collated in Table 4.28.8.

Various parameters exert varying levels of influence on the thermal conductivity of dry (air dry) paper:

- Moisture content: This is due to the high value of water thermal conductivity (0.6 W/m·K). There is, however, no literature available on the variation of the thermal conductivity of paper as a function of moisture content in the low moisture range.

TABLE 4.28.8
Thermal Conductivity of Paper

Paper Grade	Thermal Conductivity (W/m·K)
Uncalendered paper (density 600–1000 kg/m^3, chemical pulp)	0.06–0.11
Uncalendered paper (density 400–600 kg/m^3, mechanical pulp)	0.04–0.068
Calendered paper (density 1000–1100 kg/m^3, chemical pulp)	0.105–0.16
Calendered (density 625–750 kg/m^3, mechanical pulp)	0.05–0.065
Handsheets (density 420 kg/m^3, sulfite softwood pulp)	0.016–0.02
Test liner (density 625–875 kg/m^3, unbleached recycled pulp)	0.075–0.118
Coated (density 737 kg/m^3), woodfree base	0.15
Copy paper (density 625–1260) coated or uncoated	0.08–0.18
Paper (no details)	0.04
Insulating paper for windings	0.10
Pressboard	0.19

- Temperature: Little information is available in the literature on this topic. It was shown by Lavrykov and Ramarao (2007, 2012) that thermal conductivity increases with temperature, which is consistent with the behavior of wood. However, experimental results showed a small variation only, thermal conductivity increased by 0.015 W/m·K when the temperature was increased by 100°C.
- Porosity (or density): The denser the paper, the higher the value of the thermal conductivity, owing to the decrease in air content. Interested readers should refer to the studies by Burnside and Crotogino (1984) and Hestmo et al. (2001). However, porosity and density are global parameters. They do not take into account the arrangement of the solids in the microstructure (Hestmo et al., 2001) or the possible modifications of the fibers during the densification process. In addition, it should be noted that, because refining tends to increase paper density owing to the shrinkage of fiber cross sections during drying and better interfiber bonding (relative bonded area (RBA) increases), papers made of refined pulp have a higher thermal conductivity than those made of unrefined pulps.
- Fibrous composition, nature of the pulp: The influence of the nature of the fibers is not yet completely clear, and results may not be coherent. For instance, the values measured by Kartovaara et al. (1985) do not show significant differences in thermal conductivity for papers produced with different pulps at constant sheet density, whereas Guérin et al. (2001) reported that deinked pulp fibers had a very low-thermal conductivity (0.24 W/m·K) lower than that of bleached kraft pulp (0.4 W/m·K) and thermomechanical pulp (0.58 W/m·K), which explained the relatively low thermal conductivity of papers made of deinked pulp compared to the other two products. The thermal conductivity of wood fibers in the radial direction is about 0.4–0.5 W/m·K. The low value for the deinked pulp probably means degradation of the fiber walls.
- Fiber mix: The thermal conductivity of a mixture of fibers having different values of thermal conductivity can be calculated considering it is equivalent to layers of each type of fibers combined in series (Guérin et al., 2001). In addition, only little information is available in the literature on the influence of the degree of polymerization of cellulose: From measurements performed by Lopatkiewicz et al. (2011) who determined the value of the thermal conductivity of a paper impregnated with mineral oil before (DP = 1300) and after aging (DP = 300), it seems that the degree of polymerization of cellulose is not a significant factor.
- Interfiber bonds: Experimental studies (Kartovaara et al., 1985) consisting of the breakage of interfiber bonds thanks to the tensile tests have shown that the thermal conductivity of paper did not significantly change compared to the untested paper. However, the conclusion that interfiber bonds are not an influencing parameter is valid only if the bonded area remains big enough.
- Presence of mineral fillers: Mineral fillers have a greater thermal conductivity than cellulosic fibers. For example, calcium carbonate (calcite) has a thermal conductivity around 3.5 W/m·K. Therefore, the thermal conductivity of paper increases when the filler or pigment content increases (see, e.g., Lavrykov and Ramarao, 2012) but the most influencing factors remain paper density or porosity. In addition, the location of fillers in the sheet, i.e., in between the fibers or in the pores, may also affect the thermal conductivity values.

Wet paper has a higher thermal conductivity than dry paper because the pores are partly filled with water, which creates thermal bridges within the structure. In drying, the thermal conductivity of wet paper can even be greater than that of water owing to the diffusion phenomena (evaporation–condensation) that occur:

$$\lambda_{\text{eff}} = \lambda + \lambda_{\text{diff}}. \tag{4.28.15}$$

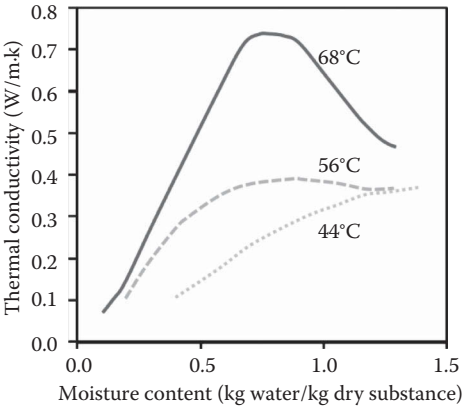


FIGURE 4.28.21 Examples of variation of the thermal conductivity of paper as a function of moisture content.

The diffusion effect is negligible at relatively low temperatures (below 50°C) and noticeable above, as shown, for example, in Figure 4.28.21. It is dominant at temperatures higher than 70°C.

Finally, it should be noted that there is no satisfactory model to predict the thermal conductivity of paper, even in the absence of diffusion, because of the complex structure of the material. Some authors refer to a lumped model considering two layers of a material where the phases in presence (solids, liquid and gas) are arranged in series in the first one and in parallel in the second one. There is, however, no robust method available to assess the relative thickness of each layer.

Conclusion

Drying is an important stage of the paper manufacturing process, which plays a significant role in determining the properties of the end product. The main phenomena taking place are understood but each case is unique, because of the diversity of the raw materials used, and the design of the drying section and the operating conditions at a paper machine. About 90% of the paper produced is dried in a multicylinder drying section that is why this chapter has mainly focused on this type of drying section.

Machine productivity, energy consumption, and product quality are major concerns today. Productivity means good runnability of the process. In that sense, many developments have been incorporated by the manufacturers to offer new solutions (single-tier drying sections, drying fabrics, ventilation systems, sheet stabilizers, etc.). Another point, which has not been developed here, is the reduction of the length of the drying section. In general, it means to be able to improve the drying rate of the cylinder dryers. The only method is to improve the contribution of convection compared to conduction. This is why the dryers may be now covered by impingement hoods (one per dryer). Heat recovery is well implemented in the paper making industry thereby leading to the sustainability. However, the process is not always optimized, diagnostics are often necessary and show that improvements are possible (Zvolinschi et al., 2006; Pettersson and Söderman, 2007; Bhutania et al., 2012).

Another aspect that has not been developed here is the modeling of paper drying. A great number of models are available having various levels of complexity depending on the modeling of the internal phenomena taking place. The relevant literature up to 1999 has been critically analyzed by Sidwall et al. (1999). A current trend is to use models that do not take into account the gradients inside the paper and apply them in a drying section simulator in order to optimize the operation of this section and minimize energy consumption.

REFERENCES

Adamur, S. 1997. *Paper Machine Clothing*. Lancaster, PA: Technomic Publishing Company Inc.

Anderson, N.T., and J.L. McCarthy. 1966. Two parameter isotherm equation for fiber-water systems. *Industrial and Engineering Chemistry, Process Design and Development* 2: 103–105.

Angrilli, F., R. Basso, R. Da Forno, M. De Cecco, G. Fonti and R. Bertolina. 2000. Aerodynamic characterization of dryer fabrics in a wind tunnel. *Tappi Journal* 83(4): 2–7.

Back, E.L., and N.L. Salmen. 1982. Glass transition of wood components hold implications for molding and pulping processes. *Tappi* 65: 107–110.

Bergman, T.L., A.S. Lavine, F.P. Incropera and D.P. DeWitt. 2011. *Fundamentals of Heat and Mass Transfer, 7th edition*. New York: Wiley.

Bhutania, N., C.-F. Lindberg, K. Starr and R. Horton. 2012. Energy assessment of paper machines. *Energy Procedia* 14: 955–963.

Burnside, J.R., and R.H. Crotogino. 1984. Some thermal properties of newsprint and their variations with bulk. *Journal of Pulp and Paper Science* 10(6): J144–J150.

Choi, S.U.S., W. Yu, D.M. France and M.W. Wambsganes. 2001. A novel multiport cylinder dryer. *Tappi Journal* 4(2): 1–15.

Corte, H. 1957. The porous Structure of Paper. *Fundamentals of Papermaking Fibres: Transactions of the Symposium* edited by F. Bolam. Cambridge: British Paper and Board Makers' Association (Technical Section), 301–331.

Dreshfield, A.C., and S.T. Han. 1956. The drying of paper. *Tappi Journal* 39(7): 449–455.

Ek, M., G. Gellerstedt and G. Henriksson, eds. 2009. *Pulp and Paper Chemistry and Technology. Volume 3. Paper Chemistry and Technology*. Berlin: De Gryuter Publications.

Eklund, D., and T. Lindström. 1991. *Paper Chemistry. An Introduction*. Grankulla: DT PaperScience Publications.

Fagerholm, L. 1990. Aerodynamical properties of dryer fabrics for high speed paper machines. *Proceedings of the TAPPI 1990 Engineering Conference*. Atlanta, GA, 165–174.

Gavelin, G. 1964. *Paper and Paperboard Drying. Theory and Practice*. New York: Lockwood Trade Journal Co., Inc.

Gavelin, G., and G.A. Smook. 1998. *Paper Machine Design and Operation: Descriptions and Explanations*. Vancouver: Angus Wilde Publications.

Ghodbanan S., R. Alizadeh and S. Shafiei. 2015. Steady-state modeling of multi-cylinder dryers in a corrugating paper machine. *Drying Technology* 33(12): 1474–1490.

Ghosh, A.K. 2011. Fundamentals of paper drying – theory and application from industrial perspective. In *Evaporation, Condensation and Heat Transfer* edited by A. Ahsan. ISBN: 978-953-307-583-9, InTech, Rijeka, Croatia. doi: 10.5772/21594.

Green, D., and R. Perry eds. 2008. *Perry's Chemical Engineers' Handbook, 8th Edition*. New York: Mc Graw Hill.

Guérin, D., V. Morin, D. Chaussy and J.L. Auriault. 2001. Conductivity of Handsheets, Papers and model coating layers. *Proceedings of the 12th Fundamental Research Symposium, the Science of Papermaking*, Oxford, pp. 927–946.

Heikkilä, P. 1993. A study on the drying process of pigment coated paper webs. PhD thesis. Helsinki University of Technology.

Hestmo, R.H., E. Mikkelsen and M. Lamvik. 2001. Determination of thermal conductivity of newsprint under compression. Thermal conductivity 26. *Proceedings of the 26th International Thermal Conductivity Conference*, Cambridge, MA.

Holik, H., ed. 2013. *Handbook of Paper and Board, 2nd Edition, Volumes 1 and 2*. Weinheim: Wiley VCH.

Horwath, B., P. Peralta, C. Frazier and I. Peszlen. 2011. Thermal softening of transgenic Aspen. *Bioresources* 6(2): 2125–2134.

Hubbe, M.A. 2006. Bonding between cellulosic fibers in the absence and presence of dry-strength agents: a review. *Bioresources* 1(2): 281–318.

Hubbe, M.A., H. Nanko and M.R. McNeal. 2009. Retention aid polymer interactions with cellulosic surfaces and suspensions: A review. *Bioresources* 4(2): 850–906.

Hubbe, M.A., R.A. Venditti and O.J. Rojas. 2007. What happens to cellulose fibers during papermaking and recycling? A review. *Bioresources* 2(4): 739–788.

Hung, J.Y., R.J. Wimberger and A.S. Mujumdar. 2014. Drying of coated webs. Chapter 46. In *Handbook of Industrial Drying, 4th edition* edited by A.S. Mujumdar. Boca Raton, FL: CRC Press.

Jokio, M., ed. 1999. Papermaking. Part 1: Finishing. Book 10. In *Papermaking Science and Technology* Series edited by J. Gullichsen and H. Paulapuro. Helsinki, FI: Fapet Oy.

Jones, G.T. 1969. The hot surface drying of paper. PhD dissertation. University of Manchester, Great Britain.
Karlsson, M., ed. 2009. Papermaking. Part 2: Drying. Book 9. In *Papermaking Science and Technology* Series edited by J. Gullichsen and H. Paulapuro. Helsinki, FI: Fapet Oy.
Kartovaara, I., R. Rajala, M. Luukkala and K. Sipi. 1985. Conduction of heat in paper. *Proceedings of 8th Fundamental Research Symposium*, Oxford. 381–412.
Kerekes, R. 1980. A simple method for determining the thermal conductivity and contact resistance of paper. *Tappi* 63: 137–140.
Lang, I. 2009a. Effect of the dryer fabric on energy consumption in the drying section. *Pulp & Paper Canada* 15(3): 33–37.
Lang, I. 2009b. Dryer fabrics and energy consumption in the drying section. *Proceedings of TAPPI PaperCon'09*, Saint Louis, MO. 7p.
Lavrykov, S., and B.V. Ramarao. 2007. Parameters for heat and moisture transport in paper sheets subjected to a temperature pulse in a hot roll nip. *Proceedings of the 61st Appita Annual Conference and Exhibition*, Gold Coast.
Lavrykov, S., and B.V. Ramarao. 2012. Thermal properties of copy paper sheets. *Drying Technology* 30(3): 297–311.
Leuk, P., M. Schneeberger, U. Hirn and W. Bauer. 2016. Heat of sorption: a comparison between isotherm models and calorimeter measurements of wood pulp. *Drying Technology* 34(5): 563–573.
Lindström, T., L. Wägberg and T. Larsson. 2005. On the nature of joint strength in paper: a review of dry and wet strength resins used in paper manufacturing. *Proceedings of the 13th Fundamental Research Symposium*, Cambridge, pp. 457–562.
Lopatkiewicz, R., Z. Nadolny and P. Przybylek. 2011. The influence of chosen factors on thermal conductivity of paper used as transformer winding insulation. *Proceedings of the XII International Symposium on High Voltage*, Hannover, 5p.
Lundell, F., L.D. Södeberg and P.H. Alfredsson. 2011. Fluid mechanics of papermaking. *Annual Review of Fluid Mechanics* 43: 195–217.
McIvor, A., C. Dahl and R. Lindström. 1999. Paper, board and tissue machines. In *Energy Cost Reduction in the Pulp and Paper Industry* edited by Browne T.C. and P.N. Williamson. Pulp and Paper Research Institute of Canada, Toronto: Agra Simons.
Molin, C., S. Ossin and F. Guillet. 2010. *Maîtrise de l'énergie dans l'industrie papetière*. Grenoble: Centre technique du papier; Angers: Ademe.
Neimo, L., ed. 1999. Papermaking chemistry. Book 4. In *Papermaking Science and Technology* Series edited by J. Gullichsen and H. Paulapuro. Helsinki, FI: Fapet Oy.
Nelson, R.M. 1982. A model for sorption of water vapor by cellulosic materials. *Wood and Fiber Science* 15(1): 8–12.
Niskanen, K. Ed. 1999a. Paper physics. Book 16, Chapter 16, pp. 55–87. In *Papermaking Science and Technology* Series edited by J. Gullichsen and H. Paulapuro. Helsinki, FI: Fapet Oy.
Niskanen, K. Ed. 1999b. Paper physics. Book 16, Chapter 4, pp. 13–53. In *PaperScience and Technology* Series edited by J. Gullichsen and H. Paulapuro. Helsinki, FI: Fapet Oy.
Nissan, A.H., and D. Hansen. 1960. Heat and mass transfer transients in cylinder drying: Part I. Unfelted cylinders. *AIChE Journal* 6(4): 606–611.
Nissan, A.H., and H.H. George. 1961. Heat and mass transfer transients in cylinder drying: Part II. Felted cylinders. *AIChE Journal* 7(4): 635–641.
Paes, S.S., S. Sun, W. MacNaughtan, R. Ibett, J. Ganster, T.J. Foster and J.R. Mitchell. 2010. The glass transition and crystallization of ball milled cellulose. *Cellulose* 17: 693–709.
Paulapuro, H., ed. 1999. Papermaking. Part 1: Stock preparation and wet end. Book 8, pp. 12–252, 285–382. In *Papermaking Science and Technology* Series edited by J. Gullichsen and H. Paulapuro. Helsinki, FI: Fapet Oy.
Paulapuro, H. 2000. *Paper and Board Grades*. Helsinki: Fapet Oy (co-edition with TAPPI Press Atlanta).
Pettersson, F., and F. Söderman. 2007. Design of robust heat recovery systems in paper machines. *Chemical Engineering and Processing* 46: 910–917.
Polat, O., and A.S. Mujumdar. 2014. Drying of pulp and paper. Chapter 39, pp. 773–796. In *Handbook of Industrial Drying,* 4th edition edited by A.S. Mujumdar. Boca Raton, FL: CRC Press.
Pulkowski, J.H., and G.I. Wedel. 1988. The effect of spoiler bars on dryer heat transfer. *Pulp Paper Magazine of Canada* 89(8): 61.
Rhim, J.W., and J.H. Lee. 2009. Thermodynamic analysis of water vapor sorption isotherms and mechanical properties of selected paper-based food packaging materials. *Journal of Food Science* 74(9): E502–E511.
Roberts, J.C. 1996. *The Chemistry of Paper.* Cambridge: The Royal Society of Chemistry.

Rojas, O.J., and M.A. Hubbe. 2004. The dispersion science of papermaking. *Journal of Dispersion Science and Technology* 25(6): 713–732.

Roonprasang, K. 2008. Thermal analysis of multi-cylinder drying section with variant geometry. PhD thesis. Technical University of Dresden.

Rouhiainen, P. 2008. State-of-the-art developments in dryer section runnability. *Proceedings of the TAPPI 2008 PaperCon*. Dallas, TX. 36p.

Scott, W.E. 1996. *Principles of Wet End Chemistry.* Atlanta: TAPPI Press.

Scott, W.E., J.C. Abbott and S. Trosset. 1995. *Properties of Paper: An Introduction*. Atlanta: Tappi Press.

Seaborg, C.O., and A.J. Stamm. 1931. Sorption of water vapor by paper-making materials. I. Effect of beating. *Industrial & Engineering Chemistry* 23: 1271–1275.

Shanbhag, A.S. 1981. Specific heat of wet cellulosic. PhD dissertation. University of Massachusetts Lowell.

Shin, J.H., K.S. Hwang, S.P. Jang and J.W. Lee. 2011. Flow and thermal characteristics of condensing steam in a single horizontal mini-channel of a multiport cylinder dryer. *Drying Technology* 29: 47–54.

Sidwall, S., M. Sadeghi, and W.J.M. Douglas. Comparative structures for simulation of paper drying. *Proceedings of the 1999 TAPPI Engineering/Process and Product Quality*, Cincinnati, OH, 239–270.

Smook, G.A. 1992. *Handbook for Pulp and Paper Technologists, 2nd Edition*. Vancouver: Angus Wilde Publications.

Stenström, S. 2004. Product engineering by the paper dryer. *Proceedings of the 14th International Drying Symposium (IDS 2004)*, São Paulo, A, 89–98.

Szczesniak, L., A. Rachocki and J. Tritt-Goc. 2008. Glass transition temperature and thermal decomposition of cellulose powder. *Cellulose* 15: 445–451.

Tim, J., G. Vedel and D.M. France. 2010. Evaluation of multiport cylindrical dryer. *Proceedings of the TAPPI PaperCon 2010 Conference*. Atlanta, GA.

Wahlström, T. 1999. Influence of shrinkage and stretch during drying on paper properties. Licenciate thesis. Royal Institute of Technology, Stockholm.

Widlund, O.N.G., H.S. Ragvald, C.B. Halldin and N.L.O. Lindquist. 1997. Aerodynamics of high-speed paper machines. *Tappi Journal* 80(4): 113–117.

Wilhelmsson, B. 1990. *A Literature Review of the Contact Coefficient in Multi-Cylinder Drying of Paper.* Lund, Sweden: Lund Institute of Technology.

Yao, F., Q. Wu, Y. Lei, W. Guo and T. Xu. 2008. Thermal decomposition kinetics of natural fibers: activation energy with dynamic thermogravimetric *analysis*. *Polymer Degradation and Stability* 93(1): 90–98.

Zvolinschi, A., E. Johannessen and S. Kjelstrup. 2006. The second-law optimal operation of a paper drying machine. *Chemical Engineering Science* 61: 3653–3662.

4.29 DRYING OF MATERIALS

Pawel Wawrzyniak, Ireneusz Zbicinski, and Mariia Sobulska

Introduction

Drying is one of the most common processes encountered in real life and in industry. Solar drying of fruits, meat, and fish for food preservation was known already in antiquity. In chemical, pharmaceutical, food, ceramic, and textile industries, thermal drying is one of the unit operations controlling quality of the final product and economic efficiency of the technology. The main aim of drying is to decrease the mass of material in order to reduce costs of transportation, change and improve mechanical and physicochemical properties, and increase shelf life of a product. Drying is a complex process comprising simultaneous heat, mass, and momentum transfer accompanied by structural, mechanical, and physicochemical changes in product properties.

The chapter presents basic information about drying of materials. The most common drying techniques, types of dried materials, drying kinetics and equilibria, mechanism of moisture evaporation, and practical applications are described. Principles of the theory of drying covering the properties of humid air, wet materials, drying curves, heat and mass transfer are illustrated and discussed. On the basis of the method of heat delivery to the material (convective, contact, steam, energy field) and hydrodynamics of the gas–material system (disperse, tunnel, belt, rotary, drum, etc.), a classification of dryers is presented.

Drying is an old and well-established technology; relatively few new techniques have been successfully implemented into practice in recent years. One of the examples of novel drying methods is steam drying and disperse techniques like flame, pulse combustion, foam spray drying, which are described in the chapter.

For better understanding of the mechanism of drying process, examples of calculation of selected dryers based on heat and mass balances, equilibria, and drying kinetics are shown.

The chapter delivers also suggestions how to select an optimal drying method and specific dryer for a group of materials.

Principles of Drying

Properties of Air–Water Vapor Mixtures

In most dryers, hot air supplies the heat for evaporation and removal of evaporated moisture from the product. Dryers where heat for drying is delivered in another way are not so common due to high capital and operating costs (freeze, contact, and steam). During drying with hot air, humidification of air and air temperature drop is always observed. Therefore, thermal properties of humid air are essential for the design calculations of such dryers. Table 4.29.1 summarizes the basic thermodynamic and transport properties of an air–water system.

Psychrometry

The capacity of air for moisture removal depends on temperature and humidity. The basic concepts used in psychrometry are as follows:

Total Pressure

The total pressure of a gas mixture, such as nitrogen, oxygen, and water vapor, is made up from the sum of pressures of its constituents called partial pressures. The partial pressure of each component results from the number of moles of the constituent, temperature, and total volume of the system. The partial pressures are added to obtain total pressure.

TABLE 4.29.1
Thermodynamic and Transport Properties of Air–Water System

Property	Expression
P_v	$P_v = 100 \cdot \exp\left[27.0214 - \left(6887 / T_{abs}\right) - 5.32 \cdot \ln\left(T_{abs} / 273.15\right)\right]$
Y	$Y = 0.622 \cdot RH \cdot P_v / \left(P - RH \cdot P_v\right)$
c_{pg}	$c_{pg} = 1.00926 \times 10^3 - 4.0403 \times 10^{-2} T + 6.1759 \times 10^{-4} \cdot T^2 - 4.097 \times 10^{-7} \cdot T^3$
k_g	$k_g = 2.425 \times 10^{-2} - 7.889 \times 10^{-5} \cdot T - 1.790 \times 10^{-8} \cdot T^2 - 8.570 \times 10^{-12} \cdot T^3$
ρ_g	$\rho_g = P \cdot M_g / \left(R \cdot T_{abs}\right)$
μ_g	$\mu_g = 1.691 \times 10^{-5} - 4.984 \times 10^{-8} \cdot T - 3.187 \times 10^{-11} \cdot T^2 - 1.319 \times 10^{-14} \cdot T^3$

Source: Pakowski, Z. and Mujumdar, A.S., *Handbook of Industrial Drying*, Marcel Dekker, New York, 1995.

Dry-Bulb Temperature

Dry-bulb temperature is the temperature measured by a thermometer exposed to humid air. The tip of the thermometer should not be moist.

Wet-Bulb Temperature

Wet-bulb temperature is the temperature reached by a water surface when exposed to humid air passing over. It is usually registered by a thermometer sensor surrounded by a wet wick. The wick and, therefore, the thermo-sensor temperature decreases below the dry-bulb temperature until the rate of heat transfer from the hotter air to the wick is equal to the rate of heat transfer needed for the evaporation of water from the wick to the air stream.

Humidity

Absolute humidity (Y) is the measure of water content of the air. The absolute humidity, called sometimes the humidity ratio, is the mass of water vapor per unit mass of dry air. The units are then kg/kg and this parameter will be later called *humidity*.

Relative Humidity

Relative humidity (RH) is defined as the ratio of the partial pressure of water vapor in the air (p) to the partial pressure of saturated water vapor at the same temperature (p_v). Therefore,

$$RH = \frac{p}{p_v} \tag{4.29.1}$$

RH is often expressed as percentage:

$$RH = \frac{p}{p_v}.100\% \tag{4.29.2}$$

Air is saturated with water vapor at a given temperature and pressure when its humidity reaches maximum under these conditions. Then, the excess of liquid water appears in the form of mist or droplets. In saturation conditions, the partial pressure of water vapor in the air is equal to the pressure of saturated vapor of water at that temperature.

Enthalpy

Enthalpy as an extensive property could be calculated as the sum of partial enthalpies of dry air I_B and water vapor I_{Av}. The enthalpy of humid air I_G is related to the unit mass of dry gas.

$$I_G = I_B + Y \cdot I_{Av} \tag{4.29.3}$$

At temperature T_g, the enthalpy of dry air is

$$I_B = c_B \cdot T_g \tag{4.29.4}$$

and enthalpy of water vapor

$$I_{Av} = c_{Av} \cdot T_g + \Delta H \tag{4.29.5}$$

where c_B and c_{Av} are the mean heat capacities of dry gas and water vapor, respectively, and ΔH is the latent heat of vaporization at 0°C.

Psychrometric Chart

Humidity charts are a convenient way to store humid air data in a graphic form. One of the most popular enthalpy–humidity charts is attributed to Mollier (Figure 4.29.1). This chart, especially suitable for thermal drying calculations, plots the enthalpy of humid air vs. absolute humidity. To increase the area of unsaturated air, the oblique enthalpy line is used.

Properties of Wet Materials

Types of Materials for Drying

The objects of drying are moist bodies with different physicochemical, structural, mechanical, and other properties. Knowledge of the properties of wet materials allows for selection of appropriate drying methods and process parameters. In drying technology, three groups of materials are distinguished:

1. *Capillary porous*: low shrinkage, fragile, e.g., detergents, ceramics
2. *Colloidal*: high shrinkage, e.g., elastic gelatin
3. *Colloidal–capillary porous*: often puffing, maintain elasticity, e.g., wood, paper, leather

This classification has a conventional nature; many materials have characteristics that combine properties of the aforementioned groups. It should be kept in mind that material properties can

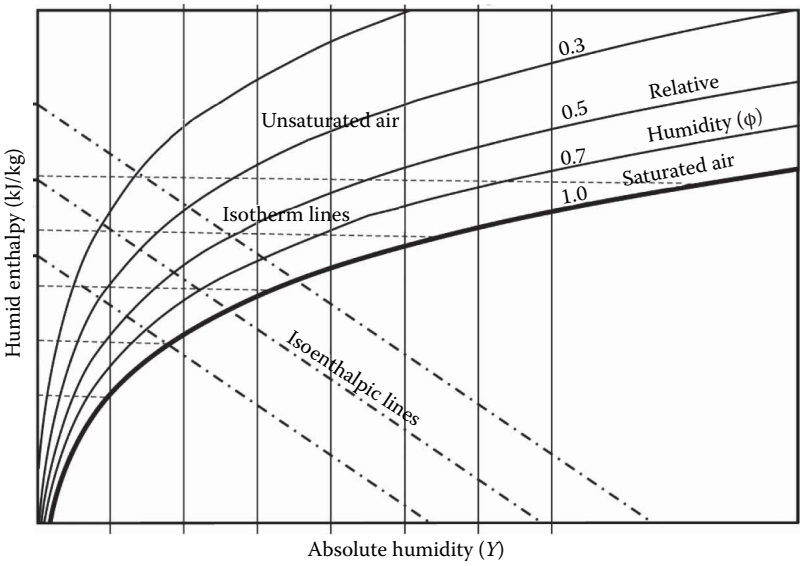


FIGURE 4.29.1 Mollier chart.

change in the drying process and one of the objectives of the drying process is to improve their properties. Broadly understood technological properties of a product depend on the type of moisture bond with the material and nature of these bonds affects the drying process. The removal of liquid from wet material is accompanied by breaking of liquid–solid bonds as a result of energy supply.

There are three groups of bonds of moisture with the material:

1. *Chemical*: Remain stable during drying process, chemically bound water cannot be removed during drying process
2. *Physicochemical*: Absorption and osmotic type bonds. Water bound due to adsorption is adsorbed mainly on inner body surfaces
3. *Physicomechanical*: Structural bonds can be removed by drying

In practice of drying, there are specific definitions of material moisture. Moisture in wet material is commonly expressed as

Absolute moisture, kg water/kg of dry material (X).

Free moisture, the moisture that is in excess in relation to equilibrium moisture. Free moisture can be removed during drying.

Equilibrium moisture (X_e, kg/kg), moisture content of the material that is in equilibrium with vapor of specified pressure and temperature. The equilibrium moisture content is the lowest moisture content to which the material can be dewatered by drying for given air temperature and humidity.

Critical moisture (X_{cr}, kg/kg), moisture content of the material at which the falling drying rate period starts due to increase in diffusion resistance of moisture flow in the material.

Drying Equilibrium

Moisture can diffuse from wet material to drying gas, if equilibrium vapor pressure over the material is higher than partial vapor pressure in the drying agent. Moisture evaporation and desorption proceeds until the system reaches the state of equilibrium (equal vapor pressure over the wet material and partial vapor pressure in the drying air). If partial vapor pressure in the ambient air is bigger than vapor pressure over the material, moisture sorption will occur, i.e., moisture will penetrate to the material by diffusion. The equilibrium moisture content in the material can be determined by a prolonged contact of the material with air at a given temperature and humidity. The relationship obtained at different air humidities at constant temperature is called sorption isotherm. An example of sorption–desorption isotherms is shown in Figure 4.29.2.

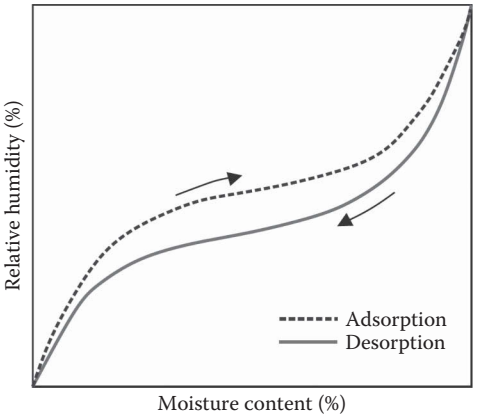


FIGURE 4.29.2 Sorption–desorption isotherms.

Figure 4.29.2 shows drying hysteresis between sorption and desorption of moisture for the same ambient air temperature. The desorption curve always has a bigger value than the equilibrium moisture sorption curve at a given temperature. Knowledge of sorption/desorption isotherms is one of the key data in drying to determine boundary values of moisture content below which the material cannot be further dried and the danger of rehydration due to moisture resorption from the air, in the case of product storage, cannot be avoided.

Drying Curves

Heat required to evaporate moisture is transferred from the drying gas. Heat transfer between the phases depends not only on material properties but also on the properties of ambient air (density, temperature, pressure, humidity, velocity, etc.). Moisture contained in the material may be in the bound or unbound form (Strumillo and Kudra, 1986). Water present in solid body capillaries, absorbed on a solid surface, or chemically bounded with a solid body falls into the category of bound moisture. Free moisture is defined as moisture that constitutes excess in relation to equilibrium moisture. During drying only unbound moisture is evaporated.

The mechanism of drying depends on the type of moisture contained in a particle. As long as unbound moisture is in the particle and the rate of moisture diffusion is sufficient to keep the state of saturation on the material surface, the particle is in the constant drying rate period (Charlesworth and Marshall, 1960).

Figure 4.29.3 shows drying curve with two characteristic drying periods: the constant drying rate period and falling drying rate period.

In the initial phase, the dried droplet has lower temperature than equilibrium temperature; as a result, in zone *A* the drying rate increases until equilibrium is reached. Typically, the initial period is so short that it is not taken into account in the considerations of drying time.

Segment *B* corresponds to equilibrium conditions, the whole amount of heat transferred to the particle from the environment is used to evaporate moisture. The evaporation rate is constant; this period is called the constant drying rate period. The rate of evaporation of moisture in this period depends on the initial concentration of dry matter; the higher concentration of the solids the lower drying rate due to decreasing equilibrium vapor pressure, which reduces the driving force of the process.

If the moisture of the material drops below a critical value (point X_{cr}), the amount of liquid reaching the surface begins to gradually decrease. As a result of an increase in diffusion resistance, the drying rate decreases. This period is called the falling drying rate period (*C*).

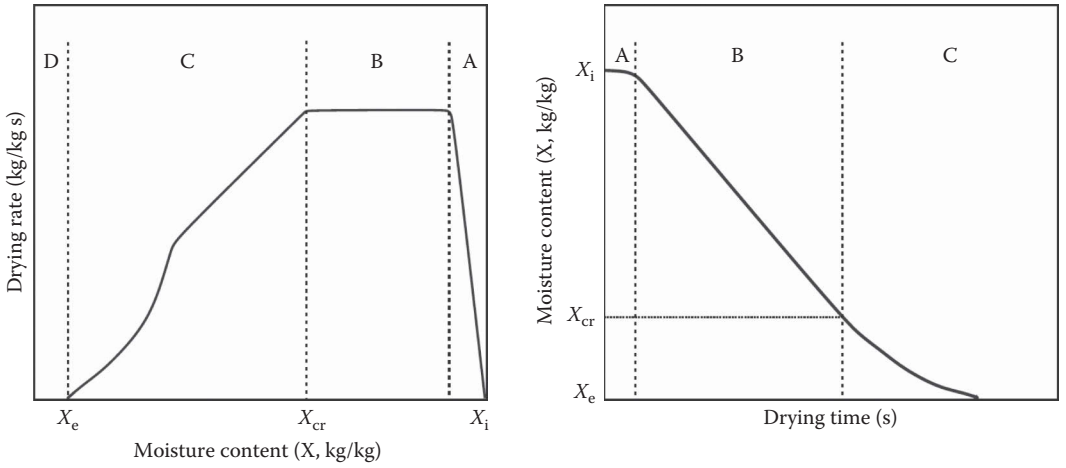


FIGURE 4.29.3 Drying rate vs. moisture content and moisture content vs. time.

The formed surface of a particle might provide an additional resistance to mass transfer and change linear shape of drying rate. The evaporation rate decreases until achieving a moisture content, which is in equilibrium with ambient air (X_e).

The mechanism of moisture removal during the falling rate drying period depends on the type of and the properties material. If the surface of the material is not porous and represents a significant resistance to the flow of moisture, the drying rate decreases and the time of moisture evaporation significantly increases.

Heat and Mass Transfer

During drying, heat of evaporation to remove water or other solvent must be provided. Heat and mass transfer occurs inside dried material and in the boundary layer between solid and drying agent. During drying, the internal structure of dried material and external conditions have various impacts in different periods of the drying process.

External Conditions

Heat may be supplied to the surface of the dried material by radiation, convection, or conduction. In most cases, all three heat transfer mechanisms are active, while one is predominant. The vast majority of drying processes is convective, hot air is used to supply energy for drying and removing water vapor. While wet material is in contact with the hot air flowing around the dried body, moisture from its surface is evaporated. Mass transfer during the evaporation of liquid is perpendicular to the direction of air flow, which affects the boundary layer. This in turn induces a change in heat transfer coefficient.

Heat and mass transfer is usually described with heat and mass transfer coefficients related to the heat and moisture flux normal to the evaporation surface, to the difference between the wall temperature or vapor pressure at the wall and reference temperature or vapor pressure.

$$q_{cw} = \alpha \cdot \left(T_w - T_{ref}\right) \tag{4.29.6}$$

$$w_{cw} = k_y \cdot \left(p_{v,w} - p_{v,ref}\right) \frac{M}{RT} \tag{4.29.7}$$

In practice, heat and mass transfer coefficients are often estimated by means of empirical correlations with the air speed, shape and size of evaporation surface, nature of the gas–vapor mixture, pressure, temperature, physical properties of the liquid, and gas component concentrations.

Internal Conditions

Drying is a diffusional process in which the transfer of moisture from the dried material to the surrounding medium is performed by surface moisture evaporation. Once some of the surface moisture evaporates, more moisture is transferred from the center of a solid to its surface. The transport of moisture within a solid is governed by a variety of mechanisms depending upon the nature and type of the solid and its form. The mechanism of moisture transport in different solids may be broadly classified into the following:

Diffusional transport of liquid or vapor
Capillary suction
Pressure induced

The type of dried materials, the structure of pores, and even rate of drying decide which mechanism is predominant. However, at different stages of drying of the same material different mechanisms may play a major role (Strumillo and Kudra, 1986).

Drying Techniques

Convective Dryers

Chamber Dryers

Scheme of a typical chamber dryer is shown in Figure 4.29.4. In most cases, chamber dryers operate in batch mode. Dried material is distributed on trays stacked inside a well-insulated drying chamber, while air circulates over the material. Usually, the stacked trays are placed on a trolley for convenient loading and unloading of the dryer. Each trolley carries the same number of trays loaded with material to be dried. Part of the moist air is recirculated with appropriately designed fans, guide vanes, or ducts. The remaining moist air continuously vented through exhaust duct is replaced with fresh air. Usually, drying times in the chamber dryers are long (from 10 to 100 h). Chamber dryers are sensitive to air flow pattern distribution over the trays as the residence time required for the slowest drying tray decides the dryer capacity.

Due to a large amount of labor to load and unload the product, the operation of chamber dryers is expensive. Chamber dryers are useful for small production lines but can be used to dry a wide range of materials, among them valuable products. In a modified design, with indirect heating of the trays, vacuum is used to remove evaporated moisture.

Tunnel Dryers

A tunnel dryer may be considered as development of the chamber dryer in which trays stacked on the trolleys move through a tunnel (long drying chamber with entrance and exit doors at opposite ends). The schematic of a tunnel dryer is depicted in Figure 4.29.5. The dried material is entering

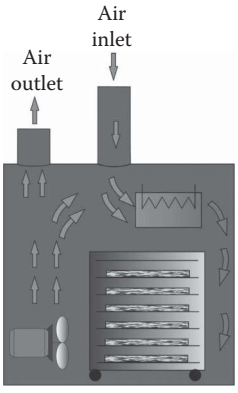


FIGURE 4.29.4 Schematic of chamber dryer.

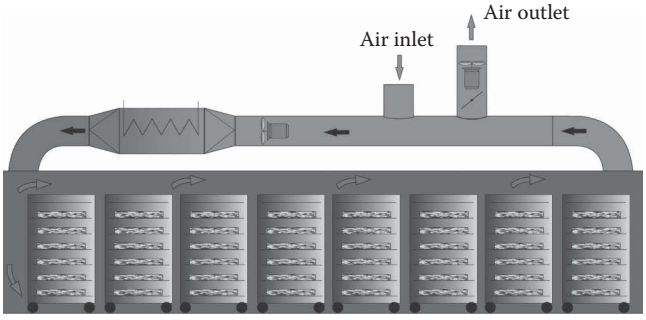


FIGURE 4.29.5 Schematic of tunnel dryer.

at one end and the product is collected at the other end of the tunnel. Trolleys with wet materials are pushed through the tunnel. In most cases, air supplying the heat for evaporation and removal of vapor circulates. Usually, part of the moist air is exhausted and replaced with fresh air. Trolleys with dried material on trays can move through the dryer either con- or countercurrently to the air flow direction. Also air cross-flow can be used but in that case the tunnel should be specially designed to control the flow path.

Tunnel dryers are often used for drying packaging materials, ceramics pottery, glass containers, etc. The drying process can be finished with sterilization of the containers. Tunnel dryers offer continuous drying of a large amount of materials.

Due to a large amount of labor to load and unload the product, the operation of a tunnel dryer is expensive. Drying rate is small, so long residence time is necessary and the process is time consuming.

Belt Dryers

Belt dryers were developed for gentle processing of materials susceptible to physical damage. Figure 4.29.6 depicts schematic of a single-belt dryer. Wet material is continuously spread as a thin layer on a usually horizontal mesh or perforated belt. The drying chamber can be divided into sections with a separate fan and heat exchanger to achieve different conditions of drying. The heat for moisture evaporation can be supplied directly or indirectly, and all heating media, such as steam, hot water, or hot gas can be used. The temperature of drying and air flow rate can be controlled separately in each section. Hot air passes through or over the material and supplies the heat for moisture evaporation and conveying humidity. The speed of a conveyer belt can be controlled, which enables easy control of drying time.

For the large depth of a dried material bed, moisture profile buildup in the layer of the material can be observed and dried material can be overheated. In order to ease the problem, the change in the air flow direction along the dryer is used. The moisture profile becomes then more uniform and the drying rate is higher. Another solution is the mixing of material on the belt. Also the application of multibelt dryers with several belts installed one above the other offers mixing of the material falling from the higher to the lower belt.

Belt dryers are fairly versatile; complex shape, relatively big, heat-sensitive, and fragile particles can be successfully processed. Schematic of a multipass band dryer is shown in Figure 4.29.7.

Superheated Steam Dryers

Drying with superheated steam is about a 100-year-old concept, but for a long time was not implemented in the industry. At the beginning of the twenty first century, the potential of this technique has been recognized again and many new research programs started (Devahastin, 2008).

Superheated steam drying (SSD) is replacing hot air in a convective drying process with superheated steam to provide energy for moisture evaporation and removal. Most constructions of the convective dryers can be used as a base for the superheated steam dryer; however, the change is not straightforward.

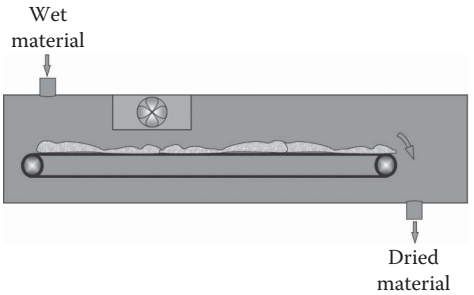


FIGURE 4.29.6 Schematic of single-belt dryer.

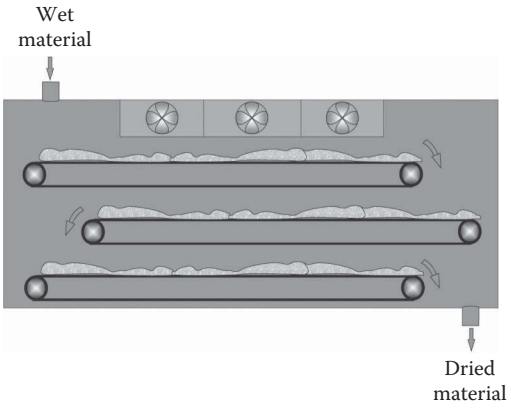


FIGURE 4.29.7 Schematic of multibelt dryer.

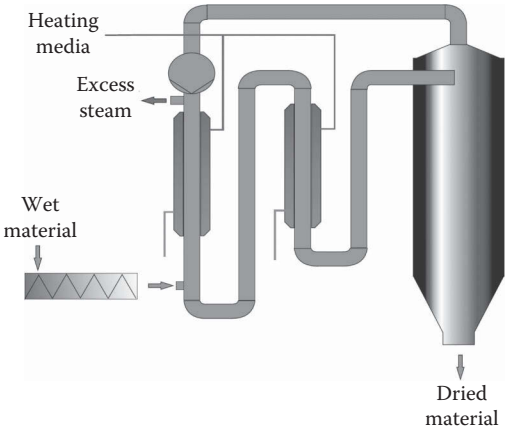


FIGURE 4.29.8 Schematic of superheated steam dryer.

SSD has a number of significant advantages over air drying. The obvious benefit is that the dryer exhaust is steam. Most of the latent heat supplied to the SSD dryer can be reclaimed by condensing the exhaust stream or by steam compression. Recovered steam can be reused in the dryer and excess can be used for other applications in the process plant. An example of a superheated steam dryer is shown in Figure 4.29.8.

Superheated steam drying eliminates the environmental problem of emission of volatile compounds produced during drying because separation of volatile organic compounds takes place at a condenser (Soponronnarit et al., 2006). The absence of oxygen in superheated steam dryers prevents oxidation or combustion reactions and eliminates the risk of fire or explosion. Lack of oxygen also inhibits oxidative reaction, which impairs the product quality (Elustondo et al., 2001). Wood chips dryer in superheated steam are softened and pelletization is improved (Pakowski et al., 2007).

Superheated steam drying can be considered as a good choice if at least one of the following conditions applies:

High-energy cost and low product value.
Product quality is significantly superior if dried in steam rather than air.
Risk of fire or explosion in the presence of oxygen is high.
Capacity and quantity of water to be removed is high.

Due to possible problems with steam condensation on the product during start-up and shutdown, steam dryers operate in continuous mode.

Disperse Dryers

Spray Dryers

Spray drying technique is commonly applied in food, chemical, ceramic, and pharmaceutical industries for mass production of powders. In spray drying, liquid feed is atomized in hot gas, dried, and collected in a form of powder, granules, or agglomerates.

First patent on spray drying (US Patent 125 406) was granted in 1872 to Samuel Percy who elaborated a concept of the process to contact dispersed liquid with hot air (Masters, 1991). Vast development of this technique was accelerated during the World War II due to high demand for dewatered powder (food, chemicals) to reduce costs of transportation. Most of existing types of spray dryers were constructed in the 40s. There are more than 500 materials already dried in spray dryers and more than 1100 patents related to this technique (Masters, 1991).

There are three stages of spray drying process (Hasheminya and Dehghannya, 2013):

1. *Atomization of feed* with pressure, pneumatic, ultrasonic or acoustic nozzles, or rotary atomizers. Figure 4.29.9 shows basic types of atomizers applied in spray drying.

   Atomization process always produces a polydisperse, log-normal particle size distribution (PSD). Typical shape of PSD with arithmetic diameter and Suter diameter characterizing the spray is shown in Figure 4.29.10. PSD occurs independently of the type of atomizer or physicochemical properties of sprayed liquid.

   In practice, for scaling-up or calculation of spray drying process, experimental distributions such as Rosin–Rammler or Nukiyama–Tanasawa are often applied (Masters, 1991). New constructions of spraying nozzles to produce narrow PSD and low velocity spray, like ultrasonic or acoustic nozzles are still under development. Up to now, no ultrasonic or acoustic nozzles have been applied in industry for mass production of powder.

   Droplets leaving the atomizer move at different speed that may lead to collisions and agglomeration. The most important parameters characterizing the spray and affecting final powder properties are atomization angle, atomization ratio (AR), and PSD. Atomization angle determines the shape of spray. Pressure nozzles are characterized by small atomization angle in comparison with pneumatic nozzles and rotary discs. One of the most important parameters determining the quality of atomization is AR, which is equal to the mass flow of atomized air to mass flow of liquid (kg/kg). The higher the AR, the smaller the droplet diameter. In practice, AR covers the range from 0.1 to 10, for AR lower than

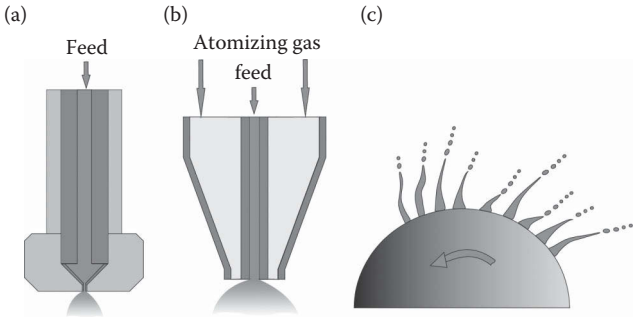


FIGURE 4.29.9 Schemes of basic types of atomizers applied in spray drying (a) pressure nozzle, (b) pneumatic nozzle, and (c) rotary disc.

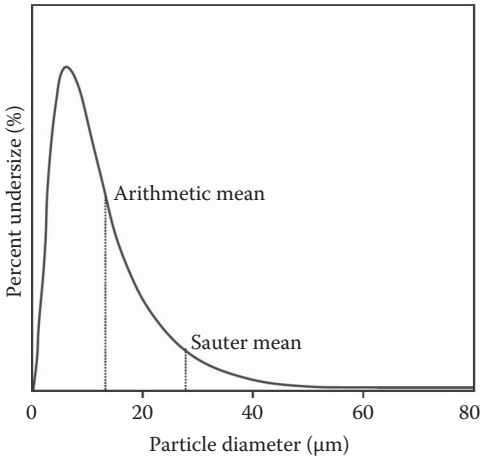


FIGURE 4.29.10 Log-normal PSD.

0.1 atomization is coarse and unstable, for AR higher than 10 no more decrease of droplet diameter can be achieved.

Atomization is accompanied by losses of volatile substances as well as flavor and aroma of dried food products (King, 1995), which occur until the formation of an impermeable layer on the particle surface. Atomization process has a significant impact on the particle size, morphology, and the organoleptic properties of the final product.

2. *Drying of droplets with hot air*

The way in which drying air is contacted with the sprayed material has a significanlt influence on the residence time of particles in the dryer and properties of the final product. Depending on sprayed material-drying agent contact, three types of spray dryers are distinguished:

Concurrent,
Countercurrent
Mixed flow

Figure 4.29.11 shows schematic diagrams of spray dryers with regard to the way of contact with drying air (Masters, 1991).

In countercurrent dryers, drying air is supplied to the dryer axially and tangentially, and the air and material being dried flow in the same direction. In countercurrent systems, raw material can be supplied into the drying chamber from the bottom or top and flows in the direction opposite to the drying air. In mixed flow dryers, the raw material is usually supplied from the top and the drying agent is fed into the drying chamber from the top and bottom.

In concurrent dryers, a so-called atomization zone is distinguished in which a stream of material penetrates the turbulent air flow and free fall zone, where the atomized material is gradually mixed with air. In the next step, the dried material is discharged from the drying chamber to cyclones where it is separated from the drying agent.

The processes of concurrent and countercurrent spray drying were developed in parallel. However, a simple construction and easy process control cause that the concurrent system is most often applied in industry. On the other hand, a better use of energy in the countercurrent system, comprising drying and agglomeration and lower production costs, also has led to a broader implementation of this system in the industry, despite the greater difficulty in controlling the process (Rahse and Dicoi, 2001). The countercurrent spray

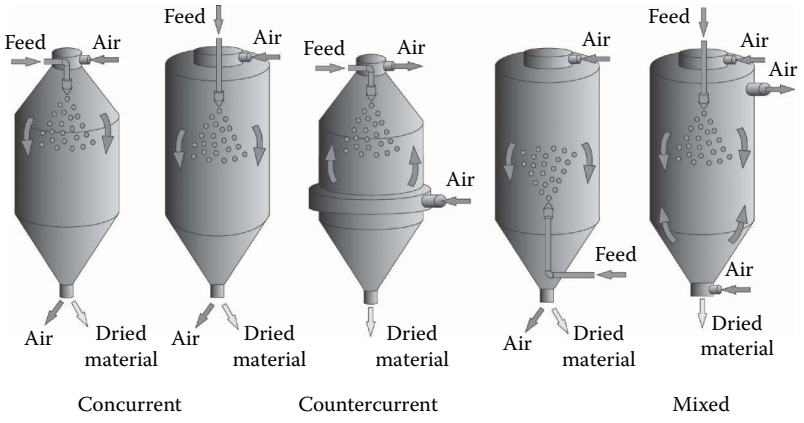


FIGURE 4.29.11 Types of spray dryers (concurrent, countercurrent, mixed).

drying method is widely used in the production of detergents, fertilizers, and certain food substances.

The countercurrent process has the following advantages over the concurrent technique:

High-energy efficiency
Intensive agglomeration, large average particle size, less dust
Control of bulk density of the powder
Drying of viscous products
Integration of upper filters
Multilevel spraying, high throughput
Lower costs of the drying process

Table 4.29.2 shows a comparison of basic powder properties dried in concurrent and countercurrent spray dryers.

Intensive agglomeration in countercurrent spray drying resulted in almost doubled average particle diameter in comparison with the concurrent process.

3. *Separation of dried powder from drying gas*

In most cases high-volume industrial dryers work in a quasi-closed system of raw material, which means that the dried product is discharged from the dryer bottom part, while particles with small diameters are separated from the stream of drying air in bag filters placed in external cyclones.

In a countercurrent system, particles entrained with the drying agent are returned to the dryer and introduced directly into the drying/atomization zone by shaking off powder

TABLE 4.29.2
Morphology of Powders Dried in Concurrent and Countercurrent Spray Dryer

Property	Concurrent Drying Process	Countercurrent Drying Process
Moisture content (%)	13	11
Bulk density (kg/m^3)	235	280
Particle diameter (μm)	170	310

Source: Taylor, T. Concurrent and countercurrent spray drying and particle morphology. Paper No. 9, *12th International Drying Symposium*, Noordwijkerhout, The Netherlands, August 28–31, 2000.

from the bag filter or by pneumatic transport of material from the cyclones. This solution can reduce the loss of material and at the same time speed up the process of agglomeration of particles in the dryer. The final purification of air usually takes place in the bag filter, where the finest particles are separated and in scrubbers where volatile impurities are removed (e.g., aromatic substances).

The spray dried materials are divided into the following three groups depending on the particle structure (Walton, 2000):

Skin forming
Agglomerate
Crystalline

Skin-forming materials are characterized by formation of an impermeable film, which covers the entire material surface, while liquid is trapped inside the particles. The particles of spray-dried materials often contain air bubbles (Verhey, 1973), which is a result of desorption of air initially present in the liquid feed or absorbed during atomization.

In view of the diversity of materials to be dried in spray dryers, it is difficult to determine general relationships presenting the kinetics of moisture removal (Masters, 1991). In a typical drying scheme, the particle stops shrinking and a solid structure is formed upon reaching the critical moisture content. From that moment particles cannot change their size, can break, agglomerate, etc.

Process conditions of spray drying have a significant effect on product morphology. Figure 4.29.12 shows the changes of bulk density of skin forming, agglomerate, and crystalline product as a function of drying air temperature (Kwapinska, 2002). Air temperature growth from 150°C to 250°C is decreased by about 75% the bulk density of the skin forming material and by approximately 30% the bulk density of the agglomerate (detergent). In the skin-forming materials, moisture is entrapped inside due to crust formation. If the particle surface is not porous, then the forming water vapor causes stretching or breaking of the crust. Temperature of the drying agent has no effect on bulk density of crystalline material.

Air bubbles in skin-forming particles as well as crashes and craters in agglomerate powder can be detected. No effect of spray drying on the morphology of crystalline material and no impact of bulk density can be found.

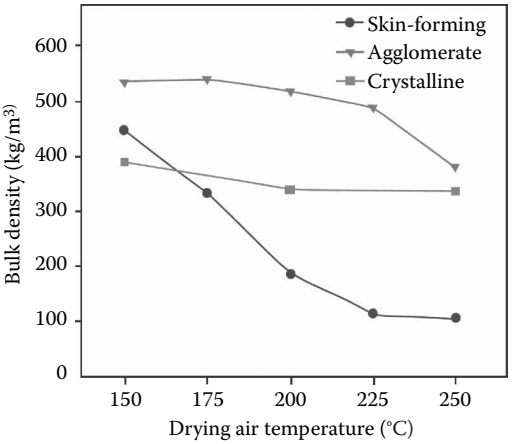


FIGURE 4.29.12 Effect of drying gas temperature on bulk density of the product (Kwapinska, 2002).

Microencapsulation

One of the most spectacular applications of spray drying is microencapsulation of products. Microencapsulation is a method of trapping an active core substance in a matrix material. This technique is used to protect specific properties of the encapsulated substance against unfavorable environmental conditions and to control release of the core material. Despite the diversity of microencapsulation techniques, spray drying is the most common method in the food industry mainly because of low cost. When removing the solvent (usually water), the process of core trapping in the matrix material proceeds at the same time. The spray drying method can be used for solutions of high viscosity, such as emulsions, pastes, and slurries. This method allows for drying of thermolabile materials without significant losses of product quality.

Microencapsulation enables extension of product shelflife, production of new flavor and aroma compositions, and improvement of nutritional values of food products. Main groups of products protected by the microencapsulation technique include spices, fats and essential oils, vitamins and minerals, bioactivators (probiotic bacteria and enzymes), and live microorganisms (acidifying bacteria and yeast) (Gibbs et al., 1999). Like in food industry, also in cosmetic industry the microencapsulation made it possible to develop production techniques in order to obtain a product with new functional properties.

Beginnings of the microencapsulation technique date back to the 30s of the 20th century (Thies, 1996). In the second half of the 20th century, the process of microencapsulation was used mainly in the food processing and pharmaceutical industries, whereas in the 1990s it was also used in cosmetic, biotechnological, and agrochemical industries (Gouin, 2004).

Analysis of the literature shows that the efficiency of microencapsulation depends on the initial temperature of the drying air, residence time, and content of the encapsulated material relative to the amount of carrier.

Table 4.29.3 gives examples of surface oil content and cardamom essential oil retention in powders at different oil concentration in the emulsion (Beristain et al., 2001).

A small ratio of matrix to core material causes a deterioration of encapsulation and increased concentration of essential oil not entrapped in the microcapsule. The highest oil retention of 83.6% was reported for the 1:4 ratio, which proves the existence of an optimum concentration of the coating material for reaching the best possible retention of the core material (Risch, 1995). On the other hand, a decrease in the retention for the oil/matrix material ratio of 1:5 was due to the shape and location of the core material in matrix material.

Fluidized Bed Dryers

A fluidized bed drying is one of the most frequently applied techniques for drying of particulate materials. High rates of heat and mass transfer cause that fluidized bed dryers are characterized by high thermal efficiency and have smaller size than other, alternative types of dryers. In fluid bed dryers thermally sensitive materials can be dried because of easy control of material temperature in

TABLE 4.29.3
Total and Surface Oil Content and Retention in Powders at Different Oil Concentrations in the Emulsion

Oil to Rubber Ratio	Oil Content in Emulsion (g oil/kg rubber)	Total Oil Content (g oil/kg rubber)	Surface Oil Content (g oil/kg rubber)	Retention (%)
1:5	166.6	111.4	1.90	66.9
1:4	200.0	167.1	2.59	83.6
1:3	250.0	185.6	8.08	74.2

Source: Beristain, C.I. et al., *Lebensmittel-Wissenschaft und Technologie*, 34, 398, 2001.

different sections of the dryer. Low costs of maintenance, few moving parts, and simple devices for receiving the product are the advantages of fluidized bed drying. Its disadvantages include a limited range of materials that can be dried, e.g., this technique is not suitable for materials with high initial moisture content, with a wide range of particle sizes (polydisperse materials), and with low critical fluidization rate (Kunii and Levenspiel, 2013). Besides dryers where the classical fluidization is applied, there is a number of modifications of this technique such as vibrofluidized bed, spouted bed, or spouted bed with inert bodies.

To obtain a fluidized bed, a stream of gas should be introduced from the bottom of the dryer through a screen baffle on which the bed of particulate material is located (Figure 4.29.13). At a certain gas flow rate, the layer of material is loosened and particles move within the bed. This gas–solid particles system has many features of a fluid and so this state is called the fluidized state. Velocity at which a packed bed transfers into the fluid bed is called the critical fluidization velocity.

Further increase of the gas flow rate causes expansion of the bed. Depending on gas flow, different states of fluidization can be observed: smooth fluidization, bubbling fluidization, and turbulent fluidization. Finally, fluidization turns into pneumatic conveying and powder is carried out from the dryer. A schematic diagram of different stages of fluidization is shown in Figure 4.29.14.

In practice, the gas flow rate is maintained at a level that provides a uniform fluidization of the particles. Above the bed there is a free space (1–2 m), which allows lighter particles entrained from the bed to fall down. Fluidization of the bed of wet material develops favorable conditions for carrying out the process of drying due to good development of the interfacial contact area, which contributes to achieving high heat and mass transfer coefficients.

There are various dryers in which the fluidized bed is used. From the point of view of the type of material, fluidized bed dryers are applied to dehydrate loose materials as well as pastes, solutions,

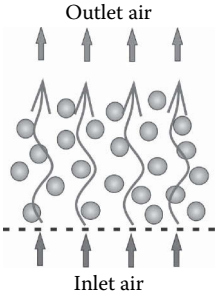


FIGURE 4.29.13 Basic mechanism of fluidized bed drying.

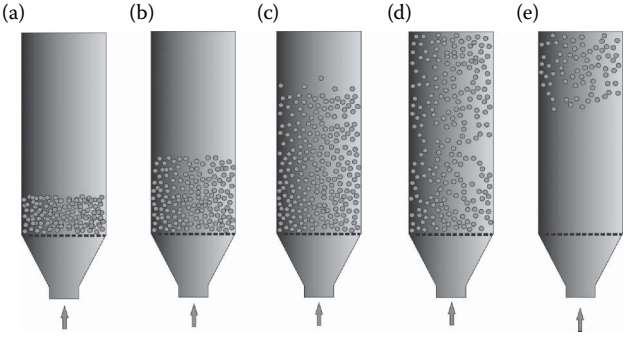


FIGURE 4.29.14 Flow regimes in fluidization, (a) packed bed, (b) smooth fluidization, (c) bubbling fluidization, (d) turbulent fluidization, and (e) conveying.

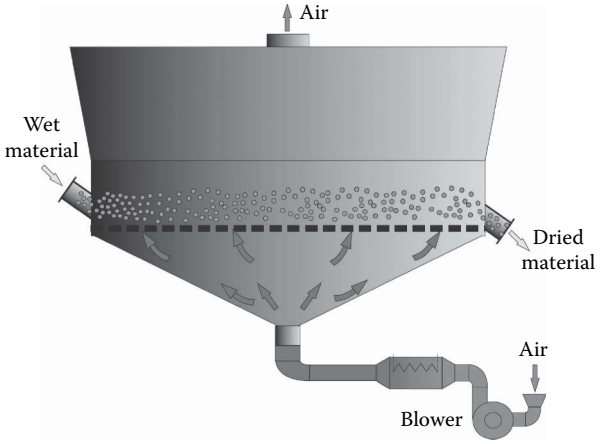


FIGURE 4.29.15 Single-stage fluidized bed dryer with a circular cross section.

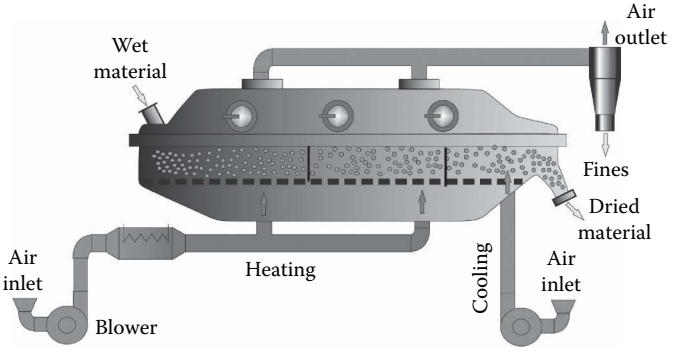


FIGURE 4.29.16 Single-stage fluidized bed dryer with rectangular cross section.

and suspensions. From the viewpoint of a drying regime, there are continuous, semicontinuous, as well as single- and multistage fluidized bed dryers. In industrial practice, the most popular are continuous dryers. Figures 4.29.15 and 4.29.16 show two typical schematic diagrams of continuous fluidized bed dryers—a classical single-stage dryer of circular cross and rectangular sections. Dryers of this type operate at a cross flow of the material and drying agent. Dried material is received directly from the dryer and dusted gas is cleaned initially in cyclones and then in bag filters. To reduce the entrainment velocity of fine particles, the diameter of a drying chamber can be expanded in the upper part (Figure 4.29.15).

Figure 4.29.16 shows a typical industrial continuous fluidized bed dryer with a rectangular cross section, divided into heating and cooling sections to maintain a predried product at a lower temperature to avoid overheating.

Baffles in the bed are used to control of residence time of particles in the dryer. Development of the fluidized bed drying technique is vibrofluidized beds. The vibrofluidized system enables uniform drying of materials with high adhesion and prone to sticking together (e.g., sugar) at high process intensity.

Pneumatic Dryers

Pneumatic drying takes place during pneumatic transport of fine particles of wet material in a hot air stream. During pneumatic drying, there is an intensive heat and mass transfer between continuous and discrete phases resulting from a significant expansion of the interfacial surface area, high

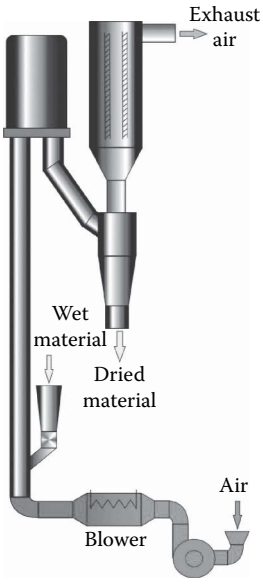


FIGURE 4.29.17 Schematic of pneumatic dryer.

heat, and mass transfer driving force (gas temperature above 500°C) and large gas flow velocities (25 m/s). A typical diagram of a pneumatic dryer is shown in Figure 4.29.17. Wet material is fed to a vertical tube in which hot air flows. Air flow velocity must be higher than settling velocity of dried material particles. The particles are entrained upwards quickly giving moisture back. Dry product is received in the cyclone and dust formed during the pneumatic transport due to abrasion of particles is collected in a bag filter or electrofilter (Borde and Levy, 2007). The residence time of material in the tube is from 1 to 5 s, which restricts the use of pneumatic dryers to materials containing mostly surface moisture. In industrial applications, the material is recirculated to prolong residence time in the dryer even to 1 min. The recommended diameter of material particles ranges from 1 to 2 mm.

Pneumatic dryers are used in many branches of chemical, ceramic, agricultural, and wood processing industries. Due to a simple construction, pneumatic dryers are characterized by low capital and operating costs as compared to other types of dispersion dryers. Owing to short residence time of material in the dryer, high inlet gas temperatures can be used and thermally sensitive materials can be dried. High temperature of the drying agent also contributes to high evaporation rates and thermal efficiency of the process. Pneumatic dryers require little space and can be installed in the open air.

A disadvantage of the pneumatic dryers is often the necessity of pretreatment of the raw material to obtain sufficiently small particles suitable for pneumatic transportation. Another problem is the abrasion of particles during pneumatic transport and high dust emission requiring efficient dust collection equipment.

Pneumatic dryers are a modern drying technique, which is continuously developed and permanently present in the industry.

Rotary Dryers

Rotary dryers are among these the most commonly used in industrial practice. Rotary dryers are characterized by simple construction, high throughput (up to hundreds tons per hour), and satisfactory thermal efficiency. Basic mechanism of the rotary drying process is as follows: material being dried is lifted due to drum rotation, then falls in the hot air stream moving in co- or countercurrent to the direction of material movement. Dried material moves along the dryer due to rotations of a slightly inclined drum.

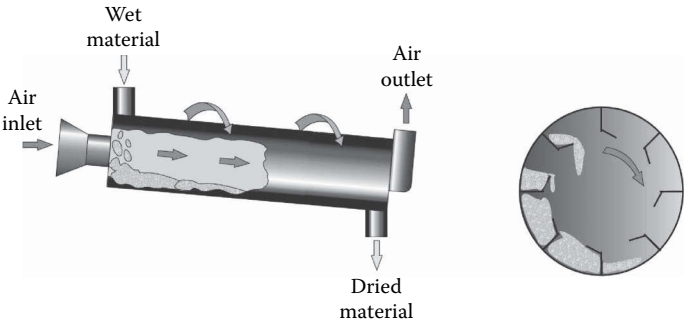


FIGURE 4.29.18 Schematic of rotary dryer.

A scheme of the rotary dryer with cross section of the kiln is shown in Figure 4.29.18.

The residence time of material in the dryer, which ranges from several minutes to about an hour in industrial conditions, depends on the speed of drum rotations (normally: several rpm), kiln inclination (up to several degrees), diameter of the drum (1–3 m), design and number of fins, load regime (normally loaded regime = 10%–12% of the dryer volume, but dryers work also in overloaded and underloaded regime), air velocity and flow direction, and also properties of the material (including moisture content). The drying time of the material must be smaller or equal to the residence time of material in the rotary kiln. Rotary dryers are applied in chemical industry (e.g., fertilizers, ores, limestone), agriculture (wood chips), for drying industrial and municipal wastes and sludges, etc. Rotary dryers, being one of the oldest dryer designs, are continually developed and find new applications due to high thermal efficiency and elasticity of the process.

Contact Dryers

In contact dryers heat for drying is supplied from the hot surface to the dried material by conduction. The contact surface may be heated up with steam, hot water, organic liquids, combustion gases, or electricity. Thermal efficiency of contact dryers is higher comparing to the convective drying because all heat from the hot surface is transferred to the wet material and there are no heat losses with exhaust air. For comparison, energy consumption of a common contact dryer is about 2800–3600 kJ/kg of water evaporated, whereas energy consumption of a convective dryer is 4000–6000 kJ/kg.

The main disadvantage of the contact drying is the higher capital costs in relation to convective drying systems; to increase throughput, the heating areas must be increased, which makes contact drying more suitable for batch operation.

Depending on the shape of a heating surface, contact dryers can be divided into drum dryers, thin-film dryers, cone dryers, chamber dryers, and agitated bed dryers. Contact dryers may operate under vacuum, e.g., in drying of heat-sensitive materials (Strumillo and Kudra, 1986).

Batch chamber contact dryers consist of a drying chamber and heating shelves where dried material is placed; the number of drying shelves depends on the dryer throughput (Figure 4.29.19a). The limitations of this type of dryers are high labor demand to load and discharge the dryer; moreover, for scaling up usually the installation of additional drying units is required.

Contact rotary dryers consist of a rotating cylindrical drying chamber with tubes installed inside, which are usually heated by steam (Figure 4.29.19b). This type of dryers might be applied when the contact of dried material with hot drying gas should be avoided due to product quality issues or when dried material is dusty and difficult to separate from the dried gas.

In the contact drum dryers, the viscous dried medium (paste or slurry) is spread on the rotating cylinder, which is heated up from the inside (Figure 4.29.19c). Depending on the construction, drum dryers are divided into single- and double-drum dryers. Different feeding methods can be applied in drum dryers: spray feed, top or bottom roller feed, “dip” feed. The choice of a suitable feeding

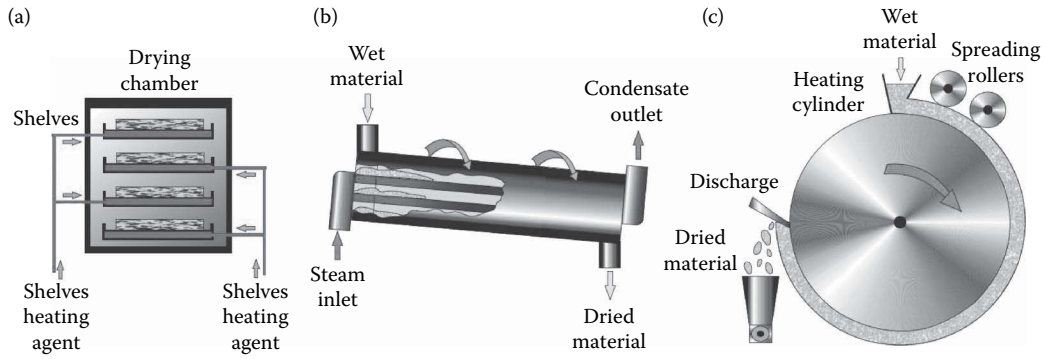


FIGURE 4.29.19 Contact dryers: (a) chamber dryer, (b) rotary dryer, and (c) drum dryer.

method depends on wet material properties such as viscosity or solids content. High heat flux and rapid moisture evaporation from the thin layer of wet material develop a porous structure of dried product making drum drying suitable for instant products. The capacity of a drum dryer is controlled by temperature of the heating surface, drum rotation speed, feeding method, and material properties (Daud, 2006).

Drying in Energy Fields

Dielectric Drying: Microwave and Radio Frequency Dryers

Dielectric drying is the drying technique were heat is generated inside the drying material due to electromagnetic energy. Depending on the frequency of electromagnetic energy, two types of dielectric drying is distinguished: microwave (MW, frequency from 300 MHz to 300 GHz) and radio frequency (RF, frequency from 3 kHz to 300 MHz).

Figure 4.29.20 shows the mechanism of dielectric drying, which can be described in the following way: polar liquids, e.g., water, absorb electromagnetic energy, which causes a change in the orientation of dipoles and heat generation due to molecular friction. Solid parts of the dried materials are not affected by the electric field; therefore, heat is generated at wet regions of the material (Zhang et al., 2016).

Dielectric properties of the wet material, strength of the electric field and frequency of the dielectric field are the main parameters which affect the rate of absorption of electric power per unit volume (Strumillo and Kudra, 1986). Energy efficiency and heating rate of MW drying depends on the dielectric property of material, which is correlated with material moisture content.

A microwave dryer usually comprises three elements (Figure 4.29.21), a magnetron where MW power is generated, waveguide transmitting the MW energy from the magnetron to dried material, and a microwave cavity where wet material is dried.

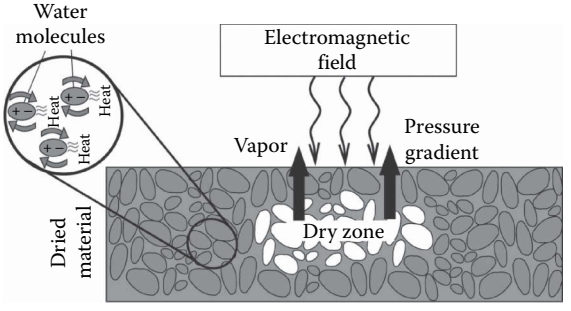


FIGURE 4.29.20 Mechanism of the microwave drying process.

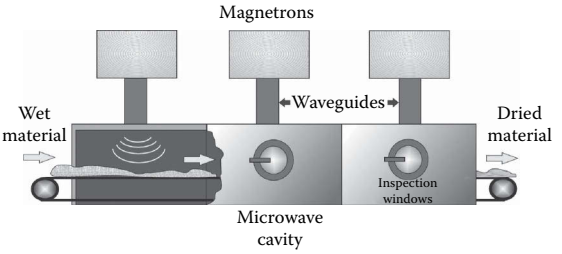


FIGURE 4.29.21 Schematic of microwave dryer.

During last decades microwave drying has found application in drying of different types of materials such as food products, pharmaceuticals, wood, ceramic, biological materials, etc. MW is suitable for materials that have good dielectric properties, i.e., ability to convert electromagnetic energy into heat.

The main advantages of MW drying are high drying rate and short drying time (Rattanadecho and Makul, 2016). The disadvantages of MW drying are high costs per unit of energy and high capital costs and nonuniform heating of dried material, which can deteriorate product quality.

Infrared Dryers

In infrared (IR), drying thermal energy is transferred from the heating elements to wet material by IR radiation (wavelength from 0.76 to 400 μm). Depending on the wavelength band, the IR radiation can be divided into three categories: near-infrared (NIR; 0.78–1.4 μm), middle-infrared (MIR; 1.4–3 μm), and far-infrared (FIR; 3–1000 μm) (Riadh et al., 2015). NIR is more efficient for thicker materials and FIR for thin materials and food products because NIR drying gives higher heat fluxes than MIR and FIR.

Figure 4.29.22 illustrates the mechanism of IR drying. The IR radiation may be reflected, absorbed, or transmitted through the product. The amount of IR energy absorbed by the product depends on the product absorption characteristics and moisture content. The mechanism of heat and moisture transport within the material is identical during the convective drying.

The scheme of an IR dryer is presented in Figure 4.29.23. IR radiators are the main elements of the dryer. In industrial practice, two types of IR radiators are commonly used: electric and gas-fired heaters. The gas-fired heaters produce temperatures in the range from 343°C to 1100°C, whereas electric heaters from 343°C to 2200°C (Krishnamurthy et al., 2008). Capital costs of electric heaters are lower but operating costs are higher comparing to the gas-fired heaters.

The IR drying technique has found wide application in industry for drying materials in the form of thin sheets, e.g., textiles, paper, building materials, paints, and coatings.

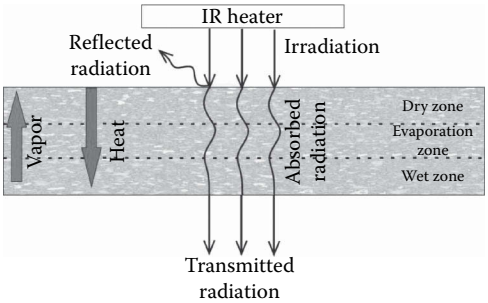


FIGURE 4.29.22 Mechanism of infrared drying process.

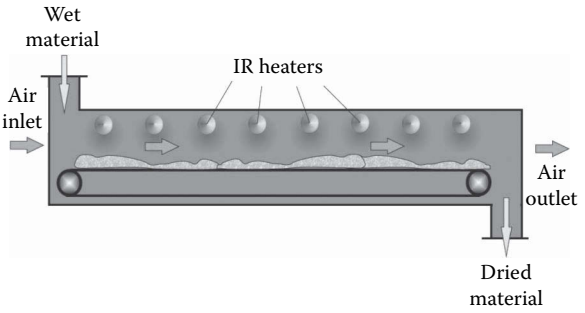


FIGURE 4.29.23 Schematic of infrared belt dryer. (From Strumillo, C. and Kudra, T., *Drying Principles, Applications and Design,* Gordon and Breach Science Publishers, New York, 1986.)

The main advantages of IR dryers is higher energy efficiency than in convective drying, reduced drying time, uniform and direct heating of the product, and low-cost technology. The disadvantages include potential fire hazards, limitation to thin layers drying and necessity to define optimal parameters, i.e., the wavelength of an IR emitter individually for each dried material. In recent years, different hybrid IR drying techniques were developed which use combination of IR with hot air, microwave, freeze-drying, vibration, etc.

Ultrasound Dryers

In ultrasound (US) drying, the energy of acoustic waves at frequencies above 20 kHz is utilized to generate air vibrations and intensify moisture evaporation. In order to enhance the dehydration process high US power intensity (>1W/cm^2) and frequencies ranging from 20 to 40 kHz are required (Witrowa-Rajchert et al., 2014). Two different types of US applications for drying intensification are described in the literature: US-assisted pretreatment of dried material and direct application of US in the drying process (Rodriguez et al., 2014, 2015). The scheme of an US dryer is presented in Figure 4.29.24. Commonly, US dryers consist of an air supplying unit, drying chamber with trays for dried material and main element—an US transducer.

Kowalski and Mierzwa (2015) described the mechanism of US drying distinguishing vibrating and heating effects of US. The propagation of acoustic wave through the gas medium generates high-frequency vibrations of air in the vicinity of dried product surface, which enhances turbulence in the evaporation zone. Additionally, the absorbed acoustic energy generates heating effect, i.e., temperature increase of dried material.

Main parameters affecting the US drying process are acoustic wave frequency, sound intensity, pressure, as well as air velocity and wet material properties (moisture content and porosity).

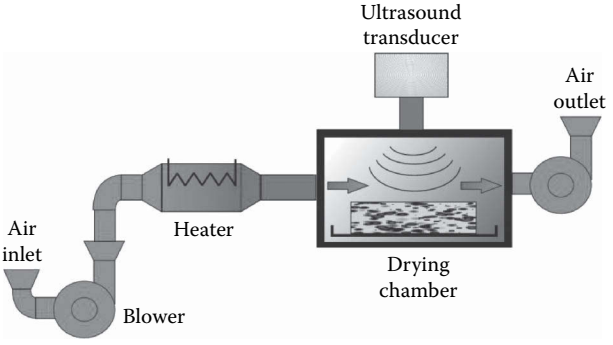


FIGURE 4.29.24 Schematic of ultrasonic dryer.

The US drying was successfully applied in the laboratory scale for drying of food materials such as fruits and vegetables (Rodriguez et al., 2014).

US drying enables one to increase the effective moisture diffusivity and mass transfer coefficient, which results in a shorter drying time. Comparison of the standard convective drying with US-assisted drying showed energy saving of up to 10% for US applications (Kowalski and Mierzwa, 2015). The demerits of US drying are immaturity of the technique, high energy cost, and destruction of cell tissues for some biological materials due to acoustic power. In the literature, there are no reports concerning the application of US drying in the industrial scale.

Vacuum Freeze-Drying

In vacuum freeze-drying, solvent is removed from the frozen material as vapor by sublimation at the pressure below triple point. The freeze-drying process includes the following steps: (1) material freezing, (2) sublimation or primary drying step where frozen "free" water is removed, and (3) secondary drying stage where the product is warmed up to remove bounded water. In the vacuum freeze-drying method, heat for drying is supplied by conduction from the heating plates or IR. In the primary drying step, the product temperature is kept below 0°C and in the secondary drying process the product temperature can be increased above 0°C, but it should be low enough to meet the product quality requirements. Material freezing is the most important step of the freeze-drying because it determines the size of ice crystals formed in the product matrix, which affects dried product morphology and drying rate (Pisano et al., 2011). Slow freezing results in the formation of large ice crystals and large pores in the material matrix.

The mechanism of freeze-drying is as follows: heat is supplied by conduction from the heated shell to the dried material, the solvent vapor is sublimating from the outer surface of the product. A thin layer of the porous dried material is formed at the surface at the beginning of the process and then rises as the sublimation front moves toward the frozen layer (Figure 4.29.25).

Freeze dryers capable of working in laboratory, pilot, or industrial scale are available on the market. Regardless of the dryer capacity all freeze dryers include the same basic elements: a drying chamber, ice condenser, and vacuum pump (Figure 4.29.26). Laboratory- and pilot-scale freeze dryers can be divided into two groups: dryers with shelves where product is placed on the heated plates located in the same chamber as the ice condenser, and dryers with flasks where product is dried in the vessels outside of the ice condenser chamber. Industrial dryers are classified using a similar principle: in single-chamber dryers the ice condenser is located at the drying chamber and in double-chamber dryers the ice condenser and drying chamber are separated. Double-chamber dryers are used when dried products have special requirements, for example, sterile pharmaceuticals.

Vacuum freeze-drying produces the best quality products in comparison to all the drying techniques. Vacuum freeze-drying, however, has high capital and operating costs. Therefore, this drying method can be suitable only for drying products of high value whose properties can be deteriorated during high-temperature drying, i.e., food products with high nutritional value, biotechnological products, pharmaceuticals, or valuable nanomaterials. The major operating cost is energy for heating, vacuuming, and condensing of the removed water.

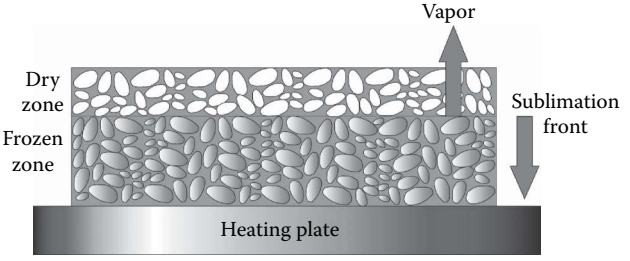


FIGURE 4.29.25 Mechanism of freeze-drying process.

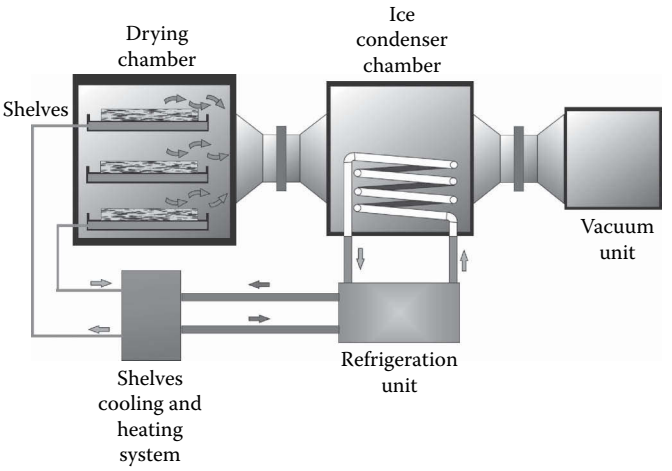


FIGURE 4.29.26 Vacuum freeze dryer with separate drying and ice condenser chambers (Liapis and Bruttini, 2007).

Atmospheric Freeze-Drying

In order to overcome the disadvantages of vacuum freeze-drying, i.e., high capital and operating costs, the atmospheric freeze-drying (AFD) technique was developed. In the AFD, the wet material is dried by convection at temperature below freezing point. As the ability of air to remove moisture is reduced when temperature decreases, the AFD is commonly carried out at temperatures ranging from –3°C to –10°C. One of the limitations of AFD is the long drying time even when comparing to vacuum freeze dryers.

Fluidized bed, tunnel and spray freeze dryers are the most common AFD systems (Reyes et al., 2010). A few authors studied drying of food products in a laboratory-scale fluidized bed atmospheric freeze dryer and find out that a moving bed intensifies heat and mass transfer and reduces drying time. The AFD technique found an application in industrial scale for drying of vegetables at AromaDry plant in Hungary (Claussen et al., 2007). Spray freeze-drying at atmospheric pressure is an emerging technique, which might reduce energy and process costs in production of pharmaceuticals comparing to conventional vacuum freeze-drying.

Special Drying Techniques

Pulse Combustion Drying

Pulse combustion drying (PCD) is the process where heat for drying and energy needed for liquid stream atomization come from intermittent combustion of gaseous fuel. The acoustic oscillations generated in the pulse combustor cause pressure and air velocity pulsations, which significantly increase the heat and mass transfer rate. Pulse combustors operate with frequency from 50 to 200 Hz, initial gas temperature is around 600°C, drying time is less than 0.1 s that enable to dry heat-sensitive products (Zbicinski et al., 2014a). The pulse combustion spray dryer has lower energy consumption (3300 kJ/kg H_2O) than a conventional spray dryer (5501 kJ/kg H_2O). The pulse combustion technique can be applied for spray dryers, fluid bed dryers, flash dryers, rotary dryers, and pneumatic dryers. Every pulse combustion dryer consists of air and fuel supplying units, pulse combustor, feeding system, drying chamber, and dried material collecting units. An example of the PCD system is presented in Figure 4.29.27. Pulse combustors available on the market may be classified as valved or valveless (mechanical or aerodynamic).

PC spray dryers are applied commercially to drying of food products, chemicals, wastes, and nano- and ceramic materials.

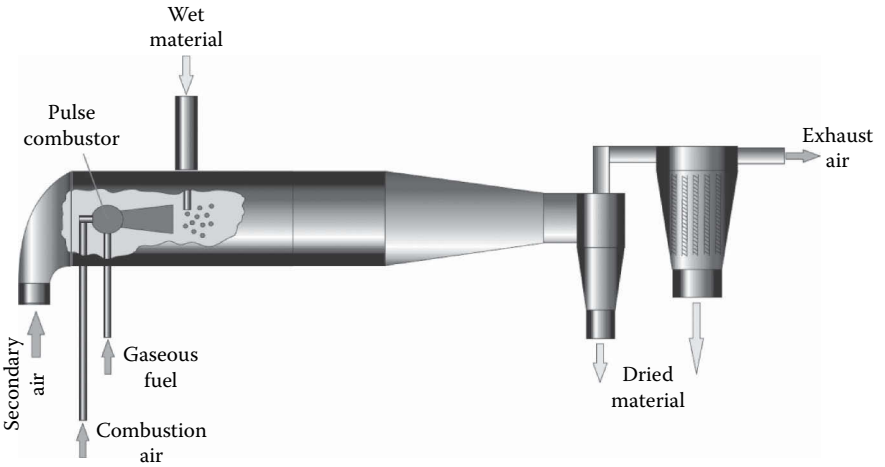


FIGURE 4.29.27 Pulse combustion spray drying system (Meng et al., 2016).

Foam Spray Drying

Foam spray drying is a modified type of the spray drying process where foaming gas is introduced into the raw material to control final product properties. Two types of the foam spray drying process are distinguished: gas desorption and gas admixing processes. In gas admixing foam spray drying, the foaming gas, for example, nitrogen, is introduced in the pipeline connecting the pump and atomizing nozzle as shown in Figure 4.29.28 (Zbicinski and Rabaeva, 2009). Feed foaming before atomization produces droplets with gas bubbles trapped inside which during contact with hot air expand producing hollow dried particles with voids.

Foam spray drying produces powders with specific properties: low bulk density, high porosity and particle size, enhanced solubility, and wettability (Zbicinski et al., 2014 b). A comparison of the bulk density of powders dried using conventional spray drying and foam spray drying methods is presented in Table 4.29.4 (Rabaeva, 2012). An increase of foaming air flow rate results in a significant reduction of the bulk density of dried products.

Flame Spray Drying

The flame spray drying (FSD) is a modification of the spray drying process where heat for drying is provided by combustion of a flammable component added to the feed. Basic principle of flame spray drying is presented in Figure 4.29.29.

Distinct features of the FSD set up is the absence of drying air heating system and presence of additional elements such as ignition system, flame detector, and safety membranes.

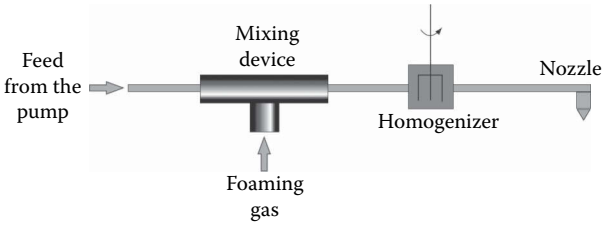


FIGURE 4.29.28 Schematic of foaming gas introduced in foam spray drying process (Zbicinski and Rabaeva, 2009).

TABLE 4.29.4
Comparison of Powder Bulk Density for Foamed and Nonfoamed Feed

Feed Rate (kg/h)	Foaming Air Flow Rate (kg/h)	Maltodextrin Bulk Density (kg/m³)	Foaming Air Flow Rate (kg/h)	Detergent Bulk Density (kg/m³)
9	0[a]	267	0[a]	346
9	40	181	42	236
12	0[a]	321	0[a]	312
12	21	205	22	185
12	75	179	74	171

Source: Rabaeva, Y., Kinetics of foam-spray drying process, PhD thesis, Lodz University of Technology, Poland, 2012.
[a] Standard spray drying process.

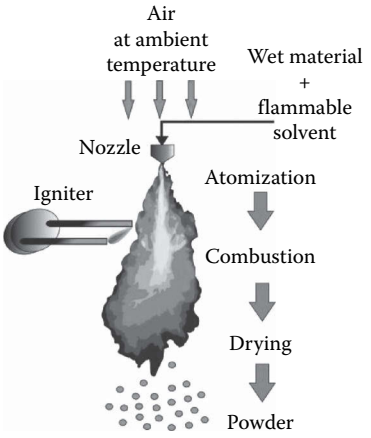


FIGURE 4.29.29 Mechanism of flame spray drying (Piatkowski et al., 2014).

For comparison, energy consumption of the FSD process was estimated to be 23.92 MJ/kg of evaporated moisture, whereas energy consumption of an equivalent standard spray drying process with electrical heating of the drying air was 36.11 MJ/kg water (Piatkowski et al., 2015).

Comparing to the standard spray drying process, dried powder obtained in FSD is characterized by reduced bulk density and increased particle diameters due to agglomeration and puffing which take place in the combustion zone. Table 4.29.5 shows a comparison of bulk density for maltodextrin powder produced using the standard spray drying process and FSD.

FSD is a new technique, without industrial applications, especially suitable for heat-resistant products.

Principles of Dryer Design

Acquisition of Drying Kinetics

An important step in the design of industrial dryers is the analysis of drying kinetics for a particular material, which can be carried out in laboratory or pilot-plant dryers. Kemp et al. (2001) distinguished four methods used to acquire experimental data on drying kinetics: periodic sampling, continuous weighing, intermittent weighing, and indirect methods. In periodic sampling, the drying process is interrupted to determine the weight of dried sample. In continuous weighing, the sample is located on a balance inside the drying chamber and sample weight loss is recorded continuously through the whole drying process. Intermitted weighing is a combination of the two aforementioned

TABLE 4.29.5
Comparison of Maltodextrin Powder Bulk Density for FSD and Standard Spray Drying Process

Fuel Content in the Feed (mass%)	Solids Content in the Fuel-Free Solution (mass%)	Maltodextrin Bulk Density (kg/m³)	Maltodextrin Apparent Density (kg/m³)
0[a]	30	567	1410
35	30	228	1021
45	30	145	793
0[a]	40	588	1358
35	40	153	937
45	40	126	849

[a] Standard spray drying process.

methods, i.e., the weight of drying sample is determined inside the drying chamber; however, at the moment of weighing, the drying process is interrupted for system stabilization. In indirect methods, data on drying kinetics are obtained from exhaust air humidity measurements.

Accurate determination of drying kinetics for disperse systems is still a challenging issue due to very short drying time. For example, drying time in an industrial-scale spray drying tower varies in the range from 1 to 10 s. Drying time in a laboratory-scale spray dryer is much longer due to large size of the analyzed droplets (from 500 to 2000 μm), therefore drying kinetics data obtained in the small scale does not represent the real spray drying kinetics (Zbicinski et al., 2005). Experimental methods for determination of spray drying kinetics can be divided into two groups: single droplet drying methods or fast drying of thin film methods. In single droplet technique, the droplet is fixed in one position due to acoustic field produced by acoustic levitator or hanged on glass filament. The data on drying kinetics of droplet are obtained from the exhaust air humidity (Groenewold et al., 2002).

In thin film method, spray drying kinetics is determined on the basis of fast drying (drying times comparable to industrial-scale drying towers) of thin layers of wet material. An example of the experimental installation, i.e., drying tunnel, elaborated for the determination of spray drying kinetics in thin film is presented in Figure 4.29.30 (Zbicinski et al., 2005). The thin layer of wet material is placed on the measuring pan mounted on a balance to record weight loss of the sample continuously during drying. Two electric heating systems provide high-temperature drying air (from 100°C to 500°C) to ensure fast drying of the sample (seconds).

The measurement of drying air temperature and humidity in the spray drying chamber is also a challenging issue as particles present in the air may hit the sensor making incorrect temperature reading. To separate gaseous phase from dispersed droplets/particles and conduct measurement in the clean air, microseparators are used (Wawrzyniak et al., 2012a).

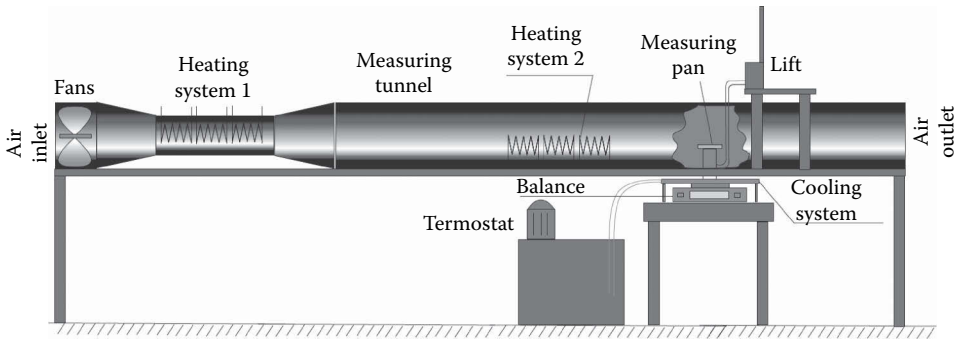


FIGURE 4.29.30 Drying tunnel for spray drying kinetics determination (Zbicinski et al., 2005).

Calculations in Drying

Drying is a complex process involving heat and mass transfer, flow of the air, water evaporation, dried material moisture diffusion, dryer geometry, and many other factors. In the steady-state operation, drying conditions are changing along the dryer, whereas in batch operation drying conditions are also changing with time.

Energy for Moisture Evaporation

During drying, moisture in the dried material is evaporated, so the heat of evaporation must be supplied. The minimum amount of energy necessary for drying process can be calculated as

$$E_{\mathrm{v,min}} = M_{\mathrm{v}} \cdot H_{\mathrm{v}} \tag{4.29.8}$$

or

$$Q_{\mathrm{v,min}} = W_{\mathrm{v}} \cdot H_{\mathrm{v}} \tag{4.29.9}$$

where

$E_{\mathrm{v,min}}$ minimum heat energy for evaporation; kJ
M_{v} evaporation load; kg
ΔH_{v} heat of vaporization; kJ/kg
Q_{v} heat supply rate; kW
W_{v} evaporation rate; kg/s

For steady state, Equation 4.29.9 takes the form

$$Q_{\mathrm{v,min}} = W_{\mathrm{s}} \left(X_{\mathrm{in}} - X_{\mathrm{out}} \right) \cdot H_{\mathrm{v}} \tag{4.29.10}$$

Heat of evaporation is the function of temperature. Considering water being the most common moisture in drying, the latent heat changes from 2501 kJ/kg at 0°C to 2256 kJ/kg at normal boiling point.

Energy for Drying

Drying always uses more energy than just the latent heat of evaporation. Considering a continuous convective dryer using hot air as a drying agent, the amount of energy necessary for heating the inlet air can be calculated with

$$Q_{\mathrm{heating}} = W_{\mathrm{g}} \cdot c_{\mathrm{pg}} \left(T_{\mathrm{g,in}} - T_{\mathrm{g,a}} \right) \tag{4.29.11}$$

where

$T_{\mathrm{g,in}}$ air temperature at the inlet of the dryer; °C
$T_{\mathrm{g,a}}$ ambient air temperature; °C

The general heat balance of a continuous dryer proposed by Kemp (2012) reads

$$W_{\mathrm{g}} \cdot c_{\mathrm{pg}} \left(T_{\mathrm{g,in}} - T_{\mathrm{g,out}} \right) \approx W_{\mathrm{s}} \cdot c_{\mathrm{ps}} \left(X_{\mathrm{in}} - X_{\mathrm{out}} \right) \cdot H_{\mathrm{v}} + W_{\mathrm{s}} \cdot c_{\mathrm{ps}} \left(T_{\mathrm{s,out}} - T_{\mathrm{s,in}} \right) + Q_{\mathrm{loss}} \tag{4.29.12}$$

Combining both equations, one can receive

$$Q_{\mathrm{heating}} = \frac{\left(T_{\mathrm{g,in}} - T_{\mathrm{g,a}} \right)}{\left(T_{\mathrm{g,in}} - T_{\mathrm{g,out}} \right)} \left[W_{\mathrm{s}} \cdot c_{\mathrm{ps}} \left(X_{\mathrm{in}} - X_{\mathrm{out}} \right) \cdot H_{\mathrm{v}} + Q_{\mathrm{s,\,sens}} + Q_{\mathrm{loss}} \right] \tag{4.29.13}$$

Equation 4.29.13 represents heat energy for a perfect adiabatic dryer. Comparing to evaporative load (Equation 4.29.12), heat lost with the exhaust gas, energy for heating the solids, and heat losses from the dryer are added.

In the last years, computational fluid dynamics (CFD) technique has found a wide application in designing of drying units, including scaling up, optimization of drying process parameters,

improvement of energy efficiency, reduction of wall deposits, etc. Commercial solvers available on the market make CFD technique a standard tool in drying modeling and scaling up.

Safety in Drying

Many dryers, especially disperse dryers, produce dust that might explode in a violent combustion reaction if it is dispersed in the air. If the dust combustion takes place within the process equipment, the heat of combustion generates a rapid development of pressure, releasing large amount of heat, and finally ends with explosion (Skjold et al., 2006). To explode, the dust–air mixture must be ignited and the concentration of dust must be between the lower (LEL) and the upper explosion limit (UEL). If the dust concentration is below LEL, the explosion will not occur because of too small amount of dust particles in the mixture to promote combustion. If dust concentration is above the UEL, there is not enough air to sustain combustion.

Procedures to determine the main parameters that describe dust explosibility (LEL, UEL, Dust Explosion Index, Minimum Ignition Energy, etc.) are given in the European Standard EN 14034.

Figure 4.29.31 shows examples of pressure distribution in an industrial spray tower for detergent drying obtained from a CFD simulations of dust–air mixture explosion (Wawrzyniak et al., 2012b). Analysis of Figure 4.29.31 shows a fast increase of pressure due to a quick propagation of shock wave in the tower (5.3 bar after 2.5 s).

Quick increase of pressure in the dryer shows how dangerous for the staff and equipment can be the results of dust explosion. The lack of safety vents or their improper configuration can lead to significant damage of equipment and even loss of life.

Explosion venting is the most common method to mitigate dust explosions. For each dryer, process or product, vent area sizing, and the efficiency of the venting system should be determined. Violence of the explosion falls as the moisture content of the dust increases; above 25%, the dust is not exposible (Barton, 2002).

Selection of Drying Method

Currently, there are many different drying methods encountered in the industry. During over 130 years of progress, each of the methods was developed to many design solutions so finally there are several hundred drying techniques in the industry.

In many cases, drying is the removal of water or other volatiles from granular or paste-like materials, but drying of solutions or liquid suspensions is also common. Analysis of the drying problem usually limits the choice of the dryer to two or three possibilities; it does not indicate the type of dryer that must be used. Selection of the dryer is then made on the basis of capital and operating

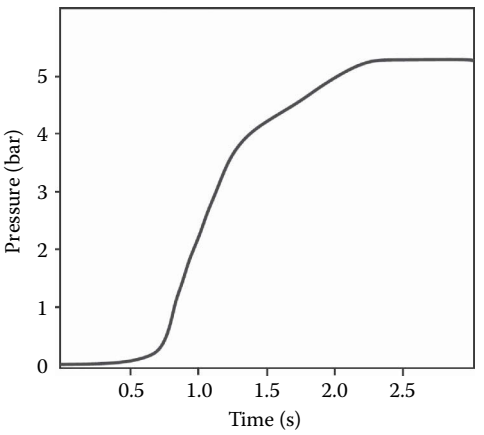


FIGURE 4.29.31 Propagation of pressure wave in the spray tower after dust explosion (Wawrzyniak, 2012).

costs. Attention must be paid, however, to the costs of the entire drying system, not just the drying unit alone. There is no universal method of dryer selection, some general rules exist, but they are not rigid and exceptions are not unusual.

Mujumdar advised to consider the following quantitative information to arrive at a suitable dryer (Mujumdar, 2007):

Dryer throughput; mode of feedstock production (batch/continuous)
Physical, chemical, and biochemical properties of the wet feed
Desired product specifications; expected variability in feed characteristics
Upstream and downstream processing operations
Moisture content of the feed and product
Drying kinetics; moist solid sorption isotherms
Quality parameters (physical, chemical, biochemical)
Safety aspects, e.g., fire and explosion hazards, toxicity
Value of the product
Need for automatic control
Toxicological properties of the product
Turndown ratio, flexibility in capacity requirements
Type and cost of fuel, cost of electricity
Environmental regulations
Space in plant

An example of classifications of drying methods based on type of dryer operation, way of heat transfer to the material, and form of dried material is shown in Figure 4.29.32.

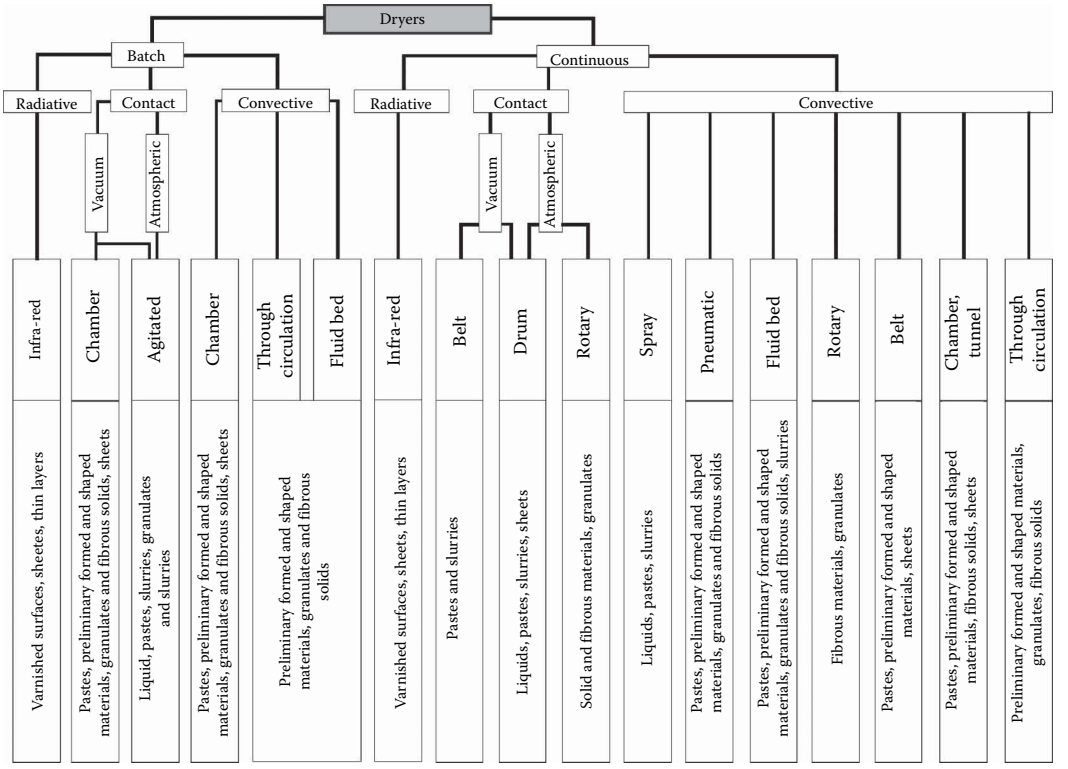


FIGURE 4.29.32 Classification of drying methods. (From Strumillo, C. and Kudra, T., *Drying Principles, Applications and Design,* Gordon and Breach Science Publishers, New York, 1986.)

Batch dryers are used when product capacity is below 200 kg/h, whereas continuous equipment is appropriate for capacity greater than 1 ton/h. The dryer should operate safely and reliably. Capital, operating, and maintenance costs must be reasonable and energy consumption must be minimized. At the same time, all the emissions (hot air, dust) must be controlled. In some cases, an auxiliary system for dust separation could be as expensive as the dryer itself.

Nomenclature

c_{Av}	mean heat capacity of water vapor, J/(kg K)
c_B	mean heat capacity of dry air, J/(kg K)
c_{pg}	specific heat, J/(kg K)
$E_{v,min}$	minimum heat energy for evaporation, kJ
ΔH	latent heat of vaporization, J/kg
I_{Av}	enthalpy of water vapor, J/kg
I_B	enthalpy of dry air, J/kg
I_G	enthalpy of humid air, J/kg
k_g	thermal conductivity, W/(m K)
k_y	convective mass transfer coefficient, m/s
M	molar mass, kg/mol
M_v	evaporation load; kg
P_v	vapor pressure of pure water, Pa
q_{cw}	heat flux at wall, W/m^2
Q_v	heat supply rate, kW
R	universal gas constant, 8.314 J/(mol K)
RH	relative humidity, fraction
T	temperature, °C
T_w	wall temperature, °C
w_{cw}	moisture flux at wall, kg/(m^2 s)
W	flow rate, kg/s
W_v	evaporation rate, kg/s
X	total moisture content, kg water/kg dry solid,
X_{cr}	critical moisture content, kg water/kg dry solid,
X_e	equilibrium moisture content, kg water/kg dry solid,
Y	absolute air humidity, kg water vapor/kg dry air
$Q_{v,min}$	heat supply rate, kW

Greek letters

α	heat transfer coefficient, W/(m^2 K)
ρ_g	density, kg/m^3
μ_g	dynamic viscosity, kg/m s

Subscripts

a	ambient
c	convection
s	solid
v	vapor
w	Wall
Ref	Reference

REFERENCES

Barton, J. 2002. Guide to dust explosion prevention and protection, Parts 1–3, ISBN: 08295 293 7/8/9, Institution of Chemical Engineers.

Beristain, C.I., Garcia, H.S., and Vermon-Carter, E.J. 2001. Spray-dried encapsulation of cardamom (*Elettaria cardamomum*) essential oil with mesquite (*Prosopis juliflora*) gum. *Lebensmittel-Wissenschaft und Technologie* 34: 398–401.

Borde, I. and Levy, A. 2007. Pneumatic and flesh drying. In: *Handbook of Industrial Drying*. A. S. Mujumdar (Ed.). 3rd edition. Taylor & Francis CRC Press, Boca Raton, 397–410.

Charlesworth, D.H. and Marshall, W.R. 1960. Evaporation from drops containing dissolved solids. *American Institute of Chemical Engineers Journal* 6 (1): 9–23.

Claussen, I.C., Ustad, T.S., Strommen, I., and Walde, P.M. 2007. Atmospheric freeze drying—A review. *Drying Technology* 25 (6): 947–957.

Daud, W. R. W. 2006. Drum dryers. In: *Handbook of Industrial Drying*. A. S. Mujumdar (Ed.). 3rd edition. Taylor & Francis CRC Press, Boca Raton, 203–213.

Devahastin, S. 2008. Superheated steam drying—Special issue of drying technology. *Drying Technology* 26 (2): 157–230.

Elustondo, D., Elustondo, M.P., and Urbicain, M.J. 2001. Mathematical modelling of moisture evaporation from foodstuffs exposed to sub-atmospheric pressure superheated steam. *Journal of Food Engineering* 49: 15–24.

Gibbs, B.F., Kermasha, S., Alli, I., and Mulligan, C.N. 1999. Encapsulation in the food industry: A review. *International Journal of Food Sciences and Nutrition* 50: 213–224.

Gouin, S. 2004. Microencapsulation: Industrial appraisal of existing technologies and trends. *Trends in Food Science and Technology* 15: 330–347.

Groenewold, C., Möser, C., Groenewold, H., and Tsotsas, E. 2002. Determination of single-particle drying kinetics in an acoustic levitator. *Chemical Engineering Journal* 86 (1–2): 217–222.

Hasheminya, S.M. and Dehghannya, J. 2013. Spray dryers: Applications, performance, essential parts and classifications. *International Journal of Farming and Allied Sciences* 2 (19): 756–759.

Kemp, I.C., Fyhr, B.C., Laurent, S., Roques, M.A., Groenewold, C.E., Tsotsas, E., Sereno, A.A., Bonazzi, C.B., Bimbenet, J.J., and Kind, M. 2001. Methods for processing experimental drying kinetics data. *Drying Technology* 19 (1): 15–34.

Kemp, I.C. 2012. Fundamentals of energy analysis of dryers. In *Modern Drying Technology Volume 4: Energy Savings*. E. Tsotsas and A.S. Mujumdar (Ed.). Wiley-VCH Verlag GmbH & Co., Weinheim, Germany, 1–45.

King, C.J. 1995. Spray drying: Retention of volatile compounds revised. *Drying Technology* 13 (5–7): 1221–1240.

Kowalski, S.J. and Mierzwa, D. 2015. US-assisted convective drying of biological materials. *Drying Technology* 33 (13): 1601–1613.

Krishnamurthy, K., Khurana, H.K., Jun, S., Irudayaraj, J., and Demirci, A. 2008. Infrared heating in food processing: An overview. *Comprehensive Reviews in Food Science and Food Safety* 7 (1): 2–13.

Kunii, D. and Levenspiel, O. 2013. *Fluidization Engineering, Butterworth-Heinemann series in chemical engineering*. Elsevier, Stoneham, MA.

Kwapinska, M. 2002. Effect of drying and atomization parameters on product morphology (in Polish), PhD thesis. Lodz university of Technology, Poland.

Liapis, A.I. and Bruttini, R. 2007. Indirect dryers. In: *Handbook of Industrial Drying*. A. S. Mujumdar (Ed.). 3rd edition. Taylor & Francis CRC Press, Boca Raton, 257–283.

Masters, K. 1991. *Spray Drying Handbook*. John Wiley & Sons, New York.

Meng, X., De Jong, W., and Kudra, T. 2016. A state-of-the-art review of pulse combustion: Principles, modeling, applications and R&D issues. *Renewable and Sustainable Energy Reviews* 55, 73–114.

Mujumdar, A.S. 2007. Principles, classification, and selection of dryers. In: *Handbook of Industrial Drying*. A. S. Mujumdar (Ed.). 3rd edition. Taylor & Francis CRC Press, Boca Raton, 4–20.

Pakowski, Z. and Mujumdar, A.S. 1995. Basic process calculations in drying. In: *Handbook of Industrial Drying*. A. S. Mujumdar (Ed.). 2nd edition. Marcel Dekker, New York, 71–112.

Pakowski, Z., Krupinska, B., and Adamski, R. 2007. Prediction of sorption equilibrium both in air and superheated steam drying of energetic variety of willow *Salix viminalis* in wide temperature range. *Fuel* 86: 1749–1757.

Piatkowski, M., Taradaichenko, M., and Zbicinski, I. 2014. Flame spray drying. *Drying Technology* 32 (11): 1343–1351.

Piatkowski, M., Taradaichenko, M., and Zbicinski, I. 2015. Energy consumption and product quality interactions in flame spray drying. *Drying Technology* 33 (9): 1022–1028.

Pisano, R., Barresi, A.A., and Fissore, D. 2011. Innovation in monitoring food freeze drying. *Drying Technology* 29 (16): 1920–1931.

Rabaeva, Y. 2012. Kinetics of foam-spray drying process, PhD thesis. Lodz University of Technology, Poland.

Rahse, W. and Dicoi, O. 2001. Spray drying in the detergent industry, *Spray Drying '01 and Related Processes*. Universitat Dortmund, Dortmund: 83–87.

Rattanadecho, P. and Makul, N. 2016. Microwave-assisted drying: A review of the state-of-the-art. *Drying Technology* 34 (1): 1–38.

Reyes, A., Vega, R.V., and Bruna, R.D. 2010. Effect of operating conditions in atmospheric freeze drying of carrot particles in a pulsed fluidized bed. *Drying Technology* 28 (10): 1185–1192.

Riadh, M.H., Ahmad, S.A.B., Marhaban, M.H., and Soh, A.Ch. 2015. Infrared heating in food drying: An overview. *Drying Technology* 33 (3): 322–335.

Risch, S.J. 1995. Encapsulation: Overview of uses and techniques. In: *Encapsulation and Controlled Release of Food Ingredients*, ACS Symposium Series 590. S.J. Risch and G.A. Reineccius (Ed.). American Chemical Society, Washington, 2–7.

Rodriguez, O., Santacatalina, J.V., Simal, S., Garcia-Perez, J.V., Femenia, A., and Rossello, C. 2014. Influence of power ultrasound application on drying kinetics of apple and its antioxidant and microstructural properties. *Journal of Food Engineering* 129: 21–29.

Rodriguez, O., Llabres, P.J., Simal, S., Femenia, A., and Rossello, C. 2015. Intensification of predrying treatments by means of ultrasonic assistance: Effects on water mobility, PPO activity, microstructure, and drying kinetics of apple. *Food Bioprocess Technology* 8 (3): 503–515.

Skjold, T., Arntzenb, B.J., Hansena, O.R., Storvika, I.E., and Eckhoff, R.K. 2006. Simulation of dust explosions in complex geometries with experimental input from standardized tests. *Journal of Loss Prevention in Process Industries* 19: 210–217.

Soponronnarit, S., Prachayawarakorn, S., Rordprat, W., Nathakaranakule, A., and Tia, W. 2006. A superheated steam fluidized bed dryer for parboiled rice: Testing of pilot-scale and mathematical model development. *Drying Technology* 24 (11): 1457–1467.

Strumillo, C., and Kudra, T. 1986. *Drying Principles, Applications and Design, Topics in Chemical Engineering*. R. Hughes (Ed.). Gordon and Breach Science Publishers, New York.

Taylor, T. 2000. Concurrent and countercurrent spray drying and particle morphology. Paper No. 9, *12th International Drying Symposium*, Noordwijkerhout, The Netherlands, August 28–31, 2000.

Thies, C. 1996. A Survey of microencapsulation processes. In: *Microencapsulation: Methods and Industrial Applications*. S. Benita (Ed.). Marcel Dekker Inc., New York, 1–19.

Verhey, J.G.P. 1973. Vacuole formation in spray powder particles. 3. Atomization and droplet drying. *Netherlands Milk and Dairy Journal* 27: 3–18.

Walton, D.E. 2000. The morphology of spray-dried particles a qualitative view. *Drying Technology* 18 (9): 1945–1986.

Wawrzyniak, P., Podyma, M., Zbicinski, I., Bartczak, Z., Polanczyk, A., and Rabaeva, J. 2012a. Model of heat and mass transfer in an industrial counter-current spray-drying tower. *Drying Technology* 30 (11–12): 1274–1282.

Wawrzyniak, P., Polanczyk, A., Zbicinski, I., Jaskulski, M., Podyma, M., and Rabaeva, J. 2012b. Modeling of dust explosion in the industrial spray dryer. *Drying Technology* 30 (15): 1720–1729.

Witrowa-Rajchert, D., Wiktor, A., Sledz, M., and Nowacka, M. 2014. Selected emerging technologies to enhance the drying process: A review. *Drying Technology* 32 (11), 1386–1396.

Zbicinski, I., Piatkowski, M., and Prajs, W. 2005. Determination of spray-drying kinetics in small scale. *Drying Technology* 23 (8): 1751–1759.

Zbicinski, I. and Rabaeva, J. 2009. Analysis of gas admixing foam spray-drying process. *Drying Technology* 28 (1): 103–110.

Zbicinski, I., Kudra T., and Liu, X. 2014a. Pulse combustion drying. In: *Modern Drying Technology*. Process Intensification. E. Tsotsas and A.S. Mujumdar (Ed.). Wiley-Vch Verlag Weinheim, Germany, 27–56.

Zbicinski, I., Rabajewa, J., and Lewandowski, A. 2014b. Drying of foamed materials. In: *Modern Drying Technology*. Process Intensification. E. Tsotsas and A.S. Mujumdar (Ed.). Wiley-Vch Verlag, Weinheim, Germany, 163–190.

Zhang, N., Shi, L., Qi, N., Xie, Y., and Cai, L. 2016. Effect of radio frequency (RF) drying technology on dehydration rate and energy consumption of Australia lignite. *Drying Technology* 34 (2): 161–166.

4.30 HEAT TRANSFER IN ROTARY KILNS

P.S. Ghoshdastidar

Introduction

A rotary kiln consists of a long refractory lined cylindrical shell mounted at a slight incline from the horizontal plane. The kiln is slowly rotated about its longitudinal axis and the solid charge is fed into the upper end of the cylinder and heated by countercurrent or cocurrent flow of hot combustion gases. During the process, the material moves to the lower end, where it is discharged. Rotary kiln reactors are used in the production of cement and sponge iron, several chemicals such as lime, phosphates and polyphosphates, zinc white, lithopone, titanium dioxide, soda ash, alumina, catalyst, and catalyst carriers. Incineration of solid waste and drying of wet solids are also performed in such kilns. Food products can be dried by superheated steam in rotary dryers. Figure 4.30.1 shows a schematic diagram of a rotary kiln used for drying and preheating of wet solids. The cross section of a rotary kiln is shown in Figure 4.30.2. The angle, α, is called the fill angle, which signifies the region of volume containing the solid. The primary advantages of rotary kilns are its simple design and great flexibility. It can be easily adaptable to changing operating conditions such as varying kiln speed, kiln inclination angle, and mass flow rates of gas and solids. Very-high-temperature processes can be carried out in rotary kilns. Solid granules of various shapes and sizes can be handled in such reactors. The main disadvantages are (1) low space–time yield, (2) inefficient energy utilization, and (3) difficult control of temperature and composition of the reactant mixture in the middle zone of the kiln [1]. Thus, there is an urgent need for optimization of rotary kilns with respect to their length, specific power input, and production capacity.

Salient Features of a Rotary Kiln

Rotary kilns are high-temperature reactors often operating in the temperature zone above 900 K. The required reaction temperature of the solid bed is achieved by burning gas or oil in the kiln (internally fired kilns) above the solid surface. Usually, a burner is positioned at the end of the kiln and flow of the hot combustion gas mixture takes place countercurrent to the solid movement. However, examples of externally fired kilns and cocurrent gas movement are also found in the industry. Only a small part of the cross section of the kiln (typically, 10%–20%) is occupied by the solids. Industrial rotary kilns often have very high length-to-diameter ratios. They may be as long as 200 m. The diameter of a rotary kiln can be as large as 5 m. The kiln inclination angle and kiln rotational speed normally does not exceed 5° and 5 rpm, respectively. The small inclination angle and low speed ensures long residence time of solids in the kiln, and enhanced contact of the solid with the hot gas. As regards the internal features of the kiln, refractory lining is used in order to insulate and protect the outer shell from thermal damage and save energy [2]. Kilns may often be equipped with lifters and tumblers to ease the motion of the solid particles axially and improve particle mixing [2].

Firing of Rotary Kilns

Kilns are usually fired by burning natural gas (CH_4) or fuel oil or pulverized solid fuels such as coal or petroleum coke. Sometimes a combination of all of these is also used [2]. Because of high temperature in the kiln the thermal radiation becomes the dominant heat transfer mode. Hence, fuels having high luminosity due to soot formation during combustion are preferred. Natural gas is normally the last choice of fuel selection except in places with strict environmental regulations on CO_2 emissions because of its high cost and transparent nature of its flame. Most rotary kilns use coal or petroleum coke as the main fuel for combustion.

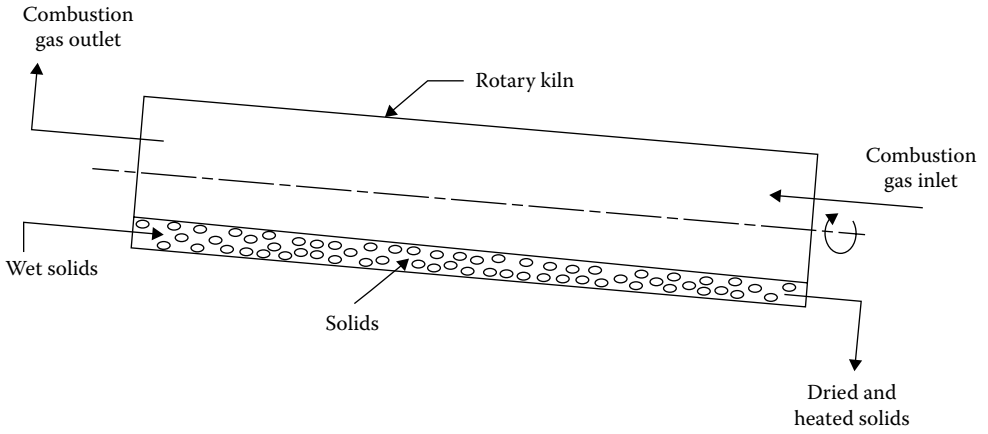


FIGURE 4.30.1 Schematic diagram of a rotary kiln used as a dryer.

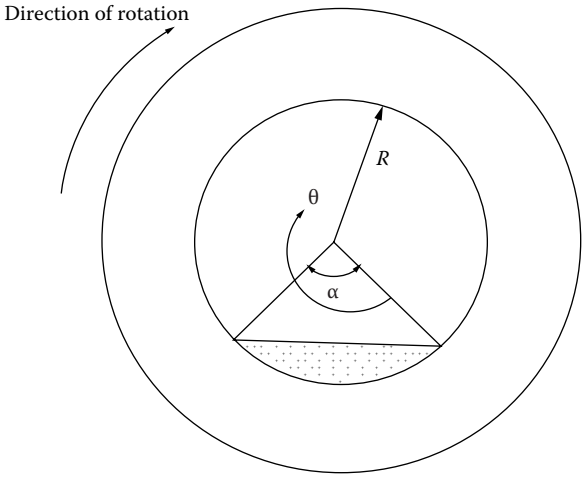


FIGURE 4.30.2 Cross section of a rotary kiln. (From Ghoshdastidar, P.S., Agarwal, A., *ASME J. Therm. Sci. Eng. Appl.*, 1, 024501, 2009. With permission.)

Phenomenon within the Solid Bed

The solid bed is composed of granular materials of various sizes and shapes. The bed is heated by the gas as well as the hot kiln wall. Heat transfer path consists of particle-to-particle conduction and radiation and interstitial gas-to-particle convection. However, transverse motion of the particles themselves results in energy transport by convection that can be a dominant heat transfer mode. Thus, modeling of heat transfer in the solid bed is a challenging task. Most simulations consider the solid bed as well mixed in the transverse plane and at a uniform temperature for the sake of simplicity.

Transverse Bed Motion

Based on the rotational speed of the kiln, the bed motion in the transverse plane can be classified into the following main categories:

1. *Centrifuging*: This is an extreme condition visualized at critical and high speeds in which all the bed materials rotate with the drum.

2. *Cascading*: This also occurs at high speeds. In this mode the height of the leading edge of the particles rises above the bed surface and granules cascade or shower down on the free surface.
3. *Rolling*: At very low rotational speeds the bed slips against the wall. In this mode, mixing is maximum. There are two distinct regions, namely, the shearing region, called the active layer, constituted by the particles near the free surface, and the passive or plug flow region at the bottom having zero shear rate. Most rotary kilns operate in the rolling mode. A rolling bed is usually flat and subtends an angle with the horizontal, called the *angle of repose*. The active layer is thinner than the passive layer because particles there move faster.

Mixing in the Solid Bed

In the rolling mode, most of the mixing (and dissociation reactions in case of chemical production) occurs in the active layer of the bed cross section. A deeper active region induces better mixing of solid particles. Increasing the kiln speed increases the depth of the active layer. It has been experimentally observed that particles move forward down the kiln through the active region [2]. Because the kiln inclination angle is small and far below the angle of repose the forward movement is not sustained by the component of gravity in the axial direction alone. Actually, the rotational speed also plays a significant role in the axial motion of the granules. For every revolution, the bed material makes several circulations in the cross section, thereby producing the axial motion [2]. Increased kiln rpm results in enhanced mixing but results in less residence time of the material in the kiln. Thus, the operator of the kiln has to strike a right balance between the two opposite effects and therein lies the need for optimization in kiln design.

Axial Velocity of Solid Particles

The axial motion of a solid particle is dependent on the kiln inclination angle and transverse motion, which in turn is dependent on the kiln rotational speed. Ghoshdastidar and Agarwal [3] showed the derivation of an expression for the axial velocity, V_z, of the solids in the kiln, based on the velocity vector diagram of a solid particle. Figure 4.30.3 shows the diagram for $\phi = 5°$.

The expression for V_z is written as

$$V_z = U \tan \varphi \tag{4.30.1}$$

where U is the circumferential speed (m/s) of the kiln and ϕ is the kiln inclination angle in degrees.

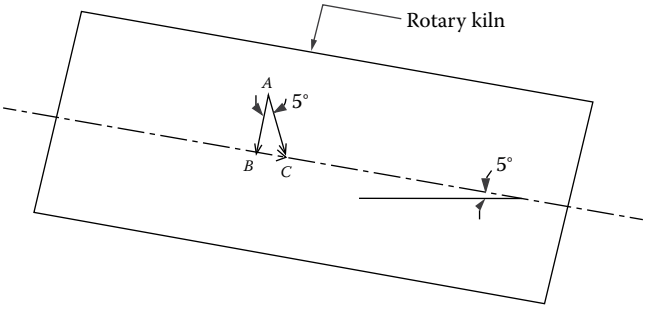


FIGURE 4.30.3 Velocity vector diagram for a solid particle in a rotary kiln. (From Ghoshdastidar, P.S., Agarwal, A., *ASME J. Therm. Sci. Eng. Appl.*, 1, 024501, 2009. With permission.)

Note that

$$U = \frac{\pi DS}{60} \tag{4.30.2}$$

where S is the kiln rotational speed (rpm).

Several other expressions for axial velocity of the solids are also found in the literature [2].

Heat Transfer Processes in a Rotary Kiln

Figure 4.30.4 shows a schematic of heat transfer processes in a rotary kiln.

The refractory wall elements are alternately heated and cooled by the hot gas and solid, respectively. Hence, quasi-steady heat conduction is present. The gas is transferring heat to both inside wall surface and solid bed by convection and radiation. The solid is heated by the gas as well as by the hot wall. The gas convection is usually turbulent. The modeling of gas convection is challenging because of the irregular geometry of the flow cross section. The wall to solid heat transfer is generally modeled using a contact heat transfer coefficient. The outer wall is losing heat by both convection and radiation. Cartesian coordinates are used because the refractory thickness is much smaller than the kiln diameter.

Rotary Kiln Modeling: Various Approaches

From the description of heat transfer processes in a rotary kiln, it is evident that modeling of rotary kilns is quite complex. The assumption of well-mixed freeboard and bed led to the development of one-dimensional (1D) model that predicts the axial solids and gas temperature distribution from the inlet to the outlet of the kiln by solving ordinary differential equations as proposed by Sass [4]. A quasi-three-dimensional (3D) model combines both an axial model (1D]) and a cross-sectional model (two-dimensional [2D]). The surface heat flux determined from the 1D model is used as a thermal boundary condition for the cross-sectional model. The 2D model is employed to determine the temperature distribution in the solid bed material and the kiln wall over axial segments. Very few models have the capability of predicting kiln length for a given set of input parameters. In the next section, modeling of a rotary food dryer is given in some detail. This is a quasi-3D model that predicts the kiln length and calculates axial solids and gas temperature distributions. However, the solid bed and gas temperatures have been considered as lumped in each axial segment.

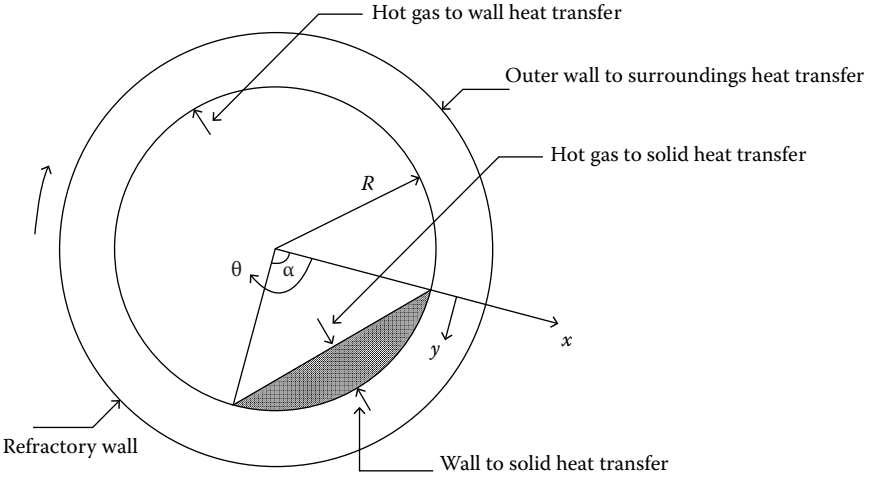


FIGURE 4.30.4 Schematic cross section of a rotary kiln showing heat transfer processes.

Drying of Food Products in a Rotary Kiln

In this section, the model of a rotary kiln used for drying and preheating of apple and carrot pieces with superheated steam at 1 bar developed by Sinhal et al. [5] is presented. The advantages of superheated steam as a drying medium are (1) energy recovery through condensation of surplus steam and (2) drying of food products in a sterilized atmosphere. The input data for the food dryer of Sinhal et al. [5] are listed in Table 4.30.1.

The dryer can be considered to comprise three zones. In the first zone, the food pieces are heated to the saturation temperature of the entrained water. In the middle zone, the water evaporates at constant temperature until the food pieces are completely dry. In the third or the last section, the dry solids are heated to some prescribed temperature and are released from the kiln.

TABLE 4.30.1
Input Data for the Food Dryer

1. Rotary Kiln	
a. Diameter (inner)	2.29 m
b. Refractory	
i. Thickness	0.23 m
ii. Thermal conductivity	0.6057 W/m K
iii. Specific heat	692.56 J/kg K
iv. Density	3000 kg/m^3
v. Emissivity	0.9
c. Rotational speed[a]	5 rpm
d. Angle of inclination[a]	5°
2. Solid	
a. Inlet temperature	288.6 K
b. Outlet temperature	413 K
c. Mass flow rate (dry)[a]	0.3 kg/s
d. Percent water with respect to dry solid (by weight)[a]	19
e. Emissivity	0.95
3. Gas (superheated steam at 1 bar)	
a. Outlet temperature[a]	416 K
b. Mass flow rate[a]	3.9 kg/s
c. Specific heat	2060.4 J/kg K
4. Water	
a. Latent heat of vaporization	2.26×10^6 J/kg
b. Specific heat (vapor)	2083.2 J/kg K
c. Specific heat (liquid)	4211.2 J/kg K
4. Heat transfer coefficients	
a. Contact heat transfer coefficient	150 W/m^2 K
b. Convective heat transfer coefficient from the kiln wall to surrounding	10 W/m^{2}K
5. Ambient temperature outside the kiln	300 K

Source: Sinhal, K. et al., *ASME J. Therm. Sci. Eng. Appl.*, 4, 011009, 2012.

[a] Base values.

Problem Formulation

Thermal Radiation among Hot Gas, Refractory Wall, and Solid Surface

Because the temperature involved is high thermal radiation plays a significant role in the exchange of heat among the hot gas, the inner wall of the kiln and the solid surface. The wall is divided into the surface elements to form an enclosure as shown in Figure 4.30.5. Each axial segment is further subdivided into 20 circumferential elements of equal size and five solid surface elements.

The surfaces are assumed to be diffuse and gray. It is assumed that the surface elements exchange radiation only with surfaces of the same axial segment because the segment lengths (Δz in Figure 4.30.5) are sufficiently long.

The theory of Hottel [6] is used to calculate radiation heat transfer in a gray enclosure containing a gas which emits, absorbs, and transmits radiation. Thus, in an enclosure having N surfaces at different temperatures, the net heat loss/gain, q_j, at a particular surface can be expressed as

$$q_j = \frac{\varepsilon+1}{2}\left[A_g F_{gj}\varepsilon_g(T_g)E_{bg} + \sum_{i=1}^{N} A_i F_{ij}\tau_g(T_i)E_{bi} - E_{bj}A_j\right] \tag{4.30.3}$$

where E is the emissive power (in W/m^2) and ε is the emissivity of the enclosure walls. Equation 4.30.3 is valid for $\varepsilon > 0.8$. The emissivity of the refractory wall is taken as 0.9 [7]. The emissivity of the solid (apple or carrot) is taken as 0.95 [8]. ε_g (T_g) and τ_g (T_i) are emissivity and transmissivity of the gas at T_g and T_i, respectively. For the calculation of the gas emissivity and other characteristics, Mean Beam Length Theory is used (Holman [6]). The temperature of the solid surface element is assumed to be uniform in each axial segment.

The view or shape factors between 20 surface elements and five solid surface elements are determined assuming the surface as infinite parallel strips [9]. The formulas for shape factor calculation are shown next [10].

Calculation of Shape Factors

The shape factors between the circumferential elements and solid surface elements are determined assuming the surfaces as infinite parallel strips. Surfaces (1) and (2) in Figure 4.30.6 are strips perpendicular to the XY plane. The shape factor F_{1-2} is given by Hottel (Jacob [9]) as follows:

$$F_{1-2} = \frac{\sin\phi' - \sin\phi''}{2} \tag{4.30.4}$$

where ϕ' is the angle subtended by the bottom point of surface (2) with the normal to surface (1), and ϕ'' is the angle subtended by the top point of surface (2) with the normal to surface (1). Both angles are measured clockwise from the normal to surface (1).

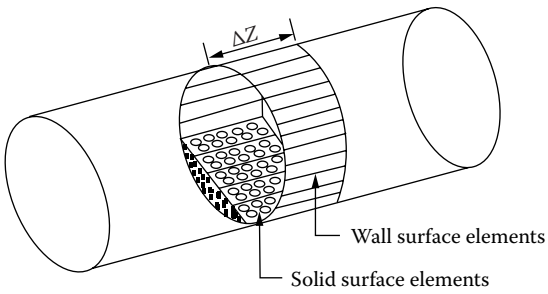


FIGURE 4.30.5 Section of the kiln showing surface elements of the wall and the solid. (From Sinhal, K. et al., *ASME J. Therm. Sci. Eng. Appl.*, 4, 011009, 2012. With permission.)

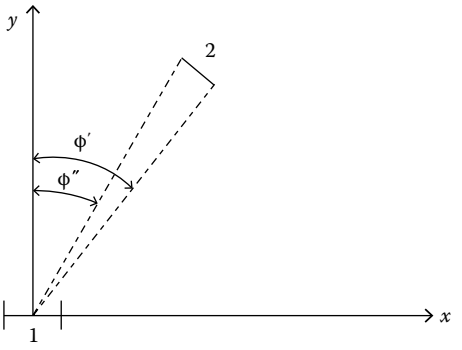


FIGURE 4.30.6 Shape factor between infinite parallel surfaces. (From Sinhal, K. et al., *ASME J. Therm. Sci. Eng. Appl.*, 4, 011009, 2012. With permission.)

Shape Factors between the Elements of the Kiln Wall

Equation 4.30.4 is written in terms of the circumferential location of the two surfaces. For the notations shown in Figure 4.30.7, the shape factor between surfaces (1) and (2) is given by

$$F_{1-2} = \frac{\Delta\theta \cos\beta}{4} \tag{4.30.5}$$

where $\Delta\theta$ is the arc angle of surface (2) in radian. The angle β is defined as

$$\beta = \frac{\pi - |\theta_2 - \theta_1|}{2} \tag{4.30.6}$$

where θ_1 and θ_2 are the circumferential positions of surface elements (1) and (2), respectively, with respect to the reference axis. For convenience, the reference axis is chosen perpendicular to the solid surface. Because of the absolute bracket in Equation 4.30.6, the shape factors F_{1-2} and F_{2-1} are equal for the same $\Delta\theta$ as required by the reciprocity relations of the shape factors.

Shape Factors between Elements of the Kiln Wall and Solid Surface

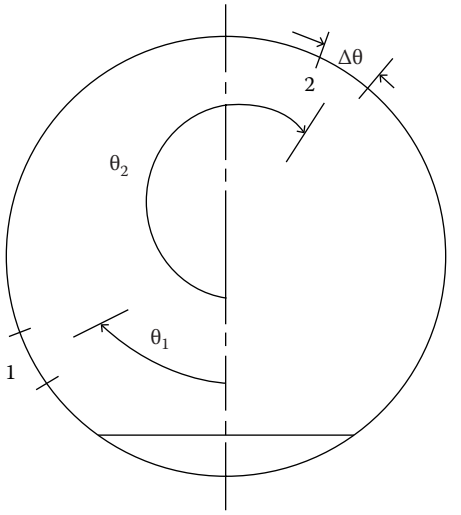


FIGURE 4.30.7 Shape factor nomenclatures when both surfaces are on refractory wall.

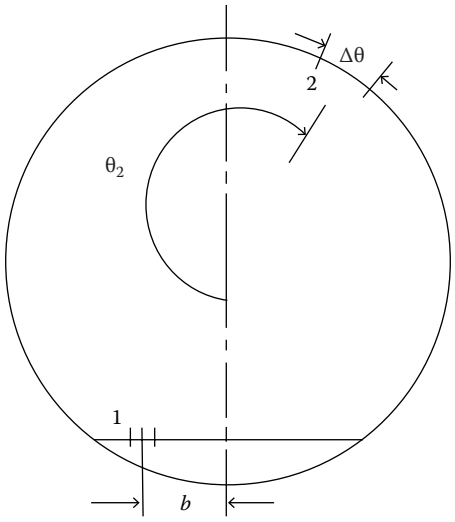


FIGURE 4.30.8 Shape factor nomenclatures when one surface is on refractory wall and other on solid.

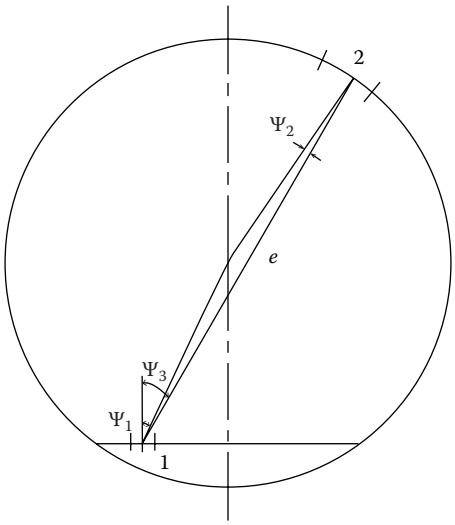


FIGURE 4.30.9 Parameters for shape factor equation.

In Figure 4.30.8, surface (1) is a solid surface element and surface (2) is a refractory wall element. Surface (1) is located at a distance b from the reference axis and θ_2 is the circumferential angle for surface (2). Angle α is the fill angle of the solid material. Additional terms that are shown in Figure 4.30.9 are needed to determine the shape factor.

The shape factor F_{1-2} is given by

$$F_{1-2} = \frac{D \cos \psi_1 \cos \psi_2 \Delta\theta}{4e} \tag{4.30.7}$$

where D is the inside diameter of the kiln. The angles ψ_1 and ψ_2 are shown in Figure 4.30.9 and are, respectively, the angles between the normal to surfaces (1) and (2) and the ray extending between

the surfaces. The distance e is the length of this ray. Expressions for ψ_1, ψ_2, and e in Equation 4.30.7 are related by expressions in the following equations:

$$\psi_3 = \tan^{-1}[2b/D\cos(\alpha/2)] \tag{4.30.8}$$

$$C = \frac{D\cos(\alpha/2)}{2\cos\psi_3} \tag{4.30.9}$$

$$e = \left[C^2 + D^2/4 - CD\cos(2\pi - \theta_2 + \psi_3)\right]^{1/2} \tag{4.30.10}$$

$$\psi_2 = \sin^{-1}[C\sin(2\pi - \theta_2 + \psi_3)/e] \tag{4.30.11}$$

$$\psi_1 = \theta_2 - \psi_2 - \pi \tag{4.30.12}$$

Calculation of Gas Radiation Properties

The composition (by volume) of hot gas for the rotary kiln used in this study is superheated steam (100% H_2O). The pressure inside the kiln is assumed to be 1 bar. Because water vapor is a radiatively participating medium, using the mean beam length theory [6], the emissivity and transmissivity of the superheated steam as a function of temperature are calculated.

Gas Convection

Brimacombe and Watkinson [11] have indicated that in a rotary kiln operating at solid temperatures up to 1100 K gas convection may be significant. In the present work, the gas flow is turbulent. The convective heat transfer coefficient is estimated using Dittus–Boelter correlation in each axial segment assuming isothermal wall and hydrodynamically and thermally fully developed flow.

Conduction Heat Transfer in the Refractory Wall

In the present study, only heat conduction in the radial direction of the wall is taken into account as the circumference and the length of the kiln are much greater than the thickness of the kiln.

The coordinate system is shown in Figure 4.30.4. Because the kiln thickness to diameter ratio is small, Cartesian coordinates are used. The energy equation for the wall is

$$U\frac{\partial T}{\partial y} = a_{\text{rf}}\frac{\partial^2 T}{\partial x^2} \tag{4.30.13}$$

where U is the circumferential velocity of the wall and a_{rf} is the thermal diffusivity of the refractory.

Elements exposed to the gas are heated by the radiation as well as convection. Heat transfer from the hot wall to the solid bed is determined from

$$q_j = h_{\text{cs}}A_{j\text{i}}(T_{j\text{i}} - T_{\text{s}}) \tag{4.30.14}$$

where q_j is the heat transfer rate, h_{cs} is the contact heat transfer coefficient between the refractory wall and the solid, $A_{j\text{i}}$ is the inner surface element area, $T_{j\text{i}}$ is the inner refractory wall surface temperature, and T_{s} is the solid temperature. Because, the heat transfer coefficient, h_{cs}, is not available, an estimated value is used.

The contact heat transfer coefficient from kiln wall to solid is estimated by Sass [4] as

$$h_{\text{cs}} = h_{\text{CONV}} + h_{\text{RAD}}$$

where

$$h_{\text{CONV}} = 118.266(\dot{m}_{\text{g}}/A)^{0.67}$$

$$h_{\text{RAD}} = \sigma f\varepsilon_{\text{s}}(T_{\text{wall}} + T_{\text{s}})(T_{\text{wall}}^2 + T_{\text{s}}^2)$$

Here, $\dot{m}_g$ is the mass flow rate of gas (kg/s), A is the cross-sectional area of kiln (m^2), f is the correction factor taken as 0.5, ε_s is the emissivity of solid mass, T_{wall} is the mean temperature of inner surface of rotary kiln (in K), and T_s is the temperature of solid mass (in K) and σ is the Stefan–Boltzmann constant having a value of 5.67×10^{-8} W/m^2 K^4. In the present problem, the contact heat transfer coefficient is estimated to be 150 W/m^2 K.

The boundary condition for the outer surface of the refractory wall is the heat loss by convection and radiation. In the present investigation, convective heat transfer coefficient, h_c, is taken as 10 W/m^2 K considering air is blowing over the kiln and using the relevant graphs of Suryanarayana et al. [12]. The radiation heat transfer coefficient is calculated from

$$h_r = \sigma\varepsilon_{rf}(T_{jo}+T_\infty)(T_{jo}^2+T_\infty^2) \tag{4.30.15}$$

where T_{jo} and T_∞ are the outer wall temperature and ambient temperature in kelvin, respectively.

The combined heat loss from the outer surface is written as

$$q_j = (h_c + h_r)A_{jo}(T_{jo}-T_\infty) \tag{4.30.16}$$

where A_{jo} is the area of the outer surface element.

False transient approach is used to solve Equation 4.30.13. The alternative direction implicit finite-difference scheme has been used to march ahead in time. Because the initial temperature distribution of the refractory wall is not known, an arbitrary temperature is assumed at all grid points in the refractory wall. The solution converges when there is no further change in the temperature at each grid point as $t \to \infty$. This temperature distribution represents the steady-state temperature distribution of wall.

Energy Transport in Solid

The solid is considered to be well mixed, and small rotational speed of the kiln gives rise to rolling bed. The gas is also assumed to be well mixed in each transversal section as the flow is highly turbulent. Thus, only the axial solid and gas temperatures are computed.

Mass and energy balances on an element of solid contained in an axial segment either in the first or in the third section of the kiln give the expression for $T_{s,z+\Delta z}$ (Equation 4.30.19). The same performed on an element of the wet solids contained in an axial segment in the second section of the kiln produces the expression for the rate of evaporation, $\dot{m}_v$ (Equation 4.30.20). In the second section, the temperature of the solids remains constant at the saturation temperature of water. The end of the second section is indicated where the cumulative $\dot{m}_v$ is equal to the total predetermined rate of evaporation of water.

Because there are five elements on the surface of the solid, total radiation heat transfer q_1 is

$$q_1 = \sum_{j=1}^{5} q_j \tag{4.30.17}$$

Similarly, the heat transfer, q_2, from the refractory inner wall, is given by

$$q_2 = \sum_{j=1}^{5} h_{cs}A_j(T_{ji}-T_s) \tag{4.30.18}$$

Sections 1 and 3 (Figure 4.30.10)

Energy in:

$$\dot{m}_s C_{ps}T_{s,z} + \dot{m}_{w,s}C_{pl}T_{s,z} + q_1 + q_2$$

Energy out:

$$\dot{m}_s C_{ps} T_{s,z+\Delta z} + \dot{m}_{w,s} C_{pl} T_{s,z+\Delta z}$$

From energy balance,

$$T_{s,z+\Delta z} = T_{s,z} + \frac{q_1 + q_2}{\dot{m}_s C_{ps} + \dot{m}_{w,s} C_{pl}} \tag{4.30.19}$$

Section 2 (Figure 4.30.11)
Energy in:

$$\dot{m}_s C_{ps} T_{s,z} + \dot{m}_{w,s,z} C_{pl} T_{s,z} + q_1 + q_2$$

Energy out:

$$\dot{m}_s C_{ps} T_{s,z+\Delta z} + \dot{m}_{w,s,z+\Delta z} C_{pl} T_{s,z+\Delta z} + \dot{m}_v (C_{pl} T_{s,z} + h_{fg})$$

For this section,

$T_{s,z} = T_{s,z+\Delta z} = 373$ K = Saturation temperature of water at 1 bar

From energy balance,

$$\dot{m}_v = \frac{q_1 + q_2}{h_{fg}} \tag{4.30.20}$$

Energy Transport in Gas

Mass and energy balances on an element of hot gas contained in an axial segment in the first or third section of the kiln give the expression for $T_{g,z+\Delta z}$ (Equation 4.30.21). The same performed on an element of the hot gas contained in an axial segment of the second section of the kiln produces another expression for $T_{g,z+\Delta z}$ (Equation 4.30.22).

Sections 1 and 3 (Figure 4.30.10)

Energy in:

$$\dot{m}_g C_{pg} T_{g,z+\Delta z} + \dot{m}_{w,g} C_{pv} T_{g,z+\Delta z}$$

Energy out:

$$\dot{m}_g C_{pg} T_{g,z} + \dot{m}_{w,g} C_{pv} T_{g,z} + q_r$$

where $q_r = \sum_{j=1}^{25} q_j$

It may be noted q_j that includes both radiation and convection.

From the mass balance,

$$\dot{m}_{w,s} = \dot{m}_{w,g}$$

From energy balance,

$$T_{g,z+\Delta z} = T_{g,z} + \frac{q_r}{\dot{m}_g C_{pg} + \dot{m}_{w,g} C_{pv}} \tag{4.30.21}$$

Section 2 (Figure 4.30.11)
Energy in:

$$\dot{m}_{\mathrm{g}} C_{pg} T_{\mathrm{g},z+\Delta z} + \dot{m}_{\mathrm{w,g},z+\Delta z} C_{pv} T_{\mathrm{g},z+\Delta z} + \dot{m}_{\mathrm{v}} C_{pv} T_{\mathrm{s},z}$$

Energy out:

$$\dot{m}_{\mathrm{g}} C_{pg} T_{\mathrm{g},z} + \dot{m}_{\mathrm{w,g},z} C_{pv} T_{\mathrm{g},z} + q_r$$

From mass balance,

$$\dot{m}_{\mathrm{w,g},z+\Delta z} = \dot{m}_{\mathrm{w,g},z} - \dot{m}_{\mathrm{v}}$$

From energy balance,

$$T_{\mathrm{g},z+\Delta z} = \frac{\dot{m}_{\mathrm{g}} C_{pg} + \dot{m}_{\mathrm{w,g},z} C_{pv}}{\dot{m}_{\mathrm{g}} C_{pg} + \dot{m}_{\mathrm{w,g},z} C_{pv} - \dot{m}_{\mathrm{v}} C_{pv}} T_{\mathrm{g},z} + \frac{q_r - \dot{m}_{\mathrm{v}} C_{pv} T_{\mathrm{s},z}}{\dot{m}_{\mathrm{g}} C_{pg} + \dot{m}_{\mathrm{w,g},z} C_{pv} - \dot{m}_{\mathrm{v}} C_{pv}} \tag{4.30.22}$$

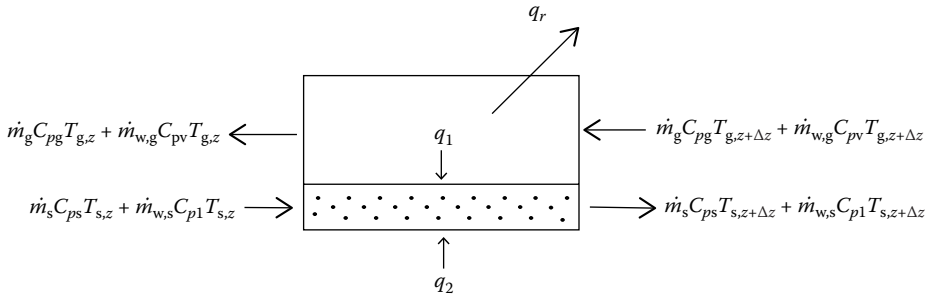


FIGURE 4.30.10 Energy transport in the solid and gas in an axial segment in the first and third sections of the kiln. (From Sinhal, K. et al., *ASME J. Therm. Sci. Eng. Appl.*, 4, 011009, 2012. With permission.)

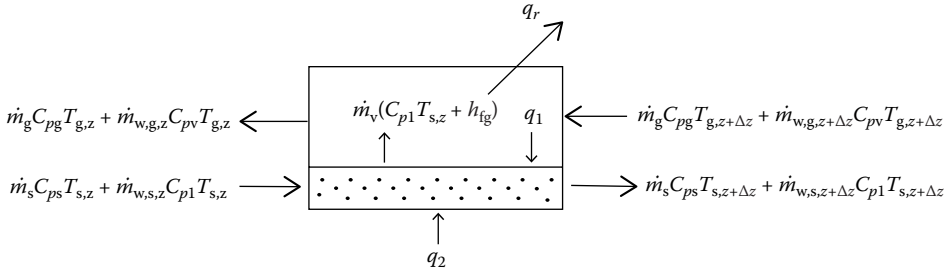


FIGURE 4.30.11 Energy transport in the solid and gas in an axial segment in the second section of the kiln. (From Sinhal, K. et al., *ASME J. Therm. Sci. Eng. Appl.*, 4, 011009, 2012. With permission.)

Basic Solution Methodology

The solution is initiated at the inlet of the kiln and proceeds to the exit. For solving the heat transfer problem, the kiln is divided into segments of equal length Δz, in the axial direction. Input data

required for the computer program are shown in Table 4.30.1 for superheated steam at 1 bar as drying medium. The property values of superheated steam are taken from Ghoshdastidar [8]). The solids properties have been calculated from the correlations given by Choi and Okos [13] based on the percent composition and temperature of apple and carrot [5]. The density and specific heat of apple at the inlet of the kiln (at 288.6 K) are 1437 kg/m^3 and 2080 J/kg K, respectively. The corresponding values for carrot are 1428 kg/m^3 and 2094 J/kg K, respectively. The source of kiln wall property data is the work of Ghoshdastidar et al. [7].

A mixture density relation is used to calculate the effective density of wet solids. The speed of rotation and the kiln inclination angle in Table 4.30.1 are typical of rotary dryers (van't land [14]).

Based on a grid independence test, the optimum number of grid points for apple drying is 25×11, the first number indicating number of grid points in the circumferential direction and the second in the radial direction. For carrot drying, the optimum number of grid points is 25×16. The axial segment length, Δz, is 1.3 m for apple drying and 1.2 m for carrot drying.

Results

Figure 4.30.12 shows axial solid and gas temperature distribution for apple and carrot for the input data in Table 4.30.1. The result shows no difference between the two. This may be due to little difference in density and specific heat. The predicted kiln length to dry and heat the apple pieces from 288.6 to 413 K is 93.6 m. The corresponding prescribed kiln length for carrot is 86.4 m. The inlet steam temperature required for both is 460 K. The higher kiln length for apple drying and preheating may be attributed to its higher density and hence, lower fill angle compared to carrot. Hence, there is less radiation heat transfer from the wall. The time to complete the drying and heating process is 29.8 min for apple and 27.5 min for carrot (drying and preheating time = kiln length/axial velocity, V_z of the solids).

Summary

This chapter presents a short overview of rotary kiln reactors widely used in chemical, metallurgical, and drying industry, emphasizing on the heat transfer processes occurring within it. The physics of transverse bed motion and mixing within the solid bed have been discussed. An expression for

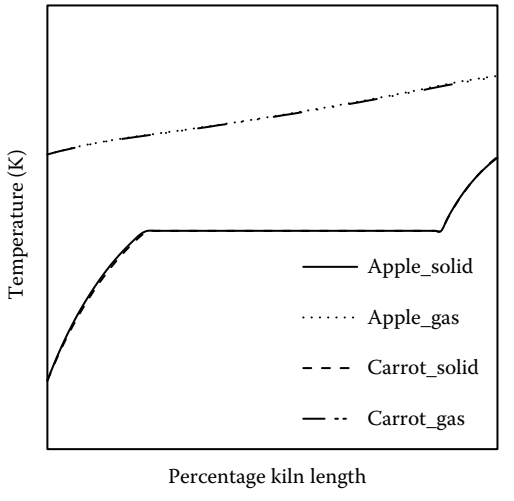


FIGURE 4.30.12 Axial solid and gas temperature distributions versus percent kiln length for apple and carrot drying based on the input data in Table 4.30.1. (From Sinhal, K. et al., *ASME J. Therm. Sci. Eng. Appl.*, 4, 011009, 2012. With permission.)

the axial velocity of solid particles has been given. Various approaches used to model rotary kilns have been introduced. Finally, a detailed numerical simulation of drying and preheating of food products such as apple and carrot with superheated steam in a rotary kiln has been presented in order to make the readers appreciate the enormous complexity and challenges in modeling rotary kilns. The main difficulties in the modeling of rotary kilns appear to be the estimation of contact heat transfer coefficient, modeling of gas convection, as well as integration of transverse solid bed motion in the overall heat transfer model.

ACKNOWLEDGMENTS

The author acknowledges with thanks the assistance received from Mr Ashish Agrawal, a doctoral student of the Department of Mechanical Engineering, Indian Institute of Technology Kanpur, Kanpur, Uttar Pradesh, India, in drawing Figures 4.30.1 and 4.30.4.

REFERENCES

1. Helmrich, H., and Schügerl, K. 1980. Rotary kiln reactors in chemical industry. *Ger. Chem. Eng.* 3: 194–202.
2. Boeteng, A. A. 2008. *Rotary Kilns*. Elsevier, Amsterdam, The Netherlands.
3. Ghoshdastidar, P. S., and Agarwal, A. 2009. Simulation and optimization of drying of wood chips with superheated steam in a rotary kiln. *ASME J. Therm. Sci. Eng. Appl.* 1: 024501.
4. Sass, A. 1967. Simulation of the heat-transfer phenomena in a rotary kiln. *Ind. Eng. Chem. Process Des. Dev.* 6 (4): 532–535.
5. Sinhal, K., Ghoshdastidar, P. S., and Dasgupta, B. 2012. Computer simulation of drying of food products with superheated steam in a rotary kiln. *ASME J. Therm. Sci. Eng. Appl*. 4: 011009.
6. Holman, J. P. 1997. *Heat Transfer.* 8th edition, McGraw-Hill, New York.
7. Ghoshdastidar, P. S., Bhargava, G., and Chhabra, R. P. 2002. Computer simulation of heat transfer during drying and preheating of wet iron ore in a rotary kiln. *Drying Technol*. 20 (1): 19–35.
8. Ghoshdastidar, P. S. 2004. *Heat Transfer.* Oxford University Press, New Delhi, India.
9. Jacob, M. 1957. *Heat Transfer.* 1st edition, John Wiley & Sons, New York, vol. 2, pp. 19–21.
10. Ghoshdastidar, P. S., Rhodes, C. A., and Orloff, D. I. 1985. Heat transfer in a rotary kiln during incineration of solid waste. *23rd ASME National Heat Transfer Conference*, Denver, CO, August 4–7, ASME Paper No.85-HT-86.
11. Brimacombe, J. K., and Watkinson, A. P. 1978. Heat transfer in a direct-fired rotary kiln: I. Pilot plant and experimentation. *Metall. Trans. B*. 9: 201–208.
12. Suryanarayana, N. V., Lyon, J. E., and Kim, N. M. 1986. Heat shield for high temperature kiln. *Ind. Eng. Chem. Process Des. Dev.* 25: 843–849.
13. Choi, Y., and Okos, M. R. 1986. Effects of temperature and composition on the thermal properties of foods. In *Food Engineering and Process Applications*, Le. M. Maguer and P. Jelen, eds., Elsevier, New York, vol. 1, pp. 93–101.
14. van't Land, C. M. 1991. *Industrial Drying Equipment*. Marcel Dekker Inc., New York.

4.31 HEAT TRANSFER IN GLASS MANUFACTURING PROCESSES

Naveen Tiwari and Rajappa Tadepalli

Introduction

Glass is one of the oldest man-made materials still in use without interruption since the recorded history. The applications of glass are widespread in today's day-to-day life starting from window glass at home or in cars, glass bottles and other glasswares, and electric components. The modern technologies demand enormous supply of glass, for example, radiation shield for nuclear technologies, glass fibers for communication, etc. Glassmaking is an integrated manufacturing process where raw minerals are transformed into a variety of finished products within a single manufacturing setup. It is an energy intensive process as it involves high temperatures at various stages of the process. This chapter discusses the concepts of heat transfer involved during the steps of glass production.

The fundamental operations to form useful glass objects appear very simple and have been practiced for more than 3000 years. With industrialization and increased consumption of glass, the manufacturing process has been streamlined. Figure 4.31.1 shows a typical process flow sheet of manufacturing a glass product in current practice. As indicated the raw materials are crushed and put through high-temperature processes. Melting of the raw materials takes place at around 1500°C. The subsequent process is to cool the melt to form a solid. When a melted substance is cooled below the melting point, crystallization is expected to occur. In glass, the cross-linkage between tetrahedron SiO_4 structures needs to be broken in order to form crystal nuclei which can only happen at low temperatures. However, at low temperatures, the high viscosity of the melt impedes the restructuring of new molecules, thus preventing the crystal growth. In general, the tendency to crystallize (or "devitrification" in the context of glass) decreases with the increasing rate of cooling and with the number of components in the glass batch. The devitrification is only desirable in glass-ceramics which is not discussed here. The "glassy" state is in continuation of the liquid state with no discontinuities in physical properties and structural arrangements. The glasses are therefore sometimes referred to as "supercooled liquids."

Composition and Its Effect on Properties

The majority of the glasses in use today are silicate based in which chief component is SiO_2, and other components are added to impart various desired characteristics to the glass. For example, lime, CaO, and small amounts of alumina (Al_2O_3) and Magnesia (MgO) improve the chemical durability of the glass. Some metal oxides can also be added to incorporate color in the glass product. For glass products that are desired to be photosensitive oxides of metals such as gold, silver, or copper are introduced into the glass composition in addition to other optical sensitizers and thermal-reducing agents. Glass products for which optical design is an essential requirement for the application better transparency, control over refractive index and freedom from defects such as bubbles and stones is necessary. Thus the glass must have a good degree of homogeneity, and the annealing should be done under conditions that will reduce internal strains in order to make an optically stable glass. Further, during the melting process components such as soda, boric oxide, and lead oxide evaporate to a certain extent from the glass melt. The change in the composition which is caused in this way must be compensated for by corrections to the raw feed if the physical properties of the glass are to be maintained. Changes in composition caused by solution of furnace envelope (refractory) must also be corrected.

In addition to change the desired properties of the glass product, they also affect the physical properties of the glass melt that affect its processing during manufacturing. Table 4.31.1 shows the effect of the glass composition on the liquidus temperature, the temperature at which crystallization

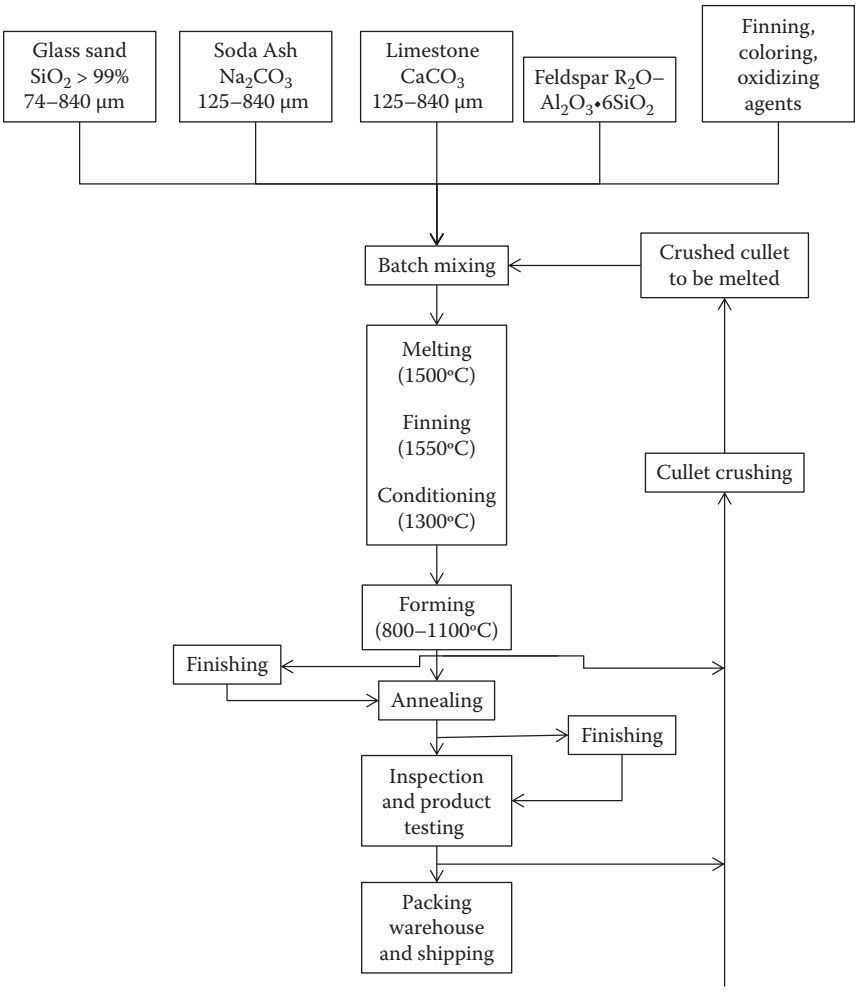


FIGURE 4.31.1 General flow of glass manufacturing. (Redrawn from Rawson, H., *Phys. Technol.*, 5, 91–114, 1974.)

may start happening provided the cooling rate is slow enough for that to initiate. It is worthy to note that the liquidus temperature can be depressed from 1710°C for pure silica to about 800°C in the presence of other metal oxides.

An important property to consider while manufacturing process is the viscosity of the melt as it continuously varies from the molten to the lowest temperature at which desired structural changes can be made. Figure 4.31.2 shows the viscosity (μ) as a function of temperature for glasses with different compositions. These curves are typically fit by the following expression:

$$\log(\mu) = -A + \frac{B}{T - T_0}, \tag{4.31.1}$$

where the constants A, B, and T_0 are determined by curve fitting for the respective glass melt. The viscosity data as a function of temperature for different compositions is important in order to assess the balance between the desired product properties versus the challenges during melting, fabrication, and annealing operations.

Similarly specific heat capacity at constant pressure C_p (J/mol/K) is a function of temperature given by Equation 4.31.2 where the coefficients are given for the respective components. Specific

TABLE 4.31.1
Effect of Composition on Liquidus Temperature

	SiO_2	Na_2O	K_2O	CaO	MgO	B_2O_3	Al_2O_3	Liquidus Temperature (°C)
Silica glass	>99.5							1710
Soda lime	73.6	16	0.6	5.2	3.6		1.5–2.5	860
Aluminosilicate	57	1		5.5	12	4	20.5	1130

Source: Shand, E.B., *Glass Engineering Handbook*, McGraw-Hill Book Company Inc., New York, 1958.

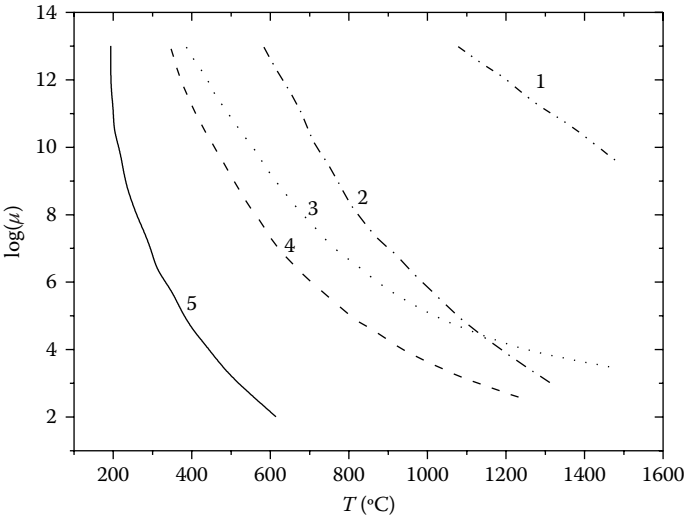


FIGURE 4.31.2 Viscosity (μ, dPa) of glass melt as a function of temperature. (1) Fused silica, (2) alumina-silicate, (3) borosilicate glass, (4) soda-lime glass, and (5) lead borate solder glass. (Redrawn from Pfaender, H.G., *Schott Guide to Glass*, Chapman and Hall, London, 1996.)

TABLE 4.31.2
Coefficients of Partial Molar Heat Capacities of Oxides in Silicate Glasses

	a_i	$10^3\ b_i$	$10^{-5}\ c_i$	d_i	Temperature Interval (K)	Range of Mole Fractions
SiO_2	127.20	–10.78	4.31	–1463.9	270–1600	0.33–1
Na_2O	70.88	26.11	–3.58	0	270–1170	0–0.33
K_2O	84.32	0.73	–8.30	0	270–1190	0–0.17
CaO	39.16	18.65	–1.52	0	270–1130	0–0.5
MgO	46.70	11.22	–13.28	0	270–1080	0–0.5
Al_2O_3	175.49	–5.84	–13.47	–1370	270–1190	0–0.33
TiO_2	64.11	22.59	–23.02	0	300–800	0–0.17
FeO	31.77	38.52	–0.012	0	300–800	0–0.4
Fe_2O_3	135.35	12.31	–39.10	0	300–800	0–0.13
H_2O	81.34	0.098	–31.09	0	298–423	0.18

Source: Richet, P., *Chem. Geol.*, 62, 111–124, 1987.

heat is an intensive property and thus the net C_p of the mixture can be obtained by summating the C_p values of all the components weighted by their respective mole fractions. The values of the coefficients and their validity range are given as follows:

$$C_{pi} = a_i + b_i T + c_i / T^2 + d_i + T^{1/2}. \tag{4.31.2}$$

In the next sections, we present some details of glass manufacturing process focusing on the heat transfer aspects of it.

Melting Process

Glass is mostly melted in a continuous process in large furnaces made up of refractory blocks. The furnaces are heated from above by the flames created by burning a fuel. There are smaller versions of the furnace called pots and batch or day tanks that are used for manufacturing small amounts of glass, particularly for special composition and properties. The holding capacity of batch-type units is of a few tons, while that for the continuous tanks or furnaces ranges from about 1 to 1500 tons. The outputs from such large continuous furnaces go up to several tons of glass per day. Typically natural gas, fuel oil, and raw producer gas are the economical fuels used in the industry. In last few years, the use of electrical energy in the glass melting process as a supplement to conventional heating means is increasing. The benefits that it provides are ease of application and better control though electrical heating costs are higher.

Figure 4.31.3 shows a schematic for the working of a continuous tank for melting glass. The tanks are generally rectangular in shape. As shown in the figure, the compartment into which the batch or raw material is introduced is the "melting end," and is the bigger compartment, while the smaller compartment is called the "working end" where the glass is cooled and distributed for use. These two compartments are separated by a permanent "bridge" wall or by floating refractory baffles. The lower part of the furnace called "bath," where the melted glass is present, is constructed of a refractory block. The role of the refractory block is to minimize the heat losses due to conduction and radiation. Also, the material should be inert to the chemicals involved in glass manufacturing, and should be able to withstand the high temperatures involved. In order to meet these requirements, the refractories may be made of fireclay with addition of burned flint clay, or the typical refractory aluminous compositions. Above the walls of the furnace, combustion takes place in a crown that also covers the furnace from the top. The crown consists of an arched roof made up of silica brick resting upon side walls of silica, or fusion cast refractory. Batch is introduced mechanically by either a screw or by a pusher bar near the back of the melting end. The "doghouse," a slanted projection, is provided near the entrance to assist in introducing the batch into the furnace. In the melting end of the furnace, this batch is heated until the fluxes melt, and the sand and other ingredients dissolve.

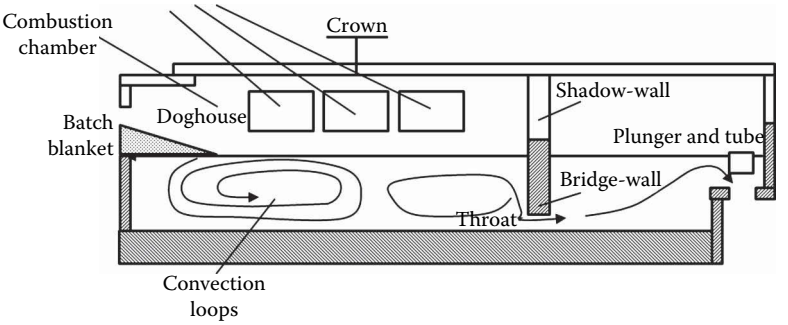


FIGURE 4.31.3 Schematic showing the melting and working compartments of the furnace. The two compartments are divided by the bridge wall.

Due to the gases such as Carbon dioxide and water vapors released during the reactions in the melt as well as air trapped in the voids of the batch, several tiny bubbles are produced in the melt due to which the molten glass is initially a foamy mass. Most of the gas bubbles are eliminated at the high temperature of the melting end. While the molten glass passes through a hole, or throat, in the bridge wall the remaining foam or scum is skimmed off from the surface of the molten glass, which is then cooled and distributed by the working-end compartment. Due to the heating of the top by the flames, cooling by side walls and addition of cold batch thermal gradients develop inside the melt. These thermal gradients lead to the formation of natural convection currents that help in making the melt homogeneous. In the work of Curlet et al. (1984), the Grashof number (ratio of buoyancy force to viscous force) is of the order of 10^5 indicating significant natural convection currents. The homogenization inside the furnace thus prevents the partially melted glass from entering into the work compartment. At such elevated temperatures, the solution of the refractory blocks may form in the glass melt leading to undissolved refractory stones, inhomogeneous glass, and bubbles. With the lower-viscosity glasses, such difficulties may limit the melting temperature used to bring the batch into solution.

When melting of sand and other batch components is completed along with the elimination of bubbles, the molten glass is cooled until its viscosity has increased to the value suitable for fabrication. In continuous furnaces, the glass melt is first cooled in the working chamber, which is maintained at a temperature between 100°C and 200°C lower than that in the melting chamber. Subsequently, the molten glass flows from the working compartment to the forming equipment through the refractory channels called "forehearths." The final cooling to desired temperature takes place at these forehearths. The uniformity in temperature is ensured such that the glass exhibits a uniform viscosity in the forming operation. Mechanical stirring by using refractory screws is one method to make the temperature homogeneous as natural convection is not very strong in this region. The forehearths' construction varies greatly and depends upon the machine to be supplied with the glass melt.

Figure 4.31.4 shows the melting chamber of the furnace where batch blanket is being introduced at the top of the melt. The granular batch mixture is usually dispensed on the surface of the melt; a picture of the real process is shown in Figure 4.31.5. Due to the lower density of the mixture, it floats on the surface of the melt until it gets transformed into glass melt. Heat is transferred from the combustion space, where fuels are being burnt, to the top of the batch blanket primarily by radiation. The amount of heat flux from the combustion space to the top of the batch blanket and to the surface of the glass melt depends on the position in the furnace. Radiation heat from the combustion space consists of wavelengths between 600 and 4500 nm and the heat flux is typically of the order of 300–500 kW/m^2. The radiation from the flames is a function of flame thickness, flame emissivity, and flame temperature. The crown also participates in this radiation heat transfer by emission and reflection of flame radiation. The highest radiation heat flux is observed

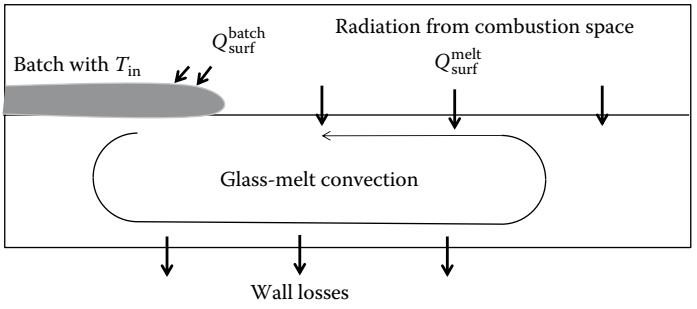


FIGURE 4.31.4 Heat flux from combustion space to glass and batch surface. (Redrawn from Krause, H., Krause, D., *Mathematical Simulation in Glass Technology*, Springer, Berlin and Heidelberg, 2002.)

FIGURE 4.31.5 The batch being introduced on top of the glass melt in the melting chamber of furnace. (Courtesy of Saint-Gobain Glass.)

at the locations with the highest flame temperatures, and to the relatively cold batch-blanket surface. The glass melt flowing underneath this blanket also transfers heat to the batch by radiation and conduction from the bottom. It is to be noted that the unreacted batch has a high reflectivity. Furthermore, the batch is a poor conductor of heat and thus develops high-temperature gradients upon heating. High heat flux densities are therefore realized due to high-temperature gradients at the start of heating. As the hot dissociation gases are released at this hot batch-bottom side the inner parts of the batch blanket are also heated up due to the rising gases. The heat penetration into the batch material thus depends on several factors such as the thermal conductivity of the batch, the nature of the reactions (endothermic or exothermic), and the nature of the released gases. The batch blanket undergoes phase change during these thermal processes. Due to the chemical and thermal processes, the composition and density are changed. The batch-to-melt conversion consists of a series of sequential and parallel chemical reactions and the conversion rate is mass transfer and heat transfer limited. It is fair to say then that the degree of batch-to-melt conversion is dependent on the local temperature T_b of the melt and also on the thermal history of the glass batch. As the heating, reactions and melting proceeds the batch parameters, such as porosity, density, conductivity, and reactivity, also change with time. The energy balance for the batch blanket can be given by

$$\frac{\partial(\rho_b C_{p,b} T_b)}{\partial t} + \nabla \cdot (\rho_b v_b C_{p,b}) = \nabla[k_b \nabla T_b] + S(T,t) + hA_f(T_{gas} - T_b), \tag{4.31.3}$$

where ρ_b is the density of the glass batch (kg/m^3); $C_{p,b}$ is the heat capacity of the glass batch (J/kg/K) that will depend on its constituents and can be computed using Table 4.31.2 (conversion needs to be done from molar to mass basis); k_b is the temperature-dependent effective batch heat conductivity (W/m/K) that depends on the constituents, porosity, and gas convection; S is the net heat sink rate due to all the chemical reactions involved (J/m^3/s), h is the heat transfer coefficient between the batch and enclosed gases (W/m^2/K); A_f is the specific heat transfer area between the batch and enclosed gases (m^2/m^3); T_{gas} is the temperature of the gases enclosed in the batch (K); T_b is the temperature of the glass batch (K); and t is time (s). The last term in Equation 4.31.3 is often neglected and the effect of released gases on the heat transfer is taken into account by defining an effective thermal conductivity $k_b(T,t)$, which can be determined by experiments. For the batch blanket, the convection terms are usually omitted from the left-hand side assuming a small contribution to energy transport due to bulk motion. The term S as a function of time and temperature is derived from laboratory tests using the complete batch. It should be mentioned that the effective batch heat conductivity k_b also takes into account the fraction of the liquid phase present in the partly molten glass batch as it moves into the furnace. In the work of Faber et al. (1991), experimental technique to measure the effective thermal conductivity and the challenges are presented.

The energy distribution in the glass melt is also governed by Equation 4.31.3, where the properties such as density and specific heat capacity are that for the melt. Further, the contribution from the convection cannot be neglected in the glass melt. The source term in general also includes the contribution due to radiation in the glass which is a semitransparent medium for the radiative heat transfer. This mode of transport in semitransparent media such as glass melts generally complicates the analysis. The contribution due to the radiation transport is typically modeled by the radiative transfer equation (RTE) for participating medium, in which the change in radiation intensity i_λ, over a distance δ is given by

$$\frac{\partial i_\lambda}{\partial \delta} = -a_\lambda i_\lambda + a_\lambda i_{\lambda b} - \sigma_{s\lambda} i_\lambda + \frac{\sigma_{s\lambda}}{4\pi} \int_{\omega_i}^{4\pi} i_\lambda(\delta,\omega)\Phi(\lambda,\omega,\omega_i)\,d\omega_i. \tag{4.31.4}$$

The right-hand side terms are loss by absorption, gain by emission, loss by scattering, and, finally, gain by scattering, respectively. In the Equation, $a\lambda$ is the absorption coefficient of the medium, $\sigma_{s\lambda}$ is the scattering coefficient, ω is the solid angle, and Φ is the scattering distribution function that depends on the both incident and scattering direction.

Note that all coefficients appearing in this equation depend on the wavelength of the radiation, and that has a strong implication for analyzing the radiation transport in glass melts, for example, the absorption coefficient for glass is shown in Figure 4.31.6 as a function of wavelength and temperature. Several integrodifferential equations need to be solved for radiation transport due to contributions from several wavelengths which need advanced numerical algorithms. Also as expected, before computations can be carried out the radiation characteristics of the glass melt as a function of wavelength must be accurately known. The presence of sand particles and bubbles may even further complicate the analysis. The boundary conditions need to be well-defined for reflection, refraction, and absorption at primary melt to reacting batch, primary melt to melt in the tank volume, and primary melt to combustion space. More details on radiation modeling in semitransparent media can be found in Frank and Klar (2010). In order to avoid such complexities associated with the analysis of radiation, the Rosseland approach is frequently used (another approach is to use a two-band model for absorption coefficient as mentioned in Roychowdhury and Srinivasan (1994)). In this approach, an effective thermal conductivity coefficient is defined by taking into account the contribution of radiation as

$$\lambda_{eff} = \lambda_{conduction} + \lambda_{radiation}, \tag{4.31.5}$$

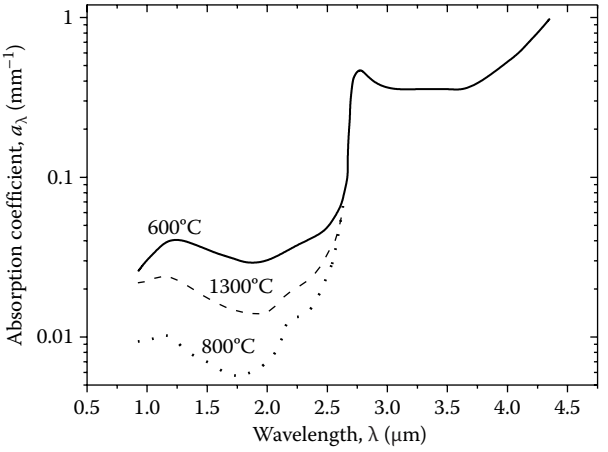


FIGURE 4.31.6 Absorption coefficient for window glass as a function of wavelength of the radiation. (Redrawn from Rawson, H., *Phys. Technol.*, 5, 91–114, 1974.)

where

$$\lambda_{\text{radiation}} = \frac{16}{3}\frac{n^2\sigma_{\text{SB}}T^3}{a_{\text{Ross}}}. \tag{4.31.6}$$

This apparent conductivity is then used in the Fourier equation of heat conduction. In the above equations, n is the refractive index, σ_{SB} is the Stefan–Boltzmann constant, and a_{Ross} is a gray absorption coefficient that can be computed from the spectral absorption of the glass melt. The effective thermal diffusivity for soda-lime silicate glass melt was experimentally measured by Van Zee and Babcock (1951), as function of temperature, and can be converted to give the effective thermal conductivity (see Pilon et al., 2006) as given in Equation 4.31.7

$$\lambda_{\text{eff}} = 213 - 0.37T + 1.66\times10^{-4}T^2, \tag{4.31.7}$$

where T is the temperature in K and the expression is valid above 1000 K (the melt will exist above this temperature anyway). The typical wavelength range of radiation in glass melting is 0.5–5 μm. Furthermore, the sand-grain diameters are smaller than 500 μm, gas bubbles range from a few microns to millimeters, and the cullet pieces are in the range of millimeter to centimeter. Due to such large sizes as compared to the radiation wavelength, diffraction and refraction need to be taken into account. Significant scattering due to bubbles and sand particles causes the melt to behave as an opaque medium, and the above analysis can be further simplified.

Fining and Bubble Dynamics

The glass-melt furnace is like a large chemical reactor where in the very first stage a large number of chemical reactions occur at different temperatures and in different phases. This stage occurs primarily in or around the batch blanket. The glass melt is left with incompletely dissolved sand grains or grains of other refractory oxides after the reactions are over in addition to the dissolved gases, and gas bubbles. In the next stage of the melting process, the residual grains are completely dissolved in the melt. Stirring of the melt is desired at high temperatures to enhance the process of dissolution. The freshly molten glass contains dissolved gases, smaller and large bubbles, and dissolved polyvalent ions. At high temperatures, above a certain fining onset temperature level, the polyvalent ions get converted to lower redox state and gases such as oxygen or SO_2 are generated in the melt. These fining gases, owing to their low solubility, diffuse toward the bubbles present in the melt. Inside the bubbles, the original gas is diluted by the fining gases. Fining gases and other dissolved gases start to diffuse into the bubbles, driven by the difference in the equilibrium pressure of these gases in the melt and the partial pressure in the bubbles. The bubbles thus growing in size ascend to the glass-melt surface with a velocity proportional to the square of the diameter and proportional to the reciprocal value of the glass-melt viscosity, in accordance with the Stokes law. The bubble-removal and gas stripping process from the melt is called fining. It requires temperatures above the fining onset temperature and almost no mixing in the melt to avoid the dispersion of the ascending bubbles throughout the whole melt. In the final stages of fining (refining), slow cooling of the melt occurs. Now, the residual tiny bubbles or the bubbles that have grown in size can completely dissolve in the melt. This dissolution (also called secondary fining) requires an increasing solubility as the melt cools down. Bubbles containing gases dissolving chemically in the melt (e.g., O_2, CO_2, SO_2) can be reabsorbed because the solubility of these gases in the melt will increase as temperature reduces. The dissolution rates of these gases and bubbles also depend on the diffusion coefficients of the gases in the melt, and as expected these values decrease with temperature reduction. This implies that the dissolution of bubbles mainly occurs in a certain temperature range. Below these temperatures, the dissolution rates become too small and the bubbles "freeze-in" in the rigid glass.

After this stage, the melt is conditioned to the viscosity level required for forming of the glass product. The viscosity should be uniform prior to forming in order to obtain a uniform glass distribution during blowing, pressing, drawing, or floating of the glass. Proper control of the glass-melt viscosity and the uniformity of the viscosity in the feeder are essential prerequisites for an optimum forming process. In the next section, description of the thermal processes and their importance is described in one of the forming processes called the float process.

Heat Treatment for Glass Transformation

Heat-treatable float glasses can be broadly classified into two categories, namely, annealed float glasses (clear or tinted glass) and coated glasses (low emissivity, solar control, and enameled/ceramic-frit printed). Before the heat treatment of glass, preprocessing steps such as cutting the glass into desired size, edge grinding, and washing are essential. The main heat treatment process depends on the amount of residual stress imparted onto the glass and can be classified as

1. Thermal tempering
   a. Convection tempering
   b. Bend tempering

2. Heat strengthening
   Typically heat strengthening increases the strength of annealed glass by a factor of two and tempering imparts a 4× increase in strength of the annealed glass. In addition, chemical tempering by ion exchange method, i.e., exchange of K (larger) ion with Na (smaller) ion is a possible option to impart higher strength to the glass. However, it is a slow batch process and is not suited for large-scale industrial applications.

The processing flow diagram for heat treatment of various types of glasses is shown in Figure 4.31.7.

A comparison of glass performance after heat treatment with various processes is shown in Table 4.31.3.

For reference, the residual stress profile comparison of glasses after different heat-treatment processes is shown in Figure 4.31.8.

Thermal tempering, being the most widely used process for toughening the glass, is covered in some detail in this chapter.

A typical tempering furnace model is shown in Figure 4.31.9 with different sections.

Glass is heated up to 600°C–700°C, depending on the glass thickness and is soaked at the tempering temperature for approximately 40 s/mm of thickness (for clear glass), followed by quenching with air jets. The nonlinear thermal expansion of the glass creates a gradient in density of the glass across thickness due to the difference in thermal gradient during cooling.

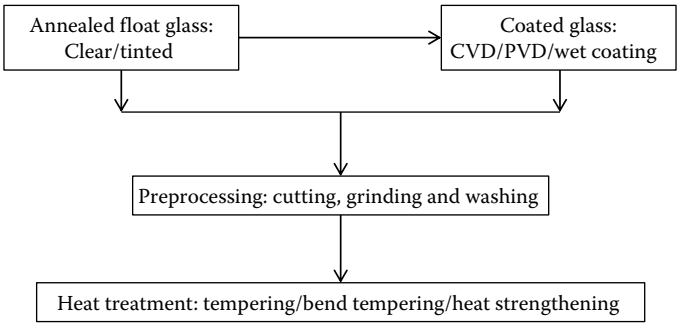
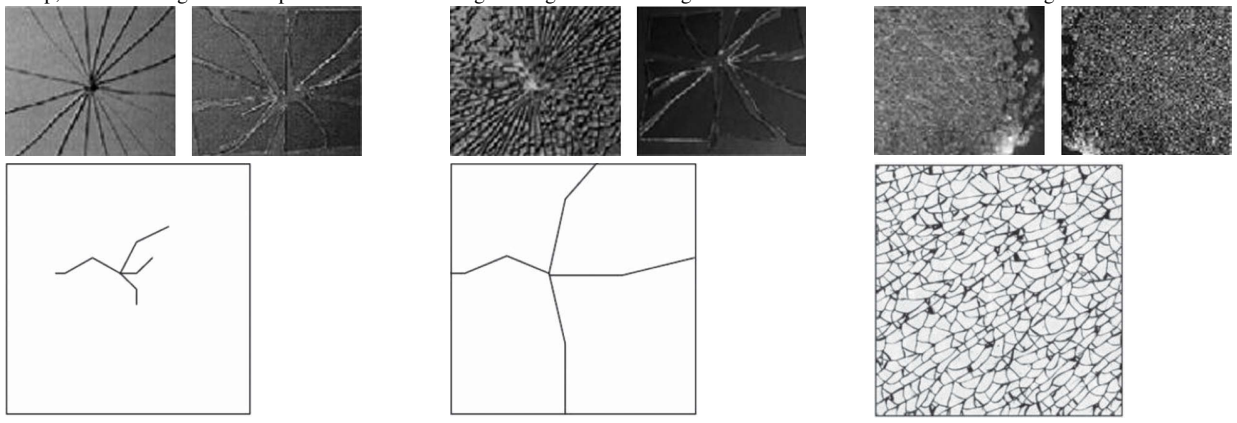


FIGURE 4.31.7 Glass processing heat-treatment flow.

TABLE 4.31.3
Comparison of Glass Performance vs. Various Heat-Treatment Processes

Parameter	Annealed Glass	Heat-Strengthened Glass	Tempered/Safety Glass
Glass strength	Bend strength: 30–40 MPa Compression strength: 1800–2200 MPa Tensile strength: 20–70 MPa Brittle and very strong in compression, Low toughness	Moderate bend strength of ~50 MPa	Very high bend strength ~120 MPa
Resistance to thermal shock	Low thermal shock resistance.	Moderate resistance to thermal shock ~100°C	Higher resistance to thermal shock ~ 200°C
Residual stress	Evenly distributed low residual stress ~15–20 MPa	Residual stress 47–52 MPa	Residual stress 78–92 MPa
Breakage pattern	Sharp, knife-like fragmentation pattern	Edge-to-edge and some fragments	Numerous small fragments



Source: Glass Academy-Premier Glass Portal. http://www.glass-academy.com/elearning/.

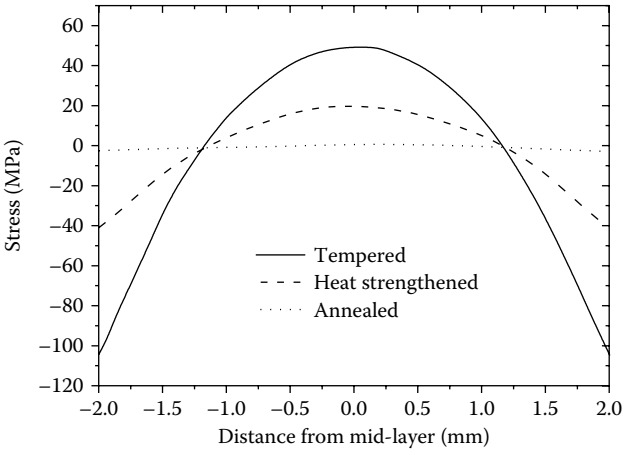


FIGURE 4.31.8 Typical thickness-wise residual stress profile of annealed, heat-strengthened and tempered glasses. (From Rantala, M. Heat transfer phenomena in float glass heat treatment processes. PhD Thesis, Tampere University of Technology, Finland, 2015. With permission.)

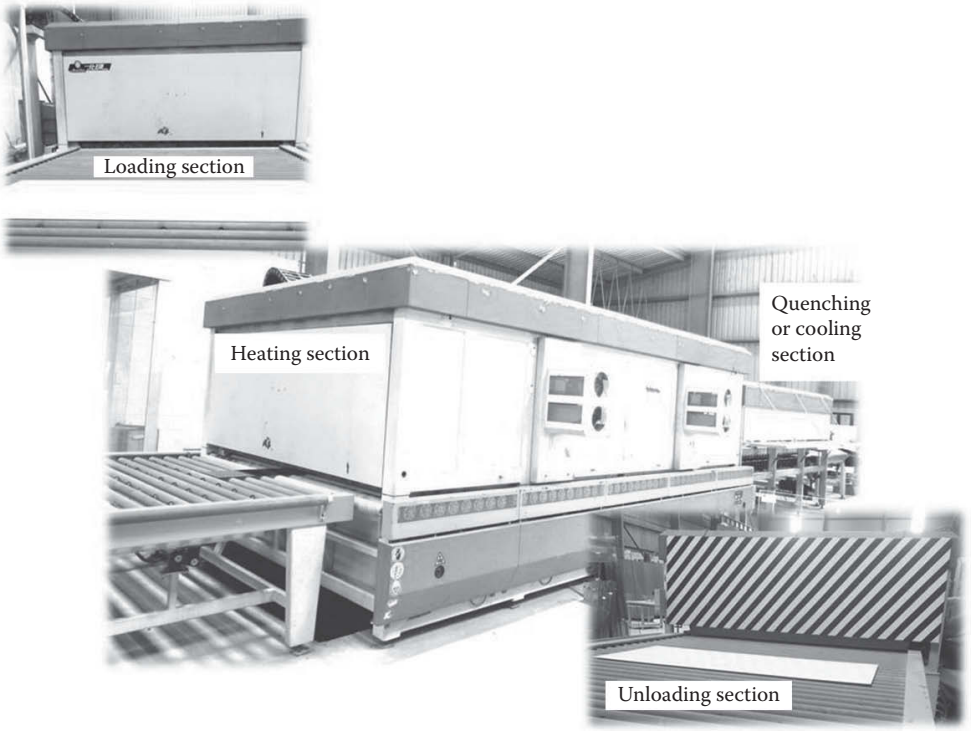


FIGURE 4.31.9 Model depicting different zones of tempering furnace. (From Glass-academy.in. With permission.)

Below T_g thermal expansion of glass is practically three times lower than that above T_g, as shown in Figure 4.31.10.

The stress during tempering develops when the glass is cooled from 670°C to 470°C and the glass is not viscous enough to move freely. When cooled rapidly, the surface experiences a higher

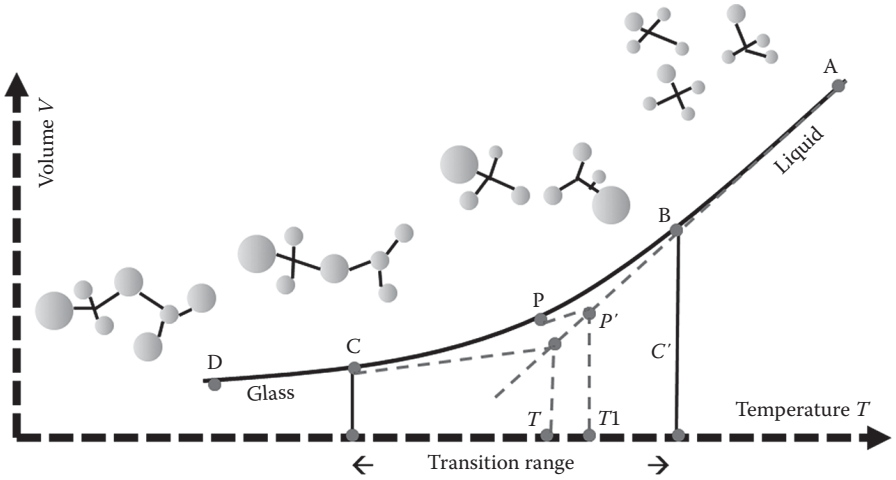


FIGURE 4.31.10 Graphical representation of thermal expansion coefficient of glass vs. temperature. (From Glass-academy.in. With permission.)

thermal gradient than the mid plane of the glass due to low-thermal conductivity of the glass. This differential rate of cooling induces a density gradient through the thickness of the glass. Rapid quenching solidifies the surface of the glass and the molecules are not free to move. But the molecules in the bulk are free to move and hence rearrange themselves resulting in a density increase.

The following modes of heating are typically used during thermal tempering:

1. Radiation
   a. Open coil heating element
   b. Framed heating element

2. Convection
   a. Combination radiation/forced convection
   b. High convection
   c. Dual chamber system with convection preheating

Details of these modes of thermal tempering are discussed now.

Radiation Tempering

Heating of glass takes place by means of radiation heat transfer from heating coils located above and below the glass as shown in Figure 4.31.11. Upon heating, the ceramic rollers get heated up and aid in heating the bottom of the glass through conductive heat transfer between roller and glass. The heating

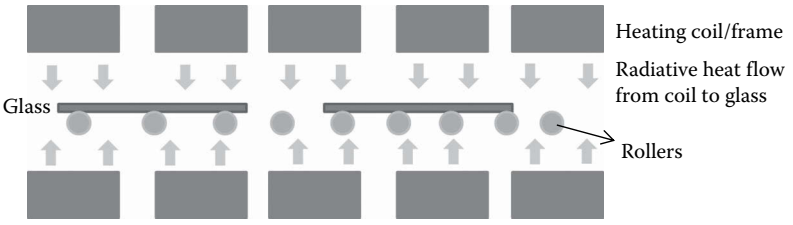


FIGURE 4.31.11 Schematic diagram of radiation tempering furnace heating setup. (From Glass-academy.in. With permission.)

elements can be an open coil arrangement or framed heating element. In open coil arrangement, the heating is not focused and the glass is kept away (20–100 cm) from the coil to avoid overheating. Framed heating elements, on the other hand, have focused heating and therefore, lower temperature is used compared to the open coil arrangement. The main drawback of these types of furnaces is that they cannot be used to heat-treat low-emissivity coated glass due to selective reflection of Infrared radiation.

Convection Tempering

In this model of tempering, the heating takes place by three modes:

1. Radiation heat transfer from heater, rollers, other surfaces and glass itself
2. Convective heat transfer to glass by hot air inside the furnace
3. Conductive heat transfer between ceramic roller and glass

The advantage of these types of furnaces over radiation furnace is that they can be used to heat-treat low-emissivity and other sensitive glasses due to relatively lower intensity and more uniform heating profile.

Various techniques for heat treatment can be deployed for convection tempering and a few examples are detailed below.

a. Combination of radiation and forced convection

The top heating chamber of the furnace is equipped with convection air jets. Heat transfer between top and bottom is maintained the same resulting in better optical quality of heat-treated glass as shown in Figure 4.31.12.

b. High convection

Heating of the air is achieved by gas or electrical heating elements and the hot air is blown against the glass by means of blowers and air nozzle arrangement thus eliminating the effect of emissivity of the glass as shown in Figure 4.31.13. Heating of glass by radiation is reduced and the peak temperature is achieved faster.

c. Dual chamber system with convection preheating

In this case, the first chamber is equipped with high convection preheating followed by forced convection in the second chamber as shown in Figure 4.31.14. This type of two-stage heating will eliminate glass bending during tempering.

The main parameters influencing thermal tempering are heating time, temperature inside the furnace, glass transfer speed, oscillation speed, distance of nozzles in chiller and percentage of nozzle opening or air pressure inside the furnace. Heating time depends on the glass thickness, also it varies according to the furnace type. Heating time for convection furnace is ~40–45 s/mm thickness of glass, and is ~15% higher for radiation tempering compared to convection furnace. The tempering temperature depends on the tint of the glass, since the heat absorption varies with the tint of the glass. Figure 4.31.15 shows the comparison of tempering temperature vs. tint of glass.

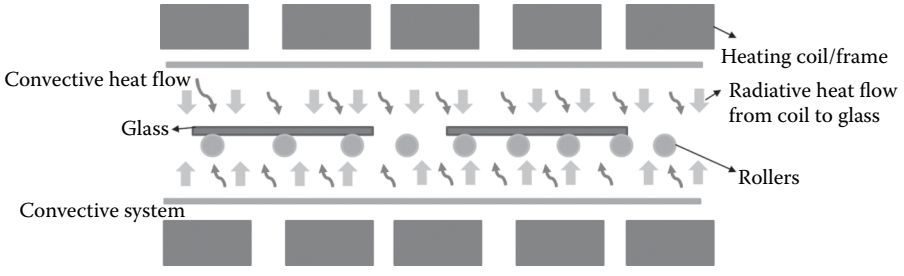


FIGURE 4.31.12 Schematic representation of combination of radiation/forced convection system. (From Glass-academy.in. With permission.)

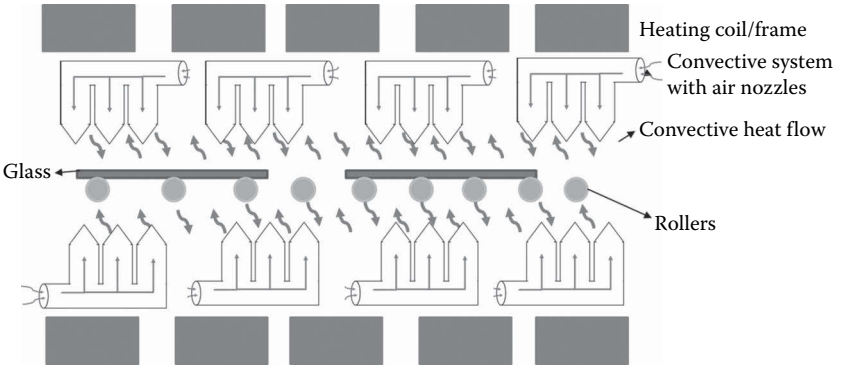


FIGURE 4.31.13 Schematic representation of highly convective system with air nozzles. (From Glass-academy.in. With permission.)

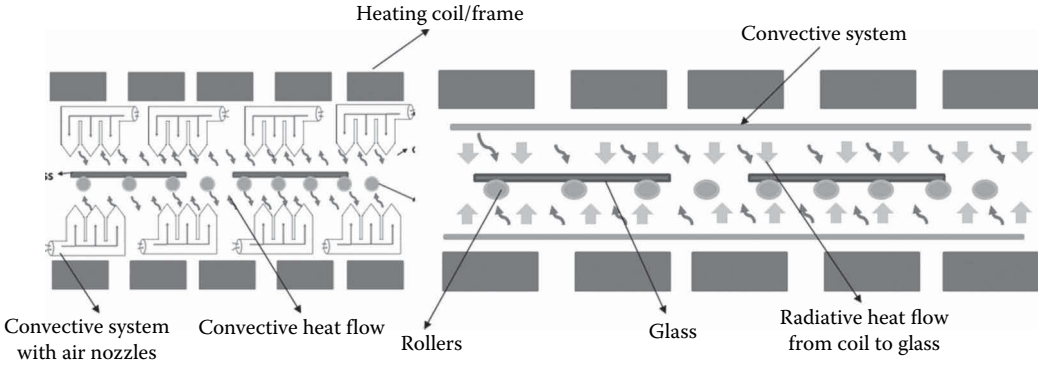


FIGURE 4.31.14 Schematic representation of dual chamber system. (From Glass-academy.in. With permission.)

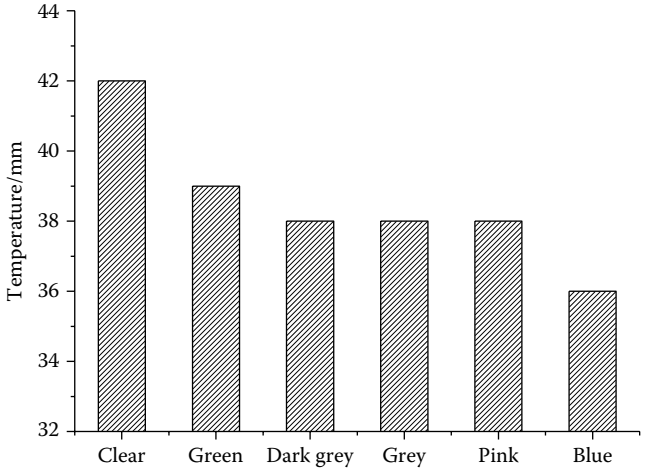


FIGURE 4.31.15 Heating requirement per mm of glass vs. tint of glass. (From Glass-academy.in. With permission.)

Heat Strengthening

Heat strengthening is almost the same process as thermal tempering except that cooling is much slower than that in the regular tempering. Thus the residual stress developed is half of that generated during thermal tempering (~45 MPa). The fragmentation structure is almost the same as that of annealed glass and the strength is twice that of the annealed glass. Heat strengthened glass is used when optical distortion requirement is very stringent, owing to its flatter surface compared to tempered glass. The residual stress comparison of heat strengthened and tempered glass is shown in Figure 4.31.16.

Heat Transfer during Tempering

During tempering, the glass is exposed to radiation emitted by radiant heaters, rollers and other surfaces inside the furnace. Also, the glass is emitting radiation, and the emission increases when glass gets hotter. Convection heat transfer between hot air and glass affects both surfaces. In addition, conductive heat transfer between glass and rollers takes place during the process. The velocity of glass is equal to the circumferential speed of the rollers. The heat transfer phenomena in a tempering furnace are represented in Figure 4.31.17.

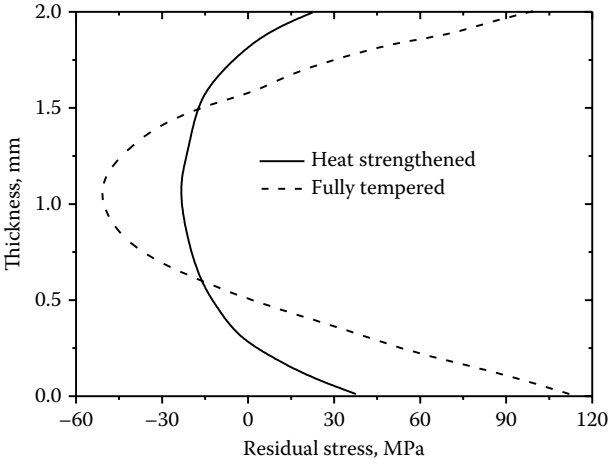


FIGURE 4.31.16 Residual stress comparison of fully tempered and heat-strengthened glass. (From Glass-academy.in.)

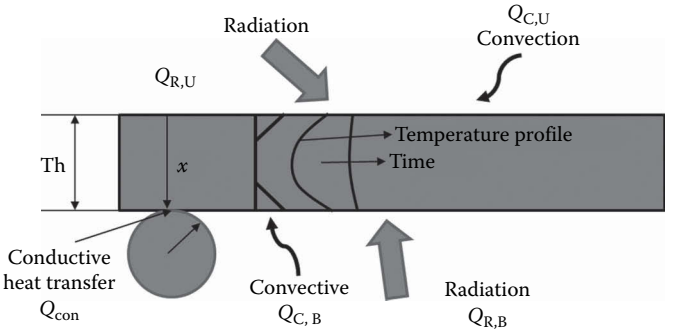


FIGURE 4.31.17 Heat transfer phenomena in a tempering furnace. (From Rantala, M. Heat transfer phenomena in float glass heat treatment processes. PhD Thesis, Tampere University of Technology, Finland, 2015. With permission.)

The equation representing the overall heat transfer phenomena as shown in Figure 4.31.17 is given by

$$\rho C_p \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) + \frac{\partial Q_R}{\partial x}, \tag{4.31.8}$$

where ρ is the density, C_p is the specific heat, and k the thermal conductivity of the glass, respectively. Further, *x is* a measure of the distance into the glass from its top surface and t is the overall time elapsed in the tempering line.

From Equation 4.31.8, the overall heat transfer through the glass in a tempering line can be classified into two categories. First, glass behaves like a lumped mass storing the overall incident heat on its surface from conduction by rollers $\left(k\frac{\partial T}{\partial x}\right)$, and radiation heat transfer from its surroundings $\left(\frac{\partial Q_R}{\partial x}\right)$ as given in Equation 4.31.8. $\rho C_p \frac{\partial T}{\partial t}$ is the overall heat stored by the glass (as lumped mass), hence resulting in an increase in its temperature *T*.

Further, an increase in glass temperature T would result in heat transfer from its surface into the ambient of tempering line via radiation and convection (with the surrounding air). The Equations representing net heat transfer by convection and radiation from glass surface are represented in Equations 4.31.9 and 4.1.10. Equation 4.31.9 represents the heat transfer from glass top surface ($x=0$) facing the tempering line. The net heat from this surface is transferred through radiation ($Q_{R,}$ $x=0$) and convection (Q_C,$x=0$). Further Equation 4.31.10 represents the heat transfer from bottom surface ($x=L$) of glass resting on the rollers. The net heat transfer from this surface is through radiation (Q_R,$x=L$), convection (Q_C,$x=L$), and conduction (Q_{Con},$x=L$) to the rollers:

$$k\frac{\partial T}{\partial x}(0,t) = Q_{R,U} + Q_{C,U}, \tag{4.31.9}$$

$$-k\frac{\partial T}{\partial x}(L,t) = Q_{R,B} + Q_{C,B} + Q_{Con}. \tag{4.31.10}$$

Representing the overall conduction heat transfer at surface $x=0$ and $x=L$ from Equations 4.31.9 and 4.31.10 in Equation 4.31.8, the net temperature of the glass in the tempering line can be estimated. The temperature, along with other parameters, determines the extent of strengthening via residual stress generation, as detailed in the chapter.

Summary

A short overview is presented in this chapter on manufacturing of glass with focus on application of heat transfer concepts. Glass manufacturing is evidently a high-temperature process in which raw material is melted, followed by reactions involving the components of the raw mixture and cooling of the melt. Radiative heat transfer plays a dominant role during melting and thermal annealing as the temperatures involved are very high. Temperature variation between the bulk of the melt and its surface generates natural convection within the melt. This convection helps to improve its quality by bringing the tiny bubbles that are formed in the melt to the surface. It also leads to better mixing of the raw material introduced at the surface with the bulk melt. After melting and cooling, thermal treatment is required to manufacture glass for a given specific application. In this chapter, float glass manufacturing is presented as an example and the thermal tempering process is discussed in some detail.

ACKNOWLEDGMENTS

R. Tadepalli wishes to acknowledge the contributions of Sasank Bethapudi and Mohammed Ashik toward the section on heat treatment for glass transformation.

REFERENCES

Curlet, N. W. E, Won, K. J., Clomburg, L. A., Jr., and Sarofim A. F. 1984, Experimental and mathematical modeling of three-dimensional natural convection in an enclosure, *AIChE J.*, 30(2):249–257.

Faber, A. J., Beerkens, R. G. C. and de Waal H., 1991, Thermal behaviour of glass batch on batch heating, *Glastech. Ber.*, 65:177–185.

Frank, M. and Klar, A. 2010, *Mathematical Models in the Manufacturing of Glass,* CIME Summer School, Springer, Terme, Italy.

Glass Academy-Premier Glass Portal. http://www.glass-academy.com/elearning/

Krause, H., and Krause, D. 2002, *Mathematical Simulation in Glass Technology*, Springer-Verlag Berlin Heidelberg.

Pfaender, H. G. 1996. *Schott Guide to Glass*, Chapman and Hall, London.

Pilon, L., Zhao, G. and Viskanta, R. 2006, Three-dimensional flow and thermal structures in glass melting furnaces. Part I. Effects of the heat flux distribution, *Glass Sci. Technol.*, 75(2):55–68

Rantala, M. 2015. Heat transfer phenomena in float glass heat treatment processes. PhD Thesis, Tampere University of Technology, Finland.

Rawson, H. 1974. Physics of glass manufacturing, *Phys. Technol.*, 5:91–114.

Richet, P. 1987. Heat capacity of silicate glasses, *Chem. Geol.*, 62:111–124.

Roychowdhury, A. P. and Srinivasan, J. 1994. The modelling of radiation heat transfer in forehearths units in glass melting, *Heat Mass Transfer*, 30(2): 71–75.

Shand, E. B. 1958. *Glass Engineering Handbook*, McGraw-Hill Book Company Inc., New York.

van Zee, A. F., and Babcock, C. L. 1951. A method for the measurement of thermal diffusivity of molten glass, *J. Am. Ceram. Soc.*, 34:244–250.

4.32 SOLAR HYDROGEN AS A "RENEWABLE REDUCTANT": POINTS AND COUNTERPOINTS

Raj Ganesh S. Pala

Introduction and Scope

Unlike most of the previous energy crises [1], which were more often than not, driven by artificial constraints on the *supply side* of fuels, the energy crisis of the future will be driven by significant increase on the *demand side*. It is not only the overall population, but also the longevity of people that is increasing and this is coupled to rising prosperity, all of which is surely bound to increase the demand for energy [2,3] Energy *per se* is not the problem as Earth is inundated with solar energy [4–6], but the quintessential search in energy science and engineering is for *energy conversion* technologies that are *environmentally sustainable*. Solar energy falling on Earth has to be converted to more usable forms like electrical energy or mechanical energy, and as this energy has significant diurnal and seasonal variation, energy conversion technologies like solar cells are often connected to energy storage technologies like batteries [4,5]. Energy storage technologies are more critical in countries like India, wherein even if electrical grid connectivity exists, certain electrical grids are often operated only during the night to meet the requirement during nights (especially in the form of energy for lighting). This imposes the constraint wherein solar photons are only available during the daytime and even if the solar photons are utilized to generate electricity via solar cells, it is impossible to store the generated electricity, for later use at night, if inexpensive energy storage technologies are unavailable.

From the time humans discovered fire, they have come to depend on burning carbon containing oxidizable biochemical or chemical feedstocks to obtain energy. These feedstocks were generated naturally over millions of years using solar energy and biomass, but are being consumed due to human activity at exceptionally faster rates, compared to the natural timescale for their generation. While the gradual drift has been from carbon-rich fuel feedstocks like wood to coal to petroleum to carbon-poor fuel feedstocks like natural gas, the generation of CO_2, which is correlated to the harmful greenhouse effect, will not stop unless zero-carbon fuels are utilized. While nuclear energy may offer a solution to near zero-CO_2 generating pathways for energy needs [3], its boon of it being *very high energy density*, also proves to be its bane and it is not clear if the socio-economical and political challenges of nuclear energy will be surmounted in the near future. If hydrogen can be generated using renewable energy and using renewable feedstocks, hydrogen as a *renewable fuel* can be "burnt" either thermochemically via hydrogen turbine/internal combustion engine to obtain mechanical energy or electrochemically via a fuel cell to obtain electrical energy [4,5].

Hydrogen, in addition to being an energy carrier and source of energy, can also be viewed as a "*renewable reductant*" to mitigate/chemically reduce/chemically fix/chemically store the CO_2 that has been generated due to the dramatic increase in energy demands following the Industrial Revolution. *For hydrogen to be a "renewable reductant," it has to be generated via a near-abundant feed like water, and via renewable sources like solar energy.* The hydrogen so generated can be reacted with CO_2 to generate useful chemicals like methanol, which can, in turn, be used for generating other commodity chemicals like ethylene via the methanol-to-olefin process [7].

Solar-based hydrogen generation can be achieved via two broad electrochemical methods. In the "indirect" electrochemical method, an electrolyzer is utilized to drive the endothermic water splitting process and generate hydrogen and oxygen, with the energy required for electrolysis supplied by solar cells [4]. There are certain advantages in this process—physicochemically, as the solar energy absorption and current generation is undertaken by a solar cell, this device can be optimized independently of a second device, the electrolyzer, which takes the current/voltage/electrical power from the solar cell and generates hydrogen/oxygen. The feasibility of both these technologies—solar

cells and electrolyzers—has been demonstrated at industrial scales for the generation of electrical power and hydrogen/oxygen, respectively, albeit not being the cheapest method for generating electrical power/hydrogen/oxygen.

In the second "direct" electrochemical method, solar light falls onto the electrode(s) of an electrolyzer, which generates the necessary potential for splitting water. Such a process, broadly referred to as the "photoelectrochemical" (PEC) process [8–14], has captured the imagination of chemical scientists and engineers for more than four decades and has been a Holy Grail problem, waiting to be the elixir for sustainable energy [15,16], if the device is made with inexpensive Earth-abundant elements [17,18] via low-cost manufacturing techniques. This elusive search toward an engineered "photochemical" solution predates the last four decades and has been seen as a panacea for the Faustian bargain of relying on carbon-based energy feedstocks even a century ago [19,20]. The allure of this process is partly because of its *apparent* simplicity [8–14] and partly because of the possibility presented by Nature, which has mastered solar energy-based generation of carbohydrates from the simplest of raw materials—carbon dioxide and water—via intricate self-organization and self-assembly of biomacromolecules [16,21].

Due to the relatively low energy density of solar radiation, all solar technologies involve large-area devices. To obtain a feel for the areas involved, let us consider an order of magnitude analysis to estimate the areas involved in the indirect method of hydrogen generation via electrolyzer powered by solar cells. Large-area electrodes are required for generating H_2 that is to be used for reducing all the CO_2 generated in a given year. As these electrodes can be vertically stacked, they will not occupy large land area, but significant area is required if inexpensive solar cells are to power these electrolyzers. For the year 2014, the global CO_2 emissions were ~35.9 Gt ($= 34\times10^{12}$ kg), which amounts to 0.816×10^{15} g mol. Suppose, we convert CO_2 to methanol via the following reaction: $CO_2+3H_2 \rightarrow CH_3OH+H_2O$, thereby requiring 2.448×10^{15} mol of H_2 per year. Let us assume that electrocatalysts made from Earth-abundant materials are used for the production of H_2, and that the efficiency of H_2 production is 68.3% ($\eta = 1.23/1.8$). Let us also assume that amorphous Si solar cells having an energy efficiency of 10% are to power these electrolyzers. Since the solar power flux is 100 mW/cm^2, the current density obtained from the electrolyzer, powered by amorphous Si solar cells, will be 5.55 mA/cm^2 ($0.1\times0.683\times100\,/\,1.23$ mA/cm^2). The total mol of H_2 produced in a year, assuming six hours of continuous solar light in a day, per square km will be 226.75×10^7 ($0.5\times5.55\times10^{-3}\times365\times6\times3600\times10^{10}/96{,}485$ mol). As 2.448×10^{15} mol of H_2 is required for converting CO_2 to methanol, the area of solar panels required to power the water electrolyzer will be 1,079,729 km^2. Hence, a land area of 1,295,675 km^2 (1.2 times the solar panel area) will be required, which is approximately equal to the land area of the state of New Mexico!

Land area requirement for direct PEC H_2 generation processes has been estimated in solar insolation rich regions [22] and in the vicinity of refineries [23], wherein the hydrogen generated can be used for other petrochemical processes without the need for expensive hydrogen storage devices.

While photocatalytic (PC) routes have been explored for many chemical systems like photoreduction of CO_2 and wastewater treatment [4,12], most of the work has been focused on the direct water splitting. The scope of this review chapter will be restricted to photoelectrochemically driven low-temperature water splitting alone; no attempt has been made to address other approaches to renewable hydrogen [21]. The chapter is designed to serve as a qualitative introduction to the field, but is also intended to give pointers on issues that have been addressed less extensively. In addition, the chapter has carefully chosen references to books and recent reviews of the field, wherein the subject matter is discussed at a much greater depth. In the section on physicochemical principles of photoelectrochemical systems, we review the elementary principles of photoelectrochemical systems. In the section on model identification and refinement, we discuss some common methods to discern the models that have been proposed for PEC systems. The issues of efficiency and process design are addressed in "Aspects of Efficiency Metrics, Photoreactor Design, and Process Design." Materials are central to the realization of PEC technologies and aspects of material design are discussed in "Realizable Efficiencies are via Materials: Some Hope, Some Hype, and the New Hope."

The section "Roads Less Travelled: Some Challenges in Attaining the Holy Grail" has certain personal and speculative viewpoints on different aspects of PEC systems that I feel have not been adequately addressed. The chapter ends with "Concluding Remarks."

Physicochemical Principles of PEC Systems

Energetics of Water Splitting

A water electrolyzer/electrolytic cell as well as a water splitting PEC system/photocatalytic semiconducting particle are engineered to drive the following endothermic reaction:

$$H_2O \rightarrow \frac{1}{2}O_2 + H_2, \Delta G = 237 \text{ kJ/mol} \quad \text{(R1)}$$

This overall *redox* reaction is a combination of a reduction reaction:

$$2H^+ + 2e^- \rightarrow H_2 \quad \text{(R2)}$$

and an oxidation reaction:

$$H_2O \rightarrow \frac{1}{2}O_2 + 2H^+ + 2e^- \quad \text{(R3)}$$

The oxidation and reduction reaction occurs at a solid–liquid material interface, in regions often referred to as the anodic electrode surface of an electrolytic cell (or anodic surface sites of a single particle) and cathodic electrode surface (or cathodic surface sites of a single particle). In addition to the nature of the reaction occurring at these surfaces (or surface sites of a single particle), these surfaces (or surface sites of a single particle) are characterized by their distinct electrochemical potential. Without a highly technical discussion on the nature of electrochemical potential, which requires an extended detour to the fundamentals of electrochemistry [24,25], it is useful to develop qualitative insights via classical electrostatics.

In electrochemical reactions, the electrode potentials play a role akin to temperature in a thermochemical reaction, in the sense that the variation of electrode potentials modulates electron transfer between the electrode and electro-active species in the electrolyte, thereby controlling the rate of the electrochemical reaction. Modulation of *rates/kinetics* is possible as electrode potentials modulate the *energetics/thermodynamics* of the electro-oxidation and electro-reduction reactions.

As the overall reaction (Equation R1) is endothermic and considering reaction R3 at "zero electrode potential," the energetics does not favor dissociation of water. However, as the electrode potential is made positive (or shifted to an *anodic* potential), the dominant effect of the *positive* electrode potential is in stabilizing the electron appearing on the right-hand side of Equation R3. Greater positive potential favors water splitting to a greater extent; *negative* potential (or shifting to a *cathodic* potential) has the opposite effect and it destabilizes the electron appearing on right side of Equation R3. The more endothermic the overall reaction is, greater will be the electrical bias, which is the difference between the positive and negative electrode potential, required to drive the overall reaction. The free energies of the reaction are measured in kJ/mol, but V_{bias} (defined as $V_{bias} = V_{cathode} - V_{anode}$) is in volts, and both are connected by the equation $\Delta G_{reaction} = -nFV_{bias}$, where n = number of moles of electrons involved in the electrochemical reaction and F = Faraday constant = 96,485 C/mol. In the case of water splitting reaction, $V_{bias} = -1.23$ (as $V_{cathode}$ is negative with respect to V_{anode}) and $n = 2$.

Typically, in an electrolytic cell, the anodes and cathodes are metallic in nature, with their primary purpose being to change the energetics of the electro-oxidation and electro-reduction reactions. In addition, different metallic electrodes can chemically interact with the reactants/products in a different manner and hence, have different "*electrocatalytic*" properties [26].

In contrast to electrically biased electrolytic cells for water splitting, a PEC water splitting cell has at least one of the electrodes made of a semiconductor material. In contrast to metals, which have unoccupied electronic states for "small" energy electronic excitations, semiconductors, lack unoccupied electronic states for "small" excitations [27]. Here, "small" energies are of the order of characteristic energy scale kT, where k is the Boltzmann constant, T is the temperature of the material [27], and kT at room temperature is ~25 meV. In a semiconductor, the lowest unoccupied electronic state is separated from the highest occupied state by a band gap, which is much larger than kT. The highest "very closely" spaced set of energy levels of filled electronic states is called the valence band and the lowest very closely spaced energy levels of empty electronic states is called the conduction band. Here, "very closely" spaced indicates that the separation between electronic energy levels is much less than the characteristic energy scale kT. The highest energy level of the valence band is called the valence band maximum (VBM) and the lowest energy level of the conduction band is called the conduction band minimum (CBM). The difference between CBM and VBM is called the band gap. The band gap of Si is 1.12 eV (which is much larger than kT at room temperature); the band gap of typical semiconductors that have been explored in the context of PEC is given in Refs. [8] and [12] amongst others. The distribution in energies of the photons is dependent on the black body spectrum of the sun [8] and if the semiconductor has a small band gap, it absorbs more solar photons. However, not all of these photons can contribute toward water splitting.

The minimum V_{bias} required for splitting water in an electrolyzer is 1.23 V; but in addition to this, other losses like the ohmic drop in the electrolyte and losses associated with overcoming the electrochemical reaction activation barrier [24,26] are lumped together under various overpotentials [24,26]; the sum of these various overpotentials amounts to about 0.4 V. Hence, a typical electrolyzer operates at a potential of $V_{op} = 1.63$ V so as to have appreciable reaction rates and current density.

PEC systems are very similar to a simple electrolytic cell, but it contains at least one semiconductor electrode. The two electrodes are physically separated by an electrolyte, but are electrically connected. When light is incident on the semiconducting electrode, a potential is generated in the semiconductor electrode and if this potential bias is of the order $V_{op} = 1.63$ V of an electrolytic cell, then there will be sufficient water splitting activity. But alas, there would have not been the frenetic activity in the field of PEC over the last four decades if the generated photopotential were so high.

In contrast to PEC systems, water splitting can be attempted via a single semiconductor particle too; we will refer to such systems as *photocatalytic* (*PC*) water splitting semiconducting particles (Figure 4.32.1). When light falls on such particles, potential bias is generated within these particles such that some parts of the particle are anodically biased and some other parts are cathodically biased. The anodic surface sites of the PC particles host the oxidation reaction (Equation R3) and the cathodic surface sites host the reduction reaction (Equation R2).

Band Bending in the Semiconductor–Electrolyte Interface

When a voltmeter is connected between two different metals, it measures the difference between the highest energy level of the electron (called the Fermi level) in the different metals. Fermi level is a thermodynamic quantity with rigorous microscopic interpretation from statistical mechanics and it takes into account the effect of temperature/chemical nature of the material [27]. Lower the Fermi level of a metal, greater is the energy required to remove an electron from the metal; this energy is called the work function of the material. When two metals are electrically connected to each other, electrons will flow from the metal that has lower work function (or greater Fermi level) to the metal that has higher work function (or lower Fermi level) till the Fermi level of the two metals equalizes.

In a similar manner, the Fermi level of the semiconductor can be measured with respect to other metals or other semiconductors. For an n-type semiconductor, the Fermi level is closer to and just below the CBM and for a p-type semiconductor, the Fermi level is closer to and just above the VBM. In an undoped semiconductor, the Fermi level is exactly in the middle of the band gap, despite there being no electronic energy level in the band gap [24].

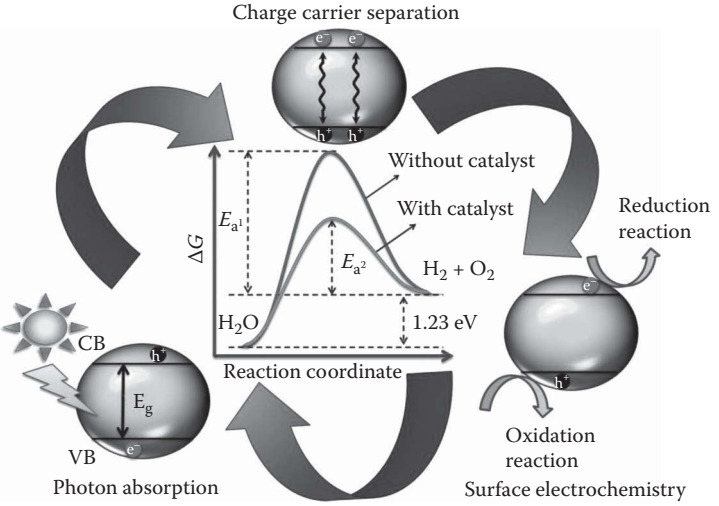


FIGURE 4.32.1 Coupling between the three fundamental physicochemical processes in a photocatalytic system.

In the electrochemical literature, the electronic energy level of redox couple in the electrolyte is given by the Nernst equation [24]; this redox energy level of the electrolyte is indicative of the tendency to transfer charges. A more positive redox level will more easily get reduced by accepting electrons from a redox level that has less positive redox level.

In a PEC cell, there is at least one interface between a semiconductor electrode and the electrolyte and in a *tandem*-PEC cell, both the electrodes are made of semiconductors. The electronic energy levels between the semiconductor and the electrolyte can be compared using a common scale; it is to be noted that the Fermi level of a semiconductor is functionally the same as the redox electrochemical potential of an electrolyte [2]. Further, when an electrical contact between the semiconductor and the electrolyte is established, charges are transferred from the higher electrochemical potential to the lower electrochemical potential. Such a transfer of charges induces bending of electronic bands in the semiconductor, which is very crucial for the separation of e^-/h^+ (Figure 4.32.2).

Dynamics of Charge Transfer between the Semiconductor and the Electrolyte

When an *n*-type semiconductor is exposed to light, it will absorb photons that have energies of the order of the band gap or greater. As these absorbed photons provide the energy for driving the endothermic water splitting reaction and as the band gap sets the bound for minimum energy for the photon that can be absorbed by the semiconductor, the band gap in turn has to be at least >1.23 V. Upon photon absorption, an electron is excited to the CBM and a hole (which is a missing electron) is at the VBM.

As the energy levels near the semiconductor surface are bent due to the lack of electroneutrality (Figure 4.32.2), the hole is "driven up" along the VBM to the surface of the semiconductor and reacts with water molecules to form O_2 and protons. Thus, the *n*-type semiconductor serves as the anode, performing oxidation, in the PEC cell. For this charge transfer to take place, the redox level of the H_2O/O_2 couple should be higher than the bent VBM of the semiconductor so that the hole can move up from the bent VBM level of the semiconductor to the occupied redox levels in the electrolyte.

As the valence band of a semiconductor has greater contribution from orbitals from the anions of the semiconductor, the more electronegative the anions, the deeper are the VBM levels and more facile is the transfer of the hole from the VBM to the less electronegative O from the water molecule. This has important consequences in terms of trade-off between band engineering to increase solar

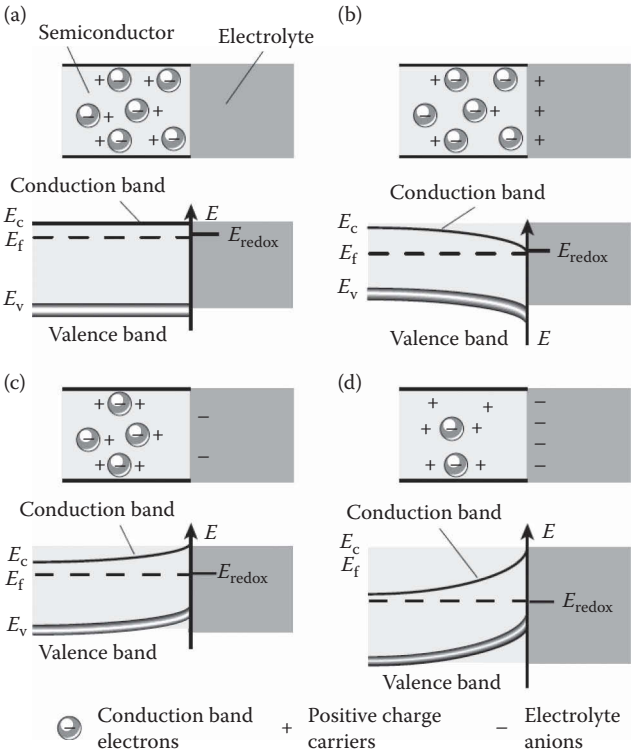


FIGURE 4.32.2 Different band-bending scenarios depending on the differences in the fermi level of the semiconductor and the redox level of the electrolyte. (Adapted from Gratzel, M., *Nature*, 414(6861): 338–344, 2001. With permission.)

photon absorption and ease of hole transfer. If the semiconductor is made of less electronegative anions, the VBM might move up and this will decrease the band gap and increase solar photon absorption, but, in turn, the driving force for hole transfer will also be decreased.

The electron along the CBM is "driven down" because of band bending and driven away from the semiconductor–electrolyte interface. *Note that band bending provides the driving force for the movement of the electrons and holes in the opposite direction.* Electron from the conduction band is transferred to the "counter electrode" (which is a metal in Figure 4.32.2), and cathodic reduction of the proton is performed at the surface of the metal electrode. The energy of the electron ejected from the counter electrode is determined by the CBM of the semiconductor photoanode and hence, for the reduction of a proton to occur at the metallic counter-electrode, the CBM level of the semiconductor should be higher than the empty energy levels available on the proton. As the orbitals of the cation make a greater contribution to the conduction band of the semiconductor, less electropositive cations will decrease the band gap, increase photon absorption, and decrease the driving force for electron transfer at the counter electrode.

Typically, the V_{bias} generated when light is incident on the semiconductor does not provide sufficient potential to split water. Hence, in addition to the photopotential generated due to exposure to light, an external electrical bias is applied in a manner similar to that of an electrolytic cell. However, the applied electrical bias is expected to be smaller than 1.6 V, which is the order of magnitude of bias in an electrolytic cell. More the generated photopotential upon light incidence, lesser will be the applied external electrical bias, and more efficient the solar energy harvesting using a PEC system.

In this section, we have attempted to provide an elementary discussion of the three physicochemical processes occurring in PEC systems—photoabsorption, electron–hole separation, and surface

electrochemistry. The physicochemical processes are much richer and many more intricate details can be added to this description to provide a quantitative description of PEC systems. The overall goal of such a description is to obtain a quantitative relationship for photocurrent as a function of solar intensity and electrical bias. These have been presented in many books and reviews in the past (see, e.g., Refs. [8,9,28,29]).

Model Identification and Refinement

The three processes central to PEC/PC—photoabsorption, electron–hole separation, and surface electrochemistry can be measured and modeled to varying degrees of certainty and success. Measurement of photoabsorption and rate of PECs reactions can be made rather conclusively [8,24]. Advancements in density functional theory (DFT) [26,30–46] have made it possible to screen a large number of materials for appropriate band gap, band edges and surface electrocatalytic properties. As of now, DFT provides the most accurate energetics for the material systems being considered for PEC/PC applications. Given this accuracy, it would be very useful to develop mathematical models via a bottom-up approach starting from the energetics based on electronic structure DFT calculations all the way to phenomenological mathematical quantities relevant to the measured photocurrents. While such advances have been made for electrochemical systems [47–49], accurate multiscale models for PEC systems has been beyond the reach of such approaches.

The most challenging aspect of PEC is to unambiguously quantify and model electron–hole separation. In this context, electrochemical impedance spectroscopy (EIS) has played an important role in yielding insights into these processes and on the dynamics of charge transfer [50,51]. In contrast to DFT, which is typically used for modeling "ideal" bulk and surface solids, EIS is used for modeling "real" experimental systems [50,51]. EIS has played an important role in discriminating between the competing models and have established the crucial role of surface states in PEC [48,52]. While the complexity of the system often demands more than one electrical circuit that fits the EIS data well, ascertaining the validity of the models via DFT seems very promising [51].

Absorption of light creates quasi-Fermi levels, which are hard to measure quantitatively. However, the "on-set potential" under light has been utilized as an approximation for quasi-Fermi levels in PEC systems [53]. Material design based on this strategy can be challenging because band positions of the materials are sensitive to impurities, defects, and surfaces [54]. Nevertheless, much progress has been made in aligning the energy levels of various molecules and semiconductors [55].

Aspects of Efficiency Metrics, Photoreactor Design, and Process Design

At a fundamental level, solar-based H_2 generation is a method for energy harvesting/energy conversion and the efficiency of this process can be captured by solar to hydrogen (STH) conversion efficiency [14,29,56]. The STH efficiency of a PEC water splitting system is defined for systems exposed to solar irradiance (1 Sun, Air Mass (AM) 1.5 global conditions) under zero bias conditions and these conditions ensures that no additional electrical bias is applied between working and counter electrodes. In order to avoid chemical potential bias, the working and counter electrodes are at the same pH and the electrolyte should not contain any sacrificial donors or acceptors. Quantitatively, STH efficiency calibrates the ratio of the chemical energy of the hydrogen produced to the incident solar energy:

$$\mathrm{STH(based\ on\ collected\ H_2\ gas)} = \frac{(\mathrm{mmol\ H_2/s}) \times (237\ \mathrm{kJ/mol})}{(P_{\mathrm{total}})(\mathrm{mW/cm^2}) \times \mathrm{Area(cm^2)}}_{\mathrm{AM1.5\,G}} \times 100 \qquad \text{(R4)}$$

Note that in this equation, Gibbs free energy per mole of H_2 is $\Delta G^\circ = 237$ kJ/mol. The precise definition of AM 1.5 G are available in many text books (see, e.g., Ref. [21]). Typically, the extent of H_2 generated is monitored by mass spectrometry or gas chromatography. It is important to ensure

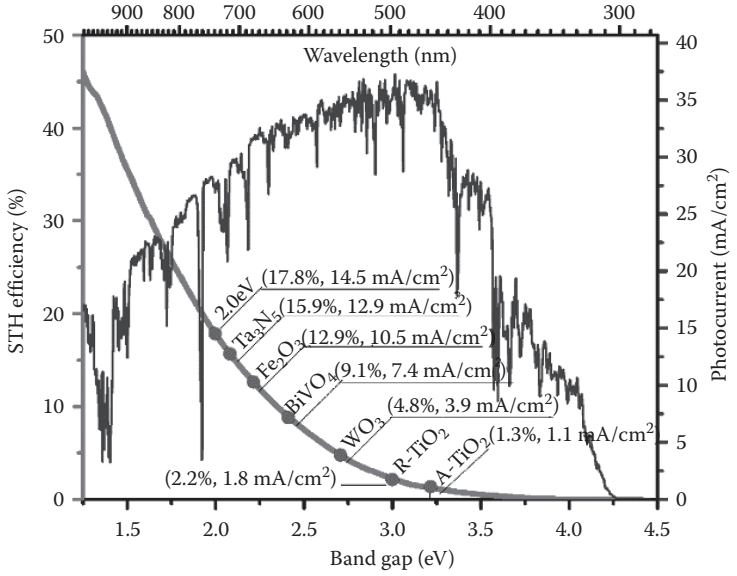


FIGURE 4.32.3 Plot of STH efficiency vs. material band-gap for important photo-materials: Right axis represents theoretical maximal photocurrent (j_{max}) and left axis represents STH efficiency are plotted against material band gap (E_g). (Reprinted from Li, J. and Wu, N., *Catal. Sci. Technol.*, 5(3), 1360–1384, 2015. Copyright © 2014, Royal Society of Chemistry with permission.)

that both H_2 and O_2 are evolved in stoichiometric proportions; deviations suggest corrosion of the electrode or other side reactions.

The short-circuit photo current (j_{SC}, mA/cm^2) can also be used to compute STH efficiency.

$$\text{STH (based on current)} = \frac{(j_{SC}\ \text{mA/cm}^2)\times(1.23\ \text{V})\times\eta_F}{(P_{total})(\text{mW/cm}^2)}\Bigg|_{\text{AM 1.5 G}}\times 100 \tag{R5}$$

where η_F is the Faradaic efficiency [24].

The most easily quantifiable factor that is correlated to the STH efficiency is the band gap of the material as the band gap determines the maximum theoretical photocurrent (j_{max}) (Figure 4.32.3). This might also be one of the reasons why most emphasis has been placed on modulating the band gap even though electron–hole recombination and surface electrochemistry also play a significant role in determining the STH efficiency. While many strategies have indeed succeeded in increasing the solar photon absorption, there has not been concomitant increase in the STH efficiency; one of the reasons might be that strategies to increase solar photon absorption have a deleterious effect on e^-/h^+ recombination [58].

As PEC systems have many coupled physiochemical processes, STH efficiency convolutes all these coupled processes and in order to deconvolute the efficiency of different processes, the following metrics are commonly used:

1. Applied bias photon to current efficiency (ABPE)
   In most semiconducting electrodes, unless external electrical bias is applied, significant current densities are not registered. Applying an electrical bias affects the extent of band bending and is expected to have a significant effect on e^-/h^+ separation and hence, the efficiency [9]. Such changes in efficiencies can be calibrated by ABPE, which is defined as:

$$\text{ABPE} = \frac{|j|(1.23 - V_{bias})}{P_{Sun}} \tag{R6}$$

where $|j|$ is the photocurrent density, V_{bias} is the electrical bias of the working electrode with respect to the counter electrode and not with respect to the reference electrode [59,60], and P_{Sun} is the incident solar power density. It is also important to thoroughly understand the difference between a "2-electrode" and "3-electrode" measurement [59,60].

2. Incident photon to current efficiency (IPCE)
While STH/ABPE does not differentiate between photons of different wavelengths, IPCE calibrates the efficiency of photons of different wavelengths, hence different energies, to generate photocurrent. It is defined as:

$$\mathrm{IPCE} = \frac{\dfrac{\mathrm{Electrons}}{\mathrm{cm^2 s}}}{\dfrac{\mathrm{Photons}}{\mathrm{cm^2 s}}} = \frac{j_{ph}\left(\dfrac{\mathrm{mA}}{\mathrm{cm^2}}\right)\times 1239.8(\mathrm{V}\times\mathrm{nm})}{P_{mono}\left(\dfrac{\mathrm{mW}}{\mathrm{cm^2}}\right)\times\lambda(\mathrm{nm})} \tag{R7}$$

where $1239.8(\mathrm{V}\times\mathrm{nm})$ arises from multiplying the Planck's constant (h) and the speed of light (c). While IPCE deconvolutes the effect of photons of different wavelengths, it does not deconvolute the three processes of photon absorption, e/h separation, and surface electrochemistry.

3. Absorbed photon to current efficiency (APCE)
To deconvolute optical processes like reflection and transmission from the overall PEC processes, APCE calibrates the photocurrent generated for every absorbed photon (of a particular wavelength), assuming that each absorbed photon generates an e/h pair. Thus, APCE convolutes only the efficiency of e/h separation and surface electrochemistry and is defined as

$$\mathrm{APCE} = \frac{j_{ph}\left(\dfrac{\mathrm{mA}}{\mathrm{cm^2}}\right)\times 1239.8(\mathrm{V}\times\mathrm{nm})}{P_{mono}\left(\dfrac{\mathrm{mW}}{\mathrm{cm^2}}\right)\times\lambda(\mathrm{nm})\times(1-10^{-A})} \tag{R8}$$

where A is the absorbance of the sample estimated from Beer's law [relates A to the logarithmic ratio of measured output light intensity (I) to initial light intensity (I_0)].

4. Intrinsic solar to chemical conversion (ISTC) efficiency
ISTC aims to calibrate the effectiveness of a photoelectrode even when the electrical bias is more than 1.23 V [56]; measurements are made within the 3-electrode system. ISTC is defined as

$$\mathrm{ISTC} = \frac{\eta_F\times 1.23\ (\mathrm{V})}{V_{dark}(\mathrm{V})}\times\frac{j_{photo}\left(\dfrac{\mathrm{mA}}{\mathrm{cm^2}}\right)\times V_{photo}(\mathrm{V})}{P_{solar}\left(\dfrac{\mathrm{mW}}{\mathrm{cm^2}}\right)}, \tag{R9}$$

where V_{photo} is the photo-voltage and is the difference between the voltage required to produce a particular current density in dark and light conditions. The voltage required to produce the same current density under illumination and η_F is the Faradaic efficiency [24]. Compared to ABPE, ISTC is more useful in terms of the solar energy conversion to Faradaic current when electrical bias is >1.23 V for some materials (please note that ABPE is not defined in this range).

When efficiency is less than 100%, as is always the case in any energy conversion device, it is important to have an insight into what possible factors contribute to losses in efficiency. In electrochemical literature [24], different "overpotentials" are indicative of physicochemical factors that contribute toward the loss of available free energy in an electrochemical energy conversion device.

Losses like concentration and activation overpotentials for water splitting are well documented in the electrochemical literature and these amount to at least 0.4 V; such losses may be present in PEC systems, but to a lesser extent as the current densities involved are smaller in PEC systems compared to a water splitting electrolytic cell. In addition to the losses mentioned earlier, there will also be collection losses as not all the products are being collected; both ohmic and collection losses might be more challenging in scaled-up photoreactors.

Many aspects of reactor and process design have been documented recently [22,23,61–63]. Among the photoreactors analyzed, the simplest and least expensive design is a slurry type reactor containing semiconducting photocatalytic particles. However, the main trade-off in this design is that the combustible mixture of hydrogen and oxygen is present in the same compartment. Pressure-swing absorption can be used for separating H_2 and O_2 mixtures from the reactor [22]. Analysis suggests that photocatalytic slurry-type reactors can be competitive with hydrogen generated from steam-reforming processes, *provided significant improvements are achieved in terms of the photocatalysts* [22].

It is beneficial to locate a solar-based hydrogen generation facility at places where there is high solar insolation. However, such places may not be sites where hydrogen is used. Due to these reasons, solar-based hydrogen economy may require additional technologies like hydrogen storage devices or fuel cells; these technologies are expensive and have not been implemented at large scales in an economical manner. In contrast to the solar-based hydrogen generation, the steam methane reforming route to generating H_2 is typically undertaken at the sites where H_2 is required. In light of these arguments, the process economics for 1 ton of solar hydrogen generation by locating the photoreactors in the vicinity of major petrochemical plants has been analyzed [23]. The solar insolation of these plants are different, but all of these plants have extensive and increasing need for H_2 [23]. The role of carbon credits was also taken into account and was used to evaluate the extent to which this will contribute to the process economics of solar-based hydrogen generation when contrasted with the steam methane reforming process that generates CO_2, but it was found that at the present levels of carbon credits, this does not significantly affect the process economics. While the critical factor that will determine the process economics is the efficiency of PC/PEC materials, we do feel, this process design of integrating solar-based hydrogen generation into traditional petrochemical complexes might be the first viable implementation of PC/PEC system [23], if these technologies are to be implemented at large scales in the future.

Realizable Efficiencies Are via Materials: Some Hope, Some Hype, and the New Hope

Literally thousands of materials have been explored in the context of PEC; it is not only impossible to review these results in this chapter, but also unnecessary as review of materials has been done many times over in many reviews [2,8–10,12,14,29,61,64–70]. Instead we focus on two classes of materials that have always posed considerable, but unrealized promise.

Titania: Black, White, and Colored

Sometimes a single class of materials, like Si for photovoltaics, zeolites for the petrochemical industry, changes the landscape of the entire field. Titania or functionalized titania, from the time of its use as a photoanode by Fujishima and Honda in 1972 [71], has appeared to be a "messiah material" for PEC, but that promise has not fructified. The main strengths of titania are that it is very stable in the electrolyte, inexpensive, non-toxic, and fairly abundant. However, the disadvantage of titania is that it is a high band gap semiconductor because of which it absorbs very little solar energy. Pure titania is "white." In addition, it does not have significant electrocatalytic activity compared to typical oxygen evolution electrocatalysts like RuO_2 and IrO_2 [72,73]. To alleviate the optical absorption, various kinds of doping [65,74,75] and sensitization strategies [2,76] have been employed. Photosensitization involves adsorption or ligation of a dye with greater

photoabsorption cross section than titania [2,76]. Doping has often involved anions that are less electronegative than oxygen, cations that are less electropositive than titanium, and, in general, such dopants introduce mid-gap states, all of which tend to increase absorption cross section. Introduction of aliovalent dopants or reduction of titania can also introduce vacancies and all this makes titania colored. In the extreme, certain treatments via hydrogenation can make it "black," which is indicative of very high photoabsorption cross section [75]. However, higher photoabsorption cross section comes at the cost of concomitant increase in e^-/h^+ recombination, due to which darker shades of titania have not showed significant improvement in overall efficiency. This points toward the fact that two of the physicochemical phenomena—photoabsorption and e^-/h^+ separation—are strongly coupled [58].

Silicon: Slender and Stable

As silicon (Si) is the main workhorse of the entire photovoltaic science and engineering, it is natural to explore its utility in PEC systems [77]. The main challenge, when it comes to PEC, is that Si is unstable in aqueous electrolytes of most pH and gets corroded, forming an insulating silica layer; numerous strategies are being attempted to overcome this issue [77]. These remedial procedures include coating Si surface with a stable overlayer like metal oxides, metals, organic compounds, and other functional groups [77]. In addition, the band gap of Si is 1.1 eV and hence, a single photon excitation from VBM to CBM does not provide the driving force (which must be ~1.6 eV) for water splitting. However, when used in a "tandem" configuration, where two Si-based photoelectrodes absorb one photon as a photocathode (wherein Si surface is stable) and one photon as a surface passivated Si photoanode, they can overcome the driving force for water splitting [78]. Also, the nanostructuring of Si, which is of much importance in a variety of modern nanostructured semiconductor devices [79], has been explored. The rationale behind the design of Si nanowires is to decrease the distance that the electronic carriers (at least one of them, either the electrons or the holes) have to travel before they can be utilized for surface electrochemistry.

In addition, Si nanowires have also been combined with titania nanowires to implement a "Z-scheme" [80], but such a design, while it enhances e^-/h^+ separation, will still be limited by photoabsorption cross section of titania.

Organic–Inorganic Perovskite: Let Us Wrap It Up

Organic–inorganic perovskite (OIP) materials are crystalline materials that have a composition ABX_3, wherein "A" is a small organic cation, "B" is a (typically a Group 14) metal like Pb, and "X" is a halogen. The most well studied OIP is methylammonium lead triiodide; OIP seems to pose revolutionary changes in the field of solar-based materials technology [81,82]. Many OIPs are anticipated to possess properties very relevant to good optoelectronic materials such as having an appropriate band structure for significant light absorption, long carrier lifetime, low carrier effective mass, very good electron and hole transport properties, relatively benign electronic defects, and most importantly, the fabrication of these materials is relatively inexpensive and requires ingredients that are Earth abundant [81,82]. The progress in these systems has been so rapid, it has even reached popular and well-respected financial news outlets [83]. The central challenges toward a wide-scale implementation of these materials are the presence of leachable Pb in the crystal framework, hysteresis, and most critically, chemical instability in the presence of moisture [82,84]. When a material is sensitive to moisture, it degrades very easily under electrolytic conditions relevant to the PEC systems. *Hence, compared to the photovoltaic applications, the PEC applications of OIP have been rather limited and have relied on wrapping up perovskites with a protective layer involving metals, polymers, "conformal" coatings, or with other functional groups* [85–87]. In any case, the science of OIP is very rich and engineering consequences of OIP are enormous; this class of materials will be explored thoroughly for its utility in PEC systems.

Roads Less Travelled: Some Challenges in Attaining the Holy Grail

Engineered e^-/h^+ Separation

Of the three physicochemical factors mentioned, optical absorption, e^-/h^+ separation, and surface electrochemistry, increasing photoabsorption cross section has been the most emphasized, probably because of the emphasis on improving titania; this has not led to a successful material. Even if an inexpensive material absorbs just 10% of the solar photons, but if all these photons can be collected toward generating H_2, we still may have a successful PEC material; in this process, being able to separate the generated electron and hole to the surface is rather crucial.

Of late, there has been a more focused demonstration of engineered e^-/h^+ separation and this often involves the formation of heterostructures with appropriate staggering of band positions so as to facilitate the migration of electron–holes in the opposite directions. While the arguments are based on band alignments at the interface of the heterostructures and impedance changes, a much more detailed experimental and computational study should be undertaken to unfold the nature of the interface. Many of the interfaces that are formed do not have an epitaxial relationship and hence, it is expected that there will be dislocations/broken bonds and associated electronic defects. It is not clear how these will influence e^-/h^+ separation and it is possible that phenomenon like pinning of electronic states, in a manner analogous to Fermi level pinning might occur at the interface.

We are addressing this problem using "isomaterial native/non-native heterostructures" [88–90]. Non-native and "native" bulk structures differ in their discrete translational symmetry—for e.g., in titania, rutile is the native structure and anatase is a non-native structure and the commercial P-25 Degussa heterostructures used as a heterogeneous catalysts might be a commonly occurring isomaterial native/non-native heterostructures. The rationale, in the context of PEC, is that the formation of an interface between a native and non-native crystal structure, but with same chemical composition (and hence, "isomaterial") may decrease electronic defects that harm carrier transport [90].

Decoupling PEC, Electrochemical, and Heterogeneous Catalytic Metrics

While photoelectrocatalysts, electrocatalysts, and heterogeneous catalysts all are functional materials aimed at catalyzing chemical reactions, it is useful to compare the level to which the active sites are characterized quantitatively in heterogeneous catalysis, electrocatalysis, and in PC/PEC. Even though heterogeneous catalysts are themselves complex, the complexity of these materials is less than electrocatalysts and photocatalysts. Significant efforts have been spent in heterogeneous catalysis, with considerable success, to identify active sites at a molecular level and to characterize the activity in terms of the turnover frequency [91] and at the level of single-site heterogeneous catalytic sites [92]. Similarly, considerable knowledge on how to identify and characterize active sites at a molecular level has also been achieved [93]. However, when it comes to PC/PEC, part of the activity depends on "optical active sites" and partly on the "surface electrochemical active sites." In addition, strategies used for enhancing optical cross section will not only influence e^-/h^+ separation but also surface electrochemistry. For instance, numerous dopants have been used for enhancing the optical absorption of titania, but it is also known in the heterogeneous catalysis community that dopants have influence in the Mars–Van-Krevelen mechanism [94,95] and also induce non-Mars–Van-Krevelen oxidation mechanisms [95,96]. Hence, it is quite possible that the changes in PEC activity is not just due to dopant-induced changes in optical cross section, but can also be due to dopant-induced changes in surface electrochemistry, and this issue has not been adequately quantified. We feel that as PEC is a strongly coupled physicochemical phenomena, it is important to decouple the interlinked effects to test the effectiveness of different material design strategies.

Are All Potentials Created Equally?

Despite enormous effort and research over the past four decades, we feel that measurements and models for photopotentials developed are still inadequately standardized, especially considering that photopotential is a very fundamental quantity. Some models for photopotential infer the

magnitude of the photopotential primarily based on band edges [37], whereas many other estimates are based on quasi-Fermi levels [97]. It is also not clear whether the photopotential is entirely determined by thermodynamics factors or based on other kinetic factors; such issues have been explored in dye-sensitized solar cells more thoroughly [98]. Also, it is not clear if the *photopotential* generated upon exposure to sunlight amounts to an additional *electrochemical potential*, although the measurement of open-circuit voltage [29] and the definition of intrinsic solar to chemical conversion [56] suggests this to be the case. As photopotential is a strong indicator of the driving force for water splitting, it would be useful to characterize this very fundamental physicochemical property as accurately as possible.

Heavily Doped Semiconductors

In the context of optoelectronics devices, what is *chemically* pure is *electronically* impure as just a few percent of chemical dopants/vacancies make a semiconductor heavily doped. In a heavily doped semiconductor, the depletion width decreases significantly and with it, the driving force that exists at the semiconductor–electrolyte interface to separate e^-/h^+ is also much suppressed. While the theory of heavy doping is well developed in the context of semiconductor devices, it has been addressed to much lesser extent in PEC systems and this gap needs to be filled for more fundamental understanding of PEC principles.

Interfacial Structure and Dynamics of Surface State

It can be anticipated that the crystal coordination structure near the electrolyte interface may be different from the bulk crystal structure. For instance, computations suggest ZnO, which is fourfold coordinated in the bulk, can form a sixfold coordinated structure near the electrolyte interface [99]. These "near-surface non-native structures" are hard to characterize experimentally, but if proved, they will have important consequence in terms of band alignments, band–band bending, and surface electrochemical properties. In general, the interfacial structure is expected to be rough. The transition from being a semiconductor to an electrolyte might be gradual, and it is not clear how this will influence the band bending and other fundamental properties of PEC systems. While recent DFT studies have advanced our understanding of surface states [51], the picture is still a static picture wherein the changes in surface atomic structure to diffusion of reactants/products is not taken into account. If such changes are incorporated, it may be possible to obtain a more detailed dynamic view of surface states and their influence on PEC processes.

Concluding Remarks

The endeavor of a chemical approach to solar energy-driven generation of renewable fuels/chemicals/hydrogen has been rich in science, promised immense engineering consequence, and has had infrequent unraveling of "breakthrough" materials; all these have kept the quest of solar-based hydrogen alive for more than four decades. Within this period, it is not uncommon to observe that a generation of scientists who have entered the field anew pose a surreal recycle of questions, whose answers may have been partially reposed by an earlier generation of scientists, who have moved on to the other fields. The apparent simplicity is beguiling as it hides the complex convolution of physicochemical events that a single material interface is supposed to accomplish. From classical semiconductor materials like Si to the newest class of materials, the organic–inorganic perovskites, materials are being investigated via newer avatars (like nanostructures) with the hope that the three fundamental properties—optical absorption, electron–hole separation, and surface electrochemistry—which are intricately coupled, are better optimized within a single material interface. However, it is not clear whether too much is being asked from a single material interface. A partial decoupling of the aforementioned three physicochemical events is possible via solar cells driving an electrolyzer, and this has been a possibility for some time now for achieving significant amount of solar-based hydrogen, but for the economics and a probable lack of scientific intrigue. The gradual

progression in understanding is hoped to pave a natural pathway for the realization of artificial photosynthesis in the not too distant future. Even if the solar-based hydrogen does not live up to its promise as a competitive energy source, we hope that it may still serve as a source of "renewable reductant" to facilitate large-scale chemical reduction of CO_2.

ACKNOWLEDGMENTS

Almost all of the work undertaken in our group in this area is in collaboration with Dr. Sri Sivakumar, Department of Chemical Engineering, IIT Kanpur and his contributions to our work have been invaluable. Discussions with professor B. Viswanathan, National Center of Catalysis Research, IIT Madras have always been enlightening. Some parts of the work have been performed in collaboration with professors. Malay Das and K. Muralidhar, Department of Mechanical Engineering, IIT Kanpur. A host of students, Arun Kumar Upadhay, Dilip Kumar Behara, Babu Radhakrishnan, Gyan Prakash Sharma, Ashok Ummireddi, Gyan Maurya, Sayantan Sasmal, Sulay Saha, and Koshal Kishor continue to contribute to my understanding in this field, for which I am very thankful. This work would not have been possible without the support of the Department of Science and Technology, India and the Indian Space Research Organization.

REFERENCES

1. Yergin, D., *The Prize*, 2008: Free Press, New York.
2. Gratzel, M., Photoelectrochemical cells. *Nature*, 2001, 414(6861): 338–344.
3. McFarland, E.W., Solar energy: Setting the economic bar from the top-down. *Energy Environ. Sci.*, 2014, 7, 846–854.
4. Bockris, J.O.M., and A.K.N. Reddy, *Modern Electrochemistry 2B: Electrodics in Chemistry, Engineering, Biology and Environmental Science*. 2nd ed, 2001: Kluwer Academic Publishers, Dordrecht, The Netherlands.
5. Deutch, J., and R.K. Lester, *Making Technology Work*, 2004: Cambridge University Press, Cambridge, UK.
6. Smalley, R.E., Future global energy prosperity: The terawatt challenge. *MRS Bull.*, 2005, 30(June): 412–418.
7. Olah, G.A., A. Goeppert, and G.K. Surya Prakash, *Beyond Oil and Gas: The Methanol Economy*. 2005: Wiley-VCH, Weinhelm, Germany.
8. Memming, R., *Semiconductor Electrochemistry*. 2nd ed, 2015: Wiley-VCH, Weinhelm, Germany.
9. Van de Krol, R., and M. Gratzel, *Photoelectrochemical Hydrogen Production*, 2012: Springer, New York.
10. Grimes, C.A., O.K. Varghese, and S. Ranjan, *Light, Water, Hydrogen*, 2008: Springer, New York.
11. Rajeshwar, K., R. McConnell, and S. Licht, eds. *Solar Hydrogen Generation*, 2008: Springer.
12. Viswanathan, B., and M.A. Scibioh, *Photoelectrochemistry*, 2014: Narosa Publishing House, New Delhi, India.
13. Pleskov, Y.V., and Y.Y. Gurevich, *Semiconductor Photoelectrochemistry*, 1986: Springer, New York.
14. Vayssieres, L. (ed.), *On Solar Hydrogen and Nanotechnology*, 2009: John Wiley & Sons, Singapore.
15. Lewis, N.S., Research opportunities to advance solar energy utilization. *Science*, 2016, 351(6271): 353.
16. Nocera, D.G., The artificial leaf. *Acc. Chem. Res.*, 2012, 45(5): 767.
17. Vesborg, P.C.K., and T.F. Jaramillo, Addressing the terawatt challenge: Scalability in the supply of chemical elements for renewable energy. *RSC Adv.*, 2012, 2(21): 7933.
18. Wadia, C., A.P. Alivisatos, and D.M. Kammen, Materials availability expands the opportunity for large-scale photovoltaics deployment. *Environ. Sci. Technol.*, 2009, 43(6): 2072.
19. Ciamician, G., The photochemistry of the future. *Science*, 1912, 36(926): 385.
20. Verne, J., *The Mysterious Island*, 1874: Mass Market Paperback, Signet Press, Kolkata, India.
21. *Hydrogen Production*. [cited 2016 May 15]; Available from: http://energy.gov/eere/fuelcells/hydrogen-production.
22. Pinaud, B.A. et al., Technical and economic feasibility of centralized facilities for solar hydrogen production via photocatalysis and photoelectrochemistry. *Energy Environ. Sci.*, 2013, 6(7): 1983.
23. Agrawal, M., R. Mishra, and R.G.S. Pala, Technoeconomic analysis of solar H_2 production in the vicinity of Indian refineries. *Indian Chem. Eng.*, 2014, 56(3): 258.

24. Bard, A., *Electrochemical Methods*. 2nd ed, 1980: John Wiley & Sons, Oxford, UK.
25. Newman, J., and K.E. Thomas-Alyea, *Electrochemical Systems*. 3rd ed, 2004: Wiley-Interscience, Hoboken, NJ.
26. Norskov, J.K. et al., *Fundamental Concepts in Heterogeneous Catalysis*, 2014: Wiley, Hoboken, NJ.
27. Goodstein, D.L., *States of Matter*, 2014: Dover, Toronto, ON.
28. Tan, M.X. et al., Principles and applications of semiconductor photoelectrochemistry. *Prog. Inorg. Chem.*, 1994, 41: 21–144.
29. Walter, M.G. et al., Solar water splitting cells. *Chem. Rev.*, 2010, 110(11): 6446.
30. Martin, R.M., *Electronic Structure: Basic Theory and Practical Methods*, 2004: Cambridge University Press, Cambridge, UK.
31. Lejaeghere, K. et al., Reproducibility in density functional theory calculations of solids. *Science*, 2016, 351(6280): 1415.
32. Michaelides, A. et al., Preface: Special topic section on advanced electronic structure methods for solids and surfaces. *J. Chem. Phys.*, 2015, 143(10): 2.
33. Landis, D.D. et al., The computational materials repository. *Comput. Sci. Eng.*, 2012, 14(6): 51–57.
34. Wu, Y., and G. Ceder, First principles study on Ta_3N_5:$Ti_3O_3N_2$ solid solution as a water-splitting photocatalyst. *J. Phys. Chem. C*, 2013, 117(47): 24710–24715.
35. Wu, Y. et al., First principles high throughput screening of oxynitrides for water-splitting photocatalysts. *Energy Environ. Sci.*, 2013, 6(1): 157–168.
36. Valdes, A. et al., Solar hydrogen production with semiconductor metal oxides: New directions in experiment and theory. *Phys. Chem. Chem. Phys.*, 2012, 14(1): 49–70.
37. Valdes, A. et al., Oxidation and photo-oxidation of water on TiO_2 surface. *J. Phys. Chem. C*, 2008, 112(26): 9872.
38. Bendavid, L.I., and E.A. Carter, First-principles predictions of the structure, stability, and photocatalytic potential of Cu_2O surfaces. *J. Phys. Chem. B*, 2013, 117(49): 15750–15760.
39. Carter, E.A., *Photocatalyst Design Concepts from Molecular Biology and Quantum Mechanics*. Abstracts of Papers of the American Chemical Society, 2009. Vol. 238.
40. Baltrusaitis, J. et al., Photoelectrochemical hydrogen production on alpha-Fe_2O_3 (0001): Insights from theory and experiments. *ChemSusChem*, 2014, 7(1): 162–171.
41. Hellman, A., and R.G.S. Pala, First-principles study of photoinduced water-splitting on Fe_2O_3. *J. Phys. Chem. C*, 2011, 115(26): 12901.
42. Iandolo, B. et al., Correlating flat band and onset potentials for solar water splitting on model hematite photoanodes. *RSC Adv.*, 2015, 5(75): 61021–61030.
43. Iandolo, B. et al., The rise of hematite: Origin and strategies to reduce the high onset potential for the oxygen evolution reaction. *J. Mater. Chem. A*, 2015, 3(33): 16896–16912.
44. Viswanathan, V., and H.A. Hansen, Unifying solution and surface electrochemistry: Limitations and opportunities in surface electrocatalysis. *Top. Catal.*, 2014. 57(1–4): 215–221.
45. Castelli, I.E. et al., New light-harvesting materials using accurate and efficient bandgap calculations. *Adv. Energy Mater.*, 2015, 5(2): 9034–9043.
46. Castelli, I.E. et al., New cubic perovskites for one- and two-photon water splitting using the computational materials repository. *Energy Environ. Sci.*, 2012, 5(10): 9034–9043.
47. Karlberg, G.S. et al., Cyclic voltammograms for H on Pt(111) and Pt(100) from first principles. *Phys. Rev. Lett.*, 2007, 99(12) 126101-1–126101-4.
48. Viswanathan, V., F. Wang, and H. Pitsch, Monte Carlo-based approach for simulating nanostructured catalytic and electrocatalytic systems. *Comput. Sci. Eng.*, 2012, 14(2): 60–69.
49. Viswanathan, V. et al., Simulating linear sweep voltammetry from first-principles: Application to electrochemical oxidation of water on Pt(111) and $Pt_3Ni(111)$. *J. Phys. Chem. C*, 2012, 116(7): 4698–4704.
50. Bertoluzzi, L. et al., Charge transfer processes at the semiconductor/electrolyte interface for solar fuel production: Insight from impedance spectroscopy. *J. Mater. Chem. A*, 2016, 4(8): 2873–2879.
51. Iandolo, B., and A. Hellman, The role of surface states in the oxygen evolution reaction on hematite. *Angew. Chem., Int. Ed. Engl.*, 2014, 53(49): 13404.
52. Viswanathan, V. et al., Simulating linear sweep voltammetry from first-principles: Application to electrochemical oxidation of water on Pt(111) and $Pt_3Ni(111)$. *J. Phys. Chem. C*, 2012, 116(7): 4698.
53. Sprunken, H.R., R. Schumacher, and R.N. Schindler, Evaluation of the flat-band potentials by measurments of anodic-cathodic photocurrent transitions. *Faraday Discuss.*, 1980, 70: 55.
54. Helander, M.G. et al., Work function of fluorine doped tin oxide. *J. Vac. Sci. Technol., A*, 2011, 29(1): 4.
55. Greiner, M.T. et al., Universal energy-level alignment of molecules on metal oxides. *Nat. Mater.*, 2012, 11(1): 76.

56. Dotan, H. et al., On the solar to hydrogen conversion efficiency of photoelectrodes for water splitting. *J. Phys. Chem. Lett.*, 2014, 5(19): 3330.
57. Li, J., and N. Wu, Semiconductor-based photocatalysts and photoelectrochemical cells for solar fuel generation: A review. *Catal. Sci. Technol.*, 2015, 5(3): 1360–1384.
58. Behara, D.K. et al., Coupled optical absorption, charge carrier separation, and surface electrochemistry in surface disordered/hydrogenated TiO_2 for enhanced PEC water splitting reaction. *Phys. Chem. Chem. Phys.*, 2016, 18(12): 8364.
59. Chen, Z., H. Dinh, and E. Miller, *Photoelectrochemical Water Splitting: Standards, Experimental Methods, and Protocols*, 2013: Springer, New York.
60. Hodes, G., Photoelectrochemical cell measurements: Getting the basics right. *J. Phys. Chem. Lett.*, 2012, 3(9): 1208–1213.
61. Seitz, L.C. et al., Modeling practical performance limits of photoelectrochemical water splitting based on the current state of materials research. *ChemSusChem*, 2014, 7(5): 1372.
62. Walczak, K. et al., Modeling, simulation, and fabrication of a fully integrated, acid-stable, scalable solar-driven water-splitting system. *ChemSusChem*, 2015, 8(3): 544–551.
63. Singh, M.R. et al., An electrochemical engineering assessment of the operational conditions and constraints for solar-driven water-splitting systems at near-neutral pH. *Energy Environ. Sci.*, 2015, 8(9): 2760–2767.
64. Osterloh, F.E., Inorganic materials as catalysts for photochemical splitting of water. *Chem. Mater.*, 2007, 20(1): 35.
65. Zhu, J.F., and M. Zach, Nanostructured materials for photocatalytic hydrogen production. *Curr. Opin. Colloid Interface Sci.*, 2009, 14(4): 260.
66. Chen, X.B. et al., Semiconductor-based photocatalytic hydrogen generation. *Chem. Rev.*, 2010, 110(11): 6503.
67. Kronawitter, C.X. et al., A perspective on solar-driven water splitting with all-oxide hetero-nanostructures. *Energy Environ. Sci.*, 2011, 4(10): 3889.
68. Sivula, K., and R.V. De Krol, Semiconducting materials for photoelectrochemical energy conversion. *Nat. Rev. Mater.*, 2016, 1: 15010–15027.
69. Chen, Z. et al., Accelerating materials development for photoelectrochemical hydrogen production: Standards for methods, definitions, and reporting protocols. *J. Mater. Res.*, 2010. 25(1): 3.
70. Maeda, K., and K. Domen, Photocatalytic water splitting: Recent progress and future challenges. *J. Phys. Chem. Lett.*, 2010, 1(18): 2655.
71. Fujishima, A., and K. Honda, Electrochemical photolysis of water at a semiconductor electrode. *Nature*, 1972, 238(5358): 37–38.
72. McCrory, C.C.L. et al., Benchmarking heterogeneous electrocatalysts for the oxygen evolution reaction. *J. Am. Chem. Soc.*, 2013, 135(45): 16977–16987.
73. Kishor, K. et al., Roughened Zn-doped Ru-Ti oxide water oxidation electrocatalysts by blending active and activated passive components. *ChemElectroChem*, 2015, 2(11): 1839–1846.
74. Liu, G. et al., Titania-based photocatalysts-crystal growth, doping and heterostructuring. *J. Mater. Chem.*, 2010, 20(5): 831.
75. Chen, X.B. et al., Increasing solar absorption for photocatalysis with black hydrogenated titanium dioxide nanocrystals. *Science*, 2011, 331(6018): 746–750.
76. Oregan, B., and M. Gratzel, A low-cost, high-efficiency solar-cell based on dye-sensitized colloidal TiO_2 films. *Nature*, 1991, 353(6346): 737–740.
77. Sun, K. et al., Enabling silicon for solar-fuel production. *Chem. Rev.*, 2014, 114(17): 8662.
78. Boettcher, S.W. et al., Photoelectrochemical hydrogen evolution using Si microwire arrays. *J. Am. Chem. Soc.*, 2011, 133(5): 1216–1219.
79. Dasgupta, N.P. et al., 25th anniversary article: Semiconductor nanowires synthesis, characterization, and applications. *Adv. Mater.*, 2014, 26(14): 2137.
80. Liu, B. et al., All inorganic semiconductor nanowire mesh for direct solar water splitting. *ACS Nano*, 2014, 8(11): 11739–11744.
81. *Perovskite Photovoltaics (MRS Bulletin)*, Vol. 40, 2015: Cambridge University Press, Cambridge, UK.
82. Berry, J. et al., Hybrid organic-inorganic perovskites (HOIPs): Opportunities and challenges. *Adv. Mater.*, 2015, 27(35): 5102–5112.
83. Crystal clear? *The Economist*, 2015, Economist Group: London.
84. Loo, L., P. Patel, and D. Mitzi, Perovskite photovoltaics: David Mitzi addresses the promises and challenges. *MRS Bull.*, 2015, 40(8): 636.

85. Kim, J.H. et al., Wireless solar water splitting device with robust cobalt-catalyzed, dual-doped $BiVO_4$ photoanode and perovskite solar cell in tandem: A dual absorber artificial leaf. *ACS Nano*, 2015, 9(12): 11820–11829.
86. Gurudayal, D.S. et al., Perovskite-hematite tandem cells for efficient overall solar driven water splitting. *Nanoletters*, 2015, 15(6): 3833.
87. Da, P. et al., High-performance perovskite photoanode enabled by Ni passivation and catalysis. *Nanoletters*, 2015, 15: 3452–3457.
88. Pandey, M., and R.G.S. Pala, Stabilization and growth of non-native nanocrystals at low and atmospheric pressures. *J. Chem. Phys.*, 2012, 136(4): 044703-1–044703-6.
89. Pandey, M., and R.G.S. Pala, Stabilization of rocksalt CdSe at atmospheric pressures via pseudomorphic growth. *J. Phys. Chem. C*, 2013, 117(15): 7643–7647.
90. Behara, D.K. et al., Design of photoelectrohemical materials via non-native nanostructures and their "Click" assembly into photoreactor. *SMC Bull.*, 2014, 5(2): 49–58.
91. Vannice, M.A., *Kinetics of Catalytic Reactions*, 2005: Springer, New York.
92. Thomas, J.M., and R. Raja, The advantages and future potential of single-site heterogeneous catalysts. *Top. Catal.*, 2006, 40(1–4): 3–17.
93. Guerrini, E., and S. Trasatti, Recent developments in understanding factors of electrocatalysis. *Russ. J. Electrochem.*, 2006, 42(10): 1017.
94. Pala, R.G.S., and H. Metiu, Modification of the oxidative power of ZnO(101h0) surface by substituting some surface Zn atoms with other metals. *J. Phys. Chem. C*, 2007, 111: 8617.
95. McFarland, E.W., and H. Metiu, Catalysis by doped oxides. *Chem. Rev.*, 2013, 113: 4391–4427.
96. CO oxidation by Ti- and Al-doped ZnO: Oxygen activation by adsorption on the dopant. *J. Catal.*, 2009, 266: 50–58.
97. Vayssieres, L., ed. *On Solar Hydrogen and Nanotechnology*, 2009, John Wiley & Sons, Singapore.
98. Pichot, F., and B.A. Gregg, The photovoltage-determining mechanism in dye-sensitized solar cells. *J. Phys. Chem. B*, 2000, 104: 6–10.
99. Pandey, M., and R.G.S. Pala, Hydroxylation induced stabilization of near-surface rocksalt nanostructure on wurtzite ZnO structure. *J. Chem. Phys.*, 2013, 138(22): 224701-1–224701-6.

4.33 PASSIVE AND ACTIVE SOLAR DISTILLATION

Desh Bandhu Singh and G.N. Tiwari

Introduction

Pure water, clean environment, and electricity are considered as the basic needs for the survival of human beings on the planet earth. The available sources of water are rivers, lakes, seawater, icebergs, and underground water reservoirs. Figure 4.33.1 shows the distribution of water on the surface of the earth. Water covers more than 67% of the earth's surface. However, the major portion (97%) of this world's water is salty. Less than 3% of available water on earth is freshwater. Out of this freshwater, more than 2% water is locked up in ice caps and glaciers of Polar Regions and a very small amount (only <1%) is within human reach which can be used for drinking, agriculture, and industrial purposes. This small quantity of water is believed to be sufficient to support life and vegetation on earth. However, the available freshwater is fixed on earth and its demand increases day by day due to increase in population. It leads to an essential and earnest need to get freshwater from saline/brackish water present on the planet earth.

According to the United Nations (2004) report, more than one billion people are not having access to fresh drinking water and it has been analyzed that 2/3rd population of the world will face a shortage of water from 2025 onward. As per UNICEF (2004) report, more than 1/6th people lack access to safe drinking water and more than 1/3rd people lack adequate sanitation. The driving force for the requirement of freshwater is domestic, agriculture, and industrial areas. The use of potable water resources fluctuates from country to country. Figures 4.33.2 and 4.33.3 present the use of potable water requirement in various sectors of low- and high-income group countries, respectively. High-income group countries have higher water requirement for industries and household sectors in comparison to that of low-income countries. However, agriculture sector needs the highest percentage of potable water in both low- as well as high-income group countries.

A number of water treatment technologies are available. However, distillation using solar still which takes input as sunlight is an environmental friendly alternative for getting freshwater in the far-flung area for cleaning underground brackish water (salinity ~10,000 ppm). It is a renewable, eco-friendly, easy to operate, and low-cost technology. Hence, it can be a promising solution to the problems of shortage of drinking water, energy security, and climate change.

Solar distillation systems can effectively remove many impurities ranging from salts to microorganisms. It can remove impurity of water mainly salinity, iron, manganese, fluorides, heavy metals, bacterial contamination, and pesticide/herbicide. As per World Health Organization, the allowable limit of salinity in water is 500 ppm and it can be up to 1000 ppm in special cases. The salinity level of most of the water available on earth in terms of total dissolved salts is up to 10,000 ppm. However, seawater normally has salinity in the range of 35,000–45,000 ppm. The excessive brackishness affects the taste, causes the stomach problems and other diseases. The permissible salinity limits of 500 ppm or less can be accomplished by the solar distillation system to desalinate water for remote habitation where salty or brackish water is available for survival.

The solar distillation can be useful for the places where power is not available and demand is <200 m^3/day. The setting up of water pipelines is uneconomical and delivery by truck is unreliable and expensive on the hilly area. The other desalination plants are uneconomical for low-capacity potable water demands at such places. Under these situations, solar stills which are self-sustainable can ensure regular supply of water. All conventional desalination plants need fossil fuels either directly (thermal) or indirectly (conversion to electricity). More than half the population of the world depends on desalinated seawater. This scenario is not sustainable for the future. Using solar thermal energy for desalination is environmental friendly, free, and sustainable.

Thus the purification of polluted water by solar distillation is the most prominent method. The solar distillation system can be used anywhere with lesser number of problems as it requires

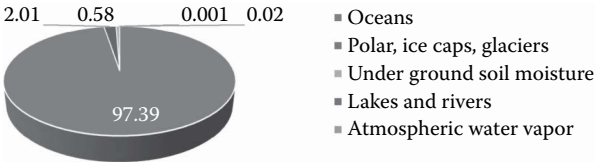


FIGURE 4.33.1 Availability of water on the earth surface (percentage).

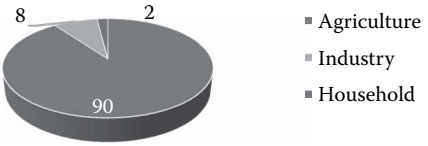


FIGURE 4.33.2 Use of freshwater requirements in low-income countries.

FIGURE 4.33.3 Use of freshwater requirements in high-income countries.

comparatively simple technology, low maintenance, and only sunshine to operate. It is suitable for places where (1) hot climatic conditions exist during major part of the year; (2) water transportation cost is high; (3) fuel supply is expensive and unreliable; (4) rainfall is <50 cm per year; and (5) only saline/brackish water is available.

Distillation Process

When the solar energy falls on the surface of brackish/saline water, the temperature of water rises; it gets evaporated and vapor is formed. Further, the vapor condenses and distillate will come as an output. This is an open cycle process as shown in Figure 4.33.4. However, solar still is based on a closed cycle.

Basic Principle of Solar Distillation

Solar distillation technology uses solar radiation as the source of heat energy to obtain potable water with the help of a device known as the solar still. It is based on the principle of greenhouse effect and evaporation/condensation of water. Figure 4.33.5 represents the schematic diagram of conventional single-slope passive solar still. It is a box-type structure made up of fiber-reinforced plastic (FRP), galvanized iron (GI) sheet, wood or concrete covered with some insulation to prevent the heat loss from the bottom and side walls. The box is covered with a glass cover which is inclined

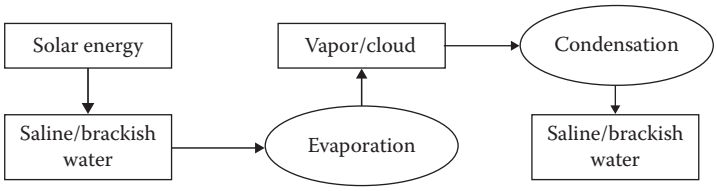


FIGURE 4.33.4 Distillation process (hydrogen open cycle).

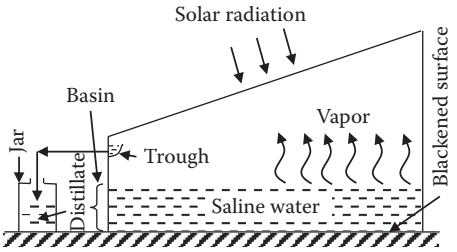


FIGURE 4.33.5 Conventional single-slope passive solar still (south oriented).

with the horizontal surface to facilitate the trickling down of water condensed on the inner surface. The condensed water at the inner surface of glass moves down due to the component of gravitational force acting along the glass cover. The bottom surface of the box is blackened from inside to absorb almost all radiation falling on it. This blackened (painted black) surface is generally known as basin liner.

The major portion of solar radiation falling on the surface of glass cover is transmitted to water mass in the basin after reflection and absorption by it. A small portion of solar radiation reaching the water surface is reflected back and absorbed by water mass and remaining portion is transmitted to basin liner where it is mostly absorbed. The attenuation of solar radiation in water mass depends on its depth and absorptivity. The temperature of basin liner increases which further transmits most of its heat to water mass by convection and a small portion of heat is lost to the atmosphere by conduction. Hence, the temperature of water increases which creates a temperature difference between water and glass surface. Heat is transferred from water surface to glass cover by three modes namely evaporation, convection, and radiation. The evaporated water undergoes film-type condensation after releasing the latent heat at the inner surface of the glass cover. The condensed water trickled down the glass cover and it is collected in a trough fixed at the lower end of the glass cover. The collected water in the trough is allowed to flow to the jar through a pipe and it is taken out for further use. Here, it should be noted that the distilled water obtained is not fit for direct use in drinking and cooking application as it lacks the minerals. However, it can be made suitable for these applications after adding the required minerals needed for human beings. The thermal energy received by glass cover by convection, radiation, and latent heat is lost to the atmosphere by convection and radiation. The daily yield from single-slope passive solar still may vary from 0.5 L/m^2 (winter) to 2.5 L/m^2 (summer).

The advantages of solar distillation system/solar still over other desalination technologies are as follows:

1. It contains no moving parts to run the system in passive mode.
2. It is an easy and cost-effective technology to provide potable water for the domestic and commercial level.
3. It requires only solar energy for its operation which is renewable and nonpolluting.
4. No skilled labor is required for its operation and maintenance.
5. It is possible to manufacture and repair locally.
6. It gives a sense of satisfaction in owning a trusted and easy to use water treatment plant.
7. It is very useful for remote settlements where only saline water is available.
8. It is self-sustainable and it can also meet electricity requirement up to certain extent when operating in active mode integrated with the photovoltaic (PV) module.

However, a solar distillation system has some demerits also which sometimes may limit the use of technology for large-scale production. It includes (1) dependency on weather for input energy, (2) requirement of large solar collection area for commercial use, (3) low interest of producer, (4) low

efficiency, (5) comparatively higher initial investment, (6) need of frequent filling and flushing of solar still, and (7) requirement of system to be robust as it is installed in open space.

The application of solar distillation technology includes (1) production of potable water as well as DC electric power for remote areas where electricity has not reached so far, (2) hospitals, (3) automobile garage, (4) workshops, (5) laboratories, (6) educational institutions, and (7) cleaning of kitchen utensils and mechanical parts.

Material Requirement of Solar Still

Although, the local material can be used for manufacturing solar still if we need a cheaper and insulating material and want to facilitate necessary repairs. However, solar still manufactured from such material will not last long. Also, a high-quality material will increase the cost of solar still but at the same time, it will survive for a longer time. In such case, one should introspect whether one need an inexpensive and short-lived solar still that requires replacement or repair every few years, or manufacture something more durable and long lasting with the hope that the distilled water produced will be cheaper in the long run.

The material used for making the solar still should be able to bear a long life and it should also be inexpensive enough so that it can be replaced after degradation. It must be strong enough to bear minor earth movements and wind pressure. It should be nontoxic. It should not release vapors or encourage an unpleasant taste to the water at elevated temperatures. It should be corrosion resistant against saline water and distilled water. It should be compact enough and easy to handle so that it can be carried by local transportation. The film-type condensation should occur at the inner surface of glass cover as the formation of water droplets there will reduce the intensity of radiation falling on the surface of water because water droplets act as a mirror which will reflect the solar intensity. Solar still may be constructed using concrete/cement, GI sheet, FRP with the top cover of transparent material like glass, plastic, etc.

Classification of Solar Distillation System/Solar Still

Solar still also known as solar distillation system is a device that converts saline/brackish water into potable water with the help of solar energy. It imitates the natural distillation process with the difference that it operates in a closed loop. It can be classified into two parts namely passive solar still and active solar still. Saline or brackish water is directly fed into the basin in passive solar stills (Kudish et al., 1982; Delyannis and Delyannis, 1983; Tiwari et al., 1986; Yadav and Tiwari, 1987; Clark, 1990) but for active solar stills, there is an additional source of heat to provide thermal energy. This thermal energy is added to the saline/brackish water into the basin of passive solar still to increase the rate of evaporation.

The additional source of heat may be included through collector/concentrator (Rai and Tiwari, 1983; Fernandez and Chargoy, 1990; Lawrence and Tiwari, 1990; Zaki et al. 1992; Kudish et al., 2003) or waste heat energy from any industrial plant (Tleimat and Howe, 1966), etc. Further, it can be divided into two parts namely single-slope and double solar still on the basis of its structure and positioning. The orientation of single-slope solar still is due south and for double-slope solar still orientation of one glass cover is due east and other is kept due west to collect maximum solar radiation. The orientation of single-slope solar still is due south for the southern globe and due north for the northern globe. The complete classification of solar distillation systems has been given in Figure 4.33.6.

The double-slope solar still (passive) may be symmetrical or unsymmetrical. In symmetrical-type double-slope solar still (passive), both glass covers have the same area and their angles of inclination with horizontal are also the same. However, the orientation of solar still glass covers is dissimilar. No external energy is given. In unsymmetrical-type double-slope solar still (passive), both glass covers do not have the same area and the angle of inclination with horizontal is not the same.

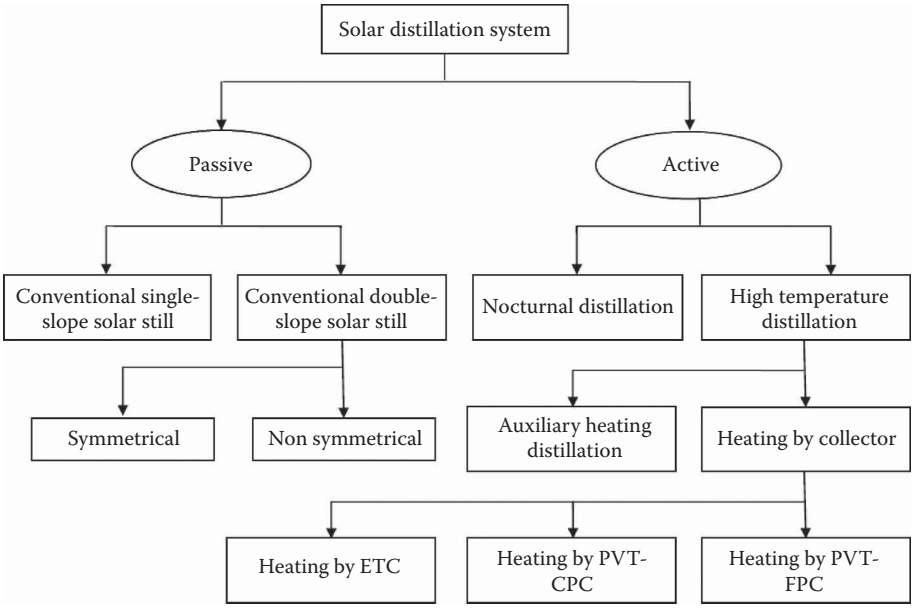


FIGURE 4.33.6 Classification of solar distillation system.

The active solar still can be divided into two parts namely nocturnal distillation and high-temperature distillation. In nocturnal distillation, solar energy is used in day time and waste thermal energy of industrial plant is used at night when the solar energy is not available. Waste heat from heat treatment plant (food, chemical and textile industry, etc.) is transmitted to the solar still with the help of heat exchanger.

Further, high-temperature distillation is divided into auxiliary heating distillation and heating by collectors (photovoltaic thermal (PVT)). The flat plate collector (FPC) may be alone or it may be fully or partially covered by photovoltaic (PVT-FPC). Further, PVT-FPC may be integrated with parabolic concentrator in which case, it is known as photovoltaic thermal-compound parabolic concentrator (PVT-CPC) collector. FPC is alone used for thermal energy generation which is supplied to the solar still. To develop a self-sustainable system, FPC is fully or partially covered by PV due to which both thermal energy and electrical energy are generated. Thermal energy is supplied to solar still and some part of electrical energy is used to run the pump for water circulation. The rest part of the electrical energy may be used for storage/other application. PVT-FPC is integrated with parabolic concentrator to reflect the beam radiation to receiver area.

Historical Background

Historically the earliest documented work on solar distillation is that of Arab alchemists in the 16th century (Mouchot, 1869). According to Arab alchemists, polished Damascus concave mirrors could be used for solar distillation. Wide earthen pots were used for distillation (Della Porta, 1589). These earthen pots were kept in sunlight to receive the solar radiation to heat, water got evaporated and condensate was collected into the container placed underneath (Nebbia and Menozzi, 1966). In 1862, Lavoisier, a French chemist, made use of large glass lenses fitted on supporting structure with an aim to converge solar energy on to the distillation flask filled with water. Mouchot (1869) has also described the use of silver- or aluminum-coated glass reflectors to concentrate solar energy for distillation purpose.

The solar still was first designed and constructed in 1872 by Carlos Wilson in Northern Chile. It was actually a basin-type solar still designed to produce potable water for meeting the drinking

water needs of the mining community in residing in Las Salina area. The system was constructed using a series of wooden bays (each of dimension 1.14 m × 61.0 m) making up an overall surface area of about 4700 m^2. The top portion was covered with glass and the bottom portion was painted black with logwood dye and alum for increasing the absorption of sunlight. The bays were filled with saline water and the same was allowed to evaporate due to solar radiation falling on it through the glass. It then condensed over the glass cover and trickled down into the container kept to store distilled water. This device worked satisfactorily for almost four decades and the daily yield of about 4.9 kg/m^2 was attained in summer season (Harding, 1883).

Not much work on solar distillation was documented between the 1880s and the end of First World War due to slow research progress in the field of solar energy. Then, with the revitalization of interest, a number of solar-based systems viz. roof type, V-covered, tilted wick, inclined tray, suspended envelope, tubular stills, etc., were reported. Solar concentrators made of metal-coated reflectors were developed for solar distillation applications (Kausch, 1920). Further, Pasteur, in 1928, reported the use of a number of concentrators to reflect the solar radiation on to a water-filled copper boiler. The steam from the boiler was then allowed to flow to the distilled water collector through a water-cooled condenser. Use of aluminum-coated cylindrical parabolic reflectors for focusing the solar energy on evacuated water tubes was reported in the literature (Abbot, 1938). An efficiency (explained in Section 4.33.9.3) of as high as 80% could be attained. Boiling of water in tubes, however, created certain problems. During Second World War, an air inflated still made of plastic was developed by Telkes (1945) for US defense forces. US Navy used as many as 200,000 of these stills during the war.

As a next stage, the forced air circulation in stills was tried with an aim to increase the vapor condensation rate so that higher distillate might be obtained. Many researchers studied the exploitation of latent heat of vaporization in multieffect systems and preheating the brine to increase the distillate output. Many researchers have suggested the integration of desalination of water with electric power generation to increase the overall operating efficiency of the plant.

The basin-type solar stills, also called greenhouse roofs, are in an advanced stage of evolution. A number of researchers have reported the effect of design, operational, and climatic parameters on the performance of such stills. Cooper investigated the performance of basin-type still and concluded that the maximum efficiency of such a system that could be achieved is 60% (Cooper, 1969a,b, 1973a,b). Frick proposed a mathematical model based on sine wave heat flow assumption and employed thermic circuits and Sankey diagrams for modeling (Frick, 1970). Hirschman and Roefler (1970) studied the relationship between of heat capacity and the performance of still considering periodic insolation. Further, Baum et al. (1970), Nayak et al. (1980), and Sodha et al. (1980) have also presented the periodic and transient analysis.

Recently, many other types of solar stills have also been proposed and studied. Talbert et al. (1970) presented a review on the solar still. Delyannis and Delyannis (1973) presented the review of major solar distillation units around the world. Oltra (1972) investigated multieffect solar stills. A tilted tray or inclined–stepped solar stills were proposed by Howe (1961) and Akhtamov (1978). Frick and Sommerfeld (1973), Sodha et al. (1981), and Moustafa et al. (1979) studied the tilted, wick-type and multiple wick-type solar distillation systems. Norov et al. (1975) and Menguy et al. (1980) have investigated the solar film covered still and wiping spherical stills. Soliman (1976), Malik and Tran (1973), and Malik et al. (1982) investigated indirectly heated solar stills. Researchers have also investigated several designs including reflectors, the effect of heat-storing materials, climatic, operational, and design parameters with modifications in the conventional-type solar stills for the last 40 years with significant improvement in terms of yield, efficiency, and reported several studies.

A review on passive solar still till 1982 was done by Malik et al. (1982) and it was updated up to 1992 by Tiwari (1992), which also incorporate active solar still. Rai and Tiwari (1983) studied the active solar still in forced mode theoretically for the first time and found that the daily yield was 24% higher than conventional solar still. At the same time, Zaki et al. (1983) studied the active solar still under natural circulation mode for the first time and reported a maximum increase in yield of

33% in comparison to conventional solar still. Tiwari and Dhiman (1991) investigated the solar still augmented with collectors experimentally and the thermal model was validated with experimental results. The heat exchanger was used in the basin for transferring the thermal energy of collector fluid (water) to water in the basin and they found the rise in yield by 12% if the length of the heat exchanger was increased by 100% (6–12 m). Lawrence and Tiwari (1990) developed an empirical relation for internal heat transfer coefficient on the basis of theoretical results for such systems. Hamadou and Abdellatif (2014) provided for heating the solar still at its bottom by a circulating heat transfer fluid and reported a considerable increase in productivity. They found that doubling the heat transfer fluid rate effected 9% improvement in distilled water yield. The relation between distilled water production and the heat transfer fluid rate is thus nonlinear.

The optimum number of collectors for maximum yield in the case of solar still coupled with collectors is 8 m^2 as the increase in gain becomes lower than thermal loss (Kumar and Tiwari, 1998). They have reported that the daily yield and overall thermal efficiency decline if basin area is increased which occurs because of the large heat content storage capacity for a given depth of water. Tiris et al. (1998) investigated basin-type solar still augmented with two flat plate solar collectors experimentally and reported an average increase of 100% in the yield of active solar still in comparison to conventional solar still. The productivity of active solar still decreases with increase in water depth; however, it increases with the increase in solar flux (Badran and Al-Tahainesh, 2005). Tripathi and Tiwari (2005) investigated the active solar still for the effect of water depth on internal heat and mass transfer and concluded that convective heat transfer depends on water depth. They also reported that the yield was higher during the off-sunshine hours than day time if the depth of water is more due to the heat content of water mass. Badran et al. (2005) studied single stage basin-type solar still augmented with FPC and reported 52% increase in yield in comparison to conventional solar still. Dwivedi and Tiwari (2010) investigated double-slope active solar still under natural circulation mode experimentally and reported that the thermal efficiency was lower and exergy efficiency was higher in comparison to conventional double-slope solar still.

Sinha et al. (1994) studied active solar distillation system as an investment alternative to a solar hot water system and concluded that the annual operating cost of the solar water heater was higher than that of the solar still. However, the cost of hot water was much less than the cost of distilled water. Kumar and Sinha (1996) studied the double-slope solar still integrated with nontracking-type cylindrical parabolic concentrator and concluded that the yield obtained was higher than FPC-coupled solar still. Kumar et al. (2000) found that for maximum annual yield, the optimum collector inclinations of a flat plate collector and still glass cover are 20° and 15°, respectively, for New Delhi climatic condition. Lourdes et al. (2002) made a comparative study of solar still integrated with solar pond, FPC, evacuated tube collector (ETC), compound parabolic collector, and parabolic trough collector for direct steam generation (DSG) and concluded that DSG was a promising technology for solar-assisted sea water desalination. Zuriat and Abu-arabi (2004) investigated the regenerative solar still and reported that the productivity of the regenerative solar still was marginally affected by the thickness of water at the top of the first condensing cover and the amount of water flowing per unit time going into the second effect. Tiwari and Tiwari (2005) calculated the annual performance of solar stills at different water depth and inclinations of condensing covers. Further, Tiwari et al. (2007) did the thermal modeling of basin-type solar still integrated with FPC, concentrating collector, ETC, and ETC with pipe and reported an overall thermal efficiency as 13.14%, 17.57%, 17.22%, and 18.26%, respectively. They found their average exergy efficiency in the range of 0.59%–1.82%. Abdel-Rehim and Lasheen (2007) studied solar still using heat exchanger and tracking concentrator experimentally and found an increase in productivity by 18% as compared to conventional solar still.

El-Sebaii et al. (2009) studied theoretically the performance of active single-basin solar still with a sensible storage material and concluded that daily productivity of the solar still with storage was 23.8% higher than that when it was used without storage. They used sand as the storing material. Arslan (2012) investigated the performance of various designs of active solar stills under closed

cycle mode experimentally and concluded that the circular box active solar still design produced highest overall daily efficiency. Taghvaei et al. (2014) investigated experimentally the long-term (more than 2 days) effects of water depth on the performance of active solar stills and recommended the higher depth of water in the basin for practical use of solar still. Lilian et al. (2014) have developed a solar distillation system in which a partly immersed slowly rotating low-weight hollow drum was led within the cavity of the solar still. They found an average improvement in water yield by 20%–30% in comparison to simple solar still. Rajaseenivasan et al. (2014) investigated solar still integrated with FPC and the basin of solar still was distributed into small sections. They found 60% increment in yield as compared to conventional solar still.

The integration of PVT with FPC results in the increase in efficiency of solar cell and utilization of residual thermal energy for distillation. The advantage is that comparatively higher electricity will be produced by PV module due to lower temperature and the system becomes self-sustainable. Hendrie (1979) presented a theoretical model on PVT systems using conventional thermal collector techniques. Bhargava et al. (1991) investigated the effect of air channel depth, length, air mass flow rate, and part of absorber plate area covered by solar cells (packing factor, PF) on a single pass. They concluded that the overall electrical efficiency of the PV module can be enhanced by increasing the packing factor and reducing the temperature of the PV module by withdrawing the thermal energy associated with the PV module. Tiwari et al. (2011) have reviewed various PV modules and their applications, especially PVT solar distillation system and thermal modeling for overall thermal energy and exergy analysis. Estahbanati et al. (2015) carried out a study to predict the effect of number of stages on productivity of multieffect active solar still in both the modes viz. continuous and noncontinuous. They concluded that the distillate production varies with number of stages in a quadratic fashion as we switch from noncontinuous to continuous mode, for instance, a five-stage system would produce 25% more freshwater in continuous mode than in noncontinuous mode.

Kumar and Tiwari (2008, 2009, 2010) designed and validated the solar still integrated with PVT-FPC and concluded that the yield of active solar still increased by more than 3.5 times than the passive solar still. They found the payback period of active solar still in the range of 3.9–23.9 years. The work of Kumar and Tiwari (2010) was further extended by Singh et al. (2011) for double-slope solar still. Eltawil and Omara (2014) investigated the performance of solar still by incorporating PVT, FPC, and hot air and found an increase in productivity by 51%–148% in comparison to conventional solar still. Calise et al. (2014) designed a novel poly-generation system employing photovoltaic/thermal collectors, absorption chiller, multieffect distiller, storage tank, auxiliary heater and balance of plant devices for seawater desalination in European Mediterranean countries, known to be rich in renewable sources but poor in fossil fuels and water resources. Tiwari et al. (2015) and Singh et al. (2016) designed, fabricated, and validated the single-slope PVT-FPC active solar distillation system and concluded that the thermal efficiency was lesser than the system proposed by Kumar and Tiwari (2010) and Singh et al. (2011). However, exergy efficiency and overall thermal efficiencies were found to be higher. Saeedi et al. (2015) have studied the optimization of PVT solar still taking energy efficiency as objective function and found the optimum value of mass flow rate and number of collectors as 0.044 and 7 kg/s. They used simulation technique for optimization. Singh et al. (2016) have presented the theoretical analysis of the effect of energy matrices on life cycle cost analysis of single- and double-slope passive solar stills for the same basin area under similar climatic condition and concluded that single slope performs better than double-slope on the basis of annual exergy gain, annual energy output, thermal efficiency, exergy efficiency, life cycle conversion efficiency (LCCE), energy payback time (EPBT) based on exergy, energy production factor (EPF) based on exergy and exergo-economic parameter. However, double performs better than single-slope passive solar stills on the basis of annual productivity, EPBT based on energy and EPF based on energy. Recently, Singh and Tiwari (2016) have studied single- and double-slope solar stills integrated with N identical partially covered PVT-CPC water collectors at 0.14 m water depth, optimum number of collectors

and mass flow rate for the same basin area under similar climatic condition to assess the life cycle cost by incorporating the effect of energy payback period. They have concluded that single slope performs better than double-slope PVT-CPC active solar distillation system on the basis of energy matrices and cost of water.

Heat Transfer in Solar Still

The performance of solar still is affected by all heat exchanges that define temperature of water, basin liner, inner surface of glass cover, outer surface of glass cover and yield. Here, it is interesting to note that there are two opposing phenomena. On one hand, the aim is to store maximum energy to increase the temperature like basin. On other hand, the aim is to lose maximum energy so that maximum production of distillate in the solar still can be ensured. The energy flow diagram of a conventional single-slope solar still with its equivalent thermal circuit diagram is shown in Figure 4.33.7 in which $\dot{q}$ and C represent rate of heat transfer and heat capacity, respectively. The various heat transfers and heat capacities at different stages such as water-glass, glass-ambient are water-ambient are considered in electrical analogy. The fraction of solar intensity at different components of solar still unit (Tiwari et al., 1989) can be written mathematically as follows:

Fraction of solar intensity absorbed by glass cover, $\alpha'_g = (1-R_g)\alpha_g$

Fraction of solar intensity reflected by water mass, $R'_w = (1-R_g)(1-\alpha_g)R_w$ Fraction of solar intensity absorbed by water mass, $\alpha'_w = (1-R_g)(1-\alpha_g)(1-R_w)\alpha_w$. Fraction of solar intensity absorbed by basin liner, $\alpha'_b = (1-R_g)(1-\alpha_g)(1-R_w)(1-\alpha_w)\alpha_b$. Fraction of solar intensity lost to the ambient through water and glass cover, $(1-R_g)(1-\alpha_g)(1-R_w)(1-\alpha_w)(1-\alpha_b)$, where R and α represent the reflectivity and absorptivity, respectively; g, w, and b are suffix for glass cover, water mass, and basin liner, respectively.

However, if attenuation of solar intensity within water mass is considered, then α'_w and α'_b is expressed as follows:

$$\alpha'_w = (1-R_g)(1-\alpha_g)(1-R_w)\left[1-\sum \mu_j \exp(-\eta_j d_w)\right]$$

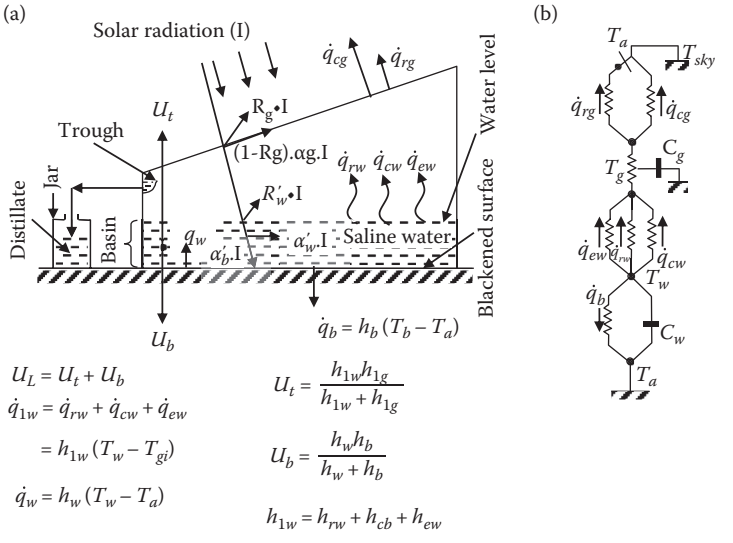


FIGURE 4.33.7 (a) Energy flow diagram of a conventional single-slope solar still and (b) equivalent thermal circuit diagram of conventional single-slope solar still.

where $\sum \mu_j \exp(-\eta_j d_w)$ is the attenuation factor in which d_w is the depth of water.

$$\alpha_b' = (1-R_g)(1-\alpha_g)(1-R_w)\alpha_b\left[\sum \mu_j \exp(-\eta_j d_w)\right]$$

Hence, the fraction of solar intensity lost to the ambient through water and glass cover if attenuation of solar intensity within water mass is considered is expressed as

$$(1-R_g)(1-\alpha_g)(1-R_w)(1-\alpha_b)\left[\sum \mu_j \exp(-\eta_j d_w)\right]$$

The heat transfer mode in the case of solar still can be divided as external and internal. The external heat transfer mode consists of conduction, convection, and radiation. It occurs at the outer surface of glass cover, bottom, and side insulation. The internal heat transfer mode takes place by radiation, convection, and evaporation. It occurs between water surface and inner surface of glass cover. The evaporative heat transfer is responsible for solar distillation. Convection and radiation are losses.

External Heat Transfer

(a) Top loss heat coefficient

The temperature in the glass may be assumed to be uniform due to the small thickness of glass. The rate of heat transfer from the glass cover to the outside atmosphere ($\dot{q}_g$) can be written as

$$\dot{q}_g = \dot{q}_{rg} + \dot{q}_{cg} \tag{4.33.1}$$

where

$$\dot{q}_{rg} = \epsilon_g \sigma\left[(T_{go}+273)^4 - (T_{sky}+273)^4\right] \tag{4.33.2}$$

$$\dot{q}_{rg} = h_{rg}\left(T_{go} - T_a\right) \tag{4.33.3}$$

with

$$h_{rg} = \frac{\epsilon_g \sigma\left[(T_{go}+273)^4 - (T_{sky}+273)^4\right]}{(T_{go}-T_a)} \tag{4.33.4}$$

$$\dot{q}_{cg} = h_{cg}\left(T_{go} - T_a\right) \tag{4.33.5}$$

Putting values of $\dot{q}_{rg}$ and $\dot{q}_{cg}$ in Equation 4.33.1, we get

$$\dot{q}_g = h_{1g}\left(T_{go} - T_a\right) \tag{4.33.6}$$

where

$$h_{1g} = h_{rg} + h_{cg} \tag{4.33.7}$$

The empirical relation for h_{1g} can be discussed for the following cases:

Case (i): The expression for h_{1g} is given by

$$h_{1g} = 5.7 + 3.8\,V$$

where V is the velocity of wind (m/s). This expression incorporates the effect of free convection and radiation from the glass cover as explained by Watmuff et al. (1977).

Case (ii): In this case, the radiation and convection losses are evaluated separately. The radiative heat transfer coefficient (h_{rg}) can be calculated using Equation 4.33.4. The convective heat transfer coefficient (h_{cg}) can be evaluated from the relation

$$h_{cg} = 2.8 + 3.0\,V \tag{4.33.8}$$

However, there is no significant change in the performance of the solar still by considering h_{1g} as represented either by Case (i) or Case (ii).

(b) Bottom and side loss coefficient

Heat is also lost from the water in the basin to the ambient through the insulation and consequently by convection and radiation from the bottom or side surface of the basin. The bottom loss coefficient (U_b) can be expressed as

$$U_b = \left[\frac{1}{h_w} + \frac{1}{h_b}\right]^{-1} = \left[\frac{1}{h_w} + \frac{L_i}{K_i} + \frac{1}{h_{cb} + h_{rb}}\right] \tag{4.33.9}$$

The side heat loss coefficient (U_e) can be written as

$$U_e = \frac{U_b A_{ss}}{A_s} \tag{4.33.10}$$

If A_{ss} is very small in comparison to A_s, for small depth of water, U_e can be neglected. Here, A_{ss} is surface area in contact with water and A_s is the area of basin of solar still. The rate of heat loss per m^2 from basin liner to the ambient can be written as

$$\dot{q}_b = h_b (T_b - T_a) \tag{4.33.11}$$

where

$$h_b = \left[\frac{L_i}{K_i} + \frac{1}{h_{cb} + h_{rb}}\right]^{-1} \tag{4.33.12}$$

Internal Heat Transfer

a. Radiative loss coefficient

In this case, the water surface and the glass cover are considered as infinite parallel planes. This approximation is valid for small inclination of the glass cover and for large width of distiller unit. The rate of radiative heat transfer ($\dot{q}_{rwg}$) from the water surface to the glass cover for these infinite parallel planes is given by

$$\dot{q}_{rwg} = \epsilon_{\text{eff}} \sigma \left[(T_w + 273)^4 - (T_{gi} + 273)^4\right] \tag{4.33.13}$$

$$\dot{q}_{rwg} = h_{rwg} (T_w - T_{gi}) \tag{4.33.14}$$

where h_{rwg} is the radiative heat transfer coefficient from water surface to the glass cover and it is given by

$$h_{rwg} = \epsilon_{\text{eff}} \sigma \left[(T_w + 273)^2 + (T_{gi} + 273)^2\right][T_w + T_{gi} + 546] \tag{4.33.15}$$

Here, it should be noted that water and glass are considered to be parallel surfaces and hence the radiation shape factor is 1. Also, $\frac{1}{\epsilon_{\text{eff}}} = \frac{1}{\epsilon_w} + \frac{1}{\epsilon_g} - 1$.

b. Convective loss coefficient
Heat transfer takes place across humid air in the solar distillation unit by free convection. It is caused by the effect of buoyancy due to variation of density in humid fluid. The variation in density is due to the temperature gradient in the fluid. Hence, the heat transfer rate from water surface to the glass covers by convection ($\dot{q}_{cwg}$) in upward direction through humid fluid can be expressed as

$$\dot{q}_{cwg} = h_{cwg}\left(T_w - T_{gi}\right) \tag{4.33.16}$$

The coefficient h_{cwg} can be evaluated by the relation given by Dunkle (1961) as

$$h_{cwg} = 0.884\left[\left(T_w - T_g\right) + \frac{(P_w - P_{gi})(T_w + 273)}{268.9\times10^3 - P_w}\right]^{1/3} \tag{4.33.17}$$

c. Evaporative loss coefficient
The evaporative heat transfer rate from water surface to glass cover can be written as

$$\dot{q}_{ew} = h_{ewg}(T_w - T_{gi}) \tag{4.33.18}$$

The value of evaporative heat transfer coefficient can be obtained by the relation given by Cooper (1973a,b) as

$$h_{ewg} = 16.273\times10^3 h_{cwg}\frac{\left(P_w - P_g\right)}{\left(T_w - T_g\right)} \tag{4.33.19}$$

where

$$P\left(T\right) = \exp\left(25.317 - \frac{5144}{T + 273}\right) \tag{4.33.20}$$

The value of h_{ewg} can be more realistic for larger value of $(T_w - T_{gi})$. The total heat transfer coefficient (h_{1w}) can be can be evaluated as

$$h_{1w} = h_{rwg} + h_{cwg} + h_{ewg} \tag{4.33.21}$$

where h_{rwg}, h_{cwg}, and h_{ewg} can be evaluated using Equations 4.33.15, 4.33.17, and 4.33.19, respectively.

Overall Heat Transfer

a. Top loss coefficient
The rate of heat loss in the upward direction of a solar still can be expressed as

$$\dot{q}_t = U_t(T_w - T_a) \tag{4.33.22}$$

where U_t is the top loss coefficient from the water surface to the ambient air and it can be written as

$$U_t = \left[\frac{1}{h_{1g}} + \frac{1}{h_{1w}}\right]^{-1} \tag{4.33.23}$$

b. Bottom loss coefficient
The rate of heat loss through the bottom of insulation from water to the ambient air can be written as

$$\dot{q}_b = U_b(T_w - T_a) \tag{4.33.24}$$

where U_b is the bottom heat loss coefficient and it is given by

$$U_b = \frac{h_w h_b}{h_w + h_b} \tag{4.33.25}$$

The rate of total heat lost per m^2 from the water surface to the ambient through the top and bottom of the system ($\dot{q}_{\text{loss}}$) can be calculated as

$$\dot{q}_{\text{loss}} = U_L(T_w - T_a) \tag{4.33.26}$$

where U_L is the overall heat transfer coefficient and it is given by $U_L = U_t + U_b$.

Modified Internal Heat Transfer

The convective heat transfer coefficient from water surface to inner glass cover surface of a conventional solar still is calculated by Equation 4.33.17 given by Dunkle (1961). However, this expression has some limitations. Firstly, this equation is valid only for normal operating temperature ≅ 50°C in a solar still and temperature difference of 17°C. Secondly, it is independent of cavity volume due to $n = 1/3$. Thirdly, it is valid only for parallel evaporative and condensing surfaces. In the new design of solar still, the operating temperature is higher and limitations mentioned above are not satisfied. Hence, there is a need to find new methodology to get the value of convective heat transfer coefficient. In this regard, Kumar and Tiwari (1996b) have proposed the methodology to evaluate convective heat transfer coefficient using linear regression analysis. This methodology is as follows:

The hourly distillate output per m^2 from a distillate unit is given by

$$\dot{m}_{ew} = \frac{\dot{q}_{ew} \times 3600}{L} = \frac{16.273 \times 10^3 h_{cw}\left(P_w - P_g\right) \times 3600}{L} \tag{4.33.27}$$

(from Equation 4.33.18 and 4.33.19)
where

$$h_{cw} = \frac{\text{Nu}\ K_f}{d_f} = \frac{K_f}{d_f} C\left(\text{GrPr}\right)^n \tag{4.33.28}$$

Putting the value of h_{cw} in Equation 4.33.27, one gets the following expression.

$$\frac{\dot{m}_{ew}}{R} = C\left(\text{GrPr}\right)^n \tag{4.33.29}$$

where $R = 16.273 \times 10^3 h_{cw}\left(P_w - P_g\right)\dfrac{K_f}{d_f}\left(\dfrac{3600}{L}\right)$.

Equation 4.33.29 represents the equation of straight line and it can be expressed as

$$y = ax^b \tag{4.33.30}$$

where $y = \dfrac{\dot{m}_{ew}}{R}$; $x = \text{GrPr}$; $a = C$; and $b = n$.

Now, taking ln on both sides of Equation 4.33.30, it can be expressed in the form of straight line equation as follows:

$$y' = b'x' + a' \tag{4.33.31}$$

where $y' = \ln(y)$; $b' = b$; $x' = \ln x$; and $a' = a$.

Using linear regression analysis, coefficients a' and b' in Equation 4.33.31 can be expressed as

$$b' = \frac{N\left(\sum x'y'\right) - \left(\sum x'\right)\left(\sum y'\right)}{N\left(\sum (x')^2\right) - \left(\sum x'\right)^2} \tag{4.33.32}$$

$$a' = \left(\frac{\sum y'}{N}\right) - b'\left(\frac{\sum x'}{N}\right) \tag{4.33.33}$$

Here, N is the number of experimental observations under steady-state condition. In a quasi-steady-state condition, N becomes $(N+1)$. After calculating b' and a' from Equations 4.33.32 and 4.33.33, respectively, C and n can be evaluated using the following expressions.

$$C = \exp(a') \text{ and } n = b'$$

Once, values of C and n are known, one can evaluate convective heat transfer coefficient from Equation 4.33.28.

Thermal Modeling of Solar Stills

Single and double-slope PVT-CPC active solar stills are shown in Figures 4.33.8 and 4.33.9, respectively. Assumptions for writing energy balance equation for single-slope and double-slope PVT-CPC active solar stills (Singh and Tiwari, 2016) are as follows:

1. The partially covered PVT-CPC water collector and solar stills are in quasi steady state.
2. The losses due to resistance in the solar cells are negligible. These are the losses occurring as a result of the interconnections and bus bar in the PV module and other connecting wires. These wires are generally made up of aluminum or copper which have very low electrical resistivity (of the order of 10^{-8} Ω-m). So, i^2r losses are negligible in comparison to the other thermal losses or heat transfers in the system. Hence, neglecting the ohmic losses will not affect the thermal performance of the solar cell in greater extent.
3. The heat capacity of solar cell, glass cover, absorbing, and insulation material of collector and solar stills is negligible as heat losses due to heat capacity are low in comparison to the other rate of heat transfer in the system. The amount of heat loss in solar cell can be written as $m \cdot c_p \cdot \Delta T$, where m represents the mass of solar cell, c_p represents specific heat capacity at constant pressure of solar cell material (silicon), and ΔT represents the difference of average cell temperature and ambient temperature. The mass of solar cell

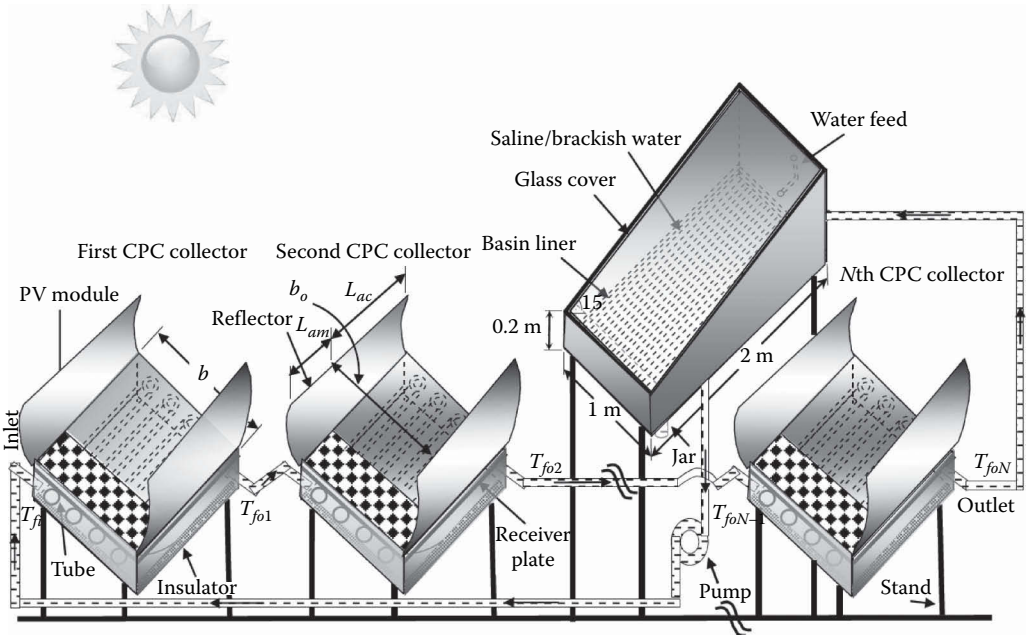


FIGURE 4.33.8 Schematic diagram of single-slope PVT-CPC active solar still.

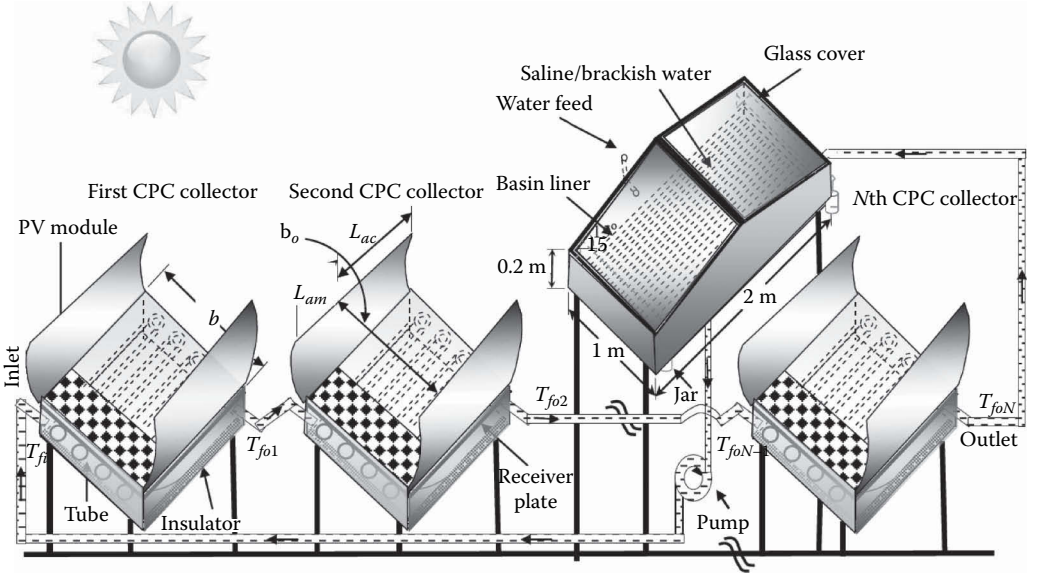


FIGURE 4.33.9 Schematic diagram of double-slope PVT-CPC active solar still.

(0.156 m×0.156 m×0.0002 m) comes out to be ~0.0113 kg taking the density of silicon as 2.33×10^3 kg/m^3. The value of c_p for silicon (Si) at room temperature is 0.705 kJ/kg K and average value of ΔT in the month of June for the system is 44.87 K (Singh and Tiwari, 2016). Therefore, heat loss due to heat capacity in solar cell comes out to be 0.36 kJ which is 0.0012% of the heat output from solar still because heat output from solar still is of the order of 10^4kJ. Further, the amount of heat loss due to heat capacity of glass cover is given by $m \cdot c_p \cdot \Delta T$, where m represents mass of glass cover, c_p represents specific heat capacity

of glass at constant pressure, and ΔT represents the temperature difference between inner surface and outer surface. The average value of temperature difference between inner and outer surfaces of glass cover is 2.84 K in the case of PVT active solar still as reported by Kumar et al. (2010). Mass of glass cover in solar still comes out to be 20.70 kg taking glass density as 2500 kg/m^3. The heat consumed by glass is 49.38 kJ taking c_p of glass as 0.84 kJ/kg K, whereas daily heat output of the system is of the order of 10^4kJ. So, heat consumed by glass due to its heat capacity comes out to be 0.17% in comparison to the amount of heat output from the solar still. In the similar way, heat consumed by absorbing and insulating materials corresponding to their heat capacities can also be computed which will come out negligibly small. Hence, the amount of heat loss due to heat capacities of solar cell, glass, absorbing, and insulating materials can be neglected as they are not going to affect the result in greater extent.

4. Solar stills are vapor leakage proof. If leakage occurs, the steady-state condition no longer exists and energy equation will become too complex due to transient condition. Hence, the leakage of vapor has been neglected as it will not affect the result much and steady-state condition exists.
5. The level of water in the basin is constant. The maximum daily yield of the solar distillation system is 30.36 kg as reported by Singh and Tiwari (2016). The corresponding change in level of water comes out to be 0.015 m during 24 h. The variation in daily yield is 0.25% as reported by Tiwari et al. (2009) if water level changes by 15.38% (0.11–0.13 m) in the case of active solar still. Hence, the small change in level of water can be neglected.
6. No stratification (i.e., no layered configuration) of water occurs in the basin of active solar stills as there is a continuous stirring of water due to its circulation by pump throughout the system which results in mixing of water.
7. The film-type condensation occurs at the inner surface of glass. A small inclination with horizontal is provided to glass for facilitating the same as the component of gravitational force $(mg\sin\theta)$ will act along the surface in the downward direction which will help in trickling down the condensate. Here, θ is the angle of inclination of the glass with horizontal. Also, the inner surface of glass is cleaned to prevent the droplet formation. Further, if the film-type condensation does not take place then water droplets will form at the inner surface of glass cover which will act as mirror and some portion of the solar radiation will be reflected. Hence, lesser amount of solar radiation will reach the surface of water which will adversely affect the yield. Therefore, the occurrence of film-type condensation is ensured during the design phase. However, small amount of droplets may still form which can be ignored as they are negligibly small.

N Identical PVT-CPC Active Solar Still

Useful Energy Gain for N Identical Partially Covered PVT-CPC Water Collectors Connected in Series

Following Tripathi et al. (2016), the rate of useful thermal output from N identical partially covered PVT-CPC water collectors connected in series can be expressed as,

$$\dot{Q}_{uN} = \frac{(1-K_k^N)}{(1-K_k)}(A\ F_R(\alpha\tau))_1 I_b(t) + \frac{(1-K_k^N)}{(1-K_k)}(A\ F_R U_L)_1(T_{fi} - T_a) \tag{4.33.34}$$

In their set-up, PVT-CPC water collectors are in open loop. However, collectors are in closed loop with solar still in PVT-CPC active solar still. The water from basin of solar still enters to the inlet of first PVT-CPC water collector and outlet of *N*th PVT-CPC water collector are fed to the basin of solar still. Hence, T_{fi} becomes equal to T_w.

The temperature at the outlet of *N*th PVT-CPC water collector (T_{foN}) is given by

$$T_{foN} = \frac{(A\ F_R(\alpha\tau))_1}{\dot{m}_f C_f}\frac{(1-K_k^N)}{(1-K_k)} I_b(t) + \frac{(A\ F_R U_L)_1}{\dot{m}_f C_f}\frac{(1-K_k^N)}{(1-K_k)} T_a + K_k^N T_{fi} \tag{4.33.35}$$

where $T_{fi} = T_w$. The temperature obtained at the outlet of *N*th PVT-CPC water collector is fed to the basin of solar still. Hence, $T_{wo} = T_{foN}$.

The analytical expression for the temperature-dependent electrical efficiency of solar cells (η_{cN}) of a number (N) of PVT-CPC water collectors is given by Evans (1981) and Schott (1985) as

$$\eta_{cN} = \eta_o\left[1-\beta_o(\overline{T}_{cN} - T_o)\right] \tag{4.33.36}$$

where η_o is the efficiency at standard test condition and $\overline{T}_{cN}$ is the average solar cell temperature of *N*th PVT-CPC water collector. The value of $\overline{T}_{cN}$ can be calculated using the expression developed by Tripathi et al. (2016) in which $T_{fi} = T_w$.

Energy Balance Equations for Single Slope Solar Still

Following Singh and Tiwari (2016), energy balance equation for single-slope solar still can be written as follows:

Inner surface of glass cover

$$\alpha'_g I_s(t) A_g + h_{1w}(T_w - T_{gi})A_b = \frac{K_g}{L_g}(T_{gi} - T_{go})A_g \tag{4.33.37}$$

where $\alpha'_g = (1 - R_g)\alpha_g$ represents the fraction of solar flux absorbed by the glass cover and $h_{1w} = h_{rwg} + h_{cwg} + h_{ewg}$ represents the rate of total heat transfer coefficient from water surface to inner surface of glass cover.

Outer surface of glass cover

$$\frac{K_g}{L_g}(T_{gi} - T_{go})A_g = h_{1g}(T_{go} - T_a)A_g \tag{4.33.38}$$

Here, $h_{1g} = h_{r,g} + h_{c,g}$ or $h_{1g} = 5.7 + 3.8\ V$

Water mass in basin

$$\dot{Q}_{uN} + \alpha'_w I_s(t) A_b + h_{bw}(T_b - T_w)A_b = h_{1w}(T_w - T_{gi})A_b + M_w C_w \frac{dT_w}{dt} \tag{4.33.39}$$

Here, $\alpha'_w = (1 - R_g)(1 - \alpha_g)(1 - R_w)\alpha_w$ represents the fraction of solar flux absorbed by water mass and $\dot{Q}_{uN}$ is the rate of useful thermal output from N identical hybrid PVT collectors connected in series.

Basin liner

$$\alpha'_b I_s(t) A_b = h_{bw}(T_b - T_w)A_b + h_{ba}(T_b - T_a)A_b \tag{4.33.40}$$

where $\alpha'_b = (1 - R_g)(1 - \alpha_g)(1 - R_w)(1 - \alpha_w)\alpha_b$ which is the fraction of solar flux absorbed by basin liner.

Using Equations 4.33.34 and 4.33.37 through 4.33.40, one can get the following differential equation for water temperature in basin.

$$\frac{dT_w}{dt} + a_1 T_w = f_1(t) \tag{4.33.41}$$

Values of a_1 and $f_1(t)$ of Equation 4.33.41 are given in Appendix.

The following assumptions have been made to get an approximate solution of differential equation (4.33.41).

1. The time interval (Δt) is small, i.e., ($0 < t < \Delta t$).
2. Values of T_a, $I_b(t)$, and $I_s(t)$ can be considered as average value between "0" and "t", i.e., $\overline{T}_a$, $\overline{I}_b(t)$, and $\overline{I}_s(t)$. Hence $f_1(t)$ becomes constant and its average value will be $\overline{f_1}(t)$.
3. a_1 is constant for interval Δt.

One can get the solution of Equation 4.33.41 with initial condition, $T_w = T_{w0}$ at $t=0$, as follows:

$$T_w = \frac{\overline{f_1}(t)}{a_1}(1 - e^{-a_1 t}) + T_{w0} e^{-a_1 t} \tag{4.33.42}$$

After evaluating T_w from Equation 4.33.42, one can obtain glass temperature from Equations 4.33.37 and 4.33.38 as follows:

$$T_{gi} = \frac{\alpha'_g I_s(t) A_g + h_{1w} T_w A_b + U_{c,ga} T_a A_g}{U_{c,ga} A_g + h_{1w} A_b} \tag{4.33.43}$$

$$T_{go} = \frac{\dfrac{K_g}{L_g} T_{gi} + h_{1g} T_a}{\dfrac{K_g}{L_g} + h_{1g}} \tag{4.33.44}$$

The hourly yield ($\dot{m}_{ew}$) can be expressed as

$$\dot{m}_{ew} = \frac{h_{ewg} A_b (T_w - T_{gi})}{L} \times 3600 \tag{4.33.45}$$

Here, L is the latent heat of evaporation and can be evaluated using the expression given by Fernandez and Chargoy (1990) and Toyama and Kagakuv (1972).

Above equations can be discussed for the following limiting condition:

1. For $\frac{A_a}{A_r} = 1$ and $I_b(t) = I(t)$, equations derived for single-slope PVT-CPC active solar still reduce to equations for single-slope PVT-FPC active solar still.
2. For $\frac{A_a}{A_r} = 1$, $A_{rm} = 0.625\ \text{m}^2$, $A_{rc} = 1.375\ \text{m}^2$, $I_b(t) = I(t)$, $A_b = 1\ \text{m}^2$, and $N=2$, equations derived for single-slope PVT-CPC active solar still reduce to expressions derived by Singh et al. (2016).
3. In the case of single-slope passive solar still $\dot{Q}_{uN} = 0$ as no external energy is fed to the basin of solar still. Hence, putting $\dot{Q}_{uN} = 0$, equations derived for PVT-CPC active solar still reduce to equations for single-slope passive solar still (Singh et al., 2016).

Energy Balance Equations for Double Slope Solar Still

Following Singh and Tiwari (2016), energy balance equation for double-slope solar still can be written as follows:

Energy balances for east cover

For inner glass cover

$$\alpha'_g I_{SE}(t) A_{gE} + h_{1wE}(T_w - T_{giE}) \frac{A_b}{2} - h_{EW}(T_{giE} - T_{giW}) A_{gE} = \frac{K_g}{L_g}(T_{giE} - T_{goE}) A_{gE} \tag{4.33.46}$$

Here, $h_{1wE} = h_{rwgE} + h_{cwgE} + h_{ewgE}$ which is called total heat transfer coefficient from water surface to glass cover and α'_g represents the fraction of solar flux absorbed by the glass cover.

For outer glass cover

$$\frac{K_g}{L_g}(T_{giE} - T_{goE})A_{gE} = h_{1gE}(T_{goE} - T_a)A_{gE} \tag{4.33.47}$$

where $h_{1gE} = h_{rgE} + h_{cgE}$ or $h_{1gE} = 5.7 + 3.8$ V.

Energy balances for west cover

For inner glass cover

$$\alpha'_g I_{SW}(t)A_{gW} + h_{1wW}(T_w - T_{giW})\frac{A_b}{2} + h_{EW}(T_{giE} - T_{giW})A_{gE} = \frac{K_g}{L_g}(T_{giW} - T_{goW})A_{gW} \tag{4.33.48}$$

where $h_{1wW} = h_{rwgE} + h_{cwgE} + h_{ewgE}$ which is called total heat transfer coefficient from water surface to glass cover on the west side.

For outer glass cover

$$\frac{K_g}{L_g}(T_{giW} - T_{goW})A_{gW} = h_{1gW}(T_{goW} - T_a)A_{gW} \tag{4.33.49}$$

where $h_{1gW} = h_{rgW} + h_{cgW}$ or $h_{1gW} = 5.7 + 3.8$ V.

For basin liner

$$\alpha'_b\left(I_{SE}(t) + I_{SW}(t)\right)\frac{A_b}{2} = h_{bw}(T_b - T_w)A_b + h_{ba}(T_b - T_a)A_b \tag{4.33.50}$$

Here, α'_b is the fraction of solar flux absorbed by basin liner.

For water mass in basin

$$(M_w C_w)\frac{dT_w}{dt} = \left(I_{SE}(t) + I_{SW}(t)\right)\alpha'_w\frac{A_b}{2} + h_{bw}(T_b - T_w)A_b - h_{1w}(T_w - T_{giE})\frac{A_b}{2} - h_{1w}(T_w - T_{giE})\frac{A_b}{2} + \dot{Q}_{uN} \tag{4.33.51}$$

Using Equation 4.33.34 and 4.33.46 through 4.33.51, one can get the differential equation for water temperature (T_w) in basin. Following assumptions made in solving Equation (4.33.41), one can get the solution of this differential equation with initial condition $T_w = T_{wo}$ at $t=0$ as

$$T_w = \frac{\overline{f_2}(t)}{a_2}(1 - e^{-a_2 t}) + T_{w0}e^{-a_2 t} \tag{4.33.52}$$

Equation 4.33.52 is similar to Equation 4.33.42 obtained for single-slope PVT-CPC active solar distillation system. However, constants of Equation 4.33.52 which has been obtained for double-slope PVT-CPC active solar distillation system are different. Their expressions are given in Appendix. After evaluating T_w from Equation 4.33.52, one can obtain glass temperature for double-slope PVT-CPC active solar distillation system using Equations 4.33.47 through 4.33.50 as follows:

$$T_{giE} = \frac{A_1 + A_2 T_w}{P} \tag{4.33.53}$$

$$T_{giW} = \frac{B_1 + B_2 T_w}{P} \tag{4.33.54}$$

$$T_{goE} = \frac{\frac{K_g}{L_g} T_{giE} + h_{1gE} T_a}{\frac{K_g}{L_g} + h_{1gE}} \tag{4.33.55}$$

$$T_{goW} = \frac{\frac{K_g}{L_g} T_{giW} + h_{1gW} T_a}{\frac{K_g}{L_g} + h_{1gW}} \tag{4.33.56}$$

Constants of Equations 4.33.53 through 4.33.56 are given in Appendix. After evaluating glass temperature from Equations 4.33.53 and 4.33.54 and water temperature from Equation 4.33.52, one can obtain hourly yield for east and west side in the case of double-slope PVT-CPC active solar distillation system as follows:

$$\dot{m}_{ew,E} = \frac{h_{ewE} \frac{A_b}{2} \left(T_w - T_{giE}\right)}{L} \times 3600 \tag{4.33.57}$$

$$\dot{m}_{ew,W} = \frac{h_{ewW} \frac{A_b}{2} \left(T_w - T_{giW}\right)}{L} \times 3600 \tag{4.33.58}$$

The hourly yield for double-slope PVT-CPC active solar distillation system can be calculated by adding the hourly yield obtained from east side ($\dot{m}_{ew,E}$) using Equation 4.33.57 and hourly yield obtained from west side ($\dot{m}_{ew,W}$) using Equation 4.33.58.

Above equations can be discussed for the following limiting conditions:

1. For $\frac{A_a}{A_r} = 1$ and $I_b(t) = I(t)$, equations derived for double-slope PVT-CPC active solar still reduce to equations for double-slope PVT-FPC active solar still.
2. In the case of single-slope passive solar still, $\dot{Q}_{uN} = 0$ as no external energy is fed to the basin of solar still. Hence, putting $\dot{Q}_{uN} = 0$, equations derived for double-slope PVT-CPC active solar still reduce to equations for double-slope passive solar still (Dwivedi and Tiwari, 2008).

Performance Parameters of Solar Still

The performance of solar still is analyzed by its energy, exergy, efficiency, energy matrices, productivity, exergoeconomic parameter, and enviroeconomic parameter. These parameters are as follows.

Energy Output

The energy output can be calculated with the help of first law of thermodynamics. Following Singh and Tiwari (2016), the annual overall thermal energy (E_{out}) of PVT-CPC active solar still can be written as

$$E_{\text{out}} = \frac{\left(M_{ew} \times L\right)}{3600} + \frac{\left(P_m - P_u\right)}{0.38} \tag{4.33.59}$$

Here, M_{ew}, P_m, P_u, and L are annual yield from solar still, annual power generated from PV module, annual power utilized by pump and L is latent heat. The value of M_{ew} can be computed with the help of Equations 4.33.45, 4.33.57, and 4.33.58. $(P_m - P_u)$ in Equation 4.33.59 represents net electrical energy for the whole year. It has been divided by 0.38 to get the equivalent thermal energy as electrical energy is high grade energy. The factor 0.38 corresponds to power generation efficiency for a conventional power plant (Huang et al., 2001). The expression for hourly electrical energy/electrical exergy $(\dot{E}x_e)$ can be expressed as

$$\dot{E}x_e = A_m I_b(t) \sum_{1}^{N} (\alpha \tau_g \eta_{cN}) \tag{4.33.60}$$

The daily electrical exergy can be computed by adding hourly electrical exergy for 10 h as beam intensity exists for 10 h only. The monthly electrical exergy can be evaluated by multiplying daily electrical exergy with number of clear days and addition of monthly electrical exergy for 12 months will give annual electrical exergy (P_m) used in Equation 4.33.59.

In the case of solar still integrated with N identical PVT-FPC, $I_b(t) = I(t)$. In the case of passive solar still and solar still integrated with FPC, the second term of Equation 4.33.59 will not appear.

Exergy Output

The exergy output can be computed using the first law (energy) and second law (entropy) of thermodynamics. Following Nag (2004), the hourly output thermal exergy $\dot{E}x_{\text{out}}(W)$ for single-slope active solar still can be written as

$$\dot{E}x_{\text{out},s} = A_b h_{ewg} \left[(T_w - T_{gi}) - (T_a + 273) \times \ln \left\{ \frac{(T_w + 273)}{(T_{gi} + 273)} \right\} \right] \tag{4.33.61}$$

where h_{ewg} can be computed with the help of Equation 4.33.19.

Following Equation 4.33.61, the expression of hourly exergy gain for double-slope active solar still $(\dot{E}x_{\text{out},d})$ can be written as

$$\dot{E}x_{\text{out},d} = h_{ewgE} \times \frac{A_b}{2} \times \left[(T_w - T_{giE}) - (T_a + 273) \times \ln \left\{ \frac{(T_w + 273)}{(T_{giE} + 273)} \right\} \right] + h_{ewgW} \times \frac{A_b}{2} \times \left[(T_w - T_{giW}) - (T_a + 273) \times \ln \left\{ \frac{(T_w + 273)}{(T_{giW} + 273)} \right\} \right] \tag{4.33.62}$$

The value of h_{ewgE} and h_{ewgW} can be calculated with the help of Equation 4.33.19.

The overall annual exergy gain $(G_{ex,\text{annual}})$ for single- and double-slope solar stills integrated with PVT collectors can be written as

$$G_{\text{ex,annual}} = \text{Ex}_{\text{out}} + (P_m - P_u) \tag{4.33.63}$$

Here, Ex_{out} is annual thermal exergy output obtained from solar still, P_m is annual electrical exergy obtained from N identical PVT collectors, and P_u is power required to drive the pump. In the case of single- and double-slope passive solar stills, the second term of Equation 4.33.63 does not exist as collectors and pump are absent in these cases.

The daily exergy for clear day can be calculated by adding hourly exergy for 10 h as solar intensity exists for sunshine hours only. The monthly exergy for clear days can be evaluated by multiplying

daily exergy with the corresponding number of clear days. The annual exergy can be calculated by adding monthly exergy for 12 month. Following Equation 4.33.61, the hourly thermal exergy output ($\dot{E}x_c(t)$) from N identical PVT-CPC/PVT-FPC water collectors can be written as

$$\dot{E}x_c(t) = (\dot{m}_f \times C_f)\left[(T_{foN} - T_{fi}) - (T_a + 273) \times \ln\frac{(T_{foN} + 273)}{(T_{fi} + 273)}\right] \tag{4.33.64}$$

Here, $\dot{m}_f$, C_f, T_{foN}, T_{fi}, and T_a represent mass flow rate of fluid in collector, specific heat capacity of fluid in collector, temperature at the outlet of Nth collector, temperature at the inlet of first collector and ambient temperature, respectively.

Efficiency of Solar Still

Efficiency of solar still is calculated on the basis of first law as well as second law of thermodynamics. Various efficiencies of solar still are as follows:

a. Thermal efficiency
It is calculated on the basis of first law of thermodynamics. The hourly and daily thermal efficiency (Tiwari, 2002) of partially covered PVT-CPC single-slope active solar still can be written as follows:

$$\eta_{\text{hourly,thermal}} = \frac{\dot{m}_{ew} \times L}{\left[\dot{Q}_{uN} + A_b \times I_s(t)\right] \times 3600} \times 100 \tag{4.33.65}$$

$$\eta_{\text{daily,thermal}} = \frac{\sum_{t=1}^{24} \dot{m}_{ew} \times L}{\sum_{t=1}^{24}\left[\dot{Q}_{uN} + A_b \times I_s(t)\right] \times 3600} \times 100 \tag{4.33.66}$$

where t represents time, $\dot{m}_{ew}$ is the hourly yield obtained from solar still, $I_s(t)$ is hourly solar intensity falling on solar still, $\dot{Q}_{uN}$ is the hourly useful energy gain from N identical partially covered PVT-CPC water collectors, and L is the latent heat of vaporization which can be calculated using expression given by Fernandez and Chargoy (1990) and Toyama and Kagakuv (1972).
Following Equations 4.33.65 and 4.33.66, the hourly and daily thermal efficiency of partially covered PVT-CPC double-slope active solar still can be written as follows:

$$\eta_{\text{hourly,thermal}} = \frac{\left(\dot{m}_{ewE} + \dot{m}_{ewW}\right) \times L}{\left[\dot{Q}_{uN} + \frac{A_b}{2}\left(I_{sE}(t) + I_{sW}(t)\right)\right] \times 3600} \times 100 \tag{4.33.67}$$

$$\eta_{\text{daily,thermal}} = \frac{\sum_{t=1}^{24}\left(\dot{m}_{ewE} + \dot{m}_{ewW}\right) \times L}{\sum_{t=1}^{24}\left[\dot{Q}_{uN} + \frac{A_b}{2}\left(I_{sE}(t) + I_{sW}(t)\right)\right] \times 3600} \times 100 \tag{4.33.68}$$

where $\dot{m}_{ewE}$, $\dot{m}_{ewW}$, $I_{sE}(t)$, and $I_{sW}(t)$ represent hourly yield obtained from east side, hourly yield obtained from west side, hourly solar intensity falling on east side, and hourly solar intensity falling on west side, respectively.

In the case of solar still integrated with N identical PVT-FPC, $\dot{Q}_{uN}$ represents the hourly energy gain from N identical PVT-FPC which can be obtained by putting $\frac{A_a}{A_r}=1$ and $I_b(t)=I(t)$ in Equation 4.33.34 or alternatively from Shyam et al. (2015). In the case of N identical FPC connected to solar still, $\dot{Q}_{uN}$ represents the hourly energy gain from N identical FPC. The value of $\dot{Q}_{uN}$ becomes zero in the case of single-slope passive solar still as no energy is fed to solar still.

b. Thermal exergy efficiency

The hourly thermal exergy efficiency of single- and double-slope PVT-CPC/PVT-FPC active solar still can be expressed as

$$\eta_{\text{hourly,exergy},s}=\frac{\dot{E}x_{\text{out},s}(t)}{\dot{E}x_c(t)+0.933\times\left(A_b\times I_S(t)\right)}\times 100 \tag{4.33.69}$$

$$\eta_{\text{hourly,exergy},d}=\frac{\dot{E}x_{\text{out},d}(t)}{\dot{E}x_c(t)+\left[0.933\times\frac{A_b}{2}\times\left(I_{SE}(t)+I_{SW}(t)\right)\right]}\times 100 \tag{4.33.70}$$

Here, the factor 0.933 has been obtained using the expression given by Petela (2003). It has been used to convert radiation to exergy. Following Equations 4.33.69 and 4.33.70, the daily thermal exergy efficiency of single- and double-slope PVT-CPC/PVT-FPC active solar still can be expressed as

$$\eta_{\text{daily,exergy},s}=\frac{\sum_{t=1}^{24}\left[\dot{E}x_{\text{out},s}(t)\right]}{\sum_{t=1}^{24}\left[\dot{E}x_c(t)+0.933\times\left(A_b\times I_S(t)\right)\right]}\times 100 \tag{4.33.71}$$

$$\eta_{\text{daily,exergy},d}=\frac{\sum_{t=1}^{24}\left[\dot{E}x_{\text{out},d}(t)\right]}{\sum_{t=1}^{24}\left[\dot{E}x_c(t)+\left[0.933\times\frac{A_b}{2}\times\left(I_{SE}(t)+I_{SW}(t)\right)\right]\right]}\times 100 \tag{4.33.72}$$

In the case of single- and double-slope passive solar stills, the value of $\dot{E}x_c(t)$ will be zero as no exergy is fed to the basin of solar stills in this case.

c. Electrical exergy efficiency

Following Tiwari et al. (2015), the electrical exergy efficiency of single- and double-slope PVT-CPC active solar stills can be written as

$$\eta_{\text{hourly,electrical exergy}}=\frac{\dot{E}x_e(t)-\dot{P}_u(t)}{0.933\times A_{am}\times N\times I_b(t)}\times 100 \tag{4.33.73}$$

$$\eta_{\text{daily,electrical exergy}}=\frac{\sum_{t=1}^{24}\left(\dot{E}x_e-\dot{P}_u\right)}{0.933\times\sum_{t=1}^{t=24}\left[A_{am}\times N\times I_b(t)\right]}\times 100 \tag{4.33.74}$$

where $\dot{P}_u(t)$, N, $I_b(t)$, and A_{am} are hourly consumption of pump, number of collectors, beam radiation, and aperture area of module, respectively. Hourly electrical exergy $\left(\dot{E}x_e(t)\right)$ can be computed using Equation 4.33.60. The value of beam radiation and electrical energy/exergy are zero during off-sunshine hours. In the case of PVT-FPC active solar stills,

$I_b(t) = I(t)$ and $A_{am} = A_m$. In the case of passive solar stills and FPC active solar stills, electrical exergy efficiency does not exist as no electrical output is obtained due to the absence of PV.

d. Overall exergy efficiency

Overall exergy represents the summation of thermal exergy and electrical exergy. Following Singh et al. (2016), the hourly overall exergy efficiency of single- and double-slope PVT-CPC active solar distillation system can be written as

$$\eta_{\text{hourly,overall exergy},s} = \frac{\dot{E}x_{\text{out},s}(t) + \left(\dot{E}x_e(t) - \dot{P}_u(t)\right)}{0.933 \times \left[\left(A_b \times I_S(t)\right) + (A_{am} + A_{ac}) \times N \times I_b(t)\right]} \times 100 \tag{4.33.75}$$

$$\eta_{\text{hourly,overall exergy},d} = \frac{\dot{E}x_{\text{out},d}(t) + \left(\dot{E}x_e(t) - \dot{P}_u(t)\right)}{0.933 \times \left[\left[\frac{A_b}{2} \times \left(I_{SE}(t) + I_{SW}(t)\right)\right] + (A_{am} + A_{ac}) \times N \times I_b(t)\right]} \times 100 \tag{4.33.76}$$

Following Equations 4.33.69 and 4.33.70, the daily overall exergy efficiency of single- and double-slope PVT-CPC active solar distillation system can be written as

$$\eta_{\text{daily,overall exergy},s} = \frac{\sum_{t=1}^{24}\left[\dot{E}x_{\text{out},s}(t) + \left(\dot{E}x_e(t) - \dot{P}_u(t)\right)\right]}{0.933 \times \sum_{t=1}^{24}\left[\left(A_b \times I_S(t)\right) + (A_{am} + A_{ac}) \times N \times I_b(t)\right]} \times 100 \tag{4.33.77}$$

$$\eta_{\text{daily,overall exergy},d} = \frac{\left[\sum_{t=1}^{t=24} \dot{E}x_{\text{out},d}(t) + \left(\dot{E}x_e(t) - \dot{P}_u(t)\right)\right]}{0.933 \times \sum_{t=1}^{24}\left[\left[\frac{A_b}{2} \times \left(I_{SE}(t) + I_{SW}(t)\right)\right] + (A_{am} + A_{ac}) \times N \times I_b(t)\right]} \times 100 \tag{4.33.78}$$

where values of electrical exergy and solar flux used in Equations 4.33.75 through 4.33.78 are zero during off-sunshine hours.

In the case of PVT-FPC active solar still, $I_b(t) = I(t)$, $A_{am} = A_m$, and $A_{ac} = A_c$. In the case of FPC active solar still and passive solar still, electrical component will not exist due to the absence of PVT collectors in these cases.

e. Overall thermal efficiency

Overall thermal energy represents the summation of thermal energy obtained from solar still and the thermal energy equivalent to electrical exergy obtained from PVT. The electrical exergy is divided by a factor 0.38 (Huang et al., 2001) to get the equivalent thermal energy. The hourly overall thermal efficiency of single- and double-slope PVT-CPC active solar stills can be written as

$$\eta_{\text{hourly,overall thermal},s} = \frac{\dot{m}_{ew} \times L}{\left[\dot{Q}_{uN} + A_b \times I_s(t)\right] \times 3600} \times 100 + \frac{\dot{E}x_e(t) - \dot{P}_u(t)}{0.38 \times A_{am} \times N \times I_b(t)} \times 100 \tag{4.33.79}$$

$$\eta_{\text{hourly,overall thermal},d} = \frac{\left(\dot{m}_{ewE} + \dot{m}_{ewW}\right) \times L}{\left[\dot{Q}_{uN} + \frac{A_b}{2}\left(I_{sE}(t) + I_{sW}(t)\right)\right] \times 3600} \times 100 + \frac{\dot{E}x_e(t) - \dot{P}_u(t)}{0.38 \times A_{am} \times N \times I_b(t)} \times 100 \tag{4.33.80}$$

Here, $\dot{m}_{ew}$, $\dot{m}_{ewE}$, and $\dot{m}_{ewW}$ can be evaluated using Equations 4.33.45, 4.33.57, and 4.33.58, respectively. $\dot{Q}_{uN}$ and $\dot{E}x_e$ can be evaluated using Equations 4.33.34 and 4.33.61, respectively. The work consumed by DC motor pump can be evaluated using its specification. Following Equations 4.33.73 and 4.33.74, the daily overall thermal efficiency of single- and double-slope PVT-CPC active solar distillation system can be written as

$$\eta_{\text{daily, overall thermal},s} = \left[\frac{\sum_{t=1}^{24} \left(\dot{m}_{ew} \times L \right)}{\sum_{t=1}^{24} \left[\dot{Q}_{uN}(t) + A_b \times I_s(t) \right] \times 3600} + \frac{\sum_{t=1}^{24} \left[\dot{E}x_e(t) - \dot{P}_u(t) \right]}{\sum_{t=1}^{24} \left[0.38 \times A_{am} \times N \times I_b(t) \right]} \right] \times 100 \tag{4.33.81}$$

$$\eta_{\text{daily,overall thermal},d} = \left[\frac{\sum_{t=1}^{24} \left[\left(\dot{m}_{ewE} + \dot{m}_{ewW} \right) \times L \right]}{\sum_{t=1}^{24} \left[\dot{Q}_{uN}(t) + \frac{A_b}{2} \left(I_{sE}(t) + I_{sW}(t) \right) \right] \times 3600} + \frac{\sum_{t=1}^{24} \left[\dot{E}x_e(t) - \dot{P}_u(t) \right]}{\sum_{t=1}^{24} \left[0.38 \times A_{am} \times N \times I_b(t) \right]} \right] \times 100 \tag{4.33.82}$$

where values of solar flux and electrical exergy used in Equations 4.33.73 through 4.33.76 are zero during off-sunshine hours.

Energy Matrices

Unlike efficiency, energy matrices are calculated on annual basis. Analysis of energy matrices consists of the evaluation of EPBT, EPF, and LCCE. These are an important parameter for renewable technologies as the use of technology does not make any sense if the energy produced by them during the whole life time is less than the energy used in their manufacturing (Tiwari and Mishra, 2012). These are as follows:

a. Energy payback time
   It represents the time period needed to recover the total energy exhausted in preparing the materials (embodied energy) required for fabrication of solar distillation system. Following Tiwari and Mishra (2012), it can be expressed as

$$\text{EPBT based on energy} = \frac{\text{Embodied energy } \left(E_{\text{in}} \right)}{\text{Annual energy output } \left(E_{\text{out}} \right)} \tag{4.33.83}$$

$$\text{EPBT based on exergy} = \frac{\text{Embodied energy } \left(E_{\text{in}} \right)}{\text{Annual exergy output } \left(G_{\text{ex, annual}} \right)} \tag{4.33.84}$$

   Here, embodied energy represents the amount of energy needed to manufacture the material required for solar distillation system. It consists of energy as minor part and exergy as major part. The numerical value of EPBT for solar distillation system should be as low as possible to make them cost effective. It means that the solar distillation system is better if its EPBT is lower.

b. Energy production factor
   It represents the overall performance of solar distillation system. It is defined as the reciprocal of EPBT and its ideal value on annual basis is 1. Following Tiwari and Mishra (2012), EPF for solar distillation system on annual basis can be written as

$$\text{EPF based on energy} = \frac{E_{\text{out}}}{E_{\text{in}}} \tag{4.33.85}$$

$$\text{EPF based on exergy} = \frac{G_{\text{ex,annual}}}{E_{\text{in}}} \tag{4.33.86}$$

where E_{out}, E_{in}, and $G_{ex,\text{annual}}$ are, respectively, overall annual energy output, embodied energy, and overall annual exergy gain for solar distillation system.

c. Life cycle conversion efficiency

It represents the net output of solar distillation system with respect to solar radiation falling on it for the entire life time. The ideal value of LCCE for solar distillation system is one. The system is considered better if LCCE is higher. Evaluation of LCCE considers the entire life of the system, whereas efficiency can be evaluated for a particular instant, hour or day. Following Tiwari and Mishra (2012), LCCE for PVT-CPC active solar distillation system can be expressed as

$$\text{LCCE based on energy} = \frac{E_{\text{out}} \times n - E_{\text{in}}}{E_{\text{sol}} \times n} \tag{4.33.87}$$

$$\text{LCCE based on exergy} = \frac{G_{\text{ex,annual}} \times n - E_{\text{in}}}{(\text{Annual solar exergy}) \times n} \tag{4.33.88}$$

Here, E_{sol}, n, E_{out}, E_{in}, and $G_{ex,\text{annual}}$ are, respectively, annual solar energy, life, overall annual energy output, embodied energy, and overall annual exergy gain for solar distillation system. Annual solar exergy can be calculated by multiplying E_{sol} with 0.933 (a factor obtained from the expression given by Petela, 2003).

Productivity of Solar Still

Productivity is the relationship between product and the factors used in attaining the product. It means getting more and more with less and less input of resources. It is generally expressed as the ratio of effectiveness and efficiency as per ILO (1979). One should keep in mind that yield and electricity gain can be considered as production in the case of solar still. It is not productivity. Production represents the volume, whereas productivity represents the effectiveness with which input of resources is used. Also, annual productivity should be more than 100% for a system to be feasible. Following Ashcroft (1950), Cox (1951), Benson (1952), and ILO (1979), productivity (η_p) of solar still can be written as

$$\eta_p = \frac{\text{Output of the system}}{\text{Input to the system}} \times 100 \tag{4.33.89}$$

The output consists of annual yield of water (kg) and annual electricity gained (kWh) and it can be expressed as

$$\text{Output} = R_w + R_e \tag{4.33.90}$$

where $R_w = \left(M_w \times (\text{SP})_w\right)$ and $R_e = E_e \times (\text{SP})_e$

Here, R_w, R_e, M_w, E_e, $(\text{SP})_w$, and $(\text{SP})_e$ are annual revenue (Rs.) earned from water, electricity gain, annual yield of water (kg), annual electricity gain (kWh), selling price of water (Rs./kg), and selling price of electricity (Rs./kWh), respectively. The daily yield can be evaluated by adding hourly yield for the period of 24 h. The monthly yield can be evaluated by multiplying daily yield with number of clear days observed in a month. The annual yield can be evaluated by adding monthly yield for 12 month.

The hourly electricity gain ($\dot{E}_e$) can be calculated with the help of Equation 4.33.60. The daily electricity gain can be evaluated by adding hourly electricity gain for 10 h because solar intensity exists for sunshine hours only. The monthly electricity gain can be evaluated by multiplying daily electricity gain with number of clear days observed in a particular month. The annual electricity gain can be evaluated by adding monthly electricity gain for 12 months.

Input represents the uniform end-of-year annual cost (UAC) of the solar still. UAC (Kumar and Tiwari, 2009; Tiwari and Ghosal, 2005) for a given initial investment of solar distillation systems can be expressed as

$$\text{UAC} = P_s \times F_{\text{CR},i,n} + P_s \times F_{\text{CR},i,n} \times M_s - S_s \times F_{\text{SR},i,n} \tag{4.33.91}$$

Here, P_s, S_s, and M_s are the net present cost, salvage value, and maintenance cost of the solar still, respectively. It is based on the concept of net present value. The factor $F_{\text{CR},i,n}$ is called capital recovery factor which converts the present cost into UAC and the factor $F_{\text{SR},i,n}$ is called sinking fund factor which converts the present cost into UAC (Tiwari, 2013). They can be computed as $F_{\text{CR},i,n} = \dfrac{i\times(1+i)^n}{(1+i)^n-1}$ and $F_{\text{SR},i,n} = \dfrac{i}{(1+i)^n-1}$, where i and n are the rate of interest and life of the system, respectively. Maintenance cost is generally taken as 10% of net present cost and hence value of M_s used in Equation 4.33.95 may be taken as 0.1. The first, second, and third terms used in Equation 4.33.95 are, respectively, the part of UAC corresponding to net present cost, maintenance cost, and salvage value of the solar still.

Following Equation 4.33.93, the hourly and daily productivity of solar still can be written as

$$(\eta_p)_{\text{hourly}} = \frac{[\text{Hourly yield}\times(\text{SP})_w]+[\text{Hourly electricity gain}\times(\text{SP})_e]}{\text{Hourly cost}}\times 100 \tag{4.33.92}$$

$$(\eta_p)_{\text{daily}} = \frac{[\text{Daily yield}\times(\text{SP})_w]+[\text{Daily electricity gain}\times(\text{SP})_e]}{\text{Daily cost}}\times 100 \tag{4.33.93}$$

The daily cost can be evaluated by dividing UAC with the number of clear days observed in a year and hourly cost can be evaluated by dividing daily cost with 24 h.

Here, productivity has been computed on the basis of energy. However, it should be noted that productivity can also be calculated on the basis of exergy.

Exergoeconomic Parameter for Solar Still

The evaluation of exergoeconomic parameter is an economic analysis method based on exergy/energy. It incorporates exergy/energy analysis with conventional cost analysis for improving the performance of energy systems (Tsatsaronis et al., 1993). Its aim is to estimate the cost-optimal structure and facilitate designers to find ways to enhance the performance of a system in a cost effective manner. It is considered as one of the best suitable methods for the design, analysis, and performance improvement studies of energy conversion systems. Exergoeconomic parameters (Tsatsaronis and Winhold, 1985) based on exergy loss (R_{ex}) and energy loss (R_{en}) can, respectively, be expressed as

$$R_{ex} = \frac{L_{ex,\text{annual}}}{\text{UAC}} \tag{4.33.94}$$

$$R_{en} = \frac{L_{en,\text{annual}}}{\text{UAC}} \tag{4.33.95}$$

where L_{ex}, L_{en}, and UAC are annual exergy loss, annual energy loss, and uniform end-of-year annual cost. UAC can be evaluated using Equation 4.33.99. The hourly exergy loss rate (L_{ex}) for solar still as reported by Tiwari et al. (2015) can be expressed as

$$L_{\text{ex}} = \left[(h_{cwg} + h_{rwg}) \left\{ (T_w - T_{gi}) - (T_a + 273) \ln \frac{(T_w + 273)}{(T_{gi} + 273)} \right\} \right] + \left[M_w \times c_w \left\{ (T_{wf} - T_{wi}) - (T_a + 273) \ln \frac{(T_{wf} + 273)}{(T_{wi} + 273)} \right\} \right] \quad (4.33.96)$$

where h_{cwg}, h_{rwg}, T_w, T_{gi}, M_w, c_w, T_{wf}, and T_{wi} are, respectively, convective heat transfer coefficient between water and inner glass surface, radiative heat transfer coefficient between water and inner glass surface, temperature of water, temperature of glass, mass of water in basin, final temperature of water, and initial temperature of water. The daily exergy loss can be computed by adding hourly exergy loss rate for 24 h. Monthly exergy loss can be evaluated as the product of daily exergy loss and number of clear days observed in a month. The sum of monthly exergy loss for 1 year will give annual exergy loss. Similarly, annual energy loss can be determined.

The definition of exergoeconomic parameter given by Equations 4.33.94 and 4.33.95 is appropriate if one wants to minimize the exergy/energy loss. However, it does not seem to be appropriate if one wants to compare different solar distillation systems as one need not to pay any penalty against exergy/energy loss. It happens because we do not have to pay anything for the corresponding input which is solar radiation. Also, solar radiation depends on climatic condition which cannot be controlled and hence we cannot provide the same input to different solar distillation systems while making comparison. In such case, the exergoeconomic parameter based on annual exergy/energy gain with an aim to maximize the exergy/energy gain seems to be more appropriate. So, exergoeconomic parameter based on annual exergy gain ($R_{g,\text{ex}}$) and energy gain ($R_{g,\text{en}}$) as reported by Singh et al. (2016) can be written as

$$R_{g,\text{ex}} = \frac{G_{\text{ex,annual}}}{\text{UAC}} \quad (4.33.97)$$

$$R_{g,\text{en}} = \frac{E_{\text{out}}}{\text{UAC}} \quad (4.33.98)$$

where $G_{\text{ex,annual}}$ represents annual overall exergy gain which can be computed with the help of Equation 4.33.63 and E_{out} represents annual energy gain which can be computed with the help of Equation 4.33.59.

Enviroeconomic Parameter for Solar Still

The evaluation of enviroeconomic parameter for solar still presents an approach to control the environmental pollution by providing economic incentives for attaining reductions in emissions of pollutants and encourages deploying renewable energy technology that does not emit carbon to the atmosphere. The enviroeconomic analysis is based on the price of CO_2 emission and emitted carbon quantity. Following Caliskan et al. (2012), the enviroeconomic parameter based on CO_2 emission reduction and CO_2 emission reduction price can be written as follows:

$$C_{CO_2} = c_{CO_2} \times x_{CO_2} \quad (4.33.99)$$

where C_{CO_2}, c_{CO_2}, and x_{CO_2} are enviroeconomic parameter or environmental cost reduction parameter, CO_2 emission reduction price per tCO_2, and reduction in CO_2 emission for the whole life of solar still (tCO_2), respectively.

If unit power is used by a consumer and the losses due to poor domestic appliances are around 20% and transmission and distribution losses are 40%, which is common in Indian conditions, then the power that has to be generated in the power plant is $\frac{1}{(1-0.2)(1-0.4)} = 2.08$ units. The average

carbon dioxide emission for electricity generation from coal is ~0.96 kg of CO_2 per kWh at source (Sovacool 2008). Thus, the amount of CO_2 emission for unit power consumption by the consumer is $2.08 \times 0.96 = 2$ kg.

The expression for reduction in CO_2 emission based on energy ($x_{CO_2,en}$) and based on exergy ($x_{CO_2,ex}$) for the whole life of solar still can, respectively, be written as

$$x_{CO_2,en} = (E_{out} \times n - E_{in}) \times 2 \times 10^{-3} \tag{4.33.100}$$

$$x_{CO_2,ex} = (G_{ex,annual} \times n - E_{in}) \times 2 \times 10^{-3} \tag{4.33.101}$$

where E_{out}, E_{in} $G_{ex,annual}$ and n are the annual energy available from the solar still, embodied energy, annual exergy gain and life of solar still, respectively.

Den Elzen et al. (2011) mentioned the international carbon price between 13\$/tCO_2 and 16\$/tCO_2 for the low and high pledge scenario. So, average value comes out to be 14.5\$/tCO_2. Hence, the value of c_{CO_2} can be taken as 14.5\$.

Production Cost of Water and Cost of Electricity Gain

The cost (C_w) of water production in Rs./kg and cost (C_e) of electricity gain in Rs./kWh for N identical PVT-FPC solar still (Singh et al., 2016), respectively, can be written as

$$C_w = \frac{\text{UAC} - R_e}{M_w} \tag{4.33.102}$$

$$C_e = \frac{\text{UAC} - R_w}{E_e} \tag{4.33.103}$$

If the value of (UAC$-R_w$) becomes negative, then it can be taken as zero because negative cost does not make any sense. It will make value of C_e zero which will mean that we are getting electricity at no cost. In the case of passive solar still, no electricity is produced and hence C_e will not appear there.

Passive Solar Stills

A conventional single-slope passive solar still has already been explained in Section 4.33.3. The overall thermal efficiency of this solar still varies from 5% to 40% throughout the year. It has been modified to enhance its performance by increasing heat gain in the basin resulting in higher evaporation/condensation and reducing heat losses to ambient. Some modified passive solar stills are as follows.

Modified Single-Slope Solar Still

The conventional single effect solar still has low efficiency due to poor dissipation of latent heat of the condensing vapor at glass cover surface. A modification of conventional solar still is multieffect solar still which utilizes the latent heat of condensing vapor. It consists of a number of basins which contains brackish water. The vapor evaporated from first basin releases its latent heat at the lower surface of second basin. This latent heat is used to heat the water in the second basin rather than being wasted to atmosphere. This process of transferring latent heat to the second basin placed above the first basin is called multieffect in solar still. The material used for additional basin is transparent so that radiation can be allowed to lower basin. These stills produce a larger quantity of water in comparison to convention single effect solar still for the same area of basin because latent heat of condensation of vapor is utilized to heat water. However, multieffect solar stills are costlier as they need additional material. Cooper and Appleyard (1967) designed a multieffect solar still for optimum utilization of heat energy and decreasing the losses by multiple condensation and

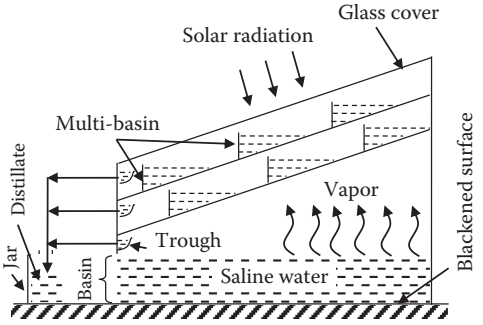


FIGURE 4.33.10 Multieffect multibasin solar still.

evaporation inside a single unit of solar still. Lobo and Araujo (1977) reported an increase in the cost of standard still per unit area by 15–25% when intermediate basins are added to it. Hence, they designed a multieffect multibasin solar still (Figure 4.33.10) to overcome this problem. They concluded that the double-effect solar still would produce 40%–55% more distillate than standard solar still at 20–25 MJ/m^2/day and cost of distillate from multieffect solar still could be 20%–25% lower because of the optimal utilization of space.

Sodha and Tiwari (1979) investigated the performance of double-basin solar still for various systems and climatic conditions. He has concluded that double basin still gives 36% higher yield than that of a single basin and observed that the yield increases rapidly as the insulation thickness is increased up to 4 cm and then it increases rather slowly. Al Mahdi (1992) predicted the effect of number of basins on the daily yield of the still. He has observed that the double basin still gives the highest peak. However, the triple and quadruple basin stills produce higher daily yield due to appreciable amounts of distillate during the night. But, addition of more than four basins does not have significant effect on yield. It may happen due to the fact that increase in temperature of water will be lesser due to higher water mass.

El-Sebaii (2005) studied the triple basin solar still and reported that the total daily yield of the still was 12.635 kg/m^2 on a typical summer day of Tanta (latitude 30° 47′ N). Tanaka et al. (2000, 2002) designed an improved single-slope solar still called single-slope single-basin multieffect solar still (Figure 4.33.11) for getting higher yield. In this case, a wetted wick is used to cover the outer portion of vertical wall to facilitate condensation on its inner surface. The condensation takes place at the inner surface of vertical wall due to its lower temperature in comparison to glass cover and latent heat is released to the vertical wall which results in heating the brine water trickling over it

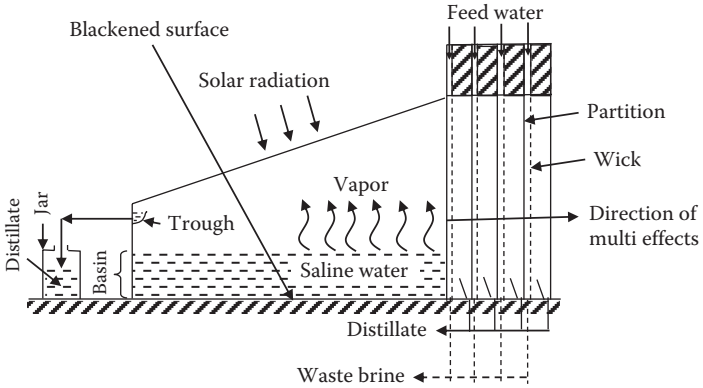


FIGURE 4.33.11 Single-slope single-basin multieffect solar still.

and further evaporation occurs. This evaporated water is condensed at other metallic wall covered with wetted wick after releasing the latent heat and this process continues.

A further improvement of single-slope solar still is regenerative-type solar still (Prakash and Kavathekar, 1986). In this case, saline water is allowed to flow over the glass surface (Figure 4.33.12, Bapeshwararao et al., 1983) which results in the improvement of yield of upper basin as well as efficiency of system. Reduction in the temperature of glass surface results in the increased circulation of air inside the still which leads to higher evaporative and convective heat transfer between water and glass (Toyama et al., 1983). Also, lower glass surface temperature increases the rate of condensation. The flowing fluid (water) over the glass cover surface is preheated before going into the basin. The glass cover temperature decreases and yield increases by intermittent flow (Madhuri and Tiwari, 1985) or continuous flow (Abu-Hijleh, 1996) of raw cooling water on the glass cover (Abu-Arabi et al., 2002). A/K Abu-Hijlew and Hasan (1997) reported that the production rate was increased by 20% when water was allowed to flow over condensing cover. Zuriat and Abu-arabi (2004) investigated the modeling and performance of regenerative solar desalination systems and concluded that the yield was increased by 34%–35% for the month of March and June for Oman weather condition.

Dutt et al. (1989) studied the effect of adding dye in the double basin still and concluded that the system efficiency was increased by 10% by adding dye in the double basin still because of increased yield due to higher increase in temperature of water. They also reported that the evaporative heat transfer coefficients were strong functions of temperature; however, variation of convective and radiative heat transfer coefficients was weak with temperature. The performance of multibasin solar still depends strongly on the mass of water in the basin. At lower mass in the basin, higher distillate is obtained because of higher increase in water temperature which results in larger difference of temperature between water surface and glass surface. Tiwari et al. (1991) studied the performance of double-basin solar still and reported that the total production of distillate was strongly dependent on depth of water in the lower basin. At lower water depth in basin, a higher yield was obtained; however, no distillation was obtained during off-sunshine hours. At higher depth of water, a lower amount of distillate was obtained, but distillation process continued for off-sunshine hours also due to heat content of water mass.

An experimental study on upward-type double-effect solar distillers with air flow through the second effect (Figure 4.33.13) was carried out by Yeh (1993). The inclination of still was 10° with the horizontal surface. A lower inclination is kept if the solar still is to be used for summer condition. He concluded that the upward-type double-effect solar distillers were more efficient than downward-type units. Hamdan et al. (1999) investigated the performance of the single, double, and triple basin

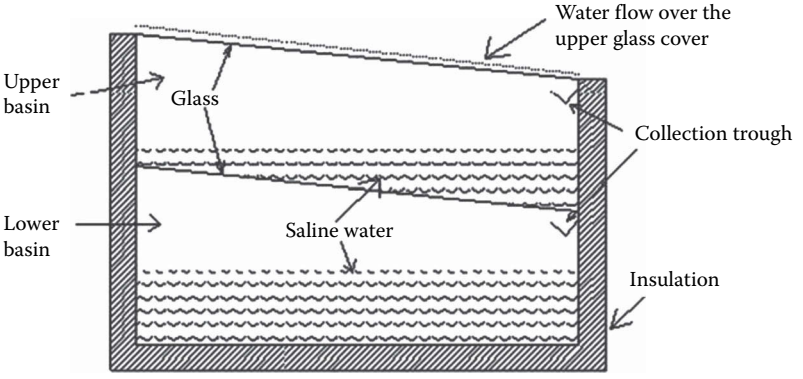


FIGURE 4.33.12 Double basin solar still with water flowing over upper glass cover. (From Bapeshwararao, V.S.V. et al., *Energy Convers. Manag.*, 23(2), 83–90, 1983. With permission.)

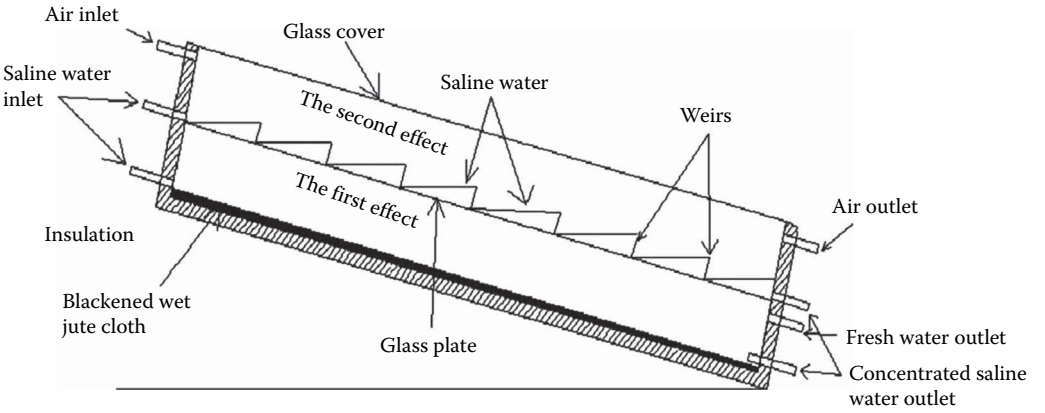


FIGURE 4.33.13 Upward-type double effect solar still. (From Yeh, H., *Energy*, 18(11), 1107–1111, 1993.With permission.)

pyramid solar stills under Jordanian climate and concluded that the triple effect solar still gave the highest daily efficiency followed by double and single effect due to similar variation in daily yield.

Cappelletti (2002) investigated experimentally the double-basin solar still (Figure 4.33.14). He designed the upper basin of the still as "V" shape and the provision was made to collect the condensed water of the lower basin in the middle. The efficiency of the still was found to be about 16% due to low temperature of the water in basin (around 50°C). Al-Hinai et al. (2002) reported the parametric investigation of a double-effect solar still and it was compared with a single effect solar still. The basin of solar still was made shallow with asphalt coating. They found an average annual yield of solar still as 4.15 and 6 kg/m^2/day for single and double-effect solar stills, respectively. Fukui et al. (2004) investigated theoretically a new maritime, life-saving, multiple-effect solar still design and concluded that the yield was increased with the use of multieffect still because of the utilization of latent heat of condensation. Al-Karaghoulia and Alnaserb (2004) investigated experimentally the performance of single and double-basin solar stills and a higher production rate was

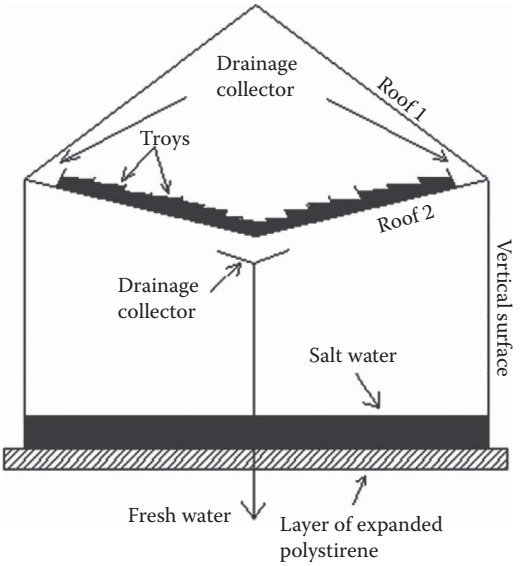


FIGURE 4.33.14 Double basin solar still. (From Cappelletti, G.M., *Desalination*, 142, 221–228, 2002. With permission.)

reported with the double basin still. It occurred because temperature of glass in the case of double basin was lower which in turn occurred due to entrapment of solar radiation in two basins in the case of double basin solar still.

They also reported that the insulation had significant effect on the productivity, especially for the double-basin type because heat losses are lower in insulation is used.

An inverted absorber solar still is another modification of single-slope solar still. It is a concentrating type of solar still and consists of a conventional solar still with a curved reflector below its metallic basin plate. Figure 4.33.15 shows the schematic diagram of inverted absorber double-basin solar still. This type of solar still receives solar radiation from top as well as from bottom. The reflector provided below the metallic basin concentrates the solar radiation on the lower portion of basin. Hence, temperature of water in this type of solar still increases at a faster rate in comparison to conventional single-slope solar still due to additional heat gain which results in higher yield. Tiwari and Suneja (1998) studied theoretically the single-basin inverted absorber single-slope solar still and reported daily yields of 5.1, 4.8, and 4.5 L/m^2 at water depths of 0.01, 0.02, and 0.03 m, respectively. Further, Sangeeta and Tiwari (1999a) studied an inverted absorber double-basin solar still to assess the effect of water depth on its and concluded that the yield of system was maximum when the depth of water was least in the lower basin. It happened due to higher increase in temperature at lower depth because heat loss corresponding to heat capacity was lower as water mass was lesser at lower depth.

Sangeeta et al. (1997) investigated the parametric study on inverted double-effect solar still and reported a 10% increase in yield if the initial water temperature was raised from 22°C to 35°C. The yield of an inverted absorber double-basin solar still was found to increase by about 99% as the absorptivity of the absorber increases from 0.55 to 0.95. It happens because loss decreases with the increase in absorptivity and more heat is transferred to water due to higher temperature difference between water mass and absorber. Sangeeta and Tiwari (1999b) studied an inverted absorber triple effect solar still and reported that the yield of inverted triple effect solar still was higher than the yield of conventional triple effect solar still because higher solar intensity was available for water mass due to reflector provided at the bottom. They also concluded that the enhancement in daily yield was significant as the number of effects increased from the first effect to the third effect. It happened due to better utilization of latent heat of condensation and better evaporation. There is a marginal increase in the yield in the case of inverted solar still if the number of basins is increased beyond seven. Also, the yield from a seven-effect inverted absorber solar still is about 4.2 times the yield from an inverted absorber single-basin solar still (Sangeeta and Tiwari, 1998).

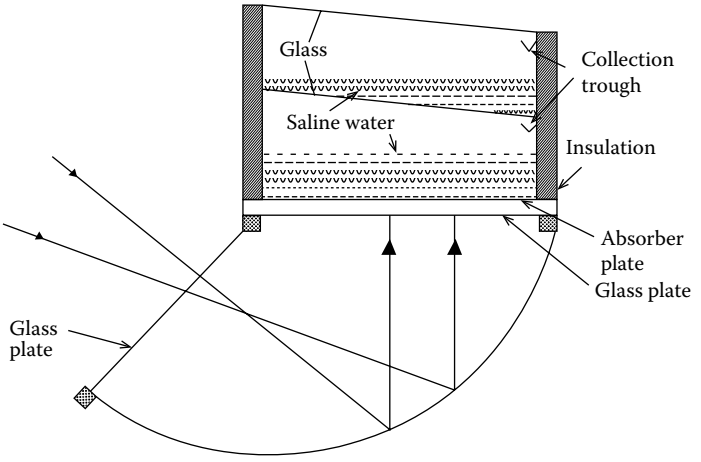


FIGURE 4.33.15 Inverted absorber double basin solar still. (From Sangeeta, S., Tiwari, G.N., *Energy Convers. Manage*., 40, 1885–1897, 1999a. With permission.)

Addition of condenser with the conventional solar still is another modification. The condensation rate is low in a simple conventional still because the inner surface of the glass cover is only area available for condensation and the difference of temperature across glass thickness is very small. Hence, an additional area for condensation is added to conventional still to increase the rate of condensation which results in higher yield. Heat capacity of the evaporation area increases due to the provision of additional area and the latent heat of condensation is used by water in upper basin of condenser area. Fath (1996) investigated the performance of a two effect solar distillation unit shown in Figure 4.33.16. In this type of solar still, the evaporated vapor from the first effect is allowed to be purged to the second effect. Fins are provided at the outer surface of the second effect cover to maintain the cover at or near-ambient temperature, and inner surface is reflective to minimize radiation losses. He reported the daily yield for the proposed unit as 10.7 kg/m^2/day under the climatic conditions of the city of Dhahran, Saudi Arabia. A higher yield is obtained due to better utilization of heat obtained from solar intensity and lower heat losses.

Madhlopa and Johnstone (2009) investigated the performance of a passive solar still with separate condenser shown in Figure 4.33.17. It consists of three basins. One basin (basin 1) lies in the

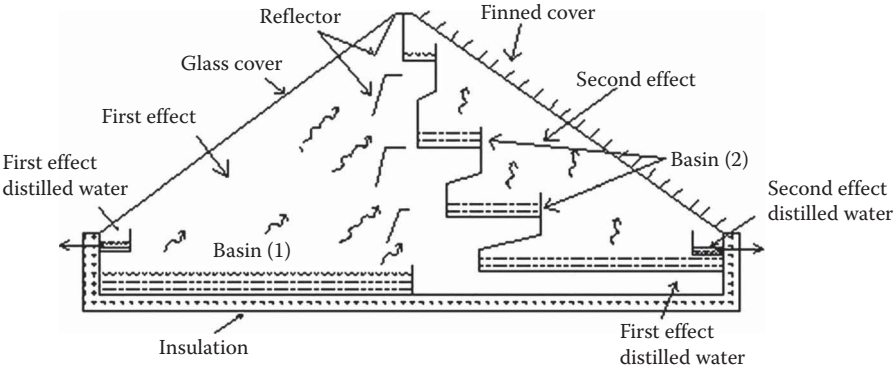


FIGURE 4.33.16 Solar still with inbuilt condenser. (From Fath, H.E.S., *Desalination*, 107, 223–233, 1996. With permission.)

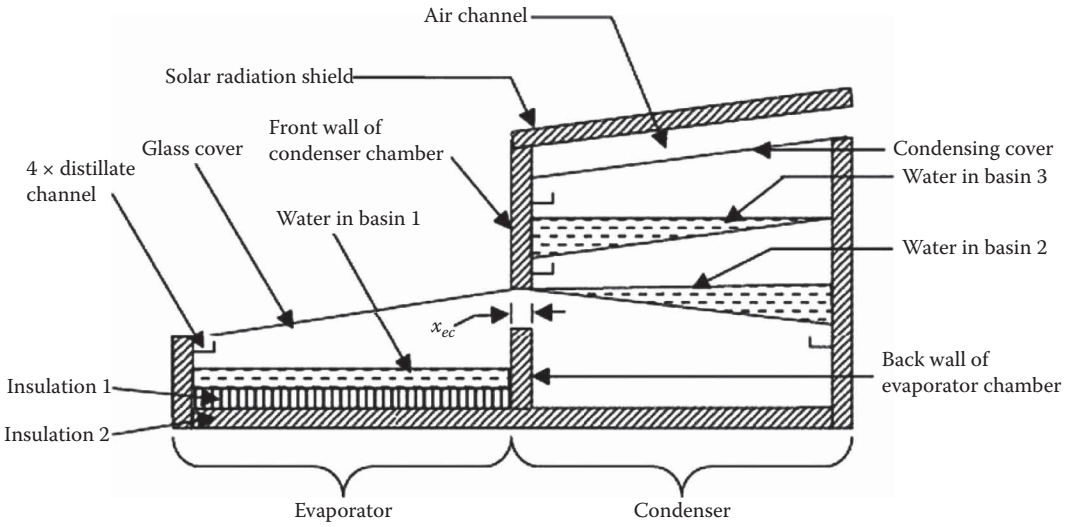


FIGURE 4.33.17 Solar still with separate condenser. (From Madhlopa, A., Johnstone, C., *Renewable Energy*, 34, 1668–1677, 2009. With permission.)

evaporation chamber and two other basins (2 and 3) lie in the condenser chamber. The evaporator basin is covered with glass cover and basin 3 is covered by an opaque condensing cover. The top part of the condensing cover is shielded from solar radiation so that the cover remains relatively cool. Some part of water vapor from the first effect condenses under the glass cover. The remainder of it flows into the condenser by purging and diffusion and condenses under the liner of basin 2. After analyzing the system, they inferred that the distillate of the still was 62% higher than that of the conventional type because of faster condensation/evaporation.

Single basin solar still with baffle absorber is another modification. In this type of solar still, aluminum or some other suitable material sheet is placed in the water mass. This plate is painted black at the top surface and thermally insulated at the bottom surface and it is commonly known as baffle. The plate floats in water and divides the unit into a shallow basin evaporative zone above the plate and heat storing zone below the plate. The plate absorbs a part of solar radiation entering the still which results in increasing the temperature of plate and heat is transferred to the brackish water in heat-storing zone by conduction. No heat transfer takes place between the two zones by convection through a gap left between the plate and side walls because hot water remains above the plate due to its lower density. Heat transfer by conduction is very small between the water above and below the plate because plate is thermally insulated at the bottom. The temperature of water above the plate becomes lower than the temperature of heat-storing zone at night. So, convection current is developed between the two zones and cold water above the plate is replaced by hot water below the plate through the gap left between the plate and side walls. This helps in the continuation of evaporation and condensation process till the temperature between water and glass is maintained during nocturnal production.

El-Sebaii et al. (2000) investigated theoretically and experimentally the effect of using aluminum plate as baffle suspended absorber in single-basin solar still and reported that the yield is about 20% higher in comparison to conventional single-slope solar still because of decrease in preheating time of the upper water column of the modified still. Further, Rahim (2003) used a painted black aluminum sheet to divide the basin water into an evaporating zone above the plate and a heat-storing zone below the plate. He reported that the system did not need any source of external power for storing and recovering processes. Srivastava and Agrawal (2013) analyzed single-slope basin-type solar still experimentally and theoretically by integrating multiple low-thermal inertia porous absorbers floating adjacent to each other. They reported about 15%–48% enhancement in yield in comparison to conventional solar still because of lower heat loss and increased thermal capacity of base fluid due to the presence of absorber plate in water mass. El-Bialy (2014) investigated experimentally the effect of basin water mass, the type of the floating absorber plate and thermal conductivity of plate on the daily productivity and efficiency of single-slope single-basin solar distiller with a floating absorber (SBDFA) under the weather conditions of Tanta, Egypt (Latitude 30° 47′ N) during typical summer days. Figure 4.33.18 shows the schematic diagram of SBDFA. The result obtained from SBDFA was compared with conventional solar still for the same weather condition. For all investigated values of the mass of water above the floating absorber, the yield of the SBDFA was found to be higher in comparison to conventional solar still under the same weather condition. The daily yield was found to be 42.2%, 15.2%, 20.1%, and 17.2% higher than conventional solar still when mica, stainless steel, aluminum, and copper were used as floating absorbers. He also indicated that SBDFA can be used for supplying potable water with reasonable amounts.

Double Slope Passive Solar Still

The cross-sectional view of double-slope solar still is shown in Figure 4.33.19. It consists of two condensing glass covers facing east and west direction and a basin which is filled with saline water to be purified. The working principle of double-slope solar still is similar to single-slope solar still. It can harness the solar radiation for the entire day unlike single-slope solar still which harnesses the solar radiation for the limited time during the day. Rubio et al. (2004) studied the performance of double-slope solar still for the climatic condition of Mexico. They developed thermal model and validated by experimental results. Further modification of double-slope solar still can be done to collect rain

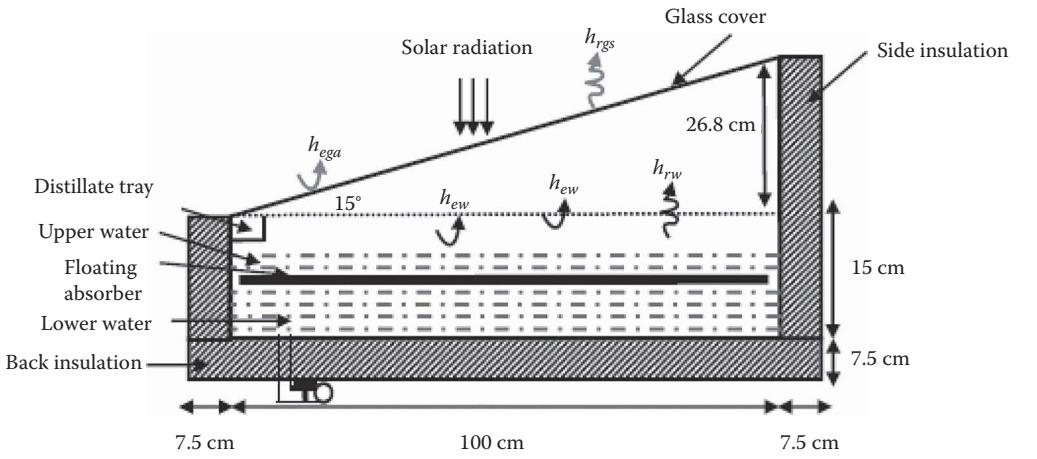


FIGURE 4.33.18 Single slope single basin solar distiller with a floating absorber. (From El-Bialy, E., *Energy*, 68, 117–124, 2014. With permission.)

water by extending the height of walls. Rain water is distilled water and can be used directly. During rainy season, solar still becomes nonfunctional due to unavailability of solar radiation. Rain water can be used during that time.

A further modification to double-slope solar still is the integration of additional basins inside the solar still. Kumar et al. (1991) investigated the effect of additional basin in double-slope solar still. Their conclusions are as follows:

1. Yield decreases with the increase in depth as the increase in temperature of water is lower at higher depth due to higher heat capacity of water mass.

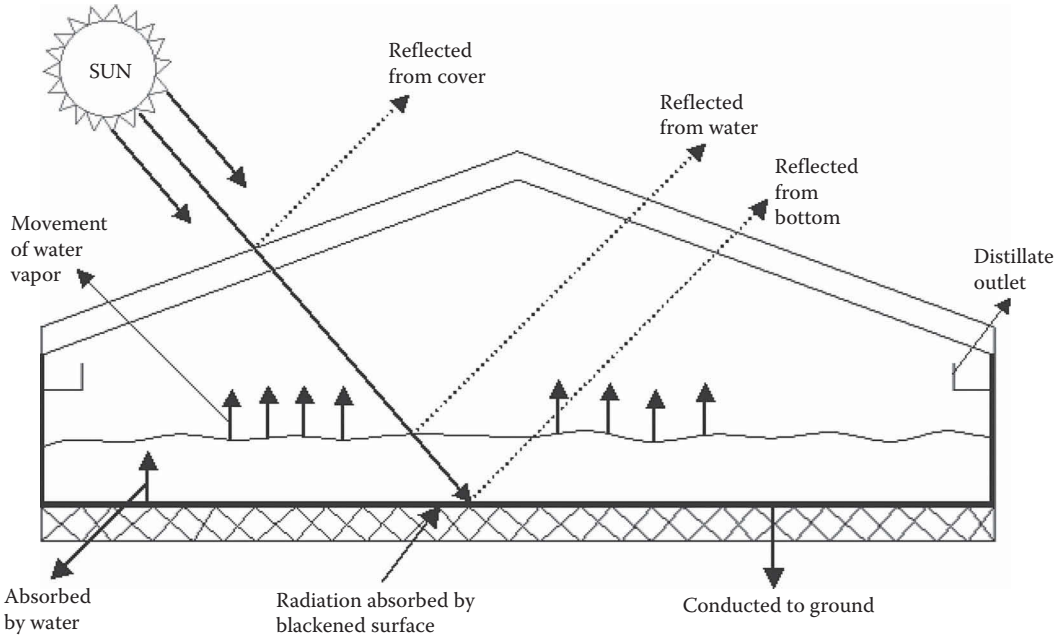


FIGURE 4.33.19 Cross-sectional view of a double slope passive solar still placed in east–west direction.

2. The daily yield increases with the decrease in depth of water in the upper basin. It happens due to better utilization of the latent heat of vaporization.
3. The lower basin produces higher distillate than upper basin.
4. Double basin solar still has higher efficiency than single-basin solar still.

A study of double-basin double-slope solar still was conducted by Rajaseenivasan et al. (2013) experimentally with diverse basin conditions. Figure 4.33.20 represents the schematic diagram of double-basin double-slope solar still. The results obtained from double-slope double basin solar still were compared with the conventional single-basin double-slope solar still. The upper basin of the double basin still was split into three compartments (each side). It was done to ensure even distribution of water mass through the entire area of the upper basin. The tests were performed with different depth of water in the basin taking mild steel pieces as energy storing material. Black cotton cloth, jute cloth, and waste cotton pieces were used as wick materials and small clay pot as porous material. They concluded that the yield of the double-basin solar still with energy storing materials in lower basin was 169.2% higher than the conventional still. Zeroual et al. (2011) and Rajamanickam and Ragupathy (2012) studied south–north oriented of double-slope passive solar still with partially cooled condenser and reported an increase of 11.82% in the yield.

Effect of Nanofluids on the Performance of Single and Double Slope Passive Solar Stills

The term nanofluid was coined by Choi (1995) for the first time. Nanofluids having water as solvent are obtained by mixing nanoparticles with water in a particular concentration along with sodium dodecyl benzene sulfonate for maintaining even suspension. They are stirred by magnetic stirrer for around 15 min and then ultrasonicated with the help of ultrasonicator. The prepared sample is then tested for the computation of stability and thermal conductivity. Zeta Potential is used for stability testing. Aluminum oxide (Al_2O_3), zinc oxide (ZnO), iron oxide (Fe_2O_3), tin oxide (SnO_2), or some other material can be used as nanoparticles. The size of particles is of the order of nanometer. The yield of solar still depends on stability and thermal conductivity of fluid. However, thermal conductivity dominates over stability and hence it is the most important factor while using nanofluid. The thermal conductivity of nanofluid can be controlled by the concentration and aspect ratio. Higher the thermal conductivity, sooner will the fluid be heated, and higher temperature difference between water surface and glass cover will be obtained which will positively affect the yield. Unlike conventional fluid used in basin, water receives heat from both basin liner and metallic nanoparticles. Elango et al. (2015) investigated the performance of single-slope passive solar still with different water nanofluids experimentally. They used nanofluids of aluminum oxide (Al_2O_3), zinc oxide (ZnO), iron oxide (Fe_2O_3), and tin oxide (SnO_2) of different concentrations. They concluded that single-slope solar still with Al_2O_3, ZnO, and SnO_2 nano-fluids had 29.95%, 12.67%, and 18.63% more production respectively than single-slope solar still with water. It happens because thermal

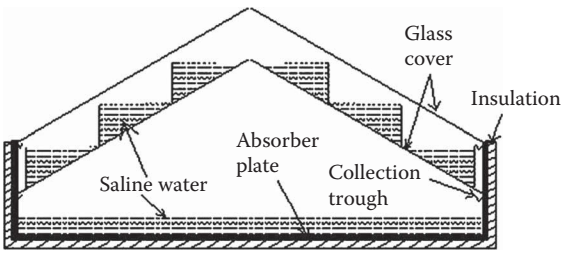


FIGURE 4.33.20 Double basin double slope solar still. (From Rajaseenivasan et al., *Desalination*, 319, 25–32, 2013. With permission.)

conductivity of Al_2O_3 nanofluid is highest followed by ZnO and SnO_2. They also observed that the payback period of nanofluid was 2.85 years which was lesser than that of the still with water as the basin fluid due to higher yield.

Sahota and Tiwari (2016a,b) investigated theoretically the performance of double-slope solar still using nanofluids. Sahota and Tiwari (2016a) investigated the effect of Al_2O_3 nanofluid on the performance of double-slope passive solar still and concluded that yield of solar still is higher by 12.2% than solar still with water at 35 kg mass of water. It happened because of increase in heat transfer coefficient and fluid temperature. They have also concluded that the daily yield increases with the increase in the concentration of nanoparticles in water because of increased surface area for heat transfer from nanoparticles to water at higher concentration of nanoparticles. Further, Sahota and Tiwari (2016b) studied theoretically the performance of double-slope passive solar still for optimized concentration (0.25%) of metallic nanoparticles namely Al_2O_3, titanium oxide (TiO_2), and copper oxide (CuO). They concluded the following on the basis of their studies.

1. The value of fluid temperature difference between nanofluid and water increases with an increase in the concentration of nanofluids because of increased surface area for heat transfer from nanoparticles to nanofluid at higher concentrations.
2. Thermal energy, exergy, thermal efficiency, and exergy efficiency are higher for Al_2O_3 nanofluid followed by TiO_2 and CuO nanofluids. It happens due to the increased thermophysical properties of Al_2O_3 nanofluid in comparison to other nanofluids under study.

Thermal–Electrical Solar Still with Boosting Reflector

Figure 4.33.21 shows the thermal–electrical solar still with boosting reflector and outside condenser proposed by Monowe et al. (2011). This solar still needs solar radiation to evaporate water from a still and a small amount of electrical energy input to run low power electrical fan (25 W/m^2 of basin area) which forces water vapor into the outside condenser. The vapor in the condenser releases its latent heat which can be used to preheat the feed water for the next batch or even for operating the still at night. This heated water can also be used for domestic purpose. Hence, loss of latent heat of condensation to the environment is minimized. The boosting reflector concentrates the solar energy to the glass. A simple railing device is provided to adjust the angle of the boosting reflector from visual observation. It can also be used to protect the glazing of the still during a hail storm or dust storm. A tank is connected to the other end of the condenser coil to collect the distilled water. Preheated saline water must be removed from the condenser every day to keep the efficiency of the still at a high level. This still is portable as saline water is charged in from the borehole reservoir using hand-operated mechanical pump.

Monowe et al. (2011) investigated the performance of thermal electrical solar still with boosting reflector and outside condenser. They have reported the efficiency of still can be as high as 77% if the preheated saline water was employed for domestic purposes. Further, efficiency can be as high as 85% if preheated saline water could be used to run the still during night times and to recharge the still by the next batch of preheated water.

Wick-Type Solar Still

A basin-type solar still takes some time to increase the temperature of water and then for evaporation of water because of heat storage due to heat capacity of water. This drawback is overcome by wick-type solar still. Figure 4.33.22 shows the schematic diagram of multi wick-type solar still (single effect). It mostly comes under inclined-type solar still. In a wick still, a water reservoir is provided at the top of the absorber and one end of a number of jute cloth pieces (wicks) are dipped into the reservoir while other ends are spread over absorber. A thin sheet of polyethylene is provided between two jute cloth pieces. The feed water flows slowly through a porous wick which is also called radiation-absorbing pad. Water gets evaporated immediately after receiving the solar

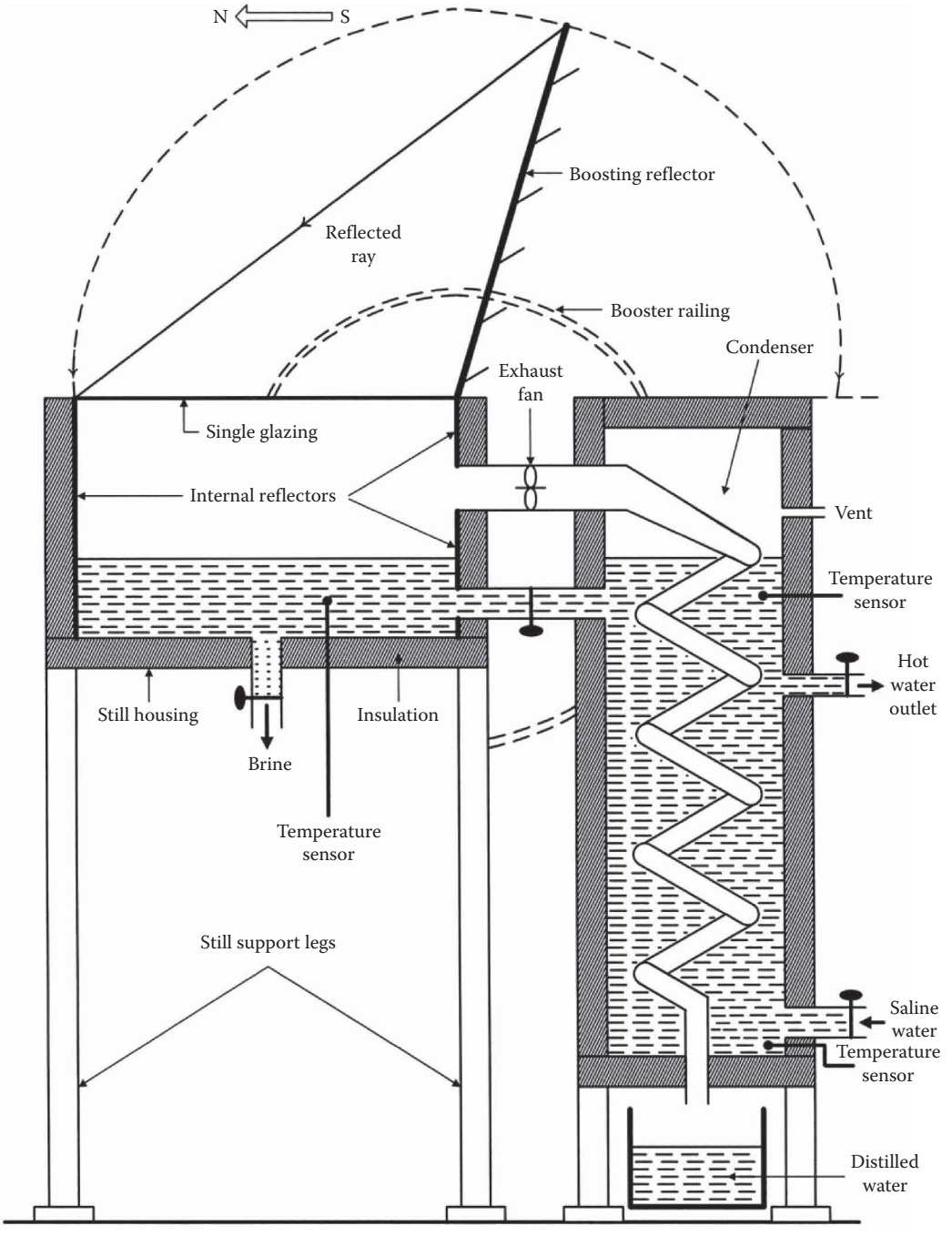


FIGURE 4.33.21 Thermal–electrical solar still with boosting reflector and outside condenser. (from Monowe P., Masale M., Nijegorodov N. and Vasilenko V., *Desalination*, 280, 332–338, 2011.)

radiation by the pad and condenses at the inner surface of glass. There are generally two advantages of wick still over basin still. First, the wick still can be made inclined so that the feed water presents a better angle to the sun. It results in the reduction of reflection. Second, depth of water is very small due to lesser feed water in the still at any time and hence the water is heated more quickly to a higher temperature. The main disadvantage of such type of still is that it does not produce distillate while

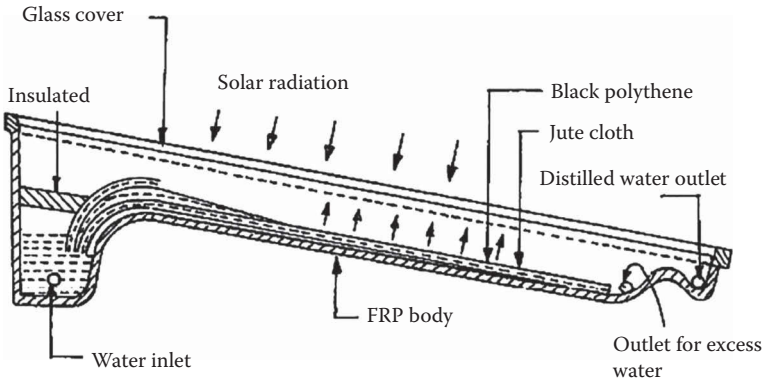


FIGURE 4.33.22 Wick-type solar still.

cloud passing or after sunset. However, basin-type solar still continues to produce distillate for some time due to heat stored in the basin water.

Sodha et al. (1981) and Tiwari and Salem (1984) investigated multiwick solar still and reported that the heat storage effect was minimized because of continuous flow of feed water through the capillary effect. The inverted multiwick double-effect solar still produced 20% higher daily yield than the conventional multieffect still. The yield was enhanced with the decrease in flow rate for a particular length of still. Also, the number of effect has a significant effect on the total yield of the still. The daily yield was increased with an increase in inlet water temperature, particularly in the second- and third-stage distillation (Tiwari et al., 1989). Singh and Tiwari (1992) studied the performance of the double-effect multiwick solar still shown in Figure 4.33.23 and reported a decreases in overall thermal efficiency with the increase in mass flow rate. Also, thermal efficiency is always higher in the first effect than that of the second effect.

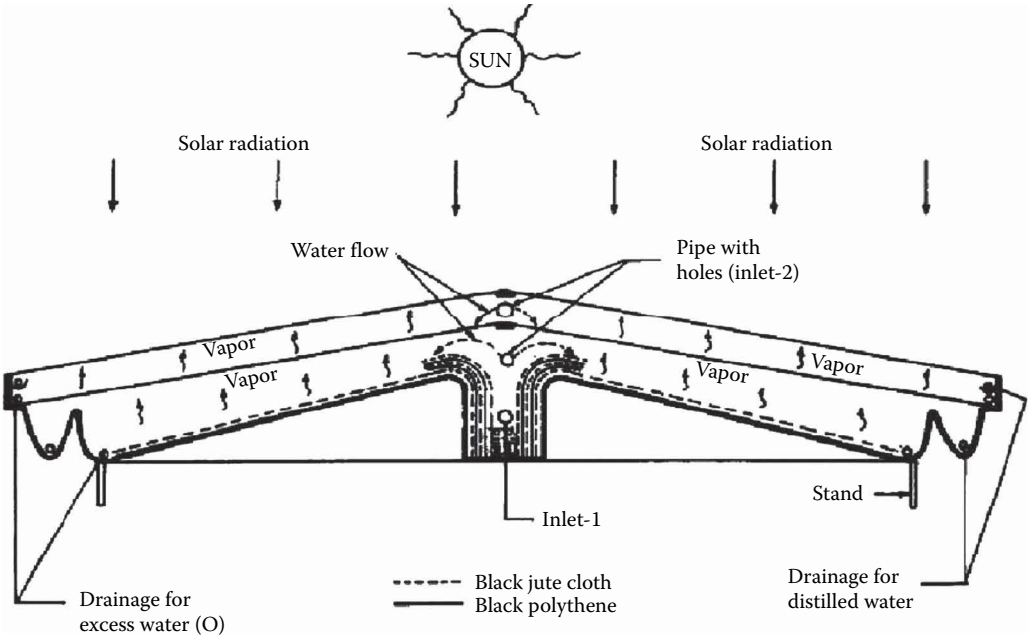


FIGURE 4.33.23 Double effect, wick-type solar still.

Yeh and Chen (1992) investigated wick-type, double-effect solar distillers with air flow through the second-effect unit experimentally. Water vapor was carried out from the second-effect unit to improve the production rate. There was a significant increase in productivity when the airflow was conducted at the optimal rate, especially for lower strengths of solar radiation. Kumar and Anand (1992) investigated the tubular multiwick solar still and found a higher distillate output of about 8%, 13%, and 18% for tubular multiwick solar still in comparison to tubular solar still, the simple multiwick solar still and conventional basin-type solar still.

Active Solar Still

If some external source is added to passive solar still to transfer thermal energy to water mass in the basin, then the solar still is called active solar still. A higher increase in the temperature of water is obtained which results in a higher temperature difference between water surface and glass cover surface. A higher evaporation is obtained as a consequence of higher temperature difference and ultimately higher distillate is produced. The thermal energy can be added to water mass either directly through solar collector/concentrator panel or indirectly through heat exchanger. Most of the passive solar stills discussed in Section 4.33.10 can be operated in active mode by integrating with source of thermal energy. This type of solar still is generally more suitable for commercial purpose as the initial investment is higher. However, it can be used for domestic purpose also.

FPC-Integrated Solar Still

The integration of FPC to basin-type solar still was proposed by Soliman (1976). The water in the basin is heated up directly by solar radiation and it also receives heat while passing through the pipe in the FPC. Pipes in FPC are provided beneath the absorber plate to gain maximum amount of solar radiation. This type of solar still operates at higher temperature of 70°C–80°C. However, the water temperature can be increased to 100°C. Some other fluids may be used in place of water like nanofluids/oils for temperature higher than 100°C. FPC-integrated solar still was also investigated by Badran et al. (2005) and reported a significant increase in yield. The yield of this solar still was 36% higher than passive solar still because of higher increase in water temperature as external heat is provided in addition to solar intensity. Further, a thermal model of FPC-integrated active solar still was developed by Rai and Tiwari (1982) and it was validated by experimental results. Performance of the system was further improved through regeneration (Bapeshwar and Tiwari, 1984; Dutt et al., 1988). In this type of solar still, water is allowed to flow over the glass cover and this preheated water is fed to basin. The glass surface temperature is reduced because of water flowing over it which results in the creation of higher temperature difference between water surface and glass surface and hence improvement in yield is obtained. The number of FPC can be optimized for getting optimum distillate output which was done by Kumar and Tiwari (1996a,b). They reported the optimum number of collectors as 10, 9, and 6 for single, double, and triple effect, respectively.

In the case of FPC-integrated active solar still, external electric source is required to run the pump and it is not self-sustainable for operation in remote areas. Hence, to make the system self-sustainable so that it can operate in remote areas also, FPC can be covered partially or fully by PV module. Such type of system is called hybrid PVT active solar still. Kumar and Tiwari (2010) designed, fabricated, and analyzed the performance of hybrid PVT-FPC single-slope active solar still. This system was equipped with PVT to operate in remote location without grid connection as the need of electrical was fulfilled by PVT output. The effective area of PV module was taken as 0.605 m2 with 36 solar cells connected in series. They integrated two FPCs in series to the basin of still. A DC motor pump was inserted between the basin and inlet to the first collector and they were connected through insulated pipe. Out of the two, only one FPC was partially covered by PV module. They conducted the experiment for different water depth and reported the maximum daily yield of system as 7.22 kg/m^2 at 0.05 m water depth. It was also observed that the system was self-sustainable to operate in remote locations. Singh et al. (2011) extended the same work for

double-slope active solar still and concluded that the yield was 1.4 times higher than that obtained for hybrid PVT-FPC single-slope active solar still. The reason being that the moderate solar intensity exists for longer period in the case of double-slope active solar still and also heat loss is lower due to lower surface area of double-slope active solar still. It was also concluded that the total cost of the system was 14% lower than hybrid PVT-FPC single-slope active solar still. It happened due to the geometrical shape of double-slope solar still as the material requirement in this case is lower if the same basin area is maintained.

The work of Kumar and Tiwari (2010) was further extended by Tiwari et al. (2015) and Singh et al. (2016). The solar still coupled with two identical partially covered PV module FPCs is shown in Figure 4.33.24a,b. Tiwari et al. (2015) investigated this active solar distillation system for exergoeconomic and enviroeconomic parameter for New Delhi climatic condition. The hourly thermal, exergy, electrical, overall exergy, and overall thermal efficiency was also computed and results were compared with the results obtained by earlier researchers. They concluded that daily thermal efficiency of this solar distillation system was less than systems designed by Kumar and Tiwari (2010) and Singh et al. (2011) because lesser area of collector is exposed to sunlight to gain thermal energy as both collectors are partially covered in this case. However, daily exergy efficiency of the system was higher than the system designed by Kumar and Tiwari (2010) but slightly lower than the system proposed by Singh et al. (2011). Further, the daily electrical exergy efficiency, the daily overall exergy efficiency and the daily overall thermal efficiency of solar distillation system under study were higher than the systems designed by Kumar and Tiwari (2010) and Singh et al. (2011). It happened because of the presence of larger area of PVT as both collectors are partially covered with PVT which produces higher electrical energy/exergy.

Singh et al. (2016) investigated active solar still integrated with two hybrid PVT collectors experimentally under the composite climate condition of New Delhi. They developed the thermal model of the system and validated with experimental results. The annual productivity was found to vary between 120.29% and 883.55% depending on the rate of interest and life of solar still. The annual productivity was obtained more than 100% which indicated that the system was feasible. There are two advantages of such type of solar still. First, it is self-sustainable to operate in remote areas where electricity has not reached so far but solar energy is available. Second, besides producing potable water, it can provide DC electrical power during sunshine hour which can be used for lighting or some other purpose.

PVT-CPC Water Collector Solar Still

Single and double solar stills integrated with N identical PVT-CPC collectors were studied theoretically by Singh and Tiwari (2016) to assess their performance by incorporating the effect of EPBT under optimized condition. They have presented the comparative analyses of single-slope PVT-CPC active solar still (Figure 4.33.8) and double-slope PVT-CPC active solar still (Figure 4.33.9) at 0.14 m water depth for the same basin area under similar climatic condition (New Delhi). They concluded the following based on their studies.

1. The optimum number of collectors and mass flow rate for these systems are 7 and 0.04 kg/s, respectively.
2. Double slope has higher daily yield than single-slope PVT-CPC active solar still if depth of water in basin is <0.19 m and vice versa. It happens because a moderate-temperature difference between water and glass surfaces exist for a longer period in the case of double-slope PVT-CPC active solar still. However, in the case of single-slope PVT-CPC active solar still, the difference of temperatures is higher but it survives for a shorter period. At higher depth, the sensible heat of water mass is a dominating factor and it is higher for single-slope PVT-CPC active solar distillation system due to higher daily solar energy.
3. The value of EPBT based on energy and exergy for double slope is lower, respectively, by 7.5% and 17.98% than single-slope PVT-CPC active solar still. Again, the value of EPF

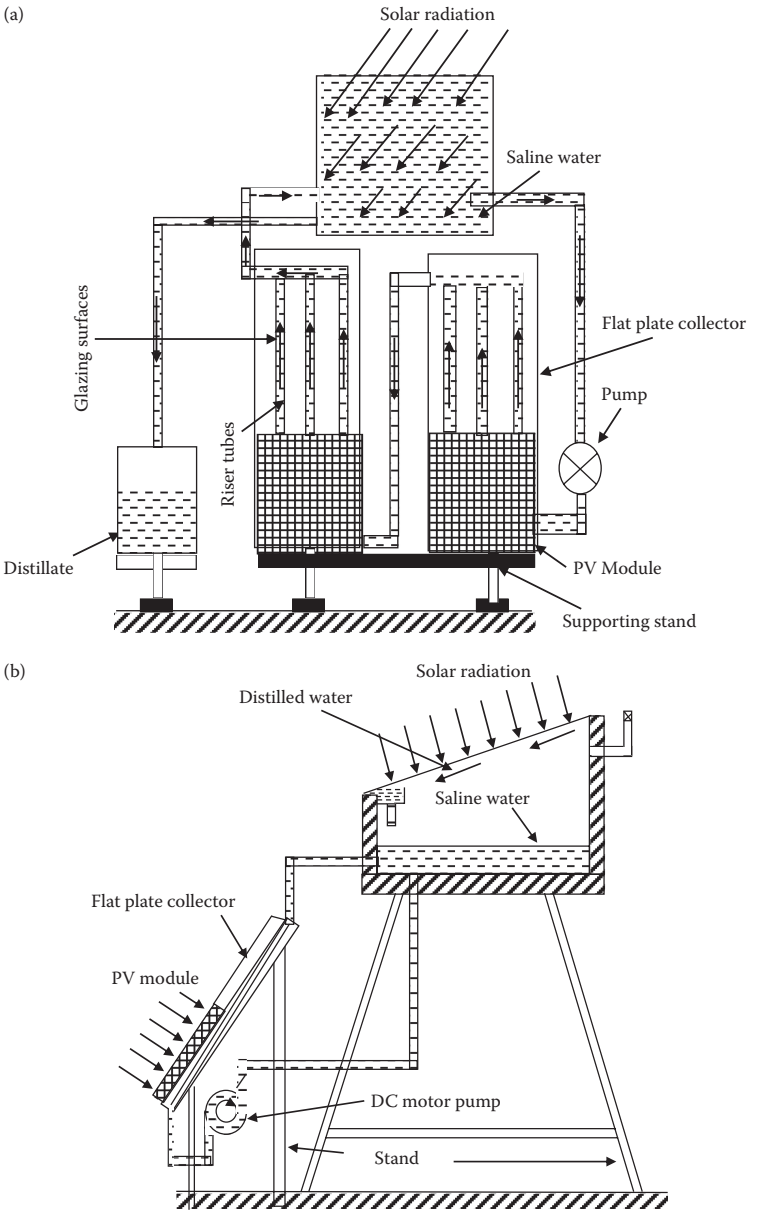


FIGURE 4.33.24 (a) Sectional top view of active photovoltaic thermal solar distillation system. (b) Sectional side view of active photovoltaic thermal solar distillation system.

based on energy and exergy for double slope is higher, respectively, by 5.12% and 12.72% than single-slope PVT-CPC active solar still. It happens because embodied energy of double slope is lower by 2.83%, but annual energy and exergy are, respectively, higher by 4.2% and 12.79% for double slope than single-slope PVT-CPC active solar still.

4. The LCCE of double slope based on energy and exergy is higher, respectively, by 5.55% and 22.22% than single-slope PVT-CPC active solar still. It occurs because overall annual energy and exergy are higher by 4.2% and 12.79%, respectively, for double slope than single-slope solar still. However, embodied energy and solar energy as well as exergy are lower by 2.83% and 1.56%, respectively, for double slope than single-slope PVT-CPC active solar still.

5. The production cost of water for single-slope PVT-CPC active solar still varies from Rs. 0.51 per kg to Rs. 2.42 per kg and for double-slope PVT-CPC active solar distillation system, the water cost varies between Rs. 0.47 and Rs. 2.21 per kg. The production cost of water for double slope is lower than single-slope PVT-CPC active solar still because yield of double slope at 0.14 m depth is higher and UAC is lower than single-slope PVT-CPC active solar still. The electricity is freely available from both single- and double-slope PVT-CPC active solar stills because revenue obtained from water is sufficient to meet the cost of system.

ETC-Integrated Solar Still

If FPC of active solar distillation system is replaced by ETC, the solar still is called ETC-integrated active solar still shown in Figure 4.33.25. ETC has two concentric tubes and there is vacuum in the annular space between tubes. The inner glass tubes are blacked at its outside surface with a selective coating to ensure the maximum absorption of solar radiation. Solar radiation incident on these tubes gets absorbed and the heat is transferred to the water inside the inner glass tube through its contact peripheral surface. ETC is inclined at some angle with horizontal (45° in this case) and the whole unit faces south to absorb maximum radiation. A pump is provided in between the basin and ETC to circulate water. Kumar et al. (2014) investigated the evacuated tube collector integrated solar still shown in Figure 4.33.25. The thermal model of the system was developed and performance of solar still under New Delhi (India) climatic condition was predicted using the thermal model. The performance of the system was optimum at mass flow rate of 0.06 kg/s and basin water depth of 0.03 m. The optimum daily yield was found to be 3.9 kg/m^2 with energy and exergy efficiencies as 33.8% and 2.6%, respectively, during typical summer day.

Climatic, Operational, and Design Parameter

The performance of solar still is affected by climatic, operational, and design parameters which are shown in Table 4.33.1.

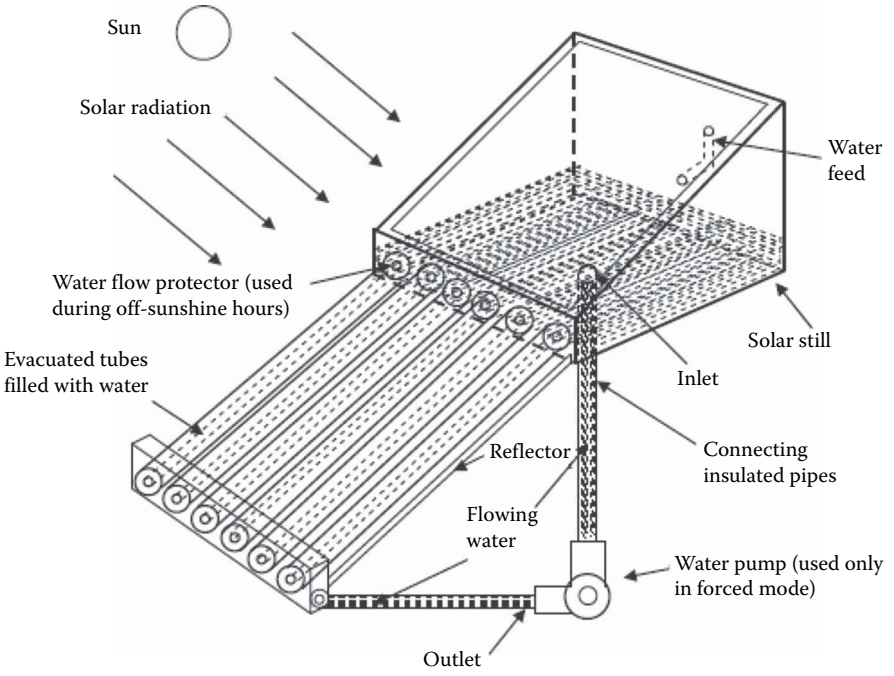


FIGURE 4.33.25 ETC-integrated solar still.

TABLE 4.33.1
Climatic, Operational, and Design Parameters

Type of parameter	Factors
Climatic parameter	Solar intensity, wind velocity, ambient temperature, humidity, and sky temperature
Operational parameter	Water depth in basin, salinity, amount of dye, and mode of operation
Design parameter	Material of construction, area of the basin, angle of inclination of condensing cover, and thickness of insulation thickness

APPENDIX

Expressions for different terms used in Equations 4.33.34 and 4.33.35 are as follows:

$$U_{tca} = \left[\frac{1}{h_o} + \frac{L_g}{K_g}\right]^{-1}; \quad U_{tcp} = \left[\frac{1}{h_i} + \frac{L_g}{K_g}\right]^{-1};$$

$$h_o = 5.7 + 3.8\,V, \ \text{W/m}^2\text{K}; \quad h_i = 5.7, \ \text{W/m}^2\text{K};$$

$$U_{tpa} = \left[\frac{1}{U_{tca}} + \frac{1}{U_{tcp}}\right]^{-1} + \left[\frac{1}{h_i'} + \frac{1}{h_{pf}} + \frac{L_i}{K_i}\right]^{-1}; \quad h_i' = 2.8 + 3\,V', \ \text{W/m}^2\text{K};$$

$$U_{L1} = \frac{U_{tcp}U_{tca}}{U_{tcp} + U_{tca}}; \quad U_{L2} = U_{L1} + U_{tpa}; \quad U_{Lm} = \frac{h_{pf}U_{L2}}{F'h_{pf} + U_{L2}}; \quad U_{Lc} = \frac{h_{pf}U_{tpa}}{F'h_{pf} + U_{tpa}};$$

$$PF_1 = \frac{U_{tcp}}{U_{tcp} + U_{tca}}; \quad PF_2 = \frac{h_{pf}}{F'h_{pf} + U_{L2}}; \quad PF_c = \frac{h_{pf}}{F'h_{pf} + U_{tpa}};$$

$$(\alpha\tau)_{1\text{eff}} = \rho(\alpha_c - \eta_c)\tau_g\beta_c\frac{A_{am}}{A_{rm}}; \quad (\alpha\tau)_{2\text{eff}} = \rho\alpha_p\tau_g^2(1-\beta_c)\frac{A_{am}}{A_{rm}};$$

$$(\alpha\tau)_{m\text{eff}} = \left[(\alpha\tau)_{1\text{eff}} + PF_1(\alpha\tau)_{2\text{eff}}\right]; \quad (\alpha\tau)_{c\text{eff}} = PF_c \cdot \rho\alpha_p\tau_g\frac{A_{ac}}{A_{rc}};$$

$$A_{rm} = b_r L_{rm}; \quad A_{am} = b_o L_{am}; \quad A_c F_{Rc} = \frac{\dot{m}_f c_f}{U_{Lc}}\left[1 - \exp\left(\frac{-F'U_{Lc}A_c}{\dot{m}_f c_f}\right)\right];$$

$$A_m F_{Rm} = \frac{\dot{m}_f c_f}{U_{Lm}}\left[1 - \exp\left(\frac{-F'U_{Lm}A_m}{\dot{m}_f c_f}\right)\right];$$

$$\left(AF_R(\alpha\tau)\right)_1 = \left[A_c F_{Rc}(\alpha\tau)_{c\text{eff}} + PF_2(\alpha\tau)_{m\text{eff}} A_m F_{Rm}\left(1 - \frac{A_c F_{Rc}U_{Lc}}{\dot{m}_f c_f}\right)\right];$$

$$(AF_R U_L)_1 = \left[A_c F_{Rc}U_{Lc} + A_m F_{Rm}U_{Lm} + A_m F_{Rm}U_{Lm}\left(1 - \frac{A_c F_{Rc}U_{Lc}}{\dot{m}_f c_f}\right)\right]$$

$$K_K = \left(1 - \frac{(AF_R U_L)_1}{\dot{m}_f c_f}\right); \quad \left(AF_R(\alpha\tau)\right)_{m1} = PF_2(\alpha\tau)_{m\text{eff}} A_m F_{Rm};$$

$$(AF_RU_L)_{m1} = A_mF_{Rm}U_{Lm};\quad K_m = \left(1-\frac{A_mF_{Rm}U_{Lm}}{\dot{m}_fc_f}\right)$$

Expressions for a_1 and $\overline{f_1}(t)$ used in Equation 4.33.41 and expressions of heat transfer coefficients used in Equations 4.33.43 through 4.33.45 are as follows:

$$a_1 = \frac{1}{M_wC_w}\left[\dot{m}_fC_f(1-K_k^N)+U_sA_b\right];$$

$$\overline{f_1}(t) = \frac{1}{M_wC_w}\left[\alpha'_{\text{eff}}A_b\overline{I_s}(t)+\frac{(1-K_k^N)}{(1-K_k)}\left(AF_R(\alpha\tau)\right)_1\overline{I_b}(t)+\left(\frac{(1-K_k^N)}{(1-K_k)}(AF_RU_L)_1+U_sA_b\right)\overline{T_a}\right];$$

$$\alpha'_{\text{eff}} = \alpha'_w+h_1\alpha'_b+h'_1\alpha'_g;\quad h_1 = \frac{h_{bw}}{h_{bw}+h_{ba}};$$

$$h'_1 = \frac{h_{1w}A_g}{U_{c,ga}A_g+h_{1w}A_b};\quad h_{1w} = h_{rwg}+h_{cwg}+h_{ewg};$$

$$h_{ewg} = 16.273\times10^{-3}h_{cwg}\left[\frac{P_w-P_{gi}}{T_w-T_{gi}}\right];$$

$$h_{cwg} = 0.884\left[(T_w-T_{gi})+\frac{(P_w-P_{gi})(T_w+273)}{268.9\times10^3-P_w}\right]^{\frac{1}{3}};$$

$$P_w = \exp\left[25.317-\frac{5144}{T_w+273}\right];\quad P_{gi} = \exp\left[25.317-\frac{5144}{T_{gi}+273}\right];$$

$$h_{rwg} = \left(0.82\times5.67\times10^{-8}\right)\left[(T_w+273)^2+(T_{gi}+273)^2\right]\left[T_w+T_{gi}+546\right];$$

$$U_s = U_t+U_b;\quad U_b = \frac{h_{ba}h_{bw}}{h_{bw}+h_{ba}};\quad U_t = \frac{h_{1w}U_{c,ga}A_g}{U_{c,ga}A_g+h_{1w}A_b};$$

$$U_{c,ga} = \frac{\frac{K_g}{l_g}h_{1g}}{\frac{K_g}{l_g}+h_{1g}};\quad h_{ba} = \left[\frac{L_i}{K_i}+\frac{1}{h_{cb}+h_{rb}}\right]^{-1};$$

$$h_{cb}+h_{rb} = 5.7\ \text{W/m}^2\text{K},\ h_{bw} = 250\ \text{W/m}^2K;$$

Expressions for a_2 and $\overline{f_2}(t)$ used in Equation 4.33.52 and expressions of different terms used in Equations 4.33.53 through 4.33.58 are as follows:

$$a_2 = \frac{1}{M_wC_w}\left[\dot{m}_fC_f(1-K_k^N)+U_bA_b+\frac{h_{1wE}(P-A_2)A_b}{2P}+\frac{h_{1wW}(P-B_2)A_b}{2P}\right];$$

$$\overline{f}_2(t)=\frac{1}{M_wC_w}\left[\begin{array}{l}\left(\frac{\alpha'_w}{2}+h_1\alpha'_b\right)A_b\left(\overline{I}_{SE}(t)+\overline{I}_{SW}(t)\right)+\frac{(1-K_k^N)}{(1-K_k)}\left(AF_R(\alpha\tau)\right)_1\overline{I}_b(t)\\+\left(\frac{(1-K_k^N)}{(1-K_k)}\left(AF_RU_L\right)_1+U_bA_b\right)\overline{T}_a+\left(\frac{h_{1wE}A_1+h_{1wW}B_1}{P}\right)\frac{A_b}{2}\end{array}\right]$$

$$A_1=R_1U_1A_{gE}+R_2h_{EW}A_{gW};\quad A_2=h_{1WE}U_2\frac{A_b}{2}+h_{EW}h_{1WW}\frac{A_b}{2};$$

$$P=\left(U_1U_2-\frac{h_{EW}^2}{A_{gE}}h_{1WW}\frac{A_b}{2}\right)A_{gW};\quad U_1=\frac{h_{1wE}\frac{A_b}{2}+h_{EW}A_{gE}+U_{c,gaE}A_{gE}}{A_{gW}};$$

$$U_2=\frac{h_{1wW}\frac{A_b}{2}+h_{EW}A_{gW}+U_{c,gaW}A_{gW}}{A_{gE}};\quad B_1=\frac{(R_2P+A_1h_{EW})A_{gW}}{U_2A_{gE}};$$

$$B_2=\frac{Ph_{1wW}\frac{A_b}{2}+h_{EW}A_{gW}A_2}{U_2A_{gE}};\quad R_1=\alpha'_gI_{SE}(t)+U_{c,gaE}T_a;$$

$$R_2=\alpha'_gI_{SW}(t)+U_{c,gaW}T_a$$

$$h_{EW}=0.034\times5.67\times10^{-8}\left[(T_{giE}+273)^2+(T_{giW}+273)^2\right]\left[T_{giE}+T_{giW}+546\right]$$

$$U_{c,gaE}=\frac{\frac{K_g}{l_g}h_{1gE}}{\frac{K_g}{l_g}+h_{1gE}};\quad U_{c,gaW}=\frac{\frac{K_g}{l_g}h_{1gW}}{\frac{K_g}{l_g}+h_{1gW}};\quad h_{1gE}=5.7+3.8V;\quad h_{1gW}=5.7+3.8\,V;$$

$$h_{1wE}=h_{rwgE}+h_{cwgE}+h_{ewgE};\quad h_{1wW}=h_{rwgW}+h_{cwgW}+h_{ewgW};$$

$$h_{ewgE}=16.273\times10^{-3}h_{c,wgE}\left[\frac{P_w-P_{giE}}{T_w-T_{giE}}\right];\quad h_{cwgE}=0.884\left[(T_w-T_{giE})+\frac{(P_w-P_{giE})(T_w+273)}{268.9\times10^3-P_w}\right]^{\frac{1}{3}};$$

$$h_{ewgW}=16.273\times10^{-3}h_{c,wgW}\left[\frac{P_w-P_{giW}}{T_w-T_{giW}}\right];\quad h_{cwgW}=0.884\left[(T_w-T_{giW})+\frac{(P_w-P_{giW})(T_w+273)}{268.9\times10^3-P_w}\right]^{\frac{1}{3}};$$

$$P_w=\exp\left[25.317-\frac{5144}{T_w+273}\right];\quad P_{giE}=\exp\left[25.317-\frac{5144}{T_{gi}E+273}\right];$$

$$P_{giW}=\exp\left[25.317-\frac{5144}{T_{giW}+273}\right];$$

$$h_{rwgE} = (0.82 \times 5.67 \times 10^{-8})\left[(T_w + 273)^2 + (T_{giE} + 273)^2\right][T_w + T_{giE} + 546];$$

$$h_{rwgW} = (0.82 \times 5.67 \times 10^{-8})\left[(T_w + 273)^2 + (T_{giW} + 273)^2\right][T_w + T_{giW} + 546];$$

Nomenclature

A_m	Area of module, m^2
A_c	Area of collector, m^2
Ag	Area of glass cover, m^2
A_b	Area of basin liner, m^2
C	Specific heat capacity, J/kg-K
F'	Flat plate collector efficiency factor, dimensionless
F_R	Heat removal factor, dimensionless
H	Heat transfer coefficient, W/m^2-K
PF_1	Penalty factor first, dimensionless
PF_2	Penalty factor second, dimensionless
$I(t)$	Incident solar intensity, W/m^2
$I_c(t)$	Incident solar intensity on collector, W/m^2
$I_s(t)$	Incident solar intensity on solar still, W/m^2
$I_b(t)$	Bean radiation incident on collector, W/m^2
I	Rate of interest, %
K	Thermal conductivity, W/m-K
L	Latent heat, J/kg
L_g	Thickness of glass, m
$\dot{m}_f$	Rate of flow of water mass in collector, kg/s
$\dot{Q}_{UN}$	Rate of useful energy transfer, kW
T	Temperature, °C
$U_{tc,a}$	Total heat transfer coefficient from solar cell to ambient through glass cover, W/m^2-K
$U_{tc,p}$	Total heat transfer coefficient from solar cell to plate through glass cover, W/m^2-K
$U_{L,m}$	An overall heat transfer coefficient from blackened surface to ambient, W/m^2-K
V_a	Velocity of air, m/s
V_{oc}	Open-circuit voltage, V
V_L	Load voltage, V
I_{sc}	Short-circuit current, A
I_L	Load current, A
h_{1w}	Total hear transfer coefficient from water surface to inner glass cover, W/m^2-K
h_{1g}	Total hear transfer coefficient from glass cover to ambient, W/m^2-K
h_{bw}	Heat transfer coefficient from blackened surface to water mass, W/m^2-K
h_{ba}	Heat transfer coefficient from blackened surface to ambient, W/m^2-K
h_r	Radiative heat transfer coefficient, W/m^2-K
h_c	Convective heat transfer coefficient, W/m^2-K
h_{ewg}	Evaporative heat transfer coefficient from water to glass, W/m^2-K
$\dot{m}_{ew}$	Hourly distillate output, kg/h
R	Reflectivity, dimensionless
FF	Fill factor, dimensionless
R_e	Annual revenue earned from electricity gain, Rs.
R_w	Annual revenue earned from distillate, Rs.
SP	Selling price, Rs.
M_w	Annual yield, kg
$\dot{E}_e$	Hourly electricity gain, kWh

(*Continued*)

E_e	Annual electricity gain, kWh
UAC	Uniform end-of-year annual cost, Rs.
$F_{CR,i,n}$	Capital recovery factor, dimensionless
$F_{SR,i,n}$	Sinking fund factor, dimensionless
P_s	Net present cost of system, Rs.
P	Initial cost of system, Rs.
C_w	Production cost of water, Rs.
C_e	Cost of electricity gain, Rs.
E_{in}	Embodied energy
E_{out}	Annual energy output
LCCE	Life cycle conversion efficiency
EPBT	Energy payback time
EPF	Energy production factor
$(G_{ex,annual})$	Annual exergy gain
R_{ex}	Exergoeconomic parameter based on exergy loss
R_{en}	Exergoeconomic parameter based on energy loss
$R_{g,ex}$	Exergoeconomic parameter based on exergy gain
$R_{g,en}$	Exergoeconomic parameter based on energy output
L_{ex}	Hourly exergy loss rate
N	Life of system
Subscripts	
A	Ambient
C	Solar cell
Eff	Effective
W	Water in solar still
F	Water in flat plate collector
G	Glass
Gi	Inner glass cover
Go	Outer glass cover
M	Module
N	Number of collectors
P	Plate
B	Blackened surface
R	Radiative
Greek Letters	
α	Absorptivity
$(\alpha\tau)_{eff}$	Product of effective absorptivity and transmittivity
β	Packing factor
β_c	Electrical efficiency
η	Efficiency
τ	Transmittivity

REFERENCES

A/K Abu-Hijlew B. and Hasan A.M., (1997), Water film cooling over the glass cover of a solar still including evaporation effects, *Energy*, 22(1), 43.

Abbot C.G., (1938), Solar distillation apparatus, US patent 2,141,330, December 27, 1938.

Abdel-Rehim Z.S. and Lasheen A. (2007), Experimental and theoretical study of a solar desalination system located in Cairo, Egypt, *Desalination*, 217, 52–64.

Abu-Arabi M., Yousef Z., Hilal A.-H. and Saif A.-H., (2002), Modeling and performance analysis of a solar desalination unit with double-glass cover cooling, *Desalination*, 143, 173–182.

Abu-Hijleh B.A.K., (1996), Enhanced solar still performance using water film cooling of the glass cover, *Desalination*, 107, 235–244.

Akhtamov R.A., (1978), Study of regenerative inclined stepped solar still, *Geliotekhnika*, 14(4), 51–55.
Al Mahdi N., (1992), Performance prediction of a multi-basin solar still, *Energy*, 17(1), 87–93.
Al-Hinai H., Al-Nassri M.S. and Jubran B.A., (2002), Parametric investigation of a double-effect solar still in comparison with a single-effect solar still, *Desalination*, 150, 75–83.
Al-Karaghoulia A.A. and Alnaserb W.E., (2004), Experimental comparative study of the performances of single and double basin solar-stills, *Applied Energy*, 77, 317–325.
Al-Sebaii A.A., (2005), Thermal performance of triple-basin solar still, *Desalination*, 174(1), 23–37.
Arslan M., (2012), Experimental investigation of still performance for different active solar still designs under closed cycle mode, *Desalination*, 307, 9–19.
Ashcroft H., (1950), The productivity of several machines under the care of one operator, *Journal of the Royal Statistical Society: Series B*, XII, 145–151.
Badran O.O., Al-Tahaineh H.A., (2005), The effect of coupling a flat plate collector on the solar still productivity, *Desalination*, 183, 137–142.
Badran A.A., Al-Hallaq A.A., Eyal Salman I.A. and Odat M.Z., (2005), A solar still augmented with a flat-plate collector. *Desalination*, 172, 227–234.
Bapeshwar V. and Tiwari G.N., (1984), Effect of water flow over the glass on the performance of solar still coupled with a flat plate collector, *International Journal of Solar Energy*, 2, 277–288.
Bapeshwararao V.S.V., Singh U. and Tiwari G.N., (1983), Transient analysis of double basin solar still. *Energy Conversion and Management*, 23(2) 83–90.
Baum V.A., Bayaramov R.B. and Malevsky Y.M., (1970), The solar still in deserts. Proceedings of International Solar Energy Congress, Melbourne.
Benson F., (1952), Further notes on the productivity of machines requiring attention at random intervals, *Journal of the Royal Statistical Society: Series B*, XIV, 200–210.
Bhargava A.K., Garg H.P. and Agarwal R.K., (1991), Study of hybrid solar system- solar air heater combined with solar cell. *Energy convers Mgmt*, 31(5), 471–479.
Calise F., d'Accadia M.D. and Piacentino A., (2014), A novel solar trigeneration system integrating PVT (photovoltaic/thermal collectors) and SW (seawater) desalination: Dynamic simulation and economic assessment, *Energy*, 67, 129–148.
Caliskan H., Dincer I. and Hepbasli A., (2012), Exergoeconomic, enviroeconomic and sustainability analyses of a novel air cooler, *Energy and Buildings*, 55, 747–756.
Cappelletti G.M., (2002), An experiment with a plastic solar still, *Desalination*, 142, 221–228.
Choi U.S., (1995), Enhancing Thermal Conductivity of Fluids with Nanoparticles, *Developments and Applications of Non-Newtonian Flows*, D.A. Siginer and H.P. Wang, eds., FED-Vol. 231/MD-Vol. 66, ASME, New York, 99–105.
Clark J.A., (1990), The steady state performance of a solar still, *Solar Energy*, 44, 43–49.
Cooper P.I., (1969a), Digital simulation of transient solar still processes, *Solar Energy*, 12(3), 313–331.
Cooper P.I., (1969b), The absorption of solar energy radiations in solar stills, *Solar Energy*, 12(3), 333.
Cooper P.I., (1973a), Digital simulation of experimental solar still data, *Solar Energy*, 14, 451.
Cooper P. I., (1973b), Maximum efficiency of a single effect solar stills, *Solar energy*, 15, 205–214.
Cooper P.I. and Appleyard J.A., (1967), The construction and performance of a three effect wick type tilted solar still, *Sun Work*, 12(1), 4–8.
Cox D.R., (1951), The productivity of machines requiring attention at random intervals. *Journal of the Royal Statistical Society: Series B*, XIII, 65–82.
Della Porta, (1589), Magiae naturalis libri XX, Napoli.
Delyannis A. and Delyannis E., (1973), Solar distillation plant of high capacity. In *Proceedings of 4th International Symposium on Fresh Water from Sea*, European Federation of Chemical Engineering, Athens, 4, 487.
Delyannis A. and Delyannis E., (1983), Recent solar distillation developments, *Desalination*, 45(2), 361–369.
Den Elzen M.G.J., Hof A.F., Mendoza Beltran A. et al., (2011), The Copenhagen Accord: Abatement costs and carbon prices resulting from the submissions, *Environmental Science & Policy*, 14(1) 28–39.
Dunkle, R.V. (1961), Solar water distilation: The roof type still and a multiple effect diffusion still. *Proceedings of International Heat Transfer Conference*, University of Colorado, Boulder, Colorado, Part V, 895–902.
Dutt D.K., Rai S.N. and Tiwari G.N., (1988), Thermal modeling of high temperature distillation, *Energy Conversion Management*, 28(2), 151–159.
Dutt D.K., Kumar A., Anand J.D. and Tiwari G.N., (1989), Performance of a double-basin solar still in the presence of dye. *Applied Energy*, 32, 207–223.

Dwivedi V.K. and Tiwari G.N., (2008), Annual energy and exergy analysis of single and double slope passive solar stills, *Applied Science Research*, 3(3), 225–241.
Dwivedi V.K. and Tiwari G.N., (2009), Experimental validation of thermal model of a double slope active solar still under natural circulation mode, *Desalination*, 250(2010), 49–55, Elsevier.
Dwivedi V.K. and Tiwari G.N., (2010), Experimental validation of thermal model of double slope active solar still under natural circulation mode, *Desalination*, 250(1), 49–55.
Elango T., Kannan A. and Kalidasa Murugavel K., (2015), Performance study on single basin single slope solar still with different water nanofluids, *Desalination*, 360, 45–51.
El-Bialy E., (2014), Performance analysis for passive single slope single basin solar distiller with a floating absorber: An experimental study, *Energy*, 68 117–124.
El-Sebaii A.A., (2005), Thermal performance of a triple-basin solar still, *Desalination*, 174(1), 23–37.
El-Sebaii A.A., Aboul-Enin S. and El-Bialy E., (2000), Single basin solar still with baffle suspended absorber, *Energy Conversion Management*, 41(7), 661–675.
El-Sebaii A.A., Yaghmour S.J, Al-Hazmi F.S., Faidah A.S., Al-Marzouki F.M. and Al-Ghamdi A.A., (2009), Active single basin solar still with a sensible storage medium, *Desalination*, 249, 699–706.
Eltawil M.A., Omara Z.M., (2014), Enhancing the solar still performance using solar photovoltaic, flat plate collector and hot air. *Desalination*, 349, 1–9.
Estahbanati M.R.K., Feilizadeh M., Jafarpur K., Feilizadeh M. and Rahimpour, M.R., (2015), Experimental investigation of a multi-effect active solar still: The effect of the number of stages, *Applied Energy*, 137, 46–55.
Evans D.L., (1981), Simplified method for predicting PV array output, *Solar Energy*, 27, 555–560.
Fath H.E.S., (1996), High performance of a simple design, two effect solar distillation unit, *Desalination*, 107, 223–233.
Fernandez, J. and Chargoy, N., (1990), Multistage indirectly heated solar still, *Solar Energy*, 44(4), 215.
Frick B., (1970), Some new considerations about solar stills. In *Proceedings of International Solar Energy Congress*, Melbourne. 395.
Frick G. and von Sommerfeld J., (1973), Solar still of inclined evaporative cloth, *Solar Energy*, 14(4), 427–431.
Fukui K., Takehiro N., Hiroshi T. and Takashi N., (2004), A new maritime lifesaving multiple-effect solar still design, *Desalination*, 160 271–283.
Hamadou O.A. and Abdellatif K., (2014), Modeling an active solar still for sea water desalination process optimization, *Desalination*, 354, 1–8, Elsevier.
Hamdan M.A., Musa A.M. and Jubran B.A., (1999), Performance of solar still under Jordanian climate, *Energy Conversion and Management*, 40, 495–503.
Harding J., (1883), Apparatus for solar distillation, *Proceedings of the Institution of Civil Engineers*, 73, 284–288.
Hendrie S.D., (1979), Evaluation of combined photovoltaic/thermal collectors, In: *Proceedings of international conference ISES*, Atlanta, GA, 3, 1865–1869.
Hirschman J.R. and Roefler S.K., (1970), Thermal inertia of solar still and its influence on performance. In *Proceeding of International Solar Energy Congress*, Melbourne, 402.
Howe E.D., (1961), Solar distillation research at the university of California, UN conference on New Sources of Energy, Rome, 1
Huang B.J., Lin T.H., Hung W.C. and Sun F.S., (2001), Performance evaluation of solar photovoltaic/thermal systems, *Solar Energy*, 70(5), 443–448.
International Labor Office, (1979), *Introduction to Work Study*. International Labor Organization, Geneva.
Jafarkazemi, F. and Ahmadifard, E., (2012), Energetic and exergetic evaluation of flat plate solar collectors. *Renewable Energy*, 56, 55–63.
Kausch O., (1920), Die unmittelbare ausnutzung de sonnenergia, Weimar, Cael Steinert.
Kudish A.I., Gale J. and Zarmi Y., (1982), A low cost design solar desalination unit, Energy Conversion and Management, 22(3), 269–274.
Kudish A.I., Evseeva E.G., Walter G. and Priebe T., (2003), Simulation study on a solar desalination system utilizing an evaporator/condenser chamber, *Energy Conversion and Management*, 44(10), 1653–1670.
Kumar A. and Anand J.D., (1992), Modeling and performance of a tubular multi-wick solar still, *Energy*, 17(11), 1067–1071.
Kumar S. and Sinha S. (1996), Transient model and comparative study of concentrator coupled regenerative solar still in forced circulation mode, *Energy Conversion and Management*, 37(5), 629–36.
Kumar S. and Tiwari G.N., (1996a), Performance evaluation of an active solar distillation system, *Energy*, 21(9),805–808.

Kumar S. and Tiwari G.N., (1996b), Estimation of convective mass transfer in solar distillation systems, *Solar Energy*, 57(6), 459–464.
Kumar S. and Tiwari G.N., (1998), Optimization of collector and basin areas for a higher yield for active solar stills, *Desalination*, 116, 1–9.
Kumar, S. and Tiwari, G.N., (2009), Life cycle cost analysis of single slope hybrid (PV/T) active solar still, *Applied Energy*, 86, 1995–2004.
Kumar, S. and Tiwari, A., (2010), Design, fabrication and performance of a hybrid photovoltaic/thermal (PV/T) active solar still, *Energy Conversion and Management*, 51, 1219–1229.
Kumar A., Anand J.D. and Tiwari G.N., (1991), Transient analysis of a double slope-double basin solar distiller, *Energy Conversion and Management*, 31(2), 129–139.
Kumar S., Dubey A., Tiwari G.N., (2014), A solar still augmented with an evacuated tube collector in forced mode. *Desalination*, 347, 15–24.
Kumar S., Tiwari G.N. and Gaur M.K., (2010), Development of empirical relation to evaluate the heat transfer coefficients and fractional energy in basin type hybrid (PVT) active solar still. *Desalination*, 250, 214–221.
Kumar S., Tiwari G.N. and Singh H.N., (2000), Annual performance of an active solar distillation system, *Desalination*, 127(1), 79–88.
Lavoisier A.L., (1862), Oeuvres de Lavoisier, Son Excellence le ministre de I'instruction publique et de cultes, 3, Takle 9.
Lawrence S.A. and Tiwari G.N., (1990), Theoretical evaluation of solar distillation under natural circulation with heat exchanger, *Energy Conversion and Management*, 30,205.
Lilian M., George M.A. and Al-Hindi M., (2014), The effect of cover geometry on the productivity of a modified solar still desalination unit', *Energy Procedia*, 50 406–413.
Lobo P.C. and Araujo S.R.D., (1977), Design a simple multi-effect basin type solar still. In *Proceedings of International Solar Energy Congress*, New Delhi, 2026.
Lourdes G.R., Palmero-Marrero A.I, and Carlos G.C., (2002), Comparison of solar thermal technologies for applications in seawater desalination, *Desalination*, 142, 135–42.
Madhlopa A. and Johnstone C., (2009), Numerical study of a passive solar still with separate condenser, *Renewable Energy*, 34 1668–1677.
Madhuri and Tiwari G.N., (1985), Performance of solar still with intermittent flow of waste hot water in the basin, *Desalination*, 52, 345.
Malik, M.A.S. and Tran, V.V. (1973). A simplified mathematical model for predicting the nocturnal output of a solar still, *Solar Energy*, 14, 371.
Malik, M.A.S., Tiwari, G.N., Kumar, A. and Sodha, M.S., (1982), *Solar Distillation: A Practical Study of a Wide Range of Stills and Optimum Design, Construction and Performance*, Pergamon Press, New York.
Menguy G. et al., (1980), New solar still (the wiping spherical still): Design and experimentation (Malik et al., 1982).
Monowe P., Masale M., Nijegorodov N. and Vasilenko V., (2011), A portable single-basin solar still with an external reflecting booster and an outside condenser, *Desalination*, 280, 332–338.
Mouchot A., (1869), *La Chaleur Solavie et ses Applications*. Gauthier-Villars, Paris.
Moustafa S.M.A., Brusewitz G.H. and Farmer D.M., (1979), Direct use of solar energy for water desalination, *Solar Energy*, 22(2), 141–148.
Nag P.K., (2004), *Basic and applied thermodynamics*, Tata McGraw-Hill, ISBN 0-07-047338-2.
Nayak J.K., Tiwari G.N. and Sodha M.S., (1980), Periodic theory of solar still, *International Journal of Energy Research*, 4(1), 41–57.
Nebbia G. and Mennozi G., (1966), A short history of water desalination. Acque Dolce Dal Mare, II Inchiesta Internazionole. In *Proceedings of Interantional Symposium*, Milano, 129.
Norov E.Z.H. et al., (1975), Results of tests on solar film covered still, *Geliotekhnika*, 11(3/4), 130–132.
Oltra F., (1972), *Saline Water Conversion and Its Stage of Development in Spain*, Publication of J.E.M., Madrid.
Pasteur F., (1928), *Compt Rend*, 30, 187
Petela R., (2003), Exergy of undiluted thermal radiation, *Solar Energy*, 86, 241–247.
Prakash J. and Kavathekar K., (1986), Performance prediction of a regenerative solar still. *Solar and Wind Technology*, 3(2), 119–125.
Rahim N.H.A., (2003), New method to store heat energy in horizontal solar desalination still, *Renewable Energy*, 28(3), 419–433.

Rai S.N. and Tiwari G.N., (1982), Single basin solar still coupled with flat plate collector, *Energy Conversion and Management*, 23(3), 145–149.

Rai S.N. and Tiwari G.N., (1983), Single basin solar still coupled with flat plate collector, *Energy Conversion and Management*, 23(3), 145–149.

Rajamanickam, M.R. and Ragupathy A., (2012), Influence of water depth on internal heat and mass transfer in a double slope solar still, *Energy Procedia*, 14, 1701–1708.

Rajaseenivasan T., Ilango T. and Murugavel K.K., (2013), Comparative study of double basin and single basin solar still, *Desalination*, 309, 27–31.

Rajaseenivasan T., Nelson P. and Srithar K., (2014), An experimental investigation on a solar still with an integrated flat plate collector, *Desalination*, 347, 131–137, Elsevier.

Rubio, E., Fernandez J.L. and Porta-Gandara M.A., (2004), Modeling thermal asymmetries in double slope solar still, *Renewable Energy*, 29,895.

Sahota L. and Tiwari G.N., (2016a), Effect of Al_2O_3 nano-particles on the performance of passive double slope solar still, *Solar Energy*, 130, 260–272.

Sahota L. and Tiwari G.N., (2016b), Effect of nanofluids on the performance of passive double slope solar still: A comparative study using characteristic curve, *Desalination*, 388 9–21.

Sangeeta S. and Tiwari G.N., (1998), Optimization of number of effects for higher yield from an inverted absorber solar still using the Runge–Kutta method, *Desalination*, 120, 197–209.

Sangeeta S. and Tiwari G.N., (1999a), Effect of water depth on the performance of an inverted absorber double basin solar still, *Energy Conversion and Management*, 40, 1885–1897.

Sangeeta S. and Tiwari G.N., (1999b), Parametric studies of an inverted absorber triple effect solar still, *Energy Conversion and Management*, 40, 1871–1884.

Sangeeta S., Tiwari G.N. and Rai S.N., (1997), Parametric study of an inverted absorber double-effect solar distillation system, *Desalination*, 109, 177–186.

Schott T., (1985), Operational temperatures of PV modules, In: *Proceedings of 6th PV Solar Energy Conference*, 392–396.

Shyam., Tiwari G.N. and Al-Helal I.M., (2015), Analytical expression of temperature dependent electrical efficiency of N-PVT water collectors connected in series, *Solar Energy*, 114, 61–76.

Singh A.K. and Tiwari G.N., (1992), Performance study of double effect distillation in a multiwick solar still, *Energy Conversion and Management*, 33(3) 207–214.

Singh D.B. and Tiwari G.N, (2016), Effect of energy matrices on life cycle cost analysis of partially covered photovoltaic compound parabolic concentrator collector active solardistillation system, *Desalination*, 397, 75–91.

Singh G., Kumar S. and Tiwari G.N., (2011), Design, fabrication and performance of a hybrid photovoltaic/thermal (PVT) double slope active solar still, *Desalination*, 277, 399–406.

Singh D.B., Yadav J.K., Dwivedi V.K., Kumar S., Tiwari G.N. and Al-Helal I.M., (2016), Experimental studies of active solar still integrated with two hybrid PVT collectors, *Solar Energy*, 130, 207–223.

Sinha S., Kumar S. and Tiwari G.N., (1994), active solar distillation system –an investment alternative to solar hot water system. *Energy Convers. Mgmt*, 35(7), 583–588.

Sodha M.S., and Tiwari G.N., (1979), Low cost portable solar still, *Sunworld*, 3, 172.

Sodha M.S., Kumar A., Singh U. and Tiwari G.N., (1980), Energy conversion of solar still. *Energy Conservation*, 200(3), 191.

Sodha M.S., Kumar A., Tiwari G.N. and Tyagi R.C., (1981), Simple multiple wick solar still: Analysis and performance. *Journal of Solar Energy*, 26(2), 127.

Soliman H.S., (1976), *Solar Still Coupled with a Solar Water Heater.* Mosul University: Mosul, 43, Iraq.

Sovacool B.K., (2008), Valuing the greenhouse gas emissions from nuclear power: A critical survey, *Energy Policy*, 36, 2940–2953.

Srivastava P.K. and Agrawal S.K., (2013), Winter and summer performance of single sloped basin type solar still integrated with extended porous fins. *Desalination*, 319, 73–78.

Taghvaei H., Taghvaei H., Khosrow J., Karimi Estahbanati M.R., Mehrzad F., Mansoor F. and SeddighArdekani A., (2014), A thorough investigation of the effects of water depth on the performance of active solar stills, *Desalination*, 347, 77–85.

Talbert S.G., Eibling J.A., and Lof G.O.G, (1970). Manual on solar distillation of saline water. R & D Progress Report No. 546, U.S. Department of Interior.

Tanaka H., Nosoko T. and Nagata T., (2000), A highly productive basin type multiple effect coupled solar still, *Desalination*, 130, 279.

Tanaka H., Nosoko T. and Nagata T., (2002), Experimental study of basin type, multi-effect, diffusion coupled solar still, *Desalination*, 150(2), 131–144.

Telkes M., (1945), Solar distillation for life raft, Unites States Office of Science, R & D, Report No. 5525, P.B. 21120.
Tiwari G.N., (1992), Recent advances in solar distillation. Chapter II, *Contemporary Physics: Solar Energy and Energy Conservation,* Wiley Eastern Ltd., New Delhi, India.
Tiwari G.N., (2013), *Solar Energy: Fundamentals, Design, Modelling and Applications*, CRC Publication/ Narosa Publishing House, New Delhi/New York.
Tiwari G.N. and Dhiman N.K., (1991), Performance study of a high temperature distillation system, *Energy Conversion and Management*, 32(3), 283–291.
Tiwari G.N., Dimri V., Singh U., Chel A. and Sarkar B., (2007), Comparative thermal performance evaluation of an active solar distillation system, *International Journal of Energy Research*, 31, 1465–1482.
Tiwari G.N., Dimri V. and Chel A., (2009), Parametric study of an active and passive solar distillation system: energy and exergy analysis, *Desalination*, 242, 1–18.
Tiwari G.N. and Ghosal M.K., (2005), *Renewable Energy Resources: Basic Principles and Applications*. Narosa Publishing House, New Delhi.
Tiwari G.N. and Mishra R.K., (2012), *Advanced Renewable Energy Sources*, Royal Society of Chemistry Publishing House, UK.
Tiwari G.N., Mishra R.K. and Solanki S.C., (2011), Photovoltaic modules and their applications: A review on thermal modeling, *Applied Energy*, 88, 2287–2304.
Tiwari G.N. and Salem G.A.M., (1984), Double slope ERP multi-wick solar still, *Solar Wind Technology*, 1, 229–235.
Tiwari G.N. and Suneja S., (1998), Performance evaluation of an inverted absorber solar distillation system, *Energy Conversion and Management*, 39(3), 173–180.
Tiwari G.N., Chetna S. and Yadav Y.P., (1991), Effect of water depth on the transient performance of a double basin solar still, *Energy Conversion and Management*, 32(3), 293–301.
Tiwari G.N., Gupta S.P. and Lawrence S.A., (1989), Transient analysis of solar still in the presence of dye, *Energy Conversion and Management*, 29(1), 59–62.
Tiwari G.N., Mukherjee, K. Yadav, Y.P., (1986), Comparisons of various designs of solar stills, *Desalination*, 60, 191–202.
Tiwari G.N., Singh H.N. and Tripathi R., (2003), Present status of solar distillation, *Solar Energy,* 75, 367–373.
Tiwari G.N., Yadav J.K., Singh D.B., Al-Helal, I.M. and Abdel-Ghany, A. M., (2015), Exergoeconomic and enviroeconomic analyses of partially covered photovoltaic flat plate collector active solar distillation system, *Desalination*, 367, 186–196.
Tleimat B.W. and Howe E.D., (1966), Nocturnal production of solar distiller, *Solar Energy*, 10, 61.
Toyama S. and Kagakuv K., (1972), Gijitsu, 24, 159, Maruzen, Tokyo.
Toyama S., Aragaki T. and Murase K., (1983), Simulation of a multi effect solar distillator. *Desalination*, 45, 101–108.
Tripathi R. and Tiwari G.N., (2005), Effect of water depth on internal heat and mass transfer for active solar distillation. *Desalination*, 173, 187–200.
Tripathi R., Tiwari G.N. and Al-Helal I.M, (2016), Thermal modeling of N partially covered photovoltaic thermal (PVT) – Compound parabolic concentrator (CPC) collectors connected in series, *Solar Energy*, 123, 174–184.
Tiris C., Tiris M., Erdalli Y., and Sohmen M., (1998), Experimental studies on a solar still coupled with a flat plate collector and a single basin still, *Energy Conversion and Management*, 39(8), 853–6.
Tsatsaronis G. and Winhold M., (1985), Exergoeconomic analysis and evaluation of energy conversion plants: A new general methodology, *Energy*, 10, 69–94.
Tsatsaronis G., Lin L. and Pisa J., (1993), *Exergy costing in Exergoeconomics*, *Journal of Energy Resource: ASME*, 115, 9–16.
UN., (2004), Millennium Project, www.unmillenniumproject.org/who/.
UNICEF/WHO., (2004), Joint Monitoring Programme for Water Supply and Sanitation; Meeting the MDG drinking water and sanitation target: a mid-term assessment of progress.
Watmuff J.H., Characters W.W.S. and Proctor D., (1977), Solar and wind induced external coefficient for solar collectors, Cooperation Mediterraneenne pour I'Energie Solaire, Revue Internationale d'Heliotechnique, 2nd Quarter, 56.
Yadav Y.P. and Tiwari G.N. (1987), Monthly comparative performance of solar stills of various designs, *Desalination*, 67, 565–578.
Yeh H., (1993), Experimental studies on upward-type double effect solar distillers with air flow through the second effect, *Energy*, 18(11), 1107–1111.

Yeh H. and Chen Z., (1992), Experimental studies on wick-type, double effect solar distillers with air flow through the second-effect unit, *Energy*, 17(3), 269–273.

Zaki G.M., Al-Turki A. and AI-Fatani M., (1992), Experimental investigation on concentrator assisted solar stills, *Solar Energy*, 11, 193.

Zaki G.M., Dali T. El. and Shafie H. El., (1983), Improved performance of solar still, *Proc. First Arab Int. Solar Energy Conf.*, Kuwait, 331–335.

Zeroual M., Bouguettaiab H., Bechkib D., Boughalib S., Bouchekimab B. and Mahceneb H., (2011), Experimental investigation on a double-slope solar still with partially cooled condenser in the region of Ouargla (Algeria), *Energy Procedia*, 6, 736–742.

Zuriat Y.H. and Abu-arabi M.K., (2004), Modeling and performance analysis of a regenerative solar desalination unit, *Applied Thermal Energy*, 24, 1061–1072.

5 Numerical Methods and Computational Tools

Atul Sharma, Salil S. Kulkarni, K. Hrisheekesh, Amit Agrawal, Shyamprasad Karagadde, Amitabh Bhattacharya, and Rajneesh Bhardwaj

CONTENTS

5.1 Computer Aided Engineering 1442
References 1444
5.2 Finite-Difference Method 1445
Introduction 1445
Main Idea of the FDM 1445
Governing Differential Equation and Boundary Conditions 1447
Grid Generation 1448
FDM-Based Algebraic Formulation 1450
Solution of the Algebraic Equations 1452
Concluding Remarks 1453
References 1453
5.3 Finite Volume Method 1454
Introduction 1454
References 1467
5.4 Finite Element Method 1468
Introduction 1468
Main Idea of the FEM 1468
Governing Differential Equation 1470
Steps in the Derivation of the Finite-Element Equations 1471
Semidiscrete Finite-Element Model in a General Case 1476
Solution Phase 1477
Numerical Evaluation of Elemental Matrices/Vectors 1481
Practical Issues in the Use of the FEM 1488
Examples 1490
References 1494
5.5 Lattice Boltzmann Method 1496
Introduction 1496
Background 1496
Lattice Boltzmann Method 1498
Example Problems 1503
Concluding Remarks 1504
References 1505
5.6 Immersed Boundary Methods For Fluid-Structure Interaction Simulations 1507
Introduction 1507
Continuous Forcing Approach 1509

Discrete Forcing Approach 1510
Applications of IB Methods for Moving Rigid Structures 1516
Applications of IB Methods with Moving Deformable Structures 1517
Concluding Remarks 1520
Acknowledgments 1521
References 1521
5.7 Numerical Methods for Multiphase Flows 1524
Introduction 1524
Continuum Modeling of Moving Interfaces 1524
Mathematical Formulation 1526
Enthalpy Method 1529
VOF Method 1533
Level Set Method 1537
Concluding Remarks 1542
References 1542
5.8 Large Eddy Simulation of Wall-Bounded Flows 1545
Introduction 1545
Properties of LES Filter 1546
Governing Equations 1548
Subgrid Stress Models 1550
Wall Modeling 1553
Grid Generation 1554
Case Study: LES of Channel Flow 1554
Summary 1556
References 1557
5.9 Software and Computer Codes 1558
Mathematical Software 1558
Computational Fluid Dynamics Software and Codes 1558

5.1 COMPUTER AIDED ENGINEERING

Atul Sharma

The *last fifty years* has seen tremendous evolution in computational fluid dynamics (CFD) from getting steady-state pictures to transient fluid dynamic movies of scientific-exciting flow properties and engineering-relevant parameters (Sharma, 2016); from solving the flow in lower to higher dimensions—2D to 3D; over or inside simple to complex shaped solid objects; involving single physics to multi-physics; and in a single to multiple parallelized computers. Recent advances in CFD involve its integration with other tools and its use for multiphysics simulations (incorporating structural, thermal, fluid, and electromagnetic analysis)—called *c*omputer *a*ided *e*ngineering.

The evolution has been possible mainly because of *fluid dynamicists* and *computer scientists*. Fluid dynamicists developed new or improved mathematical models and/or numerical methodologies, which could give results of acceptable accuracy in least computational time, whereas, computer scientists developed new or improved computer systems as well as programming paradigms to speedup the computations. Although presently CFD is more commonly used as an analysis tool, we can expect that CFD will be more often used as a design and optimization tool in future. This will lead to innovations in the development of the new products and improvements in the performance as well as efficiency of the existing products.

For fluid dynamics and heat transfer, we have a closed system of equations as there are three equations based on mass, momentum, and energy conservation laws and three dependent variables (pressure, velocity, and temperature). Laws of conservation of momentum and energy have a term

that describes the rate of change of momentum and energy; thus, they result in equations that are used to create fluid dynamics movie for velocity and temperature, respectively. Thereafter, we are left with the continuity equation based on the mass conservation law and pressure as a dependent variable. However, there is no rate of change of pressure term in the mass conservation law that could result in a fluid dynamic movie for pressure; in fact, the continuity equation does not have any pressure term. Thus, there is no explicit equation for pressure and the continuity equation needs to be converted into an equation for pressure. This created a lot of difficulties in handling pressure and coming up with a solution procedure for the Navier–Stokes (NS) equations in the early days of CFD. To circumvent these limitations, a pressure–velocity formulation, called as pressure correction approach, was proposed for incompressible flow; here, the continuity equation was employed as a constraint to derive the correct pressure field through a detailed iterative solution procedure.

The governing equations in fluid mechanics and heat transfer are *p*artial *d*ifferential *e*quations (PDEs), with flow properties (such as velocity, pressure, and temperature) as the dependent variables, and both spatial and temporal coordinates as the independent variables. The *exact solution* of the PDEs, subjected to certain *b*oundary *c*onditions (BCs), is a functional relationship for the flow property as a function of the spatial and temporal coordinates in a flow domain. The exact solution is continuous in both spatial and temporal coordinates. However, it is possible to obtain exact solutions for only a very few and simple problems in fluid dynamics and heat transfer, whereas a numerical solution is possible for almost all problems and is discrete in space and time. Thus, a numerical methodology requires two types of discretization: first, discretization of space and time and second, discretization of the PDEs and the BCs. The first one leads to *grid generation* for spatial discretization and time step Δt for the temporal discretization, whereas the second one leads to an *approximate algebraic formulation* for the equation discretization (PDEs and the BCs). The grid generation is a method to convert the infinite number of points in a flow domain to a certain fixed number of discrete points called grid points and determine their coordinates. The approximate algebraic formulation leads to a system of *l*inear *a*lgebraic *e*quations (LAEs), for the value of a flow property at all the grid points. Finally, the system of LAEs is solved by a solution methodology to obtain the spatial as well as temporal discrete solution for a flow property resulting in a fluid dynamic movie.

The exact solution is analogous to an infinite-resolution virtual video camera, and the numerical solution is analogous to a finite-resolution virtual video camera (Sharma, 2016). The camera results in a fluid dynamics movie for a flow property, with pixels analogous to the grid points and frame rate analogous to the time step. The numerical solution depends on the number of grid points for the spatial accuracy and on the magnitude of the time step for the temporal accuracy. The numerical solution results in a CFD software as a product, which is analogous to a virtual video camera.

The discretization method is an approximate algebraic formulation method to formulate a system of LAEs from the PDEs (Patankar, 1980) or physical-laws (proposed recently by Sharma, 2016) and BCs. Mainly, there are three discretization methods: *finite-difference method* (FDM), *finite-volume method* (FVM), and *finite-element method* (FEM); they are presented in the next three sections. FDM and FVM are more commonly used in CFD, whereas FEM is more popular in computational solid mechanics. Early development of CFD started with FDM for problems where the boundary of the computational domain corresponds to one (not any) of the three standard coordinate systems—Cartesian, cylindrical, and spherical; such problems were called simple (otherwise complex) geometry problems. Cartesian/cylindrical/spherical geometry problems were solved in a uniform or nonuniform grid, where one of the grid lines fits to the boundary of the computational domain—such a grid is called a body-fitted grid. With advances in CFD and more application to industrial problems, fitting the boundary of an computational domain to one of the standard coordinate systems was a big constraint; thus, there was a need to develop a method for computing flows in complex geometry. Initially, this was attempted with FDM on a body-fitted Cartesian grid, where the curved boundary of the computational domain was approximated as a stepped one approximating the small portions of the curved surface as either horizontal or vertical lines. However, for a very nonlinear curved domain, this led to loss of accuracy on a particular grid size or increase in

computational time (due to fine grid size needed at the boundary) to achieve a particular accuracy. Alternatively, an FDM-based method was proposed (Anderson, 1995), where certain equations were used to map the complex physical (x, y, t) domain to the simple Cartesian (ξ, η, τ) computational domain and then transform the PDEs in the complex domain to the simple domain. The resulting PDEs became quite complicated and leads to loss of computational efficiency and accuracy.

This led to the development of FVM where the governing PDEs are solved directly in the physical complex/curved domain (without transformation of domain and PDEs) on a body-fitted curvilinear structured grid. However, implementation of a staggered grid—a remedy to pressure–velocity decoupling—was found almost impossible for complex geometry; another remedy to obtain a physical solution in a colocated grid was proposed using momentum interpolation (Rhie and Chow, 1983). This was employed for quite some time with further development in the FVM-based solution that made it applicable in structured to multiblock structured to unstructured body-fitted grid systems. However, recently with more industrial applications of CFD where reasonably accurate results are desired in less computational time, there is a substantial increase in computational time for good-quality grid generation as very complex geometries as well as three-dimensional flow are commonly encountered. In terms of ease in code development, convergence characteristics of iterative solvers, memory requirement, computational cost, and computational accuracy, the Cartesian grid outweighs the body-fitted structured/unstructured grid. Thus, there is a renewed interest in the Cartesian grid (old is gold!) on *nonbody-fitted* systems instead of the earlier body-fitted grid, with a large body of work in the last decade on the Cartesian grid method.

One of the Cartesian grid methods, proposed for *f*luid–*s*tructure *i*nteraction problems, is the *immersed boundary method*, which is presented in Section 5.6. Other than the continuum methods based on the Navier-Stokes equations, the *lattice Boltzmann method* is presented in Section 5.5 as an example of mesoscale methods. Due to the large number and wide range of industrial applications, *numerical methods for two-phase flow* are gaining a lot of importance and are presented in *Section 5.7. Section 5.8* deals with *large eddy simulation of wall-bounded flow* for turbulent flows. Finally, this section on numerical methods and computational tools ends with a section on *software and computer codes.*

REFERENCES

Anderson, J.D. (1995). *Computational Fluid Dynamics: The Basics with Applications*, Mc-Graw Hill, New York.

Patankar, S.V. (1980). *Numerical Heat Transfer and Fluid Flow*, Hemisphere Publishing Corporation, New York.

Rhie, C.M. and Chow, W.L. (1983). A numerical study of the turbulent flow past an isolated airfoil with trailing edge separation, *AIAA Journal*, vol. 21, pp. 1525–1532.

Sharma, A. (2016). *Introduction to Computational Fluid Dynamics: Development, Application and Analysis*, John Wiley & Sons, Chichester, UK.

5.2 FINITE DIFFERENCE METHOD

Atul Sharma

Introduction

Finite-difference method (FDM) is a discretization method that uses Taylor series to convert each differential term in a partial differential equation (PDE) to a linear algebraic equation (LAE), called finite-difference equation (FDE). FDM is relatively older as compared to finite-volume method. FDM is algorithmically simple, efficient, and accurate; however, the FDM was found better suited for the simple as compared to complex geometry in a computational fluid dynamics (CFD) problem. With the advances in CFD, recently few FDM-based methods demonstrated its accurate applicability for the complex geometry as well as moving boundary problems; and created a renewed interest in the FDM. One such method is the immersed boundary method, presented in Section 5.6. In this section, an FDM based discretization as well as numerical methodology is presented for 2D steady-state heat conduction, in a Cartesian coordinate system and on a uniform grid.

Main Idea of the FDM

The algebraic formulation in FDM is based on Taylor series expansion of a flow property. Considering temperature as the flow property for the heat transfer here, the Taylor series is an *infinite series* representing temperature T at a location $x + \Delta x$, $T_{x+\Delta x}$, as an infinite sum of terms, consisting of the value and the derivatives of the function T at a *single point* x; given as

$$T_{x+\Delta x} = T_x + \left(\frac{\mathrm{d}T}{\mathrm{d}x}\right)_x \Delta x + \left(\frac{\mathrm{d}^2T}{\mathrm{d}x^2}\right)_x \frac{\Delta x^2}{2!} + \left(\frac{\mathrm{d}^3T}{\mathrm{d}x^3}\right)_x \frac{\Delta x^3}{3!} + \cdots + \left(\frac{\mathrm{d}^nT}{\mathrm{d}x^n}\right)_x \frac{\Delta x^n}{n!} \tag{5.2.1}$$

where Δx is any positive or negative spacing from the single representative location x.

For a 2D problem with x and y as the independent variables, the Taylor series expansion is given as

$$T_{x+\Delta x,y} = T_{x,y} + \left(\frac{\partial T}{\partial x}\right)_{x,y} \Delta x + \left(\frac{\partial^2 T}{\partial x^2}\right)_{x,y} \frac{\Delta x^2}{2!} + \cdots + \left(\frac{\partial^n T}{\partial x^n}\right)_{x,y} \frac{\Delta x^n}{n!} \tag{5.2.2}$$

$$T_{x-\Delta x,y} = T_{x,y} - \left(\frac{\partial T}{\partial x}\right)_{x,y} \Delta x + \left(\frac{\partial^2 T}{\partial x^2}\right)_{x,y} \frac{\Delta x^2}{2!} \cdots (-1)^n \left(\frac{\partial^n T}{\partial x^n}\right)_{x,y} \frac{\Delta x^n}{n!} \tag{5.2.3}$$

where the left-hand side (LHS) consists of T at a spacing of Δx and $-\Delta x$ from x, and same y. From *each* of the above equation, the first derivative is given as

$$\left(\frac{\partial T}{\partial x}\right)_{x,y} = \frac{T_{x+\Delta x,y} - T_{x,y}}{\Delta x} - \left[\left(\frac{\partial^2 T}{\partial x^2}\right)_{x,y} \frac{\Delta x}{2!} + \cdots\right] \tag{5.2.4}$$

$$\left(\frac{\partial T}{\partial x}\right)_{x,y} = \frac{T_{x,y} - T_{x-\Delta x,y}}{\Delta x} + \left[\left(\frac{\partial^2 T}{\partial x^2}\right)_{x,y} \frac{\Delta x^2}{2!} + \cdots\right] \tag{5.2.5}$$

Using *both* the equations, subtracting Equation 5.2.3 from Equation 5.2.2 results in an FDE for the first derivative, and adding both the equations results in an FDE for the second derivative; given as

$$\left(\frac{\partial T}{\partial x}\right)_{x,y} = \frac{T_{x+\Delta x,y} - T_{x-\Delta x,y}}{2\Delta x} - \left[\left(\frac{\partial^3 T}{\partial x^3}\right)_{x,y} \frac{\Delta x^2}{3!} + \cdots\right] \tag{5.2.6}$$

$$\left(\frac{\partial^2 T}{\partial x^2}\right)_{x,y} = \frac{T_{x+\Delta x,y} - 2T_{x,y} + T_{x-\Delta x,y}}{\Delta x^2} - 2\left[\left(\frac{\partial^4 T}{\partial x^4}\right)_{x,y} \frac{\Delta x^2}{4!} + \cdots\right] \tag{5.2.7}$$

Division of a dividend with a divisor results in a quotient and a remainder; similarly, for the right-hand side of Equations 5.2.4 through 5.2.7, the first term is called *finite-difference quotient* and the second term (within a square bracket) is called truncated term—resulting in a *truncation error*. The truncated terms in the equations show higher order derivatives multiplied by a certain power of Δx. The power of the Δx in the lowest order derivative of the truncated terms decides the order of accuracy of a finite-difference representation.

The finite-difference quotient and the order of accuracy are shown more clearly in Figure 5.2.1, for the first and second derivative. The figure shows the spatial coordinates in Equations 5.2.4 through 5.2.7 replaced by the running indices: (x, y) by (i, j), $(x + \Delta x, y)$ by $(i + 1, j)$, and $(x - \Delta x, y)$ by $(i - 1, j)$. The running indices corresponds to the element number in a 2D matrix, which is used as a data structure to store the discrete values of the temperature; discussed in the next subsection. The figure shows the various finite-difference quotients as an algebraic term. A quotient that involves one-point forward, one-point backward, and both forward and backward points is called forward,

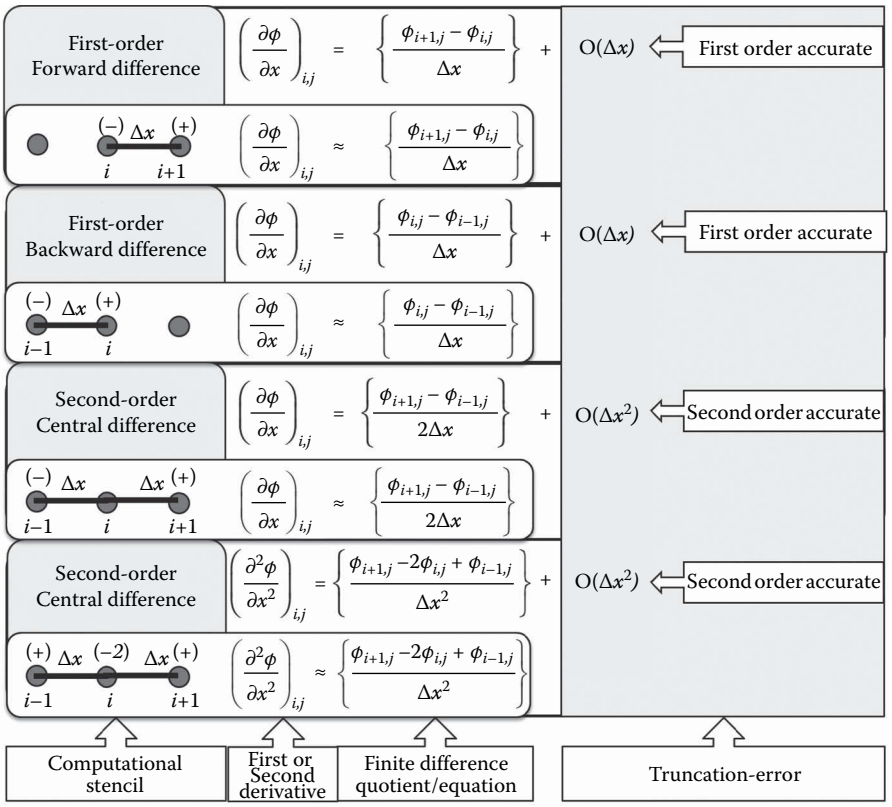


FIGURE 5.2.1 Illustration of the various types of FDM-based algebraic representation of the first and second derivatives, using Taylor series expansion.

backward, and central difference method, respectively. The points involved in the quotient are also shown schematically (joined by straight lines) in the figure, called computational stencil.

The FDEs (Figure 5.2.1) are also applicable for the derivatives in the *y*-direction, with certain replacements: $T_{i+2,j}$ by $T_{i,j+2}$, $T_{i+1,j}$ by $T_{i,j+1}$, $T_{i-1,j}$ by $T_{i,j-1}$, $T_{i-2,j}$ by $T_{i,j-2}$, Δx by Δy, and Δx^2 by Δy^2. Note that the grid point considered in the FDE for $\partial T/\partial x$ and $\partial T/\partial y$ are on a horizontal and a vertical line, with constant y and x coordinates, respectively. Further *higher order* one-sided (forward/backward) or both-sided (central) FDE can also be derived using the above procedure. Note that the procedure corresponds to the algebraic operation of the Taylor series for T at neighboring grid points.

For the first and second derivatives in the x and y directions, consider a generic FDE (Sharma 2016) as

$$\begin{aligned}\left(\frac{\partial T}{\partial x}\right)_{i,j}\Delta x,\left(\frac{\partial^2 T}{\partial x^2}\right)_{i,j}\Delta x^2 &\approx ffT_{i+2,j}+fT_{i+1,j}+cT_{i,j}+bT_{i-1,j}+bbT_{i-2,j}\\ \left(\frac{\partial T}{\partial y}\right)_{i,j}\Delta y,\left(\frac{\partial^2 T}{\partial y^2}\right)_{i,j}\Delta y^2 &\approx ffT_{i,j+2}+fT_{i,j+1}+cT_{i,j}+bT_{i,j-1}+bbT_{i,j-2}\end{aligned} \tag{5.2.8}$$

The values of the coefficients for the various neighboring temperatures in the above equation are presented in Table 5.2.1. The table shows first- and second-order forward/backward difference method for the first derivative, and second- and fourth-order central difference method for both first and second derivative. For the central difference method, note from the table that the coefficient *f* is equal to *b*, and *ff* is equal to *bb*; however, the equality is with a *change in sign* of the coefficients for the first derivative but not for the second derivative. Furthermore, the sum of the coefficients is equal to zero—the approximate representation ensures that the derivative is zero for a uniform temperature field. Finally, note that the coefficients in the table are same for the first/second derivative in the *x*-and *y*-directions.

Governing Differential Equation and Boundary Conditions

The FDM-based algebraic formulation and solution are presented in the following subsections, for 2D steady-state heat conduction with volumetric heat generation. The corresponding *governing equation* is given as

$$\text{G.E: } \frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\overline{Q}_{\text{gen}}}{k}=0 \tag{5.2.9}$$

TABLE 5.2.1
The Coefficients in the General FDE (Equation 5.2.8) for the First and the Second Derivatives in *x*/*y* Direction (Sharma 2016)

Derivatives	FDM	Order of Accuracy	*bb*	*b*	*c*	*f*	*ff*
First	Forward	First			−1	1	
		Second			−3/2	2	−1/2
	Backward	First		−1	1		
		Second	1/2	−2	3/2		
	Central	Second		−1/2	0	1/2	
		Fourth	1/12	−2/3	0	2/3	−1/12
Second	Central	Second		1	−2	1	
		Fourth	−1/12	4/3	−5/2	4/3	−1/12

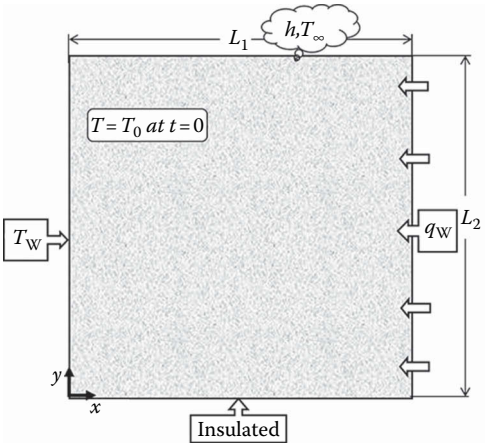


FIGURE 5.2.2 Conduction heat transfer in a rectangular plate, subjected to an initial and different types of thermal boundary conditions.

The above equation corresponds to the various heat conduction problems; however, the spatial variation in the steady-state temperature field depends on boundary conditions. Consider an example of 2D unsteady-state heat conduction in a square plate shown in Figure 5.2.2. The figure shows that the plate is subjected to the different types of thermal boundary conditions. Mathematically, the boundary conditions are classified into three types: *Dirichlet*, if the value of a flow property ($T/\vec{u}/p$) is prescribed; *Neumann*, if the gradient of the variable is prescribed; and *Robin or mixed*, if a linear combination of the variable and its gradient is prescribed at the boundary. For the example problem in Figure 5.2.2, the different types of boundary conditions (BCs) are as follows:

1. Dirichlet BC: *c*onstant *w*all *t*emperature BC at the left wall
2. Neumann BC: insulated BC at the bottom wall, and the *u*niform *h*eat *f*lux (UHF) BC at the right wall
3. Robin or mixed BC: convective BC at the top wall

The various BCs in Figure 5.2.2 are represented mathematically as

$$\begin{aligned}
&\text{BCs: At } x = 0 \Rightarrow T = T_W \\
&\text{At } x = L_1 \Rightarrow -k\left(\frac{dT}{dx}\right)_{x=L_1} = -q_W \\
&\text{At } y = 0 \Rightarrow \left(\frac{dT}{dy}\right)_{y=0} = 0 \\
&\text{At } y = L_2 \Rightarrow -k\left(\frac{dT}{dy}\right)_{y=L_2} = h(T_{y=L_2} - T_\infty)
\end{aligned} \tag{5.2.10}$$

Grid Generation

The governing equations and the BCs presented above are converted into a system of LAE, using an FDM in the next subsection. The LAEs correspond to the temperature at various discrete grid points in the computational domain—the grid points are generated by a method, presented in this subsection.

A grid generation is a method for the *spatial discretization* of the internal region of the domain as internal grid points, and the boundary region discretized as boundary grid points. The grid points are shown in Figure 5.2.3, for a simple geometry problem in a Cartesian coordinate system (Figure 5.2.2). The figure shows equispaced horizontal and vertical lines, leading to the simplest form of grid generation. The lines are called grid lines, and it is ensured that the boundary of the domain is represented by one of the grid lines—for a *body-fitted grid*.

Figure 5.2.3 shows a *uniform grid*, with the spacing between the consecutive horizontal line as Δx and the vertical line as Δy—called *grid spacing*. As discussed previously, the grid points for the application of FDM to PDE and BCs are different. Thus, the grid points are classified here into two types: boundary and internal. The figure also shows i and j as the *running indices* (consisting of integer values), which act as a tag (or address) for a grid point. The respective running indices correspond to the rows and columns of a matrix—a computationally efficient data structure to store the values of a flow property at the grid points shown in the figure. For a representative grid point (i, j), Figure 5.2.3 shows that $(i + 1, j)$, $(i - 1, j)$, $(i, j + 1)$, and $(i, j - 1)$ correspond to the adjoining east, west, north, and south neighboring grid points, respectively. The figure also shows that the respective neighboring grid points are located at $(x + \Delta x, y)$, $(x - \Delta x, y)$, $(x, y + \Delta y)$, and $(x, y - \Delta y)$.

The in-built neighboring information in the matrix makes the computer programming in CFD much easier. Using a *s*econd-order FDM on the governing PDE for a 2D steady-state conduction, it is presented below that the temperature at *any* internal grid point is represented in terms of *four* (east, west, north, and south; not all) adjoining neighboring grid points. This functional relationship is presented schematically in Figure 5.2.3, for *one* of the representative internal grid point, and is called *computational stencil*. Note that the functional dependence remains same for all the other internal grid points. Using second-order FDM for the non-Dirichlet BC, it is presented below (Equation 5.2.12) that the temperature at a boundary grid point is a function of two adjoining internal

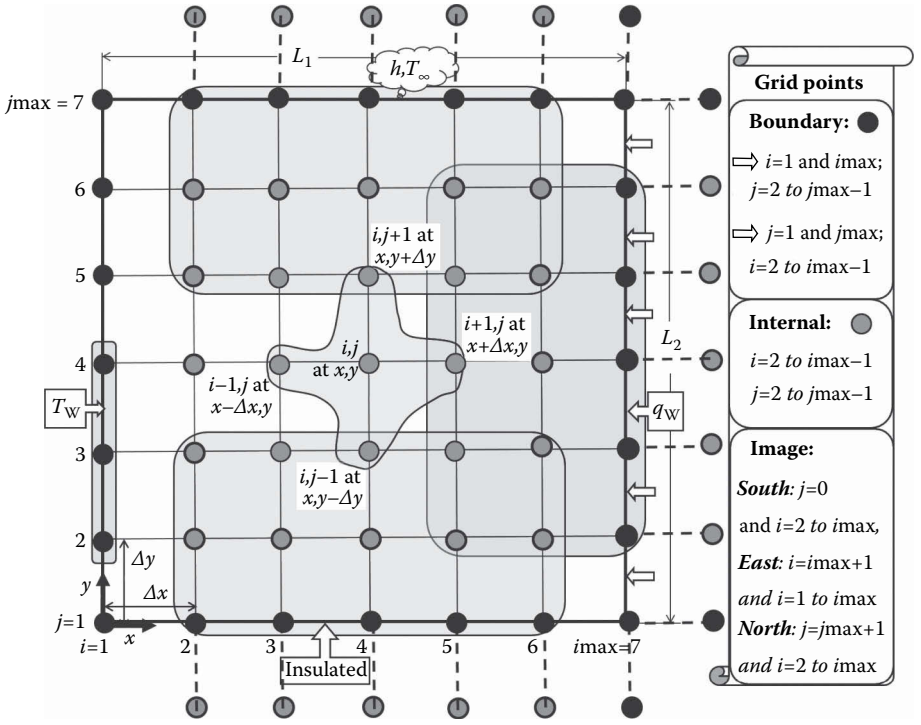


FIGURE 5.2.3 A uniform Cartesian grid generation, for an FDM. The figure also shows the various types of grid points, computational stencils (for an internal grid point and various boundary grid points), and the BCs for a 2D heat conduction problem (Figure 5.2.2).

grid points; and the temperature is constant for a Dirichlet BC. The corresponding computational stencil for the Dirichlet BC at the west boundary, and non-Dirichlet BCs at the other boundaries, are shown in Figure 5.2.3. For the grid point-by-point recursive computation, a range of values of the running indices (i and j) is shown in Figure 5.2.3; separately for the internal and boundary grid points, since their computational stencils are different. The recursive computation corresponds to scrolling through all the internal grid points for an iterative solution of the corresponding LAE, with the help of a loop in a computer program; presented below in the solution methodology.

FDM-Based Algebraic Formulation

Using the second-order FDM for both governing equations and BCs, presented in Equation 5.2.8 and Table 5.2.1, the algebraic formulation for the differential term in Equations 5.2.9 and 5.2.10 results in the FDE as

$$\text{FDE for G.E}: T_{i-1,j} + T_{i+1,j} + \beta^2 (T_{i,j-1} + T_{i,j+1}) - 2(1+\beta^2) T_{i,j} + \frac{\bar{Q}_{\text{gen},i,j} \Delta x^2}{k} = 0 \quad \text{where } \beta = \Delta x^2 / \Delta y^2 \tag{5.2.11}$$

$$\begin{aligned} &\text{FDE for BCs}: \text{At } x = 0 \quad T_{1,j} = T_{\text{W}} \equiv T_{\text{wb}} \\ &\text{At } x = L_1 \quad T_{i\text{max},j} = \frac{4T_{i\text{max}-1,j} - T_{i\text{max}-2,j}}{3} + \frac{2q_{\text{W}} \Delta x}{3k} \equiv T_{\text{eb}} \\ &\text{At } y = 0 \quad T_{i,1} = \frac{4T_{i,2} - T_{i,3}}{3} \equiv T_{\text{sb}} \\ &\text{At } y = L_2 \quad T_{i,j\text{max}} = \frac{k(4T_{i,j\text{max}-1} - T_{i,j\text{max}-2}) + 2h\Delta y T_{\infty}}{(3k + 2h\Delta y)} \equiv T_{\text{nb}} \end{aligned} \tag{5.2.12}$$

where *central difference method* is used for the second derivative in the *GE*, and the second-order *forward (or backward) difference method* for the first derivative in the *BCs*. The application of the forward difference at the south boundary, and the backward difference at the east as well as the north boundary, is decided by the one-sided forward or backward location of the internal grid points adjoining the corresponding boundary grid points (Figure 5.2.3).

For the 2D steady-state heat conduction with volumetric heat generation, on a uniform grid (with $\Delta x = \Delta y$), Equations 5.2.11 and 5.2.12 finally results in FDE as

$$\left.\begin{aligned} -T_{i-1,j} - T_{i+1,j} - T_{i,j-1} - T_{i,j+1} + 4T_{i,j} = S_{\text{T}} \\ \text{where } S_{\text{T}} = \bar{Q}_{\text{gen},i,j} \Delta x^2 / k \end{aligned}\right\} \quad \begin{aligned} &\text{for } i = 2 \text{ to } i\text{max} - 1 \\ &\text{for } j = 2 \text{ to } j\text{max} - 1 \end{aligned}$$

$$\left.\begin{aligned} T_{1,j} = T_{\text{wb}} (\text{west boundary}) \\ T_{i\text{max},j} = T_{\text{eb}} (\text{east boundary}) \end{aligned}\right\} \quad \text{for } j = 2 \text{ to } j\text{max} - 1 \tag{5.2.13}$$

$$\left.\begin{aligned} T_{i,1} = T_{\text{sb}} (\text{south boundary}) \\ T_{i,j\text{max}} = T_{\text{nb}} (\text{north boundary}) \end{aligned}\right\} \quad \text{for } i = 2 \text{ to } i\text{max} - 1$$

The boundary temperature value above is a constant for a Dirichlet BC, and gets updated (after every iteration, using Equation 5.2.12) for a non-Dirichlet BC. Since the temperature at a particular internal grid point depends on the temperature at adjoining four (not all) grid points in the domain, the system of LAEs for Equation 5.2.13 results in a *penta-diagonal* coefficient matrix. The matrix consists of five nonzero diagonal vector: one main, two lower, and two upper diagonal. The nonzero value corresponds to 4 for the main diagonal, and −1 for all the upper and lower diagonals.

For obtaining the FDE at those boundary grid points which are subjected to the non-Dirichlet BC (involves derivative of temperature), FDM involves another widely used technique—called an *image-point technique*. The technique considered image grid points outside those boundaries which are subjected to the non-Dirichlet BC, shown outside the south, east, and north boundary in Figure 5.2.3. For the south boundary ($j = 1$) in Figure 5.2.3, the second-order central difference FDM for the insulated BC is given as

$$\left(\frac{\mathrm{d}T}{\mathrm{d}y}\right)_{i,1} = \frac{T_{i,2} - T_{i,0}}{2\Delta y} = 0 \Rightarrow T_{i,0} = T_{i,2} \tag{5.2.14}$$

where $T_{i,0}$ is the temperature at the image grid point, located at a distance Δy below the south boundary for each i (Figure 5.2.3). Applying the FDE for G.E. to the south boundary grid point (Equation 5.2.11 for $T_{i,1}$), and substituting $T_{i,0} = T_{i,2}$ (Equation 5.2.14), we obtain

$$\left.\begin{aligned} T_{i-1,1} + T_{i+1,1} + 2\beta^2 T_{i,2} - 2(1+\beta^2)T_{i,1} + \frac{\bar{Q}_{\mathrm{gen},i,1}\Delta x^2}{k} = 0 \end{aligned}\right\} \text{ for } i = 1 \text{ to } i\mathrm{max} - 1 \tag{5.2.15}$$

Similarly, using the second-order FDM for both non-Dirichlet BC and GE at the east and north boundary grid points, the final FDE for the respective grid points ($T_{i\mathrm{max},j}$ and $T_{i,j\mathrm{max}}$) are given as

$$\left.\begin{aligned} 2T_{i\mathrm{max}-1,j} + \beta^2(T_{i\mathrm{max},j-1} + T_{i\mathrm{max},j+1}) - 2(1+\beta^2)T_{i\mathrm{max},j} \\ + \frac{\bar{Q}_{\mathrm{gen},i\mathrm{max},j}\Delta x^2}{k} + \frac{2q_{\mathrm{W}}\Delta x}{k} = 0 \end{aligned}\right\} \text{for } j = 1 \text{ to } j\mathrm{max} - 1 \tag{5.2.16}$$

$$\left.\begin{aligned} T_{i-1,j\mathrm{max}} + T_{i+1,j\mathrm{max}} + 2\beta^2 T_{i,j\mathrm{max}-1} - 2\left(1 + \beta^2 + \frac{\beta^2 h\Delta y}{k}\right)T_{i,j\mathrm{max}} \\ + \frac{\bar{Q}_{\mathrm{gen},i,j\mathrm{max}}\Delta x^2}{k} + \frac{2\beta^2 h\Delta y}{k}T_{\infty} = 0 \end{aligned}\right\} \text{for } i = 1 \text{ to } i\mathrm{max} - 1 \tag{5.2.17}$$

For the boundary grid points at the non-Dirichlet boundary, note that either the FDE for BCs (Equation 5.2.12) or the FDE for both GE and BCs (Equations 5.2.15 through 5.2.17; obtained by image-point technique) can be used. Both the methods are commonly used, and presented in Chapter 5 of an edited book by Muralidhar and Sundarajan (2003).

The boundary grid-point value at the corner of the domain, involved in Equations 5.2.15 through 5.2.17, are taken as T_{wb} for the south-west and north-west corner grid point ($T_{1,1} = T_{1,j\mathrm{max}} = T_{\mathrm{wb}}$); and obtained by the image-point technique for the south-east corner and north-east corner grid points (which are exposed to two non-Dirichlet BCs) as

$$\left.\begin{aligned} T_{i\mathrm{max}-1,1} + \beta^2 T_{i\mathrm{max},2} - (1+\beta^2)T_{i\mathrm{max},1} \\ + \frac{\bar{Q}_{\mathrm{gen},i\mathrm{max},1}\Delta x^2}{2k} + \frac{q_{\mathrm{W}}\Delta x}{k} = 0 \end{aligned}\right\} \text{for } i = i\mathrm{max} \text{ and } j = 1 \tag{5.2.18}$$

$$\left.\begin{aligned} T_{i\max-1,j\max} + \beta^2 T_{i\max,j\max-1} - \left(1+\beta^2+\frac{\beta^2 h\Delta y}{k}\right)T_{i\max,j\max} \\ + \frac{\bar{Q}_{\text{gen},i\max,j\max}\Delta x^2}{2k} + \frac{\beta^2 h\Delta y}{k}T_\infty + \frac{q_W\Delta x}{k} = 0 \end{aligned}\right\} \text{for } i = i\max \text{ and } j = j\max \quad (5.2.19)$$

The above equations are obtained by substituting $T_{i\max,0} = T_{i\max,2}$ in Equation 5.2.16, and $T_{i\max+1,j\max} = (2q_W\Delta x/k) + T_{i\max-1,j\max}$ in Equation 5.2.17, respectively.

Solution of the Algebraic Equations

There are two approaches for the solution of the system of LAEs: direct and indirect. *Direct method* is based on algebraic eliminations for obtaining solution in a fixed number of operations; for example, matrix inversion and Gauss elimination. *Indirect method* is used to obtain the solution asymptotically by an *iterative* procedure; for example, Jacobi and Gauss–Seidel method.

Iterative or indirect methods are generally used under four conditions (Hoffman, 1993): first, the number of equation is large (100 or more); second, most of the coefficient a_{ij} are zero (sparse matrix); third, the system of equations is diagonally dominant; and fourth, the system of equations is not ill-conditioned. A coefficient matrix is called *diagonally dominant* if the absolute value of each term on the major diagonal is equal to, or larger than, the sum of the absolute value of all the other terms in that row. These conditions are commonly encountered in CFD: the first condition due to large grid size, the second condition due to the discretization method (FDM or FVM) resulting in a LAE with sparse coefficient matrix (pentadiagonal coefficient matrix, discussed previously), and the third condition is ensured during the discretization, and the fourth condition is avoided during the discretization. Thus, it is generally more efficient in terms of computational time and storage requirements to solve the system of LAEs in CFD by an iterative than a direct method.

Jacobi and Gauss–Seidel methods are two of the simplest iterative methods; and will be presented here. For these methods, each of the LAE is converted into an explicit equation for one of the element of the unknown vector, and the values of the other elements are taken from the *previous iteration* for the *Jacobi method,* and the *improved (or latest) value* for the *Gauss–Seidel* method. The sequence of the solution of the LAEs is relevant for the Gauss–Seidel but not for the Jacobi method; this is because of the iterative solution for the system of LAEs is *simultaneous* for Jacobi and *successive* for Gauss–Seidel method. The diagonal dominance is a *sufficient* condition for the convergence of the Jacobi and Gauss–Seidel methods; for any initial guess.

Considering $N + 1$ as the *present* and N as the *previous* iteration number, and $T_{i,j}^{N+1}$ and $T_{i,j}^{N}$ as the respective iterative value, the actual *error* in the solution ($T_{i,j}^{N+1} - T_{i,j}^{\text{exact}}$) is approximated as $\Delta T_{i,j} \equiv T_{i,j}^{N+1} - T_{i,j}^{N}$ (Hoffman, 1993). This error is computed for all the N values of T and should be reduced *simultaneously* for all of them, to ensure the minimum accuracy desired (as per the value of $\in$) for the system of equations. One of the most common convergence criteria is given as

$$\left|\Delta T_{\text{for all } i,j}\right|_{\max} \le \epsilon \quad (5.2.20)$$

where the LHS represents absolute value of the maximum from all the grid-point values of $\Delta T_{i,j}$, and $\in$ is a *convergence tolerance;* corresponding to user-defined zero—say 10^{-3} or 10^{-5}. The number of iterations required to achieve convergence depends on four conditions (Hoffman, 1993): first, the dominance (in terms of magnitude) of the diagonal coefficients; second, the initial guess for the unknowns; third, the iterative method used; and fourth, the prescribed value of convergence criterion.

Concluding Remarks

Finite difference-based numerical methodology is presented in this section—starting with governing equation and boundary conditions, followed by grid generation and algebraic formulation, and ending with the iterative solution methodology.

The methodology is presented for a simpler case of 2D steady-state heat conduction, with Jacobi and Gauss–Seidel iterative method on a uniform grid. However, for the solution of various model equations in fluid dynamics and heat transfer on a nonuniform grid, similar FDM method-based methodology can be found in the book by Anderson et al. (1984), Hoffmann and Chiang (1989), and Anderson (1995). These books also present the other iterative methods; such as, conjugate-gradient method and multigrid method. The FDM-based solution can also be found in an edited book by Muralidhar and Sundarajan (2003)—Chapter 6 on stream-function vorticity method, for a 2D incompressible flow, and Chapter 7 on Marker-and-Cell method, for both heat and fluid flow in a convective heat-transfer problem.

REFERENCES

Anderson, D.A., J.C. Tannehill, and R.H Pletcher. 1984. *Computational Fluid Mechanics and Heat Transfer*, 1st edition. New York: Hemisphere Publishing Corporation.

Anderson, J.D. 1995. *Computational Fluid Dynamics: The Basics with Applications*. New York: McGraw-Hill.

Hoffman, J.D. 1993. *Numerical Methods for Engineers and Scientists*. New York: McGraw-Hill.

Hoffmann, K.A. and S.T. Chiang, 2000. *Computational Fluid Dynamics*, Vols. 1, 2 and 3. Wichita, KS: Engineering Education System.

Muralidhar, K. and T. Sundarajan, eds. 2003. *Computational Fluid Flow and Heat Transfer*, 2nd edition. New Delhi, India: IIT Kanpur Series, Narosa Publishing House.

Sharma, A. 2016. *Introduction to Computational Fluid Dynamics: Development, Application and Analysis*, John Wiley & Sons, Chichester, UK.

5.3 FINITE VOLUME METHOD

Atul Sharma

Introduction

Finite volume method (FVM) is a discretization method that starts with both volume and time integral for each term of the partial differential equation (PDE); thereafter, uses certain approximations for the spatial and temporal variation of the integrand; and finally, results in the system of linear algebraic equations (LAEs). Instead of the differential form of the governing equation in the finite-difference method (FDM), the volume integral form of the equations in the FVM gives greater flexibility in handling the irregular shape of the control volumes (CVs), for complex geometry problems in computational fluid dynamics (CFD). If all the boundaries of the computational domain are aligned along *one* (not any) of the three *standard* coordinate systems: Cartesian, cylindrical, and spherical, they are called simple geometry problem; otherwise called complex geometry problem. In this subsection, an FVM-based algebraic formulation is presented for *two-dimensional* unsteady-state convective heat transfer, in a *complex geometry* CFD problem, with a curvilinear structured grid.

Governing Differential Equations

For incompressible flow, a compact *coordinate-free* conservative form of the governing PDEs is given as

$$\begin{aligned}
&\text{Continuity}: && \nabla\cdot(\rho\vec{u})=0\\
&\text{Momentum}: && \frac{\partial(\rho\vec{u})}{\partial t}+\nabla\cdot(\rho\vec{u}\vec{u})=\mu\nabla^2\vec{u}-\overrightarrow{\nabla p}\\
&\text{Energy}: && \frac{\partial(\rho c_pT)}{\partial t}+\nabla\cdot(\rho\vec{u}c_pT)=k\nabla^2T+\vec{Q}_{\text{gen}}\\
&\text{Transport}: && \frac{\partial(\rho C\phi)}{\partial t}+\nabla\cdot(\rho\vec{u}C\phi)=\Gamma_\phi\nabla^2\phi+S_\phi
\end{aligned} \tag{5.3.1}$$

where, for this generalized transport equation, $\phi=\vec{u}$, $C=1$, $\Gamma_\phi=\mu$ and $S_\phi=-\nabla p$ for the momentum equation, and $\phi=T$, $C=c_p$, $\Gamma_\phi=k$, and $S_\phi=\overline{Q}_{\text{gen}}$ for the energy equation. The Cartesian, cylindrical, and spherical coordinate versions of these Navier–Stokes equations are obtained by substituting appropriate form of the velocity vector $\vec{u}$ and the gradient operator ∇; thus, these equations are called the coordinate-free form (Muralidhar and Biswas, 2005).

Grid Generation

For an FVM, a commonly considered grid generation is presented below first for a Cartesian and then for a complex geometry. A simplest example of grid generation is shown in Figure 5.3.1, for a simple Cartesian geometry problem. A nonuniform Cartesian grid generation is shown in the figure as a procedure of drawing nonuniformly spaced horizontal and vertical lines. The lines subdivide the computational domain to subdomains, called CVs. The points located at the centroid of the CVs are called internal grid points. The figure also shows grid points located at the boundary of the domain, called boundary grid points. Similar grid generation procedure is already introduced in Figure 5.3.1 for the FDM. However, the circles in both the figures shows that the grid points are defined differently in the FDM and the FVM—at the *vertex* of the CVs for the *FDM* (Figure 5.2.3) and at the *centroid* of the CVs for the FVM (Figure 5.3.1).

Figure 5.3.1 shows a subdivision of a 2D Cartesian domain into contiguous nonoverlapping nonuniform CVs—also called *cells* in the FVM. The CVs are generated by a nonuniform Cartesian grid

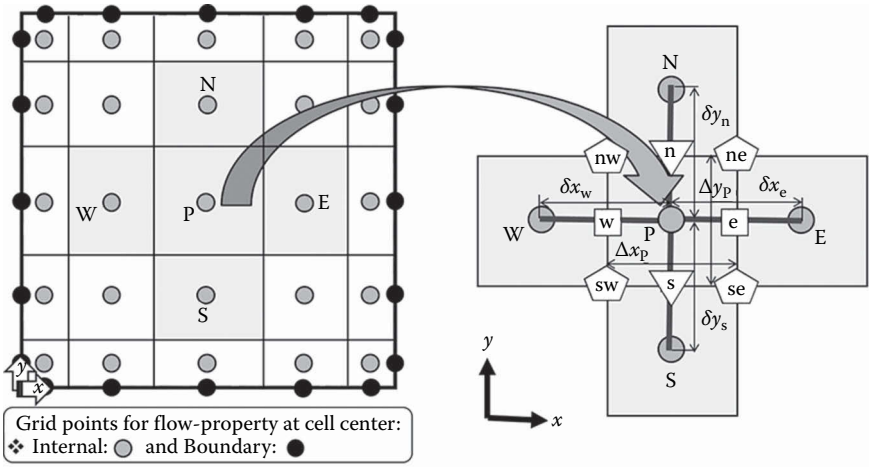


FIGURE 5.3.1 A nonuniform Cartesian grid generation, for an FVM. The figure also shows a representative Cartesian CV and its neighboring CVs.

generation method that leads to the nonuniformly spaced horizontal and vertical lines. The circles in the figure are the symbolic representation of the centroids of the CVs, which are also called *cell centers* or *internal grid points*. They are also represented by P for a representative 2D Cartesian CV; and by E, W, N, and S for the east, west, north, and south neighboring cell centers, respectively. Along with the centroid of the CVs, the figure also shows the centroids of the various faces, called *face centers*; represented by e, w, n, and s for the east, west, north, and south surfaces, respectively. They are differentiated by symbols (Sharma, 2016), as square at the vertical surfaces (for the e and w face centers), and triangle at the horizontal surfaces (for the n and s face centers). The pentagon symbol in the figure represents the corners of the representative CV and are indicated by ne, nw, se, and sw. Here, note from the figure that the capital letters (E, W, N, S) correspond to the cell centers/grid points, and small letters correspond to face centers (e, w, n, s) as well as corners (ne, nw, se, sw) of a CV.

Similar cell centers, face centers, and corners are shown in Figure 5.3.2, for a representative CV in a complex geometry problem. The grid is generated by a curvilinear grid generation method, presented in a book by Sharma (2016). Compared to the Cartesian CV for simple geometries, presented in Figure 5.3.1, Figure 5.3.2 shows that a complex CV for a complex geometry has two major differences. First, the faces for the CV are not orthogonal to each other and are not inclined along the direction of the Cartesian coordinates. Second, the lines joining the cell center P and neighboring cell centers are also not inclined along the direction of the Cartesian coordinates, do not pass through the corresponding face centers, and are not normal to the face. For the complex geometry problem, note from the figure that the domain is represented by the Cartesian coordinate system, and an FVM is presented later for the Cartesian components of the velocity field.

Main Idea of the FVM

The main idea for the FVM is that it results in algebraic formulation for the discrete heat and fluid flow field which satisfies the mass, momentum, and energy conservation laws to machine-level accuracy. The main idea in the formulation procedure for the FVM corresponds to the integral form of the governing partial differential equations, and two approximations presented later in separate subsections.

For simplicity, the main idea of the algebraic formulation is presented below for energy equation (Equation 5.3.1) and for the Cartesian CV (Figure 5.3.1). However, it is also applicable to the continuity and momentum equation (Equation 5.3.1), and for the complex CV (Figure 5.3.2) presented later.

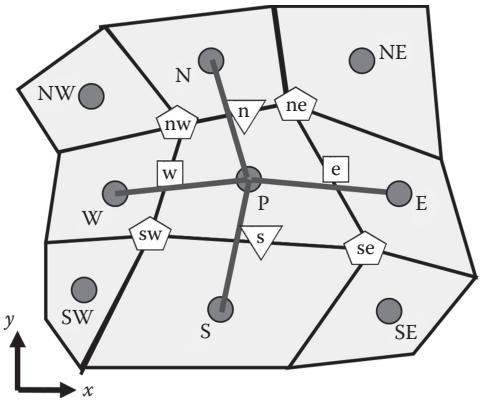


FIGURE 5.3.2 A representative complex CV for a complex geometry, with nonuniform, nonorthogonal, and curvilinear grid arrangement.

Integral Form of the Governing Equation

For an *Cartesian domain* of volume V bounded by a closed surface $\vec{S}$, the FVM starts with a volume integration of the governing partial differential equations, for a representative internal CV of volume ΔV_P (Figure 5.3.1), and a time integration between the previous time instant t and the present time instant $t+\Delta t$. This is given, for the energy equation (Equation 5.3.1), as

$$\int_{t}^{t+\Delta t}\int_{\Delta V_P}\frac{\partial(\rho c_p T)}{\partial t}\mathrm{d}V\mathrm{d}t+\int_{t}^{t+\Delta t}\int_{\Delta V_P}\nabla\cdot(\rho\vec{u}c_pT-k\overline{\overline{\nabla T}})\mathrm{d}V\mathrm{d}t=\int_{t}^{t+\Delta t}\int_{\Delta V_P}\overline{Q}_{\mathrm{gen}}\mathrm{d}V\mathrm{d}t$$

It is important to note from this integral equation that the integrands are either a volumetric term (volumetric enthalpy $\overline{\mathcal{V}_{\mathrm{T}}}=\rho c_{\mathrm{p}}T$ and volumetric heat generation $\overline{Q}_{\mathrm{gen}}$) or a flux term (enthalpy flux $\vec{h}=\rho\,\overline{\vec{u}}\,c_{\mathrm{p}}T$ and conduction flux $\vec{q}=-k\overline{\overline{\nabla T}}$). The integrands that represent a flux term are converted into surface integral—using *Gauss divergence theorem*. During this operation, one of the ∇ operator (in the integrand) pulls out, and we end up with a dot product of the remaining term of the integrand and the unit normal vector $\hat{n}$ (normal to the surface), as follows:

$$\int_{t}^{t+\Delta t}\frac{\partial}{\partial t}\int_{\Delta V_{\mathrm{P}}}(\rho c_{\mathrm{p}}T)\mathrm{d}V\mathrm{d}t+\int_{t}^{t+\Delta t}\int_{\Delta S}(\rho\vec{u}c_{\mathrm{p}}T-k\overline{\overline{\nabla T}})\cdot\hat{n}\mathrm{d}S\mathrm{d}t=\int_{t}^{t+\Delta t}\int_{\Delta V_{\mathrm{P}}}\overline{Q}_{\mathrm{gen}}\mathrm{d}V\mathrm{d}t \tag{5.3.2}$$

where $\hat{n}\,\mathrm{d}S$ represents the surface-area vector $\overline{\mathrm{d}\vec{S}}$.

Although this equation represents application of the conservation laws locally for each CV, the global conservation (for the complete domain) is built into this method. This is because the integral form of global conservation equations is obtained if we sum up this equation for all the CVs, since surface integrals over inner CV faces cancel out.

Approximations for Algebraic Formulation

Since the aim of the FVM is to obtain LAE for each CV, consisting of the cell-center temperature T_{P} and a number of neighboring cell-center values (T_{E}, T_{W}, etc.), certain approximations are needed for an algebraic representation of the integrals in Equation 5.3.2. They are broadly classified as the first and the second approximations, and are separately presented as follows.

First Approximation: Average Flux and Volumetric Term

The surface and volume integrals in Equation 5.3.2 need to be approximated using quadrature formulae. For the surface integral, it can be seen in the equation that the integrand is a flux term $\vec{f}$ (enthalpy and conduction flux), whereas, for the volume integral, the integrand is a volumetric term $\overline{\mathcal{V}}$ (volumetric enthalpy and heat generation). For both surface and volume integrals, the *I-approximation* is an approximate representation of the average value of the integrands—*the space-averaged one-point value for the spatial variation and the time-averaged one time-instant value for the temporal variation.*

For the spatial variation of the integrand, within the surface for the surface integral and inside the volume for the volume integral, the simplest one-point approximation is the midpoint rule—the space-averaged value of the integrand is approximated by its value at the centroid f of the surface for the flux term ($\vec{f}_\mathrm{f}$), and at the centroid P of the volume for the volumetric term ($\overline{\mathcal{V}}_\mathrm{P}$); the face centers f = e, w, n, and s and cell center P can be seen in Figure 5.3.1. Whereas, for the temporal variation in-between the two consecutive time instants (previous t and present $t+\Delta t$), there are three types on one-time-instant approximation—the time-averaged value of the flux is approximated by its value at previous time instant t for the explicit method, present time instant $t+\Delta t$ for the implicit method, and $t+\Delta t/2$ for the Crank–Nicolson method (follows the midpoint rule). Note that the three types of methods are considered only for the *flux* term $\vec{f}$ but not for the *volumetric* term $\overline{\mathcal{V}}$.

Thus, for a discrete representation of the integrands (general flux term $\vec{f}$ and volumetric term $\overline{\mathcal{V}}$), the *I*-approximation is given as

$$\begin{aligned}
\int_t^{t+\Delta t} \oint_{\Delta S} (\vec{f}\cdot\hat{n})\Delta S_\mathrm{f} &= \sum_{\mathrm{f=e,w,n,s}} \int_t^{t+\Delta t} \int_{\Delta S_\mathrm{f}} (\vec{f}\cdot\hat{n})\Delta S_\mathrm{f} \\
&= \sum_{\mathrm{f=e,w,n,s}} \left(\vec{f}_\mathrm{av}^{\,\mathrm{mean}}\cdot\hat{n}\right)\Delta S_\mathrm{f}\Delta t \approx \sum_{\mathrm{f=e,w,n,s}} (\vec{f}_\mathrm{f}^{\,\chi}\cdot\hat{n})\Delta S_\mathrm{f}\Delta t \qquad (5.3.3)\\
\int_{\Delta V_\mathrm{P}} \overline{\mathcal{V}}\,\mathrm{d}V &= \overline{\mathcal{V}}_\mathrm{av}\Delta V_\mathrm{P} \approx \overline{\mathcal{V}}_\mathrm{P}\Delta V_\mathrm{P}
\end{aligned}$$

where $\Delta S_{\mathrm{f=e,w,n,s}}$ is the area of the various surfaces of a CV, and ΔV_P is the volume of the CV; $\Delta V_\mathrm{P}=\Delta x_\mathrm{P}\Delta y_\mathrm{P}$, $\Delta S_\mathrm{e}=\Delta S_\mathrm{w}=\Delta y_\mathrm{P}$, and $\Delta S_\mathrm{n}=\Delta S_\mathrm{s}=\Delta x_\mathrm{P}$, for the representative Cartesian CV in Figure 5.3.1. For this equation, the subscript av represents the space-averaged value, and the superscript mean represents the time-averaged value. Note the approximate sign $\approx$ in this equation, where the average value av is represented by the centroid f = e, w, n, and s of the surfaces and the centroid P of a representative CV; the face centers f and cell center P can be seen in Figure 5.3.1 for the Cartesian CV, and Figure 5.3.2 for the complex CV. The mean value in Equation 5.3.3 is represented by the time level $\chi=n$ for the explicit method, $\chi=n+1$ for the implicit method, and $\chi=n+1/2$ for the Crank–Nicolson method. Note that $\vec{f}_\mathrm{f}^{\,n+1/2} \approx \left(\vec{f}_\mathrm{f}^{\,n}+\vec{f}_\mathrm{f}^{\,n+1}\right)\Big/2$ for Crank–Nicolson method. Also note from this equation that the time integral of the volumetric term is not considered in the *I*-approximation.

For the surface integral, this spatial approximation is second-order accurate, which can be preserved if the face-center value of the flux term $\vec{f}_\mathrm{f}$ is computed with at least second-order accuracy. For the volume integral, the spatial approximation becomes exact if the volumetric term $\overline{\mathcal{V}}$ is either constant or varies linearly within the CV; otherwise, it contains a second-order error. The time integral for this flux term is first-order accurate for the explicit and implicit method, and second order accurate for the Crank–Nicolson method.

Second Approximation: Interpolation or Extrapolation Method and Central Difference Method

For the flux term in energy equation, the integrand after the *I*-approximation in Equation 5.3.3 is $\vec{f}_{\mathrm{f}}^{\chi} \cdot \hat{n} = \vec{f}_{\eta,\mathrm{f}}^{\chi}$; where η is the outward normal direction at a surface. The integrand corresponds to *outward* enthalpy and conduction flux, at a face center f, given as

$$h_{\eta,\mathrm{f}} = \rho c_p u_{\eta,\mathrm{f}} T_{\mathrm{f}} = m_{\eta,\mathrm{f}} c_p T_{\mathrm{f}}$$

$$q_{\eta,\mathrm{f}} = -k\left(\frac{\partial T}{\partial \eta}\right)_{\mathrm{f}}$$

For the outward conduction flux $q_{\eta,\mathrm{f}}$ in the diffusion term, the *II-approximation* corresponds to FDM-based *central difference* (CD)-based representation for $(\partial T/\partial \eta)_{\mathrm{f}}$ (the normal gradient of temperature T at a face center f) in terms of the cell center values *across* the face. For the outward enthalpy flux $h_{\eta,\mathrm{f}}$ in the advection term, the *II-approximation* corresponds to an *interpolation or extrapolation method*, for the computation of face-center value of temperature T_{f} in terms of neighboring cell center values—called advection scheme.

For the Cartesian CV, shown in Figure 5.3.1, the h_η and q_η are h_x and q_x for the east surface, and $-h_x$ and $-q_x$ for the west surface; and h_y and q_y for the north surface, and $-h_y$ and $-q_y$ for the south surface. Note that η represent the outward normal to a surface, which is in the positive *x*/*y* direction for the fluxes at the east/north surface, and negative *x*/*y* direction for the fluxes at the west/south surface. The CD-based *II*-approximation for the normal gradient of temperature results in the conduction flux as

$$\begin{aligned} q_{x,\mathrm{e}}^{\chi} &\approx -k\frac{T_{\mathrm{E}}^{\chi} - T_{\mathrm{P}}^{\chi}}{\delta x_{\mathrm{e}}}, q_{x,\mathrm{w}}^{\chi} \approx -k\frac{T_{\mathrm{P}}^{\chi} - T_{\mathrm{W}}^{\chi}}{\delta x_{\mathrm{w}}} \\ q_{y,\mathrm{n}}^{\chi} &\approx -k\frac{T_{\mathrm{N}}^{\chi} - T_{\mathrm{P}}^{\chi}}{\delta y_{\mathrm{n}}}, q_{y,\mathrm{s}}^{\chi} \approx -k\frac{T_{\mathrm{P}}^{\chi} - T_{\mathrm{S}}^{\chi}}{\delta y_{\mathrm{s}}} \end{aligned} \tag{5.3.4}$$

where the δx_{e}, δx_{w}, δy_{n}, and δy_{s} are the distances between the cell centers *across* the e, w, n, and s face centers, respectively (refer Figure 5.3.1). Furthermore, for the Cartesian geometry, the enthalpy flux in the advection term is given as

$$h_{x,\mathrm{e}}^{\chi} = m_{x,\mathrm{e}} c_p T_{\mathrm{e}}^{\chi}, h_{x,\mathrm{w}}^{\chi} = m_{x,\mathrm{w}} c_p T_{\mathrm{w}}^{\chi}, h_{y,\mathrm{w}}^{\chi} = m_{y,\mathrm{n}} c_p T_{\mathrm{n}}^{\chi} \text{ and } h_{y,\mathrm{s}}^{\chi} = m_{y,\mathrm{s}} c_p T_{\mathrm{s}}^{\chi} \tag{5.3.5}$$

where the mass flux $m_x = \rho u$ and $m_y = \rho v$. Since the extrapolation or interpolation method (*II*-approximation) used to determine the temperature T_{f} in this equation depends on the direction of mass flux $m_{x/y,\mathrm{f}}$ (Patankar, 1980), the product of mass flux and temperature in the this equation is given (Sharma, 2016) as

$$\begin{aligned} & m_{x/y,\mathrm{f}} T_{\mathrm{f}} = m_{x/y,\mathrm{f}}^{+} T_{\mathrm{f}}^{+} + m_{x/y,\mathrm{f}}^{-} T_{\mathrm{f}}^{-} \Big\} \mathrm{f} = \mathrm{e,w,n,s} \\ & \text{where } T_{\mathrm{f}} = w_1 T_{\mathrm{D}} + w_2 T_{\mathrm{U}} \\ & \left.\begin{aligned} m_{x/y,\mathrm{f}}^{+} &\equiv \max(m_{x/y,\mathrm{f}}, 0) = \left|m_{x/y,\mathrm{f}}\right| \text{ if } m_{x/y,\mathrm{f}} > 0 \\ m_{x/y,\mathrm{f}}^{-} &\equiv \min(m_{x/y,\mathrm{f}}, 0) = -\left|m_{x/y,\mathrm{f}}\right| \text{ if } m_{x/y,\mathrm{f}} < 0 \end{aligned}\right\} \text{ otherwise } \quad 0 \end{aligned} \tag{5.3.6}$$

where $m_{x/y,\mathrm{f}}$ represents $m_{\mathrm{x,f}}$ at the east/west face center and $m_{\mathrm{y,f}}$ at the north/south face center. Furthermore, the temperature T_{f}^{+} is for the positive and T_{f}^{-} is for the negative *x*/*y* direction of the

fluid flow at a face center f. Also D is the downstream cell center and U is the upstream cell center: D=E and U=P if the direction of mass flux at the east face center e (refer Figure 5.3.1) is positive $m_{x,e} > 0$ and vice versa for the negative mass flux ($m_{x,e} < 0$). The various advection schemes can be found in Sharma (2016), with the weights $w_1=0$ and $w_2=1$ and $T_f^{\pm} = T_U$ (Equation 5.3.6) for the first-order-upwind (FOU) Scheme (constant extrapolation), that is, $T_e^{+} = T_P$ and $T_e^{-} = T_E$. Similarly, using the FOU scheme to determine the temperature at the other face centers, Equation 5.3.5 for the enthalpy fluxes is approximated as

$$\begin{aligned} h^{\chi}_{x,e} &\approx c_p(m^{+}_{x,e}T^{\chi}_{P} + m^{-}_{x,e}T^{\chi}_{E}), h^{\chi}_{x,w} \approx c_p\left(m^{+}_{x,w}T^{\chi}_{W} + m^{-}_{x,w}T^{\chi}_{P}\right) \\ h^{\chi}_{y,n} &\approx c_p(m^{+}_{y,n}T^{\chi}_{P} + m^{-}_{y,n}T^{\chi}_{N}) \;\;\&\;\; h^{\chi}_{y,s} \approx c_p\left(m^{+}_{y,s}T^{\chi}_{S} + m^{-}_{y,s}T^{\chi}_{P}\right) \end{aligned} \tag{5.3.7}$$

Thus, the I-approximation corresponds to the space as well as time averaged value of a flux term $\vec{f}^{\,\text{mean}}_{\text{av}}$, and space-averaged value of a volumetric term $\overline{\mathcal{V}}_{\text{av}}$; and the *II-approximation corresponds to obtaining the algebraic equation for the value or first derivative of a flow property at a face center, in terms of neighboring cell center values.*

FVM-Based Algebraic Formulation

The main idea of the FVM presented earlier for the energy equation and Cartesian CV is continued in this subsection for the Navier–Stokes (continuity, momentum and energy) equations and complex CV. Similar to the FVM for Cartesian geometry mentioned earlier, for a representative CV in an *arbitrary-shaped complex domain* of volume V bounded by a closed surface $\vec{S}$ (Figure 5.3.2), the FVM starts with the volume and the time integration of the continuity and the transport equations (Equation 5.3.1) as

$$\int_{t}^{t+\Delta t}\int_{\Delta V_P} \nabla\cdot(\rho\vec{u})\,\mathrm{d}V\,\mathrm{d}t = 0$$

$$\int_{t}^{t+\Delta t}\int_{\Delta V_P} \frac{\partial(\rho C\phi)}{\partial t}\mathrm{d}V\mathrm{d}t + \int_{t}^{t+\Delta t}\int_{\Delta V_P} \nabla\cdot(\rho\vec{u}C\phi - \Gamma_\phi\overrightarrow{\nabla\phi})\mathrm{d}V\mathrm{d}t = \int_{t}^{t+\Delta t}\int_{\Delta V_P} S_\phi\,\mathrm{d}V\,\mathrm{d}t$$

where $\phi = \vec{u}$ $C=1$, $\Gamma_\phi=\mu$ and $S_\phi=-\nabla p$ for the momentum equation, and $\phi=T$, $C=c_p$, $\Gamma_\phi=k$ and $S_\phi = \overline{Q}_{\text{gen}}$ for the energy equation. Here again, the integrand with the flux term are mass flux $\rho\vec{u}$, advection flux $\rho\vec{u}C\phi$ and diffusion flux $\Gamma_\phi\overrightarrow{\nabla\phi}$. As compared to enthalpy flux $\rho\vec{u}c_pT$ for the energy equation, the generic advection flux corresponds to x-momentum flux $\rho\vec{u}u$ for the x-momentum equation, and y-momentum flux $\rho\vec{u}v$ for the y-momentum equation. For the respective equation, the diffusion flux corresponds to viscous stress in the x- and y-direction; as compared to conduction flux for the energy equation. Using the Gauss divergence theorem for those integrands (in this integral equation) which are flux terms, we obtain

$$\begin{aligned} \int_{t}^{t+\Delta t}\oint_{\Delta S} (\rho\vec{u})\cdot\hat{n}\mathrm{d}S\mathrm{d}t &= 0 \\ \int_{t}^{t+\Delta t}\frac{\partial}{\partial t}\int_{\Delta V_P}\rho C\phi\,\mathrm{d}V\,\mathrm{d}t + \int_{t}^{t+\Delta t}\int_{\Delta S}(\rho\vec{u}C\phi - \Gamma_\phi\overrightarrow{\nabla\phi})\cdot\overrightarrow{\mathrm{d}S}\mathrm{d}t &= \int_{t}^{t+\Delta t}\int_{\Delta V_P} S_\phi\mathrm{d}V\mathrm{d}t \end{aligned} \tag{5.3.8}$$

For complex geometry, the algebraic formulation is initiated above with the coordinate-free form of the differential and integral form of the governing equations, Equations 5.3.1 and 5.3.8,

respectively. Thereafter, an FVM-based algebraic formulation is presented for the Cartesian components of velocity. Thus, for the *CVs of arbitrary/complex shape*, the *coordinate-free form* of the equations based FVM leads to the application of the momentum conservation law in the Cartesian coordinate directions: x-direction with $\varphi = u$, and y-direction with $\varphi = v$ in Equation 5.3.8. For the complex geometry problem, it is important to note that the Cartesian velocities (u and v) are obtained from the coordinate-free form, not the Cartesian form, of the momentum equation.

Applying the generic form of the I-approximation (Equation 5.3.3), an algebraic formulation of the continuity and transport equation (Equation 5.3.8) is given as

$$\sum_{\mathrm{f=e,w,n,s}} \rho\left(\vec{u}_{\mathrm{f}}^{\,n+1} \cdot \hat{n} \Delta S_{\mathrm{f}}\right) \Delta t = 0$$

$$\int_{t}^{t+\Delta t} \frac{\partial}{\partial t}(\rho C \phi_{\mathrm{P}} \Delta V_{\mathrm{P}})\mathrm{d}t + \sum_{\mathrm{f=e,w,n,s}} \left[\rho C\left(\vec{u}_{\mathrm{f}}^{\,n} \cdot \hat{n} \Delta S_{\mathrm{f}}\right)\phi_{\mathrm{f}}^{\chi} - \Gamma_{\phi}\left(\overrightarrow{\nabla \phi}_{\mathrm{f}}^{\chi} \cdot \hat{n} \Delta S_{\mathrm{f}}\right)\right] \Delta t \tag{5.3.9}$$

$$= \int_{t}^{t+\Delta t} \int_{\Delta V_{\mathrm{P}}} S_{\phi} \mathrm{d}V \mathrm{d}t$$

where ΔV_{P} is the volume of the CV, and $\Delta S_{\mathrm{f=e,\,w,\,n,\,s}}$ is the surface area at the various faces of the complex CV (Figure 5.3.2).

For the complex geometry problem, Sharma and Eswaran (2004) presented two coordinate system: global (x, y) and the local (s, t) coordinate. They are shown in Figure 5.3.3, with the local η-coordinate consisting of two nonorthogonal components: s-coordinate, along the line joining the cell centers across the surface, and t-coordinate, along the direction tangential to the surface. Note from the figure that the symbol η represents a coordinate which is normal to a surface. Since the direction of the η-coordinate varies from the one to the other surface of a complex CV, its definition is local to a surface. Thus, η-coordinate is called a local coordinate, and the Cartesian coordinate is called the global coordinate.

The Cartesian components ($\Delta S_{x,\mathrm{f}}$ and $\Delta S_{y,\mathrm{f}}$) of the surface area vector $\overrightarrow{\Delta S}_{\mathrm{f}}$ are shown in Figure 5.3.4a; the direction of the area vector corresponds to the outward normal at each surface. The figure also shows the expression for the computation of volume of the CV.

The 2D components of the vector quantities in Equation 5.3.9 are given as

$$\text{Global } x-y: \quad \vec{u}_{\mathrm{f}} = u_f + v_f; \qquad \overrightarrow{\Delta S}_{\mathrm{f}} = \Delta S_{x,\mathrm{f}} \hat{i} + \Delta S_{y,\mathrm{f}} \hat{j}$$

$$\text{Local } s-t: \quad \overrightarrow{\nabla \phi}_{\mathrm{f}} = \left(\frac{\partial \phi}{\partial s}\right)_{\mathrm{f}} \hat{s} + \left(\frac{\partial \phi}{\partial t}\right)_{\mathrm{f}} \hat{t}; \qquad \overrightarrow{\Delta S}_{\mathrm{f}} = \Delta S_{s,\mathrm{f}} \hat{s} + \Delta S_{t,\mathrm{f}} \hat{t}$$

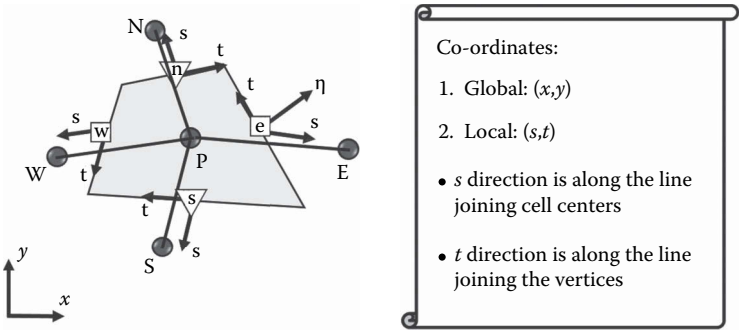


FIGURE 5.3.3 Schematic representation of the global and local coordinates, for a representative complex CV.

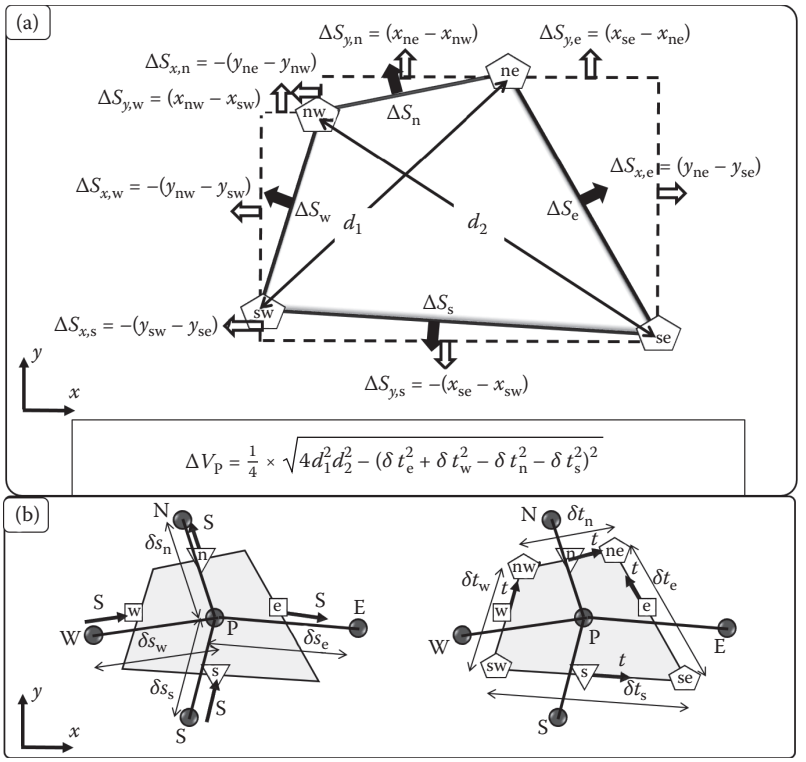


FIGURE 5.3.4 Schematic representation of the (a) surface areas $\Delta S_{x/y,f}$ and volume ΔV_P, and (b) distances between the cell centers δs_f and vertices δt_f, for a representative complex CV.

where the components for $\vec{u}_f$ and $\overrightarrow{\nabla\phi}_f$ are considered in the global x–y and local s–t coordinate system, respectively. Note that the global x–y coordinate system is considered for the computation of mass flow rate M and advection flow rate A (M and A obtained from the Cartesian velocity u/v), and local s –t coordinate is considered for the computation of diffusion D, given as

$$
\begin{aligned}
M_f^{n+1} &= \vec{m}_f^{n+1} \cdot \overrightarrow{\Delta S}_f = \rho\left(u_f^{n+1}\Delta S_{x,f} + v_f^{n+1}\Delta S_{y,f}\right) \equiv m_{\eta,f}^{n+1}\Delta S_f \\
A_{\phi,f}^{\chi} &= a_f^{\chi} \cdot \hat{n}\Delta S_f = Cm_{\eta,f}^{n}\phi_f^{\chi}\Delta S_f \equiv a_{\phi\eta f}^{\chi}\Delta S_f \\
D_{\phi,f}^{\chi} &= \vec{d}_f^{\chi} \cdot \overrightarrow{\Delta S}_f = \Gamma_\phi\left[\left(\frac{\partial\phi}{\partial s}\right)_f^{\chi}\hat{s} + \left(\frac{\partial\phi}{\partial t}\right)_f^{\chi}\hat{t}\right] \cdot \left(\Delta S_{s,f}\,\hat{s} + \Delta S_{t,f}\,\hat{t}\right) \\
&= \Gamma_\phi\left[\left(\frac{\partial\phi}{\partial s}\right)_f^{\chi}\Delta S_{s,f} + \left(\frac{\partial\phi}{\partial t}\right)_f^{\chi}\Delta S_{t,f}\right] \equiv d_{\phi\eta f}^{\chi}\Delta S_f
\end{aligned}
\tag{5.3.10}
$$

where $\vec{m}_f^{n+1} = \rho\left(\vec{u}_f^{n+1}\right)$ is the mass flux, $\vec{a}_f^{\chi} = \rho C\left(\vec{u}_f^{n}\right)\phi_f^{\chi}$ is the advection flux, and $\vec{d}_f^{\chi} = \Gamma_\phi\left(\overrightarrow{\nabla\phi}^{\chi}\right)$ is the diffusion flux. The component of respective flux in the direction η (normal to the surface) is represented as $m_{\eta,f}$, $a_{\varphi\eta,f}$, and $d_{\varphi\eta,f}$. The *II*-approximation for the mass flux corresponds to a momentum interpolation (Rhie and Chow, 1983) $m_{\eta,f}$, advection schemes for the

advection flux, and CD method for the diffusion flux. Using the two approximations, the final algebraic formulation from the continuity and transport equation are presented as follows.

Continuity Equation

After the *I*-approximation (Equation 5.3.9), the algebraic formulation for the mass source in the continuity equation (for incompressible flow) is given as

$$S_{m,P} = \sum_{\mathrm{f=e,w,n,s}} M_{\mathrm{f}}^{n+1}\Delta t = \sum_{\mathrm{f=e,w,n,s}} \rho\left(\vec{u}_{\mathrm{f}}^{n+1}\cdot\hat{n}\Delta S_{\mathrm{f}}\right)\Delta t = \sum_{\mathrm{f=e,w,n,s}} \rho u_{\eta,\mathrm{f}}^{n+1}\Delta S_{\mathrm{f}}\Delta t$$

where $u_{\eta,\mathrm{f}}$ is the outward normal velocity at the various face centers f = e, w, n, s of the complex CVs. The normal velocity is computed by a momentum interpolation method (Rhie and Chow, 1983), as the colocated (not staggered) grid is used for the complex geometry.

Transport Equations

After the *I*-approximation, the algebraic formulation for the transport equation is presented in Equation 5.39. Further application of the *II*-approximation and the resulting algebraic formulation are presented separately for the unsteady, advection, diffusion, and source terms.

Unsteady Term

The algebraic formulation for the unsteady term in Equation 5.3.9 is given as

$$\Delta V_{\phi} = \int_{t}^{t+\Delta t} \frac{\partial}{\partial t}(\rho C\phi_{\mathrm{P}}\Delta V_{\mathrm{P}})\,\mathrm{d}t \approx \rho C\Delta V_{\mathrm{P}}\frac{\phi_{\mathrm{P}}^{n+1}-\phi_{\mathrm{P}}^{n}}{\Delta t}\Delta t \tag{5.3.11}$$

where ΔV_{P} is the volume of the representative complex CV (Figure 5.3.4).

Advection Term

Advection scheme is used as the *II*-approximation, for the advection term in Equation 5.3.9. The scheme is presented for the complex geometry with the help of Figure 5.3.5 and Table 5.3.1. Advection scheme is an inward/outward direction of mass flux $m_{\eta,\mathrm{f}}$ based interpolation/extrapolation procedure. Thus, φ_{f} is expressed in terms of downstream D, upstream U, and upstream-of-upstream UU cell center values, represented by φ_{D}, φ_{U}, and φ_{UU}, respectively.

Since a face is common to two CVs, the advection scheme is presented in Table 5.3.1 for the north and east face center. *The negative value of total advection at the east face center e (of a CV P) is equal to that of the west face center* w *value of the adjoining east CV E*. The table shows a general expression for the product of the mass flux $m_{\eta,\mathrm{f}}$ and the advected variable φ_{f}, needed for computation of $A_{\phi,\mathrm{f}}^{x}$ in Equation 5.3.10. Note from the table that the equation for ϕ_{f}^{+} corresponds to positive/outward, and ϕ_{f}^{-} corresponds to negative/inward, direction of the fluid flow. This is demonstrated in the table for the east face center, where the equation for ϕ_{e}^{+} and ϕ_{e}^{-} is obtained after substituting appropriate subscript D, U, and UU in $\phi_{\mathrm{e,D}}^{\pm}$, $\phi_{\mathrm{e,U}}^{\pm}$, and $\phi_{\mathrm{e,UU}}^{\pm}$. They are also shown in Figure 5.3.5, for the FOU, CD, second-order upwind (SOU), and quadratic upwind interpolation for convective kinematics (QUICK) advection schemes. The figure and the table also show certain weights for the neighboring cell center values. Note that the equations for the weights in Table 5.3.1 correspond to a nonuniform volume of the complex CV-weighted interpolation/extrapolation procedure in the various advection schemes (Sharma and Eswaran, 2003).

Using the advection scheme, for the value of the variable φ in $a_{\varphi\eta,\mathrm{f=e,w,n,s}}$ (Equation 5.3.10), the *II*-approximation results in the algebraic formulation for the advection term in Equation 5.3.9 as

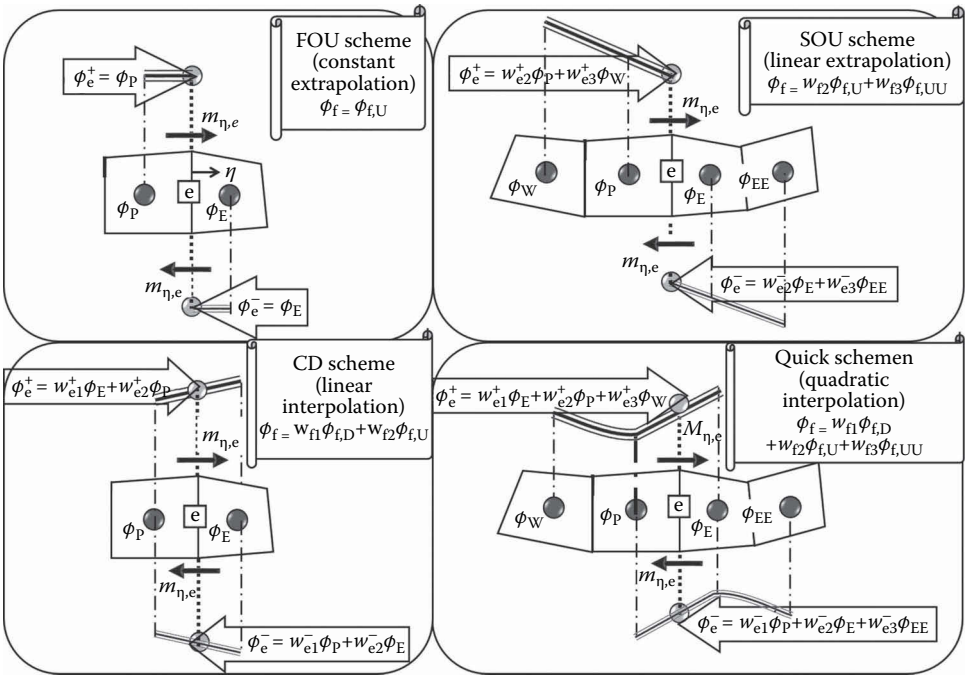


FIGURE 5.3.5 Schematic representation of various advection schemes to calculate the advected variable ϕ_f at the east face center, considering both positive/outward and negative/inward direction of the mass flux at the face. (From Sharma, A., *Introduction to Computational Fluid Dynamics: Development, Application and Analysis*, John Wiley & Sons, Chichester, UK, 2016. With permission.)

$$A_{\phi,P}^{\chi}\Delta t = \sum_{f=e,w,n,s} a_{\phi\eta,f}^{\chi}\Delta S_f \Delta t$$

$$\text{where } A_{\phi,P}^{\chi} \approx \left\{\left[m_{\eta,e}^{+,n}\phi_e^{+,\chi} + m_{\eta,e}^{-,n}\phi_e^{-,\chi}\right]\Delta S_e + \left[m_{\eta,w}^{+,n}\phi_w^{+,\chi} + m_{\eta,w}^{-,n}\phi_w^{-,\chi}\right]\Delta S_w\right\}C + \left\{\left[m_{\eta,n}^{+,n}\phi_n^{+,\chi} + m_{\eta,n}^{-,n}\phi_n^{-}\right]\Delta S_n + \left[m_{\eta,s}^{+,n}\phi_s^{+,\chi} + m_{\eta,s}^{-,n}\phi_s^{-,\chi}\right]\Delta S_s\right\}C \tag{5.3.12}$$

Diffusion Term

Using the CD method for the outward normal gradient of the variable ϕ in $d_{\phi\eta,f}^{\chi}\left[=\Gamma_\phi(\nabla\phi\cdot\hat{n})_f\right]$ (Equation 5.3.10), the *II*-approximation results in the algebraic formulation for the diffusion term in Equation 5.3.9 as

$$D_{\phi,P}^{\chi}\Delta t = \sum_{f=e,w,n,s} d_{\phi\eta,f}^{\chi}\Delta S_f \Delta t$$

$$D_{\phi,P}^{\chi} = \sum_{f=e,w,n,s} \Gamma_\phi(\nabla\phi^{\chi}\cdot\hat{n})_f \Delta S_f$$

$$\text{where } (\nabla\phi^{\chi}\cdot\hat{n})_e \Delta S_e \approx \frac{\phi_E^{\chi} - \phi_P^{\chi}}{\delta s_e}\left|\Delta S_{s,e}\right| + \frac{\phi_{ne}^{\chi} - \phi_{se}^{\chi}}{\delta t_e}\left|\Delta S_{t,e}\right| \tag{5.3.13}$$

$$(\nabla\phi^{\chi}\cdot\hat{n})_n \Delta S_n \approx \frac{\phi_N^{\chi} - \phi_P^{\chi}}{\delta s_n}\left|\Delta S_{s,n}\right| + \frac{\phi_{ne}^{\chi} - \phi_{nw}^{\chi}}{\delta t_w}\left|\Delta S_{t,n}\right|$$

$$(\nabla\phi^{\chi}\cdot\hat{n})_{\mathrm{w}}\Delta S_{\mathrm{w}} \approx \frac{\phi^{\chi}_{\mathrm{W}}-\phi^{\chi}_{\mathrm{P}}}{\delta s_{\mathrm{w}}}\left|\Delta S_{s,\mathrm{w}}\right| + \frac{\phi^{\chi}_{\mathrm{sw}}-\phi^{\chi}_{\mathrm{nw}}}{\delta t_{\mathrm{w}}}\left|\Delta S_{t,\mathrm{w}}\right|$$

$$(\nabla\phi^{\chi}\cdot\hat{n})_{\mathrm{s}}\Delta S_{\mathrm{s}} \approx \frac{\phi^{\chi}_{\mathrm{S}}-\phi^{\chi}_{\mathrm{P}}}{\delta s_{\mathrm{s}}}\left|\Delta S_{s,\mathrm{s}}\right| + \frac{\phi^{\chi}_{\mathrm{sw}}-\phi^{\chi}_{\mathrm{se}}}{\delta t_{\mathrm{s}}}\left|\Delta S_{t,\mathrm{s}}\right|$$

where the components of surface area $\Delta S_{s,\mathrm{f}}$ and $\Delta S_{t,\mathrm{f}}$ are represented in terms of their absolute values. Also δs_{f} is the distance between the adjoining cell centers, and δt_{f} is the distance between the associated vertices shown in Figure 5.3.4b. Furthermore, $\Delta S_{s,\mathrm{f=e,w,n,s}}$ and $\Delta S_{t,\mathrm{f=e,w,n,s}}$ are the local $s-t$ components of the surface area $\Delta S_{\mathrm{f=e,w,n,s}}$, calculated using the Cartesian components of the surface area. The Cartesian and local components of the surface area are correlated in terms of direction

TABLE 5.3.1
Advection Schemes for Complex Geometry (Sharma and Eswaran, 2003): Volume-Weighted Interpolation/Extrapolation-Based Equations for the Product of Mass $m_{\eta,\mathrm{f}}(\int \rho\vec{u}\,\boldsymbol{f}\cdot\hat{n})$ and Advected Variable ϕ_{f} at the East and North Face Centers

Advection scheme: $m_{\eta,\mathrm{f}}\phi_{\mathrm{f}} = m^{+}_{\eta,\mathrm{f}}\phi^{+}_{\mathrm{f}} + m^{-}_{\eta,\mathrm{f}}\phi^{-}_{\mathrm{f}}\}$ f = e, n

$$\text{where } m_{\eta,\mathrm{f}} = \vec{u}_{\mathrm{f}}\cdot\hat{n}, \quad \left.\begin{array}{l} m^{+}_{\eta,\mathrm{f}}\int \max(m_{\eta,\mathrm{f}},0) = +\left|m_{\eta,\mathrm{f}}\right| \text{ if } m_{\eta,\mathrm{f}} > 0 \\ m^{-}_{\eta,\mathrm{f}}\int \min(m_{\eta,\mathrm{f}},0) = -\left|m_{\eta,\mathrm{f}}\right| \text{ if } m_{\eta,\mathrm{f}} < 0 \end{array}\right\} \text{otherwise } 0$$

$$\text{and } \phi^{\pm}_{\mathrm{f}} = w^{\pm}_{\mathrm{f},1}\phi^{\pm}_{\mathrm{f,D}} + w^{\pm}_{\mathrm{f},2}\phi^{\pm}_{\mathrm{f,U}} + w^{\pm}_{\mathrm{f},3}\phi^{\pm}_{\mathrm{f,UU}}$$

$$\phi^{+}_{\mathrm{e}}(\text{for } m_{\eta,\mathrm{f}} > 0) = w^{+}_{\mathrm{e},1}\phi_{\mathrm{E}} + w^{+}_{\mathrm{e},2}\phi_{\mathrm{P}} + w^{+}_{\mathrm{e},3}\phi_{\mathrm{W}} \text{ and } \phi^{-}_{\mathrm{e}}(\text{for } m_{\eta,\mathrm{f}} < 0) = w^{-}_{\mathrm{e},1}T_{\mathrm{P}} + w^{-}_{\mathrm{e},2}\phi_{\mathrm{E}} + w^{-}_{\mathrm{e},3}\phi_{\mathrm{EE}}$$

$w_{\mathrm{f}}\downarrow$	FOU	CD	SOU	QUICK
$w^{+}_{\mathrm{e},1}$	0	$\dfrac{\Delta V_{\mathrm{E}}}{\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{P}}}$	0	$\dfrac{\Delta V_{\mathrm{P}}(2\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{W}})}{(\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{P}})\,(\Delta V_{\mathrm{E}}+2\Delta V_{\mathrm{P}})+\Delta V_{\mathrm{W}})}$
$w^{-}_{\mathrm{e},1}$	0	$\dfrac{\Delta V_{\mathrm{P}}}{\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{E}}}$	0	$\dfrac{\Delta V_{\mathrm{E}}(2\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{EE}})}{(\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{E}})\,(\Delta V_{\mathrm{P}}+2\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{EE}})}$
$w^{+}_{\mathrm{e},2}$	1	$\dfrac{\Delta V_{\mathrm{P}}}{\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{P}}}$	$\dfrac{2\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{W}}}{\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{W}}}$	$\dfrac{\Delta V_{\mathrm{E}}(2\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{W}})}{(\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{P}})\,(\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{W}})}$
$w^{-}_{\mathrm{e},2}$	1	$\dfrac{\Delta V_{\mathrm{E}}}{\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{E}}}$	$\dfrac{2\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{EE}}}{\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{EE}}}$	$\dfrac{\Delta V_{\mathrm{P}}(2\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{EE}})}{(\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{E}})\,(\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{EE}})}$
$w^{+}_{\mathrm{e},3}$	0	0	$-\dfrac{\Delta V_{\mathrm{P}}}{\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{W}}}$	$-\dfrac{\Delta V_{\mathrm{E}}\Delta V_{\mathrm{P}}}{(\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{W}})\,(\Delta V_{\mathrm{E}}+2\Delta V_{\mathrm{P}}+\Delta V_{\mathrm{W}})}$
$w^{-}_{\mathrm{e},3}$	0	0	$-\dfrac{\Delta V_{\mathrm{E}}}{\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{EE}}}$	$-\dfrac{\Delta V_{\mathrm{P}}\Delta V_{\mathrm{E}}}{(\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{EE}})\,(\Delta V_{\mathrm{P}}+2\Delta V_{\mathrm{E}}+\Delta V_{\mathrm{EE}})}$

Note: The equations are given for both positive/outward and negative/inward direction of the mass flux at the face centers. Subscript D, U, and UU correspond to the downstream, upstream, and upstream-of-upstream cell center of a face center f, respectively. The weights for the respective cell center value is represented as $w_{\mathrm{f},1}$, $w_{\mathrm{f},2}$, and $w_{\mathrm{f},3}$. Superscripts + and − correspond to $m_{\eta,\mathrm{f}} > 0$ and $m_{\eta,\mathrm{f}} < 0$, respectively. This equation for $\phi^{\pm}_{\mathrm{e}}$ and $w^{\pm}_{\mathrm{e}}$ is used to obtain similar equation for $\phi^{\pm}_{\mathrm{n}}$ and $w^{\pm}_{\mathrm{n}}$, by replacing EE, E, and W with NN, N, and S, respectively.

Source: From Sharma, A., *Introduction to Computational Fluid Dynamics: Development, Application and Analysis*, John Wiley & Sons, Chichester, UK, 2016.

cosines, shown in Figure 5.3.6—for the east surface of the representative CV. The direction cosines are expressed in terms of the Cartesian coordinates of the adjoining cell center and the vertices for the surface, such as the Cartesian coordinates of E, P, ne, and se used for the computation of direction cosines seen in Figure 5.3.6.

The diffusion is represented earlier (Equation 5.3.13) in terms of adjoining cell-centered diffusion and the cross diffusion. For momentum equation, this summation represents the net viscous forces acting on the CV. In case of energy equation, it represents a gain in the conduction heat-transfer rate by the fluid inside a CV. For this equation, the cell vertex values are approximated, in terms of the four adjoining cell center values (refer Figure 5.3.2), by *volume-weighted interpolation* (Sharma and Eswaran, 2003) as

$$\phi_{ne} = \frac{\Delta V_{NE}\phi_P + \Delta V_P\phi_{NE} + \Delta V_N\phi_E + \Delta V_E\phi_N}{\Delta V_{NE} + \Delta V_P + \Delta V_N + \Delta V_E}$$

$$\phi_{se} = \frac{\Delta V_{SE}\phi_P + \Delta V_P\phi_{SE} + \Delta V_S\phi_E + \Delta V_E\phi_S}{\Delta V_{SE} + \Delta V_P + \Delta V_S + \Delta V_E}$$

$$\phi_{nw} = \frac{\Delta V_{NW}\phi_P + \Delta V_P\phi_{NW} + \Delta V_N\phi_W + \Delta V_W\phi_N}{\Delta V_{NW} + \Delta V_P + \Delta V_N + \Delta V_W}$$

$$\phi_{sw} = \frac{\Delta V_{SW}\phi_P + \Delta V_P\phi_{SW} + \Delta V_S\phi_W + \Delta V_W\phi_S}{\Delta V_{SW} + \Delta V_P + \Delta V_S + \Delta V_W}$$

Pressure as a Source Term

Using the coordinate-free form of the momentum equation, the Cartesian velocity-based algebraic formulation in Equation 5.3.12 represents x- and y-momentum flow rates, for $\phi = u$ and $\phi = v$, respectively. The respective ϕ in Equation 5.3.13 results in viscous force in the x- and y-directions. The pressure forces in the x- and y-directions, for the determination of the u and v velocities from the coordinate-free form of the momentum equation, are presented as the source term $S_u = -\nabla p \cdot \hat{i}$ and $S_v = -\nabla p \cdot \hat{j}$, respectively.

The volume and time integral of the source term are approximated as

$$-\int_t^{t+\Delta t}\int_{\Delta V_P} \nabla p \cdot \hat{i}\, dV = -\int_t^{t+\Delta t}\int_{\Delta S_f} p\hat{i} \cdot \hat{n}\, dS$$

$$\approx -\sum_{f=e,w,n,s} p_f^{n+1}\Delta S_{x,f}\Delta t = S_{u,P}^{n+1}\Delta t$$

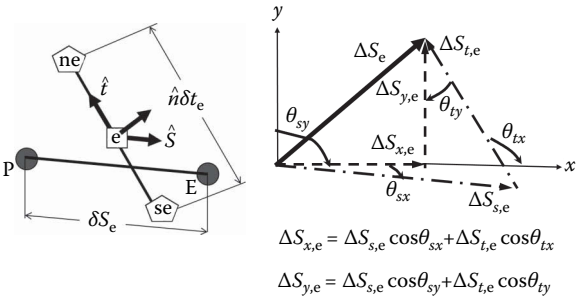


FIGURE 5.3.6 For the east face of representative complex CV, schematic representation of surface area vector and its Cartesian and local components. (From Sharma, A., *Introduction to Computational Fluid Dynamics: Development, Application and Analysis*, John Wiley & Sons, Chichester, UK, 2016. With permission.)

$$-\int_{t}^{t+\Delta t}\int_{\Delta V_{\mathrm{P}}} \nabla p\cdot\hat{j}\,\mathrm{d}V = -\int_{t}^{t+\Delta t}\int_{\Delta S_{\mathrm{f}}} p\hat{j}\cdot\hat{n}\,\mathrm{d}S$$
$$\approx -\sum_{\mathrm{f=e,w,n,s}} p_{\mathrm{f}}^{n+1}\Delta S_{y,\mathrm{f}}\Delta t = S_{v,\mathrm{P}}^{n+1}\Delta t \tag{5.3.14}$$

where Gauss divergence theorem is applied to the volume integral, as the pressure is a flux term. In this equation, the *I*-approximation is used to represent the surface-averaged pressure as p_{f} (at the face center f) and the time-averaged pressure as p^{n+1} (at present time level $(t+\Delta t)$). Furthermore, the *II*-approximation corresponds to the volume-weighted linear interpolation for p_{f} (similar to the CD method in Figure 5.3.5 and Table 5.3.1) given as

$$p_{\mathrm{f=e,w,n,s}} = \frac{\Delta V_{\mathrm{P}} p_{\mathrm{NB}} + \Delta V_{\mathrm{NB}} p_{\mathrm{P}}}{\Delta V_{\mathrm{P}} + \Delta V_{\mathrm{NB}}} \tag{5.3.15}$$

where NB = E, W, N and S for f = e, w, n, and s, respectively. Note that the Equation 5.3.14 represents the 2D Cartesian components of the total pressure force, acting on the CV.

All Terms

Substituting Equations 5.3.11 through 5.3.14 to Equation 5.3.8, the final algebraic formulation for 2D unsteady-state transport equation is given as

$$\frac{\Delta\mathcal{V}\phi}{\Delta t} + A_{\phi,\mathrm{P}}^{\chi} = D_{\phi,\mathrm{P}}^{\chi} + S_{\phi,\mathrm{P}}^{n+1}$$

$$\rho\frac{u_{\mathrm{P}}^{n+1}-u_{\mathrm{P}}^{n}}{\Delta t}\Delta V_{\mathrm{P}} + A_{u,\mathrm{P}}^{\chi} = D_{u,\mathrm{P}}^{\chi} - \sum_{\mathrm{f=e,w,n,s}} p_{\mathrm{f}}^{n+1}\Delta S_{\mathrm{fx}}$$

$$\rho\frac{v_{\mathrm{P}}^{n+1}-v_{\mathrm{P}}^{n}}{\Delta t}\Delta V_{\mathrm{P}} + A_{v,\mathrm{P}}^{\chi} = D_{v,\mathrm{P}}^{\chi} - \sum_{\mathrm{f=e,w,n,s}} p_{\mathrm{f}}^{n+1}\Delta S_{\mathrm{fy}}$$

$$\rho c_P\frac{T_{\mathrm{P}}^{n+1}-T_{\mathrm{P}}^{n}}{\Delta t}\Delta V_{\mathrm{P}} + A_{\mathrm{T,P}}^{\chi} = D_{\mathrm{T,P}}^{\chi} + \overline{Q}_{\mathrm{gen,P}}\Delta V_{\mathrm{P}}$$

where

$$A_{\phi,\mathrm{P}}^{\chi} = \left\{\left[m_{\eta,\mathrm{e}}^{+,n}\phi_{\mathrm{e}}^{+,\chi} + m_{\eta,\mathrm{e}}^{-,n}\phi_{\mathrm{e}}^{-,\chi}\right]\Delta S_{\mathrm{e}} + \left[m_{\eta,\mathrm{w}}^{+,n}\phi_{\mathrm{w}}^{+,\chi} + m_{\eta,\mathrm{s}}^{-,n}\phi_{\mathrm{w}}^{-,\chi}\right]\Delta S_{\mathrm{S}}\right\}C$$
$$+\left\{\left[m_{\eta,\mathrm{n}}^{+,n}\phi_{\mathrm{n}}^{+,\chi} + m_{\eta,\mathrm{n}}^{-,n}\phi_{\mathrm{n}}^{-}\right]\Delta S_{\mathrm{n}} + \left[m_{\eta,\mathrm{s}}^{+,n}\phi_{\mathrm{s}}^{+,\chi} + m_{\eta,\mathrm{s}}^{-,n}\phi_{\mathrm{s}}^{-,\chi}\right]\Delta S_{\mathrm{S}}\right\}C$$

$$D_{\phi,\mathrm{P}}^{\chi} = \Gamma_{\phi}\left(\frac{\phi_{\mathrm{E}}^{\chi}-\phi_{\mathrm{P}}^{\chi}}{\delta s_{\mathrm{e}}}\left|\Delta S_{s,\mathrm{e}}\right| + \frac{\phi_{\mathrm{W}}^{\chi}-\phi_{\mathrm{P}}^{\chi}}{\delta s_{\mathrm{W}}}\left|\Delta S_{s,\mathrm{w}}\right| + \frac{\phi_{\mathrm{N}}^{\chi}-\phi_{\mathrm{P}}^{\chi}}{\delta s_{\mathrm{n}}}\left|\Delta S_{s,\mathrm{n}}\right| + \frac{\phi_{\mathrm{S}}^{\chi}-\phi_{\mathrm{P}}^{\chi}}{\delta s_{\mathrm{S}}}\left|\Delta S_{s,\mathrm{s}}\right|\right)$$
$$+\Gamma_{\phi}\left(\frac{\phi_{\mathrm{ne}}^{\chi}-\phi_{\mathrm{se}}^{\chi}}{\delta t_{\mathrm{e}}}\left|\Delta S_{s,\mathrm{e}}\right| + \frac{\phi_{\mathrm{sw}}^{\chi}-\phi_{\mathrm{nw}}^{\chi}}{\delta t_{\mathrm{W}}}\left|\Delta S_{s,\mathrm{w}}\right| + \frac{\phi_{\mathrm{ne}}^{\chi}-\phi_{\mathrm{nw}}^{\chi}}{\delta t_{\mathrm{n}}}\left|\Delta S_{s,\mathrm{n}}\right| + \frac{\phi_{\mathrm{sw}}^{\chi}-\phi_{\mathrm{se}}^{\chi}}{\delta t_{\mathrm{S}}}\left|\Delta S_{s,\mathrm{s}}\right|\right) \tag{5.3.16}$$

where the transport variable at a face center $\phi_{\mathrm{f=e,w,n,s}}^{\pm}$ is presented in Table 5.3.1, and the volume-weighted pressure at the various face center p_{f} is given in Equation 5.3.15.

Concluding Remarks

FVM-based algebraic formulation is presented, for a 2D unsteady fluid dynamics and forced-convection heat transfer in a complex geometry—starting with the grid generation, and the differential and integral form of the governing equations; followed by the two approximations for the integrands in the integral equations; and ending with the resulting FVM-based approximated algebraic formulation. However, starting from the mass, momentum, and energy conservation laws instead of their differential formulation and using the two approximations (presented here), the same algebraic formulation (Equation 5.3.16) was presented recently by Sharma (2016) – called as physical law based FVM.

Due to the fact that there is no explicit equation for the pressure in an incompressible flow, the algebraic equation for the continuity equation is further converted into an equation for the pressure correction—using a pressure correction method (Sharma and Eswaran, 2003). Thereafter, the coupled system of LAEs for the velocity, pressure, and temperature are solved using a semi-explicit method (Sharma and Eswaran, 2003) or semi-implicit method (Patankar, 1980). The FVM-based solution methodology can be found in the book by Patankar (1980), Ferziger and Peric (1995), and Versteeg and Malalasekera (1995).

REFERENCES

Ferziger, J.H. and M. Peric. 2002. *Computational Methods for Fluid Dynamics*. Berlin: Springer.

Muralidhar, K. and G. Biswas. 2005. *Advanced Fluid Mechanics*, 2nd edition. New Delhi: Narosa Publishing House.

Muralidhar, K. and T. Sundararajan, eds. 2003. *Computational Fluid Flow and Heat Transfer*, 2nd edition. New Delhi: IIT Kanpur Series, Narosa Publishing House.

Patankar, S.V. 1980. *Numerical Heat Transfer and Fluid Flow*. New York: Hemisphere Publishing Corporation.

Rhie, C.M. and W.L. Chow. 1983. A numerical study of the turbulent flow past an isolated airfoil with trailing edge separation. *AIAA Journal*. 21: 1525–1532.

Sharma, A. and V. Eswaran. 2003. The finite volume method, Chapter 12 in *Computational Fluid Flow and Heat Transfer* (eds.: Muralidhar, K., and T. Sundarajan). New Delhi: IIT Kanpur Series, Narosa Publishing House.Thompson, J.F., Z.U.A. Warsi, and C.W. Mastin. 1985. *Numerical Grid Generation: Foundations and Applications*. Amsterdam: North-Holland.

Sharma, A. 2016. *Introduction to Computational Fluid Dynamics: Development, Application and Analysis*, John Wiley & Sons, Chichester, UK.

Versteeg, H.K. and W. Malalasekera. 1995. *An Introduction to Computational Fluid Dynamics: The Finite Volume Method*. Harlow: Longman Scientific & Technical.

5.4 FINITE ELEMENT METHOD

Salil S. Kulkarni

Introduction

The finite-element method (FEM) is a numerical technique for finding approximate solutions to a wide variety of boundary value problems in engineering and science. It is probably the most popular and commonly used numerical technique when compared to other techniques like the finite-difference method, finite-volume method, and the boundary element method. This popularity of the FEM stems from its applicability to solve a wide variety of problems in solid mechanics (see, e.g., Zienkiewicz et al., 2014a), fluid dynamics (see, e.g., Zienkiewicz et al., 2014b), heat transfer (see, e.g., Gartling and Reddy, 2010), electromagnetism (see, e.g., Jin, 2014), biomechanics (see, e.g., Oomens et al., 2010), acoustics (see, e.g., Ihlenburg, 2006), etc. It is completely general with respect to the material properties and can easily handle complicated geometries subject to a variety of boundary conditions. It is applicable in multiphysics problems where there is a coupling between two or more different field variables, e.g., thermomechanical problems. The mathematical theory of the FEM well matured (see, e.g., Brenner and Scott, 2008) and the conditions under which the approximate solution approaches the actual solution, especially for linear problems, are well understood. On account of the various attributes of the FEM, which are mentioned earlier, it is widely used as an analysis tool in various fields of engineering, viz., automotive engineering, aerospace engineering, Microelectromechanical systems (MEMS) industry, biomedical engineering, electrical engineering, and civil engineering.

Papers that discuss the application of the FEM for solving heat-transfer problems first appeared in the 1960s (see, e.g., Zienkiewicz and Cheung, 1965; Wilson and Nickell, 1966). Since then, the field has fairly matured and the application of the finite element to solving problems in heat transfer is now a standard topic in many books on the FEM (see, e.g., Bathe, 1996; Hughes, 2000; Gartling and Reddy, 2010; Zienkiewicz et al., 2013). The present chapter provides a brief overview of the FEM when applied to solving heat conduction problems in stationary domains and is based on the books mentioned earlier. A brief outline of the topics discussed in the chapter is given next.

The main idea of the finite element is described in Section 5.4.2. The governing differential equations for heat conduction problems in stationary domains are discussed in Section 5.4.3. Using a simple problem, the steps involved in the derivation of the finite-element equations are described in detail in Section 5.4.4. The semidiscrete form of the finite-element equations for a general problem in transient heat conduction is given in Section 5.4.5. The linear equation solvers used in the finite equations are mentioned in Section 5.4.6. The numerical evaluation of the finite-element matrices using isoparametric elements and Gauss quadrature is discussed in Section 5.4.7. Some practical issues when using the FEM are discussed in Section 5.4.8. Finally, two example problems are solved using the FEM and are presented in Section 5.4.9.

Main Idea of the FEM

The main idea of the FEM is now explained using a simple example. Let the true temperature distribution in a rod of length L be $T(x)$ and it is as shown in Figure 5.4.1.

It is desired to find an approximation, $T_h(x)$, to the true temperature distribution, $T(x)$. In general, it is extremely difficult, if not impossible, to find a function that matches $T(x)$ everywhere over the domain $[0, L]$. The main idea of the finite element is to divide the domain $[0, L]$ into smaller nonoverlapping subdomains (elements) with the exact function being approximated with simple functions, $T_h^e(x)$, over each of the elements. Then, the approximation over the entire domain is given by $T_h(x) = \sum_{e=1}^{N} T_h^e(x)$, where N is the number of elements. For example, in Figure 5.4.1, the rod

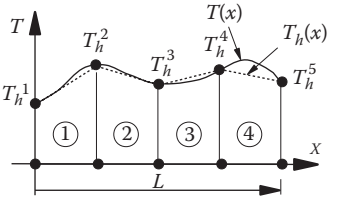


FIGURE 5.4.1 Main idea of the FEM.

is discretized into four elements. In each of the elements, the unknown function is approximated with simple (linear) functions. Then, the approximation $T_h(x)$ for the current problem is given by $\sum_{e=1}^{4} T_h^e(x)$. The basic features of the FEM can be summarized as follows:

1. The domain of the given boundary value problem is represented by a nonoverlapping collection of simple subdomains called finite elements. Each element consists of nodes, faces (in three dimensions), and edges (in two and three dimensions). This process is commonly referred to as discretization. The collection of the finite elements is called a finite-element mesh. In one-dimensional domains, the elements can be either straight or curved, in two-dimensional domains, the most common elements are either triangular or quadrilateral in shape with straight or curved edges, and in three-dimensional domains, the elements are either tetrahedral or hexahedral in shape with either straight or curved edges.
2. Over each element, the field variable, e.g., temperature, is approximated by functions of the desired type (mainly polynomials), and algebraic equations relating the physical quantities at the nodes are developed using the governing differential equations. For example, for heat conduction problems, these equations relate the temperature at the nodes to the heat flux on the element boundary.
3. The algebraic equations that are written for each element are now assembled using continuity and balance of the physical quantities. For example, for heat conduction problems, if two elements share the same edge, then the value of the temperature at the common nodes along the common edge is same for both the elements. In addition, in absence of a heat source along the edge, the outward heat flux of both the elements along the common edge is equal and opposite to each other.
4. Appropriate boundary conditions are applied and then the equations are solved for the unknown nodal quantities. The Dirichlet-type boundary conditions can be incorporated using a Lagrange multiplier approach, Nitsche's method, a penalty approach, or an elimination approach. Note that if one is dealing with a transient problem, the unknown nodal quantities are the function of time and then one has to solve a coupled set of ordinary different equations.
5. Once the equations are solved, quantities of interest, e.g., temperatures, flux, etc., are extracted, visualized, and analyzed in a process referred to as postprocessing.

Sources of Error in FEM

As mentioned earlier, the FEM is a numerical method that provides one with an approximate solution to the problem at hand. It is, therefore, essential to be aware of the sources of errors in the finite-element solution. The main sources of errors in the FEM are summarized next:

1. The division of the domain into simple classical finite elements may not be exact. For example, an attempt to represent a curved boundary with elements having straight edges leads to an error in the representation of the domain.

2. While deriving the algebraic equations for individual elements, it is usually assumed that the unknown field quantity can be represented using simple polynomials. It is possible that the actual solution may not be represented exactly using polynomials, e.g., trignometric or exponential functions.
3. The algebraic equations are frequently established using numerical integration. As the integrations may not be exact, this leads to the errors in the finite-element solution. Note that it may be possible to use symbolic computations to eliminate this error.
4. The imposition of the Dirichlet-type boundary conditions may lead to an error depending on the approach adopted to impose the boundary condition. For example, if one uses the penalty method to impose the boundary conditions, the boundary conditions are not imposed exactly and the system matrix may become ill-conditioned.
5. Additional errors may occur depending on the algorithms used for the solution of the algebraic equations.

In the next section, the governing differential equation for heat conduction in a stationary domain along with the associated initial and boundary conditions is summarized.

Governing Differential Equation

Consider a stationary three-dimensional body, Ω, with surface Γ whose temperature distribution is of interest. The surface Γ is subdivided into two disjoint surfaces, Γ_T and Γ_q such that on Γ_T the temperature is prescribed, while on Γ_q the flux is prescribed. It is assumed that the initial temperature distribution, $T_0(\mathbf{x})$, is also known. To find the subsequent temperature distribution in the body, one then solves the following (see, e.g., Bergman et al., 2011):

$$\rho C_p \frac{\partial T}{\partial t} = \nabla\cdot(\mathbf{k}\nabla T) + f \text{ in } \Omega, \tag{5.4.1}$$

subject to the following boundary conditions and initial conditions:

$$T = \overline{T}(\mathbf{x},t) \text{ on } \Gamma_T, \tag{5.4.2}$$

$$q_n = -\mathbf{n}\cdot(\mathbf{k}\nabla T) = h \text{ on } \Gamma_q, \tag{5.4.3}$$

$$T = T_0(\mathbf{x}) \quad \text{at } t = 0 \quad \text{in } \Omega. \tag{5.4.4}$$

Here, ρ is the mass density, C_p is the specific heat, T is the temperature, $\mathbf{k}$ is the conductivity tensor, f is a time, position and temperature-dependent heat source, $\mathbf{x}$ is a point in the body, $\mathbf{n}$ is an outward unit normal to the surface, and ∇ is the gradient operator. Both C_p and $\mathbf{k}$ can be functions of position and temperature. The prescribed flux boundary condition could be one of the conditions listed here or can be a combination of two or more conditions.

1. Adiabatic boundary condition: $h=0$.
2. Prescribed flux: $h = \hat{q}(\mathbf{x},t)$. Here, the prescribed flux is a function of both position and time.
3. Convection boundary condition. $h=h_c(\mathbf{x}, t)(T(\mathbf{x}, t)-T_B(\mathbf{x}, t))$. Here, $h_c(\mathbf{x}, t)$ is the known position and time-dependent convective heat-transfer coefficient, and $T_B(\mathbf{x}, t)$ is the known position and time-dependent bulk temperature.

4. Radiation boundary condition. $h = h_r'(\mathbf{x},t)\left(T^4(\mathbf{x},t) - T_S^4(\mathbf{x},t)\right)$. Here, $h_r'(\mathbf{x},t)$ is the known time-dependent radiation heat-transfer coefficient, and $T_S(\mathbf{x}, t)$ is the known time-dependent temperature of the surroundings. The radiation boundary condition is usually rewritten as

$$h = h_r(T,\mathbf{x},t)\left(T(\mathbf{x},t) - T_S(\mathbf{x},t)\right),$$

where

$$h_r(T,\mathbf{x},t) = h_r'(\mathbf{x},t)\left(T^2(\mathbf{x},t) - T_S^2(\mathbf{x},t)\right)\left(T(\mathbf{x},t) + T_S(\mathbf{x},t)\right).$$

Note that the radiation boundary condition is a nonlinear boundary condition.

The general transient heat conduction problem in a three-dimensional body is governed by Equations 5.4.1 through 5.4.4. Depending on the specific problem under consideration, one can identify it with one of the four types of problems mentioned in Table 5.4.1.

Note that a presence of a heat source that has nonlinear dependence on temperature will lead to a nonlinear problem.

Steps in the Derivation of the Finite-Element Equations

For most problems, it is not possible to solve Equations 5.4.1 through 5.4.4 exactly as the analytical solution is not available. This lack of availability of the analytical solution could be due to one or some of the following reasons: (1) complicated geometry, (2) complicated boundary conditions, and (3) complicated nature of the differential equation. In such cases, one of the numerical methods, viz., the FEM, the finite-difference method, etc., has to be used to obtain the solution. To clearly explain the steps followed in the derivation of the finite-element equations from the governing partial differential equation, one specific problem is chosen and all the necessary details of the derivation are provided. The derivation of the finite-element equations for other types of problems can be approached similarly.

There are multiple methods to derive the finite-element equations, viz., the direct method, method based on the variational approach, and the method based on the weighted residual approach (see, e.g., Zienkiewicz et al., 2013). The derivation of the finite-element equations presented in this section closely follows the approach presented in Reddy (2005). The approach is based on the application method of weighted residue to a single element, followed by the use of the divergence theorem and the Galerkin approach. This approach makes it possible to derive the finite-element equations even when a variational approach is not possible. The derivation of the elemental finite-element equations for a transient two-dimensional heat conduction in a stationary, homogeneous, isotropic media with temperature independent material properties and time- and position-dependent internal heat source is given next.

The governing partial differential equation for the problem under consideration is of the form (see Equation 5.4.1):

TABLE 5.4.1
Classification of Heat Conduction Problems

Type of Problem	Time Dependence	Radiation Boundary Condition AND/OR Temperature-Dependent Material Property
Linear steady state	X	X
Nonlinear steady state	X	✓
Linear transient	✓	X
Nonlinear transient	✓	✓

$$\rho C_p \frac{\partial T}{\partial t} = k\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) + f(x,y,t), \text{ in } \Omega. \tag{5.4.5}$$

It is assumed that on part of the boundary, Γ_T, the temperature is prescribed while on the remaining part of the boundary, and Γ_q the flux is prescribed. The initial temperature distribution, $T_0(x, y)$ is also assumed to be known. The boundary conditions and initial conditions are, therefore, written as

$$T(x,y,t) = \overline{T}(x,y,t) \text{ on } \Gamma_T \quad \text{and} \quad q_n(x,y,t) = \overline{q}(x,y,t) \text{ on } \Gamma_q, \quad T(x,y,0) = T_0(x,y) \text{ in } \Omega.$$

Figure 5.4.2 shows the domain under consideration, which is discretized using triangular finite elements. The finite-element model consists of *Nel* number of elements and *N* of nodes. Consider a typical element that is denoted by Ω^e and whose boundary is denoted by Γ^e.

The exact temperature distribution in the element, $T(x, y, t)$, is approximated using separation of variables as

$$T(x,y,t) \approx T_h^e(x,y,t) = \sum_{j=1}^{n} N_j(x,y) T_j^e(t), \tag{5.4.6}$$

where n is the number of nodes in the element, $T_j^e(t), j = 1,\ldots,n$ are the time-dependent unknown nodal temperatures and $N_j(x, y)$, $j = 1, \ldots, n$ are known as time-independent shape functions. The shape functions dictate the nature of the approximation. Because $T_h^e(x,y,t)$ is not exact, the substitution of Equation 5.4.6 in Equation 5.4.5 leads to a residual of the form:

$$R(x,y,t,T_1^e,\ldots,T_n^e) = \rho C_p \frac{\partial T_h^e}{\partial t} - k\left(\frac{\partial^2 T_h^e}{\partial x^2} + \frac{\partial^2 T_h^e}{\partial y^2}\right) - f(x,y,t) \neq 0. \tag{5.4.7}$$

Note that the residual defined in Equation 5.4.7 is a function of the location, time, and the nodal temperatures. The goal is now to find the nodal temperatures that minimize the residual is some averaged sense. The residual is first multiplied by a weighting function $w(x, y)$, followed by integrating the product over the element and then setting the integral to zero. This leads to**

$$\begin{aligned} &\int_{\Omega^e} w(x,y) R(x,y,t,T_1^e,\ldots,T_n^e)\, d\Omega^e = \\ &\int_{\Omega^e} w(x,y)\rho C_p \frac{\partial T_h^e}{\partial t} d\Omega^e - \int_{\Omega^e} w(x,y) k\left(\frac{\partial^2 T_h^e}{\partial x^2} + \frac{\partial^2 T_h^e}{\partial y^2}\right) d\Omega^e \\ &- \int_{\Omega^e} w(x,y) f(x,y,t)\, d\Omega^e = 0. \end{aligned} \tag{5.4.8}$$

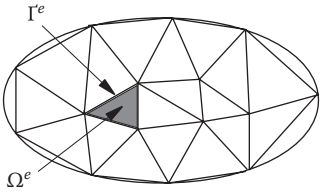


FIGURE 5.4.2 Discretization of the domain.

Equation 5.4.8 is commonly referred to as the statement of the weighted residual. Note that if one proceeds to directly evaluate Equation 5.4.8, one requires that $T_h^e(x,y,t)$ should possess continuous first derivatives with respect to both x and y. In order to reduce the continuity requirements on the approximation function, one uses the Gauss divergence theorem to obtain:

$$\int_{\Omega^e} w(x,y)\rho C_p \frac{\partial T_h^e}{\partial t} \mathrm{d}\Omega^e + \int_{\Omega^e} k\left(\frac{\partial T_h^e}{\partial x}\frac{\partial w}{\partial x} + \frac{\partial T_h^e}{\partial y}\frac{\partial w}{\partial y}\right)\mathrm{d}\Omega^e = \int_{\Omega^e} w(x,y) f(x,y,t)\ \mathrm{d}\Omega^e - \int_{\Gamma^e} w(x,y) q_n\ \mathrm{d}\Gamma^e, \tag{5.4.9}$$

where $q_n = -k\,\partial T_h^e/\partial n$. One observes that due to the use of the Gauss divergence theorem one has indeed reduced the continuity required on the approximation $T_h^e(x,y,t)$, which is now required to be only continuous. In order to obtain the required number of algebraic equations from Equation 5.4.9, in this case n, for the time derivative of the unknown nodal temperatures $T_i^e(t), i=1,\ldots,n$, one chooses n different weighting functions. The choice of $w(x, y)=N_i(x, y)$, $i=1, \ldots, n$, in Equation 5.4.9, is referred to as the Bubnov–Galerkin approximation or just Galerkin approximation. Using the Galerkin approximation and substituting $T_h^e(x,y,t) = \sum_{j=1}^{n} N_j(x,y)T_j^e(t)$ in Equation 5.4.9 leads to the following set of n coupled first-order differential equations for $T_j^e, i=1,\ldots,n$.

$$\sum_{j=1}^{n}\int_{\Omega^e} \rho C_p N_i N_j\ \mathrm{d}\Omega^e \frac{dT_j^e}{dt} + \sum_{j=1}^{n}\int_{\Omega^e} k\left(\frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial y}\frac{\partial N_j}{\partial y}\right)\mathrm{d}\Omega^e T_j^e = \int_{\Omega^e} N_i f(x,y,t)\ \mathrm{d}\Omega^e + \int_{\Gamma^e} N_i q_n(x,y,t)\ \mathrm{d}\Gamma^e, \quad i=1,\ldots,n. \tag{5.4.10}$$

Equation 5.4.10 can be rewritten in a vector form as

$$\mathbf{C}^e\dot{\mathbf{T}}^e + \mathbf{K}^e\mathbf{T}^e = \mathbf{f}^e(t) + \mathbf{q}^e(t), \tag{5.4.11}$$

where $\mathbf{C}^e$ and $\mathbf{K}^e$ and $n\times n$ matrix and $\mathbf{T}^e$, $\mathbf{f}^e$, and $\mathbf{q}^e$ are $n\times 1$ vectors. The matrix $\mathbf{C}^e$ is referred to as the elemental capacity matrix, the matrix $\mathbf{K}^e$ is referred to as the elemental stiffness matrix, the vector $\mathbf{T}^e$ is the vector of the unknown nodal temperatures, the vector $\mathbf{f}^e(t)$ is the vector due to the source term, and the vector $\mathbf{q}^e(t)$ is the vector of the nodal fluxes. The components of $\mathbf{C}^e$, $\mathbf{K}^e$, $\mathbf{f}^e$, and $\mathbf{q}^e$ are given by

$$C_{ij}^e = \int_{\Omega^e} \rho C_p N_i N_j\ \mathrm{d}\Omega^e, \qquad K_{ij}^e = \int_{\Omega^e} k\left(\frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial y}\frac{\partial N_j}{\partial y}\right)\mathrm{d}\Omega^e,$$
$$f_i^e = \int_{\Omega^e} N_i f(x,y,t)\ \mathrm{d}\Omega^e, \qquad q_i^e = \int_{\Gamma^e} N_i q_n(x,y,t)\ \mathrm{d}\Gamma^e. \tag{5.4.12}$$

The numerical evaluation of the elemental matrices and vectors is discussed in Section 5.4.7.1. Note that q_i^e given in Equation 5.4.12 is evaluated only on that part of the boundary of the body on which the flux is prescribed. This is due to the fact that the contributions to the flux vectors at the shared nodes of the common internal edge of the two adjacent elements cancel each other. This cancellation is because of the balance of flux across the shared internal edge.

From Equation 5.4.12, it is observed' that $\mathbf{C}^e=(\mathbf{C}^e)^T$ and $\mathbf{K}^e=(\mathbf{K}^e)^T$, i.e., the elemental capacity matrix and the elemental stiffness matrix are both symmetric matrices. The matrix $\mathbf{C}^e$ as defined in

Equation 5.4.12 is referred to as the consistent capacitance matrix. In practice, from the point of view of computational efficiency, one usually uses a diagonal form of the capacitance matrix. Various methods for obtaining the diagonal matrix are available in literature (see, e.g., Zienkiewicz et al., 2013). On assembly of the elemental equations using the continuity of the temperatures at the nodes and balance of flux across the internal shared element boundaries, one obtains the global set of equations as

$$\mathbf{C}\dot{\mathbf{T}} + \mathbf{K}\mathbf{T} = \mathbf{F}(t). \tag{5.4.13}$$

where

$$\mathbf{C} = \mathop{\mathbf{A}}_{e=1}^{Nel} \mathbf{C}^e, \mathbf{K} = \mathop{\mathbf{A}}_{e=1}^{Nel} \mathbf{K}^e, \mathbf{F}(t) = \mathbf{f}(t) + \mathbf{q}(t), \mathbf{f}(t) = \mathop{\mathbf{A}}_{e=1}^{Nel} \mathbf{f}^e(t), \mathbf{q}(t) = \mathop{\mathbf{A}}_{e=1}^{Nel} \mathbf{q}^e(t).$$

Note that $\dot{\mathbf{T}} = d\mathbf{T}/dt$. Here $\mathop{\mathbf{A}}_{e=1}^{Nel}$ is the assembly operator and *Nel* is the total number of elements in the model. Matrices **C** and **K** are referred to as the global capacitance matrix and the global stiffness matrix, respectively. Due to the fact that the elemental capacity matrix and the elemental stiffness matrix are both symmetric, the global capacity matrix and the global stiffness matrix are also symmetric, i.e., $\mathbf{C} = \mathbf{C}^T$ and $\mathbf{K} = \mathbf{K}^T$. Vector $\mathbf{F}(t)$ is referred to as the global load vector. Note that **C** and **K** are $N \times N$ matrices and $\mathbf{T}(t)$ and $\mathbf{F}(t)$ are $N \times 1$ vectors, where N is the total degrees of freedom in the model. For the present problem, the number of degrees of freedom is equal to the number of nodes. Equation 5.4.13 is referred to the semidiscrete finite-element model as it is obtained by discretizing the partial difference equation in only the spatial dimensions. Once the global set of equations is obtained (see Equation 5.4.13), the specified temperature boundary conditions are applied at appropriate nodes. Some of the methods used to apply the temperature boundary conditions include the elimination method, the Lagrange multiplier method, and the penalty method (see, e.g., Bathe, 1996 for details). Once the boundary conditions are applied, the resulting set of ordinary differential equations is solved to find the unknown nodal quantities.

The assembly process to form the global matrices and vectors from the elemental matrices and vectors is illustrated with a simple example shown in Figure 5.4.3.

The figure shows a discretized rectangular plate whose steady-state temperature distribution is to be found out. It is assumed that there is a distributed time independent heat source in the plate and the plate is subjected to some prescribed temperature and prescribed flux boundary conditions. As seen from the figure, the plate is discretized with two straight edge triangular elements, 1 and 2, which share a common edge. Element 1 has nodes 1–2–3, whereas element 2 has nodes 1–3–4. This node numbering is referred to as global node numbering. The three nodes of each element are also labeled 1–2–3 and are shown boxed in Figure 5.4.3. These numbers are referred to as local node numbers. Note that the elemental matrices and vectors are calculated as per the local node numbering. To aid the assembly process, an element connectivity table is formed as shown in Table 5.4.2.

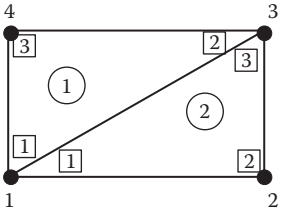


FIGURE 5.4.3 Assembly of elements.

TABLE 5.4.2
Element Connectivity for Problem Shown in Figure 5.4.3

Element No.	Local Node Numbers		
	Node 1	Node 2	Node 3
1	1	3	4
2	1	2	3

This table helps in relating the local node numbers of the element with the global node numbers. The elemental equation for element 1 in expanded form is given by (see Equation 5.4.11)

$$\begin{array}{c} \\ 1 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{ccc} 1 & 3 & 4 \end{array} \\ \begin{bmatrix} k_{11}^1 & k_{12}^1 & k_{13}^1 \\ & k_{22}^1 & k_{23}^1 \\ \text{symm} & & k_{33}^1 \end{bmatrix} \end{array} \begin{bmatrix} T_1^1 \\ T_2^1 \\ T_3^1 \end{bmatrix} = \begin{array}{c} 1 \\ 3 \\ 4 \end{array}\begin{bmatrix} f_1^1 \\ f_2^1 \\ f_3^1 \end{bmatrix} + \begin{array}{c} 1 \\ 3 \\ 4 \end{array}\begin{bmatrix} q_1^1 \\ q_2^1 \\ q_3^1 \end{bmatrix},$$

where the components of the elemental matrices and the vectors are calculated using Equation 5.4.12. Here, the subscripts denote the local node numbers, and the superscript denotes the element number. Note that the components of the flux vector, i.e., q_1^1, q_2^1, and q_3^1, have to be evaluated over only those edges of the element that are on the part of the boundary of the plate where the flux is prescribed. Similarly, the elemental equation for element 2 is written as

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} k_{11}^2 & k_{12}^2 & k_{13}^2 \\ & k_{22}^2 & k_{23}^2 \\ \text{symm} & & k_{33}^2 \end{bmatrix} \end{array} \begin{bmatrix} T_1^2 \\ T_2^2 \\ T_3^2 \end{bmatrix} = \begin{array}{c} 1 \\ 2 \\ 3 \end{array}\begin{bmatrix} f_1^2 \\ f_2^2 \\ f_3^2 \end{bmatrix} + \begin{array}{c} 1 \\ 2 \\ 3 \end{array}\begin{bmatrix} q_1^2 \\ q_2^2 \\ q_3^2 \end{bmatrix}.$$

Here, again the components of the flux vector, i.e., q_1^2, q_2^2, and q_3^2, have to be evaluated over only those edges of the element that are on the part of the boundary of the plate where the flux is prescribed. Noting that $T_1^1 = T_1^2 = T_1, T_2^1 = T_3^2 = T_3, T_3^1 = T_4$, and $T_2^2 = T_2$, one then assembles the two sets of the elemental equations to obtain the global set of equations as follows:

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} k_{11}^1 + k_{11}^2 & k_{12}^2 & k_{12}^1 + k_{13}^2 & k_{13}^1 \\ & k_{22}^2 & k_{23}^2 & 0 \\ & & k_{22}^1 + k_{33}^1 & k_{23}^1 \\ & \text{symm} & & k_{33}^1 \end{bmatrix} \end{array} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{bmatrix} = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{bmatrix} f_1^1 + f_1^2 \\ f_2^2 \\ f_2^1 + f_3^2 \\ f_3^1 \end{bmatrix} + \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{bmatrix} q_1^1 + q_1^2 \\ q_2^2 \\ q_2^1 + q_3^2 \\ q_3^1 \end{bmatrix}.$$

The matrix occurring in the above equation is referred to as the global stiffness matrix, while the vectors appearing on the right side are referred to as the global load vectors. For more details about the assembly procedure, see, e.g., Reddy (2005). The global capacitance matrix that occurs in transient problems is also assembled in a similar manner.

Semidiscrete Finite-Element Model in a General Case

In this section, the semidiscrete finite-element model arising in the solution of the problem of transient heat conduction in a three-dimensional body with position and temperature-dependent material properties and time-dependent boundary conditions (see Equations 5.4.1 through 5.4.4) is presented. To develop the semidiscrete finite-element model in the general case, one essentially follows the same steps as described in Section 5.4.4. One proceeds by assuming an approximation for the temperature in a typical element of the form given in Equation 5.4.6. Following the steps outlined in Section 5.4.4, i.e., using the method of weighted residual followed by the use of Gauss divergence theorem, and finally the Galerkin method, one obtains the elemental equations as

$$\mathbf{C}^e(\mathbf{T}^e)\dot{\mathbf{T}}^e + \hat{\mathbf{K}}^e(t,\mathbf{T}^e)\mathbf{T}^e = \mathbf{F}^e(t,\mathbf{T}^e), \tag{5.4.14}$$

where

$$\hat{\mathbf{K}}^e(t,\mathbf{T}^e) = \mathbf{K}^e(\mathbf{T}^e) + \mathbf{H}^e(t) + \mathbf{R}^e(t,\mathbf{T}^e) \text{ and } \mathbf{F}^e(t,\mathbf{T}^e) = \mathbf{f}^e(t,\mathbf{T}^e) + \mathbf{q}^e(t) + \mathbf{h}_c^e(t) + \mathbf{h}_r^e(t,\mathbf{T}^e).$$

Here, $\mathbf{C}^e(\mathbf{T}^e)$ and $\mathbf{K}^e(\mathbf{T}^e)$ are the temperature-dependent elemental capacitance and stiffness matrix, respectively, and the elemental vector $\mathbf{f}^e(t, \mathbf{T}^e)$ is due to the source term. The elemental matrix $\mathbf{H}^e(t)$ and the vector $\mathbf{h}_c^e(t)$ are related to the convective boundary condition, while the elemental matrix $\mathbf{R}^e(t, \mathbf{T}^e)$ and the vector $\mathbf{h}_r^e(t,\mathbf{T}^e)$ are related to the radiation boundary condition. The components of the domain integrals $\mathbf{C}^e(\mathbf{T}^e)$, $\mathbf{K}^e(\mathbf{T}^e)$, and $\mathbf{f}^e(t, \mathbf{T}^e)$ are given by

$$C_{ij}^e = \int_{\Omega^e} \rho C_p N_i N_j \, \mathrm{d}\Omega^e, K_{ij}^e = \int_{\Omega^e} \left(\frac{\partial N_i}{\partial x} k_x \frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial y} k_y \frac{\partial N_j}{\partial y} + \frac{\partial N_i}{\partial z} k_z \frac{\partial N_j}{\partial z} \right) \mathrm{d}\Omega^e,$$

$$f_i^e = \int_{\Omega^e} N_i f \, \mathrm{d}\Omega^e.$$

The components of boundary integrals $\mathbf{H}^e(t)$, $\mathbf{R}^e(t, \mathbf{T}^e)$, $\mathbf{h}_c^e(t)$, $\mathbf{h}_r^e(t,\mathbf{T}^e)$, and $\mathbf{q}^e(t)$ are given by

$$H_{ij}^e = \int_{\Gamma^e} h_c N_i N_j \, \mathrm{d}\Gamma^e, R_{ij}^e = \int_{\Gamma^e} h_r N_i N_j \, \mathrm{d}\Gamma^e, \tag{5.4.15}$$

and

$$h_{ci}^e = \int_{\Gamma^e} h_c N_i T_B \, \mathrm{d}\Gamma^e, \; h_{ri}^e = \int_{\Gamma^e} h_r N_i T_S \, \mathrm{d}\Gamma^e, \; q_i^e = \int_{\Gamma^e} N_i q_n \, \mathrm{d}\Gamma^e. \tag{5.4.16}$$

These integrals are to be evaluated on those portions of the boundary where the prescribed flux boundary conditions are given. Observe that $\mathbf{C}^e(\mathbf{T}^e) = (\mathbf{C}^e(\mathbf{T}^e))^T$, $\mathbf{K}^e(\mathbf{T}^e) = (\mathbf{K}^e(\mathbf{T}^e))^T$, $\mathbf{H}^e(t) = (\mathbf{H}^e(t))^T$, and $\mathbf{R}^e(t, \mathbf{T}^e) = (\mathbf{R}^e(t, \mathbf{T}^e))^T$. The numerical evaluation of the elemental matrices is discussed in Section 5.4.7.1. Note that one of the simplest approach to evaluate the coefficients of the various temperature-dependent matrices and vectors appearing in Equation 5.4.14 is to assume that the temperature is constant in the element and is equal to the average of the nodal temperatures. The material properties corresponding to this temperature are then used in the numerical evaluation of the integrals. A more accurate but computationally intensive approach is to evaluate the temperature at the Gauss points (see Section 5.4.7.3) using the element shape functions and then use it to evaluate the material properties.

On assembly of the elemental equations (see Equation 5.4.14), one obtains the global system of equations as

$$\mathbf{C}(\mathbf{T})\dot{\mathbf{T}} + \hat{\mathbf{K}}(t,\mathbf{T})\mathbf{T} = \mathbf{F}(t,\mathbf{T}), \tag{5.4.17}$$

where

$$\hat{\mathbf{K}}(t,\mathbf{T}) = \mathbf{K}(\mathbf{T}) + \mathbf{H}(t) + \mathbf{R}(t,\mathbf{T}) \text{ and } \mathbf{F}(t,\mathbf{T}) = \mathbf{f}(t,\mathbf{T}) + \mathbf{q}(t) + \mathbf{h}_c(t) + \mathbf{h}_r(t,\mathbf{T}).$$

Observe that due to the symmetry of the elemental capacity matrix and the elemental stiffness matrix, one has $\mathbf{C}(\mathbf{T}) = (\mathbf{C}(\mathbf{T}))^T$ and $\hat{\mathbf{K}}(t,\mathbf{T}) = (\hat{\mathbf{K}}(t,\mathbf{T}))^T$. Here, $\mathbf{C}(\mathbf{T})$ and $\hat{\mathbf{K}}(t,\mathbf{T})$ are $N \times N$ matrices and $\mathbf{T}$ and $\mathbf{F}(t, \mathbf{T})$ are $N \times 1$ vectors, where N is the total degrees of freedom in the model. For the present problem, the number of degrees of freedom are equal to the number of nodes. Equation 5.4.17, which is a set of coupled nonlinear first-order ordinary differential equations, is the semidiscrete finite-element model of the problem defined by Equations 5.4.1 through 5.4.4. Based on the problem classification given in Table 5.4.1, one can write a general algebraic or a semidiscrete form of the finite-element equations for each type of problem as follows:

1. Linear steady-state problem

$$\hat{\mathbf{K}}\mathbf{T} = \mathbf{F}, \tag{5.4.18}$$

   where

$$\hat{\mathbf{K}} = \mathbf{K} + \mathbf{H} \text{ and } \mathbf{F} = \mathbf{f} + \mathbf{q} + \mathbf{h}_c.$$

2. Nonlinear steady-state problem

$$\hat{\mathbf{K}}(\mathbf{T})\mathbf{T} = \mathbf{F}(\mathbf{T}), \tag{5.4.19}$$

   where

$$\hat{\mathbf{K}}(\mathbf{T}) = \mathbf{K}(\mathbf{T}) + \mathbf{H} + \mathbf{R}(\mathbf{T}) \text{ and } \mathbf{F}(\mathbf{T}) = \mathbf{f}(\mathbf{T}) + \mathbf{q} + \mathbf{h}_c + \mathbf{h}_r(\mathbf{T}).$$

3. Linear transient problem

$$\mathbf{C}\dot{\mathbf{T}} + \hat{\mathbf{K}}(t)\mathbf{T} = \mathbf{F}(t), \tag{5.4.20}$$

   where

$$\hat{\mathbf{K}}(t) = \mathbf{K} + \mathbf{H}(t) \text{ and } \mathbf{F}(t) = \mathbf{f}(t) + \mathbf{q}(t) + \mathbf{h}_c(t).$$

4. Nonlinear transient problem. This is the same as Equation 5.4.17 and is repeated here for completeness.

$$\mathbf{C}(\mathbf{T})\dot{\mathbf{T}} + \hat{\mathbf{K}}(t,\mathbf{T})\mathbf{T} = \mathbf{F}(t,\mathbf{T}),$$

   where

$$\hat{\mathbf{K}}(t,\mathbf{T}) = \mathbf{K}(\mathbf{T}) + \mathbf{H}(t) + \mathbf{R}(t,\mathbf{T}) \text{ and } \mathbf{F}(t,\mathbf{T}) = \mathbf{f}(t,\mathbf{T}) + \mathbf{q}(t) + \mathbf{h}_c(t) + \mathbf{h}_r(t,\mathbf{T}).$$

Solution Phase

Once the set of algebraic equations (see Equation 5.4.18 or Equation 5.4.19) or the semidiscrete equations (see Equation 5.4.205.4.17**) are assembled, the temperature boundary conditions are applied using, for example, the penalty method (see, e.g., Bathe, 1996 for details). A brief description of the

penalty method when it is used in linear steady-state problems is given next. Let i denote the degree of freedom, which is prescribed, i.e., $T_i = \overline{T}_i$. Then, a large constant α is added to the diagonal entry $\hat{K}_{ii}$ of the $\hat{\mathbf{K}}$ matrix and ith entry of the vector $\mathbf{F}$ is modified to $F_i + \alpha\overline{T}_i$. As a result of these modifications, the ith equation is changed to

$$(\hat{K}_{ii} + \alpha)T_i + \sum_{\substack{j=1 \\ j\neq i}}^{N} \hat{K}_{ij}T_j = F_i + \alpha\overline{T}_i.$$

Note that as α is chosen to be a very large number, the solution to the equation given earlier yields $T_i \approx \overline{T}_i$ as desired. After the boundary conditions are incorporated, the equations are then solved to obtain the solution. In case the problem is of the type given by Equation 5.4.18, the linear algebraic equations can be directly solved using the methods discussed in Section 5.4.6.1. If the problem is of type given by Equation 5.4.20, one needs to use a time integration algorithm to obtain a set of linear algebraic equations from the set of linear ordinary differential equations. Some of the most commonly used algorithms for carrying out the time integration are described in Section 5.4.6.2. If the problem is of the type given by Equation 5.4.19, then one can use a nonlinear equation solver of the type described in Section 5.4.6.3 to find the solution. If the problem is of type given by Equation 5.4.17, then one needs to combine the solvers described in Sections 5.4.6.2 and 5.4.6.3. One particular solver to solve Equation 5.4.17 is described in Section 5.4.6.4.

Linear Equation Solver

The solution of equations of the form, represented by

$$\mathbf{Ax} = \mathbf{b}, \tag{5.4.21}$$

is required when solving all the four types of problems mentioned in Section 5.4.5. Here, $\mathbf{A}$ is a $N\times N$ coefficient matrix, $\mathbf{b}$ is a $N\times 1$ right-hand side vector and $\mathbf{x}$ is a $N\times 1$ unknown vector. N depends on the degrees of freedom in the model, which for most problems in heat conduction is equivalent to the number of nodes in the model. This one-to-one equivalence between the degrees of freedom and the number of unknown hold good when the only nodal unknown is the temperature. For industrial problems, N can range from a few hundred thousand to a few millions and hence the solution of finite-element equations can require a huge computational effort. For example, if each entry of $\mathbf{A}$ is a double precision floating point number, then the storage of the full matrix $\mathbf{A}$ requires $8N^2$ bytes. If N= 100,000, a relatively small number in the present day simulations, then one needs 80 GB to store the matrix. In addition, the computational efforts (floating point operations) required to solve these systems can range from $O(N^2)$ to $O(N^3)$ depending on the type of algorithms used. Fortunately, matrix $\mathbf{A}$ has a special structure in the context of the FEM, which can be exploited to reduce the computational burden. The coefficient matrix is typically sparse and symmetric. The sparseness is due to the fact that one uses a local approximation to the unknown field variable. The matrix $\mathbf{A}$ is symmetric in those cases when one deals with self-adjoint equations and one uses the Galerkin method. The sparseness and symmetric nature of the matrix $\mathbf{A}$ translates into tremendous saving in the memory requirement as the full matrix does not have to be stored. Special data structures, which avoid the storing the numerically zero entries, are used to store the entries of the $\mathbf{A}$. The types of algorithms available in literature to solve Equation 5.4.21 arising in the context of a finite-element analysis can be broadly classified as follows:

1. *Direct solvers*: These solvers directly solve the system of equations without the need of an iterative process. They involve the following phases: (1) decomposition and (2) solution. The decomposition phase is based on some special implementation of Cholesky, QR, and LU decompositions (see, e.g., Golub and Van Loan, 2012), which exploits the sparsity of the matrix. This involves the use of graph theory to minimize the both the storage requirement

and computational effort. The decomposition phase is then followed by the solution phase which involves forward elimination followed by back substitution. More details of these methods is founds in Davis (2006). A detailed comparison of these algorithms is available (Gould and Scott, 2007). In addition to these solvers, frontal solvers (see, e.g., Irons, 1970) that are specifically tailored to exploit the assembly process of the FEM and avoid the assembly of the coefficient matrix **A** are also commonly used.

2. *Iterative solvers.* In these solvers, one starts with an initial guess $\mathbf{x}_0$ to the solution of Equation 5.4.21 and generates a sequence of approximations $\mathbf{x}_i$, which converges to the exact solution of Equation 5.4.21. The main advantage of these method is that they do not need any decompositions and primarily involve matrix vector multiplications. For symmetric positive definite matrices, one can use the conjugate gradient method (see Hestenes and Stiefel, 1952), whereas for a general matrix one can use the Generalized minimal residual method (GMRES) method (see Saad and Schultz, 1986). The rate of convergence of an iterative method is heavily influenced by the choice of the preconditioners. For symmetric positive definite matrices, some of the preconditioners that can be used include incomplete Cholesky and diagonal scaling. For a general matrix, the preconditioners that are commonly used are based on the incomplete LU factorizations. More details of the iterative methods are found in Saad (2003). The detailed survey on the types of preconditioners for large systems is given in Benzi (2002).

Time Integration Algorithm

To carry out the time integration from $t=0$ to $t=T$, one first discretizes the time interval [0, T] into Nt time intervals. Let t_1, t_2, …, $t_{Nt}=T$ denote the discrete instances where the temperature is to be obtained. Assuming that one knows the temperature up to t_n, one uses a time integration algorithm to obtain the temperature at t_{n+1}. The most commonly used algorithms for numerically integrating Equation 5.4.20 belong to the family of the generalized trapezoidal methods. The generalized trapezoidal methods are described by the following equations (see, e.g., Hughes, 2000):

$$\mathbf{C}_{n+1}\dot{\mathbf{T}}_{n+1} + \hat{\mathbf{K}}_{n+1}\mathbf{T}_{n+1} = \mathbf{F}_{n+1}, \tag{5.4.22}$$

$$\mathbf{T}_{n+1} = \mathbf{T}_n + \Delta t\dot{\mathbf{T}}_{n+\alpha}, \tag{5.4.23}$$

$$\dot{\mathbf{T}}_{n+\alpha} = (1-\alpha)\dot{\mathbf{T}}_n + \alpha\dot{\mathbf{T}}_{n+1}, \tag{5.4.24}$$

where $\mathbf{T}_n$, $\dot{\mathbf{T}}_n$, $\mathbf{T}_{n+1}$, and $\mathbf{T}_{n+1}$ are approximation to $\mathbf{T}(t_{n+1})$, $\dot{\mathbf{T}}(t_{n+1})$, $\mathbf{T}(t_{n+1})$, and $\dot{\mathbf{T}}(t_{n+1})$, respectively, $\mathbf{F}_{n+1}=\mathbf{F}(t_{n+1})$, $\hat{\mathbf{K}}_{n+1}=\hat{\mathbf{K}}(t_{n+1})$ and $\Delta t=t_{n+1}-t_n$. One particular implementation of the generalized trapezoidal methods leads to the following equation for $\mathbf{T}_{n+1}$, the unknown temperature at t_{n+1}, given the temperature $\mathbf{T}_n$ at t_n:

$$\left(\mathbf{C}+\alpha\Delta t\hat{\mathbf{K}}\right)\mathbf{T}_{n+1} = \left(\mathbf{C}-(1-\alpha)\Delta t\hat{\mathbf{K}}\right)\mathbf{T}_n + \Delta t\left((1-\alpha)\mathbf{F}_n + \mathbf{F}_{n+1}\right). \tag{5.4.25}$$

Here, α is a user-defined parameter with $0\le\alpha\le1$, which decides the accuracy and the stability of the algorithm and $\mathbf{F}_n=\mathbf{F}(t_n)$. For different values of α one obtains different algorithms, viz. $\alpha=0$ corresponds to the Forward Euler method, $\alpha=1$ corresponds to the Backward Euler method, $\alpha=1/2$ corresponds to the trapezoidal method (mid-point rule, Crank–Nicolson scheme), and $\alpha=2/3$ corresponds to the Galerkin method. The comparison of the four methods with respect to nature of the algorithm (explicit/implicit) and stability and accuracy is given in Table 5.4.3.

Nonlinear Equation Solver

The most commonly used technique to solve the problem represented by Equation 5.4.19 is the Newton–Raphson technique (see, e.g., Bathe, 1996). The basic idea of the method is to start with

TABLE 5.4.3
Comparison of the Different Time Integration Schemes

α	Method	Explicit/Implicit	Stability	Accuracy
0	Forward Euler	Explicit[a]	Conditionally stable	First order
1	Backward Euler	Implicit	Unconditionally stable	First order
1/2	Trapezoidal	Implicit	Unconditionally stable	Second order
2/3	Galerkin	Implicit	Unconditionally stable	Second order

[a] Diagonal capacitance matrix.

an initial guess of the solution and successively improve the estimate by using the first-order Taylor series expansion of the nonlinear function. To aid convergence, it is a usual practice to apply the total load in increments. The procedure to obtain the solution, $\mathbf{T}_{n+1}$ at the $(n+1)$th load increment given the converged solution, $\mathbf{T}_n$ at the nth increment using the Newton–Raphson technique is explained next. Let $\mathbf{T}_{n+1}^i$ denote the temperature at the ith iteration at load step $(n+1)$. From Equation 5.4.19, it is observed that the residue at the ith iteration, $\mathbf{R}_{n+1}^i$ is given by

$$\mathbf{R}_{n+1}^i = \hat{\mathbf{K}}_{n+1}^i \mathbf{T}_{n+1}^i - \mathbf{F}_{n+1}^i, \tag{5.4.26}$$

where $\hat{\mathbf{K}}_{n+1}^i = \hat{\mathbf{K}}(\mathbf{T}_{n+1}^i)$ and $\mathbf{F}_{n+1}^i = \mathbf{F}(\mathbf{T}_{n+1}^i)$. Using Taylor series expansion, the first-order approximation of the residue at the $i+1$th iteration is written as

$$\mathbf{R}_{n+1}^{i+1} \approx \mathbf{R}_{n+1}^i + (\mathbf{K}_T)_{n+1}^i \Delta\mathbf{T}_{n+1}^{i+1}, \tag{5.4.27}$$

where $\mathbf{K}_T = \frac{\partial \mathbf{R}}{\partial \mathbf{T}}$ is referred to as the tangent stiffness matrix. Note that at the element level, the components of a $n \times n$ tangent stiffness matrix are given by

$$R_{ab}^e = \frac{\partial R_a^e}{\partial T_b^e} = \frac{\partial}{\partial T_b^e}\left(\sum_{k=1}^{n} \hat{K}_{ak}^e T_k^e - F_a^e\right) = \hat{K}_{ab}^e + \sum_{k=1}^{n} \frac{\partial \hat{K}_{ak}^e}{\partial T_b^e} T_k^e - \frac{\partial F_a^e}{\partial T_b^e}, \quad a,b = 1,\ldots,n. \tag{5.4.28}$$

Setting the residue at the $i+1$th iterate to zero, the increment in the temperature, $\Delta\mathbf{T}_{n+1}^{i+1}$, is obtained by solving

$$(\mathbf{K}_T)_{n+1}^i \Delta\mathbf{T}_{n+1}^{i+1} = -\mathbf{R}_{n+1}^i. \tag{5.4.29}$$

Then, the temperature at the $i+1$th iterate is obtained as

$$\mathbf{T}_{n+1}^{i+1} = \mathbf{T}_{n+1}^i + \Delta\mathbf{T}_{n+1}^{i+1}. \tag{5.4.30}$$

This procedure is continued till the norm of the residue is less than some prescribed user tolerance and the norm of the increment is also less than some user prescribed tolerance. Once the converged solution is obtained, the load or the forcing is again incremented. The procedure is repeated till the entire load/forcing is applied.

The Newton–Raphson method shows a quadratic convergence rate near the true solution. As a result, it converges quickly to the true solution provided one starts sufficiently close to the true solution. The disadvantage of the method is the requirement to calculate the tangent stiffness matrix (see Equation 5.4.29) at every iteration. In addition, as the tangent stiffness matrix is not symmetric, the highly efficient symmetric linear equations solvers cannot be used. To avoid these drawbacks, a number of modifications have been suggested to the Newton–Raphson method.

These modifications include symmetric linearization (Winget and Hughes, 1985), Modified Newton method (see, e.g., Bathe, 1996), wherein the tangent stiffness matrix is calculated only at the beginning of the load increment and the quasi-Newton methods (see, e.g., Bathe, 1996), which the tangent stiffness matrix is approximated with a secant matrix whose approximation can be computed efficiently.

Transient Nonlinear Equation Solver

To carry out the time integration from $t=0$ to $t=T$, one first discretizes the time interval $[0, T]$ into Nt time intervals. Let $t_1, t_2, \ldots, t_{Nt}=T$ denote the discrete instances where the temperature is to be obtained. A solver obtained by combing the Backward Euler method (see Section 5.4.6.2) with the Newton–Raphson method (see Section 5.4.6.3) to obtain the temperature at t_i, $i=1, \ldots, Nt$ is briefly described next. One assumes that the temperature up to time t_n is known and that the temperature at time t_{n+1} is to be obtained. Next, using the time discrete form of Equation 5.4.17 coupled with the Backward Euler method, the nonlinear equation for the unknown temperature $\mathbf{T}_{n+1}$ at t_{n+1} is written as follows:

$$\mathbf{C}_{n+1}\left(\mathbf{T}_{n+1}-\mathbf{T}_n\right)=\Delta t\left(\mathbf{F}_{n+1}-\hat{\mathbf{K}}_{n+1}\right), \tag{5.4.31}$$

where $\mathbf{C}_{n+1}=\mathbf{C}(\mathbf{T}_{n+1})$, $\mathbf{F}_{n+1}=\mathbf{F}(\mathbf{T}_{n+1}, t_{n+1})$, and $\hat{\mathbf{K}}_{n+1}=\hat{\mathbf{K}}(\mathbf{T}_{n+1}, t_{n+1})$. Now, one essentially follows the procedure outlined in Section 5.4.6.3 to solve Equation 5.4.31. Note that in the present case, the residual is calculated using Equation 5.4.31. Once the converged solution is obtained at t_{n+1}, the time variable is incremented by Δt. One then repeats the procedure till one reaches time $t=T$.

Numerical Evaluation of Elemental Matrices/Vectors

The different expressions for elemental matrices and vectors given in Section 5.4.5 need to be numerically integrated over the element or its boundaries. The details of the numerical evaluation of elemental matrices and vectors are given in this section.

Isoparametric Elements

The numerical integration is carried out using the concept of isoparametric elements and is explained next. To carry out the numerical integrations, it is convenient to assume that the actual elements are obtained from some simple geometric shapes by using a coordinate transformation. The actual element (Ω^e) is usually referred to as a mapped element, and the element with the simple shape is referred to as a parent element ($\hat{\Omega}$). The integration is then carried out by using some standard integration rules over the parent element. This coordinate transformation or mapping is carried out using appropriate shape functions defined on the parent element. In case of isoparametric elements, the function to be approximated is also expressed using the same shape functions, which are used for the coordinate transformation. The basic idea of using an isoparametric element to carry to out the numerical integration is now illustrated with the help of an example. For concreteness, it is assumed that one is interested in integrating an integral of the form

$$I=\int_{\Omega^e} f(x,y)\,\mathrm{d}x\,\mathrm{d}y, \tag{5.4.32}$$

over the three-node element shown in Figure 5.4.4. The coordinate transformation from the parent element to the mapped element is assumed to be of the form

$$x(\xi,\eta)=\sum_{i=1}^{3}\hat{N}_i(\xi,\eta)x_i^e, \quad y(\xi,\eta)=\sum_{i=1}^{3}\hat{N}_i(\xi,\eta)y_i^e, \tag{5.4.33}$$

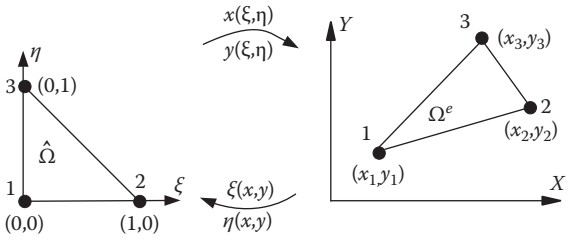


FIGURE 5.4.4 Coordinate transformation.

where the shape functions are given by

$$\hat{N}_1(\xi,\eta) = 1-\xi-\eta,\ \hat{N}_2(\xi,\eta) = \xi,\ \hat{N}_3(\xi,\eta) = \eta. \tag{5.4.34}$$

Here, ξ_i, η_i, $i=1, 2, 3$ are corner coordinates of the parent element and x_i, y_i, $i=1, 2, 3$ are the corner nodes of the mapped element. Note that the mapping between the parent element and the mapped element is chosen in such a way that it is one to one and onto, i.e., it is invertible.

The Jacobian matrix of the transformation is then given by

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial x}{\partial \xi} & \dfrac{\partial x}{\partial \eta} \\ \dfrac{\partial y}{\partial \xi} & \dfrac{\partial y}{\partial \eta} \end{bmatrix}, \tag{5.4.35}$$

while the Jacobian of the transformation is given by

$$J = \det \mathbf{J} = \frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta} - \frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \xi}. \tag{5.4.36}$$

Using the transformation and the corresponding Jacobian given in Equations 5.4.33 and 5.4.36, respectively, the integral given in Equation 5.4.32 is evaluated as

$$I = \int_{\hat{\Omega}} \hat{f}(\xi,\eta) J \, d\xi \, d\eta. \tag{5.4.37}$$

Note that because the mapping given in Equation 5.4.33 is assumed to be invertible the integrand $f(x, y)$ in Equation 5.4.32 can be written as

$$f(x,y) = f(x(\xi,\eta), y(\xi,\eta)) = \hat{f}(\xi,\eta).$$

The evaluation of the terms of the stiffness matrix is discussed next. For two-dimensional problems, the integrands in the elemental stiffness matrix consists of the derivatives of the shape functions in the mapped domain with respect to x and y. To proceed, one introduces the isoparametric element concept in which the approximation to a function in the parent domain is defined using the same shape functions, which are used for the coordinate transformation (see Equation 5.4.34). Hence, the approximation to the temperature in the parent domain, $\hat{T}(\xi,\eta)$, is written as

$$\hat{T}(\xi,\eta) = \sum_{i=1}^{3} \hat{N}_i(\xi,\eta)\hat{T}_i^e, \tag{5.4.38}$$

where $\hat{T}_i^e = T_i^e, i = 1,2,3$. Note that because the mapping given in Equation 5.4.33 is assumed to be invertible one has

$$\hat{T}(\xi,\eta) = \hat{T}(\xi(x,y),\eta(x,y)) = T(x,y).$$

Similarly one has

$$\hat{N}_i(\xi,\eta) = \hat{N}_i(\xi(x,y),\eta(x,y)) = N_i(x,y), \quad i = 1,2,3. \tag{5.4.39}$$

Using Equation 5.4.39 and the chain rule, one obtains the derivatives of the shape functions in the mapped domain with respect to x and y as

$$\frac{\partial N_i}{\partial x} = \frac{\partial \hat{N}_i}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial \hat{N}_i}{\partial \eta}\frac{\partial \eta}{\partial x} \text{ and } \frac{\partial N_i}{\partial y} = \frac{\partial \hat{N}_i}{\partial \xi}\frac{\partial \xi}{\partial y} + \frac{\partial \hat{N}_i}{\partial \eta}\frac{\partial \eta}{\partial y}. \tag{5.4.40}$$

The entries $\partial \hat{N}_i/\partial \xi$ and $\partial \hat{N}_i/\partial \eta$ can be evaluated using Equation 5.4.34, while $\partial\xi/\partial x$, $\partial\eta/\partial x$, $\partial\xi/\partial y$, and $\partial\eta/\partial y$ are given by

$$\frac{\partial \xi}{\partial x} = \frac{1}{J}\frac{\partial y}{\partial \eta}, \frac{\partial \eta}{\partial x} = -\frac{1}{J}\frac{\partial y}{\partial \xi}, \frac{\partial \xi}{\partial y} = -\frac{1}{J}\frac{\partial x}{\partial \eta}, \frac{\partial \eta}{\partial y} = \frac{1}{J}\frac{\partial x}{\partial \xi}. \tag{5.4.41}$$

The entries in Equation 5.4.41 can be readily evaluated using the coordinate transformation given in Equation 5.4.33.

Element Technology

The types of elements used in a finite-element simulation play a very important role in deciding the accuracy of the solution. Elements differ in their shape in the parent domain and also by the number of nodes. Both these attributes decide the type of boundary the elements are capable of modeling accurately. More importantly, these attributes also decide the shape functions that are used to approximate the field variable within an element. The shape functions in turn decide the nature of the approximate solution that is obtained. In order to precisely define the conditions under which the obtained approximate solution converges to the true solution as the mesh is refined, one first finds a quantity referred to as the variational index, *r*. The variational index is the highest order of the derivative appearing in the element equations (see Equation 5.4.9). The convergence requirements can be stated as

1. *Completeness.* The element shape functions must represent exactly all polynomial terms of order r in the Cartesian coordinates. Note that this requirement applies to individual elements.
2. *Compatibility.* The approximation must be C^r continuous in the interior of the elements and C^{r-1} continuous at the interface of the elements. Note that the first part of this requirement applies to individual elements and the second part applies to those elements, which share a common interface.

Note that a function is said to be C^m continuous if the function and all its derivatives up to order m are continuous but its derivative of order $m+1$ is discontinuous. Observe that the variational index for the problem at hand is $r=1$ (see Equation 5.4.9). Hence, as per the convergence requirements, the shape functions must be able to represent all polynomials of degree 1 in the Cartesian coordinates and the approximation must be C^0 continuous at the interface of the elements and must be C^1 continuous in the interior of the elements. For problems with variational index 1, shape functions

that satisfy the Kronecker delta property and the partition of unit property, automatically satisfy completeness and compatibility requirements. The Kronecker delta property and the partition of unity property are explained next.

Consider a n node parent element with nodes located at $\boldsymbol{\xi}_i$, $i=1$, …, n and shape functions $\hat{N}_i(\boldsymbol{\xi}), i=1,\ldots,n$. The Kronecker delta property is given

$$\hat{N}_i(\boldsymbol{\xi}_j)=\begin{cases} 1 & \text{if } i=j, \\ 0 & \text{if } i\neq j. \end{cases} \tag{5.4.42}$$

which states that the shape function associated with node i when evaluated at node i is 1 and when evaluated at any other node is 0. The partition of unity property is given by

$$\sum_{i=1}^{n}\hat{N}_i(\xi)=1, \tag{5.4.43}$$

which states that the sum of all the shape functions evaluated at any point in the element is 1. The most common two and three-dimensional elements, which can be used in problems where a C^1 continuity is required in the interior of the element and a C^0 continuity is required at the interface of the elements are described next.

Two-Dimensional Elements

The most common two-dimensional elements used in practice are either triangular or quadrilateral in shape.

Triangular Elements

The simplest triangular element is a three-node triangular element. The parent three-node triangular element and the mapped three-node triangular element are shown in Figure 5.4.5. Note that for the parent element, $0\leq\xi\leq1$, $0\leq\eta\leq1$, and $\xi+\eta\leq1$.

As can be seen from the figure, the element is capable of representing only straight edges exactly. The shape functions of the element are linear in ξ and η and are as follows:

$$\hat{N}_1=1-\xi-\eta,\ \hat{N}_2=\xi,\ \hat{N}_3=\eta. \tag{5.4.44}$$

An improved approximation is given by a six-node triangular element. The parent six-node element and the mapped six-node element are shown in Figure 5.4.6. Nodes 4, 5, and 6 are mid-side nodes. Note that for the parent element, $0\leq\xi\leq1$, $0\leq\eta\leq1$, and $\xi+\eta\leq1$.

The shape functions are quadratic in ξ and η and are as follows:

$$\hat{N}_1=\lambda(2\lambda-1),\ \hat{N}_2=\xi(2\xi-1),\ \hat{N}_3=\eta(2\eta-1),\ \hat{N}_4=4\xi\lambda,\ \hat{N}_5=4\xi\eta,\ \hat{N}_6=4\eta\lambda, \tag{5.4.45}$$

where $\lambda=1-\xi-\eta$.

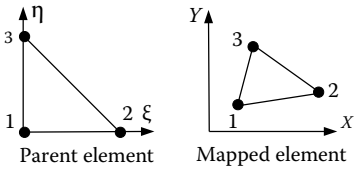


FIGURE 5.4.5 Three-node triangle.

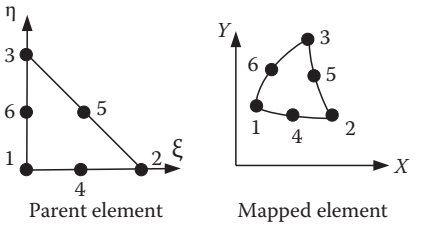


FIGURE 5.4.6 Six-node triangle.

Quadrilateral Elements

The simplest quadrilateral element is a four-node quadrilateral element. The parent four-node square element and the mapped four-node quadrilateral element are shown in Figure 5.4.7. Note that for the parent element, $-1 \le \xi \le 1$ and $-1 \le \eta \le 1$.

As can be seen from the figure, the element is capable of representing only straight edges exactly. The shape functions of the element are bilinear in ξ and η and are as follows:

$$\hat{N}_i = \frac{1}{4}(1+\xi_i\xi)(1+\eta_i\eta), \quad i = 1,2,3,4. \tag{5.4.46}$$

Here, (ξ_i, η_i) are the coordinates of the ith node. An improved approximation is given by a nine-node quadrilateral element. The parent nine-node square element and the mapped nine-node element are shown in Figure 5.4.8. Nodes 5, 6, 7, 8 are mid-side nodes. Note that for the parent element, $-1 \le \xi \le 1$ and $-1 \le \eta \le 1$. The shape functions are quadratic in ξ and η and are as follows:

$$\begin{aligned}
\hat{N}_i &= \frac{1}{4}(\xi^2+\xi_i\xi)(\eta^2+\eta_i\eta), \quad i = 1,2,3,4, \\
\hat{N}_i &= \frac{1}{2}\xi_i^2(\xi^2+\xi_i\xi)(1-\eta^2)+\frac{1}{2}\eta_i^2(\eta^2+\eta_i\eta)(1-\xi^2), \quad i = 5,6,7,8, \\
\hat{N}_9 &= (1-\xi^2)(1-\eta^2).
\end{aligned} \tag{5.4.47}$$

Three-Dimensional Elements

The most common three-dimensional elements used in practice are either tetrahedral or hexahedral in shape.

Tetrahedral Elements

The simplest tetrahedral element is a four-node tetrahedral element. The parent four-node tetrahedral element is shown in Figure 5.4.9. Note that for the parent element, $0 \le \xi \le 1$, $0 \le \eta \le 1$, $0 \le \zeta \le 1$, and $\xi + \eta + \zeta \le 1$.

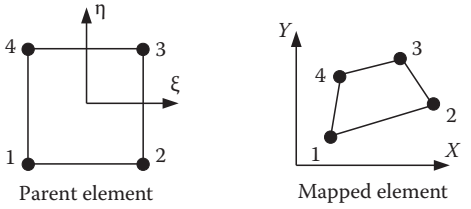


FIGURE 5.4.7 Four-node quadrilateral.

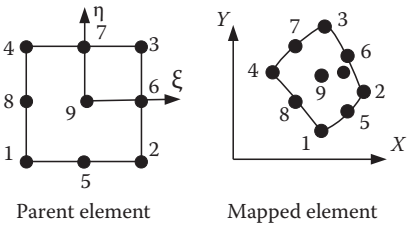


FIGURE 5.4.8 Nine-node quadrilateral.

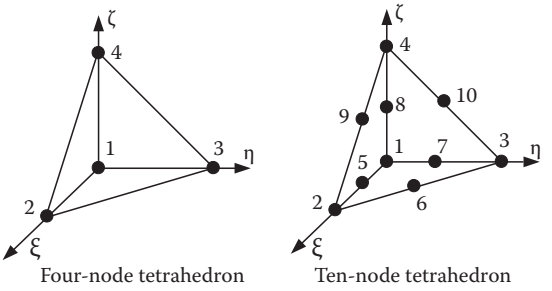


FIGURE 5.4.9 Tetrahedron elements.

The shape functions of the element are linear in ξ, η, and ζ and are as follows:

$$\hat{N}_1 = 1-\xi-\eta-\zeta,\ \hat{N}_2 = \xi,\ \hat{N}_3 = \eta,\ \hat{N}_4 = \zeta. \tag{5.4.48}$$

An improved approximation is given by a 10-node tetrahedral element. The parent 10-node element is shown in Figure 5.4.9. Note that for the parent element, $0 \le \xi \le 1$, $0 \le \eta \le 1$, $0 \le \zeta \le 1$, and $\xi+\eta+\zeta \le 1$. Here, nodes 5, 6, 7, 8, 9, and 10 are mid-side nodes. The shape functions are quadratic in ξ, η, and ζ and are as follows:

$$\begin{aligned} &\hat{N}_1 = \lambda(2\lambda-1), && \hat{N}_2 = \xi(2\xi-1), && \hat{N}_3 = \eta(2\eta-1), && \hat{N}_4 = \zeta(2\zeta-1), && \hat{N}_5 = 4\xi\lambda, \\ &\hat{N}_6 = 4\xi\eta, && \hat{N}_7 = 4\eta\lambda, && \hat{N}_8 = 4\zeta\lambda, && \hat{N}_9 = 4\eta\zeta, && \hat{N}_{10} = 4\eta\zeta, \end{aligned} \tag{5.4.49}$$

where $\lambda = 1-\xi-\eta-\zeta$.

Hexahedral Elements

The simplest hexahedral (brick) element is an eight-node hexahedral element. The parent eight-node element is shown in Figure 5.4.10. Note that for the parent element, $-1 \le \xi \le 1$, $-1 \le \eta \le 1$, and $-1 \le \zeta \le 1$. The shape functions of the element are trilinear in ξ, η and ζ and are as follows:

$$\hat{N}_i = \frac{1}{8}(1+\xi_i\xi)\ (1+\eta_i\eta)(1+\zeta_i\zeta), \quad i = 1,2,3,4,5,6,7,8. \tag{5.4.50}$$

Here, (ξ_i, η_i, ζ_i) are the coordinates of the ith node. An improved approximation is given by a 20-node hexahedral element. The parent 20-node element is shown in Figure 5.4.10. Note that for the parent element, $-1 \le \xi \le 1$, $-1 \le \eta \le 1$, and $-1 \le \zeta \le 1$. Here, nodes 9–20 are mid-side nodes. The shape functions of the element are as follows:

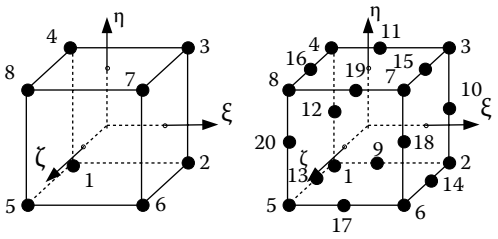


FIGURE 5.4.10 Hexahedral elements.

$$
\begin{aligned}
\hat{N}_i &= \frac{1}{8}(1+\xi_i\xi)(1+\eta_i\eta)(1+\zeta_i\zeta)(\xi_i\xi+\eta_i\eta+\zeta_i\zeta-2) && i=1,2,3,4,5,6,7,8,\\
\hat{N}_i &= \frac{1}{4}(1-\xi^2)(1+\eta_i\eta)(1+\zeta_i\zeta), && i=9,11,17,19,\\
\hat{N}_i &= \frac{1}{4}(1-\eta^2)(1+\xi_i\xi)(1+\zeta_i\zeta), && i=10,12,18,20,\\
\hat{N}_i &= \frac{1}{4}(1-\zeta^2)(1+\xi_i\xi)(1+\eta_i\eta), && i=13,14,15,16.
\end{aligned}
\tag{5.4.51}
$$

For a detailed list of the types of elements available in two and three dimensions and their properties, see, e.g., Zienkiewicz et al. (2013).

Numerical Integration

As discussed earlier, the integrations, which are to be carried out over the mapped element for calculation of various terms in the element equations, are first transformed to integrations over the parent element using appropriate coordinate transformations. These integrations are of the following type:

1. *One-dimensional integrals.* These have to be performed when solving either one-dimensional problems or while evaluating the boundary terms in two-dimensional problems. The integrals are of the form:

$$
I_1 = \int_{-1}^{+1} \hat{f}(\xi) J(\xi)\, \mathrm{d}\xi. \tag{5.4.52}
$$

2. *Two-dimensional integrals.* These have to be performed when solving either two-dimensional problems or while evaluating the boundary terms in three-dimensional problems. The integrals are of the form:

$$
I_{2a} = \int_{0}^{+1}\int_{0}^{1-\xi} \hat{f}(\xi,\eta) J(\xi,\eta)\, \mathrm{d}\xi\, \mathrm{d}\eta, \quad \text{Triangular elements}, \tag{5.4.53}
$$

$$
I_{2b} = \int_{-1}^{+1}\int_{-1}^{+1} \hat{f}(\xi,\eta) J(\xi,\eta)\, \mathrm{d}\xi \mathrm{d}\eta, \quad \text{Quadrilateral elements}. \tag{5.4.54}
$$

3. *Three-dimensional integrals*. These have to be performed when solving three-dimensional problems. The integrals are of the form:

$$I_{3a} = \int_0^{+1} \int_0^{1-\xi} \int_0^{1-\xi-\eta} \hat{f}(\xi,\eta,\zeta) J(\xi,\eta,\zeta)\, d\xi\, d\eta\, d\zeta, \quad \text{Tetrahedral elements,} \tag{5.4.55}$$

$$I_{3b} = \int_{-1}^{+1} \int_{-1}^{+1} \int_{-1}^{+1} \hat{f}(\xi,\eta,\zeta) J(\xi,\eta,\zeta)\, d\xi\, d\eta\, d\zeta, \quad \text{Hexahedral elements.} \tag{5.4.56}$$

These integrals are evaluated using an appropriately chosen numerical integration scheme. The basic idea of any integration scheme is to evaluate the integrand at a few discrete points in the domain, multiply each evaluation with a specific weighting function and finally sum all the weighted evaluations. For example, the one-dimensional integral given by Equations5.4.52 is evaluated as

$$I_1 \approx \sum_{p=1}^{N_g} \hat{f}(\xi_p) J(\xi_p) w_p, \tag{5.4.57}$$

where N_g denotes the number points where the integrand is to be evaluated, ξ_p, $p=1, \ldots, N_g$, denotes the locations of the N_g points, and w_p denotes the weight associated with point ξ_p. The most commonly used scheme to carry out the numerical integration is the Gauss integration scheme. For a one-dimensional problem, a N_g point Gauss integration scheme is able to integrate a polynomial of degree $2N_g - 1$ exactly. The locations of the sampling points and the corresponding weights to carry out the integration in one-dimensional, two-dimensional (both square and triangular) domains, and three-dimensional (both tetrahedral and hexahedral) domains are readily available in standard books (see, e.g., Zienkiewicz et al., 2013). It is important to note that as the order of the integration scheme increases, the number of function evaluations also increase, leading to an increase in the computational cost.

Practical Issues in the Use of the FEM

Some practical issues concerning the use of the FEM for analyzing problems are discussed next.

Meshing

The quality of the finite elements used in the discretization process has a profound effect on the accuracy of the solution (see, e.g., Shewchuk, 2002). The quality of the element is usually measured in terms of the element volume or area, aspect ratio, included angles, etc. Elements with large aspect ratios, large interior angles, and with high degree of distortions should be avoided.

Choice of Elements

The choice of elements dictates the order of approximation. Elements with linear and quadratic shape functions are sufficient for most types of heat conduction analysis. Higher order elements usually lead to higher accuracy at the expense of higher computational cost. In addition, the bandwidth of the stiffness matrix increases that in turn increases the computational effort to solve the system of equations.

Mesh Refinement

The usual procedure followed in performing the finite-element analysis is to first solve the problem using a relatively coarse mesh. Then based on the quality of the result, the mesh is selectively

refined to improve the accuracy of the solution in regions of interest. Mesh refinement is usually required in the regions of high gradients. The idea of mesh refinement is based on the following a priori error estimate (see, e.g., Zienkiewicz et al., 2013)

$$\| \mathbf{e} \|_E \leq Ch^p, \tag{5.4.58}$$

where $\mathbf{e} = \mathbf{T} - \hat{\mathbf{T}}$ is the difference between the true solution, $\mathbf{T}$, and the approximate solution, $\hat{\mathbf{T}}$, $\| (\cdot) \|_E$ is the energy norm, C is a constant, h is the size of a typical element in the mesh, and p is the order of approximation used. As can be seen from Equation 5.4.58, the error in the obtained solution can be reduced by either reducing the element size and/or increasing the order of approximation. Based on this observation, there are three widely used methods for mesh refinement: h-refinement, p-refinement, and h–p refinement. In h-refinement, the type of element is kept the same but the size of the element is reduced. For example, if the original mesh consisted of four-node quadrilateral, the h-refined mesh also consists of four-node quadrilaterals but with decreased size. In p-refinement, the size of the element is kept fixed but the order of approximation of the unknown function within an element is increased. For example, if the original mesh consisted of four-node quadrilateral, the p-refined mesh will consists of nine-node quadrilaterals of the same size. h–p refinement is a combination of h-refinement and p-refinement wherein not only the size of the element is decreased but also the order of approximation of the unknown function within an element is increased. Of the three methods of refinement, the h–p method is found to more effective.

Order of Gauss Integration

As discussed in Section 5.4.7.3, numerical integration is frequently used to carry out the integration of terms appearing in the elemental vectors and matrices. Because exact integration is not carried out, an error is introduced in the finite-element solution. Hence, there is a tendency to use higher order integration schemes, but this leads to a significant increase in the computational costs. Depending on the type of element that is used, one can estimate the order of Gauss integration required for the exact integration of the elemental terms if certain conditions are satisfied, namely, the material properties are constant and the mapped element is undistorted. For more discussion on the choice of the order Gauss integration, see, e.g., Bathe (1996) and Zienkiewicz et al. (2013). In practice, even if the conditions mentioned earlier are not satisfied, one still uses the estimate of the order of Gauss integration to be used based on undistorted elements and constant properties.

Note that symbolic evaluation techniques of the element matrices and vectors that avoid the use of numerical integration are a part of current research toward improving the speed and accuracy of the FEM (see, e.g., Russell and Kelly, 2013).

Choice of Solvers

As mentioned in Section 5.4.6.1, one can use either a direct solver or an iterative solver to solve the set of linear equations arising in the FEM. The choice of the solver is problem dependent. Iterative solvers require less memory compared to direct solvers as there is no fill-in. They can be considerably faster compared to direct solvers for large problems. On the other hand, direct solvers are more robust and are suitable for "difficult" problems. As mentioned before, the performance of an iterative solver is heavily dependent on the choice of the preconditioners. For some problems, it may be computationally more expensive to find a preconditioner compared to using a direct solver. In addition, the direct methods are more effective when a set of equations is to be solved for multiple right-hand side vectors. In such case, the matrix factorization needs to be calculated only once.

Examples

Two examples illustrating the use of the FEM to solve problems in heat conduction are presented next.

A Typical Two-Dimensional Problem

The purpose of this example to demonstrate in sufficient details the different steps involved in using the FEM to solve a typical two-dimensional heat conduction problem shown in Figure 5.4.11. This is an example of a linear steady-state two-dimensional problem.

The figure shows a two-dimensional body of unit thickness with a constant heat source of $f = 36$ W/m^3 that is uniformly distributed. The body is homogeneous and isotropic and the thermal conductivity is $k = 10$ W/(m °C). The temperature on the bottom edge is maintained at 50°C. The right edge is subjected to a zero flux condition, whereas there is an outward flux of 10 W/m^2 from the left edge. Convective heat transfer takes place from the top edge with constant convective heat-transfer coefficient, $h_c = 6$ W/m^2°C and bulk temperature $T_B = 10$°C.

For simplicity, the body is discretized with 4 equal 3 node triangular elements as shown in Figure 5.4.11. The element connectivity matrix is given below (Table 5.4.4):

The element stiffness matrix is calculated for each of the elements using the expression for the stiffness matrix given by

$$K_{ij}^{e} = \int_{\Omega^e} k\left(\frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial y}\frac{\partial N_j}{\partial y}\right) d\Omega^e .$$

The integrand of this integral expression is evaluated using Equation 5.4.40 and the shape functions given by Equation 5.4.44. Note that for this problem the element stiffness matrix can be calculated analytically. The element stiffness matrices for element 1, 2, and 4 are given as

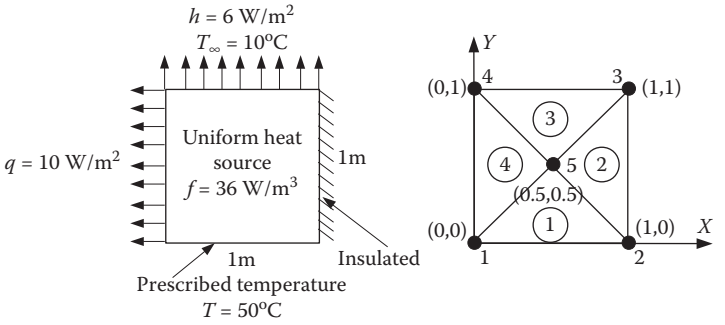


FIGURE 5.4.11 Geometry and boundary conditions of Problem 1.

TABLE 5.4.4
Element Connectivity for Problem 1

Element No.	Local Node Numbers		
	Node 1	Node 2	Node 3
1	1	2	5
2	2	3	5
3	4	5	3
4	5	4	1

$$\mathbf{K}^1 = \begin{array}{c} \\ 1 \\ 2 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 5 \end{array} \\ \begin{bmatrix} 5 & 0 & -5 \\ & 5 & -5 \\ \text{symm} & & 10 \end{bmatrix} \end{array}, \quad \mathbf{K}^2 = \begin{array}{c} \\ 2 \\ 3 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccc} 2 & 3 & 5 \end{array} \\ \begin{bmatrix} 5 & 0 & -5 \\ & 5 & -5 \\ \text{symm} & & 10 \end{bmatrix} \end{array},$$

$$\mathbf{K}^4 = \begin{array}{c} \\ 5 \\ 4 \\ 1 \end{array} \begin{array}{c} \begin{array}{ccc} 5 & 4 & 1 \end{array} \\ \begin{bmatrix} 10 & -5 & -5 \\ & 5 & 0 \\ \text{symm} & & 5 \end{bmatrix} \end{array}.$$

In the present example, the edge 3–5 of element 3 is subjected to convection. Hence, one needs to also calculate the elemental matrix related to convection for element 3. The expression for this matrix is given by

$$H_{ij}^e = \int_{\Gamma^e} h_c N_i N_j \, \mathrm{d}\Gamma^e.$$

Note that this matrix is to be evaluated only along those boundaries on which convection boundary condition is given. The element stiffness matrix and the matrix due to convection for element 3 are given by

$$\mathbf{K}^3 = \begin{array}{c} \\ 4 \\ 5 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 4 & 5 & 3 \end{array} \\ \begin{bmatrix} 5 & -5 & 0 \\ & 10 & -5 \\ \text{symm} & & 5 \end{bmatrix} \end{array}, \quad \mathbf{H}^3 = \begin{array}{c} \\ 4 \\ 5 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 4 & 5 & 3 \end{array} \\ \begin{bmatrix} 2 & 0 & 1 \\ & 0 & 0 \\ \text{symm} & & 2 \end{bmatrix} \end{array}.$$

The element vectors due to the distributed heat source are calculated using

$$f_i^e = \int_{\Omega^e} N_i f \, \mathrm{d}\Omega^e$$

and are given by

$$\mathbf{f}^1 = \begin{array}{c} \\ 1 \\ 2 \\ 5 \end{array} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{f}^2 = \begin{array}{c} \\ 2 \\ 3 \\ 5 \end{array} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{f}^3 = \begin{array}{c} \\ 4 \\ 5 \\ 3 \end{array} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}, \quad \text{and} \quad \mathbf{f}^4 = \begin{array}{c} \\ 5 \\ 4 \\ 1 \end{array} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

Next, the element vectors due to the prescribed flux are calculated using

$$q_i^e = \int_{\Gamma^e} N_i q_n \, \mathrm{d}\Gamma^e.$$

For the present problem, a nonzero flux is prescribed on the edge 4–1 of element 4. Hence, one needs to calculate the element vector due to the prescribed nonzero flux for element 4. The element load vector is obtained as

$$\mathbf{q}^4 = \begin{matrix} 5 \\ 4 \\ 1 \end{matrix} \begin{bmatrix} 0 \\ 5 \\ 5 \end{bmatrix}.$$

Note that because the temperatures are prescribed at nodes 1 and 2 ($T_1 = T_2 = 50°C$), the flux at nodes 1 and 2 (q_1 and q_2) will be treated as unknowns. Finally, the element vectors due to the prescribed convective heat transfer are calculated using

$$h_{\hat{e}}^e = \int_{\Gamma^e} h_c T_B N_i \, d\Gamma^e.$$

As the edge 3–5 of element 3 is subjected to convection, one needs to calculate the elemental vector related to convection only for element 3. This vector is given by

$$\mathbf{h}_c^3 = \begin{matrix} 4 \\ 5 \\ 3 \end{matrix} \begin{bmatrix} 30 \\ 0 \\ 30 \end{bmatrix}.$$

Based on the discussion presented in Section 5.4.4, the elemental matrices and the vectors are assembled to get the equations of the form given in Equation 5.4.18, viz.

$$(\mathbf{K} + \mathbf{H})\mathbf{T} = \mathbf{f} + \mathbf{q} + \mathbf{h}_c,$$

where

$$\mathbf{K} = \begin{matrix} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{matrix} \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{bmatrix} 10 & 0 & 0 & 0 & -10 \\ & 10 & 0 & 0 & -10 \\ & & 10 & 0 & -10 \\ & & & 10 & -10 \\ \text{symm} & & & & 40 \end{bmatrix} \end{matrix}, \quad \mathbf{H} = \begin{matrix} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{matrix} \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 \\ & & 2 & 1 & 0 \\ & & & 2 & 0 \\ \text{symm} & & & & 0 \end{bmatrix} \end{matrix},$$

$$\mathbf{T} = \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \end{bmatrix}, \quad \mathbf{f} = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{bmatrix} 6 \\ 6 \\ 6 \\ 6 \\ 12 \end{bmatrix}, \quad \mathbf{q} = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{bmatrix} q_1 \\ q_2 \\ 0 \\ 5 \\ 0 \end{bmatrix}, \text{ and } \mathbf{h}_c = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{bmatrix} 0 \\ 0 \\ 30 \\ 30 \\ 0 \end{bmatrix}.$$

As mentioned earlier, because the temperatures are prescribed at nodes 1 and 2, the flux at nodes 1, q_1, and 2, q_2, are treated as unknowns. On simplifying and taking into account the prescribed temperature boundary conditions, one obtains,

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}\begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 10 & 0 & 0 & 0 & -10 \\ & 10 & 0 & 0 & -10 \\ & & 12 & 1 & -10 \\ & & & 12 & -10 \\ & \text{symm} & & & 40 \end{bmatrix} \end{array}\begin{bmatrix} T_1 = 50 \\ T_2 = 50 \\ T_3 \\ T_4 \\ T_5 \end{bmatrix} = \begin{array}{c} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \\ \begin{bmatrix} \tilde{q}_1 \\ \tilde{q}_2 \\ 36 \\ 41 \\ 12 \end{bmatrix} \end{array}.$$

Here, $\tilde{q}_1 = q_1 + 6$ and $\tilde{q}_2 = q_2 + 6$. On solving the above set of equations for the unknown temperatures T_3, T_4, and T_5, one obtains $T_3 = 36.2°C$, $T_4 = 36.7°C$, and $T_5 = 43.5°C$. Once the temperature at the nodes is obtained, one can then obtain the temperature at any point in the body by using the element shape functions (see Equation 5.4.6). Note the same problem is solved using ANSYS, a commercial finite-element software. The geometry is discretized using 200 six-node triangular elements. A sparse direct solver is used to solve the set of linear equations. The temperatures obtained in this case are $T_3 = 36.0°C$, $T_4 = 35.5°C$, and $T_5 = 43.3°C$.

Bi-material Problem with Imperfect Interface

The purpose of this problem is to illustrate the versatility of the FEM to handle nontrivial problems in heat conduction. The two-dimensional problem considered was solved using the boundary element method in Bialecki and Kuhn (1993) and is shown in Figure 5.4.12.

The figure shows a plate with a hole made up of two different materials. The two materials share two edges: one with a perfect interface and the other with an imperfect interface. The thermal conductivity of the two materials is as follows:

Material 1:

$$k_1(T) = 70 \text{ W/mK} \quad \text{for} \quad T \le 300 \text{ K},$$
$$= 90 \text{ W/mK} \quad \text{for} \quad T \ge 600 \text{ K}.$$

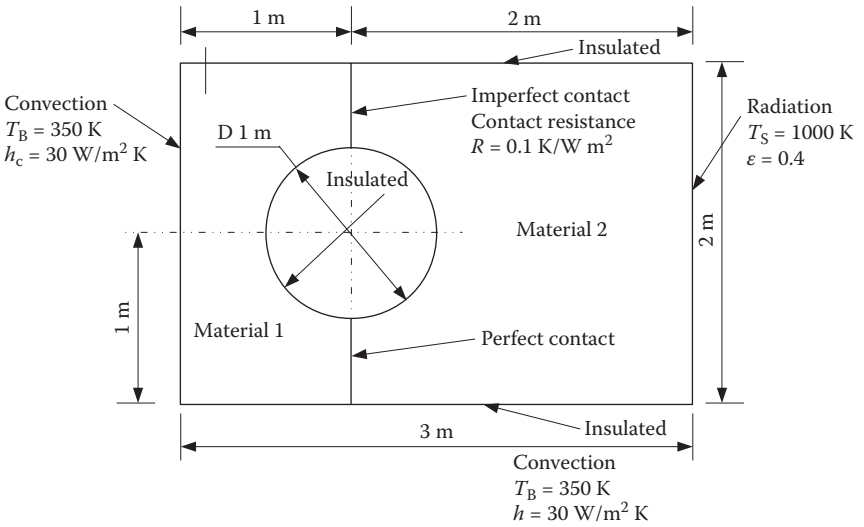


FIGURE 5.4.12 Geometry and boundary conditions of Problem 2.

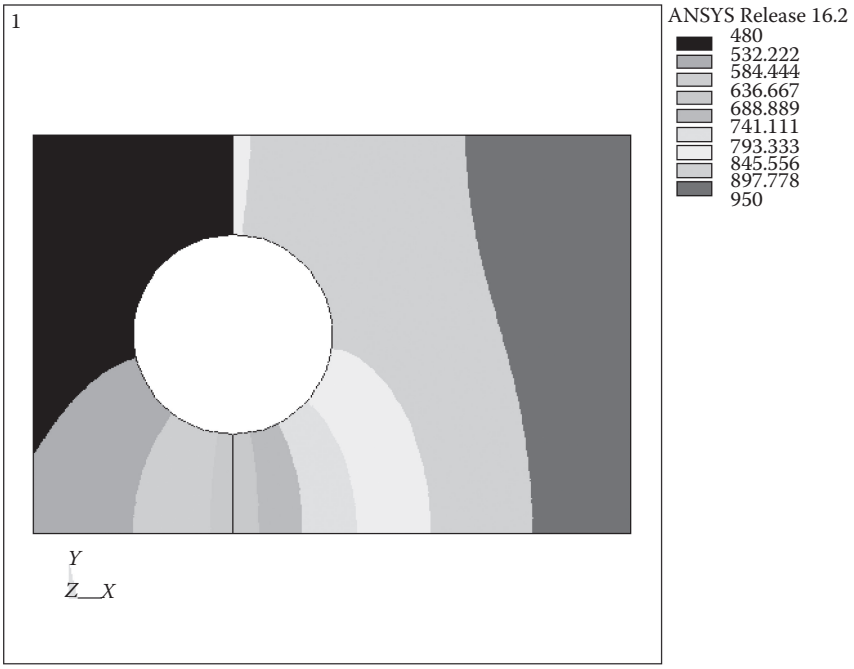


FIGURE 5.4.13 Temperature distribution in the bi-metal plate.

Material 2:

$$k_2(T) = 30 \text{ W/mK} \quad \text{for} \quad T \le 300 \text{ K},$$
$$= 70 \text{ W/mK} \quad \text{for} \quad T \ge 1000 \text{ K}.$$

The thermal conductivity $k_1(T)$ in the interval between [300 K 600 K] is obtained by linear interpolation. Similarly, the thermal conductivity $k_2(T)$ in the interval between [300 K 1000 K] is obtained by linear interpolation. Convective heat transfer takes place from the left edge of the plate, whereas radiation heat transfer takes place at the right edge. All other edges are insulated. It is desired to find the steady-state temperature distribution in the plate. Note that the presence of temperature-dependent heat conductivity and the radiation heat transfer make the problem nonlinear.

The problem is solved using ANSYS, a commercial finite-element software, and the results are shown in Figure 5.4.13.

The geometry is discretized using 4656 six-node triangular elements. The perfect and imperfect interface are modeled using contact elements. A sparse direct solver coupled with the Newton–Raphson scheme is used to solve the set of nonlinear equations. The discontinuity in the temperature field at the imperfect interface is clearly visible in Figure 5.4.13.

REFERENCES

Bathe, K.-J. 1996. *Finite Element Procedures*. Prentice Hall, Upper Saddle River, NJ.

Benzi, M. 2002. Preconditioning techniques for large linear systems: A survey. *Journal of Computational Physics* 182:418–477.

Bergman, T. L., A. S. Lavine, F. P. Incropera, and D. P. DeWitt. 2011. *Fundamentals of Heat and Mass Transfer*. 7th ed. Wiley, New York.

Bialecki, R. and G. Kuhn. 1993. Boundary element solution of heat conduction problems in multizone bodies of non-linear materials. *International Journal of Numerical Methods in Engineering* 36:799–809.

Brenner, S. C. and L. R. Scott. 2008. *The Mathematical Theory of Finite Element Methods*. 3rd ed. Springer-Verlag, New York.

Davis, T. A. 2006. *Direct Methods for Sparse Linear Systems*. Fundamentals of Algorithms. Society for Industrial Mathematics, Philadelphia, PA.

Gartling, D. K. and J. N. Reddy. 2010. *The Finite Element Method in Heat Transfer and Fluid Dynamics*. 3rd ed. CRC Press, Boca Raton, FL.

Golub, G. H. and C. F. Van Loan. 2012. *Matrix Computations*. 4th ed. Johns Hopkins University Press, Baltimore, MD.

Gould, N. I. M. and J. Scott. 2007. A numerical evaluation of sparse direct solvers for the solution of large sparse symmetric linear systems of equations. *ACM Transactions on Mathematical Software* 2 (10): 1–32.

Hestenes, M. and E. Stiefel. 1952. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards* 49 (6): 409–436.

Hughes, T. J. R. 2000. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover Publications, New York.

Ihlenburg, F. 2006. *Finite Element Analysis of Acoustic Scattering*. Springer-Verlag, New York.

Irons, B. 1970. A frontal solution method for finite element analysis. *International Journal of Numerical Methods in Engineering* 2:5–32.

Jin, J.-M. 2014. *The Finite Element Method in Electromagnetics*. 3rd ed. Wiley, Hoboken, NJ.

Oomens, C., M. Brekelmans, and F. Baaijens. 2010. *Biomechanics: Concepts and Computation*. Cambridge University Press, Cambridge, UK.

Reddy, J. N. 2005. *An Introduction to the Finite Element Method*. 3rd ed. Tata McGraw Hill Education Private Limited, New Delhi, India.

Russell, F. P. and P. H. Kelly. 2013. Optimized code generation for finite element local assembly using symbolic manipulation. *ACM Transactions on Mathematical Software* 39:26:1–26:29.

Saad, Y. 2003. *Iterative Methods for Sparse Linear Systems*. 2nd ed. Society for Industrial/Applied Mathematics, Philadelphia, PA.

Saad, Y. and M. Schultz. 1986. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* 7 (3): 856–869.

Shewchuk, J. 2002. What is a good linear finite element? Interpolation, conditioning, anisotropy, and quality measures. In *Proceedings of the 11th International Meshing Roundtable*, 15–126. Albuquerque, NM.

Wilson, E. L. and R. E. Nickell. 1966. Application of the finite element method to heat conduction analysis. *Nuclear Engineering and Design* 4:276–286.

Winget, J. M. and T. J. R. Hughes. 1985. Solution algorithms for nonlinear transient heat conduction analysis employing element-by-element iterative strategies. *Computer Methods in Applied Mechanics and Engineering* 52 (1): 711–815.

Zienkiewicz, O. C. and Y. K. Cheung. 1965. Finite elements in the solution of field problems. *The Engineer* 220:507–510.

Zienkiewicz, O. C., R. L. Taylor, and D. D. Fox. 2014a. The Finite *Element Method for Solid and Structural Mechanics*. 7th ed. Butterworth-Heinemann, Woburn, MA.

Zienkiewicz, O. C., R. L. Taylor, and P. Nithiarasu. 2014b. *The Finite Element Method for Fluid Dynamics*. 7th ed. Butterworth-Heinemann, Woburn, MA.

Zienkiewicz, O. C., R. L. Taylor, and J. Z. Zhu. 2013. *The Finite Element Method: Its Basis and Fundamentals*. 7th ed. Butterworth-Heinemann, Woburn, MA.

5.5 LATTICE BOLTZMANN METHOD

K. Hrisheekesh and Amit Agrawal

Introduction

The lattice Boltzmann method (LBM) is a relatively new approach that offers certain advantages over the traditional computational fluid dynamics to solve the Navier–Stokes (NS) equations. Unlike conventional solvers that directly discretize the NS equations, LBM recovers the NS equations at a macroscopic scale, using simplified kinetic models to capture the essential physics, often with second-order accuracy in both space and time [1]. The inherent parallelism, Poisson freedom, ease of solving multiphase flows, and handling mesoscopic features of the flow make LBM an attractive alternate computational strategy. LBM is suitable for certain classes of problems encountered in practice such as microflows, porous flows, acoustically coupled problems, and so on [2].

Background

The idea that matter was made up of atoms was controversial even at the start of the twentieth century. Philosophers such as Ernst Mach and scientists such as Wilhem Ostwald were opposed to the idea due to the lack of observability of atoms. However, gradually the overwhelming evidence made way for the atomic theory and the kinetic theory of gases was one of its early successes.

In the kinetic theory of gases, molecules are treated as point particles without structure that obey the deterministic Newtonian laws of dynamics. The particles travel in a free path and collide when they meet each other, conserving momentum and energy. Since the rules are deterministic, theoretically, it is possible to track all the molecules in a gas if one knows the position and momentum at the starting time. But the large number of molecules of the order of Avogadro number makes it impossible to track all information about the molecules. Fortunately, all the detailed dynamics is not required for practical purpose. The macroscopic behavior of the system, rather than the motion of individual particles, is of primary interest. Macroscopic quantities can be found by using a suitable averaging process on individual particles. For example, density at a point can be interpreted as the sum of the mass of particles within a small volume located at that point. Similarly, pressure experienced at a wall is interpreted as the momentum exchanged during collision of the molecules with the wall, whereas temperature gives the average kinetic energy of the molecules over a small volume. Transport coefficients can also be assigned a similar interpretation. For example, the momentum exchange between layers of moving fluid gives rise to viscosity while diffusive transport is directly attributed to the random Brownian motion of the atoms.

Lattice Gas Cellular Automaton

Historically, the LBM method evolved from the lattice gas cellular automaton (LGCA) simulations, which in itself was inspired by the kinetic theory of gases. Though no longer used, it is beneficial to understand the basic processes involved in LGCA to appreciate the LBM method itself. Note that the LBM method is at present evolved into a research field of its own and no longer relies on LGCA.

Consider a small control volume of a gas, such as shown in Figure 5.5.1. In the small volume, one can observe atoms that travel in a free path (streaming step) and collide (collision step) when they meet another atom. Under the assumption that the Newtonian mechanics holds and that the interatomic forces are conservative in nature, the collisions conserve both momentum and energy. In real gas, the molecules can collide at any angle and at any velocities indicated by the random arrows in Figure 5.5.1.

In LGCA, the region under consideration is discretized into a lattice, as shown in Figure 5.5.2. This simple lattice, shown in Figure 5.5.2, with a rather square tiling of space with streaming and collision that obey Newtonian dynamics, is known as the Hardy, de Pazzis, and Pomeu automaton, or simply

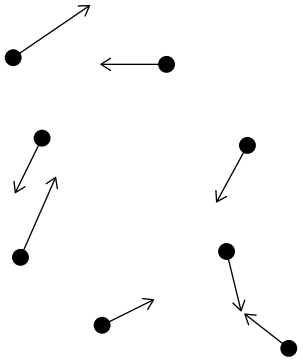


FIGURE 5.5.1 Microscopic picture of real gas molecules colliding and streaming randomly.

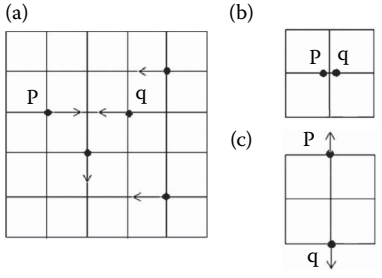


FIGURE 5.5.2 (a) Microscopic picture of a simple lattice gas—collision and streaming at discrete angles, (b) and (c) show the collision process of particles p and q shown in (a). Also known as HPP lattice, microscopic dynamics does not capture the NS equations at macroscopic level.

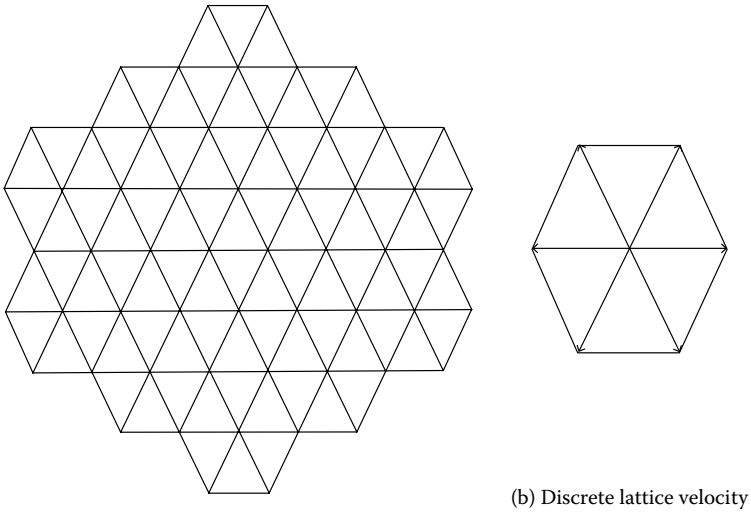
known as HPP automaton [3]. Particles occupy the nodes and have the same mass, m and same velocity magnitude, c. At each time step, they travel from one node to the other through the connected lattice edges along the prescribed direction (streaming/propagation step). When two particles meet at a node, momentum exchange happens and the direction of velocity changes (collision step). During this momentum exchange, the particles obey the mass, momentum, and energy conservation laws.

The crux of LGCA is that the evolution of a system such as the HPP automaton whose underlying microdynamics is more like a cartoon copy of the real molecular dynamics with limited angles of collision and limited magnitudes of speed, can mimic the behavior of fluid on the macroscopic scale. Then suitable microscopic averages can be applied to recover the macroscopic quantities such as pressure and velocity. However, soon after its introduction, it was found that the square lattice was insufficient to capture NS equations on the macroscopic scale. The reason is that such a square lattice does not have enough "angles" or isotropy to capture all the tensor components of the NS equation.

Frisch et al. [4] however showed that a simple automaton can reproduce the NS equations, provided that the lattice has enough symmetry (known as Frisch-Hasslacher-Pomeau (FHP) automaton [4]). It was shown that a hexagon lattice such as the one shown in Figure 5.5.3 has enough symmetry to capture all the tensor components of the 2D NS equation. The essential mechanism remains the same as streaming and collision but the collision rules are more complicated than the HPP automaton. The interested reader can find a comprehensive description of LGCA in the book by Galdrow [19].

However, the LGCA method still had several drawbacks such as the following:

1. Statistical noise—There are statistical fluctuations that may adversely affect evaluation of macroscopic quantities.



(a) FHP lattice

(b) Discrete lattice velocity vectors, FHP

FIGURE 5.5.3 FHP lattice, a lattice with sufficient symmetry to capture the NS equations at the macroscopic scale

2. Lack of Galilean invariance—When the flow Mach number becomes high, the LGCA model displays anomalous problems because of lack of Galilean invariance.
3. Low Reynolds number—The viscosity of the fluid depends on the mean free path of the fluid. In the case of lattice gases, the mean free path is the length of its edges. This dependence of viscosity on lattice geometrical parameters restricts the accessible Reynolds number.
4. Exponential complexity—There is no lattice with enough symmetry for 3D simulations. Thus, a 4D lattice has to be used. This causes the running time to be enormous even for simple problems.
5. Spurious invariants—The invariants for NS equations without sources and sinks are mass, momentum, and energy. But it was found that the LGCA has spurious invariants other than mass, momentum, and energy.

Lattice Boltzmann Method

It was noted that several of the problems of LGCA arise because of the direct use of tracking information for each particle. As mentioned previously, this is mostly unnecessary and one can employ statistical methods to obtain accurate macroscopic information with reduced computational complexity. In the statistical theory of atomistic mechanics, one is no longer interested in individual particles but in the probability of finding an atom of a particular momentum, at a particular location.

Kinetic Theory, Boltzmann Equation, and Chapman–Enskog Procedure

Based on the kinetic theory of gases, Ludwig Boltzmann gave the equation that represents the evolution of the probability density function over time as

$$\underbrace{\frac{\partial f}{\partial t}+\frac{\vec{P}}{m}\frac{\partial f}{\partial \vec{x}}}_{\text{I}}+\underbrace{\vec{F}\frac{\partial f}{\partial \vec{v}}}_{\text{II}}=\underbrace{\left(\frac{\partial f}{\partial t}\right)_{\text{coll}}}_{\text{III}}. \tag{5.5.1}$$

Equation 5.5.1 is central to the kinetic theory of gas and is known as the Boltzmann equation [2]. Here, f denotes the probability of finding a particle at a particular point in phase space, $\vec{P}$ is the momentum of the particle, m is the mass of the particle, and $\vec{F}$ is the externally acting force. In this equation, term I represents the streaming step similar to LGCA, while term II incorporates the effect of external forces such as gravity on the particle distribution function. This term is similar to the one in LGCA but will be omitted for the sake of simplicity. Term III is an important part of the equation that incorporates the effect of intermolecular collisions. In the absence of body forces, the only way for a probability density function to change is through intermolecular collisions. Each of these terms is significant in LBM also and their details will be discussed in a later section.

Solving Equation 5.5.1 is again an arduous task especially because of the nonlinear collision term. Unless an explicit expression for the collision term is given, Equation 5.5.1 would be futile. Boltzmann gave an approximation to the collision term, as a complex integral over phase space, which is however not readily amenable to an analytical solution. Bhatnagar et al. [5] presented a simple way to approximate the collision operator as

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = \frac{f - f^{\text{eq}}}{\tau}. \tag{5.5.2}$$

In Equation 5.5.2, τ is the time taken for the system to attain thermodynamic equilibrium and is known as the relaxation time. The value of τ depends on the viscosity of the fluid. This timescale is independent of the timescales in macroscopic flow and represents the timescale for transport phenomena at the microscopic level.

The term f^{eq} is known as the equilibrium distribution function. Equation 5.5.1 can be exactly solved for an isolated system in thermal equilibrium. Here, the collision term must vanish as the gain or loss of distribution function due to intermolecular collisions at a point in phase space must be in balance for a system in thermodynamic equilibrium. By imposing mass, momentum, and energy conservation as constraints on Equation 5.5.1, one can obtain the equilibrium distribution function, if one knows the macroscopic variables at that particular location. This is given by the expressions

$$f^{\text{eq}} = \left(\frac{1}{2\pi mkT}\right)^{3/2} \exp\left(-\frac{P_x^2 + P_y^2 + P_z^2}{2mKT}\right), \tag{5.5.3}$$

where k is the Boltzmann constant, T is the temperature of the fluid, P denotes the momentum of the particle, and subscripts denote the direction of each momentum component. Equation 5.5.3 is the well-known Maxwell–Boltzmann distribution function.

The formal procedure to recover the NS equation from Equation 5.5.1 is known as the Chapman–Enskog expansion. The procedure involves multiscale expansion, which separates the microscopic and macroscopic time and length scales and then substitutes the multiscale operators in the relevant moments of the distribution function to obtain the spatiotemporal variations of relevant macroscopic variables. The procedure is too cumbersome to be described here. The reader can find a relatively simple, yet detailed derivation procedure in Ref. [6]. It should be noted that all the concepts that are relevant in the kinetic theory of gases, which are used for obtaining the NS equations, are also relevant to the lattice-based kinetic theory of gases, which will be discussed in the next section.

Lattice Boltzmann Method—D2Q9 as an Example

Next, we will examine the D2Q9 lattice and a relatively simple LBM method to model 2D NS equations. The D in D2Q9 represents the dimensions and 2 represents the dimensionality of the problem; further, Q9 implies that the probability density of nine velocity vectors in the phase space are being tracked. The reader may wonder why there cannot be a D2Q6 lattice, which uses the FHP lattice constraints with six velocity vectors, as in LGCA. It is indeed possible to recover the NS equations with just six velocities in the velocity phase space using a D2Q6 scheme. However, the D2Q9 comes with

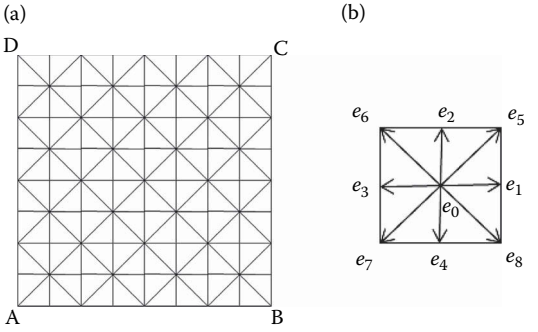


FIGURE 5.5.4 D2Q9 lattice with velocity vectors. The evolution of probability density of these velocity vectors are tracked over time and macroscopic information is recovered as moments of distribution function. (a) D2Q9 Lattice (b) Lattice Velocity Vectors, D2Q9 Scheme.

a rectangular grid that makes it easier to tile the domain and improves the accuracy of the obtained solution. Note that any arbitrary number of velocities is not admissible due to the requirements of lattice symmetry to recover the NS equations. Some of the other well-studied LBM schemes are D1Q3, D1Q5, D2Q9, D3Q15, D3Q19, and D3Q27. The list is not exhaustive. But in general, including more number of velocities in the phase space of LBM increases the computational time at the cost of increase in complexity of implementation. One, therefore, has to strike a balance between cost and accuracy.

The D2Q9 model is one of the popular methods used for solving 2D NS equations using LBM. For the sake of simplicity, it is assumed that no body forces are acting on the fluid. The 2D square lattice has nine discrete velocities for this particular model. The lattice and its discrete velocities are shown in Figure 5.5.4. The crux of LBM lies in calculating the evolution of the velocity distribution function for each of these discrete velocities.

$$\vec{e}_i = \left\{ \begin{array}{lll} (0,\ 0), & & \text{for } i\ = 0 \\ c\cos\left((i-1)\ \dfrac{\pi}{2}\right), & c\sin\left((i-1)\dfrac{\pi}{2}\right), & \text{for } i = 1-4 \\ \sqrt{2c}\cos\left((2i-9)\dfrac{\pi}{4}\right), & \sqrt{2c}\sin\left((2i-9)\dfrac{\pi}{4}\right), & \text{for } i = 5-9 \end{array} \right\}, \tag{5.5.4}$$

where $c = \delta_x / \delta_t$, δ_x is the grid spacing and δ_t is the discrete time step. Usually, unit grid spacing and unit time are used.

The evolution of the particle velocity distribution function, f_i for each of the lattice velocities for a D2Q9 model is given by the following algorithm:

- Initialize: At the start of the simulation, the initial values of macroscopic variables are given. The corresponding values of the equilibrium distribution functions can be calculated from these macroscopic variables over the entire domain.
- Streaming step: From the first step onward, the distribution functions advect from one node to the other along the direction of the corresponding velocity vector. The process is similar to the free motion of molecules in its path without collision (for example, term I in Equation 5.5.1). This step is presented as shown in Equation 5.5.5.

$$f_i\left(\vec{x}+\vec{e}_i, t+\delta_t\right) = f_i\left(\vec{x}, t\right). \tag{5.5.5}$$

- Collision step: The particles at the node collide and similar to that in LGCA, momentum exchange occurs and change in the velocity distribution functions takes place. The change in

the velocity distribution function for each velocity, due to collision, is accounted for by applying the collision operator at the local node. Here, the Bhatnagar-Gross (BGK) approximation, similar to that shown in Equation 5.5.2, is applied at each lattice. The collision step is given as

$$f_i(\vec{x},t) = f_i(\vec{x},t) - \frac{f_i(\vec{x},t) - f_i^{\text{eq}}(\vec{x},t)}{\tau}. \tag{5.5.6}$$

- The parameter τ is known as the relaxation time and is linked to the kinematic viscosity as $\vartheta = (2\tau - 1)/6$. $f_i^{\text{eq}}(\vec{x},t)$ is the equilibrium distribution function for the corresponding velocity vector $\vec{e}_i$ and can be computed once the macroscopic variables at that node are known.

$$f_i^{\text{eq}}(\vec{x},t) = \rho w_i \left(1 + \frac{(\vec{e}_i \cdot \vec{u})}{C^2} + \frac{1}{2}\frac{(\vec{e}_i \cdot \vec{u})^2}{C^4} - \frac{1}{2}\frac{(\vec{u} \cdot \vec{u})}{C^2}\right), \tag{5.5.7}$$

where ρ is the density, $\vec{u}$ is the macroscopic velocity vector at location $\vec{x}$, and w_i is a weightage factor that depends on the velocity magnitude (w_i=4/9 for i=0, w_i=1/9 for i=1 – 4, w_i=1/36 for i=5 – 8). The macroscopic variables can be computed once all the f_is are known at each node

$$\rho = \sum_{i=0-8} f_i, \rho\vec{u} = \sum_{i=0-8} f_i \vec{e}_i, P = \rho C^2. \tag{5.5.8}$$

Here, $C = 1/\sqrt{3}$ is the speed of sound for the D2Q9 lattice and P is the hydrostatic pressure. Note that the pressure is recovered through its relationship with density. It recovers the NS equations with an error proportional to the square of the Mach number. As long as the velocity scales in the lattice are negligible compared to the lattice speed of sound, the D2Q9 scheme recovers the NS equations in the incompressible limit.

- The propagation and collision steps are repeated until the specified time.

It can be shown using the Chapman–Enskog procedure applied to the lattice that the algorithm recovers the fully compressible NS equation as given here:

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} = 0, \tag{5.5.9}$$

$$\frac{\partial \rho u}{\partial t} + \frac{\partial \rho uu}{\partial x} + \frac{\partial \rho uv}{\partial y} = -\frac{\partial p}{\partial x} + \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y}, \tag{5.5.10}$$

$$\frac{\partial \rho v}{\partial t} + \frac{\partial \rho uv}{\partial x} + \frac{\partial \rho vv}{\partial y} = -\frac{\partial p}{\partial x} + \frac{\partial \sigma_{yx}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y}. \tag{5.5.11}$$

In the aforementioned equations, the pressure and velocity are linked through an isothermal equation of state $p = \rho C^2$. One advantage of the compressible formulation is that the LBM can be used in direct acoustic simulations.

Note that Equation 5.5.7 used for calculating the equilibrium function is a truncated series expansion of Equation 5.5.3. This way of calculating the equilibrium preserves the lattice symmetry requirement needed to recover the NS equation and is simpler than calculating the exponential function. However, this introduces a small error in the viscous term in the derived equations, which is proportional to Ma^3, where Ma denotes the Mach number of the flow. It should be noted that the

LBM method is fully compressible in nature and can therefore describe compressible phenomenon such as pressure wave propagation as well. However, the Mach number is restricted to low subsonic flow ($Ma < 0.3$) due to the isothermal nature of the flow and also due to the error in the viscous term owing to the approximation in the equilibrium distribution.

Accuracy and Stability

The LBM method is more faithful to the original kinetic equation 5.5.1 than to the NS equation [7]. Further, unlike conventional solvers, LBM does not discretize the spatiotemporal terms of the NS equations directly. Essentially, there are two ways to describe the accuracy for LBM. In the first case, LBM is viewed as a discretization of the Boltzmann equation and its accuracy is evaluated with respect to the Boltzmann equation. It is second-order accurate in space and time, when considered as the discretization of the Boltzmann equation [1]. However, when applied to macroscopic problems, the accuracy is compared with similar problems that are solved with conventional solvers. Succi [1] stated that for simple problems with turbulence, LBM gives quasi-spectral accuracy. Marie et al. [8] have studied the LBM method for the acoustic problem and found that the LBM is less dissipative than high-order schemes and less dispersive than a second-order scheme in space with a three-step Runge–Kutta scheme in time. Suffice to say that by using the LBM method, one can obtain, in general, second-order accuracy in both space and time for NS equations as well, though in certain cases, it can give exceedingly better accuracy.

It should be noted that similar to conventional solvers, the LBM method is also subjected to stability issues. For example, in the D2Q9 scheme with the BGK collision term, the problem becomes unstable if the value of relaxation time τ is less than 0.5. This lower limit on relaxation time simply implies that viscosity cannot be zero. In practice, the stability issue arises at a finite viscosity close to zero. One way to improve the stability of the LBM method is to use a collision operator that uses multiple relaxation times (see d'Humières [18]), instead of a single relaxation time employed in the BGK model [9]. In addition, there is also the entropic LBM that uses the counterpart of the H-theorem in kinetic theory, which asserts that kinetics of atomic motion proceeds to decrease the quantiy H, defined as $\int f \ln f d^3v$, to improve the stability of the algorithm [1].

Boundary Conditions in LBM

In a fluid dynamics problem, proper implementation of boundary conditions (BCs) is essential for obtaining the flow field solution. Usually certain macroscopic variables are specified on the boundary. One simple way to implement the BCs is to compute the equilibrium distribution function at the boundaries and propagate those distribution functions to the interior. Usually, such approaches have first-order accuracy. There are several other approaches as well, such as the Zou–He method, Inamuro method, regularized BC (see Zou–He [21] and Latt [10]). In all methods, the primary issue in implementing the BCs is computing the incoming particle velocity distribution functions. Some of the commonly used macroscopic BCs and their implementation are discussed in the following subsections.

No-Slip Wall

The no-slip BC is obtained using the bounce-back scheme. In the bounce-back scheme, the particle hits the wall and reverses the velocity as shown in Figure 5.5.5. Therefore, the corresponding distribution functions also reverse. The bounce-back scheme on the macroscopic scale retrieves the no-slip BC. Usually, the wall is assumed to be in mid-grid so that the bounced-back particle reaches exactly back into the adjacent fluid node. The mid-grid scheme has second-order accuracy [2]. The on-grid scheme is also possible but the accuracy reduces to first order.

Slip Wall

The slip BC is implemented easily using specular reflection (see Figure 5.5.6). Often in microflows, partial bounce back and reflection are employed to obtain Maxwell's BCs at the wall.

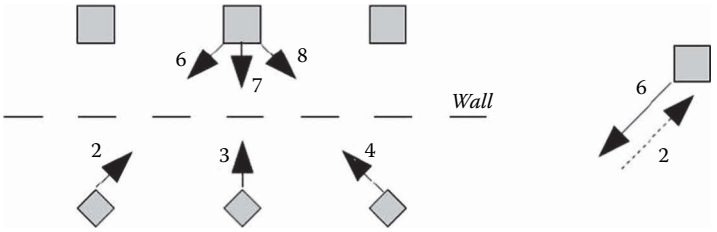


FIGURE 5.5.5 The shaded squares represent the fluid nodes adjacent to wall. The diamond nodes are solid nodes. The population f_6 strikes the wall reverses the momentum and bounces back as f_2 at a macroscopic level this is equivalent to no slip boundary condition.

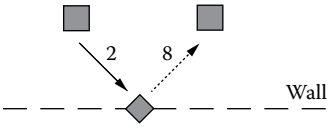


FIGURE 5.5.6 The shaded squares represent the fluid nodes adjacent to the wall. The diamond nodes are solid nodes. The population f_2 strikes the wall and reflects in a specular manner to reach f_8 at the adjacent fluid node. It retrieves slip boundary conditions at the macroscopic scale.

Velocity BCs

At velocity inlet, all velocity components are known. Once again, the equilibrium distribution functions can be calculated and propagated inward. However, there are also other methods such as the Zou–He, Inamuro BCs that can give better accuracy than simply using equilibrium distribution functions. Description of these methods is beyond the scope of the present work. The reader can refer to Latt [10] for details.

Periodic BC

In the periodic BC, the particles that leave from part of the system enter through the other part of the system. This is most readily implemented by supposing the periodic boundary nodes as adjacent to each other and exchanging distribution functions accordingly. For example in Figure 5.5.5, the particles that leave face BC enter through AD when periodicity along horizontal direction is required.

It should be noted that the BCs listed here are some of the simple BCs encountered in practice. When more complex BCs are encountered, one can always use the corresponding macroscopic information and use suitable interpolation and extrapolation schemes to achieve higher order accuracy. Characteristic-based methods and incorporation of acoustic BCs for aeroacoustic calculations are some of the new developments in BC implementation of LBM methods.

Example Problems

The LBM method eliminates almost all the difficulties associated with LGCA. The simplicity in implementation and high parallelizability of both streaming and collision steps involved in the algorithm imply that LBM can be efficiently used in modern-day parallel machines. Already several open-source implementations of LBM such as OpenLB [12], Palabos [13], and Sailfish [14] provide fully parallel, efficient CPU/GPU (central processing unit/graphics processing unit) implementation of several LBM models. Snapshots of some of the practical and complex problems which were solved using LBM are given in Figures 5.5.7 through 5.5.9 taken from references [15–17] respectively. The interested reader can get hands-on experience of the LBM method with the use of open-source programs (see Refs. [12–14]) that are freely available and fairly well documented.

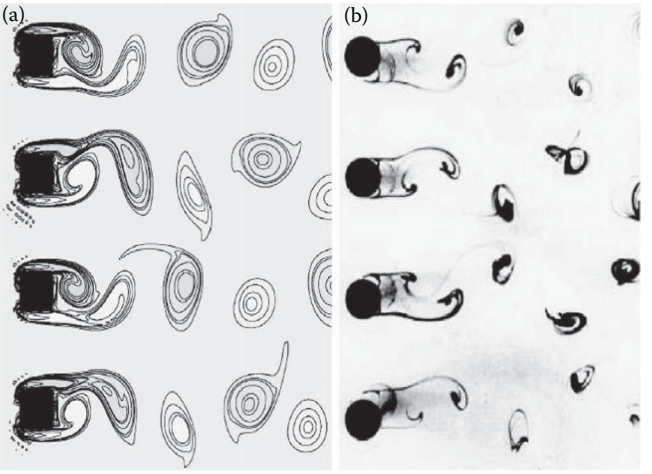


FIGURE 5.5.7 Comparison of vortex shedding from a row of cylinders (s/d = 2.0 and Re = 200), (a) LBM computations, (b) experiments of Kobayashi (1984) (taken from Mizushima & Akinaga 2003). (From Kumar, S.R., Sharma, A., and Agrawal, A., *J. Fluid. Mech*., 606, 369–392, 2008. With permission.)

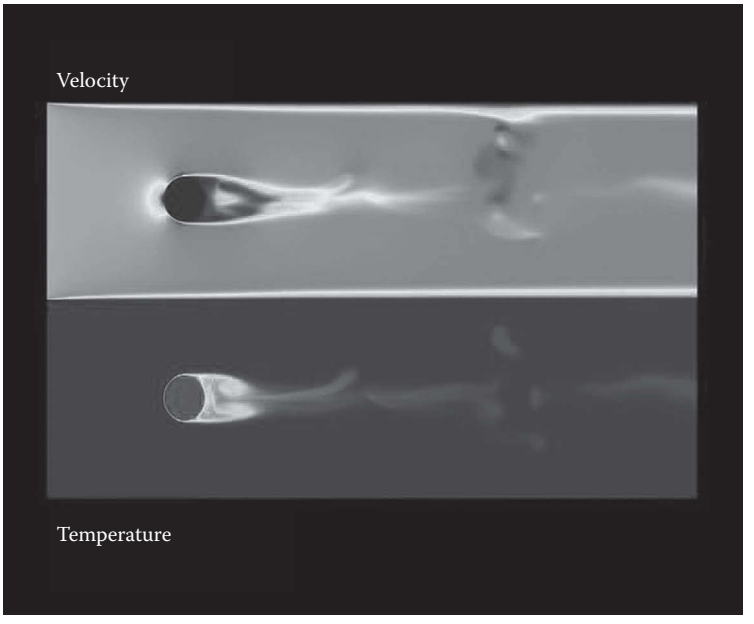


FIGURE 5.5.8 3D simulation of thermal flow past a sphere. (From FlowKit Ltd, Palabos User Guide, 2011, http://www.palabos.org.)

Concluding Remarks

During the past three decades, LBM has transgressed from just a novel approach to solve NS equations to a solver that is capable of solving industrial level problems. Already both commercial and open-source libraries (see Refs. [11–13]) are available in the market, which use the inherent strengths of LBM such as simplicity and ease of parallelization. The research aspect of LBM is also currently strong, with the use of novel thermal models, multiphase flows, turbulence modeling,

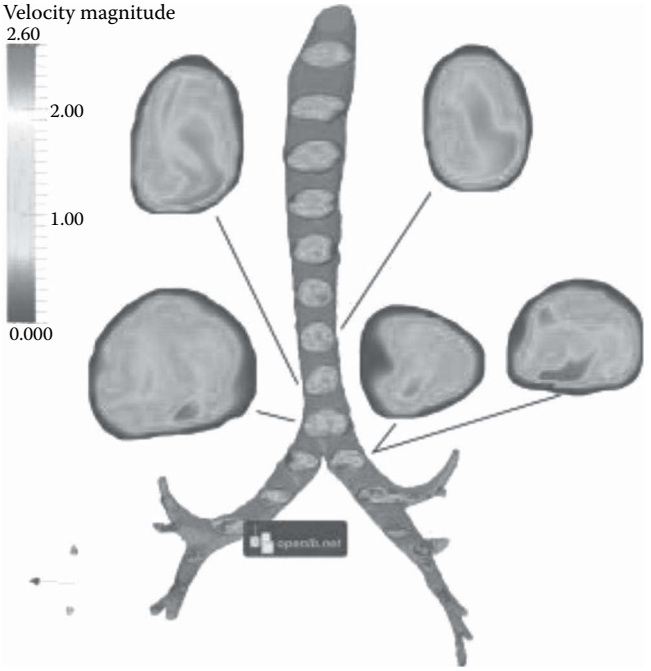


FIGURE 5.5.9 3D simulation of airflow through human trachea (From Krause, M.J., Fluid flow simulation and optimisation with lattice Boltzmann methods on high performance computers, PhD Dissertation, Karlsruhe Institute of Technology, 2010. With Permission.)

non-Newtonian fluid flows, and so on. Owing to the several advantages of LBM discussed here, we expect the use of this technique for computation of fluid flow and thermal problems to grow further with time.

REFERENCES

1. Chen, S. and Doolen, G. D., Lattice Boltzmann method for fluid flows, *Annual Review of Fluid Mechanics*, 30 (1998), 329–364.
2. Succi, S., *The Lattice Boltzmann Equation for Fluid Dynamics and Beyond*, Oxford University Press, Oxford (2001).
3. Hardy, O. J., de Pazzis, O., and Pomeau, Y., Molecular dynamics of a lattice gas: transport properties and time correlation functions, *Physical Review A*, 13 (1973), 1949.
4. Frisch, U., Hasslacher, B., and Pomeau, Y., Lattice-gas automata for the Navier–Stokes equation, *Physical Review Letters*, 56 (1986), 1505–1508.
5. Bhatnagar, P. L., Gross, E. P., and Krook, M., A model for collision process ingases. I. Small amplitude processes in charged and neutral one-component system, *Physical Review,* 94 (1954), 511–525.
6. Viggen, E.M., The lattice Boltzmann method: fundamentals and acoustics, PhD thesis, Norwegian University of Science and Technology (2014).
7. He, X. and Luo, L. S., Theory of the lattice Boltzmann equation: from Boltzmann equation to lattice Boltzmann equation, *Physical Review E*, 56 (1997), 6811–6817.
8. Marié, S., Ricot, D., and Sagaut, P., Accuracy of lattice Boltzmann method for aeroacoustics simulations, *AIAA-Paper,* (2007), 3515.
9. d'Humières, D., Ginzburg, I., Krafczyk, M., Lallemand, P., and Luo, L.-S., Multiple relaxation-time lattice Boltzmann models in three dimensions, *Philosophical Transactions of the Royal Society of London A*, 360 (2002), 437–451.
10. Latt, J., Straight velocity boundaries in the lattice Boltzmann method, *Physical Review E*, 77 (2008), 056703.

11. Exa Corporation, PowerFlow product brief, 2015, http://www.exa.com.
12. FlowKit Ltd, Palabos User Guide, 2011, http://www.palabos.org.
13. Krause M. J., OpenLB User Guide v1.0, 2016, http://www.optilb.org.
14. Michal J., Sailfish Reference Manual 2013.1, 2014, http://www.sailfish.us.edu.pl.
15. Kumar, S. R., Sharma, A., and Agrawal, A., Simulation of flow around a row of square cylinders, *Journal of Fluid Mechanics*, 606 (2008), 369–392.
16. Krause, M. J. Fluid flow simulation and optimisation with lattice Boltzmann methods on high performance computers, PhD Dissertation, Karlsruhe Institute of Technology (July 2010).
17. He, X. and Doolen, G. D., Lattice Boltzmann method on a curvilinear coordinate system: vortex shedding behind a circular cylinder, *Physical Review E*, 56 (1997), 434.
18. d'Humières, D., Ginzburg, I., Krafczyk, M., Lallemand, P., and Luo, L.-S., Multiple-relaxation-time lattice Boltzmann models in three dimensions, *Philosophical Transactions of the Royal Society A*, 360 (2002), 437–451.
19. Wolf-Gladrow, D. A., *Lattice-Gas Cellular Automata and Lattice Boltzmann Models: An Introduction*, Springer-Verlag, Berlin Heidelberg (2005).
20. Philliova, O., Grid refinement for lattice-BGK models, *Journal of Computational Physics*, 147 (1998), 219–228.
21. Zou, Q. and He, X., On pressure and velocity boundary conditions for the lattice Boltzmann BGK model, *Physics of Fluids*, 9 (1997), 1591.

5.6 IMMERSED BOUNDARY METHODS FOR FLUID-STRUCTURE INTERACTION SIMULATIONS

Rajneesh Bhardwaj and Atul Sharma

Introduction

Immersed boundary (IB) methods refer to a class of methods in which the flow and/or structure dynamics of an "immersed" or "embedded" structure, inside a fluid domain, is solved on a standard coordinate (mostly Cartesian)-based grid—the grid does not conform to the shape of the structure [1]. The method is more commonly used for moving as compared to stationary structures. The computational modeling of such *moving boundary* or *fluid–structure interaction* (FSI) problems involves calculation of fluid dynamic forces exerted on the immersed complex geometry structures, which may result in large-scale flow-induced structural deformation. The structure typically exhibits geometric as well as material nonlinearity and should be accounted for in the modeling of such systems. In the case of biological systems, the structures may be multilayered, anisotropic, and heterogeneous, and this increases the complexity of modeling the structure dynamics. Typical examples include flow-induced vibration of vocal folds in the larynx [2], cardiac hemodynamics [3], fish fins during propulsion [4,5], deformation of mitral and aortic valves by blood flow [6], deformation of insect wings [7], interaction of a blast wave with the human eye [8,9], energy harvesting devices [10], thermoregulation in elephants via flapping of their large ears [11], thermal transport enhancement in microchannels using oscillating synthetic cilia [12], and thermal augmentation using vortices generated due to flow-induced deformation of thin structures [13,14].

The other class of methods, known as arbitrary Lagrangian–Eulerian (ALE) methods, involves body-fitted grids conforming to the shape of the body. The grids for ALE and IB methods in the fluid domain are compared in Figure 5.6.1. The ALE method uses the Lagrangian formulation of the Navier–Stokes equations for structure dynamics and the Eulerian formulation of the Navier–Stokes equations for fluid dynamics [15–17]. The two formulations are employed with a weak coupling or a strong coupling. The main issue, however, is the time-wise distortion of the body-fitted grid during a remeshing procedure—needed to ensure a body-fitted grid around the moving structure. In this regard, Tezduyar et al. [18] proposed an ALE finite-element formulation for the fluid as well as the structure domain in which the remeshing criterion was based on element size. Furthermore, Souli et al. [19] used an ALE formulation with smoothing algorithms to control the distortion of the mesh. The time-wise varying grid generation requires more computational time, and it may introduce numerical errors due to the mapping of the solution from the old to the new grid.

On the other hand, in IB methods, the governing flow and structure equations are solved on a time-wise fixed non-body-fitted Cartesian grid. Fluid–structure interface is tracked in Lagrangian framework; except some recent work, such as that by Shrivastava et al. [20], which uses the level-set method on an Eulerian equation for the interface evolution. IB method was proposed by Peskin [21,22] to simulate cardiovascular dynamics, and this method and its variants such as distributed Lagrangian multiplier/fictitious domain (DLM/FD) methods, have been reported and utilized for solving several problems (see review on IB methods in Refs. [1,23,24]). Note that the IB method can easily be integrated into the existing flow solvers based on finite-difference methods [25–30].

In the case of the body-fitted grid, since the solid–fluid interface coincides with the boundary of the solid, the implementation of the solid wall boundary condition is straightforward. This is also true for the finite-volume method based on implementation of the conservation laws in the fluid control volumes (CVs); whereas, in the case of the non-body-fitted Cartesian grids in IB methods, partially filled (containing both fluid and solid) CVs that are encountered here need special treatment to ensure compliance with the conservations laws. Special treatment is also needed to accurately implement a boundary condition (BC) at the moving solid–fluid interface. Furthermore, a solid CV,

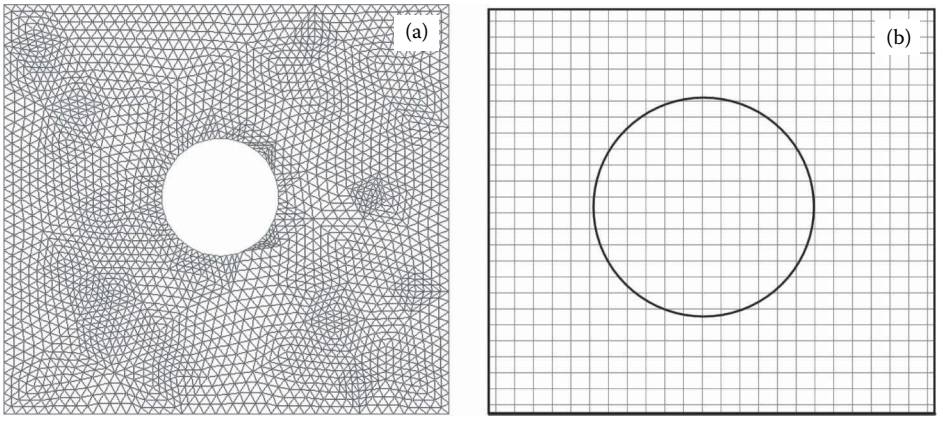


FIGURE 5.6.1 (a) Body-fitted time-wise varying unstructured grid for the ALE method and (b) non-body-fitted fixed Cartesian grid for the IB method.

a fluid CV, and a partially filled CV may interchange their identification as time progresses in a moving boundary problem. This is because certain solid CVs get exposed to fluid, and vice versa. Thus, the type of CV changes with time and needs to be properly accounted for in the IB method.

The advantages of the IB method are the simplicity of the grid (Cartesian), ease of programming due to well-defined discretization stencil, ease of computation due to numerically well-behaved coefficient matrix and reduced computation effort per grid point, and no remeshing needed for a moving boundary problem. On the other hand, some of the disadvantages include special treatment for the imposition of BCs and no control on the grid for local clustering (refinement of the grid locally).

Miital and Iaccarino [1] classified IB methods into two broad categories, namely, *continuous forcing approach* and *discrete forcing approach*. In the former, the effect of fluid–structure interface is modeled using a source term in the momentum equation. On the other hand, in the latter, the governing equations are first discretized on the Cartesian grid without treating the IB; further, discretization in the cells near the IB is adjusted for the presence of IB [1]. Due to the smoothening of the forcing function that is required across the IB in the continuous forcing approach, sharp representation of IB is not possible [1] and sometimes, this approach is referred to as the *diffuse-interface IB method*. In the discrete forcing approach, since the IB is directly treated, it is referred to as the *sharp-interface IB method*.

The sharp-interface discrete-forcing approach is further classified mainly as *ghost-cell method* and *cut-cell method*; the latter method is based on the finite-volume approach and the former method is mostly dependent on the finite-difference method. Furthermore, in the former method, the relevant boundary conditions (BC) at the IB are applied (on the ghost grid points) through an interpolation stencil; while in the latter, cells that are "cut" by IB are identified and absorbed by the neighboring cells. The cut-cell method uses the body-fitted grid for a layer of CVs enveloping the solid body and Cartesian CVs in the other regions. The main advantage of the cut-cell method over the ghost-cell method is the improvement in the global and local conservation of mass and momentum. However, extending the cut-cell method to three dimensions, reshaping/reconstruction of CVs, and discretization of the full Navier–Stokes equations on such CVs can prove to be difficult for complex shaped structures. Both the ghost-cell and the cut-cell methods use a Lagrangian method for the time-wise variation of the solid–fluid interface; whereas, a recent level set-based IB method [20] uses an Eulerian level-set equation for the temporal evolution of the interface.

The scope of this chapter is to present both CFD development and application aspects of the IB method. For the CFD development, both the diffused-interface continuous-forcing and sharp-interface discrete-forcing approaches are presented. Furthermore, the discrete forcing approach is presented for three different methods: ghost-cell, cut-cell, and level set-based IB methods.

Thereafter, advanced modeling issues in the IB method are presented for the treatment of moving boundaries and coupling with a structural solver. The CFD application of the IB method is treated for both rigid and deformable moving structures, considering some scientifically exciting and engineering-relevant fluid dynamics and heat-transfer problems.

Continuous Forcing Approach

In this approach, the fluid flow is governed by the incompressible Navier–Stokes equations, which are solved on a stationary Cartesian grid. The governing equations are given as

$$\frac{\partial v_i}{\partial x_i} = 0, \tag{5.6.1}$$

$$\frac{\partial v_i}{\partial t} + \frac{\partial v_i v_j}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 v_i}{\partial x_j^2}, \tag{5.6.2}$$

where $i, j = 1, 2$, and 3; v_i are the velocity components, p is the pressure, ρ is the fluid density, and ν is the kinematic viscosity. In this approach, the forcing function (f_b) is included as a source term in the aforementioned continuous governing equations, $L(U) = f_b$. The resulting equations are solved for the whole domain ($\Omega_f + \Omega_b$). Note that $f_b = (f_p, f_m)$, where f_p and f_m are forcing functions applied to the continuity and momentum equations, respectively. Thus, in the continuous forcing approach, forcing is applied to the continuous governing equation before discretization. The application of this (continuous forcing) approach to the elastic and rigid structures are presented separately in the following subsections.

Elastic Structure

The continuous forcing IB method was proposed by Peskin [21,22] to simulate a coupled FSI involving blood flow and muscle contraction in a beating heart; it is generally suitable for flows with immersed elastic boundaries. The IB (wetted surface of the blood vessels) is represented by a set of elastic fibers, and the location of these fibers is obtained in a Lagrangian fashion. The effect of IB on the surrounding fluid is captured by a force distribution f on these fluid particles, calculated using a constitutive model, say Hooke's law, and it can be represented in terms of the Dirac delta function, δ, as follows [1]:

$$\vec{f}(\vec{x}, t) = \sum_k \vec{F}(t)\delta\left(\left|\vec{x} - \vec{X}_k\right|\right), \tag{5.6.3}$$

where the coordinate X_k of the kth Lagrangian particle is expressed as [1]

$$\frac{\partial \vec{X}_k}{\partial t} = \vec{u}(\vec{X}_k, t). \tag{5.6.4}$$

The localized force term is used as a source term in the Navier–Stokes equation. Providing sharp interface at the IB by the Dirac delta function creates a numerical instability during the numerical solution of the Navier–Stokes equation. Thus, as discussed by Xu and Wang [31], the communication between the fluid and the IB is realized by spreading the force from the Lagrangian grid to the Cartesian grid and interpolating the velocity from the Cartesian grid to the Lagrangian grid.

As discussed by Mittal and Iaccarino [1], since the locations of the Lagrangian particles do not coincide with the Cartesian grid, the forcing function is distributed over a band of cells around each Lagrangian particle and requires smoothing of the forcing function across the IB. Thus, the representation of the IB is diffused and not sharp. The choices of several distribution functions are available in a review paper [23].

Rigid Structure

The aforementioned method for the elastic structure is applied for the rigid structure by considering the fiber to have high stiffness. Consequently, the governing equations of the fluid motion become very stiff and pose numerical difficulties in obtaining their solution.

Another approach to solve for a rigid structure is to employ a specific version of the model developed by Goldstein et al. [32]. This model gives good results at low Reynolds number but leads to stability problems at high Reynolds numbers.

Discrete Forcing Approach

The governing equations for the flow are 3D, unsteady, incompressible, Navier–Stokes equations, described by Equations 5.6.1 and 5.6.2. As discussed earlier, this approach can be classified into three categories, namely, ghost-cell, cut-cell, and level set-based IB method; each of these is dealt with separately in the ensuing subsections.

Ghost-Cell-Based IB Method

In a significant paper, Mittal et al. [27] proposed a sharp-interface IB-based method, where the Navier–Stokes equation (Equation 5.6.2) are discretized in space using a cell-centered, colocated (non-staggered) arrangement of primitive variables u_i and p. A second-order central difference method is used for all spatial derivatives, and the time integration is performed with the second-order Crank–Nicolson method for the convective and viscous terms. In addition, the face-center velocities, U_i, are computed as proposed by Zang et al. [33], to improve the mass conservation.

The unsteady Navier–Stokes equation is time-marched using a fractional step scheme, which involves two steps: solving an advection–diffusion equation followed by a pressure Poisson equation. During the first step, both the convective and the viscous terms are treated implicitly using the Crank–Nicolson scheme to improve the stability. In the second step, the pressure Poisson equation is solved with the constraint that the final velocity be divergence-free. The surface of the immersed body is represented as an unstructured surface mesh that consists of triangular elements (Figure 5.6.2), immersed in a nonuniform Cartesian grid-based fluid domain.

The ghost-cell method involves first a Lagrangian technique (ray tracing algorithm) for the representation of an instantaneous interface (by identifying the solid and fluid grid points); then, the method leads to the identification of the ghost from the solid grid points; and finally, obtaining the image of the ghost grid points. The resulting solid, fluid, ghost, boundary intercept, and image grid points are shown in Figure 5.6.3. The figure shows the ghost grid points as those solid grid points that have at least one neighboring CV in the fluid. In the method of Mittal et al. [27], a "normal probe" is extended from a ghost grid point to intersect with the immersed boundary at a point, called as the "body intercept", shown in Figure 5.6.3. The probe is extended into the fluid to the

FIGURE 5.6.2 Typical surface mesh with triangular surface elements is shown for a plate attached to a cylinder and can be used in the sharp-interface IB-based method proposed by Mittal et al. [27].

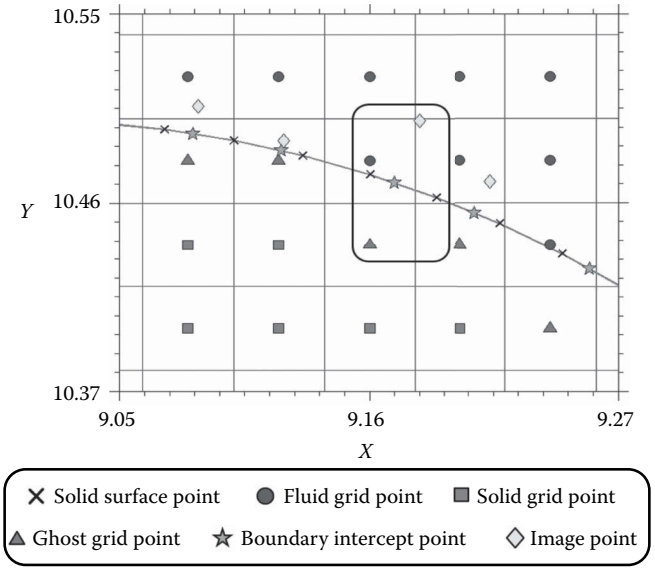


FIGURE 5.6.3 Different types of grid points defined for the implementation of the ghost-cell-based IB method [26].

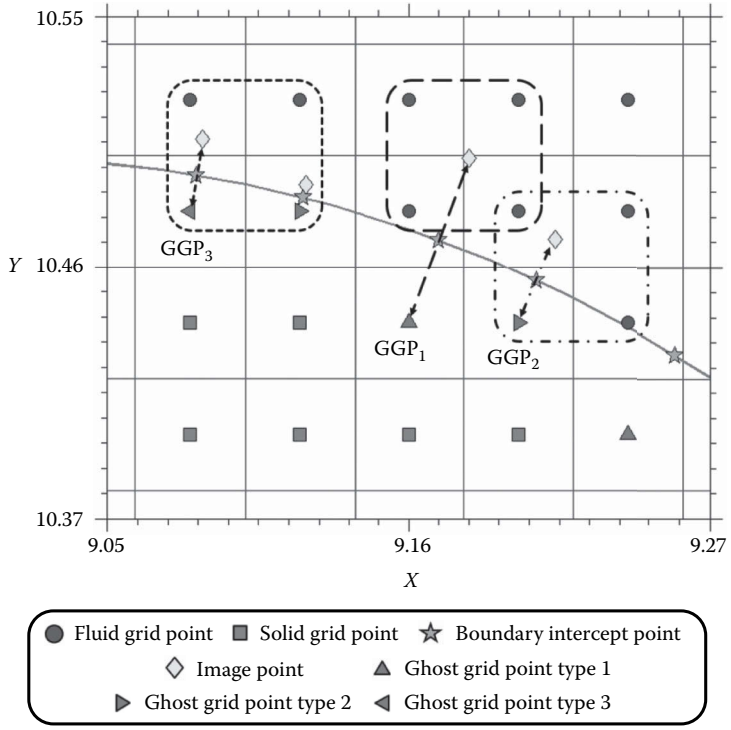


FIGURE 5.6.4 Three different types of ghost grid points, corresponding to the explicit or implicit computational stencil for the bilinear interpolation based on determination of their image point values [26].

"image point" such that the body intercept lies midway between the image and the ghost points, shown in Figure 5.6.3.

Figure 5.6.4 shows three different types of *computational stencils,* for the computation of the flow properties at the image point, which results in the three types of ghost grid points. The various types of ghost grid points, shown as GGP_1 and GGP_2 (and GGP_3) in the figure, lead to an explicit and implicit computation for the flow properties at the image point, respectively. The computation involves bilinear (for 2D here) or trilinear (for 3D) interpolation procedure, using the neighboring grid point values seen in the stencils. The interpolated value of the flow properties at the image point is used to determine the ghost grid point values for the imposition of the solid wall boundary condition (Dirichlet or Neumann type). As shown by Mittal et al. [27], this procedure leads to a nominally second-order accurate specification of the boundary condition of the immersed boundary.

Cut-Cell-Based IB Method

All the IB methods discussed earlier are not designed to satisfy the governing equations for the cells in the vicinity of the IB. The local and global conservation of mass and momentum can only be achieved by application of the finite-volume method to discretize the governing equations. This is the primary motivation of the cut-cell method.

In this method, the cells cut by IB and their centroids are identified. The region (solid or fluid) in which the cell center of the cut cell lies is treated as an independent cell and the other part of the cut cell is merged with the neighboring (fluid or solid) cell. This results in the formation of control volumes, which are of trapezoidal shape near the IB.

For the finite-volume method, one needs to evaluate mass, momentum, and pressure flux across the cell faces; the problem arises in evaluating these fluxes for the trapezoidal faces. To overcome this difficulty, Ye et al. [34] proposed a method to express any generic variable (ϕ) at the cell centers in terms of the neighboring cell values by a polynomial interpolation. The local accuracy of this method near the IB depends upon the choice of the interpolation scheme.

The cut-cell finite-volume approach was first introduced by Clarke et al. [35] and was later applied to viscous flows [34,36–38]. This method improves the global and local mass and momentum conservation as compared to the other methods. The extension of the cut-cell approach to 3D geometries is not easy because it leads to complex polyhedral cells and discretization of the Navier–Stokes equations on such cells is difficult [1]. Recently, Seo and Mittal [39] used a 3D cut-cell-based approach to enforce the geometric conservation law near the IB thereby reducing the spurious pressure oscillations due to the freshly cleared cells.

Level Set-Based IB Method

Level-set method [40] (LSM) is a widely used numerical technique for the unsteady computations of moving boundaries, not only for fluid dynamics but also for various other problems in science and engineering. For fluid dynamics, LSM was first used for the simulation of the unsteady two-phase interface in a multiphase flow situation (presented in the next section). More recently, the level-set method has been used in the IB method as the Eulerian method for tracking the unsteady solid–fluid interface; this is because the LSM is relatively easier to implement numerically as compared to the Lagrangian approach (commonly used in the ghost-cell and cut-cell methods); and it is relatively easier to implement the BC at the moving solid wall.

Choi et al. [41] and Tan et al. [42] have used the level-set method, using a discrete forcing function in the momentum equation on a Cartesian grid. Using the cut-cell method, Cheny and Botella [43] presented the level set-based IB method where the level-set function is used to represent the irregular boundary and to efficiently compute the geometrical parameters of the cut cell.

More recently, for the Cartesian grid, Shrivastava et al. [20] presented a *level set-based finite-volume IB method* where the level-set function was used for the sharp representation of the time-wise varying solid–fluid interface, and for an efficient implementation of Neumann BC at the moving boundary. Furthermore, the level-set function was used for the efficient computation of

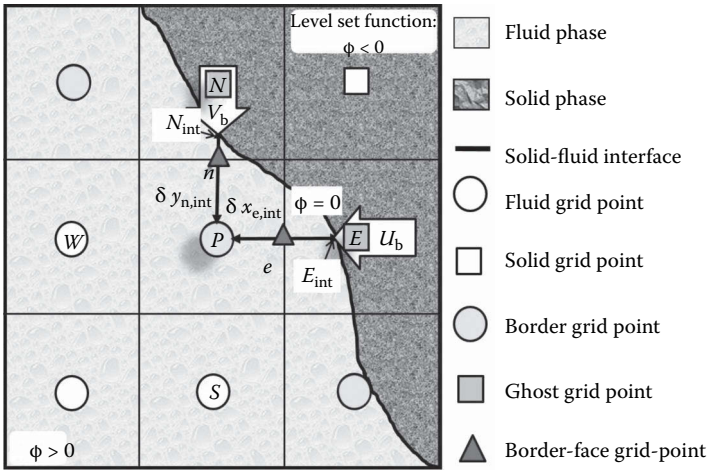


FIGURE 5.6.5 Level set-based representation of the interface and identification of the various types of grid points for the level set-based IB method [20].

the volume flux, advection flux, and diffusion flux during the implementation of the finite-volume method for the CV at a *border grid point* (Fig. 5.6.5)—the grid point in the fluid phase with at least one neighboring point in the solid phase.

For the computation of the various fluxes at those faces of the border CV adjoining the solid boundary, the solid wall BC is *directly* implemented, as discussed in the following. This is in contrast to that in the ghost-cell method, which involves an indirect implementation of the BC through the ghost grid point—not directly using the BC at the solid surface grid point (Fig. 5.6.3). Considering both the border and the ghost grid points, called interfacial points, Marella et al. [44] presented a level set-based IB method.

The interface along with the different types of grid points considered in the level set-based IB methods [20] are shown in Figure 5.6.5. For a border CV, the figure shows $\delta x_{e,int}$ and $\delta y_{n,int}$ as the distances used to compute the diffusion flux at the east and north face center—defined as the border face grid point. For these face centers, the solid wall velocity, U_b and V_b, shown in the figure are considered for the computation of momentum flux and volume flux during the solution of the momentum and continuity equations, respectively. Furthermore, during the solution of the pressure Poisson equation at the border face grid point, they [20] considered the temporal variation of the solid wall velocity and the distances ($\delta x_{e,int}$ and $\delta y_{n,int}$) while implementing a normal gradient BC for the pressure at the solid wall during the solution of the pressure Poisson equation. Thus, the solid wall BC is implemented directly in the level set-based IB method as opposed to the indirect implementation of BC in the ghost-cell-based IB method.

For the LS-based IB method, Shrivastava et al. [20] used the finite-volume method for the Navier–Stokes equations and the finite-difference method for the level-set equations; and a fully implicit and explicit method for the solution of the respective equations on a colocated Cartesian grid. They used the projection method for the solution of the pressure. For the spatial discretization, they used the fifth order Weighted Essentially Non-Oscillatory (WENO) scheme for the level-set equations, Quadratic Upstream Interpolation for Convective Kinematics (QUICK) scheme for the advection term, and central-difference scheme for the diffusion term of the momentum equations.

Numerical Modeling Issues for the IB Methods

Treatment of Moving Boundaries

Due to the motion of the solid–fluid interface in a fluid domain, a grid point that is in the solid at the previous time level n, emerges in the fluid at the present time level $(n+1)$; or vice versa [27]. This is

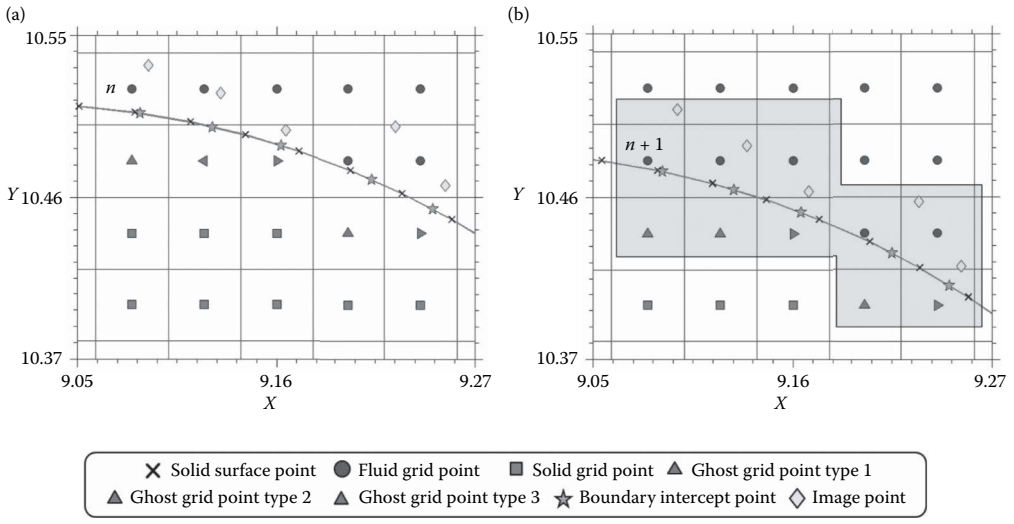


FIGURE 5.6.6 Boundary position and different points at (a) nth and (b) $(n+1)$th time step, for a downward motion of the solid body [26].

shown in Figure 5.6.6 for a downward motion of the interface, along with the time-wise change in the identity of the various grid points: ghost (of three different types), image, and boundary intercept. Note that the grid points that change their identity are shown inside the enclosed region in the figure, for the present time level $(n+1)$.

During the time interval Δt, the change in the identity of grid points from the solid phase to the fluid phase is called as the freshly cleared grid points. They are shown in Figure 5.6.7, where a freshly cleared grid point in the present time level $(n+1)$ was a ghost grid point in the previous time level n. The previous time level value of the flow property, needed for the computation of the property at the

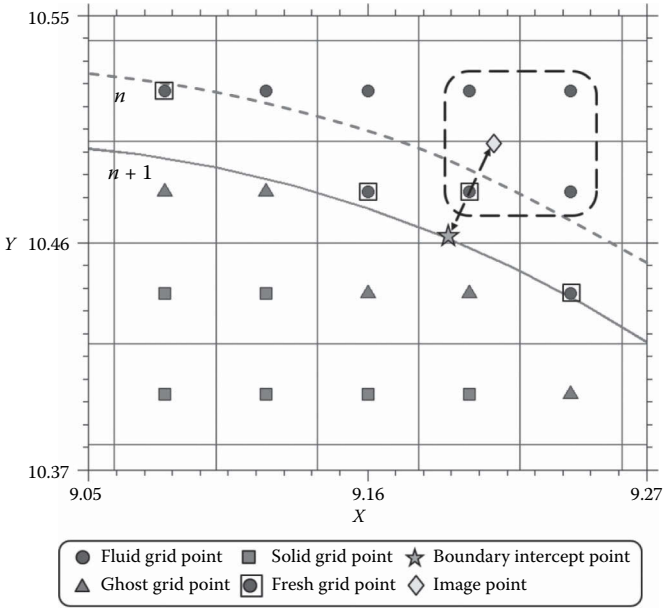


FIGURE 5.6.7 Appearance of freshly cleared grid points, after a time step, due to the motion of the boundary; and the associated computational stencil for the interpolation [26].

present time level, is obtained by the bilinear interpolation [27]; the computational stencil for the interpolation procedure is also shown in Figure 5.6.7. This treatment of the freshly cleared cell is known to result in spurious pressure oscillations. Seo and Mittal [39] used a cut-cell-based approach to enforce geometric conservation near the immersed boundary, which reduces the spurious oscillations.

Coupling with a Structure Dynamics Solver

Many previous studies have either considered rigid moving structure boundaries or ignored the internal stresses in the structure. In the latter case, a membrane of zero mass was introduced for the sake of computational tractability. However, if the structure dynamics is important and coupled with fluid dynamics, the structure solver needs to be coupled with the IB-based fluid dynamics solver. In general, for the coupling of the two solvers, there are two approaches: monolithic approach and partitioned (or segregated) approach (see Figure 5.6.8a). In the monolithic approach, the governing equations for the flow and structure domains are discretized together, and the nonlinear system of equations is solved as a whole. In the partitioned approach, the existing flow and structures solvers are utilized. However, implementation of the data exchange between the two solvers in the partitioned approach is not trivial and in general, a two-way coupling is needed to ensure the numerical stability of the FSI solver. The data exchange and algorithm in the partitioned approach are shown using a schematic in Figure 5.6.8b and c.

In order to couple the fluid and structure solvers, two coupling methods are generally utilized—explicit (or weak, one-way) coupling, and implicit (or strong, two-way) coupling, which integrate the governing equations of the flow and the structure domain explicitly and implicitly in time, respectively. The explicit coupling is computationally inexpensive and may be subject to stability constraints, which depends on the structure–fluid density ratio (ρ_s/ρ_f) [30]. The implicit coupling is needed for cases where the structure will respond strongly even with small perturbations from the fluid and vice versa; it is computationally expensive.

Some examples utilizing such coupling schemes are described here. Zheng et al. [30] proposed an FSI method, which weakly couples a sharp-interface IB method-based flow solver with a

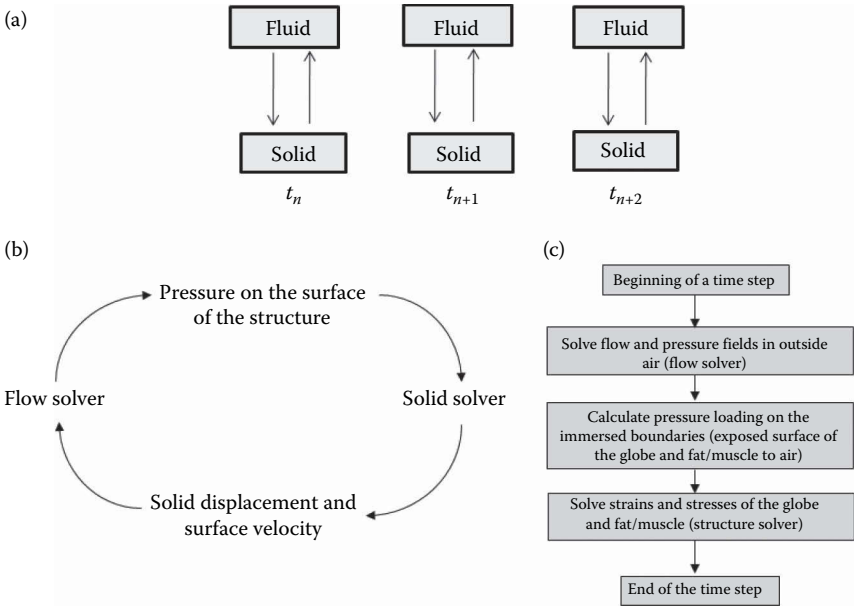


FIGURE 5.6.8 (a) Partitioned approach, (b) data exchange between flow and solid solver, and (c) algorithm of FSI solver. (Copyright Springer (2013). Reprinted with permission from Bhardwaj, R., Zeigler, K., Seo, J.H., Ramesh, K.T., and Nguyen, T.D., *Biomech. Model. Mechanobiol.*, 13, 123–140, 2014.)

finite-element structure solver. In this study, the constitutive model for the structure was based on the Kelvin–Voigt linear viscoelastic material [45]. However, this FSI solver was limited to handle small deformations of the structure and did not account for the geometric and material nonlinearity of the structure. Vigmostad et al. [6] proposed an FSI model for investigating flow fields in heart valve dynamics, using a locally refined Cartesian mesh with the sharp-interface IB method. This study utilized the implicit partitioned approach, and the 2D rod theory was used to model the structure. Recently, Bhardwaj and Mittal [46] and Tian et al. [47] have shown validation of their respective implicit partitioned IB-based solvers against an FSI benchmark proposed by Turek and Hron [48]. The FSI benchmark corresponds to the development of a self-sustained oscillation of an elastic plate attached to the lee side of a rigid cylinder, subjected to a 2D laminar flow inside a channel.

Applications of IB Methods for Moving Rigid Structures

Flow across a Linearly Oscillating Cylinder in a Stationary Fluid

The linear harmonic oscillation of a cylinder is given as $x=-A\sin(2\pi ft)$, where A is the amplitude and f is the frequency of oscillation. The diameter D and the maximum velocity u_{max} of the cylinder are taken as the nondimensional length and velocity scales, respectively. Other than the Reynolds number $\mathrm{Re}=u_{max}D/\nu$, the Keulegan–Carpenter number $K = 2\pi A/D$ is also a governing parameter here.

This problem is the widely used benchmark moving boundary problem to test the efficacy of an IB method. The numerical results are compared with the experimental and numerical results, reported by Dutsch et al. [49]. The results considered for the code validation are the variation of axial and transverse velocities at different axial locations of $-0.6D$, $0.0D$, $0.6D$, and $1.2D$, for different phase angles of the oscillating cylinder, 180°, 210°, and 330°. The benchmark result corresponds to a Reynolds number of 100 and a Keulegan–Carpenter number of 5. The code validation can be found in Ref. [20] for the LS-based IB method.

Flow across a Fish-Like Stream-Wise Undulating Foil

Using the LS-based IB method, Shrivastava et al. [20, 50] carried out 2D simulations for the free-stream flow across an undulating fish-like body modelled as an NACA0012 hydrofoil. The free-stream velocity u_∞ and chord length of the NACA0012 hydrofoil C are taken as the nondimensional velocity and length scales, respectively. The shape of the hydrofoil is transformed into a fish-like body by adding time-varying undulations (ΔY), at various X locations from the head of the hydrofoil along the axis, given as

$$\Delta Y = A(X)\sin\left\{2\pi\left(\frac{X}{\lambda}-\frac{\mathrm{St}\,\tau}{2A_{max}}\right)\right\}, \tag{5.6.5}$$

where $\mathrm{St}=\dfrac{F2A_{max}}{u_\infty}$ and A_{max} is the maximum amplitude of oscillation at the tail, X is the distance from the head of the hydrofoil along the axis, λ is the wavelength of the wave traveling across the hydrofoil, F is the frequency of undulation, τ is the nondimensional time, and St is the Strouhal number. Thus, other than the Reynolds number ($\mathrm{Re}=\rho u_\infty C/\mu$), the other governing parameters for a flow generated by a fish-like locomotion are equations of motion for undulation $A(X)$, wavelength of the wave traveling across the hydrofoil λ, frequency of undulation F, and maximum amplitude of oscillation A_{max}.

The results obtained by Shrivastava [50] for the free-stream flow across single, two in tandem, three side by side and school of fish-like undulating foil are shown in Figure 5.6.9, for vorticity contours. The figures show a reversed von Karman street, which acts as a signature for the generation of the thrust.

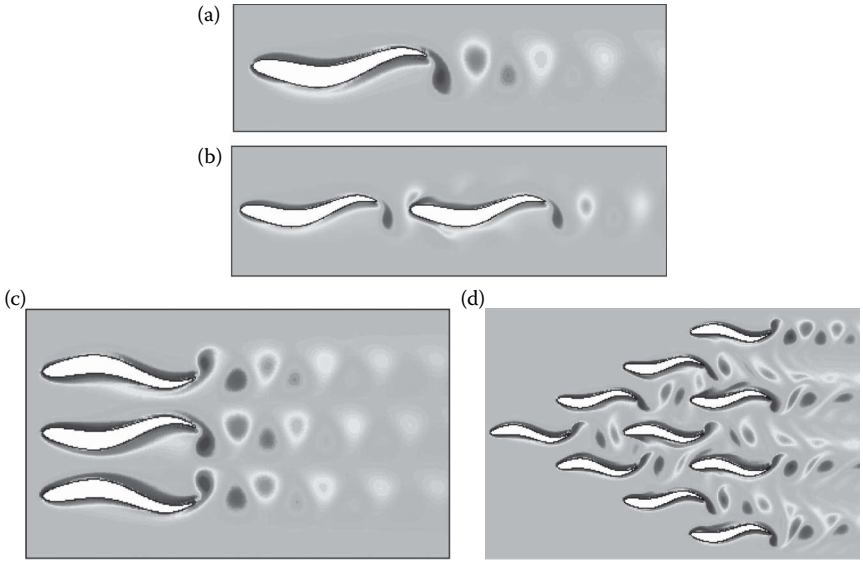


FIGURE 5.6.9 Vorticity contour for a free-stream flow across a (a) single, (b) two in tandem, (c) three side by side, and (d) school of fish-like undulating foil [50], at a Reynolds number of (a–c) 400 and (d) 1000. The other governing parameters are $\lambda = 1$, $A_{max} = 0.1$, and the Strouhal number of (a, c, d) 0.8 and (b) 0.6.

Applications of IB Methods with Moving Deformable Structures

Fluttering of an Elastic Flag Attached to a Rigid Cylinder

An FSI benchmark involving large-scale flow-induced deformation was first proposed by Turek and Hron [48], in which an elastic plate attached to the lee side of a rigid cylinder develops self-sustained oscillations in 2D laminar channel flow (shown in Figure 5.6.10). The circular cylinder is rigid and stationary while the elastic plate is deformable. The fluid is considered to be Newtonian and incompressible, while the structure is assumed to be elastic and compressible. The constitutive law for the structure is chosen as the Saint Venant–Kirchhoff law, in which the elasticity of the structure is characterized by the Poisson ratio and the Young modulus. The plane strain condition and geometric nonlinearity were considered in the structural solver.

The nondimensionalization of the parameters is based on the mean inflow velocity U_{mean} and cylinder diameter D. The Reynolds number and dimensionless Young's modulus are therefore defined

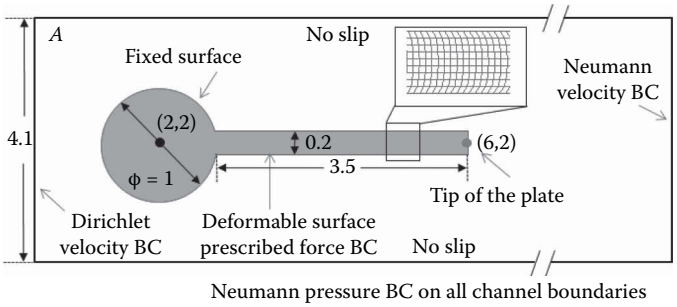


FIGURE 5.6.10 Schematic and BC of the benchmark problem. All numbers shown are dimensionless with respect to the cylinder diameter. The inset shows the finite-element mesh for the plate. (Copyright Prof. R. Mittal (2011). Reprinted with permission from Bhardwaj, R. and Mittal, R., *AIAA J.*, 50, 1638–1642, 2012.)

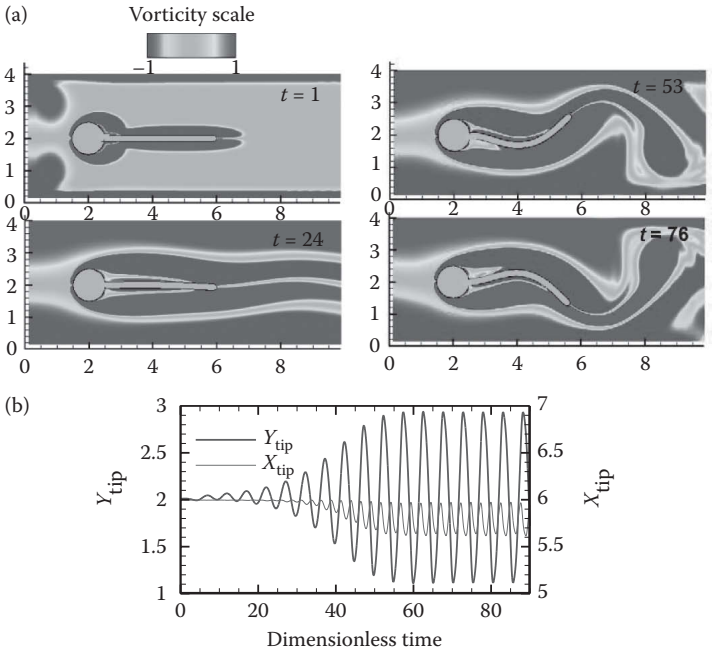


FIGURE 5.6.11 (a) The vorticity field and deformation of the elastic plate shown at different times for the baseline case. (b) Temporal variation of x and y displacements of the tip of the plate. (Copyright Prof. R. Mittal (2011). Reprinted with permission from Bhardwaj, R. and Mittal, R., *AIAA J.*, 50, 1638–1642, 2012.)

as $\mathrm{Re} = U_{\mathrm{mean}} D/\nu$ and $E = E^*/\rho_f U^{-2}$, where ν is the kinematic viscosity of the fluid [$m^2 s^{-1}$] and E^* is the Young's modulus [Nm^{-2}]. A parabolic inflow velocity profile is prescribed at the left boundary of the channel [48]:

$$v(0, y) = 1.5 U_{\mathrm{mean}} \frac{4y(H - y)}{H^2}, \tag{5.6.6}$$

where H is channel height. The usual no-slip BC are applied at the top and bottom walls. A homogeneous normal zero gradient velocity boundary condition is applied at the outflow. The other relevant BC are shown in Figure 5.6.10.

Figure 5.6.11a shows the calculated vorticity field and deformation of the plate at different time instances in Ref. [46]. Due to fluid dynamic forces, the plate starts to deform and attains a self-sustained periodic oscillation with plateau amplitude after a short time. Due to the motion of the plate, vortices shed from the tip of the plate with frequency of the order of the oscillation frequency of the plate. The time-varying y and x displacements of the plate's tip are shown in Figure 5.6.11b. The frequency of oscillation of x_{tip} is about twice of that for the y_{tip} because the plate oscillates with respect to the x-axis. The calculated values of the time period and the maximum plate displacement in Ref. [46] are about 5.155 and 0.917, respectively.

Thermal Augmentation Using Vortex-Induced Vibration

FSI of flexible thin structures immersed in a flow coupled with convective heat transfer enhances thermal transport, which could be utilized to improve the thermal performance of engineering as well as biological systems [13,14]. Soti et al. [13] numerically demonstrated large-scale flow-induced deformation as an effective passive heat-transfer enhancement technique. The authors utilized an in-house, strongly coupled FSI solver in which the flow and structure solvers are based on

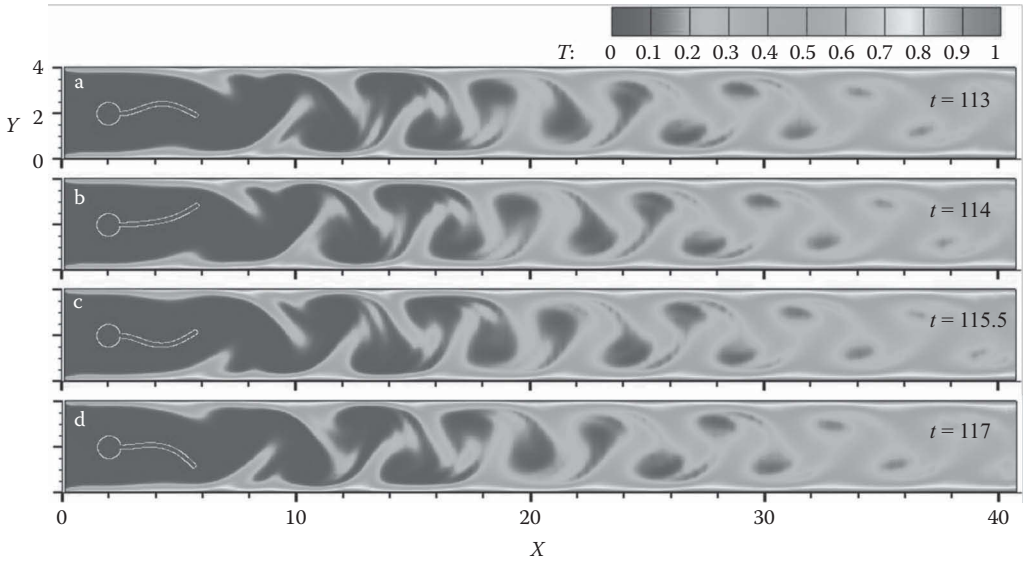


FIGURE 5.6.12 Isotherms in a channel with a cylinder attached to an elastic plate at different time instances for Re = 100 and Pr = 1. (Copyright Elsevier (2014). Reprinted with permission from Soti, A. K., Bhardwaj, R., Sheridan, J. 2015., *Int. J. Heat Mass Trans.*, 84, 1070–1081.)

the sharp-interface immersed boundary and finite-element method, respectively. They considered convective heat transfer in the FSI benchmark problem proposed by Turek and Hron [48]. The isotherms in the channel with a cylinder attached to an elastic plate at different time instances are plotted in Figure 5.6.12. The thermal augmentation is described as follows. The wake vortices past the plate sweep higher sources of vorticity generated on the channel walls out into the high-velocity regions—promoting the mixing of the fluid. The self-sustained motion of the plate assists in convective mixing, augmenting convection in bulk and near the walls, and thereby reducing the thermal boundary layer thickness as well as improving the value of the Nusselt number at the channel walls. In this study, the heat-transfer enhancement was quantified with respect to the pumping power needed and the effects of Reynolds number, Prandtl number, and material properties of the plate were investigated.

Blast Loading on Deformable Human Eye

A sharp-interface immersed boundary method-based compressible flow solver was employed by Bhardwaj et al. [8] to evaluate the influence of facial features on blast wave loading on the human eye. The flow field around the structure is compressible, and pressure fields on the structure are nonuniform and highly transient. The method proposed by Bhardwaj et al. [8] consists of the following components:

- Finite difference compressible flow solver for the propagation of the blast wave
- Finite element elasto-dynamic solver with finite deformation of the globe
- Sharp-interface immersed boundary method for FSI

Bhardwaj et al. [8] simulated the propagation of the blast wave in 3D coordinates to obtain the transient flow fields and pressure loading on the human eye. The flow fields in the sagittal plane are shown in Figure 5.6.13a, and the extensive reflections of the blast wave by the nasal and brow ridges are also shown in this figure. The blast wave reflections also produced an asymmetric pressure

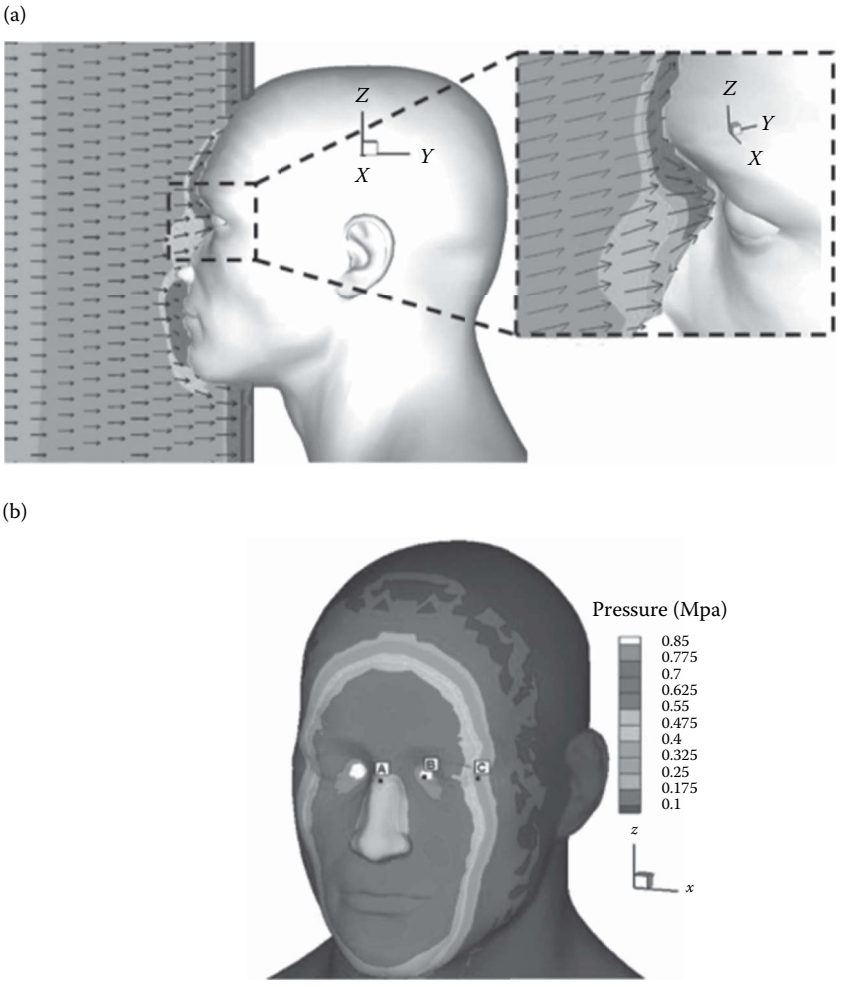


FIGURE 5.6.13 (a) Velocity vectors at every 10th grid point and pressure contours are shown in the sagittal plane on a human face in the event of a 2.5 kg TNT blast. Zoomed in view of the brow area shows the blast wave reflections. (b) Pressure contours on the face at the instance of the maximum pressure loading during the impact of the blast wave. (Copyright Springer (2013). Reprinted with permission from Bhardwaj, R., Zeigler, K., Seo, J.H., Ramesh, K.T., and Nguyen, T.D., *Biomech. Model. Mechanobiol.*, 13, 123–140, 2014.)

loading on the eye (Figure 5.6.13b) that was higher, close to the nasal bone than the zygomatic bones. This study also predicted possible eye injuries and associated injury risks using the calculated intraocular pressure (IOP) and von Mises stress in the sclera.

Concluding Remarks

The continuous forcing approach is very attractive especially for flows involving elastic moving boundaries due to the ease of implementation. However, this approach treats the physically sharp interface as a diffused interface. Furthermore, this approach is generally not well suited numerically for rigid structures and/or at high Reynolds numbers. Hence, it is difficult to use the continuous forcing approach for flows with rigid bodies. Moreover, since the governing equations are solved for both fluid and solid, in this approach, there is an additional computational effort involved.

On the other hand, since BC are implemented directly in the discrete equations, numerical implementation is not so straightforward in the discrete as compared to the continuous forcing

approach. This approach allows a sharp representation of the immersed boundary and is generally more accurate for moving rigid as well as deformable boundaries. However, there is an issue of spurious pressure oscillations due to the freshly cleared cells created by moving boundaries in this approach. In order to reduce these oscillations, additional interpolations near the immersed boundary are required.

While solving the structure dynamics with the flow field in the aforementioned approaches, the numerical stability depends on the structure–fluid density ratio. The implicit coupling between the flow and structural solver may be employed to ensure the numerical stability at low-density ratio.

ACKNOWLEDGMENTS

We would like to thanks our students involved in the present developments and applications: Dr. Mukul Shrivastava, Mr. Harshil Patel, and Mr. Atul K. Soti.

REFERENCES

1. Mittal, R., and Iaccarino, G. 2005. Immersed boundary methods. *Annual Reviews of Fluid Mechanics*, 37: 239–261.
2. Mittal. R., Erath, B. D., and Plesniak, M. W. 2013. Fluid dynamics of human phonation and speech. *Annual Reviews of Fluid Mechanics*, 45: 437–467.
3. Mittal, R., Seo, J. H., Vedula, V., Choi, Y. J., Liu, H., Huang, H. H., Jain, S, Younes, L., Abraham, T., and George, R. T. 2016. Computational modeling of cardiac hemodynamics: current status and future outlook. *Journal of Computational Physics*, 305: 1065–1082.
4. Bozkurttas, M., Tangorra, J., Lander, G., and Mittal, R. 2008. Understanding the hydrodynamics of swimming: from fish fins to flexible propulsors for autonomous underwater vehicles. *Advances in Science and Technology,* 58: 193–202.
5. Dong, H., Bozkurttas, M., Mittal, R., Madden, P. and Lauder, G. V. 2010. Computational modeling and analysis of the hydrodynamics of a highly deformable fish pectoral fin. *Journal of Fluid Mechanics*, 645: 345–373.
6. Vigmostad, S. C., Udaykumar, H. S., Lu, J. and Chandran, K. S. 2010. Fluid–structure interaction methods in biological flows with special emphasis on heart valve dynamics. *International Journal for Numerical Methods in Biomedical Engineering*, 26: 435–470.
7. Mengesha, T. E., Vallance, R. R., Barraja, M. and Mittal, R. 2009. Parametric structural modeling of insect wings. *Bioinspiration and Biomimetics*, 4: 036004.
8. Bhardwaj, R., Zeigler, K., Seo, J.H., Ramesh, K.T., and Nguyen, T.D. 2014. A computational model of blast loading on human eye. *Biomechanics and Modeling in Mechanobiology*, 13: 123–140.
9. Bailoor, S., Bhardwaj, R., Nguyen, T.D. 2015. Effectiveness of eye armor during blast loading. *Biomechanics and Modeling in Mechanobiology*, 14: 1227–1237.
10. Beeby S. P., Tudor, M. J., and White, N. M. 2006. Energy harvesting vibration sources for microsystems applications. *Measurement Science and Technology* 17: 175–195.
11. Weissenbock, N.M., Weiss, C.M., Schwammer, H.M., and Kratochvil, H. 2010. Thermal windows on the body surface of African elephants (*Loxodonta africana*) studied by infrared thermography. *Journal of Thermal Biology*, 35: 182–188.
12. Mills, Z.G., Aziz, B., and Alexeev, A., 2012. Beating synthetic cilia enhance heat transport in microfluidic channels. *Soft Matter*, 8: 11508–11513.
13. Soti, A. K., Bhardwaj, R., Sheridan, J. 2015. Flow-induced deformation of a elastic plate as manifestation of heat transfer enhancement in laminar channel flow. *International Journal of Heat and Mass Transfer*, 84: 1070–1081.
14. Joshi, R. U., Soti, A., Bhardwaj, R. 2015. Numerical study of heat transfer enhancement by deformable twin plates in laminar heated channel flow. *Computational Thermal Sciences*, 7(1): 10.
15. Dunne, T. and Rannacher, R. 2006. Adaptive finite element approximation of fluid-structure interaction based on an Eulerian variational formulation. *Lecture Notes in Computational Science and Engineering* (Eds: Bungartz H.-J. and Schaefer M.), Springer Verlag, Berlin, Germany, 110–145.
16. Hron, J. and Turek, S. 2006. A monolithic FEM/multigrid solver for an ALE formulation of fluid-structure interaction with applications in biomechanics. *Lecture Notes in Computational Science and Engineering* (Eds: Bungartz H.-J. & Schaefer M.), Springer Verlag, Berlin, Germany, 146–170.

17. Sahin, M. and Mohseni, K. 2009. An arbitrary Lagrangian–Eulerian formulation for the numerical simulation of flow patterns generated by the hydromedusa Aequorea victoria. *Journal of Computational Physics* 228: 4588.
18. Tezduyar, T.E., Sathe, S., Stein, K., and Aureli, L. 2006. Modeling of fluid-structure interactions with the space-time techniques. *Lecture Notes in Computational Science and Engineering* (Eds: Bungartz, H.-J. and Schaefer, M.), Springer Verlag, Berlin, Germany, 50–81.
19. Souli, M., Ouahsine, A. and Lewin, L. 2000. ALE formulation for fluid-structure interaction problems. *Computer Methods in Applied Mechanics and Engineering* 190: 659–675.
20. Shrivastava, M., Agrawal, A., and Sharma, A., 2013. A novel level set-based immersed boundary method for CFD simulation of moving boundary problems. *Numerical Heat Transfer-B*, 63: 304–326.
21. Peskin, C. S. 1972. Flow patterns around heart valves: a numerical method. *Journal of Computational Physics*, 10: 220–252.
22. Peskin, C. S. 1977. Numerical analysis of blood flow in the heart. *Journal of Computational Physics*, 25: 220–252.
23. Peskin, C. S. 2002. The immersed boundary method. *Acta Numerica*, 11: 1–39.
24. Hou, G., Wang, J., Layton, A., 2012. Numerical methods for fluid-structure interaction. *Communications in Computational Physics*, 12: 337.
25. Luo, H., Mittal, R., Zheng, X., Bielamowicz, S. A., Walsh, R. J., and Hahn, J. K. 2008. An immersed-boundary method for flow-structure interaction in biological systems with application to phonation. *Journal of Computational Physics*, 227: 9303–9332.
26. Patel, H. V., 2015. Development of fluid structure interaction solver using ghost cell based immersed boundary method, M.Tech. Thesis, IIT Bombay.
27. Mittal, R., Dong, H., Bozkurttas, M., Najjar, F.M., Vargas, A. and vonLoebbeck, A. 2008. A versatile immersed boundary methods for incompressible flows with complex boundaries. *Journal of Computational Physics*, 227: 10.
28. Seo, J. H. and Mittal, R. 2011. A higher-order immersed boundary method for acoustic wave scattering and low-Mach number flow-induced sound in complex geometries. *Journal of Computational Physics*, 230: 1000–1019.
29. Zheng, X., Bielamowicz, S., Luo, H., and Mittal, R. 2009. Computational study of the effect of false vocal folds on glottal flow and vocal fold vibration during phonation. *Annals of Biomedical Engineering*, 37: 625–642.
30. Zheng, X., Xue, Q., Mittal, R. and Beilamowicz, S. 2010. A coupled sharp-interface immersed boundary-finite-element method for flow-structure interaction with application to human phonation. *Journal of Biomechanical Engineering*, 132: 111003.
31. Xu, S. and Jane Wang, Z. 2006. An immersed interface method for simulating the interaction of a fluid with moving boundaries. *Journal of Computational Physics* 216(2): 454–493.
32. Goldstein, D., Handler, R., Sirovich, L. 1993. Modeling a no-slip flow boundary with an external force field. *Journal of Computational Physics* 105: 354–366.
33. Zang, Y., Streett, R. L., and Koseff, J. R. 1994. A non-staggered fractional step method for time-dependent incompressible Navier–Stokes equations in curvilinear coordinates. *Journal of Computational Physics* 114: 18–33.
34. Ye, T., Mittal, R., Udaykumar, H.S., Shyy, W. 1999. An accurate Cartesian grid method for viscous incompressible flows with complex immersed boundaries. *Journal of Computational Physics* 156: 209–240.
35. Clarke, D., Salas, M., Hassan, H. 1986. Euler calculations for multi-element airfoils using Cartesian grids. *AIAA Journal* 24: 1128–1135.
36. Udaykumar, H.S., Mittal, R., Shyy, W. 1999. Computation of solid-liquid phase fronts in the sharp interface limit on fixed grids. *Journal of Computational Physics* 153: 534–574.
37. Udaykumar, H. S., Mittal, R., Rampunggoon, P., and Khanna, A., 2001. A sharp interface Cartesian grid method for simulating flows with complex moving boundaries. *Journal of Computational Physics*, 174: 345–380.
38. Udaykumar, H. S., Mittal, R., and Rampunggon, P. 2002. Interface tracking finite volume method for complex solid-fluid interactions on fixed meshes. *Communications in Numerical Methods in Engineering*, 18: 89–97.
39. Seo, J. H. and Mittal, R. 2011. A sharp-interface immersed boundary method with improved mass conservation and reduced spurious pressure oscillations. *Journal of Computational Physics,* 230: 7347–7363.

40. Sussman, M., Smereka, P. and Osher, S., 1994. A level set approach for computing solutions to incompressible two-phase flow. *Journal of Computational Physics*, 114: 146–159.
41. Choi, J. I., Oberoi, R. C., Edwards, J.R. and Rosati, J.A., 2007. An immersed boundary method for complex incompressible flows. *Journal of Computational Physics*, 224: 757–784.
42. Tan, Z., Lim, K.M. and Khoo, B.C., 2010. A level set-based immersed interface method for solving incompressible viscous flows with the prescribed velocity at the boundary. *International Journal for Numerical Methods in Fluids*, 62: 267–290.
43. Cheny, Y. and Botella, O., 2010. The LS-STAG method: A new immersed boundary/level-set method for the computation of incompressible viscous flows in complex moving geometries with good conservation properties. *Journal of Computational Physics*, 229: 1043–1076.
44. Marella, S., Krishnan, S., Liu, H. and Udaykumar, H.S., 2005. Sharp interface cartesian grid method I: An easily implemented technique for 3D moving boundary computations. *Journal of Computational Physics*, 210: 1–31.
45. Fung, Y. C. 1965. *Biomechanics*, 2nd ed., Englewood Cliffs: Prentice-Hall.
46. Bhardwaj, R. and Mittal, R., 2012. Benchmarking a coupled immersed-boundary-finite-element solver for large-scale flow-induced deformation. *AIAA Journal*, 50: 1638–1642.
47. Tian, F. B., Dai, H., Luo, H., Doyle, J. F., Rousseau, B. 2014. Fluid-structure interaction involving large deformations: 3D simulations and applications to biological systems. *Journal of Computational Physics* 258: 451–569.
48. Turek, S. and Hron, J. 2006. Proposal for numerical benchmarking of fluid-structure interaction between an elastic object and a laminar incompressible flow. *Lecture Notes in Computational Science and Engineering* (Eds: Bungartz, H.-J. and Schaefer, M.), Springer Verlag.
49. Dutsch, H., Becker, S., and Lienhart, H., 1998. Low-Reynolds-number flow around an oscillating circular cylinder at low Keulegan-Carpenter numbers, *Journal of Fluid Mechanics*, 285: 249–271.
50. Shrivastava, M., 2015. A novel level set method based immersed boundary method and its application to hydrodynamics of fish-like locomotion of single and in-group flexible hydrofoils, Ph.D. Thesis, IIT Bombay.

5.7 NUMERICAL METHODS FOR MULTIPHASE FLOWS

Shyamprasad Karagadde and Atul Sharma

Introduction

Multimaterial interfaces invariably exist in studies involving phase-change problems and dispersed flows. These interfaces or phase boundaries introduce discontinuities in thermophysical properties and matter. Therefore, numerical solutions require additional attention to accurately track the location and shape of evolving interfaces. Moreover, the evolution of these interfaces is primarily governed by heat and mass transfer across the interfaces, which themselves are functions of both the field variables and their gradients. Evaluation of these gradients at the discontinuity (interface) poses several numerical challenges and thus needs special treatment. Two well-known phase-change problems, namely (1) boiling and (2) solidification, are used as examples throughout the remainder of this chapter. During boiling phenomena, vapor bubbles undergo growth due to the vaporization of the liquid and are also transported under the influence of forced flow or buoyancy. Some of the key aspects that dominate the evolution of the vapor bubbles are surface tension, latent heat, density, and viscosity, which have strong variation around the bubble–liquid interface. Further complications are introduced by the presence of Marangoni convection, a rigid solid wall, etc. In solidification problems, additional species conservation and differential solubility of the species in solid and liquid phases pose an added difficulty. Furthermore, when simulating microstructures, the solid–liquid interface often undergoes anisotropic growth and can be highly complex in structure.

Since multiphase flows consist of a large variation in the scales of interfaces, various classes of numerical methods have been developed in the literature. For example, in a bubble column, the scale of the interface can vary from the diameter of the bubble column to that of the smallest bubble. Such problems are simulated by a numerical method, such as two-fluid model (Wörner, 2003), which uses empirical correlations—consisting of certain modeling constants, which are determined from experimental results. Whereas, an interface resolving method is used for the multiphase flow systems, which consist of interfaces in mini- and microscales, predominantly laminar flow, and lesser number of discontinuous dispersed phase (Wörner, 2012). For such multiphase flow systems, the *scope* of this chapter is to present three interface resolving numerical methods: first, enthalpy method; second, volume of fluid (VOF) method; and third, level set method (LSM). The third method is recent as compared with the other two methods. The first is a popular method for solid–liquid phase change (melting/solidification) and the other two methods are commonly used for liquid–gas (boiling/condensation) phase-change processes, but they have also been applied to microstructure evolution during liquid to solid transformation.

The outline of this chapter is as follows: it starts with a discussion of various methods for the continuum modeling of the moving interfaces. Thereafter, a common mathematical formulation is presented for different numerical methods considered here. Finally, the three methods are presented starting with enthalpy method, followed by VOF method, and ending with the LS method (LSM). Each of the following subsections on these methods consists of their formulation, numerical methodology, and a representative application for each of the methods.

Continuum Modeling of Moving Interfaces

When heat and fluid flow equations are conventionally solved on a fixed grid in the space–time domain, the boundary/interface will, in general, be between two grid points at any given time. Therefore, special formulae are needed to evaluate the gradients of variables at the interfaces and also to account for variation in grid sizes. Alternatively, the grid itself needs to be deformed or variables to be transformed so that the interface is always on a grid line or is fixed in the transformed domain. This leads to a classification based on the methodology of characterizing the interface,

namely (1) front tracking, (2) front fixing, and (3) fixed grid methods (Crank, 1987). Front tracking methods compute the position of the interface at each step in time. Typically, front tracking methods involve field solutions on a fixed Eulerian background mesh and Lagrangian reconstruction of the interface, which is located on the fixed mesh. This interface is tracked in time by explicitly satisfying discrete forms of the interface mass, momentum, and energy balance conditions (Voller, 2008). A major drawback of this approach has been the difficulties in achieving the required level of accuracy when the interfaces evolve at much faster rates or develop complex morphologies. Furthermore, under certain conditions, the interface can also be fixed by a suitable choice of new coordinates. These approaches have found limitations in employing the method to nonanalytical interface shapes and in developing a common numerical solution platform that can address a range of multiphase flow problems. Furthermore, it may even be difficult to track the moving interface if the evolution is not smooth or monotonic with time, and particularly when sharp peaks exist or when they disappear.

Alternatively, fixed grid or fixed domain techniques have shown remarkable advantages in solving a variety of complex multiphase flow problems. These typically involve either explicit or implicit "capture" of interfaces using a bounded variable known as "order parameter" or "phase indicator" φ, $a \leqslant \varphi \leqslant b$. A nodal value of $\phi_P = a,b$ indicates the node point P in single-phase regions, and, a value $a < \varphi_P < b$ indicates that P is in the vicinity of the interface. They can also be used to evaluate interface properties such as position and curvature. Typically, the application of these methods requires differential equations to describe the conservation laws and the evolution of the order parameter. The well-known principles of volume averaging are applied for the governing equations to obtain "single-phase" formulation, and the solution of variables (velocity, pressure, temperature, concentration, etc.) is obtained by the various discretization and numerical solution techniques.

Hyman (1984) proposed a distinction based on the increasing flexibility and computational complexity. Accordingly, these are categorized into surface tracking methods, volume-tracking methods, and moving mesh (front tracking) methods. Surface tracking involves interpolation between marker particles along the interface to obtain the interface location, which is similar to the front tracking approach. Accurate resolution of subgrid structures is relatively easy due to the low dimension of such methods. These have been observed to be relatively easy to implement for simple interface evolution problems. Volume-tracking methods are particularly used where topology changes are present. The domain is subdivided into union of disjoint solution regions and the separating region is the interface, identified by marker points or volume fraction change based on the solution of an auxiliary evolution of a PDE. The interface can be reconstructed based on the phase or volume fraction information.

In the aforementioned approaches, the interface can be represented as either *sharp* or *diffuse* depending upon the width spanned by the interface. Sharp-interface representation allows the two-phase cell to be represented by only one computational grid across the interface. These methods accurately conserve mass and preserve the shape of the interface but need algorithms such as interface reconstruction, which contribute to the overall computational effort. In diffuse interface representation, the interface is spread over a few (about 2 or 3) cells around the actual interface and sharp variations in properties across the interface are smeared over this thickness. Although smearing reduces certain spurious errors (or instabilities) which occur during numerical computations, these methods are also prone to mass conservation problems. In general, for fixed domain methods, multiple phases are distinguished by an order parameter and a lower-dimensional evolution equation is solved to analyze the interface motion. This evolution equation is usually a function of the interfacial mass, momentum and heat transfer, which describes the "transport" of the interface (by advecting the phase-indicator variable).

The complete numerical solution of multiphase flows requires coupled solution of the evolution equation along with the conservation laws which may differ in respective phases (due to variation in material and thermophysical properties). The bridging of the different forms at the material interface is achieved by defining the appropriate interfacial mass, momentum and energy balance. It is generally helpful to implement the solution method if these "multifluid" relations of a particular

conservation can be merged together to obtain a single equation that is valid at every computational cell. Such forms are known as *single-field* (also referred as *single-fluid*) formulation, and are obtained by the help of volume-averaging principles (Wörner, 2003).

Mathematical Formulation

Principles of Volume Averaging

A brief overview of the volume-averaging principle and its application to governing equations in a multiphase system is presented below (Wörner, 2003). Consider a general representative volume element V bounded by surface ∂V, invariant with respect to time and space. Let the volume element contain two phases separated by a boundary. A phase-indicator function $\phi(\mathbf{x},t)$ is defined in such a way that it carries a value of 1 for phase k and 0 for the other. Let $\mathbf{x_c}$ denote the centroid of V. Any location in V is identified relative to the centroid, i.e., $\mathbf{x} = \mathbf{x}_c + \mathbf{z}$. The volume fraction of phase k within the volume element V is written as

$$f_k = \frac{1}{V}\int_V \phi(\mathbf{x},t)\mathrm{d}\mathbf{x} = \frac{V_k}{V} \tag{5.7.1}$$

Following expressions show definitions of the volume-averaging operators for any vector or scalar quantity ψ. The term ψ_k denotes the quantity in phase k.

$$\overline{\psi}_k^{\,k} = \frac{1}{V_k}\int_V \psi_k(\mathbf{x},t)\phi(\mathbf{x},t)\mathrm{d}\mathbf{x} \tag{5.7.2a}$$

$$\overline{\psi}_k^{\,V} = \frac{1}{V}\int_V \psi_k(\mathbf{x},t)\phi(\mathbf{x},t)\mathrm{d}\mathbf{x} \tag{5.7.2b}$$

There exists a linear relationship between the volume-averaged values of the quantity defined in the phase k (Equation 5.7.2a) and the overall domain (Equation 5.7.2b), i.e.,

$$\overline{\psi}_k^{\,V} = f_k \overline{\psi}_k^{\,k} \tag{5.7.3}$$

We assume that the flow is incompressible in each of the phases, and the density ρ_k in each phase k is constant. Fluid velocities in each phase are represented by $\mathbf{u}_k$. The local mass and momentum conservation equations are written as follows.

$$\frac{\partial \rho_k}{\partial t} + \nabla\cdot(\rho_k \vec{\mathbf{U}}_k) = 0 \tag{5.7.4}$$

$$\frac{\partial \rho_k \vec{\mathbf{U}}_k}{\partial t} + \nabla\cdot\rho_k\vec{\mathbf{U}}_k\vec{\mathbf{U}}_k = -\nabla p_k + \rho_k\mathbf{g} + \nabla\cdot\Omega_k \tag{5.7.5}$$

Multiplying the above equations by the respective phase-indicator function $\phi(\mathbf{x},t)$ and performing the volume averaging over V, we get,

$$\frac{\partial f_k\rho_k}{\partial t} + \nabla\cdot(f_k\rho_k\overline{\vec{\mathbf{U}}_k}^{\,k}) = \Gamma_k \tag{5.7.6}$$

$$\frac{\partial f_k\rho_k\overline{\vec{\mathbf{U}}_k}^{\,k}}{\partial t} + \nabla\cdot f_k\rho_k\overline{\vec{\mathbf{U}}_k\vec{\mathbf{U}}_k}^{\,k} = -\nabla f_k\overline{p_k}^{\,k} + f_k\rho_k\mathbf{g} + \nabla\cdot f_k\overline{\Omega_k}^{\,k} + M_k \tag{5.7.7}$$

The additional terms in the above expressions arise after simplification due to the volume averaging. Transfer of mass across the interface is represented by,

$$\Gamma_k = \overline{\rho_k(\vec{\mathbf{U}}_{ki} - \vec{\mathbf{U}}_i)\cdot\nabla\phi_k}^{V} \tag{5.7.8}$$

Due to mass conservation, jump condition is satisfied, $\Gamma_1 + \Gamma_2 = 0$. The momentum transfer across the interface is given by,

$$M_k = -\overline{(-p_k\mathrm{I} + \Omega_k)\cdot\nabla\phi_k}^{V} \tag{5.7.9}$$

The sum of M_1 and M_2 satisfies the pressure jump condition,

$$M_1 + M_2 = \frac{1}{V}\int_{S_i\cap V}(2\kappa\sigma\mathbf{n}_1 + \nabla_s\sigma)\mathrm{d}S \tag{5.7.10}$$

The second term inside the integral is zero when surface tension is assumed to be constant. By summing up the equations of conservation for individual phases, a single-field formulation can be written as.

$$\sum\left[\frac{\partial f_k\rho_k}{\partial t} + \nabla\cdot(f_k\rho_k\overline{\vec{\mathbf{U}}_k}^{k})\right] = 0 \tag{5.7.11}$$

Writing the mixture density and center-of-mass velocity for a two-phase system,

$$\rho_\mathrm{m} = f_1\rho_1 + f_2\rho_2 \tag{5.7.12}$$

$$\vec{\mathbf{U}}_\mathrm{m} = \frac{1}{\rho_\mathrm{m}}\sum_{k=1}^{2} f_k\rho_k\overline{\vec{\mathbf{U}}_k}^{k} \tag{5.7.13}$$

On simplification, we get the volume-averaged mass conservation equation as,

$$\frac{\partial\rho_\mathrm{m}}{\partial t} + \nabla\cdot(\rho_\mathrm{m}\vec{\mathbf{U}}_\mathrm{m}) = 0 \tag{5.7.14}$$

Similarly summing up and simplifying the equations for the momentum conservation and performing necessary approximations and assumptions, we obtain,

$$\begin{aligned}\frac{\partial\rho_\mathrm{m}\vec{\mathbf{U}}_\mathrm{m}}{\partial t} + \nabla\cdot(\rho_\mathrm{m}\vec{\mathbf{U}}_\mathrm{m}\vec{\mathbf{U}}_\mathrm{m}) &= -\nabla p_\mathrm{m} + \rho_\mathrm{m}\mathbf{g} + \nabla\cdot\mu_\mathrm{m}(\nabla\vec{\mathbf{U}}_\mathrm{m} + (\nabla\vec{\mathbf{U}}_\mathrm{m})^T) \\ &+ \frac{1}{V}\int_{S_i\cap V} 2\kappa\sigma\mathbf{n}\mathrm{d}S\end{aligned} \tag{5.7.15}$$

The last term on the right-hand side is often simplified to the form $\mathbf{f}_{\mathbf{sv}} = \sigma\kappa\hat{\mathbf{n}}\delta_\mathrm{s}$, where δ_s is the Dirac Delta function, representing the interfacial area per unit volume in a computational cell, and all the geometric terms in this expression are evaluated in terms of the phase-indicator function.

The energy and species conservation equations are also derived in a similar fashion to obtain the single-field governing equations, based on the interfacial heat and mass transfer that are specific to a particular process such as boiling and solidification, further details can be found elsewhere

(Wörner, 2012; Dantzig and Rappaz, 2009). For the remainder of this text, the phase volume fraction of a phase k is represented by the notation f_k. From mass conservation, the sum of phase fractions of all $n$phases in a cell is unity, i.e.,

$$\sum_{k=1}^{n} f_k = 1 \tag{5.7.16}$$

In general, any mixture property β (such as density, conductivity, and diffusivity) is represented by the volume-averaged expression

$$\beta = \sum_{i=1}^{n} \beta_i f_i \tag{5.7.17}$$

In this subsection, the notation for phases is denoted by subscripts l, s, and b representing liquid, solid, and bubble phases, respectively. Any combination of these subscripts represents the interface.

Nondimensional Governing Equations

For a single-phase flow, the well-known Navier–Stokes equations alone are solved to obtain the velocity and pressure fields. However, for two-phase flow, a governing equation for the interface is also solved using one of the interface resolving methods. The Navier–Stokes equations are coupled with the equation for the interface by invoking a *single-field formulation*, wherein a single velocity and pressure field is defined for both the fluids. Typically, constant material properties are assumed, but not equal for each phase, i.e., the bulk fluids are *incompressible*. Moreover, the interfacial mass transfer (due to phase change) and the surface tension force at interface is modeled as a volumetric source term in the continuity and momentum equation, respectively, these source terms are nonzero only at the interface. The surface tension coefficient is assumed to be constant and its tangential variation along the interface is neglected. It is assumed that interface is thin and massless with no slip in tangential velocity. The effect of radiation, viscous dissipation and the energy contribution due to interface stretching are neglected.

The analysis with dimensionless form of equations can be beneficial in studying a larger class of problems in multiphase flow and heat-transfer. The following discussion presents the dimensionless formulation of a typical boiling heat-transfer problem, and for the other cases, slight variations in the dimensionless numbers and scales may appear. The nondimensional single field governing equations for the VOF- and LSM-based simulation of multiphase flow are as follows:

Volume conservation (continuity) equation:

$$\nabla \cdot \vec{\mathbf{U}} = \left(\frac{1-\chi}{\chi}\right)\dot{M}\delta_\varepsilon \quad \left\{\dot{M} = \frac{\mathrm{Ja}}{\gamma \mathrm{Re}_l \mathrm{Pr}_l}[(\nabla\theta)_1 - \zeta(\nabla\theta)_2]\cdot\hat{\mathbf{n}}\right\} \tag{5.7.18}$$

The right-hand side of the above equation represents the change in volume of a phase ($\dot{M}$), which typically occurs during phase change or diffusion-driven phase growth. In problems involving purely advection of phases, this growth term would be zero.

Momentum Conservation Equation:

$$\frac{\partial(\rho_m \vec{U})}{\partial\tau} + \nabla\cdot(\rho_m \vec{\mathbf{U}}\vec{\mathbf{U}}) = -\nabla P + \frac{1}{\mathrm{Re}_l}\nabla\cdot(\mu_m(\nabla\vec{\mathbf{U}} + (\nabla\vec{\mathbf{U}})^T)) - \frac{\rho_m}{\mathrm{Fr}^2}\hat{j} + \frac{1}{\mathrm{We}}\kappa\hat{\mathbf{n}}\delta_\varepsilon(\phi) \tag{5.7.19}$$

where ρ_m and μ_m are the mean nondimensional density and viscosity, respectively.

Energy Conservation Equation:

$$\frac{\partial(\rho_m c_{pm}\theta)}{\partial\tau} + \nabla\cdot(\rho_m c_{pm}\vec{U}\theta) = \frac{1}{\mathrm{Re}_l \mathrm{Pr}_l}\nabla\cdot(k_m\nabla\theta) \tag{5.7.20}$$

where c_{pm} and k_m are the mean (or mixture) nondimensional specific heat and thermal conductivity, respectively.

For the above equations, using l_c as characteristic length and u_c as velocity scale, the nondimensional variables are defined as

$$\vec{U} = \frac{\vec{u}}{u_c},\ X = \frac{x}{l_c}, \tau = \frac{tu_c}{l_c}, P = \frac{p}{\rho_l u_c^2}, \theta = \frac{T - T_{SAT}}{T_W - T_{SAT}}, \dot{M} = \frac{\dot{m}}{\rho_l u_c}$$

where T_W is the constant wall temperature (as a BC), and T_{SAT} is the saturation temperature in a phase-change problem. Furthermore, $\dot{m}$ is the interfacial mass flux, obtained from the Stefan condition: $\dot{m} = (q_{int,2} - q_{int,l})/(h_2 - h_l)$—ratio of jump of conduction heat flux and jump of enthalpy at the interface (due to phase change). Furthermore, nondimensional governing parameters (Reynolds number, Re, Prandtl number, Pr, Froude number, Fr, Weber number, We, and Jakob number, Ja) are defined as

$$\mathrm{Re}_l = \frac{\rho_l u_c l_c}{\mu_l}, \mathrm{Pr}_p = \frac{\mu_l c_{p,l}}{k_l}, \mathrm{Fr} = \frac{u_c}{\sqrt{g l_c}}, \mathrm{We} = \frac{\rho_l u_c l_c^2}{\sigma}, \mathrm{Ja} = \frac{c_{p,2}(T_W - T_{SAT})}{h_2 - h_l} \tag{5.7.21}$$

Since the fluid-1 is considered as the reference fluid, note that the mean thermophysical properties in the above equation are all equal to unity for fluid-1, and are equal to the property ratio for the fluid-2; given as

$$\rho_m = \chi = \frac{\rho_2}{\rho_l}, \mu_m = \eta = \frac{\mu_2}{\mu_l}, k_m = \zeta = \frac{k_2}{k_l}, c_{pm} = \gamma = \frac{c_{p,2}}{c_{p,l}}$$

Although the surface tension force is a surface force, Brackbill et al. (1992) proposed a continuum surface force (CSF) model to represent it as a localized body force at interface—considered as the source term in Equation 5.7.19.

Enthalpy Method

During the phase change (liquid to solid) process for a pure substance, a sharp interface exists separating the liquid and solid regions. For multicomponent systems (e.g., alloys), the separation is represented by the mushy zone whose width is governed by the difference in solidus and liquidus temperatures, schematically shown in Figure 5.7.1. The first calculation of the heat flow during phase change was suggested by Eyres et al. (1946). Since then numerous modeling strategies have been developed and one of the most widely used techniques, particularly for macroscopic solidification modeling, is the enthalpy method (Voller et al., 1987, 1989). A significant advantage of the enthalpy methodology is that it does not involve a separate PDE for interface evolution, and the interface is implicitly captured by linking the temperature field with the enthalpy and liquid phase. This has been successfully implemented for a wide range of solidification problems involving alloy systems, melt convection, etc., and prediction of parameters (such as solidification time and macrosegregation) has been found to match closely with experiments.

Formulation

The principle of the enthalpy technique for pure metal solidification is as follows. The enthalpy is defined as $H = c_{p,l}T + f_l L_f$, where $c_{p,l}$ is the specific heat of the liquid, T is the temperature, and L_f

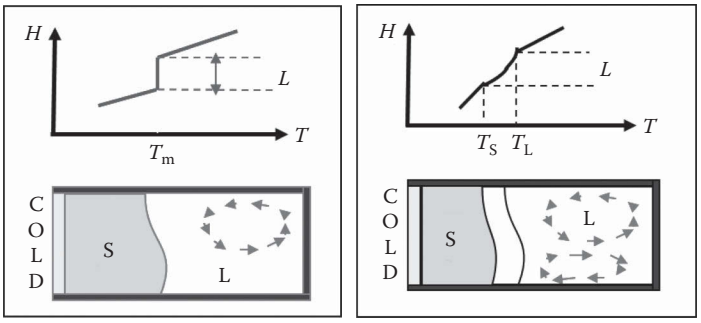


FIGURE 5.7.1 Schematic of solidification processes in (a) pure and (b) multicomponent system.

is the latent heat of fusion. The rejection of latent heat occurs at the solidifying interface, which is conducted away due to the cooling from the boundaries. The heat jump or the enthalpy is monitored to obtain the location of the interface. The classical Stefan's condition (interfacial heat balance) is given by the following relation:

$$\rho L_f \vec{\mathbf{V}}_{s-l} = k_s \nabla T\big|_{x=\text{solid}} - k_l \nabla T\big|_{x=\text{liquid}} \tag{5.7.22}$$

In this method, the volume-averaged energy conservation equation is modified to account for heat rejection at the interface, by indirectly imposing the Stefan's condition (Equation 5.7.22) via the single-field formulation, and is solved to obtain the interface location (Brent et al., 1988; Voller et al., 1989).

$$\frac{\partial(\rho H)}{\partial t} + \nabla \cdot (\rho \vec{\mathbf{U}} H) = \nabla . \left(\frac{k_m}{c_p} \nabla H \right) + S_h \tag{5.7.23}$$

where $k_m = f_s k_s + f_l k_l$ is the effective thermal conductivity of the mixture, and S_h is the source term given by the following relation.

$$S_h = \frac{\partial(\rho \Delta H)}{\partial t} + \nabla \cdot (\rho \vec{\mathbf{U}} \Delta H) \tag{5.7.24}$$

where $\Delta H = f_l L_f = f(T)$ is the latent heat content, defined as a function of temperature (for both pure systems and alloys). The fluid flow arising due to buoyancy (thermal and compositional) is modeled using appropriate (Boussinesq approximation) terms, which are generally incorporated into the momentum conservation. For isothermal phase change (pure systems), the second term on the RHS of Equation 5.7.23 vanishes. After solving Equation 5.7.23, the liquid fraction–temperature relation is utilized to obtain the interface location. Also, it may be noted that the enthalpy method is usually solved in its dimensional form along with mass and momentum conservation. Since individual growth of solid phases is not accounted for in the macro-model, the regions consisting of both solid and liquid phases are treated as a porous media with the porosity as a function of liquid (or solid) volume fraction. A well-known relation for modeling the porous medium flow is the Carman–Kozeny equation $\left(\nabla p = \left[-B \frac{(1-f_l)^2}{f_l^3} \right] \vec{\mathbf{u}}_0 \right)$, where B is a large computational number to avoid division by zero, and $\vec{\mathbf{u}}_0$ is the apparent velocity of the fluid. Hence, this popular enthalpy approach is also known as *enthalpy–porosity* formulation (Brent et al. 1988). Similarly, the conservation of species governing the mass balance of the solute (alloying element) in a binary alloy solidification problem is given as follows.

$$\frac{\partial(\rho C_m)}{\partial t} + \nabla \cdot (\rho \vec{\mathbf{U}} C_m) = \nabla \cdot (D_m \nabla C_m) + S_C \tag{5.7.25}$$

where $C_m = f_s C_s + f_l C_l$ is the mixture concentration and $D_m = f_s D_s + f_l D_l$ is the effective mixture diffusivity. The interface mass balance for the alloy element is given by a relation similar to the Stefan's condition (Equation 5.7.21) $k_p C_0 \vec{\mathbf{V}}_{s-l} = D_s \nabla C\big|_{x=\text{solid}} - D_l \nabla C\big|_{x=\text{liquid}}$, where k_p is the equilibrium partition coefficient (C_s / C_l). In contrary to the heat balance where the temperature at the interface is constant, but a discontinuity in the heat flux exists, species balance imposes discontinuity in both composition and the mass flux.

Several variations of the above formulation exist (Swaminathan and Voller, 1992; Krane et al., 1997; Chakraborty and Dutta, 2001) for macroscopic solidification problems incorporating movement of solid phase and other multicomponent alloy systems, but the essence of the method remains in estimating the temperature via enthalpy function without solving a separate equation for the evolution of phase-indicator function (liquid or solid volume fraction).

Numerical Methodology

The solution methodology for the enthalpy method is as follows:

1. Obtain the discretized forms of Equations 5.7.22 through 5.7.24 along with mass and momentum conservation equations.
2. Initialize temperature, composition and velocity fields.
3. Iteration:
   a. Update the variables (T or H, $\mathbf{u}$, C) with previous time instance
   b. Estimate the source terms for energy, mass, momentum, and species conservation
   c. Solve for T or H, $\mathbf{u}$, C
   d. Check for convergence if implicit solution and repeat steps (a)–(c).
4. Calculate the update for H or T (from the definition of ΔH).
5. Proceed to next time step.

Microscopic Model

In solidification problems, particularly in the case of alloys, the interface morphology is strongly dependent on the local cooling rate (dT / dt) and composition in the liquid ahead of the interface. At very slow (~0.1°C/s) and very high (~10^6°C/s) cooling rates, the interface remains flat, and at the intermediate values, the interface shape gradually assumes cellular, columnar dendritic and equiaxed dendritic structures, at the meso and microscopic scales (to facilitate heat transfer through larger surface areas). The macroscopic enthalpy method does not resolve these features and only provides information on an averaged value of the solid and liquid volume fractions in a given computational cell (Figure 5.7.1). For resolving and predicting the evolution (growth and motion) individual microstructural features, anisotropic crystal growth and interfacial surface energy need to be locally accounted along with the interfacial heat balance, as the growth will strongly depend on the interfacial curvature, kinetics, and alloy composition. Furthermore, due to a large disparity between the thermal and compositional diffusion rates, small grid sizes often need to be considered and adaptive meshing techniques can be very beneficial. A brief overview of microstructural simulation aspects is presented below. A detailed description of the microstructural simulation of solidification can be obtained from Voller (2008).

The microscopic modeling of solidification involves handling of more complex aspects such as anisotropy in microstructures, disparity in diffusion length scales of thermal and compositional fields and grain motion. Fully realistic simulations continue to remain a challenge for predicting accurate alloy solidification microstructures for large castings. Methods such as phase-field (Tong et al., 2001), cellular automata (Gandin and Rappaz, 1997; Wang et al., 2003; Karagadde et al., 2014), and LS (Tan and Zabaras, 2006; Theillard et al., 2014) have been extensively used for modeling microstructural evolution during solidification (Hecht et al., 2004; Rappaz, 2016). Of late, enthalpy method has also found applications in simulating the microstructural growth (Al-Rawahi

and Tryggvason, 2002; Pal et al., 2006; Voller, 2008; Karagadde et al., 2012a; Bhattacharya et al., 2014). During microstructure evolution, the interface temperature is lower than the melting point (liquidus temperature) of the alloy and the difference is given by the sum of the curvature, compositional and kinetic undercooling (Dantzig and Rappaz, 2009),

$$T_{\text{int}} = T_{\text{l}} - \kappa_{\text{s-l}}\Gamma - m\Delta C - \lambda \vec{\mathbf{V}}_{\text{s-l}} \cdot \vec{\mathbf{n}} \tag{5.7.26}$$

where $\kappa_{\text{s-l}}$ is the curvature of the solid–liquid interface (positive for a convex surface), Γ is Gibbs–Thompson coefficient which gives the extent of additional undercooling for unit curvature, m is the slope of liquidus line from the phase diagram, ΔC is the change in composition ahead of the s–l interface with respect to the equilibrium value, and λ is the kinetic coefficient (Voller, 2008). The interface temperature can be explicitly imposed while solving the enthalpy equation along with the species conservation. It is to be noted that the microscopic models represent the individual interfaces; whereas, in macroscopic solidification studies, a volume-averaged representation of the overall solid–liquid mixture is evaluated.

Application of Enthalpy Method

As an application problem, we demonstrate the simulation of solidification of aluminum in the presence of a hydrogen bubble. The hydrogen is a dissolved gas in liquid aluminum. Upon solidification, an equivalent amount of hydrogen exceeding the solubility limit is rejected into the liquid, and further diffuses into the nucleated bubbles. Subsequently, the bubbles undergo growth. There are two interfaces in this problem (solid–liquid and bubble–liquid), which are modeled using two different interface tracking methods, namely, enthalpy and LS, respectively, through a coupled approach. In this example, a 100×100 2D discretization of the domain was performed. Solidification is initiated from the left boundary, with rest of the boundaries being insulated. The evolution of solidification, hydrogen species conservation, and resulting bubble growth in liquid are captured, see Figure 5.7.2. The simulation shows segregation during solidification (distribution of species) of hydrogen (in cc/100 g of Al) ahead of the solid–liquid interface, and subsequent diffusion of hydrogen into

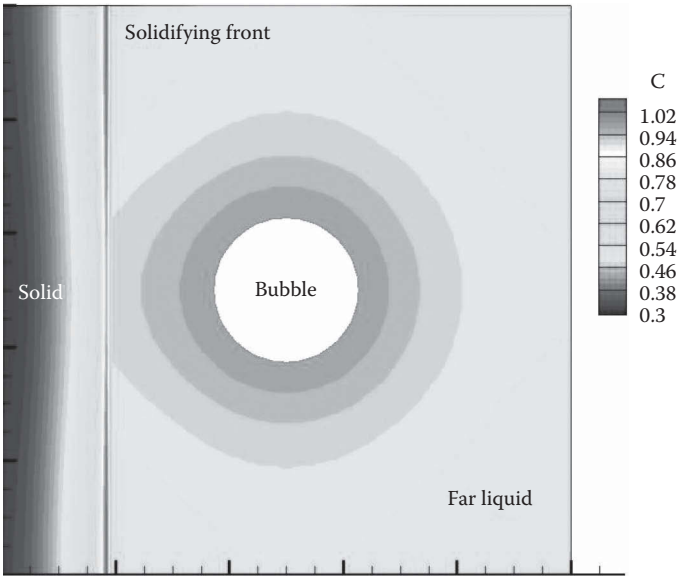


FIGURE 5.7.2 Simulation of solidification of aluminum, hydrogen species balance, and resulting growth of a hydrogen bubble.

the bubble. Higher concentration of the species can be noticed ahead of the solid–liquid interface, whereas, hydrogen depleted regions exist in the solid and also around the bubble (hydrogen is diffusing into the bubble and causing growth). Furthermore, the bubble can move under buoyancy effects and interact with the solidifying front (Karagadde et al., 2012b). This problem addresses, simultaneous mass, momentum, heat and species transfer, and demonstrates the potential of using the coupled interface tracking methods to simulate complex multiphase problems.

VOF Method

The VOF method is used primarily for sharp-interface multiphase flow problems involving both free and moving boundaries. In this method, the volume of each fluid (phase) is tracked in all cells containing the portions of the interface. The phases are distinguished by a volume or void-fraction function. This is defined for one particular fluid inside a cell and is given by material volume of that phase divided by the cell volume. The quantification is similar to any order-parameter representation, and the value is either unity or zero in any pure phase cell. It carries a value between these limits for multifluid cells. VOF methods involve solution of the evolution equation. Besides the solution of the evolution equation, which is also known as the *VOF advection equation* (Hirt and Nichols, 1981), VOF methods are successful with a good *interface reconstruction algorithm*. The reconstruction is necessary for obtaining the explicit description of the interface in each multifluid cell based on void fractions.

The development of VOF formulation is summarized by considering the methods of interface reconstruction (Gerlach et al., 2006). Some of the initial approaches consisted of representing the interface by a piecewise-constant line in every two-fluid cell. They were mainly SLIC (simple line interface calculation) of Noh and Woodward (1976) and Hirt and Nichols (1981). Youngs (1982) proposed a much improved first-order accurate method by introducing a piecewise-linear method which was shown to be very efficient and robust. Puckett et al. (1997) developed a more accurate technique known as the LVIRA (linear VOF interface reconstruction algorithm) having a second-order accuracy. Renardy and Renardy (2002) presented a technique based on piecewise parabolic segments. VOF has been successfully applied for various multiphase problems involving bubble dynamics, free surface flows, etc., particularly for incompressible flows. This implementation ensures volume (as well as mass) conservation by obtaining nondiffusive solution of the advection equation. Geometrically calculated void-fraction fluxes are estimated at the cell phases based on the reconstructed interfaces. Figure 5.7.3 shows a representation of this method and the geometric consideration of fluxes for volume fraction advection into the neighboring control volumes.

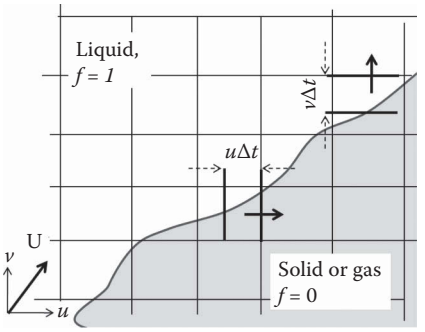


FIGURE 5.7.3 VOF advection—amount of phases advected to neighboring control volumes in *u* and *v* directions are represented.

VOF Equation

For incompressible flow, the fluid density does not change along the particle paths. The advection equation for the volume fraction f is given by,

$$\frac{\partial f}{\partial t} + \vec{\mathbf{U}}_{\text{adv}} \cdot \nabla f = 0 \tag{5.7.27}$$

Typically, the value of void fraction of liquid is chosen as $f = 1$ and of the other phase is chosen as $f = 0$. The void-fraction field f is advected using the velocity of the fluid $\vec{\mathbf{U}}_{\text{adv}}$, where $\vec{\mathbf{U}}_{\text{adv}}$ is the sum of bulk velocity $\vec{\mathbf{U}}$ (obtained from the solution of Navier–Stokes equations) and interfacial velocity due to phase change $\vec{\mathbf{U}}_{\text{PC}}$ (obtained from the Stefan condition), where $\vec{\mathbf{U}}_{\text{adv}} = \vec{\mathbf{U}} + \vec{\mathbf{U}}_{\text{PC}}$ $\left\{\vec{\mathbf{U}}_{\text{PC}} = 0.5\left(\frac{1+\chi}{\chi}\right)\dot{M}\hat{\mathbf{n}}\right\}$. The flow velocity field, along with phase-change-driven growth rate are obtained by solving the mass, momentum, and energy conservation equations; Equations 5.7.18 through 5.7.20. For cases involving advection alone, Equation 5.7.27 can be rewritten using the incompressibility condition as,

$$\frac{\partial f}{\partial t} + \nabla \cdot (\vec{\mathbf{U}}\, f) = 0 \tag{5.7.28}$$

Numerical Methodology

Figure 5.7.4 shows the interface representation in a scalar control volume. Figure 5.7.4a shows the representation for calculating the fluxes into the neighboring cells, and Figure 5.7.4b shows the representation of interface normal and its distance from the center used for the reconstruction.

The solution of VOF equation is carried out using separate update procedures for x- and y-direction. Accordingly, Equation 5.7.27 is discretized as follows:

$$f_{i,j}^{*} = f_{i,j}^{n} + \frac{\Delta t}{\Delta x}\left(\text{Flux}_{i-1/2,j} - \text{Flux}_{i+1/2,j}\right) \tag{5.7.29}$$

$$f_{i,j}^{n+1*} = f_{i,j}^{*} + \frac{\Delta t}{\Delta x}\left(\text{Flux}_{i,j-1/2} - \text{Flux}_{i,j+1/2}\right) \tag{5.7.30}$$

Here, the term *Flux* defined as $\text{Flux}_{i+1/2,j} = u_{i,j} f_{i+1/2,j}$, is computed using geometric fluxing following Puckett et al. (1997) (also see Gerlach et al., 2006) as represented in Figure 5.7.4a. In the operator split algorithm employed here, fluxing is split into x- and y-directions where the x-sweep is done first (Equation 5.7.29) in an explicit way and y-sweep is done using an implicit formulation. The directions of the sweep are swapped every alternative time step to make the advection scheme second-order accurate

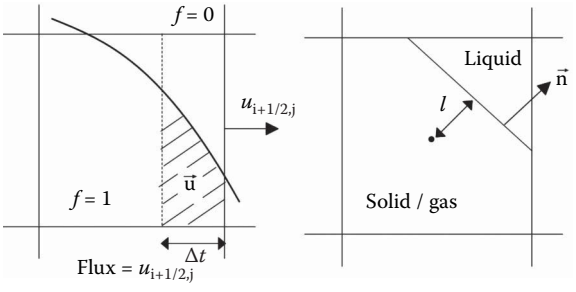


FIGURE 5.7.4 Representation of the interface in an individual control volume (From Gerlach, D. et al., *Int. J. Heat Mass Transfer*, 49(3–4), 740–754, 2006.With permission.): (a) estimation of the amount of flux into the neighboring cell and (b) linear reconstruction of the interface, showing the interface normal and the distance from the center of the cell.

in time. The details of the algorithm employed here are given in Gerlach et al. (2006). For advection of the dendrites using VOF, an extended velocity field is used. Fluid velocities in the cells adjacent to the solid boundary are imposed with the extended velocities from the dendrite particle to maintain a consistent volume fraction advection. The void-fraction field, $f_{i,j}$, by virtue of being advected geometrically, is always conserved and has values between zero and one only. However, an additional check is used to ensure that the void-fraction field does not become negative or greater than one in any grid cell.

Interface reconstruction is required after every advection sweep in order to compute the geometric fluxes accurately. Here, least-squared VOF interface reconstruction algorithm (LVIRA) of Puckett et al. (1997) is presented. The interface, in each computational grid cell, is represented by a linear segment and can be defined in 2D by the unit normal $\hat{\mathbf{n}}$ pointing into the liquid region and a distance, l, of the interface from the center (centroid in 3D) of the computational cell, as shown in Figure 5.7.4b. For e.g., in 2D calculations, a nine-point stencil around a two-phase cell is considered. A line is drawn through the stencil by using a guess $\hat{\mathbf{n}}$ (obtained by Youngs' method (Youngs, 1982)) and l, such that the void fraction at the center cell of the stencil matches the void fraction obtained after advection. The normal is estimated using the following relation.

$$\begin{aligned} n_x &= \frac{1}{8}(f_{i+1,j+1} + 2f_{i+1,j} + f_{i+1,j-1} - f_{i-1,j+1} - 2f_{i-1,j} - f_{i-1,j-1}) \\ n_y &= \frac{1}{8}(f_{i+1,j+1} + 2f_{i,j+1} + f_{i-1,j-1} - f_{i+1,j-1} - 2f_{i,j-1} - f_{i-1,j-1}) \end{aligned} \tag{5.7.31}$$

This initial guess is used to obtain the actual interface reconstruction using an optimization procedure. The least-square error (LSE) in the void-fraction field for the whole stencil is minimized to obtain the unit normal, $\hat{\mathbf{n}}$.

$$\mathrm{LSE} = \sum_{k,l=-1}^{1} (f_{i+k,j+l} - \tilde{f}_{i+k,j+l})^2 \tag{5.7.32}$$

where $f_{i+k,j+l}$ is the updated void fraction, and $\tilde{f}_{i+k,j+l}$ is the approximation estimated from the linear interface through the control volume. Geometrical construction of the interface using the normal and the distance l gives the approximation. Thus, a linear segment representing the interface in the center cell of the stencil is fixed, and this process is repeated for all the two-phase cells. This gives the correct interface normal and its location from the cell center.

Since VOF function abruptly changes across the interface, standard discretization schemes for evaluating the curvature, normal, etc., would result in large inaccuracies. Therefore, the function is smoothed in a region of finite thickness around the interface. Several methodologies exist in literature (Gerlach et al., 2006; Voller, 2008), and one of the simplest approaches is to induce the smearing by constructing a weighted averaged liquid fraction field over the nearest neighboring cells, i.e., if $\sum_{\mathrm{nb}} \beta_{\mathrm{nb}} = 1$, and $\beta_{\mathrm{nb}} \geq 0$, we have

$$g_{i,j} = \left(1 - \sum_{\mathrm{nb}} \beta_{\mathrm{nb}}\right) f_{i,j} + \sum_{\mathrm{nb}} \beta_{\mathrm{nb}} f_{\mathrm{nb}} \tag{5.7.33}$$

The Dirac delta function is evaluated as $|\mathbf{n}|$ (Gerlach et al., 2006). Similarly, interface curvature is evaluated by discretizing the following equation (for a 2D domain):

$$\frac{\left(\frac{\partial g}{\partial y}\right)^2 \frac{\partial^2 g}{\partial x^2} - 2\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\frac{\partial^2 g}{\partial x \partial y} + \left(\frac{\partial g}{\partial x}\right)^2 \frac{\partial^2 g}{\partial y^2}}{\left(\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2\right)^{3/2}} \tag{5.7.34}$$

The solution algorithm for two-phase flow simulation, using the VOF method, is as follows:

1. Initialize the volume fraction field in the domain as per the initial interface configuration ($f=1$ in cells containing liquid and 0 elsewhere). Initialize velocities, pressure, temperature, and interfacial mass flux (if any) as zero.
2. Solve the continuity and momentum equations, to obtain new time-level velocity and pressure field. In interfacial cells, the values of curvature, normal and Dirac-delta function are evaluated from the smeared volume fraction *f*.
3. Solve the energy equation, to obtain new time-level temperature field.
4. Solve advection of VOF equation (operator split approach), by calculating the amount of volume fraction "fluxed" to the neighboring cells (in the direction of flow).
5. Perform interface reconstruction using piecewise linear interface segments.
6. Go to step 2 till the stopping criterion is met.

Application of VOF Method

An application is presented here for the simultaneous growth and motion of an equiaxed dendrite in a uniform flow field. When the fluid flow velocities are comparable to the growth rate of the solid phase (in the form of a fourfold symmetric dendrite), simultaneous advection and growth of the solid can be simulated. A coupled Enthalpy and VOF approach is used in this study. A situation of this kind can occur in solidifying melts with strong convective currents. A domain of size $1000\times1000\ \mu m^2$, with $\Delta x = 10\ \mu m$ and $\Delta t = 10^{-7}$ s, is considered. Flow is introduced in the left boundary with a velocity of 0.01 m/s and is made to exit the domain at the right face. Figure 5.7.5 shows the location of the dendrite at various time instances. This example demonstrates the capability of

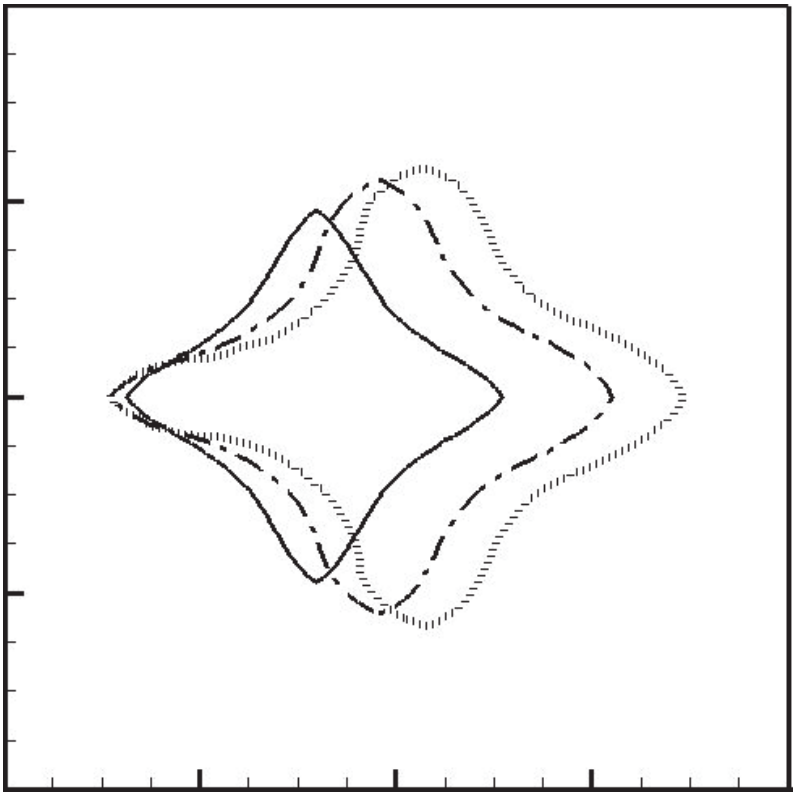


FIGURE 5.7.5 Advection of an evolving solid phase in a uniform velocity field using VOF method.

the VOF method in capturing the sharp interface, with coupled solidification and VOF advection. Further details for this implementation can be obtained from Karagadde et al. (2012a).

Level Set Method

LSM was proposed by Osher and Sethian (1988) as a general Eulerian computational technique to simulate moving interfaces and shapes—applicable to computational geometry, computer vision, material science, and computational fluid dynamics (Sethian, 1999; Osher and Fedkiw, 2003). Later, for the simulation of incompressible two-phase flow, Sussman et al. (1994) introduced the LSM along with the stream function–vorticity method.

LSM is a numerical methodology for the representation of an interface, and modeling the temporal evolution of interface, surface tension force, and interfacial mass transfer. It is based on the concept of implicit surfaces where a *normal distance function* called LS function φ is defined in a complete domain, i.e., for both fluids. The function corresponds to $\varphi = 0$ at the interface, with positive values for one of the fluids and negative values in the other fluid, shown in Figure 5.7.6a. For a circular interface of radius 5 units, Figure 5.7.6b shows the LS contours as concentric circles with an increase in its absolute value away from the interface.

The contour of $\varphi = 0$ in the domain represents the physically relevant *sharp* interface, and a numerically relevant *diffused* interface is represented by finite thickness ($-0.2 < \varphi < 0.2$, in Figure 5.7.6b), across which the fluid properties are smoothened (or averaged). This is done to avoid the numerical difficulties encountered due to the sharp change of thermophysical properties, across the interface, and nonzero surface tension force as well as volume generation/depletion (during phase change) at the interface.

The thickness across the interface is defined by normal distance across the interface and since the function φ is a normal distance function, constant thickness of 2ε across the interface is defined by the value of the function $-\varepsilon < \varphi < \varepsilon$; where ε is related to grid size and is usually taken as a factor of the grid spacing, e.g., $1.5\Delta x$. Note that, the interface thickness will be negligible on a fairly fine grid.

For LSM as well various other methods and their applications, an excellent review was presented by Wörner (2012), for multiphase flows in microfluidics and microprocess engineering. Recently, Sharma (2015) presented a detailed review of development as well as application of LSM for various types of multiphase flow.

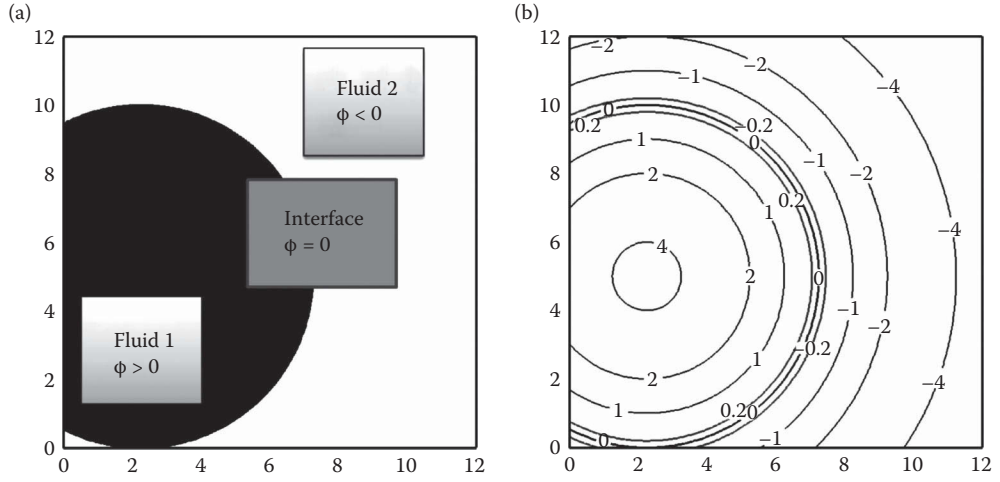


FIGURE 5.7.6 (a) LS function representation and (b) its contours for a circular interface of unit radius.

LS Functions and Equations

Along with the ease in the prescription of the thickness of the interface, the normal-distance LS function directly evaluates the normal unit vector at the interface and the curvature of the interface as follows:

$$\hat{\mathbf{n}} = \nabla\phi / |\nabla\phi| \text{ and } \kappa = -\nabla \cdot \hat{\mathbf{n}} \tag{5.7.35}$$

In addition to the LS function, the LSM considers two more mathematical functions: Heaviside and Dirac-delta functions. However, instead of their sharp definition in mathematics, they are considered diffused in LSM, called as *smoothened* Heaviside and Dirac-delta functions; their respective expressions are given as,

$$H_\varepsilon(\phi) = \begin{cases} 1 & \text{if} \quad \phi > \varepsilon \\ \dfrac{\phi+\varepsilon}{2\varepsilon} + \dfrac{1}{2\pi}\sin\left(\dfrac{\pi\phi}{\varepsilon}\right) & \text{if} \quad |\phi| \le \varepsilon \\ 0 & \text{if} \quad \phi < -\varepsilon \end{cases} \tag{5.7.36}$$

$$\delta_\varepsilon(\phi) = \frac{dH_\varepsilon(\phi)}{d\phi} = \frac{1}{2\varepsilon} + \frac{1}{2\varepsilon}\cos\left(\frac{\pi\phi}{\varepsilon}\right) \quad \text{if} \quad |\phi| \le \varepsilon$$
$$\delta_\varepsilon(\phi) = 0 \quad \text{otherwise} \tag{5.7.37}$$

The smoothened Dirac-delta function is used to model the nonzero surface tension force and interfacial mass transfer, at the diffused interface, in the momentum and continuity equation (Equations 5.7.18 and 5.7.19), respectively. Whereas, the smoothened Heaviside function is used for the gradual variation of the thermophysical properties from one to the other fluid, across the diffused interface,

$$\begin{aligned} \rho_m &= H_\varepsilon(\phi) + \chi(1 - H_\varepsilon(\phi)) \\ \mu_m &= \mu_1 H_\varepsilon(\phi) + \eta(1 - H_\varepsilon(\phi)) \end{aligned} \tag{5.7.38}$$

where $\chi = \rho_2/\rho_1$ is the density ratio and $\eta = \mu_2/\mu_1$ is the viscosity ratio; considering fluid 1 as the reference fluid. Note that the Heaviside and Dirac-delta functions are smoothened to avoid numerical instabilities, and make the numerical implementation a lot easier.

Gada and Sharma (2009) presented a physical interpretation of the value of H_ε at the centroid (face center) of a CV as volume (area) fraction, i.e., fraction of volume (surface) of the CV occupied by the fluid 1. Furthermore, the value of $\delta_\varepsilon(\phi)$ at the centroid of a CV was interpreted as interfacial area concentration, the ratio of interface area inside the CV and the volume of the CV. Furthermore, they used the physical interpretation (of the various functions) for a control volume-based derivation (commonly used in fluid mechanics and heat transfer) of the continuity equation (Equation 5.7.18) and LS advection equation (Equation 5.7.39), by applying volume and mass conservation laws to the CVs, respectively. The LS advection equation along with some additional equations needed in LSM are presented as follows:

Mass conservation (LS advection) equation:

$$\frac{\partial\phi}{\partial\tau} + \vec{\mathbf{U}}_{adv} \cdot \nabla\phi = 0$$
$$\text{where } \vec{\mathbf{U}}_{adv} = \vec{\mathbf{U}} + \vec{\mathbf{U}}_{PC} \quad \left\{ \vec{\mathbf{U}}_{PC} = 0.5\left(\frac{1+\chi}{\chi}\right)\dot{M}\hat{n} \right\} \tag{5.7.39}$$

where $\vec{\mathbf{U}}_{adv}$ is the sum of bulk velocity $\vec{\mathbf{U}}$ (obtained from the solution of Navier–Stokes equations) and interfacial velocity due to phase change $\vec{\mathbf{U}}_{PC}$ (obtained from the Stefan condition).

Reinitialization Equation:
Figure 5.7.6b shows that the line contour for $\varphi=-0.2$ and $\varphi=0.2$ are exactly in-between the contour for $\varphi=0$—before solving the LS advection equation; however, after the solution, the boundary of the numerically diffused interface does not lie exactly in-between the physically relevant sharp interface. Thus, the LS field no more remains the normal distance function, and leads to an inaccurate calculation of thermophysical properties and interface normal/curvature. This requires a repair of the LS function as the normal distance function, without affecting the sharp interface at $\varphi=0$. This is achieved by a reinitialization step, using a pseudo-transient partial differential equation based reinitialization equation. The equation, proposed by Sussman et al. (1994, 1999), is given as

$$\frac{\partial \phi}{\partial \tau_s} + S_\varepsilon(\phi_0)(1-|\nabla\phi|) = 0 \tag{5.7.40}$$

where τ_s is the pseudo time, $\phi_0\left(\vec{x}\right) = \phi(\vec{x}, \tau_s = 0)$, and $S_\varepsilon(\phi_0)$ is a smoothened sign function. The sign function was approximated numerically by Sussman et al. (1994) as $S_\varepsilon(\phi_0) = \phi_0 / \sqrt{\left(\phi_0^2 + \Delta x^2\right)}$; Δx is used to avoid dividing by zero. This equation is solved till steady state within the thickness of the interface so that $|\nabla\phi| = 1$. Moreover, the pseudo time step is taken as one tenth of LS node spacing (Sussman et al., 1994). The reinitialization is carried out in regular time intervals, often after each time step; however, less frequent reinitialization (for example, after every 10th time-step) are also common.

Sharp-Interface Treatment of the Energy Equation

Although the momentum equation (Equation 5.7.19) is solved using the averaged thermophysical properties, a sharp-interface formulation is commonly used for the solution of the energy equation; proposed by Son and Dhir (2007). Thus, Equation 5.7.20 for energy conservation is modified as

$$\frac{\partial \theta}{\partial \tau} + \nabla \cdot (\vec{\mathbf{U}}\theta) = \frac{1}{Re_p Pr_p} \nabla^2 \theta \tag{5.7.41}$$

where the subscript $p=1$ if $\varphi > 0$ and $p=2$ if $\varphi < 0$, and $\theta=0$ if $\varphi=0$. The application of the sharp-interface treatment of the energy equation for a phase-change problem leads to the calculation of the normal temperature gradient at the near-interface liquid and vapor points. However, the interfacial mass flux $\dot{M}$ needs to be calculated at all points inside the diffused-interface—for its efficient implementation in the continuity equation (Equation 5.7.18). Thus, a constant extrapolation (in the normal direction across the interface) of the near-interface normal gradient is used, so that both liquid and vapor temperature gradients are available at all points inside the diffused-interface. This is achieved by an *extension equation*, again a pseudo-transient partial differential equation (Son and Dhir, 2007), given as

$$\frac{\partial}{\partial \tau_s}\left(\frac{\partial \theta}{\partial n}\right)_p + S(\phi)\hat{n} \cdot \nabla\left(\frac{\partial \theta}{\partial n}\right)_p = 0 \tag{5.7.42}$$

where subscript $p=1$ ($p=2$) when extrapolating the liquid (vapor) side temperature gradient. Moreover, τ_s is the pseudo time, and $S\left(\phi\right)$ is a sharp sign function; equal to –1, 0, and 1 for $\varphi < 0$, $\varphi > 0$, and $\phi=0$.

Numerical Methodology

One of the biggest disadvantages of the LSM is mass error, i.e., an error in the conservation of mass (Chang et al., 1996; Sussman and Fatemi, 2000, Yap et al., 2006), with a timewise unphysical

increase or decrease in the amount of particular fluid (enclosed within an interface) as the interface evolves with time. One of the reasons for the mass error is numerical diffusion in the discretization of LS advection equation, which can be handled by using higher order advection scheme. Gada (2012) mentioned that mass error acts as an indicator of the numerical inaccuracy in the solution of LS advection equation (derived from mass conservation law; Gada and Sharma, 2009); whereas, in VOF method, there is no such indicator for the numerical inaccuracy in the solution of volume fraction advection equation.

One of the simplest approaches to reduce the mass error is the grid refinement; however, the other approaches using a higher order advection scheme and better reinitialization procedure are also available. Furthermore, without using the reinitialization equation and combining LSM with another method, hybrid methods were introduced. Since the mass error reduces with grid refinement, adaptive and dual-grid LSMs (Gada and Sharma, 2011) are found to be effective.

During the early days of LSM (Sussman et al., 1994), both LS advection and reinitialization equations were solved by second-order Essentially Non-Oscillatory (ENO) scheme (Shu and Osher, 1989). However, nowadays, LS advection equation is mostly solved with high-order schemes in time and space; third-order total variation diminishing (TVD) Runge–Kutta scheme and third-order or fifth-order Weighted Essentially Non-Oscillatory (WENO) scheme (Jiang and Peng, 2000), respectively.

One of the recent and significant developments in the LSM is the introduction of ghost fluid method, as a numerical technique for sharply enforcing the boundary or matching conditions at the interface (without being smeared out over several grid points); called *as sharp-interface LSM*, proposed by Fedkiw et al. (1999). Recently, Shaikh et al. (2016) presented a detailed numerical methodology for the sharp-interface LSM and its application to a range of heat and/or mass transfer induced Stefan problems.

The solution algorithm for two-phase flow simulation, using the LSM, proceeds as follows:

1. Initialize the LS field in the domain as per the initial interface configuration. Initialize velocities, pressure, temperature, and interfacial mass flux as zero.
2. Calculate the Heaviside function (Equation 5.7.36) and Dirac-delta function (Equation 5.7.37) from the LS field and the property field.
3. Solve the continuity and momentum equations (Equations 5.7.18 and 5.7.19), to obtain new time-level velocity and pressure field.
4. Solve the energy equation (Equation 5.7.41) using the sharp-interface LSM, to obtain the new time-level temperature field.
5. Calculate the normal temperature gradient at near-interface cells, extrapolate it to the diffused interface (Equation 5.7.42) and calculate interfacial mass flux (Equation 5.7.18) and interface velocity due to phase change.
6. Add the background velocity (from step 3) and velocity due to phase change (from step 5) at LS cell centers to obtain the total advecting velocity of LS function. Solve the LS advection equation (Equation 5.7.39) to obtain the new time-level LS field.
7. Solve the reinitialization equation (Equation 5.7.40)—to reinitialize the LS field.
8. Go to step 2 till the stopping criterion is met.

Application of LSM

The physical situation of an application for 3D film boiling is shown in Figure 5.7.7, wherein a rod is located in the pool of saturated liquid. The rod is subjected to uniform volumetric heat generation and its surface temperature is high enough to ensure an enveloping vapor film, which restricts any contact of liquid with the rod surface. The heat generated from the rod is transferred across the vapor layer to the interface; leading to phase change. Due to continuous generation of vapor, the thickness of vapor film increases, and beyond a certain thickness, the interface becomes unstable. The vapor

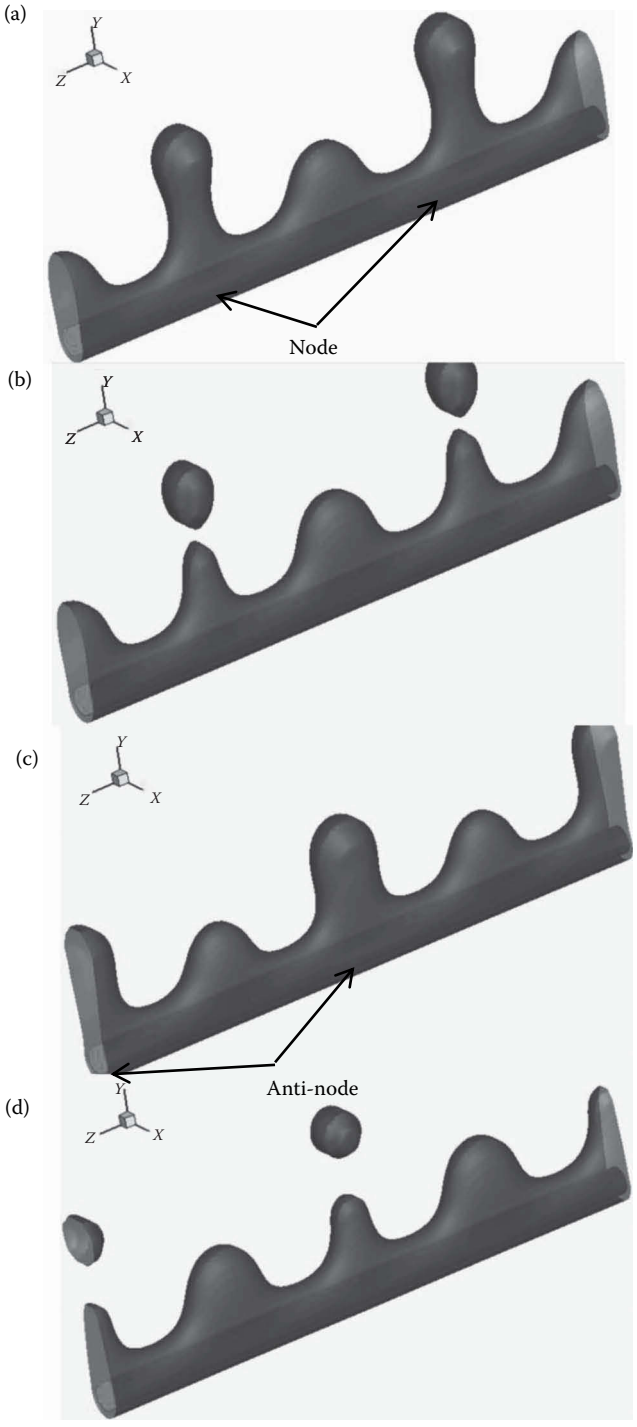


FIGURE 5.7.7 Application of a dual-grid LSM: Interface dynamics for conjugate heat transfer and 3D film boiling over a heat-generating rod (From Gada, V. H. A novel level-set based CMFD methodology in 2D/3D Cartesian and cylindrical coordinates and its application for analysis of stratified flow and film boiling. PhD Thesis, Mechanical Engineering, IIT Bombay, Mumbai, 2012.).

begins to accumulate near the crest/node of the wavy interface. Furthermore, the vapor bulb at the node grows to form a vapor bubble (Figure 5.7.7a), which ultimately breaks away beyond a certain size (Figure 5.7.7b). Thereafter, the interface retracts towards the heater rod and the bubble-release process starts at an alternate location, i.e., the antinode (Figure 5.7.7c and d).

Here, the situation is inherently 3D; where the vapor is generated at the interface enveloping the complete cylinder. However, the vapor rises up continuously due to buoyancy, leading to thinner film (and increased generation of vapor) at the bottom of cylinder. The accumulation of vapor on top of the cylinder gives rise to instability in the axial direction, leading to periodic formation and release of bubbles, which deform in all the directions. This periodic process of bubble formation and release at node and antinode, affects the temperature on the rod.

Concluding Remarks

Both numerical methodology and application of three important computational methods for simulating multiphase processes are presented here. Enthalpy method, presented here, has been widely employed for solving solid–liquid phase-change problems in multicomponent systems in macroscopic framework. VOF method as compared to LSM has a much less mass-conservation error; whereas, LSM is much easier to implement and also straightforward to extend from 2D to 3D. Thus, probably VOF method has a better formulation and LSM is numerically friendly.

Most of the numerical methods for multiphase flows exist in the context two-phase systems. Several approaches exist for obtaining phenomenological simulations of processes involving more than two phases, such as a gas, solid and liquid or multimaterial flows. Some of the recent efforts have particularly addressed such problems, for example, Karagadde et al. (2012b) presented a coupled LS and enthalpy approach to simulate growth and motion of hydrogen bubbles and their interaction with an evolving solidifying interface; Starinshak et al. (2014) presented a LSM for simulating flows involving multiple immiscible fluids. Hence, multiphase modeling approaches have demonstrated their potential to numerically simulate a variety of complex problems and processes. In future, there is a scope for numerically modeling multiphase fluid-structure interaction problems (modeling timewise varying two-fluid interface as well as solid–fluid interface), with appropriate wetting models. Furthermore, there is scope for improvements in the numerical methods to incorporate the microscopic interaction involved in moving contact line and nucleate boiling.

REFERENCES

Al-Rawahi, N., and Tryggvason, G., 2002. Numerical simulation of dendritic solidification with convection: Two-dimensional geometry. *Journal of Computational Physics*, 180(2): 471–496.

Bhattacharya, A., Kiran, A., Karagadde, S., and Dutta, P., 2014. An enthalpy method for modeling eutectic solidification. *Journal of Computational Physics*, 262: 217–230.

Brackbill, J. U., Kothe, D. B., and Zemach, C., 1992. A continuum method for modeling surface tension. *Journal of Computational Physics*, 100: 335–354.

Brent, A. D., Voller, V. R., and Reid, K. J., 1988. Enthalpy-porosity technique for modeling convection-diffusion phase change: application to the melting of a pure metal. *Numerical Heat Transfer*, 13(3): 297–318.

Chakraborty, S., and Dutta, P., 2001. A generalized formulation for evaluation of latent heat functions in enthalpy-based macroscopic models for convection-diffusion phase change processes. *Metallurgical and Materials Transactions B: Process Metallurgy and Materials Processing Science*, 32(3), 562–564.

Chang, Y. C., Hou T. Y., Merriman, B., and Osher, S., 1996. A level set formulation of Eulerian interface capturing methods for incompressible fluid flows. *Journal of Computational Physics*, 124: 449–464.

Crank, J., 1987. *Free and Moving Boundary Problems*, Clarendon Press, Oxford.

Dantzig, J. A., and Rappaz, M., 2009. *Solidification*, EFPL/CRC Press, Lausanne, Switzerland.

Eyres, N. R. et al., 1946. The calculation of variable heat flow in solids. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 240(813): 1–57.

Fedkiw, R., Aslam, T., Merriman, B., and Osher, S., 1999. A non-oscillatory Eulerian approach to interfaces in multimaterial flows (the ghost fluid method). *Journal of Computational Physics*, 152(2): 457–492.

Gada, V. H., 2012. A novel level-set based CMFD methodology in 2D/3D Cartesian and cylindrical coordinates and its application for analysis of stratified flow and film boiling. PhD Thesis, Mechanical Engineering, IIT Bombay, Mumbai.

Gada, V. H, and Sharma, A., 2009. On derivation and physical-interpretation of level set method based equations for two-phase flow simulations. *Numerical Heat Transfer, Part B: Fundamentals*, 56: 307–322.

Gada, V. H., and Sharma, A., 2011, On novel dual-grid level-set method for two-phase flow simulation. *Numerical Heat Transfer, Part B: Fundamentals*, 59: 26–57.

Gandin, C.-A., and Rappaz, M., 1997. A 3D cellular automaton algorithm for the prediction of dendritic grain growth. *Acta Materialia*, 45(5): 2187–2195.

Gerlach, D., Tomar, G., Biswas, G., and Durst, F., 2006. Comparison of volume-of-fluid methods for surface tension-dominant two-phase flows. *International Journal of Heat and Mass Transfer*, 49(3–4): 740–754.

Hecht, U. et al., 2004. Multiphase solidification in multicomponent alloys. *Materials Science and Engineering: R: Reports*, 46(1–2), 1–49.

Hirt, C. W., and Nichols, B. D., 1981. Volume of fluid (VOF) method for the dynamics of free boundaries. *Journal of Computational Physics*, 39(1): 201–225.

Hyman, J. M., 1984. Numerical methods for tracking interfaces. *Physica D: Nonlinear Phenomena*, 12(1–3): 396–407.

Jiang, G. S., and Peng, D., 2000. Weighted ENO schemes for Hamilton Jacobi equations. *SIAM Journal of Scientific Computing*, 21: 2126–2143.

Karagadde, S. Yuan, L., Shevchenko, N., Eckert, S., and Lee, P. D., 2014. 3-D microstructural model of freckle formation validated using in situ experiments. *Acta Materialia*, 79: 168–180.

Karagadde, S., Bhattacharya, A., Tomar, G., and Dutta P., 2012a. A coupled VOF–IBM–enthalpy approach for modeling motion and growth of equiaxed dendrites in a solidifying melt. *Journal of Computational Physics*, 231(10): 3987–4000.

Karagadde, S., Sundarraj, S., and Dutta, P., 2012b. A model for growth and engulfment of gas microporosity during aluminum alloy solidification process. *Computational Materials Science*, 65: 383–394.

Krane, M. J. M., Incropera, F. P., and Gaskell, D. R., 1997. Solidification of ternary metal alloys—I. Model development. *International Journal of Heat and Mass Transfer*, 40(16): 3827–3835.

Noh, W.F., and Woodward, P., 1976. SLIC (Simple Line Interface Calculation), in *Lecture Notes in Physics*, vol. 59: Proc. Fifth Int. Conf. on Numerical Methods in Fluid Dynamics, A.I. van de Vooren and P.J. Zandbergen, eds. Springer, Berlin.

Osher, S., and Fedkiw, R., 2003. *Level Set Methods and Dynamic Implicit Surfaces*. Springer-Verlag, New York.

Osher, S., and Sethian, J. A., 1988. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79(1): 12–49.

Pal, D., J. Bhattacharya, P. Dutta, and S. Chakraborty., 2006. An enthalpy model for simulation of dendritic growth. *Numerical Heat Transfer, Part B: Fundamentals*, 50(1): 59–78.

Puckett, E. G. et al., 1997. A high-order projection method for tracking fluid interfaces in variable density incompressible flows. *Journal of Computational Physics*, 130(2): 269–282.

Rappaz, M., 2016. Modeling and characterization of grain structures and defects in solidification. *Current Opinion in Solid State and Materials Science*, 20(1): 37–45.

Renardy, Y., and Renardy, M., 2002. PROST: a parabolic reconstruction of surface tension for the volume-of-fluid method. *Journal of Computational Physics*, 183: 400–421.

Sethian, J. A., 1999. *Level Set Methods and Fast Marching Methods*. 2nd edition, Cambridge University Press, New York.

Shaikh, J., Sharma, A., and Bhardwaj, R., 2016. On sharp-interface level-set method for heat and/or mass transfer induced Stefan problem. *International Journal of Heat and Mass Transfer*, 96: 458–473.

Sharma, A., 2015, Level set method for computational multi-fluid dynamics: A review on developments, applications and analysis, *Sadhana - Academy Proceedings in Engineering Science*, 40(3): 627–652.

Shu, C. W., and Osher, S., 1989. Efficient implementation of essentially non-oscillatory shock-capturing schemes II. *Journal of Computational Physics*, 83: 32–78.

Son, G., and Dhir, V. K., 2007. A level set method for analysis of film boiling on an immersed solid surface. *Numerical Heat Transfer, Part B: Fundamentals*, 52(2): 153–177.

Starinshak, D. P., Karni, S., and Roe, P. L., 2014. A new level set model for multimaterial flows. *Journal of Computational Physics*, 262: 1–16.

Sussman, M., and Fatemi, E., 2000. An efficient interface preserving level set re-distancing algorithm and its application to interfacial incompressible flow. *SIAM Journal of Science Computing*, 20: 1165–1191.

Sussman, M., Almgren, A. S., Bell, J. B., Colella, P., Howell, L. H., and Welcome, M. L., 1999. An adaptive level set approach for incompressible two-phase flows. *Journal of Computational Physics*, 148: 81–124.

Sussman, M., Samereka, P., and Osher, S., 1994. A level set approach for computing solutions to incompressible two-phase flows. *Journal of Computational Physics*, 114: 146–159.

Swaminathan, C. R., and Voller, V. R., 1992. A general enthalpy method for modeling solidification processes. *Metallurgical Transactions B*, 23(5): 651–664.

Tan, L., and Zabaras, N., 2006. A level set simulation of dendritic solidification with combined features of front-tracking and fixed-domain methods. *Journal of Computational Physics*, 211(1): 36–63.

Theillard, M., Gibou, F., and Pollock, T., 2014. A sharp computational method for the simulation of the solidification of binary alloys. *Journal of Scientific Computing*, 63(2): 330–354.

Tong, X., Beckermann, C., Karma, A., & Li, Q., 2001. Phase-field simulations of dendritic crystal growth in a forced flow. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 63(6 I): 061601/1–061601/16.

Voller, V. R., 2008. An enthalpy method for modeling dendritic growth in a binary alloy. *International Journal of Heat and Mass Transfer*, 51(3–4): 823–834.

Voller, V. R., Brent, A. D., and Prakash, C., 1989. The modelling of heat, mass and solute transport in solidification systems. *International Journal of Heat and Mass Transfer*, 32(9): 1719–1731.

Voller, V. R., Cross, M., and Markatos, N. C., 1987. An enthalpy method for convection/diffusion phase change. *International Journal for Numerical Methods in Engineering*, 24(1): 271–284.

Wang, W., Lee, P. D., and McLean, M., 2003. A model of solidification microstructures in nickel-based superalloys: Predicting primary dendrite spacing selection. *Acta Materialia*, 51(10): 2971–2987.

Wörner, M., 2003. A compact introduction to the numerical modeling of multiphase flows, Forschungszentrum Karlsruhe.

Wörner, M., 2012. Numerical modeling of multiphase flows in microfluidics and micro process engineering: a review of methods and applications. *Microfluidics and Nanofluidics*, 12(6): 841–886.

Yap, Y. F., Chai, J. C., Wong, T. N., and Toh, K. C., 2006. A global mass correction scheme for level set method. *Numerical Heat Transfer, Part B*, 50: 455–472.

Youngs, D. L., 1982. Time-dependent multi-material flow with large fluid distortion, in *Numerical Methods for Fluid Dynamics*, edited by K. W. Morton and M. J. Baines. Academic Press, New York, Vol. 24, pp. 273–285.

5.8 LARGE EDDY SIMULATION OF WALL-BOUNDED FLOWS

Amitabh Bhattacharya

Introduction

Turbulence is ubiquitous in both nature and engineering flows, and is characterized by velocity fields that contain a range of time and length scales (Pope, 2000). In this chapter, we will consider numerical simulation of incompressible isothermal flows, with velocity field $\mathbf{u}(\mathbf{x},t)$ and pressure field $p(\mathbf{x},t)$ that satisfy the Navier–Stokes equation (Cartesian tensor notation is used for Greek indices in this text):

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{5.8.1}$$

$$\frac{\partial u_i}{\partial t} + u_k \frac{\partial u_i}{\partial x_k} = -\frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_k^2} \tag{5.8.2}$$

Here, $\mathbf{x}$ denotes spatial coordinates, t denotes time, and ν is the kinematic viscosity of the fluid. The computational domain (Figure 5.8.1) is denoted as $\mathcal{D}$, with walls $\partial\mathcal{D}_w$ and inlets/outlets $\partial\mathcal{D}_f$. The boundary conditions for $\mathbf{u}$ are as follows:

$$\mathbf{u}(\mathbf{x},t) = 0 \quad \text{for} \quad \mathbf{x} \in \partial\mathcal{D}_w \quad \text{(No slip BC)} \tag{5.8.3}$$

$$\mathbf{u}(\mathbf{x},t) = \mathbf{u}^f(\mathbf{x},t) \quad \text{for} \quad \mathbf{x} \in \partial\mathcal{D}_f \quad \text{(Inlet/outlet BC)} \tag{5.8.4}$$

Solving Equations 5.8.1 and 5.8.2 via direct numerical simulation (DNS) requires a fine grid, which resolves the flow down to the Kolmogorov scales (Moin and Mahesh, 199). Particularly, at large Reynolds numbers, the ratio of large (integral scales) to small (Kolmogorov) length scales increases, requiring very large number of computational grid points and long computational times. For isotropic homogeneous turbulence, DNS computational time scales with Re_t^3 (Pope, 2000), while, for wall-bounded turbulent flows, DNS computational time scales with $Re_L^{2.4}$ (Chapman, 1979). Here, Re_t is the turbulent Reynolds number and Re_L is the Reynolds number based on streamwise length of the turbulent boundary layer, L. In large eddy simulation (LES), the goal is to reduce the computational time required to solve Equations 5.8.1 and 5.8.2 by evolving only the large scales, and model the effect of the small, "subgrid" scales on the large scales.

"Large scales" in the flow are defined via a spatial low pass filter of $\mathbf{u}(\mathbf{x},t)$ and $p(\mathbf{x},t)$, with a characteristic cutoff wavenumber $k_c = \pi/\Delta$, where Δ is the "filter width." The filtered velocity field $\tilde{\mathbf{u}}(\mathbf{x},t)$, where small-scale features of the flow have been eliminated, can be expected to have flow features with higher eddy turnover times (compared to the Kolmogorov scales). Thus, coarser spatial grid and larger time steps may be used to numerically evolve $\tilde{\mathbf{u}}(\mathbf{x},t)$. If k_c is chosen correctly, so that it lies within the inertial range of the energy spectra of the flow, the effect of the unresolved smaller scales on the larger scales can be expected to be universal [e.g., via scale similar cascade of energy from large to small scales (Kolmogorov, 1941)]. At the same time, evolving the large scales explicitly allows us to capture the effect of complex boundary conditions on the flow accurately. This makes LES models more accurate in comparison with Reynolds averaged Navier–Stokes (RANS) models, which explicitly evolve only the mean velocity $\langle\mathbf{u}(\mathbf{x},t)\rangle$, while trying to model the effect of *all* the fluctuating scales (i.e., including largest fluctuating scales) in the flow on $\langle\mathbf{u}(\mathbf{x},t)\rangle$.[1] Thus, while RANS models are computationally inexpensive compared to LES models, the model

[1] $\langle[\cdot]\rangle$ denotes ensemble average.

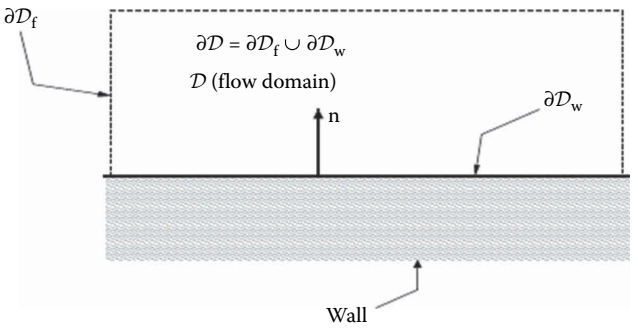


FIGURE 5.8.1 Schematic of computational domain.

constants in RANS often need to be recalibrated for different flow geometries to maintain accuracy, thus limiting its predictive capabilities.

Properties of LES Filter

For any field $\phi(\mathbf{x},t)$, the filtering operator, $\widetilde{[\cdot]}$ implies

$$\tilde{\phi}(\mathbf{x},t) = \int_{\mathcal{D}} \phi(\mathbf{x}',t) G(\mathbf{x},\mathbf{x}'-\mathbf{x}) d\mathbf{x}' \tag{5.8.5}$$

Here, $G(\mathbf{x},\mathbf{r})$ is the filter kernel (Ghosal and Moin, 1995; Sagaut, 2006), which smoothens $\phi(\mathbf{x},t)$ spatially and $\mathcal{D}$ is the flow domain (Figure 5.8.1). At any $\mathbf{x}$ location, $|G(\mathbf{x},\mathbf{r})|$ decays to zero beyond $|\mathbf{r}| > \Delta/2$. Therefore, for a given $\mathbf{x}$, the filter kernel is associated with a filter width Δ (Figure 5.8.2). In general, Δ itself will be a function of $\mathbf{x}$, i.e., $\Delta = \Delta(\mathbf{x})$. When Δ is independent of $\mathbf{x}$, the filter is said to be "homogeneous." Otherwise, the filter is called "inhomogeneous." In general, LES filters are inhomogeneous, especially in complex geometries. On a structured Cartesian computational grid, the filter kernel $G(\mathbf{x},\mathbf{r})$ is expressed as a product of three one-dimensional filter kernels $G_\alpha^{1D}(x,r)$:

$$G(\mathbf{x},\mathbf{r}) = \prod_{\alpha=1}^{3} G_\alpha^{1D}(x_\alpha, r_\alpha) \tag{5.8.6}$$

In general, $G_\alpha^{1D}(x,r)$ may be different for each α. Two popular filters are "Box" and "Gaussian" (Table 5.8.1). The important properties of homogeneous LES filters are listed here. Inhomogeneous

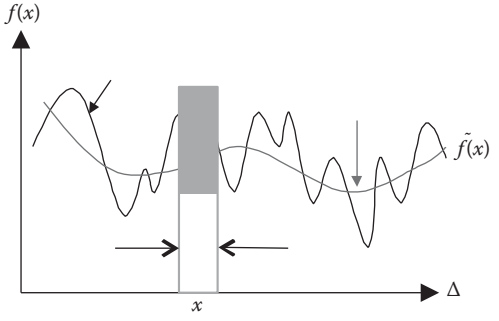


FIGURE 5.8.2 Operation of $\widetilde{[\cdot]}$ on any function $f(x)$ leads to the "filtered" function $\tilde{f}(x)$. The filter kernel $G(r)$ is small for $r > \Delta/2$. Thus, the filtering acts like a moving average and leads to the elimination of scales with size less than Δ. $\tilde{f}(x)$ is smoother compared to $f(x)$. Δ can be a function of x for inhomogeneous filters.

TABLE 5.8.1
Types of Filters. Here *H(x)* Is the Heaviside Function

Name	Filter Kernel ($G^{1D}(x,r)$)
Box	$\frac{1}{\Delta}\left(\frac{1}{2}\Delta(x)-\lvert r\rvert\right)$
Gaussian	$\left(\frac{6}{\pi\Delta(x)}\right)^2 \exp\left(-\frac{6r^2}{\Delta(x)^2}\right)$

filters will be discussed later in the text. We will therefore suppress the **x** dependence of G in the following.

The filtered field should have the same global average as the original unfiltered field. This implies $\int_D G(\mathbf{r})d\mathbf{r} = 1$. It is also easy to prove that the filtering operator is linear, i.e., $\widetilde{\phi(\mathbf{x})+\psi(\mathbf{x})} = \tilde{\phi}(\mathbf{x})+\tilde{\psi}(\mathbf{x})$. Next, if the flow domain D is infinite in space, then it is possible to show that

$$\frac{\partial\tilde{\phi}}{\partial x_k} = \widetilde{\frac{\partial\phi}{\partial x_k}} \tag{5.8.7}$$

In other words, for homogeneous filters, in the absence of any walls, the filtering operator $[\widetilde{\cdot}]$ and differentiation $\partial/\partial x_k$ commute. This property is not true for inhomogeneous filters, for which the "commutation error" $\frac{\partial\tilde{\phi}}{\partial x_k} - \widetilde{\frac{\partial\phi}{\partial x_k}}$ will not be zero in general. In the presence of boundaries, even for homogeneous filters, it can be shown that (Bhattacharya et al., 2008)

$$\widetilde{\frac{\partial\phi(\mathbf{x})}{\partial x_i}} = \frac{\partial\tilde{\phi}(\mathbf{x})}{\partial x_i} - \int_{\partial\mathcal{D}_f} G(\mathbf{x}'-\mathbf{x})\phi(\mathbf{x}')n_i d^2\mathbf{x}' - \int_{\partial\mathcal{D}_w} G(\mathbf{x}'-\mathbf{x})\phi(\mathbf{x}')n_i d^2\mathbf{x}' \tag{5.8.8}$$

The last two terms in the above equations involve values of ϕ (and not $\tilde{\phi}$) at the wall. These terms can be neglected if **x** is located at a distance much larger than Δ from the boundary, where the contribution of G to these integrals will be zero.

Filtering due to Numerical Discretization

Filtering of a spatial field essentially reduces the magnitude of the higher wavenumber modes in the field. It can be demonstrated that numerical discretization itself acts like a filter. Consider a function $\phi(x)$ that is periodic over $x \in [0,L]$ and which can be decomposed into its Fourier modes: $\phi(x) = \sum_{n=-\infty}^{\infty} \hat{\phi}_n \exp(ik_n x)$, where $k_n = 2\pi n/L$. A uniform computational grid with cell width Δ^g ensures that Fourier modes with wavenumbers $k_n > \pi/\Delta^g$ are not being represented on the grid. Thus, due to the reduction in the number of degrees of freedom, the very act of numerical discretization is equivalent to applying a low pass spatial filter on $\phi(x)$, with filter width Δ^g (Langford and Moser, 1999). Moreover, discretized differential operators acting on $\phi(x)$ may cause further attenuation of Fourier modes with higher wavenumber. For instance, if we discretize the derivative $\mathrm{d}\phi/\mathrm{d}x$ using a second-order finite-difference scheme, then we observe that

$$\left.\frac{\mathrm{d}\phi}{\mathrm{d}x}\right|_{x_i} = \frac{\phi(x_{i+1})-\phi(x_{i-1})}{2\Delta^g} = \sum_{n=-\infty}^{\infty} ik_n \frac{\sin(k_n\Delta^g)}{k_n\Delta^g}\hat{\phi}_n \exp(ik_n x)$$

Since $\sin(x)/x$ is small for $x > \pi$, it can be seen that $\sin(k_n\Delta^g)/k_n\Delta^g$ causes an attenuation of Fourier modes with $k_n \gg \pi/\Delta^g$. Thus, a finite-difference discretization of the derivative operator itself acts like a filter, although the corresponding shape of the filter kernel will depend on the specifics of the discretization scheme.

Governing Equations

We can now apply $\widetilde{[\cdot]}$ on Equations 5.8.1 and 5.8.2 and use the properties of LES filters (Equations 5.8.5–5.8.8) to obtain (Bhattacharya et al., 2008)

$$\frac{\partial \tilde{u}_k}{\partial x_k} = 0 \tag{5.8.9}$$

$$\frac{\partial \widetilde{u_i}}{\partial t} + \frac{\partial \widetilde{\tilde{u}_i \tilde{u}_k}}{\partial x_k} = -\frac{\partial \tilde{p}}{\partial x_i} + \nu \frac{\partial^2 \tilde{u}_i}{\partial x_k \partial x_k} + B_i^f + M_i + B_i \tag{5.8.10}$$

where

$$M_i = -\frac{\partial \tau_{ik}}{\partial x_k}, \tau_{ij} = -\left[\widetilde{\tilde{u}_i \tilde{u}_j} - \widetilde{u_i u_j}\right] \tag{5.8.11}$$

$$B_i(\mathbf{x}) = -\int_{\partial \mathcal{D}_w} G(\mathbf{x}' - \mathbf{x}) \sigma_{ik}(\mathbf{x}') n_k d^2\mathbf{x}' \tag{5.8.12}$$

$$B_i^f(\mathbf{x}) = -\int_{\partial \mathcal{D}_f} G(\mathbf{x}' - \mathbf{x})[u_i(\mathbf{x}')u_k(\mathbf{x}')n_k + \sigma_{ik}(\mathbf{x}')]n_k d^2\mathbf{x}' \tag{5.8.13}$$

$$\sigma_{ij} = \nu \frac{\partial u_i}{\partial x_j} - \delta_{ij} p \tag{5.8.14}$$

In Equation 5.8.10, the terms M_i, B_i, and B_f need to be modeled or specified, since they do not involve $\tilde{\mathbf{u}}$ or its gradients. M_i is the subgrid force term and is equal to the divergence of the subgrid stress τ_{ij} (Equation 5.8.11). If $\mathbf{u}$ is decomposed into its resolved $\left(\tilde{\mathbf{u}}\right)$ and unresolved ($\mathbf{u}''$) parts as $\mathbf{u} = \tilde{\mathbf{u}} + \mathbf{u}''$, then we can express the subgrid stress (SGS) as $\tau_{ij} = \widetilde{\tilde{u}_i u_{j''}} + \widetilde{\tilde{u}_j u_{i''}} + \widetilde{u_{i''} u_{j''}}$. Thus, τ_{ij} is an unclosed term since it contains information about unresolved scales; modeling this subgrid stress is one of the big challenges in LES. In the following text, we will discuss ways to model τ_{ij} in terms of $\tilde{\mathbf{u}}$. The body force due to filtered stresses at the wall, B_i, involves viscous and pressure stresses at the wall. For pressure, typically $p = \tilde{p}$ is used at the wall, which is a good approximation since pressure does not vary significantly across the height of the boundary layer. Modeling the viscous traction $\tau_i^{\nu} = \nu \frac{\partial u_i}{\partial n}$ at the wall can be more challenging since velocity gradients are steep within the viscous layer and may not be captured if the filter width Δ is much larger than the viscous length scale at the wall, $l^+ = \nu/u\tau$, where $u_\tau = \sqrt{\left|\nu \frac{\partial \langle u \rangle_i}{\partial n}\right|}$ is the wall friction velocity. One of two approaches may be taken to model σ_{ij} at the wall (Figure 5.8.3). In the first approach, $\Delta \gg l^+$ is chosen, even closer to the wall, and the unresolved viscous traction τ_i^w is modeled in terms of $\tilde{u}_i$ using the "log law" for turbulent boundary layers. This approach, commonly referred to as "wall-modeled LES," is computationally inexpensive, but may not be accurate for nonequilibrium boundary layers, where the log law is not valid in general. In the second approach, Δ is reduced near the wall to $\Delta \sim l^+$ so

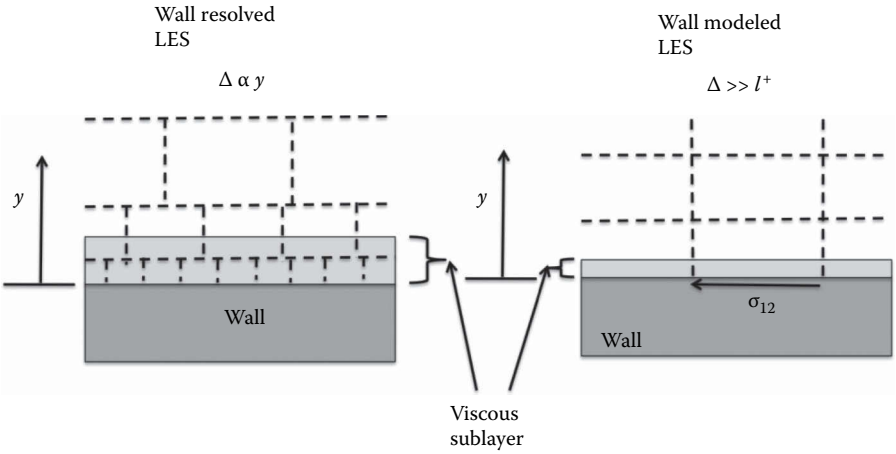


FIGURE 5.8.3 Schematic of wall-resolved (left) versus wall-modeled LES (right). The dashed lines indicate the grid cells in the mesh; note that Δ^g, the size of the grid cell may be different from Δ. The viscous sublayer, of thickness $O(l^+)$, is resolved by the grid in wall-resolved LES. In wall-modeled LES, Δ is much larger than the thickness of the viscous sublayer; in this case, the wall traction components, σ_{i2}, need to be modeled in terms of $\tilde{\mathbf{u}}$.

that the velocity gradients in the near-wall viscous layer are well captured by $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{u}} = 0$ can be applied at the wall. This approach, commonly referred to as "wall-resolved LES", while yielding higher levels of accuracy, is much more computationally expensive, and requires inhomogeneous filters, which in turn may lead to large commutation errors. The body force term due to stresses at inlet/outlet boundaries, B_i^f, does not usually require any modeling, since viscous stresses are negligible away from the wall, $p = \tilde{p}$ may be assumed, and $u_i(\mathbf{x}')u_k(\mathbf{x}') = \tilde{u}_i(\mathbf{x}')\tilde{u}_k(\mathbf{x}')$ may be assumed far from the wall.

Explicit Filtering

The filter kernel $G(\mathbf{x})$ appears explicitly in the second left-hand side term in Equation 5.8.10, where $\widetilde{\tilde{u}_i\tilde{u}_k}$ has to be obtained. The filter kernel is also required in Equations 5.8.12 and 5.8.13, besides being of importance in the "dynamic Smagorinsky model" (explained later on) for obtaining C_s dynamically. In LES, like in any numerical technique, the velocity field is discretized on a computational grid. Since filtering such a discretized field using Equation 5.8.5 essentially amounts to a numerical integration over a region with size Δ, explicit application of the kernel requires the cell width of the computational grid, Δ^g, to be much less than the filter width Δ. In such approaches, the filtering operation has to be discretized. In a single direction, using a three point stencil with $\Delta / \Delta_g = \sqrt{6}$, it can be shown that both the Gaussian and the box filters can be represented as (Garnier et al., 2009)

$$\tilde{\phi}_i = \frac{1}{4}(\phi_{i-1} + 2\phi_i + \phi_{i+1})$$

This discretization has an $O(\Delta_g^2)$ error associated with it. For inhomogeneous filters, alternate representations can be used to minimize commutation error to an arbitrary order of accuracy (Vasilyev et al., 1998). Usage of explicit filtering leads to extra computational cost, since the cell width Δ^g is smaller than the filter width Δ for such an approach. In practice, $\Delta^g = \Delta$ is used more often, especially since, as explained in the section "Filtering Due to Numerical Discretization," numerical discretization itself leads to spatial low-pass filtering, with $\Delta \approx \Delta^g$.

Subgrid Stress Models

A major focus of LES has been to construct models for the SGS, $\tau_{ij} = -\left[\widetilde{\tilde{u}_i\tilde{u}_j} - \widetilde{u_i u_j}\right]$ (Equation 5.8.11), in terms of the filtered velocity field $\tilde{u}_i$. Three popular SGS models are presented below.

Smagorinsky Model

The Smagorinsky model (SM) (Smagorinsky, 1963) for SGS is given by

$$\tau_{ij} = \frac{1}{3}\delta_{ij}\tau_{kk} - 2(C_s\Delta)^2 S\tilde{S}_{ij} \tag{5.8.15}$$

$$\text{where } \tilde{S} = \sqrt{2\tilde{S}_{ij}\tilde{S}_{ij}} \quad \tilde{S}_{ij} = \frac{1}{2}\left(\frac{\partial\tilde{u}_i}{\partial x_j} + \frac{\partial\tilde{u}_j}{\partial x_i}\right), C_s = 0.16 \tag{5.8.16}$$

Here, Δ is the filter width. For anisotropic flows, where the filter width may be different in the three directions (e.g., Δ_x, Δ_y, Δ_z), we can use $\Delta = \sqrt{\frac{\Delta_x^2 + \Delta_y^2 + \Delta_z^2}{3}}$, or $\Delta = (\Delta_x\Delta_y\Delta_z)^{1/3}$. In Equation 5.8.15, the trace of the subgrid stress, τ_{kk}, does not need to be modeled, as it can be absorbed into the pressure in Equation 5.8.10. The SM essentially uses an eddy viscosity to model SGS, where the length scale and velocity scale for defining the eddy viscosity are obtained from the filter width Δ and local strain rate S. This model ensures that the net dissipation rate to subgrid scales, given by $-\langle\tau_{ij}S_{ij}\rangle = (C_s\Delta)^2\langle S^3\rangle$ is always positive, as it should be, if Δ lies within the inertial range. Moreover, equating $-\langle\tau_{ij}S_{ij}\rangle = \epsilon$, where ϵ is the viscous dissipation rate $\epsilon = \nu\left\langle\frac{\partial u_{i'}}{\partial x_k}\frac{\partial u_i'}{\partial x_k}\right\rangle$ and $u_i' = u_i - \langle u_i\rangle$, it is possible to relate C_s to the universal constant in Kolmogorov's five-thirds law, valid for homogeneous isotropic turbulence. However, the energy cascade mechanism is not valid near the wall for scales that are comparable in size to the boundary layer thickness. Therefore, $C_s = 0.16$ does not give optimal results for wall-bounded flows with homogeneous filters, especially if Δ is made proportional to the boundary layer thickness.

For wall-resolved LES, where inhomogeneous filters are used and $\Delta \approx l^+$ at the wall, the known wall asymptotics of the velocity fluctuations is used to deduce how C_s needs to be damped at the wall. Continuity equation $\nabla\cdot\mathbf{u}' = 0$ at the wall implies that $v' \sim y^2$, while $u' \sim y$ and $w' \sim y$, where y is the distance from the wall. Since the mean strain rate $\langle S\rangle$ goes to a constant at the wall, it implies an $O(y^3)$ scaling for the eddy viscosity. On the other hand, the eddy viscosity from SM goes to a constant value at the wall (since Δ and S are finite at the wall). Therefore, in wall-resolved LES with SM, the eddy viscosity is multiplied by a Van Driest damping factor:

$$f_w(y) = \left[10 - \exp\left(-\frac{yu_\tau}{25\nu}\right)\right]$$

so that the modified SM model for SGS is

$$\tau_{ij} = \frac{1}{3}\delta_{ij}\tau_{kk} - 2f_w(y)(C_s\Delta)^2 S\tilde{S}_{ij} \tag{5.8.17}$$

Dynamic Smagorinsky Model

The model constant C_s in SM is calibrated against homogeneous isotropic turbulence. The dynamic Smagorinsky model (DSM) (Germano et al., 1991) uses the scale similar nature of turbulence to obtain the value of C_s from $\tilde{\mathbf{u}}(\mathbf{x},t)$ at every time step, making the model free of any imposed

constants. The basic principle behind the model is that due to the scale similar nature of turbulence, the model constant C_s is invariant to the filter width Δ. Therefore, in this model, a "test filter" with filter width $\overline{\Delta}$ is used, such that $\overline{\Delta} = \alpha\Delta$, where $\alpha > 1$. Typically, $\alpha = 2$ is used. The intrinsic shape of the test filter kernel, $\overline{[\cdot]}$, is usually kept the same as the original filter kernel $\widetilde{[\cdot]}$. A third filter, denoted as $\widehat{[\cdot]}$, is the test filter of the original filter $\widetilde{[\cdot]}$, e.g., $\hat{u}_i = \overline{\tilde{u}}_i$. The SGS model is then assumed to have the form given by SM at both $\widetilde{[\cdot]}$ and $\widehat{[\cdot]}$ levels:

$$T_{ij} = \widehat{u_i u_j} - \widehat{\hat{u}_i \hat{u}_j} = \frac{1}{3}\delta_{ij} T_{kk} - 2(C_s \overline{\Delta})^2 \hat{\bar{S}}\hat{\bar{S}}_{ij} \tag{5.8.18}$$

$$\tau_{ij} = \widetilde{u_i u_j} - \widetilde{\tilde{u}_i \tilde{u}_j} = \frac{1}{3}\delta_{ij} \tau_{kk} - 2(C_s \Delta)^2 \tilde{\bar{S}}\tilde{\bar{S}}_{ij} \tag{5.8.19}$$

where T_{ij} is the SGS if $\widehat{[\cdot]}$ is operated on Equation 5.8.2. Now, taking a test filter of Equation 5.8.19, we get

$$\overline{\tau}_{ij} = \overline{\widetilde{u_i u_j}} - \overline{\widetilde{\tilde{u}_i \tilde{u}_j}} = \frac{1}{3}\delta_{ij} \overline{\tau}_{kk} - 2(C_s \Delta)^2 \overline{\tilde{\bar{S}}\tilde{\bar{S}}_{ij}} \tag{5.8.20}$$

Subtracting Equation 5.8.20 from Equation 5.8.18, making it trace-free, and noting that $\widehat{[\cdot]} = \overline{\overline{[\cdot]}}$, we can get

$$\Rightarrow L_{ij} - \frac{\delta_{ij}}{3} L_{kk} = 2(C_s)^2 M_{ij} \tag{5.8.21}$$

where

$$L_{ij} = T_{ij} - \overline{\tau}_{ij} = \widehat{\tilde{u}_i \tilde{u}_j} - \widehat{\hat{u}_i \hat{u}_j}$$

is called as the Leonard stress and can be completely constructed from the resolved velocity field $\tilde{u}_i$, while

$$M_{ij} = \Delta^2 \left[\overline{\tilde{\bar{S}}\tilde{\bar{S}}_{ij}} - \alpha^2 \hat{\bar{S}}\hat{\bar{S}}_{ij} \right]$$

can be calculated from the strain rate of $\tilde{u}_i$. Minimizing the error between the left-hand side (LHS) and right-hand side (RHS) of Equation 5.8.21 over all statistically homogeneous directions, we get (Lilly, 1992)

$$C_s^2(\mathbf{x}) = \frac{\langle L_{ij}(\mathbf{x}) M_{ij}(\mathbf{x}) \rangle}{\langle M_{ij}(\mathbf{x}) M_{ij}(\mathbf{x}) \rangle} \tag{5.8.22}$$

In Equation 5.8.22, the averaging operator $\langle [\cdot] \rangle$ implies averaging over statistically homogenous directions at a given time. Thus, this procedure yields C_s dynamically from the instantaneous velocity field $\tilde{u}_i$. In simulations of fully developed turbulent channel flow, it has been found that C_s reaches a value of 0.16 at the center of the channel, where the flow is homogeneous and isotropic, while it reduces to zero near the wall, as desired. Note that a positive value of C_s^2 is required to keep the eddy viscosity positive, without which the scheme becomes unstable. Equation 5.8.22 does not however guarantee a positive value of C_s^2, and, in practice, its value is clipped at zero if it becomes negative.

Most engineering flows do not have statistically homogeneous directions. In such cases, a Lagrangian dynamic averaging procedure (Meneveau et al., 1996) can be used, where the error

between LHS and RHS of Equation 5.8.21 is minimized for a fluid tracer particle over some time T_{lag}. This involves solving two additional equations for the fields $\mathcal{J}_{LM}(\mathbf{x},t)$ and $\mathcal{J}_{MM}(\mathbf{x},t)$:

$$\frac{D\mathcal{J}_{LM}}{Dt} = \frac{1}{T_{\text{lag}}}(L_{ij}M_{ij} - \mathcal{J}_{LM})$$

$$\frac{D\mathcal{J}_{MM}}{Dt} = \frac{1}{T_{\text{lag}}}(M_{ij}M_{ij} - \mathcal{J}_{MM})$$

where $T_{\text{lag}}(\mathbf{x},t) = 15\Delta(\mathcal{J}_{MM}\mathcal{J}_{LM})^{-1/8}$. The model constant is then calculated dynamically as

$$C_s^2(\mathbf{x},t) = \frac{\mathcal{J}_{LM}}{\mathcal{J}_{MM}}$$

This procedure still does not guarantee that C_s^2 is positive (and therefore requires negative values to be clipped at zero); it is also computationally more expensive, because of the two additional equations that need to be solved.

Wall Adaptive Local Eddy Viscosity Model

One of the challenges in using DSM or SM models is that the subgrid stress may not be correctly damped at lower Reynolds numbers, especially since conventional wall damping functions (Equation 5.8.4.1) require distance from the wall in "+" units as an input. This can be a serious issue in flows where transition is important (e.g., flat plate boundary layer transitioning from laminar to turbulent state in the streamwise direction). In the wall adaptive local eddy viscosity model (WALE) (Nicoud and Ducros, 1999), an eddy viscosity is devised that vanishes in an appropriate way near the wall (i.e., like $\nu_t \sim y^3$) and at low Reynolds numbers. The model, which can only be used in a wall-resolved LES, is given as follows:

$$\tau_{ij} = \frac{\tau_{kk}}{3}\delta_{ij} - 2\nu_t\tilde{S}_{ij}$$

$$\nu_t = (C_w\Delta)^2 \frac{(\mathcal{L}_{ij}\mathcal{L}_{ji})^{3/2}}{(S_{kl}S_{kl})^{5/2} + (\mathcal{L}_{mn}\mathcal{L}_{mn})^{5/4}}$$

$$\mathcal{L}_{ij} = S_{ik}S_{kj} + \Omega_{ik}\Omega_{kj} - \frac{1}{3}(S_{ik}S_{ki} + \Omega_{ik}\Omega_{ki})\delta_{ij}$$

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \Omega_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}\right)$$

where $C_w = 0.5$ is the only model constant. Unlike SM or DSM, the subgrid stress also depends on both strain rate and vorticity. Moreover, the WALE model does not explicitly depend on the distance from the wall, and works better than SM or DSM in transitional flows.

Errors in LES

Several types of errors may be present in any practical LES simulation:

Modeling error: This is the most obvious error present when τ_{ij}, B_i, and B_i^f are not well represented by models. Usually, the *a priori* modeling errors between subgrid stresses calculated from DNS and LES can be quite large. Interestingly, LES performs well in spite of such large modeling errors (Meneveau and Katz, 1999). This issue is addressed by the "Optimal LES" (OLES) formulation (Langford and Moser, 1999), in which it is argued that the "best" subgrid stress model minimizes the *a priori* error, even though this error itself may be quite large. While

a practical OLES model has been formulated for homogeneous isotropic turbulence (Moser et al., 2009), this approach requires DNS data for wall-bounded flows (Völker et al., 2002; Bhattacharya et al., 2008).

Numerical discretization error: Numerical discretization errors (Ghosal, 1996) may be present due to the order of discretization of the differential operators in Equations 5.8.9 and 5.8.10. If a second-order discretization is used for the advection term, for instance, then the error will be $O((\Delta^g)^2)$. Since the subgrid stress models in LES are usually $O(\Delta^2)$ (e.g., Equation 5.8.15), higher order (third order or higher) discretization should ideally be used to discretize the terms in Equations 5.8.9 and 5.8.10, to ensure that the numerical discretization errors do not dominate the modeled subgrid stress terms. This issue is somewhat less serious if explicit filtering with large value of Δ/Δ^g is used.

Commutation error: As mentioned earlier, the filtering operator and spatial derivative operator do not commute for inhomogeneous filters; the resulting errors (e.g., $\widetilde{\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}x_i}} - \frac{\mathrm{d}\tilde{\mathbf{u}}}{\mathrm{d}x_i}$) are known as the commutation error in LES, and imply that additional terms will be present in Equations 5.8.9 and 5.8.10 for LES with inhomogeneous filters (Ghosal and Moin, 1995). It is possible to construct explicit filters carefully, so that commutation error is only $O((\Delta^g)^n)$, where n is the desired order of the error (Vasilyev et al., 1998). However, it is not possible to remove commutation error without explicit filtering approaches.

Aliasing error: This type of error occurs due to presence of nonlinear terms in Equation 5.8.10, in which the product of two functions represented by N spatial Fourier modes gives rise to a function with $2N$ Fourier modes (Ghosal, 1996). Such a product of functions cannot be represented accurately on a grid with N points due to the Nyquist sampling theorem. Explicit filtering can reduce aliasing errors since this procedure removes the high wavenumber modes ($k > \pi/\Delta$) from $\tilde{\mathbf{u}}$, thus ensuring that the computational grid can represent functions having higher number of Fourier modes compared to $\tilde{\mathbf{u}}$.

Wall Modeling

In the following discussion, we will consider flow along a wall, with the x-axis aligned along the streamwise direction, and the y-axis aligned along the wall normal direction. For simulating turbulent boundary layers at high Reynolds numbers flows, it is usually challenging to resolve the thin viscous layer near the wall. In such cases, the distance of the first off-wall grid point, y_1, is typically much larger than l^+; it is not possible to apply a no-slip boundary condition for $\mathbf{u}$ at this grid point. Instead, the wall traction $\sigma_{i2} = \nu \frac{\partial u_i}{\partial y}$ at $y=0$ has to be applied as a body force, as outlined in Equation 5.8.13. Typically, $\tilde{u}_2 = 0$ is imposed in such schemes, which a good approximation within a boundary layer. To obtain σ_{i2} for $i=1, 3$, the log law is utilized, which relates the friction velocity $u\tau$ at the wall to the average streamwise velocity $\langle u_1(\mathbf{x})\rangle|_{y=y_m}$, where the "matching distance" y_m lies in the "outer layer" ($y_m \sim 0.2\delta$) of the turbulent boundary layer, as follows:

$$\langle u_1 |_{y=y_m} \rangle = u_\tau \left[\frac{1}{\kappa} \log \frac{y_m u_\tau}{\nu} + B_i \right] \tag{5.8.23}$$

where $B_i = 5.5$. The friction velocity, $u\tau$, can be calculated from the aforementioned transcendental equation at every time step using iterative methods like the Newton–Raphson scheme (Piomelli and Balaras, 2002). In an LES simulation where the wall modeling approach is being used, typically $y_m = y_1$ is applied, and u_1 is replaced by $\tilde{u}_1$. Here y_1 is the wall normal distance of the first off-wall grid point from the wall. It has recently been reported (Kawai and Larsson, 2012) that $y_m = y_3$ (location of the third off-wall grid point) can give better results, since the SGS model does not work well near the wall. Thus, at a given x,z location, $u\tau$ can be obtained at every time step using Equation 5.8.23. $\langle \sigma_{12} \rangle$ is then obtained from (Piomelli and Balaras, 2002)

$$\langle \sigma_{12} \rangle = \text{sgn}[\langle \tilde{u}_1(y_m) \rangle] u_\tau^2 \tag{5.8.24}$$

where sgn[x] gives the sign of x. The fluctuating components $\sigma'_{i2} = \sigma_{i2} - \langle \sigma_{i2} \rangle$ are then made proportional to the wall-parallel fluctuating velocities as follows:

$$\sigma'_{i2}(x,z) = \tilde{u}'_i(\mathbf{x})\,|_{y=y_m} \frac{\langle \sigma_{12} \rangle}{\langle \tilde{u}_1 \rangle} \quad \text{for all } i = 1,3 \tag{5.8.25}$$

where $\tilde{u}'_i = \tilde{u}_i - \langle \tilde{u}_i \rangle$ is the fluctuating component of the filtered velocity. Again, $\langle [\cdot] \rangle$ may be obtained by averaging over the statistically homogeneous directions (if present). Otherwise, time averaging may also be used. The instantaneous wall-parallel tractions are then given by $\sigma_{12} = \langle \sigma_{12} \rangle + \sigma'_{12}$ and $\sigma_{32} = \sigma'_{32}$. Such wall modeling methods, while being computationally inexpensive, should ideally be used only if a high Reynolds number turbulent boundary layer is present near the wall; the method may not be accurate for separated flows and nonequilibrium boundary layers, for which the log law (Equation 5.8.23) does not apply.

Grid Generation

As mentioned earlier, the filter width, Δ, can vary with $\mathbf{x}$, and $\Delta^g(\mathbf{x})$ itself needs to be lower than or equal to $\Delta(\mathbf{x})$. Ideally, Δ should lie within the inertial range, i.e., $\mathcal{L} \gg \Delta \gg \eta$, where $\mathcal{L}$ is the local integral length scale of turbulence and η is the local Kolmogorov scale. It is, however, not easy to estimate $\mathcal{L}$ and η *a priori*. One way to set Δ is to base it on the length scale δ characterizing the flow locally. For example, for a turbulent jet, δ would correspond to the width of the jet at a given axial location, while, for a boundary layer, δ would correspond to the boundary layer thickness at that streamwise location. The variation of δ itself needs to be calculated *a priori* using standard correlations for boundary layers and free shear flows. For free shear flows, and at the outer edge of turbulent boundary layers (e.g., $y > 0.2\delta$), usually $\Delta \approx \delta/20$ can be used. While performing wall-resolved LES, one expects the most energetic scales near the wall to have size $\mathcal{L} \propto y$, and, therefore, Δ itself needs to increase linearly with y within the inner edge of the boundary later (i.e., $\Delta \propto y$) in *all* directions (Figure 5.8.3). For the first off-wall grid point, $\Delta^+ \approx 1$ is required in the wall normal direction. Since long streamwise streaks and rolls are expected in the viscous buffer layer, $\Delta^+ \approx 100$, $\Delta^+ \approx 20$ can be used in the streamwise and spanwise flow directions, respectively, at the first off-wall grid point (Choi and Moin, 2012). Again, an estimate of local $u\tau$, from known correlations, is required to obtain the values of Δ at the first off-wall grid point. Implementing a linear variation of Δ with y in all directions is usually impossible with structured grids, and is typically not implemented in practice. Usage of such a grid can lead to computational savings over DNS; the computational cost of LES for a turbulent boundary layer over a plate of length L using such a grid is proportional to $Re_L^{1.8}$, as opposed to $Re_L^{2.4}$ for DNS (Chapman, 1979). If wall modeling is used, then $\Delta \approx \delta/20$ can be used throughout the turbulent boundary layer, since, in this case, the dynamics inside the viscous buffer layer need not be resolved by $\tilde{\mathbf{u}}$ (Figure 5.8.3).

Case Study: LES of Channel Flow

Here we will briefly discuss results from LES of fully developed turbulent channel flow, in which a constant pressure gradient is used to drive the fluid in the x-direction, while the walls are present at $y=0$ and $y=2H$. The flow is statistically homogeneous in the x and z direction. The Reynolds number based on friction velocity is $Re\tau = u\tau H/\nu \approx 950$, which corresponds to a bulk Reynolds number of $Re_b = U_b H/\nu = 19{,}000$, where $U_b = \dfrac{1}{2H}\displaystyle\int_0^{2H} \langle u \rangle dy$ is the average bulk velocity. A second-order

pseudospectral solver with explicit time stepping (fourth order Runge–Kutta) is used to perform this simulation. The dynamic Smagorinsky model (DSM) is used with $\alpha=2$ for the test filter. The Smagorinsky constant C_s is calculated dynamically at every time step. In turbulent channel flow, statistical homogeneity is present in both the x- and z-directions, and therefore, Equation 5.8.22 can be used directly to compute C_s, with $\langle[\cdot]\rangle$ implying averaging in x- and z-directions. Homogeneous filter width is used, with implicit filtering, so that the grid cell width, Δ^g is the same as the characteristic filter width, Δ. The computational domain contains $167\times40\times93$ points, and periodic boundary condition is applied in the x- and z-directions. The normalized grid cell width in different directions are $\Delta_x/H=0.15$, $\Delta_z/H=0.1$, $\Delta_y/H=0.05$, $\Delta_y^+=47.5$, and Since $\Delta_y^+\gg1$, the viscous wall shear stress has to be modeled. The wall-modeled LES method is used to iteratively solve $u\tau$ (and therefore τ_w) in terms of the velocity at *iym*th off-wall grid point (Piomelli and Balaras, 2002; Kawai and Larsson, 2012). Here, we try $iym=1$ and $iym=3$ (matching at first and third off-wall grid points). The mean bulk velocity (and therefore Re_b) is fixed in the simulation, while $u\tau$ (or τ_w) is derived from the wall-modeled LES method (Piomelli and Balaras, 2002) at every time step. In Figures 5.8.4a and b, we observe that the mean velocity profile $U(y)=\langle u_1\rangle(y)$ compares well with DNS data, as well as the log law, for both the *iym* values. There is in fact not much effect of the matching location (*iym*) on the mean velocity profile.

In Figure 5.8.5, we compare the Reynolds stresses with DNS values (for the $iym=3$ case), and find a reasonable match between LES and DNS away from the wall ($y/H>0.25$). Near the wall, there is much higher discrepancy between LES and DNS Reynolds stresses, especially since the Reynolds stresses from filtered velocity is expected to be much lower than the corresponding DNS values near the wall. The disagreement between LES statistics (i.e., Reynolds stresses) with DNS statistics near the wall is not surprising, and has been observed earlier by several researchers in the past (Kawai and Larsson, 2012). In fact, if a homogeneous filter is used in wall-bounded flows, with a filter width proportional to the boundary layer thickness, then the matching of such statistics is possible only if DNS data is available *a priori* (Bhattacharya et al., 2008; Völker et al., 2002). In Figure 5.8.6a, we observe that $C_s\approx0.16$ in the log layer ($0.1<y/H<0.3$), $C_s\approx0.125$ in the center of the channel, and $C_s\approx0$ near the wall. This drastic reduction of C_s at the wall is in fact responsible for overprediction of normal Reynolds stresses near the wall, since C_s directly determines the energy dissipation rate $-\langle\tau_{ij}S_{ij}\rangle=(C_s\Delta)^2\langle S^3\rangle$ of the large-scale turbulent kinetic energy by the SGS. Nevertheless, it appears that, after an initial transient, the $u\tau$ is predicted reasonably well by the LES (Figure 5.8.6b), with less than 5% error in τ_w seen for both $iym=1$ and $iym=3$ cases.

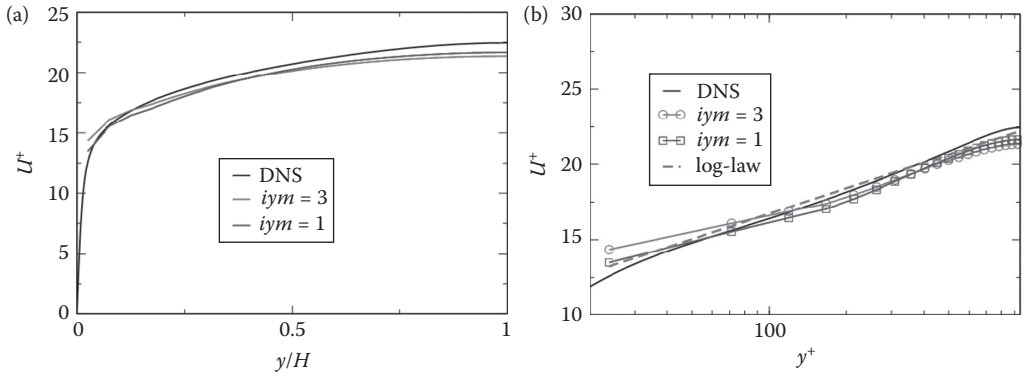


FIGURE 5.8.4 Mean velocity U^+ plotted with respect to u mean y/H on the normal scale and u mean log y^+ in the semi-log scale. Results are presented for matching points located at the first off-wall grid point ($iym=1$) and the third off-wall grid point ($iym=3$).

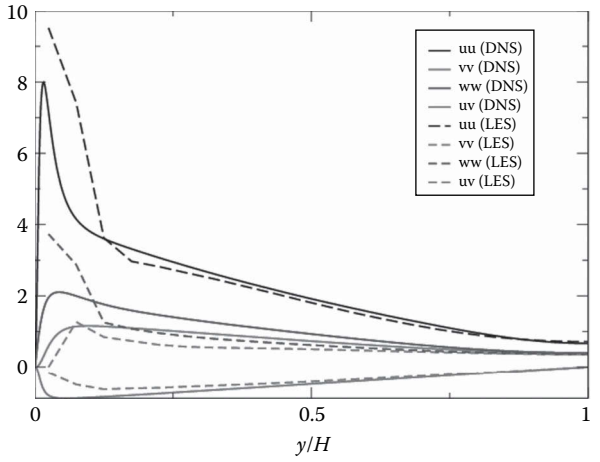


FIGURE 5.8.5 Reynolds stresses from LES compared to DNS data for turbulent channel flow.

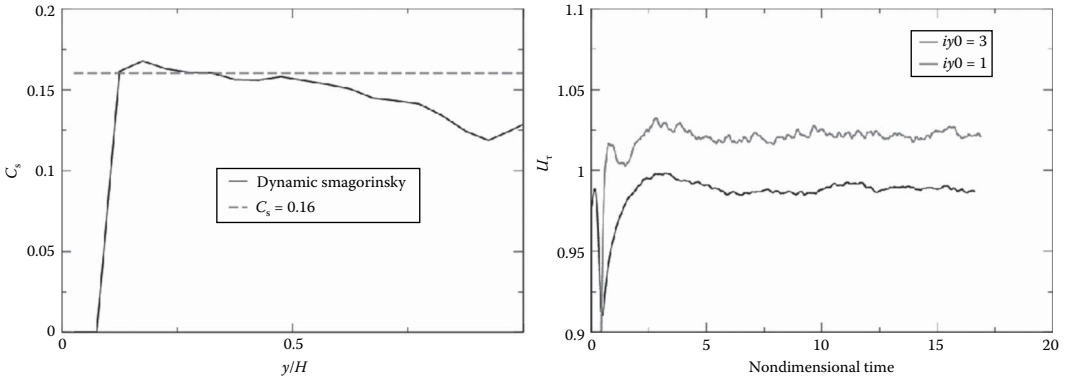


FIGURE 5.8.6 (a) C_s vs. y plotted along with reference value of C_s=0.16 and (b) $u\tau$ vs. time for matching point located at first off-wall grid point (iym=1) and third off-wall grid point (iym=3). The value of $u\tau$ = 1 in DNS.

Summary

We summarize the main points of the chapter as follows:

- LES solves equations for spatially filtered "large scales" of the velocity field, and models the effect of the small scales on the large scale, usually using eddy viscosity. It is more accurate than RANS since it attempts to model only the universal part of the energy spectrum. However, it is also much more expensive than RANS due to the larger degrees of freedom that need to be simulated.
- Ideally, explicit filtering, with a large value of Δ/Δ_g should be used to reduce errors. However, in practice, explicit filtering is not used due to the higher degrees of freedom (and therefore larger computational costs) involved.
- Wall-resolved LES models are somewhat less expensive than DNS. Choosing the correct mesh and grid size is crucial toward obtaining accurate results.
- Wall-modeled LES models are extremely inexpensive compared to wall-resolved LES. However, they may not be suitable for separated flows and equilibrium boundary layers.

REFERENCES

Bhattacharya, A., A. Das, and R. D. Moser (2008). A filtered-wall formulation for large-eddy simulation of wall-bounded turbulence. *Physics of Fluids 20* (11), 115104.

Chapman, D. R. (1979). Computational aerodynamics development and outlook. *AIAA Journal 17* (12), 1293–1313.

Choi, H. and P. Moin (2012). Grid-point requirements for large eddy simulation: Chapmans estimates revisited. *Physics of Fluids 24* (1), 011702.

Garnier, E., N. Adams, and P. Sagaut (2009). *Large Eddy Simulation for Compressible Flows*. Springer Science & Business Media, New York.

Germano, M., U. Piomelli, P. Moin, and W. H. Cabot (1991). A dynamic subgrid-scale eddy viscosity model. *Physics of Fluids A: Fluid Dynamics 3* (7), 1760–1765.

Ghosal, S. (1996). An analysis of numerical errors in large-eddy simulations of turbulence. *Journal of Computational Physics 125* (1), 187–206.

Ghosal, S. and P. Moin (1995). The basic equations for the large eddy simulation of turbulent flows in complex geometry. *Journal of Computational Physics 118* (1), 24–37.

Kawai, S. and J. Larsson (2012). Wall-modeling in large eddy simulation: Length scales, grid resolution, and accuracy. *Physics of Fluids 24* (1), 015105.

Kolmogorov, A. N. (1941). The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. In *Doklady Akademii Nauk Soyuz Sovietskich Socialisticeskich Respublik, 30*, 301–305. JSTOR.

Langford, J. A. and R. D. Moser (1999). Optimal les formulations for isotropic turbulence. *Journal of Fluid Mechanics 398* (321–346), 41.

Lilly, D. K. (1992). A proposed modification of the Germano subgrid-scale closure method. *Physics of Fluids A: Fluid Dynamics 4* (3), 633–635.

Meneveau, C. and J. Katz (1999). Dynamic testing of subgrid models in large eddy simulation based on the germano identity. *Physics of Fluids 11* (2), 245–247.

Meneveau, C., T. S. Lund, and W. H. Cabot (1996). A Lagrangian dynamic subgrid-scale model of turbulence. *Journal of Fluid Mechanics 319*, 353–385.

Moin, P. and K. Mahesh (1998). Direct numerical simulation: a tool in turbulence research. *Annual Review of Fluid Mechanics 30* (1), 539–578.

Moser, R. D., N. P. Malaya, H. Chang, P. S. Zandonade, P. Vedula, A. Bhattacharya, and A. Haselbacher (2009). Theoretically based optimal large-eddy simulation. *Physics of Fluids 21* (10), 105104.

Nicoud, F. and F. Ducros (1999). Subgrid-scale stress modelling based on the square of the velocity gradient tensor. *Flow, Turbulence and Combustion 62* (3), 183–200.

Piomelli, U. and E. Balaras (2002). Wall-layer models for large-eddy simulations. *Annual Review of Fluid Mechanics 34* (1), 349–374.

Pope, S. B. (2000). *Turbulent Flows*. Cambridge: Cambridge University Press.

Sagaut, P. (2006). *Large Eddy Simulation for Incompressible Flows: An Introduction*. Springer Science & Business Media, Berlin.

Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. The basic experiment. *Monthly Weather Review 91* (3), 99–164.

Vasilyev, O. V., T. S. Lund, and P. Moin (1998). A general class of commutative filters for les in complex geometries. *Journal of Computational Physics 146* (1), 82–104.

Völker, S., R. Moser, and P. Venugopal (2002). Optimal large eddy simulation of turbulent channel flow based on direct numerical simulation statistical data. *Physics of Fluids 14* (10), 3675–3691.

5.9 SOFTWARE AND COMPUTER CODES

Atul Sharma

Although various proprietary (or commercial) and free (or open-source) codes as well as software packages are available for problems in fluid dynamics and heat transfer, more commonly used ones are presented in this subsection—separately for the mathematical software, and CFD software and codes.

Mathematical Software

Some of the popular commercial mathematical software, for problems in heat transfer, are MATLAB® (programming, modeling, and simulation tool, developed by the MathWorks, Natick, MA, USA: http://www.mathworks.com), Maple (a general-purpose computer algebra system, developed by Maplesoft, Waterloo, Canada: http://www.maplesoft.com), and Mathematica (a computational software program based on symbolic mathematics, developed by Wolfram Research, Champaign, IL, USA: http://www.wolfram.com).

Widely used *open-source* mathematical software, for a heat-transfer study, are Scilab, acquired by ESI Group, Paris, France (for numerical computation and simulation: http://www.scilab.org), GNU Octave (mathematical modeling and simulation software http://www.gnu.org/software/octave), and Sage (System for Algebra and Geometry Experimentation: http://www.sagemath.org). Note that the mathematical software presented here are general-purpose, not only applicable to heat transfer but also applied for various problems in science and engineering.

Computational Fluid Dynamics Software and Codes

Commercial CFD software, widely used for a heat-transfer study, are FLUENT and CFX (both from ANSYS, Canonsburg, PA, USA: http://www.ansys.com), STAR-CD (from CD-adapco now part of Siemens PLM Software, Plano, TX, USA: http://mdx.plm.automation.siemens.com), CFD-ACE+ (from ESI Group, Paris, France: http://www.esi-cfd.com), TransAT (from ASCOMP GmbHa, Zurich, Switzerland: http://www.ascomp.ch), and solidworks Flow Simulation (from solidworks, Waltham, MA, USA: http://www.solidworks.com. The CFD solver for all these software packages is based on a finite-volume method-based algebraic formulation on a structured/block-structured or unstructured grid. A commercial software based on finite element method is COMSOL Multiphysics (from COMSOL, Burlington, MA, USA; for various applications including CFD: http://www.comsol.com), and a software based on the lattice Boltzmann method is PowerFLOW (from Exa Corporation, Burlington, MA, USA: http://www.exa.com).

Widely used *open-source* CFD codes/software are OpenFOAM, acquired by ESI Group, Paris, France (http://www.openfoam.com; see also http://www.extend-project.de), SU2 (Stanford University Unstructured: http://su2.stanford.edu), and PyFR (developed by the Vincent Lab at Imperial College London: http://www.pyfr.org). For multiphase flow, a freely available, very advanced state-of-the art adaptive 3D VOF method-based code is the Gerris Flow Solver, created by Stéphane Popinet, UPMC, Paris, France (http://www.gfs.sourceforge.net) and its successor Basilisk (http://basilisk.fr).

Appendix A: Properties of Gases and Vapors

Paul Norton

TABLE A.1
Properties of Dry Air at Atmospheric Pressure

Symbols and Units

K = Absolute temperature, °K

C = Temperature, °C

F = Temperature, °F

ρ = Density, kg/m^3

c_p = Specific heat capacity, kJ/kg·K

c_p/c_v = Specific heat capacity ratio, dimensionless

μ = Viscosity, N·s/m^2 × 10^6 (for N·s/m^2 (= kg/m·s) multiply tabulated values by 10^{-6})

k = Thermal conductivity, W/m·K × 10^3 (for W/m·K multiply tabulated values by 10^{-3})

Pr = Prandtl number, dimensionless

H = Enthalpy, kJ/kg

V_s = Sound velocity, m/s

Temperature			Properties							
K	°C	°F	ρ	c_p	c_p/c_v	μ	k	Pr	h	V_s
100	−173.15	−280	3.598	1.028		6.929	9.248	0.770	98.42	198.4
110	−163.15	−262	3.256	1.022	1.4202	7.633	10.15	0.768	108.7	208.7
120	−153.15	−244	2.975	1.017	1.4166	8.319	11.05	0.766	118.8	218.4
130	−143.15	−226	2.740	1.014	1.4139	8.990	11.94	0.763	129.0	227.6
140	−133.15	−208	2.540	1.012	1.4119	9.646	12.84	0.761	139.1	236.4
150	−123.15	−190	2.367	1.010	1.4102	10.28	13.73	0.758	149.2	245.0
160	−113.15	−172	2.217	1.009	1.4089	10.91	14.61	0.754	159.4	253.2
170	−103.15	−154	2.085	1.008	1.4079	11.52	15.49	0.750	169.4	261.0
180	−93.15	−136	1.968	1.007	1.4071	12.12	16.37	0.746	179.5	268.7
190	−83.15	−118	1.863	1.007	1.4064	12.71	17.23	0.743	189.6	276.2
200	−73.15	−100	1.769	1.006	1.4057	13.28	18.09	0.739	199.7	283.4
205	−68.15	−91	1.726	1.006	1.4055	13.56	18.52	0.738	204.7	286.9
210	−63.15	−82	1.684	1.006	1.4053	13.85	18.94	0.736	209.7	290.5
215	−58.15	−73	1.646	1.006	1.4050	14.12	19.36	0.734	214.8	293.9
220	−53.15	−64	1.607	1.006	1.4048	14.40	19.78	0.732	219.8	297.4
225	−48.15	−55	1.572	1.006	1.4046	14.67	20.20	0.731	224.8	300.8
230	−43.15	−46	1.537	1.006	1.4044	14.94	20.62	0.729	229.8	304.1
235	−38.15	−37	1.505	1.006	1.4042	15.20	21.04	0.727	234.9	307.4
240	−33.15	−28	1.473	1.005	1.4040	15.47	21.45	0.725	239.9	310.6
245	−28.15	−19	1.443	1.005	1.4038	15.73	21.86	0.724	244.9	313.8
250	−23.15	−10	1.413	1.005	1.4036	15.99	22.27	0.722	250.0	317.1
255	−18.15	−1	1.386	1.005	1.4034	16.25	22.68	0.721	255.0	320.2

(*Continued*)

TABLE A.1 (*Continued*)
Properties of Dry Air at Atmospheric Pressure

Temperature			Properties							
K	°C	°F	ρ	c_p	c_p/c_v	μ	k	Pr	h	V_s
260	−13.15	8	1.359	1.005	1.4032	16.50	23.08	0.719	260.0	323.4
265	−8.15	17	1.333	1.005	1.4030	16.75	23.48	0.717	265.0	326.5
270	−3.15	26	1.308	1.006	1.4029	17.00	23.88	0.716	270.1	329.6
275	+1.85	35	1.285	1.006	1.4026	17.26	24.28	0.715	275.1	332.6
280	6.85	44	1.261	1.006	1.4024	17.50	24.67	0.713	280.1	335.6
285	11.85	53	1.240	1.006	1.4022	17.74	25.06	0.711	285.1	338.5
290	16.85	62	1.218	1.006	1.4020	17.98	25.47	0.710	290.2	341.5
295	21.85	71	1.197	1.006	1.4018	18.22	25.85	0.709	295.2	344.4
300	26.85	80	1.177	1.006	1.4017	18.46	26.24	0.708	300.2	347.3
305	31.85	89	1.158	1.006	1.4015	18.70	26.63	0.707	305.3	350.2
310	36.85	98	1.139	1.007	1.4013	18.93	27.01	0.705	310.3	353.1
315	41.85	107	1.121	1.007	1.4010	19.15	27.40	0.704	315.3	355.8
320	46.85	116	1.103	1.007	1.4008	19.39	27.78	0.703	320.4	358.7
325	51.85	125	1.086	1.008	1.4006	19.63	28.15	0.702	325.4	361.4
330	56.85	134	1.070	1.008	1.4004	19.85	28.53	0.701	330.4	364.2
335	61.85	143	1.054	1.008	1.4001	20.08	28.90	0.700	335.5	366.9
340	66.85	152	1.038	1.008	1.3999	20.30	29.28	0.699	340.5	369.6
345	71.85	161	1.023	1.009	1.3996	20.52	29.64	0.698	345.6	372.3
350	76.85	170	1.008	1.009	1.3993	20.75	30.03	0.697	350.6	375.0
355	81.85	179	0.9945	1.010	1.3990	20.97	30.39	0.696	355.7	377.6
360	86.85	188	0.9805	1.010	1.3987	21.18	30.78	0.695	360.7	380.2
365	91.85	197	0.9672	1.010	1.3984	21.38	31.14	0.694	365.8	382.8
370	96.85	206	0.9539	1.011	1.3981	21.60	31.50	0.693	370.8	385.4
375	101.85	215	0.9413	1.011	1.3978	21.81	31.86	0.692	375.9	388.0
380	106.85	224	0.9288	1.012	1.3975	22.02	32.23	0.691	380.9	390.5
385	111.85	233	0.9169	1.012	1.3971	22.24	32.59	0.690	386.0	393.0
390	116.85	242	0.9050	1.013	1.3968	22.44	32.95	0.690	391.0	395.5
395	121.85	251	0.8936	1.014	1.3964	22.65	33.31	0.689	396.1	398.0
400	126.85	260	0.8822	1.014	1.3961	22.86	33.65	0.689	401.2	400.4
410	136.85	278	0.8608	1.015	1.3953	23.27	34.35	0.688	411.3	405.3
420	146.85	296	0.8402	1.017	1.3946	23.66	35.05	0.687	421.5	410.2
430	156.85	314	0.8207	1.018	1.3938	24.06	35.75	0.686	431.7	414.9
440	166.85	332	0.8021	1.020	1.3929	24.45	36.43	0.684	441.9	419.6
450	176.85	350	0.7842	1.021	1.3920	24.85	37.10	0.684	452.1	424.2
460	186.85	368	0.7677	1.023	1.3911	25.22	37.78	0.683	462.3	428.7
470	196.85	386	0.7509	1.024	1.3901	25.58	38.46	0.682	472.5	433.2
480	206.85	404	0.7351	1.026	1.3892	25.96	39.11	0.681	482.8	437.6
490	216.85	422	0.7201	1.028	1.3881	26.32	39.76	0.680	493.0	442.0
500	226.85	440	0.7057	1.030	1.3871	26.70	40.41	0.680	503.3	446.4
510	236.85	458	0.6919	1.032	1.3861	27.06	41.06	0.680	513.6	450.6
520	246.85	476	0.6786	1.034	1.3851	27.42	41.69	0.680	524.0	454.9
530	256.85	494	0.6658	1.036	1.3840	27.78	42.32	0.680	534.3	459.0
540	266.85	512	0.6535	1.038	1.3829	28.14	42.94	0.680	544.7	463.2
550	276.85	530	0.6416	1.040	1.3818	28.48	43.57	0.680	555.1	467.3
560	286.85	548	0.6301	1.042	1.3806	28.83	44.20	0.680	565.5	471.3
570	296.85	566	0.6190	1.044	1.3795	29.17	44.80	0.680	575.9	475.3

(*Continued*)

TABLE A.1 (*Continued*)
Properties of Dry Air at Atmospheric Pressure

Temperature			Properties							
K	°C	°F	ρ	c_p	c_p/c_v	μ	k	Pr	h	V_s
580	306.85	584	0.6084	1.047	1.3783	29.52	45.41	0.680	586.4	479.2
590	316.85	602	0.5980	1.049	1.3772	29.84	46.01	0.680	596.9	483.2
600	326.85	620	0.5881	1.051	1.3760	30.17	46.61	0.680	607.4	486.9
620	346.85	656	0.5691	1.056	1.3737	30.82	47.80	0.681	628.4	494.5
640	366.85	692	0.5514	1.061	1.3714	31.47	48.96	0.682	649.6	502.1
660	386.85	728	0.5347	1.065	1.3691	32.09	50.12	0.682	670.9	509.4
680	406.85	764	0.5189	1.070	1.3668	32.71	51.25	0.683	692.2	516.7
700	426.85	800	0.5040	1.075	1.3646	33.32	52.36	0.684	713.7	523.7
720	446.85	836	0.4901	1.080	1.3623	33.92	53.45	0.685	735.2	531.0
740	466.85	872	0.4769	1.085	1.3601	34.52	54.53	0.686	756.9	537.6
760	486.85	908	0.4643	1.089	1.3580	35.11	55.62	0.687	778.6	544.6
780	506.85	944	0.4524	1.094	1.3559	35.69	56.68	0.688	800.5	551.2
800	526.85	980	0.4410	1.099	1.354	36.24	57.74	0.689	822.4	557.8
850	576.85	1,070	0.4152	1.110	1.349	37.63	60.30	0.693	877.5	574.1
900	626.85	1,160	0.3920	1.121	1.345	38.97	62.76	0.696	933.4	589.6
950	676.85	1,250	0.3714	1.132	1.340	40.26	65.20	0.699	989.7	604.9
1,000	726.85	1,340	0.3529	1.142	1.336	41.53	67.54	0.702	1,046	619.5
1,100	826.85	1,520	0.3208	1.161	1.329	43.96			1,162	648.0
1,200	926.85	1,700	0.2941	1.179	1.322	46.26			1,279	675.2
1,300	1,026.85	1,880	0.2714	1.197	1.316	48.46			1,398	701.0
1,400	1,126.85	2,060	0.2521	1.214	1.310	50.57			1,518	725.9
1,500	1,220.85	2,240	0.2353	1.231	1.304	52.61			1,640	749.4
1,600	1,326.85	2,420	0.2206	1.249	1.299	54.57			1,764	772.6
1,800	1,526.85	2,780	0.1960	1.288	1.288	58.29			2,018	815.7
2,000	1,726.85	3,140	0.1764	1.338	1.274				2,280	855.5
2,400	2,126.85	3,860	0.1467	1.574	1.238				2,853	924.4
2,800	2,526.85	4,580	0.1245	2.259	1.196				3,599	983.1

Source: Tables of thermal properties of gases, National Bureau of Standards Circular 564, U.S Government Printing Office, November 1955.

TABLE A.2
Ideal Gas Properties of Nitrogen, Oxygen, and Carbon Dioxide

Symbols and Units

T = Absolute temperature, °K

$\bar{h}$ = Enthalpy, kJ/kmol

$\bar{u}$ = Internal energy, kJ/kmol

$\bar{s}^{\circ}$ = Absolute entropy at standard reference pressure, kJ/kmol K

$\bar{h}$ = Enthalpy of formation per mole at standard state = 0 kJ/kmol

T	$\bar{h}$	$\bar{u}$	$\bar{s}^{o}$	T	$\bar{h}$	$\bar{u}$	$\bar{s}^{o}$
			Part a. Ideal Gas Properties of Nitrogen, N_2				
0	0	0	0	600	17,563	12,574	212.066
220	6,391	4,562	182.639	610	17,864	12,792	212.564
230	6,683	4,770	183.938	620	18,166	13,011	213.055
240	6,975	4,979	185.180	630	18,468	13,230	213.541
250	7,266	5,188	186.370	640	18,772	13,450	214.018
260	7,558	5,396	187.514	650	19,075	13,671	214.489
270	7,849	5,604	188.614	660	19,380	13,892	214.954
280	8,141	5,813	189.673	670	19,685	14,114	215.413
290	8,432	6,021	190.695	680	19,991	14,337	215.866
298	8,669	6,190	191.502	690	20,297	14,560	216.314
300	8,723	6,229	191.682	700	20,604	14,784	216.756
310	9,014	6,437	192.638	710	20,912	15,008	217.192
320	9,306	6,645	193.562	720	21,220	15,234	217.624
330	9,597	6,853	194.459	730	21,529	15,460	218.059
340	9,888	7,061	195.328	740	21,839	15,686	218.472
350	10,180	7,270	196.173	750	22,149	15,913	218.889
360	10,471	7,478	196.995	760	22,460	16,141	219.301
370	10,763	7,687	197.794	770	22,772	16,370	219.709
380	11,055	7,895	198.572	780	23,085	16,599	220.113
390	11,347	8,104	199.331	790	23,398	16,830	220.512
400	11,640	8,314	200.071	800	23,714	17,061	220.907
410	11,932	8,523	200.794	810	24,027	17,292	221.298
420	12.225	8,733	201.499	820	24,342	17,524	221.684
430	12,518	8,943	202.189	830	24,658	17,757	222.067
440	12,811	9,153	202.863	840	24,974	17,990	222.447
450	13,105	9,363	203.523	850	25,292	18,224	222.822
460	13,399	9,574	204.170	860	25,610	18,459	223.194
470	13,693	9,786	204.803	870	25,928	18,695	223.562
480	13,988	9,997	205.424	880	26,248	18,931	223.927
490	14,285	10,210	206.033	890	26,568	19,168	224.288
500	14,581	10,423	206.630	900	26,890	19,407	224.647
510	14,876	10,635	207.216	910	27,210	19,644	225.002
520	15,172	10,848	207.792	920	27,532	19,883	225.353
530	15,469	11,062	208.358	930	27,854	20,122	225.701
540	15,766	11,277	208.914	940	28,178	20,362	226.047
550	16,064	11,492	209.461	950	28,501	20,603	226.389
560	16,363	11,707	209.999	960	28,826	20,844	226.728
570	16,662	11,923	210.528	970	29,151	21,086	227.064
580	16,962	12,139	211.049	980	29,476	21,328	227.398

(*Continued*)

TABLE A.2 (*Continued*)
Ideal Gas Properties of Nitrogen, Oxygen, and Carbon Dioxide

T	$\bar{h}$	$\bar{u}$	$\bar{s}^o$	T	$\bar{h}$	$\bar{u}$	$\bar{s}^o$
590	17,26_	12,356	211.562	990	29,803	21,571	227.728
1,000	30,129	21,815	228.057	1,760	56,227	41.594	247.396
1,020	30,784	22,304	228.706	1,780	56,938	42,139	247.798
1,040	31,442	22,795	229.344	1,800	57,651	42,685	248.195
1,060	32,101	23,288	229.973	1,820	58,363	43,231	248.589
1,080	32,762	23,782	230.591	1,840	59,075	43,777	248.979
1,100	33,426	24,280	231.199	1,860	59,790	44,324	249.365
1,120	34,092	24,780	231.799	1,880	60,504	44,873	249.748
1,140	34,760	25,282	232.391	1,900	61,220	45,423	250.128
1,160	35,430	25,786	232.973	1,920	61,936	45,973	250.502
1,180	36,104	26,291	233.549	1,940	62,654	46,524	250.874
1,200	36,777	26,799	234.115	1,960	63,381	47,075	251.242
1,220	37,452	27,308	234.673	1,980	64,090	47,627	251.607
1,240	38,129	27,819	235.223	2,000	64,810	48,181	251.969
1,260	38,807	28,331	235.766	2,050	66,612	49,567	252.858
1,280	39,488	28,845	236.302	2,100	68,417	50,957	253.726
1,300	40,170	29,361	236.831	2,150	70,226	52,351	254.578
1,320	40,853	29,878	237.353	2,200	72,040	53,749	255.412
1,340	41,539	30,398	237.867	2,250	73,856	55,149	256.227
1,360	42,227	30,919	238.376	2,300	75,676	56,553	257.027
1,380	42,915	31,441	238.878	2,350	77,496	57,958	257.810
1,400	43,605	31,964	239.375	2,400	79,320	59,366	258.580
1,420	44,295	32,489	239.865	2,450	81,149	60,779	259.332
1,440	44,988	33,014	240.350	2,500	82,981	62,195	260.073
1,460	45,682	33,543	240.827	2,550	84,814	63,613	260.799
1,480	46,377	34,071	241.301	2,600	86,650	65,033	261.512
1,500	47,073	34,601	241.768	2,650	88,488	66,455	262.213
1,520	47,771	35,133	242.228	2,700	90,328	67,880	262.902
1,540	48,470	35,665	242.685	2,750	92,171	69,306	263.577
1,560	49,168	36,197	243.137	2,800	94,014	70,734	264.241
1,580	49,869	36,732	243.585	2,850	95,859	72,163	264.895
1,600	50,571	37,268	244.028	2,900	97,705	73,593	265.538
1,620	51,275	37,806	244.464	2,950	99,556	75,028	266.170
1,640	51,980	38,344	244.896	3,000	101,407	76,464	266.793
1,660	52,686	38,884	245.324	3,050	103,260	77,902	267.404
1,680	53,393	39,424	245.747	3,100	105,115	79,341	268.007
1,700	54,099	39,965	246.166	3,150	106,972	80,782	268.601
1,720	54,807	40,507	246.580	3,200	108,830	82,224	269.186
1,740	55,516	41,049	246.990	3,250	110,690	83,668	269.763
			Part b. Ideal Gas Properties of Oxygen, O_2				
0	0	0	0	600	17,929	12,940	226.346
220	6,404	4,575	196.171	610	18,250	13,178	226.877
230	6,694	4,782	197.461	620	18,572	13,417	227.400
240	6,984	4,989	198.696	630	18,895	13,657	227.918
250	7,275	5,197	199.885	640	19,219	13,898	228.429
260	7,566	5,405	201.027	650	19,544	14,140	228.932
270	7,858	5,613	202.128	660	19,870	14,383	229.430
280	8,150	5,822	203.191	670	20,197	14,626	229.920

(*Continued*)

TABLE A.2 (*Continued*)
Ideal Gas Properties of Nitrogen, Oxygen, and Carbon Dioxide

T	$\bar{h}$	$\bar{u}$	$\bar{s}^o$	T	$\bar{h}$	$\bar{u}$	$\bar{s}^o$
290	8,443	6,032	204.218	680	20,524	14,871	230.405
298	8,682	6,203	205.033	690	20,854	15,116	230.885
300	8,736	6,242	205.213	700	21,184	15,364	231.358
310	9,030	6,453	206.177	710	21,514	15,611	231.827
320	9,325	6,664	207.112	720	21,845	15,859	232.291
330	9,620	6,877	208.020	730	22,177	16,107	232.748
340	9,916	7,090	208.904	740	22,510	16,357	233.201
350	10,213	7,303	209.765	750	22,844	16,607	233.649
360	10,511	7,518	210.604	760	23,178	16,859	234.091
370	10,809	7,733	211.423	770	23,513	17,111	234.528
380	11,109	7,949	212.222	780	23,850	17,364	234.960
390	11,409	8,166	213.002	790	24,186	17,618	235.387
400	11,711	8,384	213.765	800	24,523	17,872	235.810
410	12,012	8,603	214.510	810	24,861	18,126	236.230
420	12,314	8,822	215.241	820	25,199	18,382	236.644
430	12,618	9,043	215.955	830	25,537	18,637	237.055
440	12,923	9,264	216.656	840	25,877	18,893	237.462
450	13,228	9,487	217.342	850	26,218	19,150	237.864
460	13,535	9,710	218.016	860	26,559	19,408	238.264
470	13,842	9,935	218.676	870	26,899	19,666	238.660
480	14,151	10,160	219.326	880	27,242	19,925	239.051
490	14,460	10,386	219.963	890	27,584	20,185	239.439
500	14,770	10,614	220.589	900	27,928	20,445	239.823
510	15,082	10,842	221.206	910	28,272	20,706	240.203
520	15,395	11,071	221.812	920	28,616	20,967	240.580
530	15,708	11,301	222.409	930	28,960	21,228	240.953
540	16,022	11,533	222.997	940	29,306	21,491	241.323
550	16,338	11,765	223.576	950	29,652	21,754	241.689
560	16,654	11,998	224.146	960	29,999	22,017	242.052
570	16,971	12,232	224.708	970	30,345	22,280	242.411
580	17,290	12,467	225.262	980	30,692	22,544	242.768
590	17,609	12,703	225.808	990	31,041	22,809	243.120
1,000	31,389	23,075	243.471	1,760	58,880	44,247	263.861
1,020	32,088	23,607	244.164	1,780	59,624	44,825	264.283
1,040	32,789	24,142	244.844	1,800	60,371	45,405	264.701
1,060	33,490	24,677	245.513	1,820	61,118	45,986	265.113
1,080	34,194	25,214	246.171	1,840	61,866	46,568	265.521
1,100	34,899	25,753	246.818	1,860	62,616	47,151	265.925
1,120	35,606	26,294	247.454	1,880	63,365	47,734	266.326
1,140	36,314	26,836	248.081	1,900	64,116	48,319	266.722
1,160	37,023	27,379	248.698	1,920	64,868	48,904	267.115
1,180	37,734	27,923	249.307	1,940	65,620	49,490	267.505
1,200	38,447	28,469	249.906	1,960	66,374	50,078	267.891
1,220	39,162	29,018	250.497	1,980	67,127	50,665	268.275
1,240	39,877	29,568	251.079	2,000	67,881	51,253	268.655
1,260	40,594	30,118	251.653	2,050	69,772	52,727	269.588
1,280	41,312	30,670	252.219	2,100	71,668	54,208	270.504
1,300	42,033	31,224	252.776	2,150	73,573	55,697	271.399
1,320	42,753	31,778	253.325	2,200	75,484	57,192	272.278
1,340	43,475	32,334	253.868	2,250	77,397	58,690	273.136
1,360	44,198	32,891	254.404	2,300	79,316	60,193	273.981
1,380	44,923	33,449	254.932	2,350	81,243	61,704	274.809
1,400	45,648	34,008	255.454	2,400	83,174	63,219	275.625

(*Continued*)

TABLE A.2 (*Continued*)
Ideal Gas Properties of Nitrogen, Oxygen, and Carbon Dioxide

T	$\bar{h}$	$\bar{u}$	$\bar{s}^o$	T	$\bar{h}$	$\bar{u}$	$\bar{s}^o$
1,420	46,374	34,567	255.968	2,450	85,112	64,742	276.424
1,440	47,102	35,129	256.475	2,500	87,057	66,271	277.207
1,460	47,831	35,692	256.978	2,550	89,004	67,802	277.979
1,480	48,561	36,256	257.474	2,600	90,956	69,339	278.738
1,500	49,292	36,821	257.965	2,650	92,916	70,883	279.485
1,520	50,024	37,387	258.450	2,700	94,881	72,433	280.219
1,540	50,756	37,952	258.928	2,750	96,852	73,987	280.942
1,560	51,490	38,520	259.402	2,800	98,826	75,546	281.654
1,580	52,224	39,088	259.870	2,850	100,808	77,112	282.357
1,600	52,961	39,658	260.333	2,900	102,793	78,682	283.048
1,620	53,696	40,227	260.791	2,950	104,785	80,258	283.728
1,640	54,434	40,799	261.242	3,000	106,780	81,837	284.399
1,660	55,172	41,370	261.690	3,050	108,778	83,419	285.060
1,680	55,912	41,944	262.132	3,100	110,784	85,009	285.713
1,700	56,652	42,517	262.571	3,150	112,795	86,601	286.355
1,720	57,394	43,093	263.005	3,200	114,809	88,203	286.989
1,740	58,136	43,669	263.435	3,250	116,827	89,804	287.614
			Part c. Ideal Gas Properties of Carbon Dioxide, CO_2				
0	0	0	0	600	22,280	17,291	243.199
220	6,601	4,772	202.966	610	22,754	17,683	243.983
230	6,938	5,026	204.464	620	23,231	18,076	244.758
240	7,280	5,285	205.920	630	23,709	18,471	245.524
250	7,627	5,548	207.337	640	24,190	18,869	246.282
260	7,979	5,817	208.717	650	24,674	19,270	247.032
270	8,335	6,091	210.062	660	25,160	19,672	247.773
280	8,697	6,369	211.376	670	25,648	20,078	248.507
290	9,063	6,651	212.660	680	26,138	20,484	249.233
298	9,364	6,885	213.685	690	26,631	20,894	249.952
300	9,431	6,939	213.915	700	27,125	21,305	250.663
310	9,807	7,230	215.146	710	27,622	21,719	251.368
320	10,186	7,526	216.351	720	28,121	22,134	252.065
330	10,570	7,826	217.534	730	28,622	22,552	252.755
340	10,959	8,131	218.694	740	29,124	22,972	253.439
350	11,351	8,439	219.831	750	29,629	23,393	254.117
360	11,748	8,752	220.948	760	30,135	23,817	254.787
370	12,148	9,068	222.044	770	30,644	24,242	255.452
380	12,552	9,392	223.122	780	31,154	24,669	256.110
390	12,960	9,718	224.182	790	31,665	25,097	256.762
400	13,372	10,046	225.225	800	32,179	25,527	257.408
410	13,787	10,378	226.250	810	32,694	25,959	258.048
420	14,206	10,714	227.258	820	33,212	26,394	258.682
430	14,628	11,053	228.252	830	33,730	26,829	259.311
440	15,054	11,393	229.230	840	34,251	27,267	259.934
450	15,483	11,742	230.194	850	34,773	27,706	260.551
460	15,916	12,091	231.144	860	35,296	28,125	261.164
470	16,351	12,444	232.080	870	35,821	28,588	261.770
480	16,791	12,800	233.004	880	36,347	29,031	262.371
490	17,232	13,158	233.916	890	36,876	29,476	262.968
500	17,678	13,521	234.814	900	37,405	29,922	263.559
510	18,126	13,885	235.700	910	37,935	30,369	264.146
520	18,576	14,253	236.575	920	38,467	30,818	264.728
530	19,029	14,622	237.439	930	39,000	31,268	265.304
540	19,485	14,996	238.292	940	39,535	31,719	265.877

(*Continued*)

TABLE A.2 (*Continued*)
Ideal Gas Properties of Nitrogen, Oxygen, and Carbon Dioxide

T	$\bar{h}$	$\bar{u}$	$\bar{s}^o$	T	$\bar{h}$	$\bar{u}$	$\bar{s}^o$
550	19,945	15,372	239.135	950	40,070	32,171	266.444
560	20,407	15,751	239.962	960	40,607	32,625	267.007
570	20,870	16,131	240.789	970	41,145	33,081	267.566
580	21,337	16,515	241.602	980	41,685	33,537	268.119
590	21,807	16,902	242.405	990	42,226	33,995	268.670
1,000	42,769	34,455	269.215	1,760	86,420	71,787	301.543
1,020	43,859	35,378	270.293	1,780	87,612	72,812	302.271
1,040	44,953	36,306	271.354	1,800	88,806	73,840	302.884
1,060	46,051	37,238	272.400	1,820	90,000	74,868	303.544
1,080	47,153	38,174	273.430	1,840	91,196	75,897	304.198
1,100	48,258	39,112	274.445	1,860	92,394	76,929	304.845
1,120	49,369	40,057	275.444	1,880	93,593	77,962	305.487
1,140	50,484	41,006	276.430	1,900	94,793	78,996	306.122
1,160	51,602	41,957	277.403	1,920	95,995	80,031	306.751
1,180	52,724	42,913	278.362	1,940	97,197	81,067	307.374
1,200	53,848	43,871	279.307	1,960	98,401	82,105	307.992
1,220	54,977	44,834	280.238	1,980	99,606	83,144	308.604
1,240	56,108	45,799	281.158	2,000	100,804	84,185	309.210
1,260	57,244	46,768	282.066	2,050	103,835	86,791	310.701
1,280	58,381	47,739	282.962	2,100	106,864	89,404	312.160
1,300	59,522	48,713	283.847	2,150	109,898	92,023	313.589
1,320	60,666	49,691	284.722	2,200	112,939	94,648	314.988
1,340	61,813	50,672	285.586	2,250	115,984	97,277	316.356
1,360	62,963	51,656	286.439	2,300	119,035	99,912	317.695
1,380	64,116	52,643	287.283	2,350	122,091	102,552	319.011
1,400	65,271	53,631	288.106	2,400	125,152	105,197	320.302
1,420	66,427	54,621	288.934	2,450	128,219	107,849	321.566
1,440	67,586	55,614	289.743	2,500	131,290	110,504	322.808
1,460	68,748	56,609	290.542	2,550	134,368	113,166	324.026
1,480	69,911	57,606	291.333	2,600	137,449	115,832	325.222
1,500	71,078	58,606	292.114	2,650	140,533	118,500	326.396
1,520	72,246	59,609	292.888	2,700	143,620	121,172	327.549
1,540	73,417	60,613	292.654	2,750	146,713	123,849	328.684
1,560	74,590	61,620	294.411	2,800	149,808	126,528	329.800
1,580	76,767	62,630	295.161	2,850	152,908	129,212	330.896
1,600	76,944	63,741	295.901	2,900	156,009	131,898	331.975
1,620	78,123	64,653	296.632	2,950	159,117	134,589	333.037
1,640	79,303	65,668	297.356	3,000	162,226	137,283	334.084
1,660	80,486	66,592	298.072	3,050	165,341	139,982	335.114
1,680	81,670	67,702	298.781	3,100	168,456	142,681	336.126
1,700	82,856	68,721	299.482	3,150	171,576	145,385	337.124
1,720	84,043	69,742	300.177	3,200	174,695	148,089	338.109
1,740	85,231	70,764	300.863	3,250	177,822	150,801	339.069

Source: Moran, M.J., and Shapiro, H.N., *Fundamentals of Engineering Thermodynamics,* 3rd ed., Wiley, New York, 1995, as presented in Wark, K., *Thermodynamics,* 4th ed., McGraw-Hill, New York, 1983, based on the JANAF Thermochemical Tables, NSRDS-NBS-37, 1971.

TABLE A.3
Psychrometric Table: Properties of Moist Air at 101,325 N/m²

Symbols and Units

P_s = Pressure of water vapor at saturation, N/m²

W_s = Humidity ratio at saturation, mass of water vapor associated with unit mass of dry air

V_a = Specific volume of dry air, m³/kg

V_s = Specific volume of saturated mixture, m³/kg dry air

h_a^a = Specific enthalpy of dry air, kJ/kg

h_s = Specific enthalpy of saturated mixture, kJ/kg dry air

s_s = Specific entropy of saturated mixture, J/K·kg dry air

Temperature			Properties						
C	K	F	P_s	W_s	V_a	V_s	h_a	h_s	s_s
–40	233.15	–40	12.838	0.00007925	0.65961	0.65968	–22.35	–22.16	–90.659
–30	243.15	–22	37.992	0.0002344	0.68808	0.68833	–12.29	–11.72	–46.732
–25	248.15	–13	63.248	0.0003903	0.70232	0.70275	–7.265	–6.306	–24.706
–20	253.15	–4	103.19	0.0006371	0.71649	0.71724	–2.236	–0.6653	–2.2194
–15	258.15	+5	165.18	0.001020	0.73072	0.73191	+2.794	5.318	21.189
–10	263.15	14	259.72	0.001606	0.74495	0.74683	7.823	11.81	46.104
–5	268.15	23	401.49	0.002485	0.75912	0.76218	12.85	19.04	73.365
0	273.15	32	610.80	0.003788	0.77336	0.77804	17.88	27.35	104.14
5	278.15	41	871.93	0.005421	0.78759	0.79440	22.91	36.52	137.39
10	283.15	50	1,227.2	0.007658	0.80176	0.81163	27.94	47.23	175.54
15	288.15	59	1,704.4	0.01069	0.81600	0.82998	32.97	59.97	220.22
20	293.15	68	2,337.2	0.01475	0.83017	0.84983	38.00	75.42	273.32
25	298.15	77	3,167.0	0.02016	0.84434	0.87162	43.03	94.38	337.39
30	303.15	86	4,242.8	0.02731	0.85851	0.89609	48.07	117.8	415.65
35	308.15	95	5,623.4	0.03673	0.87274	0.92406	53.10	147.3	512.17
40	313.15	104	7,377.6	0.04911	0.88692	0.95665	58.14	184.5	532.31
45	318.15	113	9,584.8	0.06536	0.90115	0.99535	63.17	232.0	783.06
50	323.15	122	12,339	0.08678	0.91532	1.0423	68.21	293.1	975.27
55	328.15	131	15,745	0.1152	0.92949	1.1007	73.25	372.9	1,221.5
60	333.15	140	19,925	0.1534	0.94372	1.1748	78.29	478.5	1,543.5
65	338.15	149	25,014	0.2055	0.95790	1.2721	83.33	621.4	1,973.6
70	343.15	158	31,167	0.2788	0.97207	1.4042	88.38	820.5	2,564.8
75	348.15	167	38,554	0.3858	0.98630	1.5924	93.42	1,110	3,412.8
80	353.15	176	47,365	0.5519	1.0005	1.8791	98.47	1,557	4,710.9
85	358.15	185	57,809	0.8363	1.0146	2.3632	103.5	2,321	6,892.6
90	363.15	194	70,112	1.416	1.0288	3.3409	108.6	3,876	11,281

Note: The P_s column in this table gives the vapor pressure of pure water at temperature intervals of 5 °C. For the latest data on vapor pressures at intervals of 0.1°C, from 0°C to 100°C (see Wexler, A., and Greenspan, L., *J. Res. Nat. Bur. Stand.,* 75A(3):213–229, 1971).

[a] For very low barometric pressures and high wet-bulb temperatures, the values of h_a in this table are somewhat low; for corrections, see Psychrometric Tables, in "ASHRAE Handbook of Fundamentals", American Society of Heating, Refrigerating and Air-conditioning Engineers, 1972.

TABLE A.4
Water Vapor at Low Pressures: Perfect Gas Behavior $pv/T = R = 0.46151$ kJ/kg·K

Symbols and Units

t = Thermodynamic temperature, °C

T = Thermodynamic temperature, K

pv = RT, kJ/kg

u_0 = Specific internal energy at zero pressure, kJ/kg

h_0 = Specific enthalpy at zero pressure, kJ/kg

s_l = Specific entropy of semiperfect vapor at 0.1 MN/m^2, kJ/kg·K

ψ_l = Specific Helmholtz free energy of semiperfect vapor at 0.1 MN/m^2, kJ/kg

ζ_l = Specific Gibbs free energy of semiperfect vapor at 0.1 MN/m^2, kJ/kg

p_r = Relative pressure, pressure of semiperfect vapor at zero entropy, TN/m^2

v_r = Relative specific volume, specific volume of semiperfect vapor at zero entropy, mm^3/kg

c_{p0} = Specific heat capacity at constant pressure for zero pressure, kJ/kg·K

c_{v0} = Specific heat capacity at constant volume for zero pressure, kJ/kg·K

$K = c_{p0}/c_{V0}$ = isentropic exponent, $-(\partial \log p/\partial \log v)_s$

t	T	pv	u_0	h_0	s_l	ψ_l	ζ_l	p_r	v_r	c_{p0}	c_{v0}	K
0	273.15	126.06	2,375.5	2,501.5	6.8042	516.9	643.0	0.2529	498.4	1.8584	1.3969	1.3304
10	283.15	130.68	2,389.4	2,520.1	6.8711	443.9	574.6	0.2923	447.0	1.8601	1.3986	1.3300
20	293.15	135.29	2,403.4	2,538.7	6.9357	370.2	505.5	0.3363	402.4	1.8622	1.4007	1.3295
30	303.15	139.91	2,417.5	2,557.4	6.9982	296.0	435.9	0.3850	363.4	1.8647	1.4031	1.3289
40	313.15	144.52	2,431.5	2,576.0	7.0587	221.1	365.6	0.4390	329.2	1.8674	1.4059	1.3283
50	323.15	149.14	2,445.6	2,594.7	7.1175	145.6	294.7	0.4986	299.1	1.8705	1.4090	1.3275
60	333.15	153.75	2,459.7	2,613.4	7.1745	69.5	223.2	0.5642	272.5	1.8738	1.4123	1.3268
70	343.15	158.37	2,473.8	2,632.2	7.2300	−7.2	151.2	0.6363	248.9	1.8774	1.4159	1.3259
80	353.15	162.98	2,488.0	2,651.0	7.2840	−84.3	78.6	0.7152	227.9	1.8812	1.4197	1.3251
90	363.15	167.60	2,502.2	2,669.8	7.3366	−162.1	5.5	0.8015	209.1	1.8852	1.4237	1.3242
100	373.15	172.21	2,516.5	2,688.7	7.3878	−240.3	−68.1	0.8957	192.26	1.8894	1.4279	1.3232
120	393.15	181.44	2,545.1	2,726.6	7.4867	−398.3	−216.8	1.1097	163.50	1.8983	1.4367	1.3212
140	413.15	190.67	2,573.9	2,764.6	7.5811	−558.2	−367.5	1.3617	140.03	1.9077	1.4462	1.3191
160	433.15	199.90	2,603.0	2,802.9	7.6715	−720.0	−520.1	1.6564	120.69	1.9177	1.4562	1.3169

(*Continued*)

TABLE A.4 (*Continued*)
Water Vapor at Low Pressures: Perfect Gas Behavior $pv/T = R = 0.46151$ kJ/kg·K

t	T	pv	u_0	h_0	s_1	ψ_1	ζ_1	p_r	v_r	c_{p0}	c_{v0}	K
180	453.15	209.13	2,632.2	2,841.3	7.7583	–883.5	–674.4	1.9991	104.61	1.9281	1.4666	1.3147
200	473.15	218.4	2,661.6	2,880.0	7.8418	–1,048.7	–830.4	2.396	91.15	1.9389	1.4774	1.3124
300	573.15	264.5	2,812.3	3,076.8	8.2189	–1,898.4	–1,633.9	5.423	48.77	1.9975	1.5360	1.3005
400	673.15	310.7	2,969.0	3,279.7	8.5451	–2,783.1	–2,472.5	10.996	28.25	2.0614	1.5999	1.2885
500	773.15	356.8	3,132.4	3,489.2	8.8352	–3,699	–3,342	20.61	17.310	2.1287	1.6672	1.2768
600	873.15	403.0	3,302.5	3,705.5	9.0982	–4,642	–4,239	36.45	11.056	2.1980	1.7365	1.2658
700	973.15	449.1	3,479.7	3,928.8	9.3403	–5,610	–5,161	61.58	7.293	2.2683	1.8068	1.2554
800	1,073.15	495.3	3,663.9	4,159.2	9.5655	–6,601	–6,106	100.34	4.936	2.3387	1.8771	1.2459
900	1,173.15	541.4	3,855.1	4,396.5	9.7769	–7,615	–7,073	158.63	3.413	2.4078	1.9462	1.2371
1,000	1,273.15	587.6	4,053.1	4,640.6	9.9766	–8,649	–8,061	244.5	2.403	2.4744	2.0128	1.2993
1,100	1,373.15	633.7	4,257.5	4,891.2	10.1661	–9,702	–9,068	368.6	1.719	2.5369	2.0754	1.2224
1,200	1,473.15	679.9	4,467.9	5,147.8	10.3464	–10,774	–10,094	544.9	1.248	2.5938	2.1323	1.2164
1,300	1,573.15	726.0	4,683.7	5,409.7	10.5184	–11,863	–11,137	791.0	0.918	2.6431	2.1816	1.2115

Source: Keenan, J.H., Keyes, F.G., Hill, P.G., and Moore, J.G., Steam Tables, John Wiley & Sons, Inc., 1969 (International Edition–Metric Units).
For other steam tables in metric units, see "Steam Tables in SI Units", Ministry of Technology, London, 1970.

TABLE A.5
Properties of Saturated Water and Steam

		Specific Volume, m³/kg		Internal Energy, kJ/kg		Enthalpy, kJ/kg			Entropy, kJ/kg·K		
Temp., °C	Press. bars	Sat. Liquid $\upsilon_f \times 10^3$	Sat. Vapor υ_g	Sat. Liquid u_f	Sat. Vapor u_g	Sat. Liquid h_f	Evap. h_{fg}	Sat. Vapor h_g	Sat. Liquid s_f	Sat. Vapor s_g	Temp., °C
					Part a. Temperature Table						
0.01	0.00611	1.0002	206.136	0.00	2,375.3	0.01	2,501.3	2,501.4	0.0000	9.1562	0.01
4	0.00813	1.0001	157.232	16.77	2,380.9	16.78	2,491.9	2,508.7	0.0610	9.0514	4
5	0.00872	1.0001	147.120	20.97	2,382.3	20.98	2,489.6	2,510.6	0.0761	9.0257	5
6	0.00935	1.0001	137.734	25.19	2,383.6	25.20	2,487.2	2,512.4	0.0912	9.0003	6
8	0.01072	1.0002	120.917	33.59	2,386.4	33.60	2,482.5	2,516.1	0.1212	8.9501	8
10	0.01228	1.0004	106.379	42.00	2,389.2	42.01	2,477.7	2,519.8	0.1510	8.9008	10
11	0.01312	1.0004	99.857	46.20	2,390.5	46.20	2,475.4	2,521.6	0.1658	8.8765	11
12	0.01402	1.0005	93.784	50.41	2,391.9	50.41	2,473.0	2,523.4	0.1806	8.8524	12
13	0.01497	1.0007	88.124	54.60	2,393.3	54.60	2,470.7	2,525.3	0.1953	8.8285	13
14	0.01598	1.0008	82.848	58.79	2,394.7	58.80	2,468.3	2,527.1	0.2099	8.8048	14
15	0.01705	1.0009	77.926	62.99	2,396.1	62.99	2,465.9	2,528.9	0.2245	8.7814	15
16	0.01818	1.0011	73.333	67.18	2,397.4	67.19	2,463.6	2,530.8	0.2390	8.7582	16
17	0.01938	1.0012	69.044	71.38	2,398.8	71.38	2,461.2	2,532.6	0.2535	8.7351	17
18	0.02064	1.0014	65.038	75.57	2,400.2	75.58	2,458.8	2,534.4	0.2679	8.7123	18
19	0.02198	1.0016	61.293	79.76	2,401.6	79.77	2,456.5	2,536.2	0.2823	8.6897	19
20	0.02339	1.0018	57.791	83.95	2,402.9	83.96	2,454.1	2,538.1	0.2966	8.6672	20
21	0.02487	1.0020	54.514	88.14	2,404.3	88.14	2,451.8	2,539.9	0.3109	8.6450	21
22	0.02645	1.0022	51.447	92.32	2,405.7	92.33	2,449.4	2,541.7	0.3251	8.6229	22
23	0.02810	1.0024	48.574	96.51	2,407.0	96.52	2,447.0	2,543.5	0.3393	8.6011	23
24	0.02985	1.0027	45.883	100.70	2,408.4	100.70	2,444.7	2,545.4	0.3534	8.5794	24
25	0.03169	1.0029	43.360	104.88	2,409.8	104.89	2,442.3	2,547.2	0.3674	8.5580	25
26	0.03363	1.0032	40.994	109.06	2,411.1	109.07	2,439.9	2,549.0	0.3814	8.5367	26
27	0.03567	1.0035	38.774	113.25	2,412.5	113.25	2,437.6	2,550.8	0.3954	8.5156	27
28	0.03782	1.0037	36.690	117.42	2,413.9	117.43	2,435.2	2,552.6	0.4093	8.4946	28
29	0.04008	1.0040	34.733	121.60	2,415.2	121.61	2,432.8	2,554.5	0.4231	8.4739	29
30	0.04246	1.0043	32.894	125.78	2,416.6	125.79	2,430.5	2,556.3	0.4369	8.4533	30

(*Continued*)

TABLE A.5 (*Continued*)
Properties of Saturated Water and Steam

		Specific Volume, m³/kg		Internal Energy, kJ/kg		Enthalpy, kJ/kg			Entropy, kJ/kg·K		
Temp., °C	Press. bars	Sat. Liquid $\upsilon_f \times 10^3$	Sat. Vapor υ_g	Sat. Liquid u_f	Sat. Vapor u_g	Sat. Liquid h_f	Evap. h_{fg}	Sat. Vapor h_g	Sat. Liquid s_f	Sat. Vapor s_g	Temp., °C
31	0.04496	1.0046	31.165	129.96	2,418.0	129.97	2,428.1	2,558.1	0.4507	8.4329	31
32	0.04759	1.0050	29.540	134.14	2,419.3	134.15	2,425.7	2,559.9	0.4644	8.4127	32
33	0.05034	1.0053	28.011	138.32	2,420.7	138.33	2,423.4	2,561.7	0.4781	8.3927	33
34	0.05324	1.0056	26.571	142.50	2,422.0	142.50	2,421.0	2,563.5	0.4917	8.3728	34
35	0.05628	1.0060	25.216	146.67	2,423.4	146.68	2,418.6	2,565.3	0.5053	8.3531	35
36	0.05947	1.0063	23.940	150.85	2,424.7	150.86	2,416.2	2,567.1	0.5188	8.3336	36
38	0.06632	1.0071	21.602	159.20	2,427.4	159.21	2,411.5	2,570.7	0.5458	8.2950	38
40	0.07384	1.0078	19.523	167.56	2,430.1	167.57	2,406.7	2,574.3	0.5725	8.2570	40
45	0.09593	1.0099	15.258	188.44	2,436.8	188.45	2,394.8	2,583.2	0.6387	8.1648	45
50	0.1235	1.0121	12.032	209.32	2,443.5	209.33	2,382.7	2,592.1	0.7038	8.0763	50
55	0.1576	1.0146	9.568	230.21	2,450.1	230.23	2,370.7	2,600.9	0.7679	7.9913	55
60	0.1994	1.0172	7.671	251.11	2,456.6	251.13	2,358.5	2,609.6	0.8312	7.9096	60
65	0.2503	1.0199	6.197	272.02	2,463.1	272.06	2,346.2	2,618.3	0.8935	7.8310	65
70	0.3119	1.0228	5.042	292.95	2,469.6	292.98	2,333.8	2,626.8	0.9549	7.7553	70
75	0.3858	1.0259	4.131	313.90	2,475.9	313.93	2,321.4	2,635.3	1.0155	7.6824	75
80	0.4739	1.0291	3.407	334.86	2,482.2	334.91	2,308.8	2,643.7	1.0753	7.6122	80
85	0.5783	1.0325	2.828	355.84	2,488.4	355.90	2,296.0	2,651.9	1.1343	7.5445	85
90	0.7014	1.0360	2.361	376.85	2,494.5	376.92	2,283.2	2,660.1	1.1925	7.4791	90
95	0.8455	1.0397	1.982	397.88	2,500.6	397.96	2,270.2	2,668.1	1.2500	7.4159	95
100	1.014	1.0435	1.673	418.94	2,506.5	419.04	2,257.0	2,676.1	1.3069	7.3549	100
110	1.433	1.0516	1.210	461.14	2,518.1	461.30	2,230.2	2,691.5	1.4185	7.2387	110
120	1.985	1.0603	0.8919	503.50	2,529.3	503.71	2,202.6	2,706.3	1.5276	7.1296	120
130	2.701	1.0697	0.6685	546.02	2,539.9	546.31	2,174.2	2,720.5	1.6344	7.0269	130
140	3.613	1.0797	0.5089	588.74	2,550.0	589.13	2,144.7	2,733.9	1.7391	6.9299	140
150	4.758	1.0905	0.3928	631.68	2,559.5	632.20	2,114.3	2,746.5	1.8418	6.8379	150
160	6.178	1.1020	0.3071	674.86	2,568.4	675.55	2,082.6	2,758.1	1.9427	6.7502	160
170	7.917	1.1143	0.2428	718.33	2,576.5	719.21	2,049.5	2,768.7	2.0419	6.6663	170

(*Continued*)

TABLE A.5 (*Continued*)
Properties of Saturated Water and Steam

		Specific Volume, m³/kg		Internal Energy, kJ/kg		Enthalpy, kJ/kg			Entropy, kJ/kg·K		
Temp., °C	Press. bars	Sat. Liquid $\upsilon_f \times 10^3$	Sat. Vapor υ_g	Sat. Liquid u_f	Sat. Vapor u_g	Sat. Liquid h_f	Evap. h_{fg}	Sat. Vapor h_g	Sat. Liquid s_f	Sat. Vapor s_g	Temp., °C
180	10.02	1.1274	0.1941	762.09	2,583.7	763.22	2,015.0	2,778.2	2.1396	6.5857	180
190	12.54	1.1414	0.1565	806.19	2,590.0	807.62	1,978.8	2,786.4	2.2359	6.5079	190
200	15.54	1.1565	0.1274	850.65	2,595.3	852.45	1,940.7	2,793.2	2.3309	6.4323	200
210	19.06	1.1726	0.1044	895.53	2,599.5	897.76	1,900.7	2,798.5	2.4248	6.3585	210
220	23.18	1.1900	0.08619	940.87	2,602.4	943.62	1,858.5	2,802.1	2.5178	6.2861	220
230	27.95	1.2088	0.07158	986.74	2,603.9	990.12	1,813.8	2,804.0	2.6099	6.2146	230
240	33.44	1.2291	0.05976	1,033.2	2,604.0	1,037.3	1,766.5	2,803.8	2.7015	6.1437	240
250	39.73	1.2512	0.05013	1,080.4	2,602.4	1,085.4	1,716.2	2,801.5	2.7927	6.0730	250
260	46.88	1.2755	0.04221	1,128.4	2,599.0	1,134.4	1,662.5	2,796.6	2.8838	6.0019	260
270	54.99	1.3023	0.03564	1,177.4	2,593.7	1,184.5	1,605.2	2,789.7	2.9751	5.9301	270
280	64.12	1.3321	0.03017	1,227.5	2,586.1	1,236.0	1,543.6	2,779.6	3.0668	5.8571	280
290	74.36	1.3656	0.02557	1,278.9	2,576.0	1,289.1	1,477.1	2,766.2	3.1594	5.7821	290
300	85.81	1.4036	0.02167	1,332.0	2,563.0	1,344.0	1,404.9	2,749.0	3.2534	5.7045	300
320	112.7	1.4988	0.01549	1,444.6	2,525.5	1,461.5	1,238.6	2,700.1	3.4480	5.5362	320
340	145.9	1.6379	0.01080	1,570.3	2,464.6	1,594.2	1,027.9	2,622.0	3.6594	5.3357	340
360	186.5	1.8925	0.006945	1,725.2	2,351.5	1,760.5	720.5	2,481.0	3.9147	5.0526	360
374.14	220.9	3.155	0.003155	2,029.6	2,029.6	2,099.3	0	2,099.3	4.4298	4.4298	374.14

(*Continued*)

TABLE A.5 (*Continued*)
Properties of Saturated Water and Steam

		Specific Volume, m^3/kg		Internal Energy, kJ/kg		Enthalpy, kJ/kg			Entropy, kJ/kg·K		
Press. bars	Temp. °C	Sat. Liquid $\upsilon_t \times 10^3$	Sat. Vapor υ_g	Sat. Liquid u_t	Sat. Vapor u_g	Sat. Liquid h_t	Evap. h_{tg}	Sat. Vapor h_g	Sat. Liquid s_t	Sat. Vapor s_g	Press. bars
					Part b. Pressure Table						
0.04	28.96	1.0040	34.800	121.45	2,415.2	121.46	2,432.9	2,554.4	0.4226	8.4746	0.04
0.06	36.16	1.0064	23.739	151.53	2,425.0	151.53	2,415.9	2,567.4	0.5210	8.3304	0.06
0.08	41.51	1.0084	18.103	173.87	2,432.2	173.88	2,403.1	2,577.0	0.5926	8.2287	0.08
0.10	45.81	1.0102	14.674	191.82	2,437.9	191.83	2,392.8	2,584.7	0.6493	8.1502	0.10
0.20	60.06	1.0172	7.649	251.38	2,456.7	251.40	2,358.3	2,609.7	0.8320	7.9085	0.20
0.30	69.10	1.0223	5.229	289.20	2,468.4	289.23	2,336.1	2,625.3	0.9439	7.7686	0.30
0.40	75.87	1.0265	3.993	317.53	2,477.0	317.58	2,319.2	2,636.8	1.0259	7.6700	0.40
0.50	81.33	1.0300	3.240	340.44	2,483.9	340.49	2,305.4	2,645.9	1.0910	7.5939	0.50
0.60	85.94	1.0331	2.732	359.79	2,489.6	359.86	2,293.6	2,653.5	1.1453	7.5320	0.60
0.70	89.95	1.0360	2.365	376.63	2,494.5	376.70	2,283.3	2,660.0	1.1919	7.4797	0.70
0.80	93.50	1.0380	2.087	391.58	2,498.8	391.66	2,274.1	2,665.8	1.2329	7.4346	0.80
0.90	96.71	1.0410	1.869	405.06	2,502.6	405.15	2,265.7	2,670.9	1.2695	7.3949	0.90
1.00	99.63	1.0432	1.694	417.36	2,506.1	417.46	2,258.0	2,675.5	1.3026	7.3594	1.00
1.50	111.4	1.0528	1.159	466.94	2,519.7	467.11	2,226.5	2,693.6	1.4336	7.2233	1.50
2.00	120.2	1.0605	0.8857	504.49	2,529.5	504.70	2,201.9	2,706.7	1.5301	7.1271	2.00
2.50	127.4	1.0672	0.7187	535.10	2,537.2	535.37	2,181.5	2,716.9	1.6072	7.0527	2.50
3.00	133.6	1.0732	0.6058	561.15	2,543.6	561.47	2,163.8	2,725.3	1.6718	6.9919	3.00
3.50	138.9	1.0786	0.5243	583.95	2,546.9	584.33	2,148.1	2,732.4	1.7275	6.9405	3.50
4.00	143.6	1.0836	0.4625	604.31	2,553.6	604.74	2,133.8	2,738.6	1.7766	6.8959	4.00
4.50	147.9	1.0882	0.4140	622.25	2,557.6	623.25	2,120.7	2,743.9	1.8207	6.8565	4.50
5.00	151.9	1.0926	0.3749	639.68	2,561.2	640.23	2,108.5	2,748.7	1.8607	6.8212	5.00
6.00	158.9	1.1006	0.3157	669.90	2,567.4	670.56	2,086.3	2,756.8	1.9312	6.7600	6.00
7.00	165.0	1.1080	0.2729	696.44	2,572.5	697.22	2,066.3	2,763.5	1.9922	6.7080	7.00
8.00	170.4	1.1148	0.2404	720.22	2,576.8	721.11	2,048.0	2,769.1	2.0462	6.6628	8.00
9.00	175.4	1.1212	0.2150	741.83	2,580.5	742.83	2,031.1	2,773.9	2.0946	6.6226	9.00
10.0	179.9	1.1273	0.1944	761.68	2,583.6	762.81	2,015.3	2,778.1	2.1387	6.5863	10.0

(*Continued*)

TABLE A.5 (*Continued*)
Properties of Saturated Water and Steam

		Specific Volume, m^3/kg		Internal Energy, kJ/kg		Enthalpy, kJ/kg			Entropy, kJ/kg·K		
Press. bars	Temp. °C	Sat. Liquid $\upsilon_t \times 10^3$	Sat. Vapor υ_g	Sat. Liquid u_t	Sat. Vapor u_g	Sat. Liquid h_t	Evap. h_{tg}	Sat. Vapor h_g	Sat. Liquid s_t	Sat. Vapor s_g	Press. bars
15.0	198.3	1.1539	0.1318	843.16	2,594.5	844.84	1,947.3	2,792.2	2.3150	6.4448	15.0
20.0	212.4	1.1767	0.09963	906.44	2,600.3	908.79	1,890.7	2,799.5	2.4474	6.3409	20.0
25.0	224.0	1.1973	0.07998	959.11	2,603.1	962.11	1,841.0	2,803.1	2.5547	6.2575	25.0
30.0	233.9	1.2165	0.06668	1,004.8	2,604.1	1,008.4	1,795.7	2,804.2	2.6457	6.1869	30.0
35.0	242.6	1.2347	0.05707	1,045.4	2,603.7	1,049.8	1,753.7	2,803.4	2.7253	6.1253	35.0
40.0	250.4	1.2522	0.04978	1,082.3	2,602.3	1,087.3	1,714.1	2,801.4	2.7964	6.0701	40.0
45.0	257.5	1.2692	0.04406	1,116.2	2,600.1	1,121.9	1,676.4	2,798.3	2.8610	6.0199	45.0
50.0	264.0	1.2859	0.03944	1,147.8	2,597.1	1,154.2	1,640.1	2,794.3	2.9202	5.9734	50.0
60.0	275.6	1.3187	0.03244	1,205.4	2,589.7	1,213.4	1,571.0	2,784.3	3.0267	5.8892	60.0
70.0	285.9	1.3513	0.02737	1,257.6	2,580.5	1,267.0	1,505.1	2,772.1	3.1211	5.8133	70.0
80.0	295.1	1.3842	0.02352	1,305.6	2,569.8	1,316.6	1,441.3	2,758.0	3.2068	5.7432	80.0
90.0	303.4	1.4178	0.02048	1,350.5	2,557.8	1,363.3	1,378.9	2,742.1	3.2858	5.6772	90.0
100.	311.1	1.4524	0.01803	1,393.0	2,544.4	1,407.6	1,317.1	2,724.7	3.3596	5.6141	100
110.	318.2	1.4886	0.01599	1,433.7	2,529.8	1,450.1	1,255.5	2,705.6	3.4295	5.5527	110
120.	324.8	1.5267	0.01426	1,473.0	2,513.7	1,491.3	1,193.6	2,684.9	3.4962	5.4924	120
130.	330.9	1.5671	0.01278	1,511.1	2,496.1	1,531.5	1,130.7	2,662.2	3.5606	5.4323	130
140.	336.8	1.6107	0.01149	1,548.6	2,476.8	1,571.1	1,066.5	2,637.6	3.6232	5.3717	140
150.	342.2	1.6581	0.01034	1,585.6	2,455.5	1,610.5	1,000.0	2,610.5	3.6848	5.3098	150
160.	347.4	1.7107	0.009306	1,622.7	2,431.7	1,650.1	930.6	2,580.6	3.7461	5.2455	160
170.	352.4	1.7702	0.008364	1,660.2	2,405.0	1,690.3	856.9	2,547.2	3.8079	5.1777	170
180.	357.1	1.8397	0.007489	1,698.9	2,374.3	1,732.0	777.1	2,509.1	3.8715	5.1044	180
190.	361.5	1.9243	0.006657	1,739.9	2,338.1	1,776.5	688.0	2,464.5	3.9388	5.0228	190
200.	365.8	2.036	0.005834	1,785.6	2,293.0	1,826.3	583.4	2,409.7	4.0139	4.9269	200
220.9	374.1	3.155	0.003155	2,029.6	2,029.6	2,099.3	0	2,099.3	4.4298	4.4298	220.9

Source: Moran, M.J., and Shapiro, H.N., *Fundamentals of Engineering Thermodynamics,* 3rd ed., Wiley, New York, 1995, as extracted from Keenan, J.H., Keyes, F.G., Hill, P.G., and Moore, J.G., *Steam Tables*, Wiley, New York, 1969.

TABLE A.6
Properties of Superheated Steam

Symbols and Units

T = Temperature, °C

T_{sat} = Saturation temperature, °C

ν = Specific volume, m^3/kg

u = Internal energy, kJ/kg

h = Enthalpy, kJ/kg

S = Entropy, kJ/kg·K

p = Pressure, bar and μPa

T(°C)	υ, m^3/kg	u, kJ/kg	h, kJ/kg	s, kJ/kg·K	υ, m^3/kg	u, kJ/kg	h, kJ/kg	s, kJ/kg·K
	p = 0.06 bar = 0.006 MPa (T_{sat} = 36.16°C)				p = 0.35 bar = 0.035 MPa (T_{sat} = 72.69°C)			
Sat.	23.739	2,425.0	2,567.4	8.3304	4.526	2,473.0	2,631.4	7.7158
80	27.132	2,487.3	2,650.1	8.5804	4.625	2,483.7	2,645.6	7.7564
120	30.219	2,544.7	2,726.0	8.7840	5.163	2,542.4	2,723.1	7.9644
160	33.302	2,602.7	2,802.5	8.9693	5.696	2,601.2	2,800.6	8.1519
200	36.383	2,661.4	2,879.7	9.1398	6.228	2,660.4	2,878.4	8.3237
240	39.462	2,721.0	2,957.8	9.2982	6.758	2,720.3	2,956.8	8.4828
280	42.540	2,781.5	3,036.8	9.4464	7.287	2,780.9	3,036.0	8.6314
320	45.618	2,843.0	3,116.7	9.5859	7.815	2,842.5	3,116.1	8.7712
360	48.696	2,905.5	3,197.7	9.7180	8.344	2,905.1	3,197.1	8.9034
400	51.774	2,969.0	3,279.6	9.8435	8.872	2,968.6	3,279.2	9.0291
440	54.851	3,033.5	3,362.6	9.9633	9.400	3,033.2	3,362.2	9.1490
500	59.467	3,132.3	3,489.1	10.1336	10.192	3,132.1	3,488.8	9.3194
	p = 0.70 bar = 0.07 MPa (T_{sat} = 89.95°C)				p = 1.0 bar = 0.10 MPa (T_{sat} = 99.63°C)			
Sat.	2.365	2,494.5	2,660.0	7.4797	1.694	2,506.1	2,675.5	7.3594
100	2.434	2,509.7	2,680.0	7.5341	1.696	2,506.7	2,676.2	7.3614
120	2.571	2,539.7	2,719.6	7.6375	1.793	2,537.3	2,716.6	7.4668
160	2.841	2,599.4	2,798.2	7.8279	1.984	2,597.8	2,796.2	7.6597
200	3.108	2,659.1	2,876.7	8.0012	2.172	2,658.1	2,875.3	7.8343
240	3.374	2,719.3	2,955.5	8.1611	2.359	2,718.5	2,954.5	7.9949
280	3.640	2,780.2	3,035.0	8.3162	2.546	2,779.6	3,034.2	8.1445
320	3.905	2,842.0	3,115.3	8.4504	2.732	2,841.5	3,114.6	8.2849
360	4.170	2,904.6	3,196.5	8.5828	2.917	2,904.2	3,195.9	8.4175
400	4.434	2,968.2	3,278.6	8.7086	3.103	2,967.9	3,278.2	8.5435
440	4.698	3,032.9	3,361.8	8.8286	3.288	3,032.6	3,361.4	8.6636
500	5.095	3,131.8	3,488.5	8.9991	3.565	3,131.6	3,488.1	8.8342
	p = 1.5 bars = 0.15 MPa (T_{sat} = 111.37°C)				p = 3.0 bars = 0.30 MPa (T_{sat} = 133.55°C)			
Sat.	1.159	2,519.7	2,693.6	7.2233	0.606	2,543.6	2,725.3	6.9919
120	1.188	2,533.3	2,711.4	7.2693				
160	1.317	2,595.2	2,792.8	7.4665	0.651	2,587.1	2,782.3	7.1276
200	1.444	2,656.2	2,872.9	7.6433	0.716	2,650.7	2,865.5	7.3115
240	1.570	2,717.2	2,952.7	7.8052	0.781	2,713.1	2,947.3	7.4774
280	1.695	2,778.6	3,032.8	7.9555	0.844	2,775.4	3,028.6	7.6299
320	1.819	2,840.6	3,113.5	8.0964	0.907	2,838.1	3,110.1	7.7722
360	1.943	2,903.5	3,195.0	8.2293	0.969	2,901.4	3,192.2	7.9061
400	2.067	2,967.3	3,277.4	8.3555	1.032	2,965.6	3,275.0	8.0330
440	2.191	3,032.1	3,360.7	8.4757	1.094	3,030.6	3,358.7	8.1538

(*Continued*)

TABLE A.6 (*Continued*)
Properties of Superheated Steam

T(°C)	υ, m³/kg	u, kJ/kg	h, kJ/kg	s, kJ/kg·K	υ, m³/kg	u, kJ/kg	h, kJ/kg	s, kJ/kg·K
500	2.376	3,131.2	3,487.6	8.6466	1.187	3,130.0	3,486.0	8.3251
600	2.685	3,301.7	3,704.3	8.9101	1.341	3,300.8	3,703.2	8.5892
	p=5.0 bars=0.50 MPa (T_{sat}=151.86°C)				p=7.0 bars=0.70 MPa (T_{sat}=164.97°C)			
Sat.	0.3749	2,561.2	2,748.7	6.8213	0.2729	2,572.5	2,763.5	6.7080
180	0.4045	2,609.7	2,812.0	6.9656	0.2847	2,599.8	2,799.1	6.7880
200	0.4249	2,642.9	2,855.4	7.0592	0.2999	2,634.8	2,844.8	6.8865
240	0.4646	2,707.6	2,939.9	7.2307	0.3292	2,701.8	2,932.2	7.0641
280	0.5034	2,771.2	3,022.9	7.3865	0.3574	2,766.9	3,017.1	7.2233
320	0.5416	2,834.7	3,105.6	7.5308	0.3852	2,831.3	3,100.9	7.3697
360	0.5796	2,898.7	3,188.4	7.6660	0.4126	2,895.8	3,184.7	7.5063
400	0.6173	2,963.2	3,271.9	7.7938	0.4397	2,960.9	3,268.7	7.6350
440	0.6548	3,028.6	3,356.0	7.9152	0.4667	3,026.6	3,353.3	7.7571
500	0.7109	3,128.4	3,483.9	8.0873	0.5070	3,126.8	3,481.7	7.9299
600	0.8041	3,299.6	3,701.7	8.3522	0.5738	3,298.5	3,700.2	8.1956
700	0.8969	3,477.5	3,925.9	8.5952	0.6403	3,476.6	3,924.8	8.4391
	p=10.0 bars=1.0 MPa (T_{sat}=179.91°C)				p=15.0 bars=1.5 MPa (T_{sat}=198.32°C)			
Sat.	0.1944	2,583.6	2,778.1	6.5865	0.1318	2,594.5	2,792.2	6.4448
200	0.2060	2,621.9	2,827.9	6.6940	0.1325	2,598.1	2,796.8	6.4546
240	0.2275	2,692.9	2,920.4	6.8817	0.1483	2,676.9	2,899.3	6.6628
280	0.2480	2,760.2	3,008.2	7.0465	0.1627	2,748.6	2,992.7	6.8381
320	0.2678	2,826.1	3,093.9	7.1962	0.1765	2,817.1	3,081.9	6.9938
360	0.2873	2,891.6	3,178.9	7.3349	0.1899	2,884.4	3,169.2	7.1363
400	0.3066	2,957.3	3,263.9	7.4651	0.2030	2,951.3	3,255.8	7.2690
440	0.3257	3,023.6	3,349.3	7.5883	0.2160	3,018.5	3,342.5	7.3940
500	0.3541	3,124.4	3,478.5	7.7622	0.2352	3,120.3	3,473.1	7.5698
540	0.3729	3,192.6	3,565.6	7.8720	0.2478	3,189.1	3,560.9	7.6805
600	0.4011	3,296.8	3,697.9	8.0290	0.2668	3,293.9	3,694.0	7.8385
640	0.4198	3,367.4	3,787.2	8.1290	0.2793	3,364.8	3,783.8	7.9391
	p=20.0 bars=2.0 MPa (T_{sat}=212.42°C)				p=30.0 bars=3.0 MPa (T_{sat}=233.90°C)			
Sat.	0.0996	2,600.3	2,799.5	6.3409	0.0667	2,604.1	2,804.2	6.1869
240	0.1085	2,659.6	2,876.5	6.4952	0.0682	2,619.7	2,824.3	6.2265
280	0.1200	2,736.4	2,976.4	6.6828	0.0771	2,709.9	2,941.3	6.4462
320	0.1308	2,807.9	3,069.5	6.8452	0.0850	2,788.4	3,043.4	6.6245
360	0.1411	2,877.0	3,159.3	6.9917	0.0923	2,861.7	3,138.7	6.7801
400	0.1512	2,945.2	3,247.6	7.1271	0.0994	2,932.8	3,230.9	6.9212
440	0.1611	3,013.4	3,335.5	7.2540	0.1062	3,002.9	3,321.5	7.0520
500	0.1757	3,116.2	3,467.6	7.4317	0.1162	3,108.0	3,456.5	7.2338
540	0.1853	3,185.6	3,556.1	7.5434	0.1227	3,178.4	3,546.6	7.3474
600	0.1996	3,290.9	3,690.1	7.7024	0.1324	3,285.0	3,682.3	7.5085
640	0.2091	3,362.2	3,780.4	7.8035	0.1388	3,357.0	3,773.5	7.6106
700	0.2232	3,470.9	3,917.4	7.9487	0.1484	3,466.5	3,911.7	7.7571
	p=40 bars=4.0 MPa (T_{sat}=250.4°C)				p=60 bars=6.0 MPa (T_{sat}=275.64°C)			
Sat.	0.04978	2,602.3	2,801.4	6.0701	0.03244	2,589.7	2,784.3	5.8892
280	0.05546	2,680.0	2,901.8	6.2568	0.03317	2,605.2	2,804.2	5.9252
320	0.06199	2,767.4	3,015.4	6.4553	0.03876	2,720.0	2,952.6	6.1846
360	0.06788	2,845.7	3,117.2	6.6215	0.04331	2,811.2	3,071.1	6.3782

(*Continued*)

TABLE A.6 (*Continued*)
Properties of Superheated Steam

T(°C)	υ, m³/kg	u, kJ/kg	h, kJ/kg	s, kJ/kg·K	υ, m³/kg	u, kJ/kg	h, kJ/kg	s, kJ/kg·K
400	0.07341	2,919.9	3,213.6	6.7690	0.04739	2,892.9	3,177.2	6.5408
440	0.07872	2,992.2	3,307.1	6.9041	0.05122	2,970.0	3,277.3	6.6853
500	0.08643	3,099.5	3,445.3	7.0901	0.05665	3,082.2	3,422.2	6.8803
540	0.09145	3,171.1	3,536.9	7.2056	0.06015	3,156.1	3,517.0	6.9999
600	0.09885	3,279.1	3,674.4	7.3688	0.06525	3,266.9	3,658.4	7.1677
640	0.1037	3,351.8	3,766.6	7.4720	0.06859	3,341.0	3,752.6	7.2731
700	0.1110	3,462.1	3,905.9	7.6198	0.07352	3,453.1	3,894.1	7.4234
740	0.1157	3,536.6	3,999.6	7.7141	0.07677	3,528.3	3,989.2	7.5190
	p=80 bars=8.0 MPa (T_{sat}=295.06°C)				**p=100 bars=10.0 MPa (T_{sat}=311.06°C)**			
Sat.	0.02352	2,569.8	2,758.0	5.7432	0.01803	2,544.4	2,724.7	5.6141
320	0.02682	2,662.7	2,877.2	5.9489	0.01925	2,588.8	2,781.3	5.7103
360	0.03089	2,772.7	3,019.8	6.1819	0.02331	2,729.1	2,962.1	6.0060
400	0.03432	2,863.8	3,138.3	6.3634	0.02641	2,832.4	3,096.5	6.2120
440	0.03742	2,946.7	3,246.1	6.5190	0.02911	2,922.1	3,213.2	6.3805
480	0.04034	3,025.7	3,348.4	6.6586	0.03160	3,005.4	3,321.4	6.5282
520	0.04313	3,102.7	3,447.7	6.7871	0.03394	3,085.6	3,425.1	6.6622
560	0.04582	3,178.7	3,545.3	6.9072	0.03619	3,164.1	3,526.0	6.7864
600	0.04845	3,254.4	3,642.0	7.0206	0.03837	3,241.7	3,625.3	6.9029
640	0.05102	3,330.1	3,738.3	7.1283	0.04048	3,318.9	3,723.7	7.0131
700	0.05481	3,443.9	3,882.4	7.2812	0.04358	3,434.7	3,870.5	7.1687
740	0.05729	3,520.4	3,978.7	7.3782	0.04560	3,512.1	3,968.1	7.2670
	p=120 bars=12.0 MPa (T_{sat}=324.75°C)				**p=140 bars=14.0 MPa (T_{sat}=336.75°C)**			
Sat.	0.01426	2,513.7	2,684.9	5.4924	0.01149	2,476.8	2,637.6	5.3717
360	0.01811	2,678.4	2,895.7	5.8361	0.01422	2,617.4	2,816.5	5.6602
400	0.02108	2,798.3	3,051.3	6.0747	0.01722	2,760.9	3,001.9	5.9448
440	0.02355	2,896.1	3,178.7	6.2586	0.01954	2,868.6	3,142.2	6.1474
480	0.02576	2,984.4	3,293.5	6.4154	0.02157	2,962.5	3,264.5	6.3143
520	0.02781	3,068.0	3,401.8	6.5555	0.02343	3,049.8	3,377.8	6.4610
560	0.02977	3,149.0	3,506.2	6.6840	0.02517	3,133.6	3,486.0	6.5941
600	0.03164	3,228.7	3,608.3	6.8037	0.02683	3,215.4	3,591.1	6.7172
640	0.03345	3,307.5	3,709.0	6.9164	0.02843	3,296.0	3,694.1	6.8326
700	0.03610	3,425.2	3,858.4	7.0749	0.03075	3,415.7	3,846.2	6.9939
740	0.03781	3,503.7	3,957.4	7.1746	0.03225	3,495.2	3,946.7	7.0952

TABLE A.7
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	Acetylene (Ethyne)	Air (mixture)	Ammonia, anhyd.	Argon
Chemical Formula	**C_2H_2**		**NH_3**	**Ar**
Refrigerant Number	**—**	**729**	**717**	**740**
	Chemical and Physical Properties			
Molecular weight	26.04	28.966	17.02	39.948
Specific gravity, air = 1	0.90	1.00	0.59	1.38
Specific volume, ft^3/lb	14.9	13.5	23.0	9.80
Specific volume, m^3/kg	0.93	0.842	1.43	0.622
Density of liquid (at atm bp), lb/ft^3	43.0	54.6	42.6	87.0
Density of liquid (at atm bp), kg/m^3	693.	879.	686.	1,400
Vapor pressure at 25°C, psia			145.4	
Vapor pressure at 25°C, MN/m^2			1.00	
Viscosity (abs), lbm/ft·s	6.72×10^{-6}	12.1×10^{-6}	6.72×10^{-6}	13.4×10^{-6}
Viscosity (abs), centipoises[a]	0.01	0.018	0.010	0.02
Sound velocity in gas, m/s	343	346	415	322
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.40	0.2403	0.52	0.125
Specific heat, c_p J/kg·K	1,674	1,005	2,175	523
Specific heat ratio, c_p/c_v	1.25	1.40	1.3	1.67
Gas constant *R*, ft-lb/lb·°R	59.3	53.3	90.8	38.7
Gas constant *R*, J/kg·°C	319	286.8	488	208
Thermal conductivity, Btu/h-ft·°F	0.014	0.0151	0.015	0.0102
Thermal conductivity, W/m·°C	0.024	0.026	0.026	0.0172
Boiling point (sat 14.7 psia), °F	−103	−320	−28	−303
Boiling point (sat 760 mm), °C	−75	−195	−33.3	−186
Latent heat of evap (at bp), Btu/lb	264	88.2	589.3	70.
Latent heat of evap (at bp), J/kg	614,000	205,000	1,373,000	163,000
Freezing (melting) point, °F (1 atm)	−116	−357.2	−107.9	−308.5
Freezing (melting) point, °C (1 atm)	−82.2	−216.2	−77.7	−189.2
Latent heat of fusion, Btu/lb	23.	10.0	143.0	
Latent heat of fusion, J/kg	53,500	23,200	332,300	
Critical temperature, °F	97.1	−220.5	271.4	−187.6
Critical temperature, °C	36.2	−140.3	132.5	−122
Critical pressure, psia	907	550	1,650	707
Critical pressure, MN/m^2	6.25	3.8	11.4	4.87
Critical volume, ft^3/lb		0.050	0.068	0.0299
Critical volume, m^3/kg		0.003	0.00424	0.00186
Flammable (yes or no)	Yes	No	No	No
Heat of combustion, Btu/ft^3	1,450	—	—	—
Heat of combustion, Btu/lb	21,600	—	—	—
Heat of combustion, kJ/kg	50,200	—	—	—

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	Butadiene	*n*-Butane	Isobutane (2-methyl Propane)	1-Butene (Butylene)
Chemical Formula	C_4H_6	C_4H_{10}	C_4H_{10}	C_4H_8
Refrigerant Number	—	600	600a	—
	Chemical and Physical Properties			
Molecular weight	54.09	58.12	58.12	56.108
Specific gravity, air = 1	1.87	2.07	2.07	1.94
Specific volume, ft^3/lb	7.1	6.5	6.5	6.7
Specific volume, m^3/kg	0.44	0.405	0.418	0.42
Density of liquid (at atm bp), lb/ft^3		37.5	37.2	
Density of liquid (at atm bp), kg/m^3		604	599	
Vapor pressure at 25°C, psia		35.4	50.4	
Vapor pressure at 25°C, MN/m^2		0.0244	0.347	
Viscosity (abs), lbm/ft·s		4.8×10^{-6}		
Viscosity (abs), centipoises[a]		0.007		
Sound velocity in gas, m/s	226	216	216	222
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.341	0.39	0.39	0.36
Specific heat, c_p, J/kg·K	1,427	1,675	1,630	1,505
Specific heat ratio, c_p/c_v	1.12	1.096	1.10	1.112
Gas constant *R*, ft-lb/lb·°F	28.55	26.56	26.56	27.52
Gas constant *R*, J/kg·°C	154	143	143	148
Thermal conductivity, Btu/h ft·°F		0.01	0.01	
Thermal conductivity, W/m·°C		0.017	0.017	
Boiling point (sat 14.7 psia), °F	24.1	31.2	10.8	20.6
Boiling point (sat 760 mm), °C	–4.5	–0.4	–11.8	–6.3
Latent heat of evap (at bp), Btu/lb		165.6	157.5	167.9
Latent heat of evap (at bp), J/kg		386,000	366,000	391,000
Freezing (melting) point, °F (1 atm)	–164	–217	–229	–301.6
Freezing (melting) point, °C (1 atm)	–109	–138	–145	–185.3
Latent heat of fusion, Btu/lb		19.2		16.4
Latent heat of fusion, J/kg		44,700		38,100
Critical temperature, °F		306	273	291
Critical temperature, °C	171	152	134	144
Critical pressure, psia	652	550	537	621
Critical pressure, MN/m^2		3.8	3.7	4.28
Critical volume, ft^3/lb		0.070		0.068
Critical volume, m^3/kg		0.0043		0.0042
Flammable (yes or no)	Yes	Yes	Yes	Yes
Heat of combustion, Btu/ft^3	2,950	3,300	3,300	3,150
Heat of combustion, Btu/lb	20,900	21,400	21,400	21,000
Heat of combustion, kJ/kg	48,600	49,700	49,700	48,800

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	*cis*-2-Butene	*trans*-2-Butene	Isobutene	Carbon Dioxide
Chemical Formula	C_4H_8	C_4H_8	C_4H_8	CO_2
Refrigerant Number	—	—	—	744
	Chemical and Physical Properties			
Molecular weight	56.108	56.108	56.108	44.01
Specific gravity, air = 1	1.94	1.94	1.94	1.52
Specific volume, ft^3/lb	6.7	6.7	6.7	8.8
Specific volume, m^3/kg	0.42	0.42	0.42	0.55
Density of liquid (at atm bp), lb/ft^3				—
Density of liquid (at atm bp), kg/m^3				—
Vapor pressure at 25°C, psia				931.
Vapor pressure at 25°C, MN/m^2				6.42
Viscosity (abs), lbm/ft·s				9.4×10^{-6}
Viscosity (abs), centipoises[a]				0.014
Sound velocity in gas, m/s	223	221	221	270
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.327	0.365	0.37	0.205
Specific heat, c_p, J/kg K	1,368	1,527	1,548	876
Specific heat ratio, c_p/c_v	1.121	1.107	1.10	1.30
Gas constant *R,* ft-lb/lb·°F				35.1
Gas constant *R,* J/kg·°C				189
Thermal conductivity, Btu/h·ft·°F				0.01
Thermal conductivity, W/m°C				0.017
Boiling point (sat 14.7 psia), °F	38.6	33.6	19.2	–109.4[b]
Boiling point (sat 760 mm), °C	3.7	0.9	–7.1	–78.5
Latent heat of evap (at bp), Btu/lb	178.9	174.4	169	246
Latent heat of evap (at bp), J/kg	416,000	406,000	393,000	572,000
Freezing (melting) point, °F (1 atm)	–218	–158		
Freezing (melting) point, °C (1 atm)	–138.9	–105.5		
Latent heat of fusion, Btu/lb	31.2	41.6	25.3	—
Latent heat of fusion, J/kg	72,600	96,800	58,800	—
Critical temperature, °F				88
Critical temperature, °C	160	155		31
Critical pressure, psia	595	610		1,072
Critical pressure, MN/m^2	4.10	4.20		7.4
Critical volume, ft^3/lb				
Critical volume, m^3/kg				
Flammable (yes or no)	Yes	Yes	Yes	No
Heat of combustion, Btu/ft^3	3,150	3,150	3,150	—
Heat of combustion, Btu/lb	21,000	21,000	21,000	—
Heat of combustion, kJ/kg	48,800	48,800	48,800	—

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	Carbon Monoxide	Chlorine	Deuterium	Ethane
Chemical Formula	CO	Cl_2	D_2	C_2H_6
Refrigerant Number	—	—	—	170
	Chemical and Physical Properties			
Molecular weight	28.011	70.906	2.014	30.070
Specific gravity, air=1	0.967	2.45	0.070	1.04
Specific volume, ft^3/lb	14.0	5.52	194.5	13.025
Specific volume, m^3/kg	0.874	0.344	12.12	0.815
Density of liquid (at atm bp), lb/ft^3		97.3		28.
Density of liquid (at atm bp), kg/m^3		1,559		449
Vapor pressure at 25°C, psia			0.756	
Vapor pressure at 25°C, MN/m^2			0.0052	
Viscosity (abs), lbm/ft·s	12.1×10^{-6}	9.4×10^{-6}	8.75×10^{-6}	64×10^{-6}
Viscosity (abs), centipoises[a]	0.018	0.014	0.013	0.095
Sound velocity in gas, m/s	352	215	930	316
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.25	0.114	1.73	0.41
Specific heat, c_p, J/kg·K	1,046	477	7,238	1,715
Specific heat ratio, c_p/c_v	1.40	1.35	1.40	1.20
Gas constant *R*, ft-lb/lb·°F	55.2	21.8	384	51.4
Gas constant *R*, J/kg·°C	297	117	2,066	276
Thermal conductivity, Btu/h·ft·°F	0.014	0.005	0.081	0.010
Thermal conductivity, W/m·°C	0.024	0.0087	0.140	0.017
Boiling point (sat 14.7 psia), °F	–312.7	–29.2		–127
Boiling point (sat 760 mm), °C	–191.5	–34		–88.3
Latent heat of evap (at bp), Btu/lb	92.8	123.7		210
Latent heat of evap (at bp), J/kg	216,000	288,000		488,000
Freezing (melting) point, °F (1 atm)	–337	–150		–278
Freezing (melting) point, °C (1 atm)	–205	–101		–172.2
Latent heat of fusion, Btu/lb	12.8	41.0		41
Latent heat of fusion, J/kg		95,400		95,300
Critical temperature, °F	–220	291	–390.6	90.1
Critical temperature, °C	–140	144	–234.8	32.2
Critical pressure, psia	507	1,120	241	709
Critical pressure, MN/m^2	3.49	7.72	1.66	4.89
Critical volume, ft^3/lb	0.053	0.028	0.239	0.076
Critical volume, m^3/kg	0.0033	0.00175	0.0149	0.0047
Flammable (yes or no)	Yes	No		Yes
Heat of combustion, Btu/ft^3	310	—		
Heat of combustion, Btu/lb	4,340	—		22,300
Heat of combustion, kJ/kg	10,100	—		51,800

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name (s)	Ethyl Chloride	Ethylene (Ethene)	Fluorine
Chemical Formula	**C_2H_5Cl**	**C_2H_4**	**F_2**
Refrigerant Number	**160**	**1,150**	**—**
	Chemical and Physical Properties		
Molecular weight	64.515	28.054	37.996
Specific gravity, air = 1	2.23	0.969	1.31
Specific volume, ft^3/lb	6.07	13.9	10.31
Specific volume, m^3/kg	0.378	0.87	0.706
Density of liquid (at atm bp), lb/ft^3	56.5	35.5	
Density of liquid (at atm bp), kg/m^3	905.	569	
Vapor pressure at 25°C, psia			
Vapor pressure at 25°C, MN/m^2			
Viscosity (abs), lbm/ft·s		6.72×10^{-6}	16.1×10^{-6}
Viscosity (abs), centipoises[a]		0.010	0.024
Sound velocity in gas, m/s	204	331	290
	Thermal and Thermodynamic Properties		
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.27	0.37	0.198
Specific heat, c_p, J/kg·K	1,130	1,548	828
Specific heat ratio, c_p/c_v	1.13	1.24	1.35
Gas constant *R*, ft-lb/lb·°F	24.0	55.1	40.7
Gas constant *R*, J/kg·°C	129	296	219
Thermal conductivity, Btu/h·ft·°F		0.010	0.016
Thermal conductivity, W/m·°C		0.017	0.028
Boiling point (sat 14.7 psia), °F	54	−155	−306.4
Boiling point (sat 760 mm), °C	12.2	−103.8	−188
Latent heat of evap (at bp), Btu/lb	166	208	74
Latent heat of evap (at bp), J/kg	386,000	484,000	172,000
Freezing (melting) point, °F (1 atm)	−218	−272	−364
Freezing (melting) point, °C (1 atm)	−138.9	−169	−220
Latent heat of fusion, Btu/lb	29.3	51.5	11
Latent heat of fusion, J/kg	68,100	120,000	25,600
Critical temperature, °F	368.6	49	−200
Critical temperature, °C	187	9.5	−129
Critical pressure, psia	764	741	810
Critical pressure, MN/m^2	5.27	5.11	5.58
Critical volume, ft^3/lb	0.049	0.073	
Critical volume, m^3/kg	0.00306	0.0046	
Flammable (yes or no)	No	Yes	
Heat of combustion, Btu/ft^3	—	1,480	
Heat of combustion, Btu/lb	—	20,600	
Heat of combustion, kJ/kg	—	47,800	

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	**Fluorocarbons**			
Chemical Formula	**CCl_3F**	**CCl_2F_2**	**$CClF_3$**	**$CBrF_3$**
Refrigerant Number	**11**	**12**	**13**	**13 B1**
	Chemical and Physical Properties			
Molecular weight	137.37	120.91	104.46	148.91
Specific gravity, air = 1	4.74	4.17	3.61	5.14
Specific volume, ft^3/lb	2.74	3.12	3.58	2.50
Specific volume, m^3/kg	0.171	0.195	0.224	0.975
Density of liquid (at atm bp), lb/ft^3	92.1	93.0	95.0	124.4
Density of liquid (at atm bp), kg/m^3	1,475	1,490	1,522	1,993
Vapor pressure at 25°C, psia		94.51	516	234.8
Vapor pressure at 25°C, MN/m^2		0.652	3.56	1.619
Viscosity (abs), lbm/ft·s	7.39×10^{-6}	8.74×10^{-6}		
Viscosity (abs), centipoises[a]	0.011	0.013		
Sound velocity in gas, m/s				
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.14	0.146	0.154	
Specific heat, c_p, J/kg·K	586	611	644	
Specific heat ratio, c_p/c_v	1.14	1.14	1.145	
Gas constant *R*, ft-lb/lb·°F				
Gas constant *R*, J/kg·°C				
Thermal conductivity, Btu/h·ft·°F	0.005	0.006		
Thermal conductivity, W/m·°C	0.0087	0.0104		
Boiling point (sat 14.7 psia), °F	74.9	–21.8	–114.6	–72
Boiling point (sat 760 mm), °C	23.8	–29.9	–81.4	–57.8
Latent heat of evap (at bp), Btu/lb	77.5	71.1	63.0	51.1
Latent heat of evap (at bp), J/kg	180,000	165,000	147,000	119,000
Freezing (melting) point, °F (1 atm)	–168	–252	–294	–270
Freezing (melting) point, °C (1 atm)	–111	–157.8	–181.1	–167.8
Latent heat of fusion, Btu/lb				
Latent heat of fusion, J/kg				
Critical temperature, °F	388.4	233	83.9	152
Critical temperature, °C	198	111.7	28.8	66.7
Critical pressure, psia	635	582	559	573
Critical pressure, MN/m^2	4.38	4.01	3.85	3.95
Critical volume, ft^3/lb	0.0289	0.287	0.0277	0.0215
Critical volume, m^3/kg	0.00180	0.018	0.00173	0.00134
Flammable (yes or no)	No	No	No	No
Heat of combustion, Btu/ft^3	—	—	—	—
Heat of combustion, Btu/lb	—	—	—	—
Heat of combustion, kJ/kg	—	—	—	—

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	**Fluorocarbons**			
Chemical Formula	**CF_4**	**$CHCl_2F$**	**$CHClF_2$**	**$C_2Cl_2F_4$**
Refrigerant Number	**14**	**21**	**22**	**114**
	Chemical and Physical Properties			
Molecular weight	88.00	102.92	86.468	170.92
Specific gravity, air = 1	3.04	3.55	2.99	5.90
Specific volume, ft^3/lb	4.34	3.7	4.35	2.6
Specific volume, m^3/kg	0.271	0.231	0.271	0.162
Density of liquid (at atm bp), lb/ft^3	102.0	87.7	88.2	94.8
Density of liquid (at atm bp), kg/m^3	1,634	1,405	1,413	1,519
Vapor pressure at 25°C, psia		26.4	151.4	30.9
Vapor pressure at 25°C, MN/m^2		0.182	1.044	0.213
Viscosity (abs), lbm/ft·s		8.06×10^{-6}	8.74×10^{-6}	8.06×10^{-6}
Viscosity (abs), centipoises[a]		0.012	0.013	0.012
Sound velocity in gas, m/s				
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C		0.139	0.157	0.158
Specific heat, c_p, J/kg·K		582	657	661
Specific heat ratio, c_p/c_v		1.18	1.185	1.09
Gas constant *R,* ft-lb/lb·°F				
Gas constant *R,* J/kg·°C				
Thermal conductivity, Btu/h·ft·°F			0.007	0.006
Thermal conductivity, W/m·°C			0.012	0.010
Boiling point (sat 14.7 psia), °F	−198.2	48.1	−41.3	38.4
Boiling point (sat 760 mm), °C	−127.9	9.0	−40.7	3.55
Latent heat of evap (at bp), Btu/lb	58.5	104.1	100.4	58.4
Latent heat of evap (at bp), J/kg	136,000	242,000	234,000	136,000
Freezing (melting) point, °F (1 atm)	−299	−211	−256	−137
Freezing (melting) point, °C (1 atm)	−183.8	−135	−160	−93.8
Latent heat of fusion, Btu/lb	2.53			
Latent heat of fusion, J/kg	5,880			
Critical temperature, °F	−49.9	353.3	204.8	294
Critical temperature, °C	−45.5	178.5	96.5	
Critical pressure, psia	610	750	715	475
Critical pressure, MN/m^2	4.21	5.17	4.93	3.28
Critical volume, ft^3/lb	0.025	0.0307	0.0305	0.0275
Critical volume, m^3/kg	0.0016	0.00191	0.00190	0.00171
Flammable (yes or no)	No	No	No	No
Heat of combustion, Btu/ft^3	—	—	—	—
Heat of combustion, Btu/lb	—	—	—	—
Heat of combustion, kJ/kg	—	—	—	—

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	Fluorocarbons			Helium
Chemical Formula	**C_2ClF_5**	**$C_2H_3ClF_2$**	**$C_2H_4F_2$**	**He**
Refrigerant Number	**115**	**142b**	**152a**	**704**
	Chemical and Physical Properties			
Molecular weight	154.47	100.50	66.05	4.0026
Specific gravity, air=1	5.33	3.47	2.28	0.138
Specific volume, ft^3/lb	2.44	3.7	5.9	97.86
Specific volume, m^3/kg	0.152	0.231	0.368	6.11
Density of liquid (at atm bp), lb/ft^3	96.5	74.6	62.8	7.80
Density of liquid (at atm bp), kg/m^3	1,546	1,195	1,006	125
Vapor pressure at 25°C, psia	132.1	49.1	86.8	
Vapor pressure at 25°C, MN/m^2	0.911	0.3385	0.596	
Viscosity (abs), lbm/ft·s				13.4×10^{-6}
Viscosity (abs), centipoises[a]				0.02
Sound velocity in gas, m/s				1,015
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.161			1.24
Specific heat, c_p, J/kg·K	674			5,188
Specific heat ratio, c_p/c_v	1.091			1.66
Gas constant *R*, ft-lb/lb·°F				386
Gas constant *R*, J/kg·°C				2,077
Thermal conductivity, Btu/h ft·°F				0.086
Thermal conductivity, W/m·°C				0.149
Boiling point (sat 14.7 psia), °F	–38.0	14	–13	–452
Boiling point (sat 760 mm), °C	–38.9	–10.0	–25.0	4.22 K
Latent heat of evap (at bp), Btu/lb	53.4	92.5	137.1	10.0
Latent heat of evap (at bp), J/kg	124,000	215,000	319,000	23,300
Freezing (melting) point, °F (1 atm)	–149			[b]
Freezing (melting) point, °C (1 atm)	–100.6			—
Latent heat of fusion, Btu/lb				—
Latent heat of fusion, J/kg				—
Critical temperature, °F	176		387	–450.3
Critical temperature, °C				5.2 K
Critical pressure, psia	457.6			33.22
Critical pressure, MN/m^2	3.155			
Critical volume, ft^3/lb	0.0261			0.231
Critical volume, m^3/kg	0.00163			0.0144
Flammable (yes or no)	No	No	No	No
Heat of combustion, Btu/ft^3	—	—	—	—
Heat of combustion, Btu/lb	—	—	—	—
Heat of combustion, kJ/kg	—	—	—	—

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	Hydrogen	Hydrogen Chloride	Hydrogen Sulfide	Krypton
Chemical Formula	**H_2**	**HCl**	**H_2S**	**Kr**
Refrigerant Number	**702**	**—**	**—**	**—**
	Chemical and Physical Properties			
Molecular weight	2.016	36.461	34.076	83.80
Specific gravity, air = 1	0.070	1.26	1.18	2.89
Specific volume, ft^3/lb	194	10.74	11.5	4.67
Specific volume, m^3/kg	12.1	0.670	0.0930	0.291
Density of liquid (at atm bp), lb/ft^3	4.43	74.4	62	150.6
Density of liquid (at atm bp), kg/m^3	71.0	1,192	993	2,413
Vapor pressure at 25°C, psia				
Vapor pressure at 25°C, MN/m^2				
Viscosity (abs), lbm/ft·s	6.05×10^{-6}	10.1×10^{-6}	8.74×10^{-6}	16.8×10^{-6}
Viscosity (abs), centipoises[a]	0.009	0.015	0.013	0.025
Sound velocity in gas, m/s	1,315	310	302	223
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	3.42	0.194	0.23	0.059
Specific heat, c_p, J/kg·K	14,310	812	962	247
Specific heat ratio, c_p/c_v	1.405	1.39	1.33	1.68
Gas constant *R,* ft-lb/lb·°F	767	42.4	45.3	18.4
Gas constant *R,* J/kg·°C	4,126	228	244	99.0
Thermal conductivity, Btu/h·ft·°F	0.105	0.008	0.008	0.0054
Thermal conductivity, W/m·°C	0.0182	0.014	0.014	0.0093
Boiling point (sat 14.7 psia), °F	–423	–121	–76	–244
Boiling point (sat 760 mm), °C	20.4 K	–85	–60	–153
Latent heat of evap (at bp), Btu/lb	192	190.5	234	46.4
Latent heat of evap (at bp), J/kg	447,000	443,000	544,000	108,000
Freezing (melting) point, °F (1 atm)	–434.6	–169.6	–119.2	–272
Freezing (melting) point, °C (1 atm)	–259.1	–112	–84	–169
Latent heat of fusion, Btu/lb	25.0	23.4	30.2	4.7
Latent heat of fusion, J/kg	58,000	54,400	70,200	10,900
Critical temperature, °F	–399.8	124	213	
Critical temperature, °C	–240.0	51.2	100.4	–63.8
Critical pressure, psia	189	1,201	1,309	800
Critical pressure, MN/m^2	1.30	8.28	9.02	5.52
Critical volume, ft^3/lb	0.53	0.038	0.046	0.0177
Critical volume, m^3/kg	0.033	0.0024	0.0029	0.0011
Flammable (yes or no)	Yes	No	Yes	No
Heat of combustion, Btu/ft^3	320	—	700	—
Heat of combustion, Btu/lb	62,050	—	8,000	—
Heat of combustion, kJ/kg	144,000	—	18,600	—

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	Methane	Methyl Chloride	Neon	Nitric Oxide
Chemical Formula	**CH_4**	**CH_3Cl**	**Ne**	**NO**
Refrigerant Number	**50**	**40**	**720**	**—**
	Chemical and Physical Properties			
Molecular weight	16.044	50.488	20.179	30.006
Specific gravity, air = 1	0.554	1.74	0.697	1.04
Specific volume, ft^3/lb	24.2	7.4	19.41	13.05
Specific volume, m^3/kg	1.51	0.462	1.211	0.814
Density of liquid (at atm bp), lb/ft^3	26.3	62.7	75.35	
Density of liquid (at atm bp), kg/m^3	421	1,004	1,207	
Vapor pressure at 25°C, psia		82.2		
Vapor pressure at 25°C, MN/m^2		0.567		
Viscosity (abs), lbm/ft·s	7.39×10^{-6}	7.39×10^{-6}	21.5×10^{-6}	12.8×10^{-6}
Viscosity (abs), centipoises[a]	0.011	0.011	0.032	0.019
Sound velocity in gas, m/s	446	251	454	341
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.54	0.20	0.246	0.235
Specific heat, c_p, J/kg·K	2,260	837	1,030	983
Specific heat ratio, c_p/c_v	1.31	1.28	1.64	1.40
Gas constant *R*, ft-lb/lb·°F	96	30.6	76.6	51.5
Gas constant *R*, J/kg·°C	518	165	412	277
Thermal conductivity, Btu/h ft·°F	0.02	0.006	0.028	0.015
Thermal conductivity, W/m·°C	0.035	0.010	0.048	0.026
Boiling point (sat 14.7 psia), °F	–259	–10.7	–410.9	–240
Boiling point (sat 760 mm), °C	–434.2	–23.7	–246	–151.5
Latent heat of evap (at bp), Btu/lb	219.2	184.1	37	
Latent heat of evap (at bp), J/kg	510,000	428,000	86,100	
Freezing (melting) point, °F (1 atm)	–296.6	–144	–415.6	–258
Freezing (melting) point, °C (1 atm)	–182.6	–97.8	–248.7	–161
Latent heat of fusion, Btu/lb	14	56	6.8	32.9
Latent heat of fusion, J/kg	32,600	130,000	15,800	76,500
Critical temperature, °F	–116	289.4	–379.8	–136
Critical temperature, °C	–82.3	143	–228.8	–93.3
Critical pressure, psia	673	968	396	945
Critical pressure, MN/m^2	4.64	6.67	2.73	6.52
Critical volume, ft^3/lb	0.099	0.043	0.033	0.0332
Critical volume, m^3/kg	0.0062	0.0027	0.0020	0.00207
Flammable (yes or no)	Yes	Yes	No	No
Heat of combustion, Btu/ft^3	985		—	—
Heat of combustion, Btu/lb	2,290		—	—
Heat of combustion, kJ/kg			—	—

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	Nitrogen	Nitrous Oxide	Oxygen	Ozone
Chemical Formula	N_2	N_2O	O_2	O_3
Refrigerant Number	**728**	**744A**	**732**	**—**
	Chemical and Physical Properties			
Molecular weight	28.0134	44.012	31.9988	47.998
Specific gravity, air=1	0.967	1.52	1.105	1.66
Specific volume, ft^3/lb	13.98	8.90	12.24	8.16
Specific volume, m^3/kg	0.872	0.555	0.764	0.509
Density of liquid (at atm bp), lb/ft^3	50.46	76.6	71.27	
Density of liquid (at atm bp), kg/m^3	808.4	1,227	1,142	
Vapor pressure at 25°C, psia				
Vapor pressure at 25°C, MN/m^2				8.74×10^{-6}
Viscosity (abs), lbm/ft·s	12.1×10^{-6}	10.1×10^{-6}	13.4×10^{-6}	
Viscosity (abs), centipoises[a]	0.018	0.015	0.020	0.013
Sound velocity in gas, m/s	353	268	329	
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.249	0.21	0.220	0.196
Specific heat, c_p, J/kg·K	1,040	879	920	820
Specific heat ratio, c_p/c_v	1.40	1.31	1.40	
Gas constant *R*, ft-lb/lb·°F	55.2	35.1	48.3	32.2
Gas constant *R*, J/kg·°C	297	189	260	173
Thermal conductivity, Btu/h ft·°F	0.015	0.010	0.01	0.015
Thermal conductivity, W/m·°C	0.026	0.017	0.026	0.033
Boiling point (sat 14.7 psia), °F	–320.4	–127.3	–297.3	–170
Boiling point (sat 760 mm), °C	–195.8	–88.5	–182.97	–112
Latent heat of evap (at bp), Btu/lb	85.5	161.8	91.7	
Latent heat of evap (at bp), J/kg	199,000.	376,000.	213,000	
Freezing (melting) point, °F (1 atm)	–346	–131.5	–361.1	–315.5
Freezing (melting) point, °C (1 atm)	–210	–90.8	–218.4	–193
Latent heat of fusion, Btu/lb	11.1	63.9	5.9	97.2
Latent heat of fusion, J/kg	25,800	149,000	13,700	226,000
Critical temperature, °F	–232.6	97.7	–181.5	16
Critical temperature, °C	–147	36.5	–118.6	–9
Critical pressure, psia	493	1,052	726	800
Critical pressure, MN/m^2	3.40	7.25	5.01	5.52
Critical volume, ft^3/lb	0.051	0.036	0.040	0.0298
Critical volume, m^3/kg	0.00318	0.0022	0.0025	0.00186
Flammable (yes or no)	No	No	No	No
Heat of combustion, Btu/ft^3	—	—	—	—
Heat of combustion, Btu/lb	—	—	—	—
Heat of combustion, kJ/kg	—	—	—	—

(*Continued*)

TABLE A.7 (*Continued*)
Chemical, Physical, and Thermal Properties of Gases: Gases and Vapors, Including Fuels and Refrigerants, English and Metric Units

Common Name(s)	Propane	Propylene (Propene)	Sulfur Dioxide	Xenon
Chemical Formula	**C_3H_8**	**C_3H_6**	**SO_2**	**Xe**
Refrigerant Number	**290**	**1270**	**764**	**—**
	Chemical and Physical Properties			
Molecular weight	44.097	42.08	64.06	131.30
Specific gravity, air = 1	1.52	1.45	2.21	4.53
Specific volume, ft^3/lb	8.84	9.3	6.11	2.98
Specific volume, m^3/kg	0.552	0.58		
Density of liquid (at atm bp), lb/ft^3	36.2	37.5	42.8	190.8
Density of liquid (at atm bp), kg/m^3	580	601	585	3,060
Vapor pressure at 25°C, psia	135.7	166.4	56.6	
Vapor pressure at 25°C, MN/m^2	0.936	1.147	0.390	
Viscosity (abs), lbm/ft s	53.8×10^{-6}	57.1×10^{-6}	8.74×10^{-6}	15.5×10^{-6}
Viscosity (abs), centipoises[a]	0.080	0.085	0.013	0.023
Sound velocity in gas, m/s	253	261	220	177
	Thermal and Thermodynamic Properties			
Specific heat, c_p, Btu/lb·°F or cal/g·°C	0.39	0.36	0.11	0.115
Specific heat, c_p, J/kg·K	1,630	1,506	460	481
Specific heat ratio, c_p/c_v	1.2	1.16	1.29	1.67
Gas constant *R,* ft-lb/lb·°F	35.0	36.7	24.1	11.8
Gas constant *R,* J/kg·°C	188	197	130	63.5
Thermal conductivity, Btu/h·ft·°F	0.010	0.010	0.006	0.003
Thermal conductivity, W/m·°C	0.017	0.017	0.010	0.0052
Boiling point (sat 14.7 psia), °F	–44	–54	14.0	–162.5
Boiling point (sat 760 mm), °C	–42.2	–48.3	–10	–108
Latent heat of evap (at bp), Btu/lb	184	188.2	155.5	41.4
Latent heat of evap (at bp), J/kg	428,000	438,000	362,000	96,000
Freezing (melting) point, °F (1 atm)	–309.8	–301	–104	–220
Freezing (melting) point, °C (1 atm)	–189.9	–185	–75.5	–140
Latent heat of fusion, Btu/lb	19.1		58.0	10
Latent heat of fusion, J/kg	44,400		135,000	23,300
Critical temperature, °F	205	197	315.5	61.9
Critical temperature, °C	96	91.7	157.6	16.6
Critical pressure, psia	618	668	1,141	852
Critical pressure, MN/m^2	4.26	4.61	7.87	5.87
Critical volume, ft^3/lb	0.073	0.069	0.03	0.0145
Critical volume, m^3/kg	0.0045	0.0043	0.0019	0.00090
Flammable (yes or no)	Yes	Yes	No	No
Heat of combustion, Btu/ft^3	2,450	2,310	—	—
Heat of combustion, Btu/lb	21,660	21,500	—	—
Heat of combustion, kJ/kg	50,340	50,000	—	—

Note: The properties of pure gases are given at 25°C (77°F, 298 K) and atmospheric pressure (except as stated).

[a] For $N \cdot s/m^2$ divide by 1,000.

[b] Helium cannot be solidified at atmospheric pressure.

TABLE A.8
Ideal Gas Properties of Air

T	H	P_t	U	v_t	s°
		Part a. SI Units			
		T(K), H and U(kJ/kg), s°(kJ/kg·K)			
200	199.97	0.3363	142.56	1,707	1.29559
210	209.97	0.3987	149.69	1,512	1.34444
220	219.97	0.4690	156.82	1,346	1.39105
230	230.02	0.5477	164.00	1,205	1.43557
240	240.02	0.6355	171.13	1,084	1.47824
250	250.05	0.7329	178.28	979	1.51917
260	260.09	0.8405	185.45	887.8	1.55848
270	270.11	0.9590	192.60	808.0	1.59634
280	280.13	1.0889	199.75	738.0	1.63279
285	285.14	1.1584	203.33	706.1	1.65055
290	290.16	1.2311	206.91	676.1	1.66802
295	295.17	1.3068	210.49	647.9	1.68515
300	300.19	1.3860	214.07	621.2	1.70203
305	305.22	1.4686	217.67	596.0	1.71865
310	310.24	1.5546	221.25	572.3	1.73498
315	315.27	1.6442	224.85	549.8	1.75106
320	320.29	1.7375	228.42	528.6	1.76690
325	325.31	1.8345	232.02	508.4	1.78249
330	330.34	1.9352	235.61	489.4	1.79783
340	340.42	2.149	242.82	454.1	1.82790
350	350.49	2.379	250.02	422.2	1.85708
360	360.58	2.626	257.24	393.4	1.88543
370	370.67	2.892	264.46	367.2	1.91313
380	380.77	3.176	271.69	343.4	1.94001
390	390.88	3.481	278.93	321.5	1.96633
400	400.98	3.806	286.16	301.6	1.99194
410	411.12	4.153	293.43	283.3	2.01699
420	421.26	4.522	300.69	266.6	2.04142
430	431.43	4.915	307.99	251.1	2.06533
440	441.61	5.332	315.30	236.8	2.08870
450	451.80	5.775	322.62	223.6	2.11161
460	462.02	6.245	329.97	211.4	2.13407
470	472.24	6.742	337.32	200.1	2.15604
480	482.49	7.268	344.70	189.5	2.17760
490	492.74	7.824	352.08	179.7	2.19876
500	503.02	8.411	359.49	170.6	2.21952
510	513.32	9.031	366.92	162.1	2.23993
520	523.63	9.684	374.36	154.1	2.25997
530	533.98	10.37	381.84	146.7	2.27967
540	544.35	11.10	389.34	139.7	2.29906
550	554.74	11.86	396.86	133.1	2.31809
560	565.17	12.66	404.42	127.0	2.33685
570	575.59	13.50	411.97	121.2	2.35531
580	586.04	14.38	419.55	115.7	2.37348
590	596.52	15.31	427.15	110.6	2.39140

(*Continued*)

TABLE A.8 (*Continued*)
Ideal Gas Properties of Air

T	H	P_t	U	v_t	s°
600	607.02	16.28	434.78	105.8	2.40902
610	617.53	17.30	442.42	101.2	2.42644
620	628.07	18.36	450.09	96.92	2.44356
630	638.63	19.84	457.78	92.84	2.46048
640	649.22	20.64	465.50	88.99	2.47716
650	659.84	21.86	473.25	85.34	2.49364
660	670.47	23.13	481.01	81.89	2.50985
670	681.14	24.46	488.81	78.61	2.52589
680	691.82	25.85	496.62	75.50	2.54175
690	702.52	27.29	504.45	72.56	2.55731
700	713.27	28.80	512.33	69.76	2.57277
710	724.04	30.38	520.23	67.07	2.58810
720	734.82	32.02	528.14	64.53	2.60319
730	745.62	33.72	536.07	62.13	2.61803
740	756.44	35.50	544.02	59.82	2.63280
750	767.29	37.35	551.99	57.63	2.64737
760	778.18	39.27	560.01	55.54	2.66176
770	789.11	41.31	568.07	53.39	2.67595
780	800.03	43.35	576.12	51.64	2.69013
790	810.99	45.55	584.21	49.86	2.70400
800	821.95	47.75	592.30	48.08	2.71787
820	843.98	52.59	608.59	44.84	2.74504
840	866.08	57.60	624.95	41.85	2.77170
860	888.27	63.09	641.40	39.12	2.79783
880	910.56	68.98	657.95	36.61	2.82344
900	932.93	75.29	674.58	34.31	2.84856
920	955.38	82.05	691.28	32.18	2.87324
940	977.92	89.28	708.08	30.22	2.89748
960	1,000.55	97.00	725.02	28.40	2.92128
980	1,023.25	105.2	741.98	26.73	2.94468
1,000	1,046.04	114.0	758.94	25.17	2.96770
1,020	1,068.89	123.4	776.10	23.72	2.99034
1,040	1,091.85	133.3	793.36	22.39	3.01260
1,060	1,114.86	143.9	810.62	21.14	3.03449
1,080	1,137.89	155.2	827.88	19.98	3.05608
1,100	1,161.07	167.1	845.33	18.896	3.07732
1,120	1,184.28	179.7	862.79	17.886	3.09825
1,140	1,207.57	193.1	880.35	16.946	3.11883
1,160	1,230.92	207.2	897.91	16.064	3.13916
1,180	1,254.34	222.2	915.57	15.241	3.15916
1,200	1,277.79	238.0	933.33	14.470	3.17888
1,220	1,301.31	254.7	951.09	13.747	3.19834
1,240	1,324.93	272.3	968.95	13.069	3.21751
1,260	1,348.55	290.8	986.90	12.435	3.23638
1,280	1,372.24	310.4	1,004.76	11.835	3.25510
1,300	1,395.97	330.9	1,022.82	11.275	3.27345
1,320	1,419.76	352.5	1,040.88	10.747	3.29160
1,340	1,443.60	375.3	1,058.94	10.247	3.30959

(*Continued*)

TABLE A.8 (*Continued*)
Ideal Gas Properties of Air

T	H	P_t	U	v_t	s°
1,360	1,467.49	399.1	1,077.10	9.780	3.32724
1,380	1,491.44	424.2	1,095.26	9.337	3.34474
1,400	1,515.42	450.5	1,113.52	8.919	3.36200
1,420	1,539.44	478.0	1,131.77	8.526	3.37901
1,440	1,563.51	506.9	1,150.13	8.153	3.39586
1,460	1,587.63	537.1	1,168.49	7.801	3.41247
1,480	1,611.79	568.8	1,186.95	7.468	3.42892
1,500	1,635.97	601.9	1,205.41	7.152	3.44516
1,520	1,660.23	636.5	1,223.87	6.854	3.46120
1,540	1,684.51	672.8	1,242.43	6.569	3.47712
1,560	1,708.82	710.5	1,260.99	6.301	3.49276
1,580	1,733.17	750.0	1,279.65	6.046	3.50829
1,600	1,757.57	791.2	1,298.30	5.804	3.52364
1,620	1,782.00	834.1	1,316.96	5.574	3.53879
1,640	1,806.46	878.9	1,335.72	5.355	3.55381
1,660	1,830.96	925.6	1,354.48	5.147	3.56867
1,680	1,855.50	974.2	1,373.24	4.949	3.58335
1,700	1,880.1	1,025	1,392.7	4.761	3.5979
1,750	1,941.6	1,161	1,439.8	4.328	3.6336
1,800	2,003.3	1,310	1,487.2	3.944	3.6684
1,850	2,065.3	1,475	1,534.9	3.601	3.7023
1,900	2,127.4	1,655	1,582.6	3.295	3.7354
1,950	2,189.7	1,852	1,630.6	3.022	3.7677
2,000	2,252.1	2,068	1,678.7	2.776	3.7994
2,050	2,314.6	2,303	1,726.8	2.555	3.8303
2,100	2,377.4	2,559	1,775.3	2.356	3.8605
2,150	2,440.3	2,837	1,823.8	2.175	3.8901
2,200	2,503.2	3,138	1,872.4	2.012	3.9191
2,250	2,566.4	3,464	1,921.3	1.864	3.9474
		Part b. English Units			
360	85.97	0.3363	61.29	396.6	0.50369
380	90.75	0.4061	64.70	346.6	0.51663
400	95.53	0.4858	68.11	305.0	0.52890
420	100.32	0.5760	71.52	270.1	0.54058
440	105.11	0.6776	74.93	240.6	0.55172
460	109.90	0.7913	78.36	215.33	0.56235
480	114.69	0.9182	81.77	193.65	0.57255
500	119.48	1.0590	85.20	174.90	0.58233
520	124.27	1.2147	88.62	158.58	0.59172
537	128.34	1.3593	91.53	146.34	0.59945
540	129.06	1.3860	92.04	144.32	0.60078
560	133.86	1.5742	95.47	131.78	0.60950
580	138.66	1.7800	98.90	120.70	0.61793
600	143.47	2.005	102.34	110.88	0.62607
620	148.28	2.249	105.78	102.12	0.63395
640	153.09	2.514	109.21	94.30	0.64159
660	157.92	2.801	112.67	87.27	0.64902

(*Continued*)

TABLE A.8 (*Continued*)
Ideal Gas Properties of Air

T	H	P_t	U	v_t	s°
680	162.73	3.111	116.12	80.96	0.65621
700	167.56	3.446	119.58	75.25	0.66321
720	172.39	3.806	123.04	70.07	0.67002
740	177.23	4.193	126.51	65.38	0.67665
760	182.08	4.607	129.99	61.10	0.68312
780	186.94	5.051	133.47	57.20	0.68942
800	191.81	5.526	136.97	53.63	0.69558
820	196.69	6.033	140.47	50.35	0.70160
840	201.56	6.573	143.98	47.34	0.70747
860	206.46	7.149	147.50	44.57	0.71323
880	211.35	7.761	151.02	42.01	0.71886
900	216.26	8.411	154.57	39.64	0.72438
920	221.18	9.102	158.12	37.44	0.72979
940	226.11	9.834	161.68	35.41	0.73509
960	231.06	10.61	165.26	33.52	0.74030
980	236.02	11.43	168.83	31.76	0.74540
1,000	240.98	12.30	172.43	30.12	0.75042
1,040	250.95	14.18	179.66	27.17	0.76019
1,080	260.97	16.28	186.93	24.58	0.76964
1,120	271.03	18.60	194.25	22.30	0.77880
1,160	281.14	21.18	201.63	20.29	0.78767
1,200	291.30	24.01	209.05	18.51	0.79628
1,240	301.52	27.13	216.53	16.93	0.80466
1,280	311.79	30.55	224.05	15.52	0.81280
1,320	322.11	34.31	231.63	14.25	0.82075
1,360	332.48	38.41	239.25	13.12	0.82848
1,400	342.90	42.88	246.93	12.10	0.83604
1,440	353.37	47.75	254.66	11.17	0.84341
1,480	363.89	53.04	262.44	10.34	0.85062
1,520	374.47	58.78	270.26	9.578	0.85767
1,560	385.08	65.00	278.13	8.890	0.86456
1,600	395.74	71.73	286.06	8.263	0.87130
1,650	409.13	80.89	296.03	7.556	0.87954
1,700	422.59	90.95	306.06	6.924	0.88758
1,750	436.12	101.98	316.16	6.357	0.89542
1,800	449.71	114.0	326.32	5.847	0.90308
1,850	463.37	127.2	336.55	5.388	0.91056
1,900	477.09	141.5	346.85	4.974	0.91788
1,950	490.88	157.1	357.20	4.598	0.92504
2,000	504.71	174.0	367.61	4.258	0.93205
2,050	518.61	192.3	378.08	3.949	0.93891
2,100	532.55	212.1	388.60	3.667	0.94564
2,150	546.54	233.5	399.17	3.410	0.95222
2,200	560.59	256.6	409.78	3.176	0.95868
2,250	574.69	281.4	420.46	2.961	0.96501
2,300	588.82	308.1	431.16	2.765	0.97123
2,350	603.00	336.8	441.91	2.585	0.97732
2,400	617.22	367.6	452.70	2.419	0.98331

(*Continued*)

TABLE A.8 (*Continued*)
Ideal Gas Properties of Air

T	H	P_t	U	v_t	s°
2,450	631.48	400.5	463.54	2.266	0.98919
2,500	645.78	435.7	474.40	2.125	0.99497
2,550	660.12	473.3	485.31	1.996	1.00064
2,600	674.49	513.5	496.26	1.876	1.00623
2,650	688.90	556.3	507.25	1.765	1.01172
2,700	703.35	601.9	518.26	1.662	1.01712
2,750	717.83	650.4	529.31	1.566	1.02244
2,800	732.33	702.0	540.40	1.478	1.02767
2,850	746.88	756.7	551.52	1.395	1.03282
2,900	761.45	814.8	562.66	1.318	1.03788
2,950	776.05	876.4	573.84	1.247	1.04288
3,000	790.68	941.4	585.04	1.180	1.04779
3,050	805.34	1,011	596.28	1.118	1.05264
3,100	820.03	1,083	607.53	1.060	1.05741
3,150	834.75	1,161	618.82	1.006	1.06212
3,200	849.48	1,242	630.12	0.9546	1.06676
3,250	864.24	1,328	641.46	0.9069	1.07134
3,300	879.02	1,418	652.81	0.8621	1.07585
3,350	893.83	1,513	664.20	0.8202	1.08031
3,400	908.66	1,613	675.60	0.7807	1.08470
3,450	923.52	1,719	687.04	0.7436	1.08904
3,500	938.40	1,829	698.48	0.7087	1.09332
3,550	953.30	1,946	709.95	0.6759	1.09755
3,600	968.21	2,068	721.44	0.6449	1.10172
3,650	983.15	2,196	732.95	0.6157	1.10584
3,700	998.11	2,330	744.48	0.5882	1.10991
3,750	1,013.1	2,471	756.04	0.5621	1.11393
3,800	1,028.1	2,618	767.60	0.5376	1.11791
3,850	1,043.1	2,773	779.19	0.5143	1.12183
3,900	1,058.1	2,934	790.80	0.4923	1.12571
3,950	1,073.2	3,103	802.43	0.4715	1.12955
4,000	1,088.3	3,280	814.06	0.4518	1.13334
4,050	1,103.4	3,464	825.72	0.4331	1.13709
4,100	1,118.5	3,656	837.40	0.4154	1.14079
4,150	1,133.6	3,858	849.09	0.3985	1.14446
4,200	1,148.7	4,067	860.81	0.3826	1.14809
4,300	1,179.0	4,513	884.28	0.3529	1.15522
4,400	1,209.4	4,997	907.81	0.3262	1.16221
4,500	1,239.9	5,521	931.39	0.3019	1.16905
4,600	1,270.4	6,089	955.04	0.2799	1.17575
4,700	1,300.9	6,701	978.73	0.2598	1.18232
4,800	1,331.5	7,362	1,002.5	0.2415	1.18876
4,900	1,362.2	8,073	1,026.3	0.2248	1.19508
5,000	1,392.9	8,837	1,050.1	0.2096	1.20129
5,100	1,423.6	9,658	1,074.0	0.1956	1.20738
5,200	1,454.4	10,539	1,098.0	0.1828	1.21336
5,300	1,485.3	11,481	1,122.0	0.1710	1.21923

Source: Moran, M.J., and Shapiro, H.N., *Fundamentals of Engineering Thermodynamics*, 3rd. ed., Wiley, New York, 1995, as based on Keenan, J.H., and Kaye, J., *Gas Tables,* Wiley, New York, 1945.

TABLE A.9
Equations for Gas Properties

Gas	Molar Mass M kg/kmol	Gas Constant R kJ/kg·K	Specific Heats at 25°C c_p kJ/kg K	c_v kJ/kg·K	k	Temperature Range, K	Equation Coefficients for $c_p/R=a+bT+cT^2+dT^3+eT^4$ a	$b\times10^3$ K^{-1}	$c\times10^6$ K^{-2}	$d\times10^{10}$ K^{-3}	$e\times10^{13}$ K^{-4}	Critical State Properties p_c MPa	T_c K	Redlich–Kwong Constants a, kPa-m^6·K$^{0.5}$/kmol^2	b m^3/kmol	Gas
Acetylene, C_2H_2	26.04	0.319	1.69	1.37	1.232	300–1,000	0.8021	23.51	−35.95	286.1	−87.64	6.14	308	8,030	0.0362	Acetylene, C_2H_2
						1000–3,000	3.825	6.767	−3.014	6.931	−0.6469					
Air	28.97	0.287	1.01	0.718	1.400	300–1,000	3.721	−1.874	4.719	−34.45	8.531	3.77	132	1,580	0.0253	Air
						1,000–3,000	2.786	1.925	−0.9465	2.321	−0.2229					
Argon, Ar	39.95	0.208	0.520	0.312	1.667		2.50	0	0	0	0	4.90	151	1,680	0.0222	Argon, Ar
Butane, C_4H_{10}	58.12	0.143	1.67	1.53	1.094	300–1,500	0.4756	44.65	−22.04	42.07	0	3.80	425	29,000	0.0806	Butane, C_4H_{10}
Carbon dioxide, CO_2	44.01	0.189	0.844	0.655	1.289	300–1,000	2.227	9.992	−9.802	53.97	−12.81	7.38	304	6,450	0.0297	Carbon dioxide, CO_2
						1,000–3,000	3.247	5.847	−3.412	9.469	−1.009					
Carbon monoxide, CO	28.01	0.297	1.04	0.744	1.399	300–1,000	3.776	−2.093	4.880	−32.71	6.984	3.50	133	1,720	0.0274	Carbon monoxide, CO
						1,000–3,000	2.654	2.226	−1.146	2.551	−0.2762					
Ethane, C_2H_6	30.07	0.276	1.75	1.48	1.187	300–1,500	0.8293	20.75	−7.704	8.756	0	4.88	306	9,860	0.0450	Ethane, C_2H_6
Ethylene, C_2H_4	28.05	0.296	1.53	1.23	1.240	300–1,000	1.575	10.19	11.25	−199.1	81.98	5.03	282	7,860	0.0404	Ethylene, C_2H_4
						1,000–3,000	0.2530	18.67	−9.978	26.03	−2.668					
Helium, He	4.003	2.08	5.19	3.12	1.667		2.50	0	0	0	0	0.228	5.20	8.00	0.0165	Helium, He
Hydrogen, H_2	2.016	4.12	14.3	10.2	1.405	3,00–1,000	2.892	3.884	−8.850	86.94	−29.18	1.31	33.2	143	0.0182	Hydrogen, H_2
						1,000–3,000	3.717	−0.9220	1.221	−4.328	0.5202					
Hydrogen, H	1.008	8.25	20.6	12.4	1.667	300–1,000	2.496	0.02977	−0.07655	0.8238	−0.3158					Hydrogen, H
						1,000–3,000	2.567	−0.1509	0.1219	−0.4184	0.05182					

(*Continued*)

TABLE A.9 (*Continued*)
Equations for Gas Properties

Gas	Molar Mass M kg/kmol	Gas Constant R kJ/kg·K	Specific Heats at 25°C c_p kJ/kg K	c_v kJ/kg·K	k	Temperature Range, K	Equation Coefficients for $c_p/R = a + bT + cT^2 + dT^3 + eT^4$ a	$b\times10^3$ K^{-1}	$c\times10^6$ K^{-2}	$d\times10^{10}$ K^{-3}	$e\times10^{13}$ K^{-4}	Critical State Properties p_c MPa	T_c K	Redlich–Kwong Constants a, kPa-m^6·K$^{0.5}$/kmol^2	b m^3/kmol	Gas
Hydroxyl, H	17.01	0.489	1.76	1.27	1.384	300–1,000	3.874	–1.349	1.670	–5.670	0.6189					Hydroxyl, OH
						1,000–3,000	3.229	0.2014	0.4357	–2.043	0.2696					
Methane, CH_4	16.04	0.518	2.22	1.70	1.304	300–1,000	4.503	–8.965	37.38	–364.9	122.2	4.60	191	3,210	0.0298	Methane CH_4
						1,000–3,000	–0.6992	15.31	–7.695	18.96	–1.849					
Neon, Ne	20.18	0.412	1.03	0.618	1.667		250	0	0	0	0	2.65	44.4	146	0.0120	Neon, Ne
Nitric oxide, NO	30.01	0.277	0.995	0.718	1.386	300–1,000	4.120	–4.225	10.77	–97.64	31.85	6.48	180	1,980	0.0200	Nitric oxide, NO
						1,000–3,000	2.730	2.372	–1.338	3.604	–0.3743					
Nitrogen, N_2	28.01	0.297	1.04	0.743	1.400	300–1,000	3.725	–1.562	3.208	–15.54	1.154	3.39	126	1,550	0.0267	Nitrogen, N_2
						1,000–3,000	2.469	2.467	–1.312	3.401	–0.3454					
Nitrogen, N	14.01	0.594	1.48	0.890	1.667	300–1,000	2.496	0.02977	–0.07655	0.8238	–0.3158					Nitrogen, N
						1,000–3,000	2.483	0.03033	–0.01517	0.001879	0.009657					
Oxygen, O_2	32.00	0.260	0.919	0.659	1.395	300–1,000	3.837	–3.420	10.99	–109.6	37.47	5.04	155	1,740	0.0221	Oxygen, O_2
						1,000–3,000	3.156	1.809	–1.052	3.190	–0.3629					
Oxygen, O	16.00	0.520	1.37	0.850	1.612	300–1,000	3.020	–2.176	3.793	–30.62	9.402					Oxygen, O
						1000–3,000	2.662	–0.3051	0.2250	–0.7447	0.09383					
Propane, C_3H_8	44.10	0.189	1.67	1.48	1.127	300–1,500	–0.4861	36.63	–18.91	38.14	0	4.26	370	18,300	0.0626	Propane, C_3H_8
Water, H_2O	18.02	0.462	1.86	1.40	1.329	300–1,000	4.132	–1.559	5.315	–42.09	12.84	22.1	647	14,300	0.0211	Water, H_2O
						1,000–3,000	2.798	2.693	–0.5392	–0.01783	0.09027					

Source: Jones, J.B., and Dugan, R.E., *Engineering Thermodynamics,* Prentice-Hall, Englewood Cliffs, NJ, 1996 from various sources: *JANAF Thermochemical Tables,* 3rd ed., published by the American Chemical Society and the American Institute of Physics for the National Bureau of Standards, 1986. Data for butane, ethane, and propane from Kobe, K.A., and Long, E.G., *Pet. Refin.,* 28(2), 113–116, 1949.

Appendix B: Properties of Liquids

TABLE B.1
Properties of Liquid Water

Temp. °F	At 1 atm or 14.7 psia				At 1,000 psia				At 10,000 psia			
	ρ	c_p	μ	k	ρ	c_p	μ	k	ρ	c_p	μ	k^a
	Symbols and Units:											
	ρ = density, lb_m/ft^3. For g/cm^3 multiply by 0.016018. For kg/m^3 multiply by 16.018.											
	c_p = specific heat, Btu/lb_m R = cal/g·K. For J/kg·K multiply by 4186.8.											
	μ = viscosity. For lbf·s/ft^2 = slugs/s·ft, multiply by 10^{-7}. For lb_m·s·ft multiply by 10^{-7} and by 32.174. For g/s·cm (poise) multiply by 10^{-7} and by 478.80. For N·s/m^2 multiply by 10^{-7} and by 478.880.											
	k = thermal conductivity, Btu/hr-ft R. For W/m·K multiply by 1.7307.											
32	62.42	1.007	366	0.3286	62.62	0.999	365	0.3319	64.5	0.937	357	0.3508
40	62.42	1.004	323	0.334	62.62	0.997	323	0.337	64.5	0.945	315	0.356
50	62.42	1.002	272	0.3392	62.62	0.995	272	0.3425	64.5	0.951	267	0.3610
60	62.38	1.000	235	0.345	62.58	0.994	235	0.348	64.1	0.956	233	0.366
70	62.31	0.999	204	0.350	62.50	0.994	204	0.353	64.1	0.960	203	0.371
80	62.23	0.998	177	0.354	62.42	0.994	177	0.358	64.1	0.962	176	0.376
90	62.11	0.998	160	0.359	62.31	0.994	160	0.362	63.7	0.964	159	0.380
100	62.00	0.998	142	0.3633	62.19	0.994	142	0.3666	63.7	0.965	142	0.3841
110	61.88	0.999	126	0.367	62.03	0.994	126	0.371	63.7	0.966	126	0.388
120	61.73	0.999	114	0.371	61.88	0.995	114	0.374	63.3	0.967	114	0.391
130	61.54	0.999	105	0.374	61.73	0.995	105	0.378	63.3	0.968	105	0.395
140	61.39	0.999	96	0.378	61.58	0.996	96	0.381	63.3	0.969	98	0.398
150	61.20	1.000	89	0.3806	61.39	0.996	89	0.3837	63.0	0.970	91	0.4003
160	61.01	1.001	83	0.383	61.20	0.997	83	0.386	62.9	0.971	85	0.403
170	60.79	1.002	77	0.386	60.98	0.998	77	0.389	62.5	0.972	79	0.405
180	60.57	1.003	72	0.388	60.75	0.999	72	0.391	62.5	0.973	74	0.407
190	60.35	1.004	68	0.390	60.53	1.001	68	0.393	62.1	0.974	70	0.409
200	60.10	1.005	62.5	0.3916	60.31	1.002	62.9	0.3944	62.1	0.975	65.4	0.4106
250	Boiling point 212°F				59.03	1.001	47.8	0.3994	60.6	0.981	50.6	0.4158
300					57.54	1.024	38.4	0.3993	59.5	0.988	41.3	0.4164
350					55.83	1.044	32.1	0.3944	58.1	0.999	35.1	0.4132
400					53.91	1.072	27.6	0.3849	56.5	1.011	30.6	0.4064
500					49.11	1.181	21.6	0.3508	52.9	1.051	24.8	0.3836
600					Boiling point 544.58°F				48.3	1.118	21.0	0.3493

Note: From: "1967 ASME Steam Tables," American Society of Mechanical Engineers, Tables 9, 10, and 11 and Figures 6, 7, 8, and 9; The ASME compilation is a 330-page book of tables and charts, including a $2\frac{1}{2} \times 3\frac{1}{2}$ -ft Mollier chart. All values have been computed in accordance with the 1967 specifications of the International Formulation Committee (IFC) and are in conformity with the 1963 International Skeleton Tables. This standardization of tables began in 1921 and was extended through the International Conferences in London (1929), Berlin (1930), Washington (1934), Philadelphia (1954), London (1956), New York (1963), and Glasgow (1966). Based on these worldwide standard data, the 1967 ASME volume represents detailed computer output in both tabular and graphic form. Included are density and volume, enthalpy, entropy, specific heat, viscosity, thermal conductivity, Prandtl number, isentropic exponent, choking velocity, p–v product, etc., over the entire range (to 1500 psia, 1500°F). English units are used, but all conversion factors are given.

[a] At 7500 psia.

TABLE B.2
Physical and Thermal Properties of Common Liquids

Common Name	Density, kg/m³	Specific Heat, kJ/kg·K	Viscosity, N·s/m²	Thermal Conductivity, W/m·K	Freezing Point, K	Latent Heat of Fusion, kJ/kg	Boiling Point, K	Latent Heat of Evaporation, kJ/kg	Coefficient of Cubical Expansion per K
Part a. SI Units									
(At 1.0 atm pressure (0.101325 MN/m²), 300 K, except as noted)									
Acetic acid	1049	2.18	0.001 155	0.171	290	181	391	402	0.0011
Acetone	784.6	2.15	0.000 316	0.161	179.0	98.3	329	518	0.0015
Alcohol, ethyl	785.1	2.44	0.001095	0.171	158.6	108	351.46	846	0.0011
Alcohol, methyl	786.5	2.54	0.00056	0.202	175.5	98.8	337.8	1100	0.0014
Alcohol, propyl	800.0	2.37	0.00192	0.161	146	86.5	371	779	
Ammonia (aqua)	823.5	4.38		0.353					
Benzene	873.8	1.73	0.000601	0.144	278.68	126	353.3	390	0.0013
Bromine		0.473	0.00095		245.84	66.7	331.6	193	0.0012
Carbon disulfide	1261	0.992	0.00036	0.161	161.2	57.6	319.40	351	0.0013
Carbon tetrachloride	1584	0.866	0.00091	0.104	250.35	174	349.6	194	0.0013
Castor oil	956.1	1.97	0.650	0.180	263.2				
Chloroform	1465	1.05	0.00053	0.118	209.6	77.0	334.4	247	0.0013
Decane	726.3	2.21	0.000859	0.147	243.5	201	447.2	263	
Dodecane	754.6	2.21	0.001374	0.140	247.18	216	489.4	256	
Ether	713.5	2.21	0.000223	0.130	157	96.2	307.7	372	0.0016
Ethylene glycol	1097	2.36	0.0162	0.258	260.2	181	470	800	
Fluorine refrigerant R-11	1476	0.870[a]	0.00042	0.093[a]	162		297.0	180[b]	
Fluorine refrigerant R-12	1311	0.971[a]		0.071[a]	115	34.4	243.4	165[b]	
Fluorine refrigerant R-22	1194	1.26[a]		0.086[a]	113	183	232.4	232[b]	
Glycerine	1259	2.62	0.950	0.287	264.8	200	563.4	974	0.00054
Heptane	679.5	2.24	0.000376	0.128	182.54	140	371.5	318	
Hexane	654.8	2.26	0.000297	0.124	178.0	152	341.84	365	
Iodine		2.15			386.6	62.2	457.5	164	
Kerosene	820.1	2.09	0.00164	0.145				251	
Linseed oil	929.1	1.84	0.0331		253		560		
Mercury		0.139	0.00153		234.3	11.6	630	295	0.00018
Octane	698.6	2.15	0.00051	0.131	216.4	181	398	298	0.00072
Phenol	1072	1.43	0.0080	0.190	316.2	121	455		0.00090
Propane	493.5	2.41[a]	0.00011		85.5	79.9	231.08	428[b]	
Propylene	514.4	2.85	0.00009		87.9	71.4	225.45	342	

(*Continued*)

TABLE B.2 (*Continued*)
Physical and Thermal Properties of Common Liquids

Common Name	Density, kg/m³	Specific Heat, kJ/kg·K	Viscosity, N·s/m²	Thermal Conductivity, W/m·K	Freezing Point, K	Latent Heat of Fusion, kJ/kg	Boiling Point, K	Latent Heat of Evaporation, kJ/kg	Coefficient of Cubical Expansion per K
Propylene glycol	965.3	2.50	0.042		213		460	914	
Seawater	1025	3.76–4.10			270.6				
Toluene	862.3	1.72	0.000550	0.133	178	71.8	383.6	363	
Turpentine	868.2	1.78	0.001375	0.121	214		433	293	0.00099
Water	997.1	4.18	0.00089	0.609	273	333	373	2260	0.00020

[a] At 297 K, liquid.
[b] At 0.101325 MN, saturation temperature.

Common name	Density, $\frac{lb_m}{ft^3}$	Specific Gravity	Viscosity		Sound Velocity, m/s	Dielectric Constant	Refractive Index
			$lb_m/ft\ s \times 10^4$	cp			
Part b. English Units							
(At 1.0 atm pressure, 77°F (25°C), except as noted)							
Acetic acid	65.493	1.049	7.76	1.155	1584[50]	6.15	1.37
Acetone	48.98	0.787	2.12	0.316	1174	20.7	1.36
Alcohol, ethyl	49.01	0.787	7.36	1.095	1144	24.3	1.36
Alcohol, methyl	49.10	0.789	3.76	0.56	1103	32.6	1.33
Alcohol, propyl	49.94	0.802	12.9	1.92	1205	20.1	1.38
Ammonia (aqua)	51.411	0.826	—	—	—	16.9	—
Benzene	54.55	0.876	4.04	0.601	1298	2.2	1.50
Bromine	—	—	6.38	0.95	—	3.20	—
Carbon disulfide	78.72	1.265	2.42	0.36	1149	2.64	1.63
Carbon tetrachloride	98.91	1.59	6.11	0.91	924	2.23	1.46
Castor oil	59.69	0.960	—	650	1474	4.7	—
Chloroform	91.44	1.47	3.56	0.53	995	4.8	1.44
Decane	45.34	0.728	5.77	0.859	—	2.0	1.41
Dodecane	47.11	—	9.23	1.374	—	—	1.41
Ether	44.54	0.715	1.50	0.223	985	4.3	1.35
Ethylene glycol	68.47	1.100	109	16.2	1644	37.7	1.43
Fluorine refrigerant R-11	92.14	1.480	2.82	0.42	—	2.0	1.37

(*Continued*)

Common name	Density, $\frac{lb_m}{ft^3}$	Specific Gravity	Viscosity		Sound Velocity, m/s	Dielectric Constant	Refractive Index
			$lb_m/ft\ s \times 10^4$	cp			
Fluorine refrigerant R-12	81.84	1.315	—	—	—	2.0	1.29
Fluorine refrigerant R-22	74.53	1.197	—	—	—	2.0	1.26
Glycerine	78.62	1.263	6380	950	1909	40	1.47
Heptane	42.42	0.681	2.53	0.376	1138	1.92	1.38
Hexane	40.88	0.657	2.00	0.297	1203	—	1.37
lodine	—	—	—	—	—	11	—
Kerosene	51.2	0.823	11.0	1.64	1320	—	—
Linseed oil	58.0	0.93	222	33.1	—	3.3	—
Mercury	—	13.633	10.3	1.53	1450	—	—
Octane	43.61	0.701	3.43	0.51	1171	—	1.40
Phenol	66.94	1.071	54	8.0	1274[100]	9.8	—
Propane	30.81	0.495	0.74	0.11	—	1.27	1.34
Propylene	32.11	0.516	0.60	0.09	—	—	1.36
Propylene glycol	60.26	0.968	—	42	—	—	1.43
Seawater	64.0	1.03	—	—	1535	—	—
Toluene	53.83	0.865	3.70	0.550	1275[30]	2.4	1.49
Turpentine	54.2	0.87	9.24	1.375	1240	—	1.47
Water	62.247	1.00	6.0	0.89	1498	78.54[a]	1.33

Note: For viscosity in $N \cdot s/m^2$ (=kg m·s), multiply values in centipoises by 0.001.

[a] The dielectric constant of water near the freezing point is 87.8; it decreases with increase in temperature to about 55.6 near the boiling point.

Appendix C: Properties of Solids

TABLE C.1
Properties of Common Solids

Material	Specific Gravity	Specific Heat, $\frac{Btu}{lb_m \cdot {}^{o}R}$	Specific Heat, $\frac{kJ}{kg \cdot K}$	Thermal Conductivity, $\frac{Btu}{h \cdot ft \cdot {}^{o}F}$	Thermal Conductivity, $\frac{W}{m \cdot K}$
Asbestos cement board	1.4	0.2	0.837	0.35	0.607
Asbestos millboard	1.0	0.2	0.837	0.08	0.14
Asphalt	1.1	0.4	1.67		
Beeswax	0.95	0.82	3.43		
Brick, common	1.75	0.22	0.920	0.42	0.71
Brick, hard	2.0	0.24	1.00	0.75	1.3
Chalk	2.0	0.215	0.900	0.48	0.84
Charcoal, wood	0.4	0.24	1.00	0.05	0.088
Coal, anthracite	1.5	0.3	1.26		
Coal, bituminous	1.2	0.33	1.38		
Concrete, light	1.4	0.23	0.962	0.25	0.42
Concrete, stone	2.2	0.18	0.753	1.0	1.7
Corkboard	0.2	0.45	1.88	0.025	0.04
Earth, dry	1.4	0.3	1.26	0.85	1.5
Fiberboard, light	0.24	0.6	2.51	0.035	0.058
Fiber hardboard	1.1	0.5	2.09	0.12	0.2
Firebrick	2.1	0.25	1.05	0.8	1.4
Glass, window	2.5	0.2	0.837	0.55	0.96
Gypsum board	0.8	0.26	1.09	0.1	0.17
Hairfelt	0.1	0.5	2.09	0.03	0.050
Ice (32°)	0.9	0.5	2.09	1.25	2.2
Leather, dry	0.9	0.36	1.51	0.09	0.2
Limestone	2.5	0.217	0.908	1.1	1.9
Magnesia (85%)	0.25	0.2	0.837	0.04	0.071
Marble	2.6	0.21	0.879	1.5	2.6
Mica	2.7	0.12	0.502	0.4	0.71
Mineral wool blanket	0.1	0.2	0.837	0.025	0.04
Paper	0.9	0.33	1.38	0.07	0.1
Paraffin wax	0.9	0.69	2.89	0.15	0.2
Plaster, light	0.7	0.24	1.00	0.15	0.2
Plaster, sand	1.8	0.22	0.920	0.42	0.71
Plastics, foamed	0.2	0.3	1.26	0.02	0.03
Plastics, solid	1.2	0.4	1.67	0.11	0.19
Porcelain	2.5	0.22	0.920	0.9	1.5
Sandstone	2.3	0.22	0.920	1.0	1.7
Sawdust	0.15	0.21	0.879	0.05	0.08
Silica aerogel	0.11	0.2	0.837	0.015	0.02
Vermiculite	0.13	0.2	0.837	0.035	0.058

(*Continued*)

TABLE C.1 (*Continued*)
Properties of Common Solids

Material	Specific Gravity	Specific Heat		Thermal Conductivity	
		$\frac{\text{Btu}}{\text{lb}_m \cdot {}^\circ\text{R}}$	$\frac{\text{kJ}}{\text{kg} \cdot \text{K}}$	$\frac{\text{Btu}}{\text{h} \cdot \text{ft} \cdot {}^\circ\text{F}}$	$\frac{\text{W}}{\text{m} \cdot \text{K}}$
Wood, balsa	0.16	0.7	2.93	0.03	0.050
Wood, oak	0.7	0.5	2.09	0.10	0.17
Wood, white pine	0.5	0.6	2.51	0.07	0.12
Wool, felt	0.3	0.33	1.38	0.04	0.071
Wool, loose	0.1	0.3	1.26	0.02	0.3

Note: Compiled from several sources.

TABLE C.2
Miscellaneous Properties of Metals and Alloys

Common Name	Properties (Typical Only)						
	Thermal Conductivity, Btu/h ft °F	Specific Gravity	Coefficient of Linear Expansion, μ in./in. °F	Electrical Resistivity, microohm-cm	Poisson's Ratio	Modulus of Elasticity, millions of psi	Approximate Melting Point, °F
Part a. Pure Metals At Room Temperature							
Aluminum	137	2.70	14	2.655	0.33	10.0	1220
Antimony	10.7	6.69	5	41.8		11.3	1170
Beryllium	126	1.85	6.7	4.0	0.024–0.030	42	2345
Bismuth	4.9	9.75	7.2	115		4.6	521
Cadmium	54	8.65	17	7.4		8	610
Chromium	52	7.2	3.3	13		36	3380
Cobalt	40	8.9	6.7	9		30	2723
Copper	230	8.96	9.2	1.673	0.36	17	1983
Gold	182	19.32	7.9	2.35	0.42	10.8	1945
Iridium	85.0	22.42	3.3	5.3		75	4440
Iron	46.4	7.87	6.7	9.7		28.5	2797
Lead	20.0	11.35	16	20.6	0.40–0.45	2.0	621
Magnesium	91.9	1.74	14	4.45	0.35	6.4	1200
Manganese		7.21–7.44	12	185		23	2271
Mercury	4.85	13.546		98.4			–38
Molybdenum	81	10.22	3.0	5.2	0.32	40	4750
Nickel	52.0	8.90	7.4	6.85	0.31	31	2647
Niobium (Columbium)	30	8.57	3.9	13		15	4473
Osmium	35	22.57	2.8	9		80	5477
Platinum	42	21.45	5	10.5	0.39	21.3	3220
Plutonium	4.6	19.84	30	141.4	0.15–0.21	14	1180
Potassium	57.8	0.86	46	7.01			146

(*Continued*)

TABLE C.2 (*Continued*)
Miscellaneous Properties of Metals and Alloys

Common Name	Properties (Typical Only)						
	Thermal Conductivity, Btu/h ft °F	Specific Gravity	Coefficient of Linear Expansion, μ in./in. °F	Electrical Resistivity, microohm-cm	Poisson's Ratio	Modulus of Elasticity, millions of psi	Approximate Melting Point, °F
Rhodium	86.7	12.41	4.4	4.6		42	3569
Selenium	0.3	4.8	21	12.0		8.4	423
Silicon	48.3	2.33	2.8	1×10^5		16	2572
Silver	247	10.50	11	1.59	0.37	10.5	1760
Sodium	77.5	0.97	39	4.2			208
Tantalum	31	16.6	3.6	12.4	0.35	27	5400
Thorium	24	11.7	6.7	18	0.27	8.5	3180
Tin	37	7.31	11	11.0	0.33	6	450
Titanium	12	4.54	4.7	43	0.3	16	3040
Tungsten	103	19.3	2.5	5.65	0.28	50	6150
Uranium	14	18.8	7.4	30	0.21	24	2070
Vanadium	35	6.1	4.4	25		19	3450
Zinc	66.5	7	19	5.92	0.25	12	787

Part b. Commercial Metals and Alloys

Material No. (from Table 1–57)	Common Name and Classification	Properties (Typical Only)					
		Thermal Conductivity, Btu/h ft °F	Specific Gravity	Coefficient of Linear Expansion, μ in./in. °F	Electrical Resistivity, microhm-cm	Modulus of Elasticity, millions of psi	Approximate Melting Point, °F
Classification and Designation							
1	Ingot iron (included for comparison)	42	7.86	6.8	9	30	2800
2	Plain carbon steel AISI-SAE 1020	30	7.86	6.7	10	30	2760
15	Stainless-steel type 304	10	8.02	9.6	72	28	2600
19	Cast gray iron ASTM A48–48, Class 25	26	7.2	6.7	67	13	2150
21	Malleable iron ASTM A47	—	7.32	6.6	30	25	2250
22	Ductile cast iron ASTM A339, A395	19	7.2	7.5	60	25	2100
24	Ni-resist cast iron, type 2	23	7.3	9.6	170	15.6	2250
29	Cast 28–7 alloy (HD) ASTM A297–63T	1.5	7.6	9.2	41	27	2700
31	Hastelloy C	5	3.94	6.3	139	30	2350
36	Inconel X, annealed	9	8.25	6.7	122	31	2550
41	Haynes Stellite alloy 25 (L605)	5.5	9.15	7.61	88	34	2500

(*Continued*)

Classification and Designation		Properties (Typical Only)					
Material No. (from Table 1–57)	Common Name and Classification	Thermal Conductivity, Btu/h ft °F	Specific Gravity	Coefficient of Linear Expansion, μ in./in. °F	Electrical Resistivity, microhm-cm	Modulus of Elasticity, millions of psi	Approximate Melting Point, °F
43	Aluminum alloy 3003, rolled ASTM B221	90	2.73	12.9	4	10	1200
44	Aluminum alloy 2017, annealed ASTM B221	95	2.8	12.7	4	10.5	1185
49	Aluminum alloy 380 ASTM SC84B	56	2.7	11.6	7.5	10.3	1050
53	Copper ASTM B152, B124, B133, B1, B2, B3	225	8.91	9.3	1.7	17	1980
57	Yellow brass (high brass) ASTM B36, B134, B135	69	8.47	10.5	7	15	1710
61	Aluminum bronze ASTM B169, alloy A; ASTM B124, B150	41	7.8	9.2	12	17	1900
62	Beryllium copper 25 ASTM B194	7	8.25	9.3	—	19	1700
64	Nickel silver 18% alloy A (wrought) ASTM B122, No. 2	19	8.8	9.0	29	18	2030
67	Cupronickel 30%	17	8.95	8.5	35	22	2240
68	Red brass (cast) ASTM B30, No. 4A	42	8.7	10	11	13	1825
74	Chemical lead	20	11.35	16.4	21	2	621
75	Antimonial lead (hard lead)	17	10.9	15.1	23	3	554
80	Solder 50–50	26	8.89	13.1	15	—	420
82	Magnesium alloy AZ31B	45	1.77	14.5	9	6.5	1160
89	K Monel	11	8.47	7.4	58	26	2430
90	Nickel ASTM B160, B161, B162	35	8.89	6.6	10	30	2625
92	Cupronickel 55–45 (Constantan)	13	8.9	8.1	49	24	2300
95	Commercial titanium	10	5	4.9	80	16.5	3300
99	Zinc ASTM B69	62	7.14	18	6	—	785
103	Zirconium, commercial	10	6.5	2.9	41	12	3,350

Note: Compiled from several sources.

Appendix D: SI Units and Conversion Factors

GREEK ALPHABET

Greek Letter		Greek Name	English Equivalent	Greek Letter		Greek Name	English Equivalent
A	α	Alpha	a	N	ν	Nu	n
B	β	Beta	b	Ξ	ξ	Xi	x
Γ	γ	Gamma	g	O	o	Omicron	o
Δ	δ	Delta	d	Π	π	Pi	p
E	ε	Epsilon	e	P	ρ	Rho	r
Z	ζ	Zeta	z	Σ	σ ς	Sigma	s
H	η	Eta	e	T	τ	Tau	t
Θ	θ	Theta	th	Υ	υ	Upsilon	u
I	ι	Iota	i	Φ	ϕ φ	Phi	ph
K	κ	Kappa	k	X	χ	Chi	ch
Λ	λ	Lambda	ı	Ψ	ψ	Psi	ps
M	μ	Mu	m	Ω	ω	Omega	o

INTERNATIONAL SYSTEM OF UNITS (SI)

The International System of Units (SI) was adopted by the 11th General Conference on Weights and Measures (CGPM) in 1960. It is a coherent system of units built from seven *SI base units*, one for each of the seven dimensionally independent base quantities: the meter, kilogram, second, ampere, kelvin, mole, and candela, for the dimensions of length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity, respectively. The definitions of the SI base units are given below. The *SI derived units* are expressed as products of powers of the base units, analogous to the corresponding relations between physical quantities but with numerical factors equal to unity.

In the International System, there is only one SI unit for each physical quantity. This is either the appropriate SI base unit itself or the appropriate SI derived unit. However, any of the approved decimal prefixes, called *SI prefixes*, may be used to construct decimal multiples or submultiples of SI units.

It is recommended that only SI units be used in science and technology (with SI prefixes where appropriate). Where there are special reasons for making an exception to this rule, it is recommended always to define the units used in terms of SI units. This section is based on information supplied by IUPAC.

DEFINITIONS OF SI BASE UNITS

Meter: The meter is the length of path traveled by light in vacuum during a time interval of 1/299,792,458 of a second (17th CGPM, 1983).

Kilogram: The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram (3rd CGPM, 1901).

Second: The second is the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom (13th CGPM, 1967).

Ampere: The ampere is that constant current, which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross section and placed 1 m apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} N/m of length (9th CGPM, 1958).

Kelvin: The kelvin, unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water (13th CGPM, 1967).

Mole: The mole is the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kg of carbon-12. When the mole is used, the elementary entities must be specified and may be atoms, molecules, ions, electrons, or other particles, or specified groups of such particles (14th CGPM, 1971). Examples of the use of the mole are as follows:

- 1 mol of H_2 contains about 6.022×10^{23} H_2 molecules, or 12.044×10^{23} H atoms.
- 1 mol of HgCl has a mass of 236.04 g.
- 1 mol of Hg_2Cl_2 has a mass of 472.08 g.
- 1 mol of Hg_2^{2+} has a mass of 401.18 g and a charge of 192.97 kC.
- 1 mol of $Fe_{0.91}$ S has a mass of 82.88 g.
- 1 mol of e^- has a mass of 548.60 μg and a charge of –96.49 kC.
- 1 mol of photons whose frequency is 10^{14} Hz has energy of about 39.90 kJ.

Candela: The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} Hz, and that has a radiant intensity in that direction of (1/683) watt per steradian (16th CGPM, 1979).

NAMES AND SYMBOLS FOR THE SI BASE UNITS

Physical Quantity	Name of SI Unit	Symbol for SI Unit
Length	meter	m
Mass	kilogram	kg
Time	second	sec
Electric current	ampere	A
Thermodynamic temperature	kelvin	K
Amount of substance	mole	mol
Luminous intensity	candela	cd

SI DERIVED UNITS WITH SPECIAL NAMES AND SYMBOLS

Physical Quantity	Name of SI Unit	Symbol for SI Unit	Expression in Terms of SI Base Units
Frequency[a]	hertz	Hz	s^{-1}
Force	newton	N	$m \cdot kg \cdot s^{-2}$
Pressure, stress	pascal	Pa	$N \cdot m^{-2} = m^{-1} \cdot kg \cdot s^{-2}$
Energy, work, heat	joule	J	$N \cdot m = m^2 \cdot kg \cdot s^{-2}$
Power, radiant flux	watt	W	$J \cdot s^{-1} = m^2 \cdot kg \cdot s^{-3}$
Electric charge	Coulomb	C	$A \cdot s$
Electric potential, electromotive force	Volt	V	$J \cdot C^{-1} = m^2\ kg \cdot s^{-3}\ A^{-1}$
Electric resistance	Ohm	Ω	$V \cdot A^{-1} = m^2 \cdot kg \cdot s^{-3} \cdot A^{-2}$
Electric conductance	Siemens	S	$\Omega^{-1} = m^{-2} \cdot kg^{-1} \cdot s^4 \cdot A^2$
Electric capacitance	Farad	F	$C \cdot V^{-1} = m^{-2} \cdot kg^{-1} \cdot s^4 \cdot A^2$
Magnetic flux density	Tesla	T	$V \cdot s \cdot m^{-2} = kg \cdot s^{-2} \cdot A^{-1}$
Magnetic flux	Weber	Wb	$V \cdot s = m^2 \cdot kg \cdot s^{-2} \cdot A^{-1}$
Inductance	Henry	H	$V \cdot A^{-1} \cdot s = m^2 \cdot kg \cdot s^{-2} \cdot A^{-2}$
Celsius temperature[b]	degree Celsius	°C	K
Luminous flux	lumen	lm	$cd \cdot sr$
Illuminance	lux	lx	$cd \cdot sr \cdot m^{-2}$
Activity (radioactive)	Becquerel	Bq	s^{-1}
Absorbed dose (or radiation)	Gray	Gy	$J \cdot kg^{-1} = m^2 \cdot s^{-2}$
Dose equivalent (dose equivalent index)	Sievert	Sv	$J \cdot kg^{-1} = m^2 \cdot s^{-2}$
Plane angle	radian	rad	$1 = m \cdot m^{-1}$
Solid angle	steradian	sr	$1 = m^2 \cdot m^{-2}$

[a] For radial (circular) frequency and for angular velocity, the unit rad s^{-1}, or simply s^{-1}, should be used, and this may not be simplified to Hz. The unit Hz should be used only for frequency in the sense of cycles per second.

[b] The Celsius temperature θ is defined by the equation

$$q/°C = T/K = 237.15$$

The SI unit of Celsius temperature interval is the degree Celsius, °C, which is equal to the kelvin, K. °C should be treated as a single symbol, with no space between the ° sign and the letter C. (The symbol °K, and the symbol °, should no longer be used.)

UNITS IN USE TOGETHER WITH THE SI

These units are not part of the SI, but it is recognized that they will continue to be used in appropriate contexts. SI prefixes may be attached to some of these units, such as milliliter, ml; millibar, mbar; mega-electronvolt, MeV; and kilotonne, kt.

Physical Quantity	Name of Unit	Symbol for Unit	Value in SI Units
Time	minute	min	60 s
Time	hour	h	3600 s
Time	day	d	86,400 s
Plane angle	degree	°	(π/180) rad
Plane angle	minute	′	(π/10,800) rad
Plane angle	second	″	(π/648,000) rad
Length	angstrom[a]	A	10^{-10} m
Area	barn	b	10^{-28} m^2
Volume	liter	l, L	dm^3 = 10^{-3} m^3
Mass	ton	t	Mg = 10^3 kg
Pressure	bar[a]	bar	10^5 Pa = 10^5 N m^2
Energy	electronvolt[b]	eV (= e × V)	≈ 1.60218×10^{-19} J
Mass	unified atomic mass unit[b,c]	u (= m_a(12C)/12)	≈ 1.66054×10^{-27} kg

[a] The angstrom and the bar are approved by CIPM for "temporary use with SI units," until CIPM makes a further recommendation. However, they should not be introduced where they are not used at present.

[b] The values of these units in terms of the corresponding SI units are not exact, since they depend on the values of the physical constants e (for the electronvolt) and N_A (for the unified atomic mass unit), which are determined by experiment.

[c] The unified atomic mass unit is also sometimes called the dalton, with symbol Da, although the name and symbol have not been approved by CGPM.

CONVERSION CONSTANTS AND MULTIPLIERS

Recommended Decimal Multiples and Submultiples

Multiple or Submultiple	Prefix	Symbol	Multiple or Submultiple	Prefix	Symbol
10^{18}	exa	E	10^{-1}	deci	d
10^{15}	peta	P	10^{-2}	centi	c
10^{12}	tera	T	10^{-3}	milli	m
10^{9}	giga	G	10^{-6}	micro	μ (Greek mu)
10^{6}	mega	M	10^{-9}	nano	n
10^{3}	kilo	k	10^{-12}	pico	p
10^{2}	hecto	h	10^{-15}	femto	f
10	deca	da	10^{-18}	atto	a

Conversion Factors: Metric to English

To Obtain	Multiply	By
Inches	Centimeters	0.3937007874
Feet	Meters	3.280839895
Yards	Meters	1.093613298
Miles	Kilometers	0.6213711922
Ounces	Grams	$3.527396195 \times 10^{-2}$
Pounds	Kilograms	2.204622622
Gallons (U.S. liquid)	Liters	0.2641720524
Fluid ounces	Milliliters (cc)	$3.381402270 \times 10^{-2}$
Square inches	Square centimeters	0.1550003100
Square feet	Square meters	10.76391042
Square yards	Square meters	1.195990046
Cubic inches	Milliliters (cc)	$6.102374\ 409 \times 10^{-2}$
Cubic feet	Cubic meters	35.31466672
Cubic yards	Cubic meters	1.307950619

Conversion Factors: English to Metric

To Obtain	Multiply	By
Microns	Mils	25.4
Centimeters	Inches	2.54
Meters	Feet	0.3048
Meters	Yards	0.9144
Kilometers	Miles	1.609344
Grams	Ounces	28.34952313
Kilograms	Pounds	0.453 592 37
Liters	Gallons (U.S. liquid)	3.785411784
Millileters (cc)	Fluid ounces	29.57352956
Square centimeters	Square inches	6.4516
Square meters	Square feet	0.09290304
Square meters	Square yards	0.83612736
Milliliters (cc)	Cubic inches	16.387064
Cubic meters	Cubic feet	$2.831684659 \times 10^{-2}$
Cubic meters	Cubic yards	0.764554858

Conversion Factors: General

To Obtain	Multiply	By
Atmospheres	Feet of water @ 4°C	2.950×10^{-2}
Atmospheres	Inches of mercury @ 0°C	3.342×10^{-2}
Atmospheres	Pounds per square inch	6.804×10^{-2}
Btu	Foot-pounds	1.285×10^{-3}
Btu	Joules	9.480×10^{-4}
Cubic feet	Cords	**128**
Degree (angle)	Radians	57.2958
Ergs	Foot-pounds	1.356×10^{-7}
Feet	Miles	**5280**
Feet of water @ 4°C	Atmospheres	33.90
Foot-pounds	Horsepower-hours	1.98×10^{6}
Foot-pounds	Kilowatt-hours	2.655×10^{6}
Foot-pounds per minute	Horsepower	3.3×10^{4}
Horsepower	Foot-pounds per second	1.818×10^{-3}
Inches of mercury @ 0°C	Pounds per square inch	2.036
Joules	Btu	1054.8
Joules	Foot-pounds	1.355 82
Kilowatts	Btu per minute	1.758×10^{-2}
Kilowatts	Foot-pounds per minute	2.26×10^{-5}
Kilowatts	Horsepower	0.745712
Knots	Miles per hour	0.868 976 24
Miles	Feet	1.894×10^{-4}
Nautical miles	Miles	0.868 976 24
Radians	Degrees	1.745×10^{-2}
Square feet	Acres	**43,560**
Watts	Btu per minute	17.5796

[a] Boldface numbers are exact; others are given to 10 significant figures where so indicated by the multiplier factor.

Temperature Factors

$°F = 9/5(°C) + 32$

Fahrenheit temperature = 1.8(temperature in kelvins) − 459.67

$°C = 5/9[(°F) - 32]$

Celsius temperature = temperature in kelvins − 273.15

Fahrenheit temperature = 1.8(Celsius temperature) + 32

Conversion of Temperatures

From	To		From	To	
Fahrenheit	Celsius	$t_c = \frac{t_F - 32}{1.8}$	Celsius	Fahrenheit	$t_F = (t_c \times 1.8) + 32$
				Kelvin	$T_K = (t_c + 273.15)$
	Kelvin	$T_k = \frac{t_F - 32}{1.8} + 273.15$	Kelvin	Rankine	$T_R = (t_c + 273.15) \times 1.8$
				Celsius	$t_c = T_K - 273.15$
				Rankine	$T_R = T_k \times 1.8$
	Rankine	$T_R = t_F + 459.67$	Rankine	Fahrenheit	$t_F = T_R - 459.67$ F
				Kelvin	$T_K = \frac{T_R}{1.8}$

Index

Page numbers followed by "*f*" indicate figures; those followed by "*t*" indicate tables.

A

Ab initio calculations, 1057
ABPE, *see* Applied bias photon to current efficiency
ABS resin, *see* Acrylnitrite-butadiene-styrene resin
Absolute entropy, 59*t*, 60–62
Absolute humidity, 1307
Absolute moisture, 1309
Absolute pressure, 116, 624
Absolute viscosity, 130, 1018
Absorbed photon to current efficiency (APCE), 1377
Absorptance, 366
Absorption
  bands, 341
  coefficient, 366
  cycle, 922–923
Absorptivity, 341, 366
Acceleration transforms, 1116–1118, 1117, 1118*f*
Acid food product, 1017
Acid sulfite pulping, 1277
Aciduric microorganisms, 1022
Acoustic boundary conditions, 1503
Acoustic mismatch model (AMM), 1051
Acoustic phonons, 1046
Acrylnitrite–butadiene–styrene (ABS) resin, 1106–1108
Activation/threshold function, 1222
Active enhancement techniques, 947
Active heat enhancement techniques, 933
Active solar still, 1389, 1391–1393
  ETC-integrated solar still, 1429, 1429*f*
  FPC-integrated solar still, 1426–1427
  PVT-CPC water collector solar still, 1427–1429
Actual net positive suction head (NPSHA), 862
Actual volumetric efficiency, 919
Actuators, 1154
Acute fouling, 606
Adaptive mesh refinement (AMR) technique, 999
Additive pressure rule, 36
Additives
  for gases, 935
  for liquids, 935, 946
Additive volume rule, 36, 55
Adhesive films, 1103
Adhesive tapes, 1105
Adiabatic combustion, 60
Adiabatic mixing
  of pure components, 55
  of two streams, 904–907
Adiabatic vaporization, 442
Adsorption, 1280
Advection–diffusion equation, 1510
Advection scheme, 1458
Advection term, 1462–1463, 1464*f*, 1513
AdvFacSy program, 771, 773
Aerobes, 1022
Aerosol science, 426
AFD, *see* Atmospheric freeze-drying
AFPs, *see* Antifreeze proteins
AFT module, *see* Air flow-through module
Agarbatti (incense stick), 1155
Ageing, 610
Agglomerate, 1318
Agglomerations, 547, 550, 569, 570*f*
AI, *see* Artificial intelligence
Air
  cavity, 174
  changes per hour (ACH), 912
  compressor, 98, 99
    compressor types for various applications, 844–846
    maintenance requirements, 848
    performance considerations, 846–848
    reduction of energy use, 850
    thermal issues, 848–850
  conditioning, 901
  dehumidifying, 902
  distribution system heat losses, 914
  –fuel ratio, 57
  humidifying, 902
  leakage, 733
  saturated, 49
  -standard analysis, 78
  velocity, 908
Air-blast atomizer
  prefilming, 989*f*
Airborne electronics, high-altitude, 1121–1124
Air compressors, 844*f*
Air-conditioning systems
  energy distribution systems, 928
    air-and-water systems, 930–931
    all-air systems, 928
    all-water systems, 931
  load calculations, 908–909
    air distribution system heat losses, 914
    auxiliary heat gains, 914–915
    cooling load, 909–912
    heating load, 912–915
  psychrometric chart for, 901–902, 907*f*
  refrigeration, 915
    absorption cycle, 922–923
    Carnot refrigeration cycle, 915–916
    compressors, 918–922
    refrigerant selection, 923–924
    vapor compression cycle, 916–918
  schematic diagram of air-conditioned space, 906*f*
  thermal comfort conditions, 907–908
  thermodynamic properties
    of moist air, 899–901
    of saturated water, 898
Air-cooled heat exchangers, 486*f*
  arrangement, 488–489

Air-cooled heat exchangers (*cont.*)
basic equations for design of, 491–496
design envelope, 496–498
equipment configurations, 485–491
fans for, 490–491
forced-draft and induced-draft arrangements of, 489
Aircraft airfoil design, 842
Aircraft propulsion, 841
Air density, 1121*f*
Airends, 847
Air flow-through (AFT) module, 1122, 1122*f*
Airfoils, 188
Air leakage, 733
Air–water vapor mixtures, properties of
dry-bulb temperature, 1307
enthalpy, 1307–1308
humidity, 1307
psychrometric chart, 1308
psychrometry, 1306
relative humidity, 1307
total pressure, 1306
wet-bulb temperature, 1307
ALE method, *see* Arbitrary Lagrangian–Eulerian method
Algae, 479
Algebraic VoF procedure, 999
Aliasing error, in large eddy simulation (LES), 1553
All-air systems, 928–930
dual-duct systems, 930
constant volume, 930
variable air volume, 930
single-duct systems, 928
constant volume, 929
variable air volume, 929
Alumina–copper/water hybrid nanofluids, viscosity of, 564*f*
Alumina nanofluid, 558*f*
Ambient pressure, 847
American Society for Testing and Materials (ASTM), 722
American Society of Mechanical Engineering (ASME), 23
American Water Works Association (AWWA), 168
AMM, *see* Acoustic mismatch model
AMR technique, *see* Adaptive mesh refinement technique
Anaerobes, 1022
Analytical approximations, 401
Analytical efficiency, 768, 769*f*
Anemometry
constant current, 1202
hot-wire, 1202–1205
laser Doppler, 1206–1214
Angle of repose, 1340
ANNs, *see* Artificial neural networks
Annual energy costs, VRC impact on, 756–758, 759–760*f*
ANSYS software, 814, 1494
AnTherm tool, 814
Antifoulant, 611
Antifouling exchanger, 611, 612*t*
Antifreeze proteins (AFPs), 692
Antifriction bearings, 832
Antigravity reverse thermosyphon, 973
Anti-thixotropy, 226
APCE, *see* Absorbed photon to current efficiency
Apoptosis, 667
Apparent heat capacity method, 412
Apparent viscosity, 227
Applied bias photon to current efficiency (ABPE), 1376–1377
Arab alchemists, 1390
Arbitrary Lagrangian–Eulerian (ALE) method, 1507
Arbitrary thin airfoil, 207*f*
Archimedes number, 1137
Arrhenius model, 659–662, 669*f*, 670*f*, 676*f*
functional behavior of the first-order, 663
Artificial intelligence (AI)
artificial neural networks (ANNs)
additional applications in thermal engineering, 1243–1245
application to compact heat exchangers, 1225–1235
general discussion, 1245–1246
methodology, 1221–1225
thermal system dynamics and control, 1235–1243
genetic algorithms, 1248–1262
additional applications in thermal engineering, 1254–1255
applications to compact heat exchangers, 1250–1254
general discussion, 1262
methodology, 1248–1250
genetic programming, 1255–1257
additional applications in thermal engineering, 1260–1262
heat exchangers, 1219–1220
reason for, 1218
and scope, 1217–1218
Artificial neural networks (ANNs), 1217, 1237–1238*f*
fully connected multilayer, 1221*f*
and genetic algorithms, in thermal engineering, 1217–1263
heat transfer rates, 1227*t*
with internal model control, 1238–1239
nodal operation in, 1222*f*
synaptic weights and biases for heat exchanger 1, 1246*t*
Aseptic processing sterilization, 1033–1034
ASHRAE
energy performance standard, 814, 818, 901, 908, 909, 911, 914
Handbook of Fundamentals, 855, 898, 918, 1190
profiles considered in the analysis considered by Krarti et al. (1999), 737
standard 61989, 737
ASHRAE Zone Method, 811
ASME, *see* American Society of Mechanical Engineering
ASME Boiler and Pressure Vessel Code, 473
Asphaltene precipitation model, 618
Asphaltenes, 616
ASTM, *see* American Society for Testing and Materials
Asymptotic fouling, 606, 607*f*, 616
Atmospheric freeze-drying (AFD), 1328
Atomization of feed, spray drying process, 1315–1316
Atomizer(s)
design considerations, 988
effervescent, 993, 994, 994*f*
electrostatic, 994
pressure swirl, 991
rotary, 991
sprays, 988*f*
Augmentation, 933
Automobile turbochargers, 837

Autophagy, 667, 668
Auxiliary heating distillation, 1390
Avogadro's number, 660
Avoidable endogenous variable, 107
Avoidable exogenous variable, 107
Avoidance rule, 815
AWWA, *see* American Water Works Association
Axial flow compressor
  basic description, 841
  cross-sectional drawing of, 841*f*
  design characteristics, 841–842
  flow and performance characteristics, 842–843
  typical performance map of, 842*f*
  velocity vectors for, 843*f*
Axially split casings, 855
Axial velocity of solid particles, 1340–1341
Axisymmetric synthetic jets, 1154

B

Backward leaning vanes, 840
Backwater curve, 176
Back work ratio, 77
Badges, temperature-sensitive, 1173–1174
Baffles, 1029, 1321, 1420
Baffle spacing, 473
Ballistic-diffusive transport equations (BDEs), 1058
Band bending, in semiconductor–electrolyte interface, 1372–1373, 1374
Barnacles, 479
Barometric pressure, 899
Base case, 765
Basin liner, 1387
Basin-type solar still, 1392
Batch chamber contact dryers, 1323
Batchelor equation, 551
Batch retort, 1025–1026, 1033
BCs, *see* Boundary conditions
BDEs, *see* Ballistic-diffusive transport equations
BE distribution, *see* Bose–Einstein distribution
Belt drive, 861
Belt dryers, 1313, 1313–1314*f*
Bend tempering, 1360
Bernoulli equation, 134, 135, 290, 1199
Bessel function
  common definition of, 258
  modified, 257
BGK approximation, *see* Bhatnagar-Gross approximation
Bhatnagar-Gross (BGK) approximation, 1501
Bilinear interpolation, 1511*f*, 1515
Bimolecular process, 657*f*
Binary fission, 1021
Bingham number, 234–235, 331
Bingham plastics, 229, 239, 331, 333*t*, 336*t*, 1020
Biochemical model, 668*f*, 669*f*
Biofilm formation, 615
Biofouling, 479, 610
Bioheat equation, 642, 643
Bioheat transfer, 629–692
  coupling of temperature history to rate processes, 629–630
  effect of blood flow on temperature, 641–643
    continuum models, 644–645
    hybrid models, 648
    limitations of Pennes' model, 643
    thermal measurements of perfusion, 648–651
    vasculature-based models, 645–648
  human thermoregulation, 651–656
    physiological processes of thermoregulation, 652
    thermoregulatory processes, 652–653
    Wissler model of human thermoregulation, 653–656
  therapeutic heating, 656
    heat generation modalities, 656–658
    physiologic effects of local heating, 658
  tissue effects, elevated temperatures, 658–659
    application of kinetic formulations in thermal damage studies, 662–673
    damage rate process coefficients, 680–683
    histologic markers of thermal damage, 673–680
    theory of rate process descriptions, 659–662
  tissue effects, subzero temperatures, 684
    cryopreservation, 684–692
    cryosurgery, 692
  tissue thermal transport properties, 630–631
    background, 631
    estimation of heat transfer coefficient, 635–637
    measurement of thermal conductivity and diffusivity, 631–635
    temperature-dependent thermal properties, 637–641
Biological fouling, 536
Biological neural network, 1220
Biot number, 258, 262, 431, 1183
Birefringence
  loss, in myocardium, 676
  partial loss of, 675–676
  thermally induced loss of, 675–676
Birmingham wire gauge (BWG), 473
BISCO tool, 814
Blackbody, 366
  radiation, 341
Black, Joseph, 393
Blanching, 1031
Blast loading, on deformable human eye, 1519, 1520*f*
Blebs, 667
Bloch–Boltzmann–Peierls formula, 1053
Blood
  flow
    distribution, 653
    effect on temperature, 641–643
  perfusion
    data, for specific tissues, 691
    rates, 691
Blowing factor, 445
Blowoff operations, 850
Bluff bodies, 186
Body-fitted grid, 1443
Body intercept, 1510
Boiling, 1140–1142, 1524
  internal convective, 376
  nucleate, 370
  pool, 370–375
Boiling limitation, of heat pipe, 969
Boltzmann equation, 1499
Boltzmann transport equation (BTE), 1040
Bond number, 961, 1142
Boosting reflector, 1423, 1424*f*

Bose–Einstein (BE) distribution, 1040, 1041*f*, 1047
Bottom loss heat coefficient, 1396, 1398
Bounce-back scheme, 1502
Boundary conditions (BCs), 1443, 1448, 1493, 1507, 1512
  in lattice Boltzmann method, 1502
Boundary grid points, 1454
Boundary layer
  concept, 179–180
  control, 188–189
  laminar, 285, 290–291, 447–448
  temperature, 284
  thickness, 180
  turbulent, 286, 291, 449
Boundary scattering, 1055
Bound water, 1280–1281
Boussinesq approximation, 1530
Brayton cycle, 76
Breakthrough rule, 815
Breathing wall, *see* Single-layer breathing wall
Broad-crested weir, 173
Brownian agitation, 575
Brownian motion, 426, 545, 579
BSim numerical tool, 791
BTE, *see* Boltzmann transport equation
Bubbles, 1359–1360
Bubbling
  fluidized beds, 388
  regime, 388
Buck, W.E., 957
Building(s)
  characteristics, 740
  construction materials, 740
  energy audit for, 728–747
  energy retrofits, 728
  energy use, baseline for, 731
  envelope, 731–734
Building envelope systems, 739
  dynamic insulation materials, 750*f*, 778
    background, 749–750
    energy performance, 750–753
  multilayer living walls
    background, 768–770
    thermal models for advanced façade systems, 770–776
  single-layer breathing wall, 760–762
    background, 762–763
    measured and predicted thermal performance, 764–768
    products and technologies, 763–764
  utilization of PCMs in, 785–787
  variable reflectivity cool roofs (VRCS), 778
    annual energy costs impact, 756–758
    background, 753
    cooling thermal loads impact, 754–756
    cost–benefit analysis of, 758–760
    heating thermal loads impact, 754
    thermal analysis, 753–754
Building geometry, thermal bridges generated by, 803
Building materials, in thermal bridges, 803–804
Building performance standards, dealing with building thermal bridging, 814–815
Building space-conditioning systems, utilization of PCMs in, 789
Bulk density, 1018
Buoyancy, 119, 545, 1142
BWG, *see* Birmingham wire gauge

C

CAGI, *see* Compressed Air and Gas Institute
Calcein leakage, 681–682
Calendering, 1276
Calibration strip, for liquid crystals, 1174*f*
Canine arterial tissue, temperature dependence of, 639
Canine carotid artery, 639
Canine femoral artery, 639
Capacitance sensor, 1212–1213
Capacity modulation, 826
Capillary limitation, of heat pipe, 965, 968–969
Capillary porous material, 1308
Capillary viscometry, 1212
Capital-energy trade-off, 886, 887*f*
Capital investment, 102
Capitalized-cost factor, 103*t*
Capital recovery factor, 1412
Carbon-constrained energy planning, 895
Carbon nanotube (CNT) particle, 563
Cardboard, SEM of, 1281, 1282*f*
Car garages, ventilation for, 737
Carman–Kozeny equation, 1530
Carnot corollaries, 10
Carnot cycle, 82, 915–916
Carnot efficiency, 11
Carnot refrigeration cycle, 915–916, 916*f*
Cartesian coordinates, 1483
Cartesian grid, 1507, 1509, 1512, 1513
  generation, 1449*f*, 1455*f*
  methods, 1443
Cascading regime, 1294
Cascading water retort, 1026
Casimir limit, 1066
Casing orientation, in centrifugal pump, 855
Caspase enzymes, 667
Catalyst layer, 432*f*
Cauchy's momentum equation, 232
CAV systems, *see* Constant air volume systems
Cavitation, 857, 862
  best remedy for, 162
  damage, 161–162
CBM, *see* Conduction band minimum
Cell(s)
  biophysical basis of freezing processes in, 685–687
  death, 658
  in finite volume method, 1454
  membrane permeability, influence of, 687–689
  membrane transport, major challenge in applying models for, 687
Cell centers grid points, 1455
Cell death processes, 662, 667–670
Cellular phones, 1111
Cellular shrinkage, 675–676
Cellulose, 1277, 1277*f*
Cellulose microfibrils, 1277
Cellulosic fibers, during paper making, 1277–1281
Cellulosic fines, 1271
Center of pressure, 117
Central-difference scheme, 1513

Centrifugal compressors, 836*f*, 837*f*, 920–921, 921*f*
basic description, 836–837
design characteristics, 837–838
development directions, 838–839
performance at constant speed, 840*f*
velocity vectors for, 839*f*
Ceramics, 344
CFD, *see* Computational fluid dynamics
CFD-ACE+ software, 1558
CFX software, 1558
Chamber dryers, 1312, 1312*f*
Chapman–Enskog procedure, 1499
Charge transfer processes, at semiconductor–electrolyte interface, 1373–1375
Check valves, 164
Chemical cryoprotective agent (CPA), 684
Chemical exergy, 69–71
Chemical kinetics, 421
Chemical potential, 54–56
Chemical pulping, 1277
Chemical reaction fouling, 536, 609
Chen–Holmes model, 644–645
CHF, *see* Critical heat flux
Chilton–Colburn equation, 1288
Chronic fouling, 606–607
Churchill–Usagi equation (CUE), 147
applied serially, 149
prediction of, 147
CIP, *see* Cleaning-in-place
Circular foil gauges, 1179
Circulation loops, 768, 769*f*
Clapeyron equation, 31
Clausius–Clapeyron equation, 1299
Clausius inequality, 11–12
Cleaning, and fouling, 621, 625
Cleaning-in-place (CIP), 614, 615
Clearance volume, 825
Closed-loop control, 1242–1243
Closed system
energy balance, 7
entropy balance, 13
Clostridium botulinum spores, 1017–1018, 1024–1025
CMOS, *see* Complementary metal-oxide-semiconductor
CMUs, *see* Concrete masonry units
CNT particle, *see* Carbon nanotube particle
Coal, pulverized, 361–362
Coanda effect, 1292
Coating operations, paper making, 1276
Coaxial tubes, 311
Coefficient matrix, 1452
Coefficient of performance (COP), 915
Cogeneration, 739–740
building and utility data analysis
building characteristics, 740
energy use, 740
energy use baseline model, 743–744
evaluation of energy conservation opportunities, 744–746
on-site survey, 741–743
Colburn *j*-factor, 936, 939, 1257
Colebrook equation, 152
Collagen coagulation, 676
Collagen shrinkage, 677–678
Collision-dominated plasma, 1011
Collision step, 1496
Colloidal–capillary porous material, 1308
Colloidal material, 1308
Column condenser/reboiler, 893
Combined thermal bridges, 806
Combustion, 56–66
adiabatic, 60
droplet, 436–437
enthalpy of, 59–60, 64
heterogeneous, 434
property data for reactive systems, 58–65
absolute entropy, 60–62
enthalpy of formation, 58–61
Gibbs function of formation, 63–65
reaction equations, 56–58
reaction equilibrium, 64–66
Comfort zones, 908
Communication devices, handheld, 1111–1118
Commutation error, 1553
Compact heat exchangers, 461, 501–540, 1225*f*
design methodology, 527–535
flow maldistribution, 535
fouling in
accounting in heat exchangers, 537
effect and mechanisms, 535–536
fouling control and cleaning techniques, 540
importance of, 536–537
influence of operating and design variables, 537–540
heat transfer
and flow friction correlations, 522–527
and pressure drop analysis, 509–522
rating/sizing problem for crossflow plate-fin exchanger, 529–535
types and description
gas-to-fluid exchangers, 501
liquid-to-liquid exchangers, 502
plate-fin heat exchangers, 502–503
plate-type heat exchangers, 506–508
regenerators, 504–506
tube-fin heat exchangers, 503–504
Compatibility, 1483
Compensation Law, 663
Complementary metal-oxide-semiconductor (CMOS), 1036
Completeness, 1483
Complex permittivity, 1021
Composite curves, 878
Compound enhancement, 934*t*, 935, 947
Compound techniques, 945
Compressed Air and Gas Institute (CAGI), 846
Compressed Air Challenge, 850
Compressed air systems, 736–737
Compressibility factor, 34, 43
Compressible flow, 191–208
normal shock wave, 193–195
one-dimensional flow, 192–193
with area change, 198–203
with heat addition, 196–198
two-dimensional supersonic flow, 203–208
oblique shock waves, 204–206
thin-airfoil theory, 206–208
Compression
-ignition engine, 78
processes, internally reversible, 48*f*
ratio, 80

Compressor(s), 18, 818–850
air compressors
compressor types for various applications, 844–846
maintenance requirements, 848
performance considerations, 846–848
reduction of energy use, 850
thermal issues, 848–850
axial flow, 841–843
centrifugal, 836–840, 837*f*, 920–921
dynamic compressors, 835–836
axial flow compressors, 841–843
centrifugal compressors, 835–840
engine-driven, 847
multistage, 837
oil-flooded screw, 848
open drive, 824
positive displacement compressors, 818–835
compressor type by construction, 822–824
compressor type by kinematic linkages, 824–835
modifications to conventional vapor cycles, 819–822
reciprocating, 824–827, 919–920
rolling piston, 824, 827–829, 828
rotary screw, 921–922
screw, 832–835
scroll, 829–832, 920
semi-hermetic, 823
speeds, 847–848
stationary, 848
twin screw, 832–834
Computational fluid dynamics (CFD), 189, 1442, 1445
simulations, 621, 768, 769*f*
software and codes, 1558
Computational fluid-particle dynamics (CF-PD) methods, 597, 598*f*
Computational grid, 1549
Computational stencils, 1447, 1449, 1512, 1515
Computer(s)
casing materials, 1107–1109
cooled portable, 1106
housings, 1108*t*, 1109*t*
surface temperature, 1106
Computer aided engineering, 1442–1444
COMSOL multiphysics software, 1558
Comsol tool, 814
Concave mirrors, 1390
Concentrations, definitions of, 417–419
Concrete masonry units (CMUs), 804, 804*f*
Condensation, 380–385, 802, 807, 963
collapsing bubble, 1144–1145
direct contact, 1142–1144
dropwise, 380
falling-film, 381
film, 380
filmwise, 1229
flow regimes during horizontal cocurrent flow with, 383*f*
internal convective, 383–384
on jets, 1145–1146
in presence of noncondensable gas, 382–383
risks of, 800
Condensed water specific enthalpy, 899
Condensed water specific entropy, 899
Condenser, 566
Conduction band minimum (CBM), 1372
Conduction, finite difference analysis of, 259
Conduction heat transfer, 250–262
critical thickness of insulation, 254
finite difference analysis of conduction, 259–262
fins, 255–257
Fourier's law, 251
insulations, 251–252
internal heat generation, 254–255
long, cylindrical systems at steady state, 253
overall heat transfer coefficient, 253
plane wall at steady state, 252–253
transient systems with negligible internal resistance
bodies with significant internal resistance, 258–259
Conduction transfer function (CTF), 910
Conductivity wire probe, 1213–1214, 1214*f*
Configuration correction factor, 478, 478*f*
Conformational changes, in thermal coagulation, 675–676
Conservation of energy, 125–127
Conservation of mass, 15, 123–124
Constant air volume (CAV) systems, 735, 742–743
Constant drying rate period, 1310
Constant-escalation levelization factor, 103*t*
Constant volume gas thermometer, 4
Constant volume systems, 929, 930
Constitutive equation, 229–232
modified power-law, 230
Construction, compressor type by
hermetic compressors, 822
open drive compressors, 824
semi-hermetic compressors, 823
Construction subsystems, in thermal bridges, 803–804
Contact angle
definition of, 369
hysteresis, 962
Contact drum dryers, 1323–1324
Contact dryers, 1323–1324, 1324*f*
Contact rotary dryers, 1323
Continuity equation, 586, 587, 1462
Continuous forcing IB methods
elastic structure, 1509
rigid structure, 1510
Continuous jet, 1156
heat transfer performance, 1158
Continuous retort, 1026
Continuous wave (CW) lasers, 1210
Continuum modeling of moving interfaces, 1524–1526
Continuum models, 644–645
Continuum surface force (CSF) model, 1529
Contra-flux flow, 763
Controlling thermal resistance, 508
Control volume applications, 15–22
conservation of mass, 15–16
control volume energy balance, 15–16
control volume entropy balance, 16–17
control volumes at steady state, 17–21
Control volumes (CV), 766, 1507, 1512
Convection
boundary condition, 1470
correlations for external natural, 265–271
forced, 283, 305

mass, 437
particle gas, 388–392
Convection heat transfer, 263–328
correlations, 585
forced convection, external flows, 283–303
example, 300–301
flat plate, 283, 287–288
flow over cylinders, spheres, and other geometries, 294–295
flow over flat plate with zero pressure gradient, 291–293
flows with pressure gradient and negligible viscous dissipation, 289–290
flows with zero pressure gradient and negligible viscous dissipation, 283–284
heat transfer across tube banks, 295–301
heat transfer with jet impingement, 301–303
temperature boundary layer, 284–286
uniform heat flux, 286–287
uniform surface temperature, 285–286
uniform temperature, 290–291
forced convection, internal flows, 305–314
convective heat transfer coefficient, 306
fully developed velocity and temperature profiles, 305
heat transfer correlations, 306–312
mixed convection, 312–314
natural convection, 263–279
correlations for external natural convection, 265–271
correlations for open cavities, 271–273
example calculations, 277–279
heat transfer in enclosures, 273–277
nomenclature, 280
in non-Newtonian fluids, 315–338
forced convection heat transfer, 331–338
free convection flows and heat transfer, 323–331
laminar duct heat transfer, 315–321
turbulent duct flow for, 321–322
viscoelastic fluids, 322–323
Convection tempering, 1360, 1364–1365
Convective boiling/evaporation, 947–948
Convective condensation, 949
Convective dryers
belt dryers, 1313, 1313–1314*f*
chamber dryers, 1312, 1312*f*
superheated steam dryers, 1313–1315, 1314*f*
tunnel dryers, 1312–1313, 1312*f*
Convective heat transfer coefficient, 469
Convective loss heat coefficient, 1397
Conventional heat pipes, 954–956, 955*f*
design and limitations, 967–971, 970*f*
Conventional shell and tube heat exchanger, 1028–1029
Cooling
demand curves, 873–874
forced air, 1121
Newton's law of, 258, 431
systems, natural, 1262
Cooling in electronic applications, 1078–1124
computers
casing materials, 1107–1109
typical heat transfer schemes, 1109–1111
understanding temperature limits, 1105–1107
handheld communication devices
acceleration transforms, 1116–1118
casing materials, 1112, 1113*t*
impact of temperature on handheld product reliability, 1112–1116
thermal management, 1111–1112
understanding application requirements, 1111
high-altitude airborne electronics
fan selection, 1123–1124
impact of equipment location, 1121–1123
understanding environment, 1121
outdoor telecommunication electronics, 1118–1121
typical heat transfer methods, 1120–1121
understanding environment, 1119–1120
thermal characteristics of electronic packages
effect of aging on material properties of lead-free solders, 1091–1092
materials and properties, 1085–1091
package architecture, 1085
understanding basic semiconductor thermal parameters, 1092–1102
thermal characteristics of printed circuit boards, 1081–1085
thermal interface materials, 1102–1105
interpretation of material data, 1103
types of thermal interface materials, 1103–1105
understanding role of temperature in design
impact on performance, 1079
impact on reliability, 1079–1081
material limits, 1078–1079
Cooling load, 757–758*f*, 909
inside air heat balance, 911–912
heat balance
on inside surface, 911
on outside surface, 910
interior wall conduction, 910
Cooling thermal loads, VRC impact on, 754–756
Cooling tower(s), 864–874
for air-conditioning system, 865*f*
behavior, 872–873
cooling demand curves, 873–874
cost-optimal design of, 872
cross-flow packing showing coordinate system, 868*f*
mechanical-draft, 869, 872
natural-drift counterflow, 865
operating point of, 873
packings for counterflow and cross-flow cooling towers, 870*t*
packing thermal performance, 864–869
counterflow units, 864–865
cross-flow units, 865–866
sample calculation, 866–869
performance, calculation of, 872
for power plant, 865*f*, 866*f*
pressure drop correlations for, 871*t*, 872*t*
range and approach, 873
shells, pressure drop correlation for, 871
thermal-hydraulic design of, 869–872
Coordinate transformation, 1481–1482, 1482*f*
COP, *see* Coefficient of performance
Coriolis meter, 1198*f*
Correction factor, 296, 1006
Corrosion, 479
Corrosion fouling, 536, 609
Corrugated plate, 555*f*

Cost
escalation, 104
rate balance, 74
Cost–benefit analysis, of VRCs, 758–760
Cost-effectiveness, 734
Costing equations, auxiliary, 89–90
Cost targeting, for heat exchanger networks, 884–888
Counter current heat exchanger, 878*f*
Counterflow units, 864–865
CPA, *see* Chemical cryoprotective agent
CPAs, *see* Cryoprotective additives
Crank–Nicolson method, 1457, 1510
Crayons, temperature-sensitive, 1173–1174
Critical flow, 170–172
Critical fluidization velocity, 1320
Critical heat flux (CHF), 372, 385
model, 372
theories, 374
Critical moisture content, 1287, 1309
Critical pressure, 29
Critical temperature, 29
Critical thickness, of insulation, 254
Cross-contamination, 506
Crossed-strings method, 351, 354
Crossflow units, 865–866
Crude oil fouling models, 616–626
Ebert–Panchal model, 618–619
effect on shell-side flow and heat transfer, 624–625
fired heaters and two-phase flow, 625–626
temperature field plot, 619–620, 620*f*
threshold fouling, 618, 619
use in exchanger design, 620–624
baffle configuration, 621
cleaning, 621
fouling maps, 621–622
self-cleaning inserts, 622–624
Cryogenic pump, 862
Cryo-plunging, 547
Cryopreservation, 684
biophysical basis of freezing processes in cells, 685–687
influence of cell membrane permeability, 687–689
optimization of heat transfer during, 689–691
recent developments and innovations, 691–692
thermal stress effects, 691
tissues and organs, 687
by vitrification, 689
Cryoprotective additives (CPAs), 687
Cryosurgery, 692
Crystalline product, 1318
Crystallization, 479
Crystallization fouling, 536, 609
Crystals, phonons properties in, 1045–1049
CSF model, *see* Continuum surface force model
CTF, *see* Conduction transfer function
CTOWER program, 872
CUE, *see* Churchill–Usagi equation
Curtain walls, 806
Curve-fitting routine, polynomial, 1224
Cut-cell-based discrete forcing IB methods, 1508, 1512
Cutoff ratio, 80
CV, *see* Control volumes
C value, 1025
CW lasers, *see* Continuous wave lasers
Cylinder(s)
concentric, 274
dryers, 1291–1292
eccentric, 275
flows over, 294–295
fluid stream in cross flow over, 294*f*
horizontal, 268
sphere and horizontal, 328–331
unloading, 920
vertical, 268
Cylindrical cell model, 1008, 1009*f*

D

D2Q9 lattice example, 1499–1500, 1500*f*
Dalton model, 46, 49
Damage
accumulation process, 662*f*
parameter, 664
rate process coefficients
constant temperature exposures, 680–682
transient thermal history experiments, 682–683
Darcy friction factor, 211, 237, 557
Darcy's equation, 965
Darcy–Weisbach equation, 155, 156
Data mining, 1218
Daylighting controls, 734, 745
Dead states, 68
Dean effect, 1029
Dean number, 559
Death phase, of bacteria, 1021
Debye approximation, 1052
Debye frequency, 1048, 1049*f*
Debye length, 1011
Decision block system, feedback pathways in, 667*f*
Decision variables, 94
Deformation, 1518, 1518*f*
Degree of saturation, 900
Dehumidifying equipment, 902
Delayed effects, nonlethal thermal injury, 674
Density, 1018
Density factor (DF), 1073
Density functional theory (DFT), 1375
Departure from nucleate boiling (DNB), 376, 378
Deposition, 615
Designation and Safety Classification of Refrigerants, 927
Design optimization, 94–101
additional iterations, 101
case study, 96–100
iterative exergoeconomic procedure, 95–96
Desktop microprocessor packaging, 1038*f*
Deterioration, 486
Device scale thermometry, 1065
Device scaling, 1036
Devitrification, 1352
Dewar flask, 1167
Dew point temperature, 49, 56, 900
DF, *see* Density factor
DFT, *see* Density functional theory
DIABLO, *see* Second mitochondria-derived activator of caspases (SMAC)
Diagonally dominant coefficient matrix, 1452
Diamond shape orifices, 1159
Diaphragm resonance frequency, 1156

Dielectric constant, 1021
Dielectric drying, 1324–1325, 1324–1325*f*
Dielectric loss factor, 1021
Dielectric properties, in thermal processing, 1020–1021
Diesel spray, 987*f*
Differential scanning calorimetry (DSC), 393, 394, 790
Differential thermal analysis (DTA), 790
Diffuse interface, 1525
Diffuse mismatch model (DMM), 1051–1052
Diffuser, 19
Diffuse surface, 366
Diffusion
  approximation, 362
  equation, 427
  forced, 426
  mechanisms of, 422–426
  in moving medium, 434–437
  in porous catalyst, 432–434
  pumps, 861
  in stationary medium, 427
  term, 1463–1465
Diffusiophoresis, 426
Diffusivity, 631–635, 636*t*, 638*t*
Digital computer, 909
Dilatant fluids, 230, 231, 326*t*, 328*t*, 329–330*f*
Dimensional analysis, 139–146
Dimensionless groups, 139, 439
DIMM, *see* Dual in-line memory module
DIMs, *see* Dynamic insulation materials
Dirac delta function, 1509, 1527, 1535, 1538
Direct contact heat transfer, 1131–1148, 1135*f*
  with change of phase, 1140–1148
    collapsing bubble condensation, 1144–1145
    condensation on jets and sheets, 1145–1146
    direct contact condensation, 1142–1144
    direct contact condensation falling films, 1146–1148
    evaporation and boiling, 1140–1142
  between continuous parallel streams, 1131–1136
    failing films, 1131–1134
    heat transfer to jets and sheets, 1134–1136
  sensible heat transfer to dispersed media, 1136–1140
  with solidification, 1148
Direct costs, 102
Directional behavior, 344
Direct numerical simulation (DNS), 588, 995, 1545
Direct solvers, 1478–1479
Direct steam generation (DSG), 1392
Direct-type heating system, 1032
Dirichlet boundary conditions, 1448, 1450, 1512
Discharge coefficient, 988, 1189
Discrete forcing IB methods
  cut-cell-based, 1508, 1512
  ghost-cell-based, 1508, 1510–1512
  level set-based, 1508, 1512–1513
  numerical modeling issues for, 1513–1516
Discrete ordinates method (DOM), 365
Discretization
  of boundary conditions, 410–411
  process, 1469, 1472*f*, 1488
Disk refiner, 1271*f*
Dispersed drop flow, 216
Disperse dryers
  fluidized bed dryers, 1319–1321
  microencapsulation, 1319
  pneumatic dryers, 1321–1322, 1322*f*
  rotary dryers, 1322–1323, 1323*f*
  spray dryers, 1315–1318
Displaced enhancement devices, 933, 940
Displacement, 919
Displacement rate, 919
Dissolution process, 1359
Dissolution rate, 430
Distillation column, 893
Distillation process, 1387, 1387*f*
Distortion map, 1139*f*
  as function of Reynolds and Eötvos numbers, 1139
Distributed Lagrangian multiplier/fictitious domain (DLM/FD) methods, 1507
Distribution function, 1058
Dittus–Boelter correlation, 469, 1346
Dittus–Boelter equation, 585
DLS, *see* Dynamic light scattering
DMM, *see* Diffuse mismatch model
DNB, *see* Departure from nucleate boiling
DNS, *see* Direct numerical simulation
DOE-2 computer simulation, 743*f*
DOM, *see* Discrete ordinates method
Doping, 1044
Double pipe heat exchanger, 461, 463*f*, 467–471
Double-pipe helical heat exchanger, nanofluids results in, 557–560
Double slope solar still, 1389, 1403–1405, 1416–1417*f*, 1420–1422, 1421*f*
Double tube heat exchanger, 1027–1028, 1028*f*
  thermal nanofluid flow in, 590
Drag
  due to lift, 188
  frictional, 1134
Drag coefficients, 184, 207, 1005
Drag force, 1003, 1008, 1010, 1012, 1194
Drag reduction, 243*f*, 244
DRAM devices, *see* Dynamic random access memory devices
Drift-diffusion model, 1060
Drive motors, 860–861
Drop
  dynamics, 1136
  size correlations, 991
  temperature, 1138
Droplet
  combustion, 436–437
  evaporation, 435–436
Droplet–droplet collisions, 995
Dropwise condensation, 380, 385
Dry-bulb psychrometer, 443–444
Dry-bulb temperature, 899, 1307
Dryer design, principles of, 1330–1333
Dry flows, 962
Dry gypsum plaster, containing PCM, 793*t*
Drying
  of food products in rotary kiln, 1342
  of paper, *see* Paper making, thermal aspects of
Drying of materials
  air–water vapor mixtures properties
    dry-bulb temperature, 1307
    enthalpy, 1307–1308
    humidity, 1307

Drying of materials (*cont.*)
psychrometric chart, 1308
psychrometry, 1306
relative humidity, 1307
total pressure, 1306
wet-bulb temperature, 1307
atmospheric freeze-drying, 1328
contact dryers, 1323–1324
convective dryers
belt dryers, 113, 1313, 1313–1314*f*
chamber dryers, 1312, 1312*f*
superheated steam dryers, 1313–1315, 1314*f*
tunnel dryers, 1312–1313, 1312*f*
disperse dryers
fluidized bed dryers, 1319–1321
microencapsulation, 1319
pneumatic dryers, 1321–1322, 1322*f*
rotary dryers, 1322–1323, 1323*f*
spray dryers, 1315–1318
drying curves, 1310–1311
drying tunnel for spray drying kinetics determination, 1331*f*
in energy fields
dielectric drying, 1324–1325
infrared dryers, 1325–1326
ultrasound dryers, 1326–1327
equilibrium moisture content, 1309–1310
heat and mass transfer (external/internal conditions), 1311
method selection, 1333–1335, 1334*t*
principles, 1306–1311
principles of dryer design
acquisition of drying kinetics, 1330–1331
calculations in drying, 1332–1333
safety in, 1333
special drying techniques
flame spray drying, 1329–1330
foam spray drying, 1329
pulse combustion drying, 1328
techniques, 1312–1330
vacuum freeze-drying, 1327–1328
wet materials properties, 1308–1309
Dry-outs attributable, to interfacial instabilities, 965
DSC, *see* Differential scanning calorimetry
DSG, *see* Direct steam generation
DTA, *see* Differential thermal analysis
Dual chamber system, with convection preheating, 1364, 1365*f*
Dual-control strategy, 750, 752*f*
Dual-duct systems, 928, 930
Dual in-line memory module (DIMM), 1109
Dual-orifice nozzle atomizer designs, 992
Duct flow, turbulent, 321–322
Dufour effect, 553
Dulong–Petit law, 1048
Duplex nozzle atomizer designs, 992
Durability, in thermal bridging, 806–807
Dust explosibility, 1333
D value, 1023–1024
Dynamic compressors, 835, 848
axial flow compressors
basic description, 841
design characteristics, 841–842
flow and performance characteristics, 842–843
centrifugal compressors
basic description, 836–837
design characteristics, 837–838
development directions, 838–839
flow and performance characteristics, 839–840
Dynamic cool roof, 757
Dynamic insulation, 762
Dynamic insulation materials (DIMs), 778
background, 749–750
energy performance, 750–753
operating principle of, 750*f*
Dynamic light scattering (DLS), 548
particle size distributions by, 548, 549*f*
Dynamic random access memory (DRAM) devices, 1109
Dynamic Smagorinsky model, 1549, 1550–1552
Dynamic thread migration, 1038
Dynamic viscosity, 1018

E

Earthen pots, 1390
Ebert–Panchal model, 618–619, 625
Eccentric spheres, 276
Eco-Indicator 99, 106
ECOs, *see* Energy conservation opportunities
Effective conductivity, 1006
Effective heat capacity method, 412
Effervescent atomizers, 993, 994
Efficiencies
of motors, 735*t*
of solar still, 1407–1410
Egg white, 682
Egg yolk, 682
e^-/h^+ separation, 1380
Einstein equation, 551
Einstein model, for phonons, 1067
EIS, *see* Electrochemical impedance spectroscopy
Ejection stroke, in synthetic jet operation, 1153, 1153*f*
Elastomer-on-flex package, 1085, 1114
Elastomer pads, 1104
Elbows, flow through, 157
Electrical boundary layer, 1004
Electrical conductivity, 1021
Electrical energy consumption, 742*f*, 743
Electrical properties, in thermal processing, 1021
Electrical resistance, 1021
Electrical sheath, 1004
Electrical systems, 734–735
Electric field effects, on heat transfer, 1011–1013
Electric sheath, 1011
Electrochemical impedance spectroscopy (EIS), 1375
Electrochemical potential, 1381
Electrohydrodynamic enhancement, 947
Electrohydrodynamic flow, 967
Electromagnetic actuators, 1154, 1156
Electromagnetic field, 1021
Electromagnetic heating
microwave, 1030–1031
radio frequency, 1030
Electron–hole separation, 1375
Electronic cooling, 858
Electronic mold compounds (EMCs), 1086
Electronic packages
parameters for thermal performance of, 1092–1102

thermal characteristics of, 1085–1102
Electron–ion recombination, Nusselt number variation for, 1013*f*
Electron–ion transport effect, on particle heat transfer, 1004
Electron–phonon scattering, 1066, 1077
Electrons, 1040
Electrons properties
in metals, 1042–1043
in semiconductors, 1043–1045
Electroosmotic heat pipes, 967
Electrostatic atomizers, 994
Electrostatic fields, 935
Electrostatic/charge stabilization, 577, 577*f*, 578*f*
Electrothermal simulations, 1059
Electrothermal solutions, 1072, 1073*f*
Elemental matrices/vectors, numerical evaluation of
element technology, 1483–1488
isoparametric elements, 1481–1483
Element connectivity, 1474*f*, 1475*t*, 1490*t*
EMCs, *see* Electronic mold compounds
EMF–temperature (*E*–*T*), 1166*f*
Emissive power, 366
Emissivity, 343, 366, 1343
Emittance, 366
EMPL, *see* Extended modified power-law
Enclosure, idealized, 347
Energy
conservation, 125–127, 131–133
measures, 731–738
opportunities, evaluation of, 744–746
consumption, electrical, 742*f*, 743
distribution systems, 928
for drying, 1332–1333
losses, 84
management and control system (EMCS), 737–738
for moisture evaporation, 1332
rate balance, 46, 63
-saving measures, 731
sensible, 906
service company (ESCO), 728, 746
use
baseline model, 743–744
reduction of, 850
Energy audits, for buildings, 728–747, 732*t*, 733*t*, 741*f*
case study
building and utility data analysis, 740
energy use baseline model, 743–744
evaluation of energy conservation opportunities, 744–746, 745*t*
on-site survey, 741–743
recommendations, 746
common energy conservation measures
building envelope, 731–734
compressed air systems, 736–737
electrical systems, 734–735
energy management controls, 737–738
HVAC, 735–736
indoor water management, 738
new technologies, 738–746
general procedure
baseline for building energy use, 731
building and utility data analysis, 730
evaluation of energy-saving measures, 731
walk-through survey, 730–731
types
detailed energy audit, 729–730
standard energy audit, 729
utility cost analysis, 729
walk-through audit, 728
verification methods of energy savings, 746–747
regression models, 746–747
time variant models, 747
Energy audit summary
for industrial facilities, 733
for residential and commercial buildings, 732
Energy balance equations
for double slope solar still, 1403–1405
for single slope solar still, 1402–1403
Energy balance model, 766*f*
Energy conservation equation, 1529
Energy conservation opportunities (ECOs), 744–746, 745*t*
Energy consumption, effect of thermal bridging, 802
Energy equation, 586, 587
Energy fields, drying in
dielectric drying, 1324–1325, 1324–1325*f*
infrared dryers, 1325–1326, 1325–1326*f*
ultrasound dryers, 1326–1327, 1327*f*
Energy*flo* Cell, 764
Energy matrices, 1410–1411
Energy output, 1405–1406
Energy payback time (EPBT), 1393, 1410
Energy penalty, 892
Energy performance
of building envelope systems, 750–753
numerical tools for PCM thermal and, 791–792
EnergyPlus format (EPW), 771
Energy production factor (EPF), 1393, 1410–1411
Energy transport
in gas, 1348–1349
in solid, 1347–1348
Engine
-driven compressors, 847
external combustion, 83
four-stroke internal combustion, 78
spark-ignition, 78
Engineering thermodynamics fundamentals, 3–14
basic concepts and definitions, 3
equilibrium, 4
phase and pure substance, 4
process, cycle, 4
state, property, 3–4
system, 3
temperature, 4–5
entropy, 12–13
entropy balance, 13–14
first law of thermodynamics, 5–8
energy, 6–7
heat, 7
power cycles, 8
work, 6
second law of thermodynamics, 8–12
Carnot corollaries, 10
Carnot efficiency, 11
Clausius inequality, 11–12
irreversibilities, 9–10
Kelvin–Planck statement, 8–9
Kelvin temperature scale, 10–11

Enhanced surfaces, 934*f*, 935
ENO scheme, *see* Essentially Non-Oscillatory scheme
Enthalpy, 883, 901, 1307–1308
  chart, 868*f*
  of combustion, 59–60, 64
  curves, 789, 790*f*
  of formation, 58–61
  generalized charts for, 50*f*
  intervals and area targeting, 886*f*
  scheme, 407
Enthalpy method, 405, 406–407
  application of, 1532–1533
  formulation, 1529–1531
  microscopic model, 1531–1532
  numerical methodology, 1531
Enthalpy–porosity formulation, 1530
Entrainment limitation, of heat pipe, 968
Entropic lattice Boltzmann method, 1502
Entropy, 12–13
  absolute, 60–62
  balance, 13–14
  change relation, 195
  generalized charts for, 51*f*
  generation, 12, 62, 64
    in thermal nanofluid flow, 588–589
  specific, 25
  transfer, 13, 17
Enviroeconomic parameter, for solar still, 1413–1414
Environmental impact rate, 106
Enzyme deactivation model, 666–667
EoS, *see* Equation of state
EPBT, *see* Energy payback time
EPF, *see* Energy production factor
Epoxy, 1085, 1087*f*
  thermal, 1104
EPRT, *see* Equation for phonon radiative transfer
Epstein model, 618
EPW, *see* EnergyPlus format
Equal-amount money transactions, 104
Equation(s)
  Bernoulli, 134, 135, 290, 1199
  heat conduction, 251
  Kozeny–Carman, 150
  Laplace, 136
  momentum integral equation, 181
  Navier–Stokes, 131, 1132–1133, 1141
  power law, 315
  reaction equilibrium, 65
  Redlich–Kwong, 35
  of state, 27, 34
Equation for phonon radiative transfer (EPRT), 1057
Equation of state (EoS), 995
  extended modified power-law (EMPL), 230–232
  power-law model, 229–230
  rheological, 229–232
  viscoplastic models, 232
Equilibrium
  constant, 65, 66*t*
  distribution function, 1499
  moisture, 1309
  state, 4
Ericsson cycle, 82
Essentially Non-Oscillatory (ENO) scheme, 1540
ETC, *see* Evacuated tube collector
ETC-integrated solar still, 1429, 1429*f*
Eulerian–Lagrangian framework, 587
Europe
  thermal bridge calculation standards, 814
European Committee of Manufacturers of Compressors, Vacuum Pumps and Pneumatic Tools (PNEUROP), 846
Evacuated multifoil insulation, 720
Evacuated tube collector (ETC), 1392
Evaporation, 963, 1140–1142
  droplet, 435–436
Evaporative loss heat coefficient, 1397
Exchange coefficients, 441
Exchanger
  flow arrangements, 514–516*t*
  properties, 876*t*, 877*t*
Exchanger design, use of fouling models in, 620–624
Excitation frequency, 1153
Exergetic efficiency, 71–74, 75
Exergoeconomic parameter, for solar still, 1412–1413
Exergoeconomics, 67, 86–93
  auxiliary costing equations, 89–90
    auxiliary equations, 89–90
    definition of fuel and product, 89
    exergy streams, 89
  cost balance, 88–89
  exergy costing, 87–88
  general example, 90–91
  interactions of, 94
  variables and evaluation, 91–93
Exergoenvironmental analysis, 106
Exergy
  analysis, 67–75, 107, 928
  balance, 68
  balance, 68
  control volume exergy rate balance, 68–69
  costing, 74–75, 87–88
  defining, 67
  destruction ratio, 92, 96
  exergetic efficiency, 71–74
  output, 1406–1407
  removal of, 89–90
  supply of, 90
  unit, 90
Exfiltration, 763
Exit velocity, 45
Experimental efficiency
  for varying temperature gradients and airflow rates, 768*f*
Explicit formulation, 262
Explicit scheme, 410
Explosion venting, 1333
Extended modified power-law (EMPL), 318–319
  constitutive equation, 230–232, 236–238
Extended surfaces, 933, 938
Exterior air gaps, 764
External combustion engine, 83
External film condensation, 381
External incompressible flows
  airfoils, 188
  boundary layer control
    flaps and slats, 188–189
    moving surfaces, 189
    profile shaping, 188

suction and blowing, 189
boundary layers
approximate solutions, 181
concept, 179–180
effect of pressure gradient, 182–184
exact solution of laminar flat-plate boundary layer, 181
thickness definitions, 180–181
computation vs. experiment
computational fluid dynamics, 189
role of wind tunnel, 189
drag
friction and pressure drag, 186–187
friction drag, 184–185
pressure drag, 185–186
streamlining, 187
induced drag, 188
lift, 187–188
External pulsations, 962
Extrinsic semiconductors, 1044
Eyring–Polanyi equation, 660

F

Façade systems, 770–776
Face centers, 1455
Facultative anaerobes, 1022
Fahrenheit scale, 5
Falling drying rate period, 1310
Falling films, 1131–1134
condensation, 381
direct contact condensation, 1146–1148
modeling of, 1147
False transient approach, 1347
Fan, 858–861
applications, 853
characteristics of various, 860*t*
coil unit (FCU), 742–743
rotational velocity, 1123
selection, 1123–1124
slip, 1124
Fanning friction factor, 142, 240, 520, 527, 936
Fans for air-cooled heat exchangers, 490–491
Faradaic efficiency, 1376
Faraday's law of induction, 657
Far-field temperatures, 1006, 1007, 1013
Fast-circulating fluidized beds, 390–392
Fast-circulating regime, 390
Fast closing valve, 1212
Fast fluidized beds (FFBs), 391
FCG, *see* Fluid-conditioned garment
FD distribution, *see* Fermi–Dirac distribution
FDE, *see* Finite-difference equation
FDM, *see* Finite-difference method
Feedforward network, 1221–1222
FEM, *see* Finite-element method
Fermi–Dirac (FD) distribution, 1040, 1041*f*
Fermi energy, 1043
Fermi level, of semiconductor, 1372, 1374*f*
Ferrofluids, 590–593, 591*f*
FETs, *see* Field-effect transistors
FFBs, *see* Fast fluidized beds
FHP automaton, *see* Frisch-Hasslacher-Pomeau automaton
Fiberglass pipe, 725
Fiber mix, 1300
Fiber-optic blackbody temperature detector system, 1173
Fibers, 1277
Fibrous composition, 1300
Fickian diffusion, 1287
Fick's law, 423, 427
Field-effect transistors (FETs), 1036
Fighter aircraft, 1122
"Figure of merit," 969, 971*f*
Fillers, 1273
Filling ratio (FR), 566
Film(s)
boiling, 371, 385
condensation, 381, 385
countercurrent flow, 1132–1133
falling, 1131–1134
Fin(s), 255–257, 262, 517, 518*t*
circular, 273
effectiveness, 257, 262
efficiency, 256, 262, 492–493
isothermal triangular, 272–273
temperature distribution in, 492*f*
Fining gases, 1359–1360
Finite-difference equation (FDE), 1445, 1447*f*
Finite-difference method (FDM), 1443, 1445–1453
-based algebraic formulation, 1446*f*, 1450–1452
governing differential equation and boundary conditions, 1447–1448
grid generation, 1448–1450
main idea of, 1445–1447
solution of algebraic equations, 1452
Finite-difference quotient, 1446
Finite-element equations, 1471–1475
Finite-element mesh, 1469
Finite-element method (FEM), 1443
element technology, 1483–1488
numerical integration, 1487–1488
three-dimensional elements, 1485–1487
two-dimensional elements, 1484–1485
examples
bi-material problem with imperfect interface, 1493–1494
typical two-dimensional problem, 1490–1493
governing differential equation, 1470–1471
introduction, 1468
isoparametric elements, 1481–1483
main idea of, 1468–1470, 1469*f*
numerical evaluation of elemental matrices/vectors, 1481–1488
practical issues
choice of elements, 1488
choice of solvers, 1489
meshing, 1488
mesh refinement, 1488–1489
order of gauss integration, 1489
semi discrete model, in general case, 1476–1477
solution phase, 1477–1481
linear equation solver, 1478–1479
nonlinear equation solver, 1479–1481
time integration algorithm, 1479
transient nonlinear equation solver, 1481
sources of error in, 1469–1470
steps in derivation of finite-element equations, 1471–1475

Finite-volume method (FVM), 1443, 1512
FVM-based algebraic formulation, 1459–1466
advection term, 1462–1463, 1464*f*
continuity equation, 1462
diffusion term, 1463–1465
pressure as source term, 1465–1466
transport equations, 1462
unsteady term, 1462
governing differential equations, 1454
grid generation, 1454–1455
main idea, 1455–1459
approximations for algebraic formulation, 1456
average flux and volumetric term (first approximation), 1457
integral form of governing equation, 1456
interpolation/extrapolation method and central difference method (second approximation), 1458–1459
Finned surfaces, 935
Finned tube type, 485–487
Fins, 485
Fired heaters, 625–626
Firing of rotary kilns, 1338
First law of thermodynamics, 5–6, 301
First-order solution, 662
First-order-upwind (FOU) scheme, 1459
Fixed-capital investment, 102
Fixed-matrix regenerator, 504, 506
Fixed vane rotary compressor, 827, 827*f*
Flame ionization detector, 56
Flame spray drying (FSD), 1329–1330, 1330*f*
Flat fins, 526
Flat plate, 283
flows over with zero pressure gradient, 291–293
with injection, 290
with prescribed nonuniform surface temperature, 288–289
Flat plate collector (FPC), 1390
Flex-substrate BGA, 1085, 1113, 1114*f*
Float forming process, 1360
Floating bodies, 119
Flocculation level, 1273
Flooding rates, 1138
Floodplains, 171*t*
Floor slabs, heat transmission loss through, 914
Flow
axial, 858
choked, 197
coefficient, 1189
compressible, 192, 198, 841
critical, 170–172
direction, 1211–1212
dispersed drop, 216
field, 1005
free molecule, 426
gas–liquid two-phase, 212–217
gradually varied, 174–176
hypersonic, 192
incompressible, 124, 127
inviscid, 135
irrotational, 134
laminar, 306–307
Mach number, 838
maldistribution, 535
nozzles, 201, 1189, 1190–1192
characteristics of, 1190*t*
one-dimensional, 192–193, 196–198
one-dimensional flow with area change, 198–203
diffuser, 202–203
nozzle flow, 201
pattern(s)
in gas–liquid horizontal flow, 213*f*
in gas–liquid vertical flow, 213*f*
maps, 212, 214*f*
for vertical upflow of solid particles, 218*f*
potential, 123–138
reversal measurements, 1209
reversible adiabatic, 193
solid–liquid, 209
subsonic, 192, 197
turbulent, 307–308, 312, 321–322
two-dimensional supersonic, 203–208
uniform, 170
vapor–liquid, 209
zero, 858
Flow measurement, 1188–1214
classification of, 1188
classified, 1188–1214
direct methods, 1188
hot-wire anemometry, 1202–1205
constant current anemometry, 1202
constant temperature anemometry fundamentals, 1202–1203
corrections due to temperature, 1205
directional response, 1205
equations governing, 1203
multicomponent probes, 1203–1204
probe design, 1204
laser Doppler anemometry, 1206–1214
particle imaging velocimetry, 1209–1212
principle of operation, 1206–1209
viscosity measurements, 1212
void fraction measurement, 1212–1214
linear flow meters, 1192–1198
Coriolis meters, 1198
electromagnetic flow meters, 1198
laminar flow elements, 1197–1198
positive displacement meters, 1195
rotameters, 1193–1194
target flow meters, 1195
thermal mass flow meters, 1195–1197
turbine flow meters, 1195
ultrasonic flow meters, 1197
vortex shedding meters, 1195
restriction flow meters for flow in ducts, 1188–1190
flow nozzles, 1190–1192
orifice plates, 1189
venturi meters, 1192
traversing methods, 1198–1202
design of Pitot-static probes, 1201–1202
measurement of flow direction, 1202
pressure probes, 1199–1201
Flow meter, 1215
Flowsheeting software, 85
Flow simulation software, 1558
Flow velocity, 537
Flue gas, 875–876
FLUENT software, 1558

Fluid(s)
acceleration, 121
classification of, 225–228
element, equilibrium of, 115
friction, 155–157
motion, differential relations for, 127
Newtonian, 130
non-Newtonian, 130, 223, 315
power law, 320, 323, 332–335*t*
temperature, mean, 306
vibration, 935
viscoelastic, 130
Fluid-conditioned garment (FCG), 655
Fluid dynamics, 1442–1443
software and computer codes for, 1558
Fluidization, 1320
Fluidized bed dryers, 1319–1321
Fluidized state, 1320
Fluid mechanics analysis, 857*f*
Fluid statics, 115–122
buoyancy, 119
equilibrium of fluid element, 115
hydrostatic forces
in layered fluids, 119
on submerged objects, 116–119
hydrostatic pressure, 115–116
manometry, 116
pressure variation in rigid-body motion of fluid, 121–122
stability of submerged and floating bodies, 120–121
Fluid-structure interaction (FSI) simulations
immersed boundary methods for, 1507–1521, 1508*f*
Fluoroptic temperature-measuring system, 1173
Flux, definitions of, 421
Fly ash dispersions, 361–362
Foam spray drying, 1329
Foodborne botulism, 1017
Food dryer, 1342
Foods preservation and thermal processing, *see* Thermal processing system and preservation of foods
Forced air cooling, 1121
Forced convection, 283, 305, 588
Forced convective system, 1364, 1364*f*
Forced diffusion, 426
Forced-draft arrangements, of air-cooled heat exchangers, 489*f*
Forehearths, 1356
FOU scheme, *see* First-order-upwind scheme
Fouling, 459, 479, 501
and cleaning, 609, 625
in crude oil, *see* Crude oil fouling models
defined, 603
deposit, 609
in food processing, 613–616
in heat exchangers
accounting of, 537
control and cleaning techniques, 540
effect and mechanisms, 535–536
importance of, 536–537
influence of operating and design variables, 537–540
mechanisms, 610*t*
quantitative impact of, 603–604
Fouling Biot number, 604
Fouling dynamics
pressure drop, 607–608
timescales, 606–607
Fouling factor, 537, 604
and design parameters for finned tubes in fossil fuel exhaust gases, 540*t*
for various fluid streams used in heat exchangers, 538*t*
Fouling maps, 621
Fouling mechanisms, 609–610
Fouling service, designing for, 611–613
Fourdrinier paper machine, 1270, 1282
Fourier heat conduction law, 132, 953, 1359
Fourier law, 251, 262, 1039, 1058
Four-stroke internal combustion engine, 78
FPC, *see* Flat plate collector
FPC-integrated solar still, 1426–1427
FR, *see* Filling ratio
Free air delivery, 846
Free convection, from vertical isothermal plate, 144–146
Free moisture, 1309, 1310
Free molecule flow, 426
Free-stream velocity, 283
Free water, 1280–1281
Freeze desalination, 1148
Freeze-drying, vacuum, 1327–1328, 1327*f*
Freeze/thaw protocols, 689
Freezing
fouling, 536
processes, 685
time, 403
Freon's refrigerants, 923
Frequency domain thermoreflectance, 1060
Freshwater, 1386, 1387*f*
Friction drag, 184–185, 187
Friction factors, 235, 237*f*, 271, 520, 526–527, 560, 561*f*
for commercial pipes, 150
correlation, 492
Darcy, 211
fanning, 237, 240
plot, 152
for polyethylene oxide (Polyox) solutions, 243*f*
Frisch-Hasslacher-Pomeau (FHP) automaton, 1497
Front capturing methods, 405
Front head, 475
Front tracking methods, 404–405, 1525
Frozen tissue, thermal properties of, 639
FSCK, *see* Full-spectrum correlated k-distributions
FSD, *see* Flame spray drying
FSI simulations, *see* Fluid-structure interaction simulations
Fuel-air ratio, 57
Fuel stream, 86
Fugacity
coefficient, 51
generalized charts for, 52*f*
Full-spectrum correlated k-distributions (FSCK), 366
Fundamental heat conduction physics, 1039–1042
Fungi, 1021
Future value, of money, 102
F value, 1024
FVM, *see* Finite-volume method

G

Galilean invariance, 1498
Galvanometric measuring instruments, 1163

Gardon gauges, 1176
Gardon-type heat flux meters, 1179
Gas(es)
  additives for, 935
  calorically perfect, 194
  chromatograph, 56
  conduction, 717–718
  convection, 1346
  ideal, 19
  lifts, 855
  liquid two-phase flow, 212–217
  mixtures, 35–37, 419
  power cycles, 76
  scale, empirical, 5
  solubility and permeability of in solids, 428–429*t*
  turbine, 100, 988
Gas admixing processes foam spray drying, 1329
Gas desorption foam spray drying, 1329
Gasketed PHE, 506, 507*f*
Gas radiation properties, for rotary kiln, 1346
Gas-to-fluid exchangers, 501
Gauge pressure, 116
Gaugler, R.S., 957
Gauss divergence theorem, 1456
Gaussian phonon wave packet, 1055
Gauss integration scheme, 1488, 1489
Gauss method, 513
Gauss–Seidel method, 1452
GCC, *see* Grand composite curve
Gear drives, 861
Generalized Newtonian fluids (GNFs), 225
Genetic algorithms, 1248–1262, 1250*f*, 1250*t*
Genetic programming (GP), 1255–1257
  additional applications in thermal engineering, 1260–1262
  compact heat exchangers applications, 1257–1259
  computer representation, 1255–1256
  effect of penalty parameters, 1259–1260
  fitness and penalty, 1256
  methodology of, 1257
Geometry rule, 815
Gerris flow solver software, 1558
Gettering pumps, 862
Ghost-cell-based discrete forcing IB methods, 1508, 1510–1512
Ghost grid points, 1510, 1511*f*
Gibbs energy, 956
Gibbs function, 51, 54, 59*t*
Gibbs phase rule, 54, 396
Glassmaking process, 1352
Glass manufacturing processes
  composition and effect on properties, 1352–1355
  convection tempering, 1364–1365
  fining and bubble dynamics, 1359–1360
  general flow of, 1353*f*
  heat strengthening, 1366
  heat transfer in, 1352–1367
  heat treatment for glass transformation, 1360–1365, 1360*f*
  melting process, 1355–1359
  radiation tempering, 1363–1364
  tempering furnace, heat transfer phenomena in, 1366–1367
Glass tube rotameters, 1193
Glazing retrofit, 745
Global coordinates, 1460, 1460*f*
Global warming potential (GWP), 924
GNFs, *see* Generalized Newtonian fluids
GNU Octave software, 1558
Governing equations, 1512, 1548–1549
GP, *see* Genetic programming
Gradually varied flow, 174–176
Graetz solution, 1135
Gram staining method, 1022
Grand composite curve (GCC), 887, 887*f*, 888*f*, 894
Graphene, thermal conductivity of, 1068, 1069*f*
Grashof number, 145, 439, 1356
Gravity-assisted thermosyphons, 973
Gray analysis, 345, 347, 366
Greenhouse roofs, 1391
Grid generation, 1448–1450, 1554
Grid lines, 1449
Grid points, 1443, 1510, 1511*f*, 1513*f*
Grid refinement, 1540
Grid spacing, 1449
Grover, George M., 957
GWP, *see* Global warming potential
Gypsum boards, 786–788, 792

H

HAAKE RS6000 rheometer, 557
Hadamard model, 1138
Halophilic microorganisms, 1022
Handbook of Fundamentals (ASHRAE), 855, 898, 918, 1190
Handheld portable product, material properties of housing for, 1113
Handheld product reliability, 1112–1116
Hard drives, 1106
Hardy, de Pazzis, and Pomeu (HPP) automaton, 1496
Hastelloy C-276, 506
HCFCs, *see* Hydrochlorofluorocarbons
Heat
  addition, effect on subsonic flow, 197
  balance, on outside surface, 910
  of desorption, 1298–1300, 1299*f*
  engine, 8
  exchange, in presence of participating medium, 362–364
  flux, 1176–1179
    condition, minimum, 375
    gauges, 1180, 1182
    uniform, 308–310, 312
  gains, 751
  generation
    internal, 254–255
    metabolic, 642
  interaction, 7
  load, 954
  losses, 751
    calculations, 722–723
    across window cavity, 278–279
  motive power of, 3
  ratio, specific, 25
  recovery technologies, 739, 1297
  sources, electronic components as, 1254
  strengthening, 1360, 1366, 1366*f*

transmission loss, 913–914
Heat3 tool, 814
Heat balance integral, 401
Heat capacity method, 412
Heat conduction
classification of problems, 1471*t*
equation, 251, 262
Fourier's law of, 132, 953
physics, 1039–1042
Heat dissipation, region, 1072*f*
Heat Exchanger Design Handbook, 461
Heat exchanger networks (HENs), 875, 876*f*, 877*f*, 884–886, 888–891, 892*f*
integrating with other energy devices, 892–894
Heat exchangers, 96, 1219–1220
classification of, 459
cost targeting, 885
counter current, 878*f*
fouling in, 535–540
nanofluid flow in, 590
nanofluids application in, 545–570
for process and energy industries, 459–461
air-cooled, 484–498
principles of design, 461–471
shell-and-tube, 471–484
types, used in food processing operations, 615*t*
wall as, 763
Heat flow
across metal–dielectric interfaces, 1053
across metal–metal interfaces, 1053
Heat flow paths, 803
from building structure, 805
Heating
by collectors, 1390
physiologic effects of local, 658
therapeutic, 656
Heating equipment, thermal processing and foods preservation, 1025–1031
Heating load, 755–756*f*, 912
air distribution system heat losses, 914
auxiliary heat gains, 914–915
heat transmission loss, 913–914
indoor design conditions, 913
infiltration, 914
outdoor design conditions, 913
Heating seasonal performance factor (HSPF), 832
Heating thermal loads, VRC impact on, 754
Heating, ventilating, and air conditioning (HVAC) systems/applications, 735–736, 739, 858, 1217
Heat pipes
conventional, 954–956, 955*f*
design and limitations, 967–971, 970*f*
description of, 953
development history, 956–958
evaporators, nanofluids performance in, 566–569
example problem, 975–977, 975*f*
genealogy of heat pipe family, 971–974, 972*f*
literature on, 974–977, 974*f*
loop heat pipes (LHPs), 958–960
micro heat pipes, 964–965
pulsating heat pipes (PHPs), 960–964
thermophysical property data, 977–978
types, 958–964, 966–967
variable conductance heat pipes (VCHPs), 965–966, 966*f*
Heat pump, 915
with HEN analysis, 894
Heat-recovery steam generator (HRSG), 98, 99
Heat transfer, *see also* Conduction heat transfer; Convection heat transfer; Direct contact heat transfer
coefficient, 1184*f*, 1288
conductance, 441
conduction in refractory wall, 1346–1347
correlations, 306–312, 491
cost targeting, 885–886
data, correlations of nucleate pool boiling, 373
definitions for computing, 266*f*
determination of, 527
during drying, 1285, 1311
in enclosures, 273–277
exchanger and pressure drop analysis, 509–522
film coefficients, 479
fin efficiency and extended surface efficiency, 513–522
and flow friction correlations, 522–527
fluids, intermediate, 84
fouling effect on, 624–625
free convection flows and, 323–331
and friction for fully developed flow-through specified ducts, 523*t*
fully turbulent, 265
in glass manufacturing processes, 1352–1367
internally reversible, 21–22
with jet impingement, 301–303
laminar duct, 315–321
methods, 1120–1121
in nanofluid flow, 584–588
discrepancies, 589–590
experimental evidence, 584–585
natural convection, 588
numerical studies, 585–588
natural convection, 263
optimization during cryopreservation, 689–691
in plasma sprays, 1003–1013
to array of particles, 1008–1010
defining features of plasma heat transfer, 1004–1005
electric field effects, 1011–1013
to single particle, 1005–1008
plate-fin extended surfaces, 525–526
in presence of synthetic jets, 1153–1161
radiation, 718
rate, 299, 802
resistance, negligible liquid-side, 865
in rotary kilns, 1338–1351
schemes, 1109–1111, 1109*f*, 1110*f*, 1120*f*
in solar still, 1394–1399
external heat transfer, 1395–1396
internal heat transfer, 1396–1397
modified internal heat transfer, 1398–1399
overall heat transfer, 1397–1398
tube-fin extended surfaces, 526–527
unresolved problem
fouling and cleaning, 609
fouling dynamics, 606–608
quantitative impact of fouling, 603–604
robust design, 605–606

Heat transfer coefficient, 296, 478, 493, 530, 608
average, 286, 287
convective, 306
estimation of, 635–637
evaluating, 1182–1186
external, 1142
local, 285–286
overall, 253
single convective, 392
single-phase, 378
Heat transfer enhancement
convective boiling/evaporation, 947–948
convective condensation, 949
pool boiling, 945–947
single-phase forced convection, 936–945
single-phase free convection, 935–936
techniques, 545–546, 570
vapor-space condensation, 948–949
Heat transfer measurements, 1162–1186
evaluating heat transfer coefficient, 1182–1186
direct methods, 1183
indirect methods, 1185
steady state, 1183–1185
transient lumped parameter systems, 1185
heat flux
calibration, 1179
circular foil gauges, 1179
hot-film anemometry, 1178–1179
planar heat flux gauge, 1177–1178
slug calorimeter, 1176
sensor environmental errors, 1180–1182
steady-state errors
in gas-temperature measurement, 1180–1181
in heat flux gauges for convective heat transfer, 1182
in solid and surface-temperature measurements, 1181–1182
temperature measurement, 1162–1186
radiation devices, 1172, 1173
resistance temperature detectors, 1170
temperature-sensitive paints, crayons, and badges, 1173–1174
thermistors, 1168
thermocouples, 1162–1168
Heat Transfer Research, Inc. (HTRI), 479
Heavily doped semiconductors, 1381
Heaviside and Dirac-delta functions, 1538
HEC QP-300 polymer, 946
Heikkilä formula, 1299
Heisler charts, 1140
Helical baffles, 1029
Helical heat exchanger, 1029
Helically coiled tube, 555*f*
Heller cycle, 1131
Helmholtz resonance frequency, 1156
HEM, *see* Homogeneous equilibrium model
Hemicelluloses, 1277
Henry constant, 420
HENs, *see* Heat exchanger networks
Hermetically sealed compressor, 919
Hermetically sealed container, 1033
Hermetic compressors, 822
Herschel–Bulkley fluid model, 232–235, 1020
Heterogeneous catalysis, 1380
Heterogeneous nucleation, 369
Hexahedral elements, 1486–1487
Hidden pinch phenomenon, 890, 890*f*
High convective system, 1364, 1365*f*
High-finned tubes, 494*t*
Highly resolved simulations (HRSs), of sprays, 997–1000, 998*f*
High-side compressor, 822
High-temperature distillation, 1390
Hill-climbing, algorithm, 1248
Homogeneous equilibrium model (HEM), 211
Homogeneous LES filter, 1546–1547
Homogeneous nucleation, 369
Hood air system, 1296
Hooke's law, 1509
Horizontally split casings, 855
Horizontal-tube condensers, 3-D surfaces for, 949
Hornification, 1281
Hot filling process, 1032–1033
Hot-film anemometry, 1178–1179
Hot water correction, 870
Hot-wire anemometry (HWA), 1177, 1202–1205
HRSG, *see* Heat-recovery steam generator
HRSs, *see* Highly resolved simulations
H–S bounds, evaluation of, 552*f*
HSPF, *see* Heating seasonal performance factor
HTRI, *see* Heat Transfer Research, Inc.
Human arterial tissue, temperature dependence of, 638–639
Human comfort, 907
Human thermoregulation, 651
Humidity, 1307
Humidity ratio, 50, 849, 899, 900, 901, 902, 904, 905, 908, 914
HVAC systems/applications, *see* Heating, ventilating, and air conditioning systems/applications
HWA, *see* Hot-wire anemometry
HYBCELL 1.2, 773–774
Hybrid methods, 412
Hybrid models, 648
Hybrid nanofluids
in heat exchangers, 557, 569
heat-transfer coefficient vs. pumping power for, 566*f*
HTC versus flow velocity for, 564*f*
in plate heat exchanger (PHE), 563–566
pressure drop vs. volumetric flow rate for, 565*f*
thermal conductivity and viscosity of, 564*f*
Hybrid PVT active solar still, 1426
Hybrid sphere, 563
Hydraulic jump, 172–173
Hydraulic radius, 170
Hydrochlorofluorocarbons (HCFCs), 924
Hydro-drilling, 625
Hydrodynamic boundary layers, 1004, 1011
Hydrodynamic model, 1060
Hydrodynamics, 962
Hydrofluorocarbons, 819
Hydrogen
bonding, 1283
as renewable fuel, 1369
Hydrostatic forces
in layered fluids, 119
on submerged objects, 116–119

Hydrostatic pressure, 115–116
Hydrostatic retorts, 1026
Hygroexpansion, 1284
Hygrometer, 1018
Hygrothermal performance, 803
Hypersonic flow, 192

I

IB method, *see* Immersed boundary method
Ice bath, 1167*f*
Ice storage control, 744–745
Ideal fluid, with zero viscosity, 179
Ideal gas, 19
  expressions, 44
  mixtures, 46–49, 61
  model, 27, 43–45, 71
IEQ, *see* Indoor environmental quality
IIF, *see* Intracellular ice formation
Image point, 1511*f*, 1512
  technique, 1451
Image processing, and temperatures, 1176
IMC, *see* Internal model control
Immediate effects, nonlethal thermal injury, 673–674
Immersed boundary (IB) method
  applications
    with moving deformable structures, 1517–1520
    for moving rigid structures, 1516
  continuous forcing approach, 1509–1510
  cut-cell-based IB method, 1512
  discrete forcing approach, 1510–1516
  for fluid-structure interaction (FSI) simulations, 1507–1521
  ghost-cell-based IB method, 1510–1512
  level set-based IB method, 1512–1513
  numerical modeling issues for, 1513–1516
Impeller, in centrifugal pump, 854
Impinging synthetic jet, 1157
  cooling performance of, 1158
Implicit formulation, 260, 262
Incense stick (*agarbatti*), 1155
Incident photon to current efficiency (IPCE), 1377
Incoloy 825, 506
Incompressible flow, 124, 127
  vorticity in, 133–135
Incompressible model, 27, 40
Inconel 625, 506
In-container sterilization, 1033
Increase of entropy principle, 14
Indirect costs, 102
Individually finned tubes, 526
Indoor environmental quality (IEQ), 749
Induced-draft arrangements, of air-cooled heat exchangers, 489*f*
Induction period, 606
Industrial insulations, 723
Inelastic fluids, 225
Inertia, centroidal moments of, 119*f*
Infiltration, 763, 914
Infinite parallel plates, 310
In-flow sterilization, 1033
Infrared dryers, 1325–1326, 1325–1326*f*
Infrared thermography, 1172–1173
Infrared thermometry, 1060
Inhomogeneous LES filter, 1546
Injection, 935
Inorganic phase changing materials, 782, 784–785*t*
Insulation(s), 251–252
  application, 724–726
  critical thickness of, 254
  evacuated multifoil, 721
  heat transfer in, 718
  industrial, 723
  systems, 718, 721–724
  thermal, 723–724, 732
Integral conductivity, 1006
Integrated part load value (IPLV), 838
Intelligent information systems, 1218
Intensification, 933
Intensive property, 3
Interfaces, 1065, 1069
Interfacial points, 1513
Interfacial structure, 1381
Interfacial thermal conductance, 1049–1054, 1050*f*, 1069–1071
Interfiber bonds, 1300
Interior air gaps, 764
Interior wall conduction, 910
Internal convective boiling, 376–379
Internal grid points, 1455
Internal model control (IMC), 1238–1239
  structure plus integral control, 1239*f*
Internal resistance, bodies with significant, 258–259
International Nanofluid Property Benchmark Exercise, 589
International Performance Measurement and Verification Protocol, 746
International Practical Temperature Scale (IPTS), 1162
International Temperature Scale (ITS), 898
Interpolation scheme, 1512
Intracellular ice formation (IIF), 686
Intraocular pressure (IOP), 1520
Intrinsic semiconductor, 1044
Intrinsic solar to chemical conversion (ISTC) efficiency, 1377
Inverted absorber solar still, 1418, 1418*f*
Inviscid flows, 135, 136
Ion pumps, 862
IOP, *see* Intraocular pressure
IPCE, *see* Incident photon to current efficiency
IPLV, *see* Integrated part load value
IPTS, *see* International Practical Temperature Scale
Irradiation, 343, 366
Irreversibility, 9–10
Irrotational flows, 134
Isentropic efficiency, 78, 919
ISO building standards, in analysis of thermal bridges, 814, 815
Isolated bubble regime, 371
Isoparametric elements, 1481–1483
Isothermal planes method, 810
Isothermal surfaces, 332–333*t*
Isotherms, 1519*f*
ISTC efficiency, *see* Intrinsic solar to chemical conversion efficiency
Iterative solvers, 1479
ITS, *see* International Temperature Scale

J

Jacobi method, 1452
Jerk pump systems, 990
Jet(s)
  condensation on, 1145–1146
  heat transfer to, 1134–1136
  hydrodynamics of, 1135
  impingement, heat transfer with, 301–303
  pumps, 855
  Weber number, 988
Joule heating, *see* Ohmic heating
Joule–Thomson data, 39
Junction temperature, 1036

K

Kapitza resistance, 1050
Kay's rule, 36
Kelvin–Helmholtz instabilities, 1132
Kelvin–Planck statement, 8–9
Kelvin temperature scale, 11
Kelvin–Voigt linear viscoelastic material, 1516
Kinematic linkages, compressor type by
  reciprocating compressors, 824–827
  rotary compressors, 827–829
  screw compressor, 832–835
  scroll compressors, 829–832
Kinetic energy, 6, 21
Kinetic formulations, application in thermal damage studies
  cell death processes, 667–670
  first-order Arrhenius process analysis, 662–667
  modeling cell death processes, 670–673
Kinetics of reactions, in preservation of foods
  C value, 1025
  D value, 1023–1024
  F value, 1024
  process optimization, 1025
  z value, 1024
Kinetic theory, 1496, 1498
Knowledge engineering, 1218
Knudsen diffusion, 426, 1287
Kolmogorov mapping neural network existence theorem, 1224
Kolmogorov scales, 1545, 1554
Kozeny–Carman equation, 150
Kraft pulping, 1277
Krieger–Dougherty model, 550, 551*f*
Kronecker delta property, 1484
Kutateladze model, 1143

L

LAEs, *see* Linear algebraic equations
Lag phase, of bacteria, 1021
Lagrangian–Eulerian methodology, 994, 996, 1148
Lagrangian grid, 1509
Laminar boundary layer, 182, 285, 290–291, 447–448
Laminar duct heat transfer, 315–321
Laminar flow, 306–307, 546, 584, 586
  fully developed flow in tube, 232–239
    extended modified power-law model fluids, 236–238
    generalized approach of purely viscous fluids, 238–239
    Herschel–Bulkley fluids, 232–235
  in noncircular ducts, 239–240
Laminar flow element (LFE), 1197–1198
LANL, *see* Los Alamos National Laboratory
Laplace equation, 136
Laplace's capillary equation, 965
Large eddy simulation (LES)
  case study of channel flow, 1554–1555
  errors in, 1552–1553
  explicit filtering, 1549
  filtering due to numerical discretization, 1547–1548
  filter properties, 1546–1548
  governing equations, 1548–1549
  grid generation, 1554
  Smagorinsky model, 1550–1552
  subgrid stress models, 1550–1553
  wall adaptive local eddy viscosity model, 1552
  of wall-bounded flows, 1545–1556
  wall modeling, 1553–1554
Laser
  argon-ion, 1209
  continuous-wave, 1206–1207
  Doppler anemometry
    particle image velocimetry, 1209–1212
    principle of operation, 1206–1209
    viscosity measurements, 1212
    void fraction measurement, 1212–1214
Laser Doppler velocimetry (LDV), 1177, 1206, 1206*f*, 1207*f*
LAT, *see* Limiting ambient temperature
Latent heat
  of melting, 393
  of PCMs, 790, 792
Latent heat storage (LHS) materials, 781, 792
Lattice Boltzmann method (LBM), 587, 1443
  accuracy and stability, 1502
  Boltzmann equation, 1498–1499
  boundary conditions
    no-slip wall, 1502
    periodic BCs, 1503
    slip wall, 1502
    velocity BCs, 1503
  Chapman–Enskog procedure, 1498–1499
  D2Q9 lattice as example, 1499–1502
  entropic LBM, 1502
  example problems, 1503, 1504*f*, 1505*f*
  introduction, 1496
  kinetic theory, 1498–1499
  lattice gas cellular automaton, 1496–1498
Lattice gas cellular automaton (LGCA), 1496–1498, 1497*f*, 1503
LBL, *see* Line-by-line calculations
LBM, *see* Lattice Boltzmann method
LCA, *see* Life-cycle assessment
LCC, *see* Life-cycle cost
LCCE, *see* Life cycle conversion efficiency
LDV, *see* Laser Doppler velocimetry
Legionnaires' disease, 874
Leidenfront point, 375
LEM, *see* Lumped element modeling
Length effect, 512
Leonard stress, 1551

LES, *see* Large eddy simulation
Lethality, 1024
Level set-based discrete forcing IB methods, 1508, 1512–1513
Level set method (LSM), 1512–1513, 1537–1542
  application, 1540–1542
  functions and equations, 1538–1539
  numerical methodology, 1539–1540
  sharp-interface treatment of energy equation, 1539
Lewis–Randall rule, 54
Lewis relation, 908
LFE, *see* Laminar flow element
LGCA, *see* Lattice gas cellular automaton
LHPs, *see* Loop heat pipes
LHS materials, *see* Latent heat storage materials
Life-cycle assessment (LCA), 106
Life cycle conversion efficiency (LCCE), 1393, 1411
Life-cycle cost (LCC), 730
Lift
  coefficient, 187, 207
  drag due to, 188
Lifting flat plate, 207*f*
Lighting, energy-efficient, 734
Light pipe technologies, 739
Lightweight thermal mass, 782
Lignin, 1277, 1279
Limiting ambient temperature (LAT), 847
Linear algebraic equations (LAEs), 1445, 1454
Linear chains, 553–554
Linear equation solver, 1478–1479
Linear flow meters, 1192–1198
Linear fouling, 606, 607*f*
Linear interpolation, 40
"Linearizing the BTE" simplification, 1041
Linear thermal bridges, 812
Linear thermal transmittance method, 811–813, 813*f*
Linear VOF interface reconstruction algorithm (LVIRA), 1533
Line-by-line calculations (LBL), 366
Lipid oxidation, 1018
Liquid(s)
  additives for, 935, 946
  breakup characteristics, 988
  crystals, 1175*f*, 1176
Liquid atomization and spraying, 986–1000
  atomizer design considerations, 988
  atomizer types, 988–994
  spray characterization, 986–988
Liquid flooded compression, 820, 820*f*
Liquid/gas interface, 133
Liquid-to-liquid exchangers, 502
Liquid–vapor/bubble-slug system, 960
Liquid–vapor phase-change processes, 368
Living wall system, design, 770*f*
LMTD, *see* Logarithmic mean temperature difference
LMTD, *see* Log mean temperature difference
Load calculations, 909
Load vectors, 1475
Local condensation, 806, 815
Local coordinates, 1460, 1460*f*
Lockhart–Martinelli correlation, 214, 217
Logarithmic mean temperature difference (LMTD), 463
Log mean temperature difference (LMTD), 885
Log phase, of bacteria, 1021
Long, cylindrical systems, at steady state, 253
Long-wavelength radiation model, 774
Longwave radiation, 910
Loop breaking, 892
Loop heat pipes (LHPs), 958–960, 959*f*
Lorentz force, 591–592
Lorenz number, 1043
Los Alamos National Laboratory (LANL), 957
Loss
  calculations, 158
  coefficients, 157, 870
Louver fins, 525
Low-side compressor, 822
LSM, *see* Level set method
Lumped element modeling (LEM), 1154
LVIRA, *see* Linear VOF interface reconstruction algorithm
Lymphatic fluid loss, 646
Lysholm type compressor, 832

M

Machining techniques, development of advanced, 831
Mach number, 202, 203, 1199
Macrobiofouling, 479
Macroscale continuum approach, for nanofluid flow simulations, 585
Magnetic nanofluids (MNFs), 590–593, 591*f*
Magnetic seeds/nanoparticles, 658
Maillard browning, 1018
Maldistribution, flow, 535
Maltodextrin powder, 1330, 1331*t*
Maple software, 1558
Marangoni convection, 1524
Marching solution, 261
Mars–Van-Krevelen mechanism, 1380
Mass
  concentration, of species, 418
  conservation of, 15–16, 123–124, 125
  convection, 437–449
  flow rate, 15, 97*t*
  fractions, 36
  species, local concentration of, 417
  utilization, 460
Mass conservation (LS advection) equation, 1538–1539
Mass median diameter, 986
Mass transfer, 417–449, 1285
  concentrations, velocities, and fluxes, 417–422
  conductance, 446*f*
  convective heat and, 439–441
  diffusion in moving medium
    diffusion with one component stationary, 434
    droplet combustion, 436–437
    droplet evaporation, 435–436
    heterogeneous combustion, 434–435
  diffusion in stationary medium, 427–430
    diffusion in porous catalyst, 432–434
    steady diffusion through plane wall, 427–430
    transient diffusion in semi-infinite solid, 430
    transient diffusion in slabs, cylinders, and spheres, 431–434
  driving force, 437, 444
  during drying, 1311
  mass convection, 437–449

Mass transfer (*cont.*)
analogy between convective heat and mass transfer, 439–441
dimensionless groups, 439
high mass transfer rate theory, 444–445
low mass transfer rate theory, 438–439
mass and mole transfer conductances, 438
simultaneous heat and mass transfer, 441–443
variable property effects of high mass transfer rates, 446–449
wet- and dry-bulb psychrometer, 443–444
mechanisms of diffusion, 422–426
rates, high, 446–449
simultaneous heat and, 441–443
species conservation equation, 426–430
Material(s)
degradation limits, 1080
drying of, *see* Drying of materials
freezing, 1327
thermal conductivity ranges, 716
wide-band, 1176
Mathematica (mathematical software), 1558
Mathematical software, 1558
MATLAB® software, 1558
Matter, properties of, 3
Matthiessen's rule, 1042, 1045
Maxwell–Boltzmann distribution function, 1042, 1499
Maxwell–Garnet model, 554
Maxwell model, 579
Maxwell relation, 27, 37
Mazur equation, 686
MC method, *see* Monte Carlo method
McQuiston correlations, 1229
MD, *see* Molecular dynamics
Mean beam length(s)
average, 363
for radiation from gas volume, 364*t*
Mean effective pressure, 79
Mean fluid temperature, 306
Mean temperature difference (MTD) approach, 461–467, 495, 495–496*f*, 511*t*, 512, 513*t*
example applied to double pipe heat exchanger, 467–471
Measuring junction, 1165
Mechanical aids, 935
Mechanical blower, 861
Mechanical cleaning, 625
Mechanical defibering process, 1278
Mechanical-draft cooling tower, 869
Mechanical pumps, 861
Megerlin method, 401
Melting process, in glass manufacturing, 1355–1359
MEMS, *see* Microelectromechanical systems
Merkel's assumptions, 865
Merkel's method, 864, 872
Mesh refinement, 1488–1489
Meso/microscale approach, for nanofluid flow simulations, 585, 586
Mesophilic microorganisms, 1022
Metabolic heat generation, 642
Metacenter, 120
Metacentric height, 121
Metal(s)
electrons properties in, 1042–1043
framing, 811
polished, 344
sheathing, 811
Meters
Coriolis, 1198
electromagnetic flow, 1198
linear flow, 1192–1198
positive displacement meters, 1195
turbine flow, 1195
ultrasonic flow, 1197
venturi, 1192
vortex shedding, 1195
Methane, ideal gas model for, 71
Method of combination, 141
Metzner–Reed Reynolds number, 239
Microaerophiles, 1022
Microbiological integrity, 614
Micro-convection, 552
Microelectromechanical-fabricated device, 1063, 1064*f*
Microelectromechanical systems (MEMS), 1468
Microelectronic manufacturing, 1244
Microencapsulated paraffinic hydrocarbons, 785
Microencapsulated phase change materials, 792
gypsum board containing, 793*t*
Microencapsulation, 782, 1319
Microfin tubes, 949
Micro heat pipes, 964–965, 964*f*
Microorganisms, 1021–1023
Microscale interaction, 579, 580*f*
Microwave, 1030
dryers, 1324–1325, 1324–1325*f*
energy deposition, 254
frequencies, 656
heating systems, 1032
MIE, *see* Momentum integral equation
Milk stone, 613
Mineral fillers, 1300
Miner's superposition rule, 1117
Miniature sensors, 1033
Minimum heat flux, 385
Mist flow evaporation, post-dryout, 379
Mitigate fouling, 611, 612*t*, 614
Mixed boundary conditions, 1448
Mixed flow arrangement, 1028
MNFs, *see* Magnetic nanofluids
Model(s)
Arrhenius, 659–662
Chen–Holmes, 644–645
continuum, 644–645
Dalton, 46, 49
dynamic, 1235–1237
energy use baseline, 743–744
enzyme deactivation, 666–667
Hadamard, 1138
ideal gas, 27, 43–45
incompressible, 2, 40
kinetic, 659–660
Kutateladze, 1143
Pennes', 643
PRInceton Scorekeeping Method (PRISM), 746
regression, 746–747
simplified Weinbaum–Jiji, 646–647
small artery, 647–648
time variant, 747

vasculature-based, 645–648
Weinbaum–Jiji, 645–646
Wissler, 653–656
Modeling error, in large eddy simulation (LES), 1552
Modeling sprays, 994–997, 998*f*
Modified Problem Table Algorithm (MPTA), 881, 882*t*
Modified zone method, 811
Modular breathing wall, 764
Moisture content, in paper, 1299
Moisture evaporation, 1332
Moisture migration, 1018
Molar concentration, 418
Molded package electronics, 1079
Mold formation, effect of thermal bridging, 802
Molding compound, 1088–1090*f*
Molds, 1022
Molecular dynamics (MD), 1054
simulation, 587–588
Molecular gases, radiative properties of, 352
Mole fractions, 35–36
Mole transfer conductance, 438
Mollier chart, 1308*f*
Momentum conservation, 124–125, 127
equation, 1528
Momentum equation, 586, 587
Momentum integral equation (MIE), 181
Money, time value of, 102
Monte Carlo (MC) method, 1054, 1057, 1060
Montreal Protocol on Substances that Deplete the Ozone Layer, 923
Moody chart, 170, 172
Moore's law, 1036, 1072
Motion and potential flow, equations of, 123–138
analysis of rate of deformation, 129
boundary conditions, 133
conservation of energy, 125–127, 131–133
conservation of mass, 123–124
continuity equation, 127
conservation of momentum, 124–125, 127–129
integral relations for control volume, 123
inviscid irrotational flow, 136–138
Navier–Stokes equations, 131
relationship between forces and rate of deformation, 129
Reynolds transport theorem, 123–125
stream function, 135–136
vorticity in incompressible flow, 133–135
Motion, Newton's second law of, 154
Motors, 734
efficiencies of, 735*t*
replacement, 745
typical efficiencies of, 734
Moving boundary problems, 393
Moving interfaces, continuum modeling of, 1524–1526
Moving surfaces, 189
MPTA, *see* Modified Problem Table Algorithm
MTD approach, *see* Mean temperature difference approach
Mullions, 806
Multicomponent system, 52–53
Multieffect active solar still, 1393
Multieffect, in solar still, 1414
Multifoil vacuum insulation, 722*t*
Multilayer living walls, 749, 777–778*t*
background, 768–770
building materials properties, 774*t*
configurations coupled to full-scale building, 774*f*
thermal models for advanced façade systems, 771–776
Multilayer perception, 1221
Multi-orifice synthetic jet, 1160
Multiparticle interactions, 1005
Multiphase flows
continuum modeling of moving interfaces, 1524–1526
enthalpy method
application, 1532–1533
formulation, 1529–1531
microscopic model, 1531–1532
numerical methodology, 1531
fundamentals, 209–212
gas–liquid two-phase flow
flow regimes, 212–214
pressure drop, 217
void fractions, 214–217
gas–solid, liquid–solid two-phase flows, 217–222
flow regimes, 218
minimum fluidization, 218–219, 219*f*
pneumatic conveying, 220–222
terminal slip velocity, 220
level set method (LSM), 1537–1542
application, 1540–1542
functions and equations, 1538–1539
numerical methodology, 1539–1540
sharp-interface treatment of energy equation, 1539
mathematical formulation
nondimensional governing equations, 1528–1529
principles of volume averaging, 1526–1528
volume of fluid (VOF) method
application, 1536–1537
equation, 1534
numerical methodology, 1534–1536
Multiple-event paints, 1175
Multitube heat exchanger, 1028
Multiwalled carbon nanotubes (MWCNTs), 548, 556, 558*f*
Multizone air duct system, 928
Mushy zone, 396
MWCNTs, *see* Multiwalled carbon nanotubes
Mycotoxins, 1022

N

Nanodroplet–vapor–air mixture dynamics, 593–597, 595*f*
Nanofluids
applications, 590–597
in heat exchangers, 545–570
challenges in heat exchangers, 569–570
convection heat transfer in conduits with, 575–598
discrepancies among enhanced heat transfer studies using, 589–590
entropy generation, 588–589
future work, 597–598
heat transfer in, 584–588
magnetic nanofluids (MNFs), 590–593, 591*f*
nanodroplet–vapor–air mixture dynamics, 593–597
performance in heat pipe evaporators, 566–569
on performance of single and double slope solar still, 1422–1423
preparation
and characterization, 547–549

Nanofluids (*cont.*)
stabilization, 576–578
synthesis, 575–576
production methods, 547
properties of, 549–557, 550*f*
results
in double-pipe helical heat exchanger, 557–560
in PHE, 560–566
thermal conductivity of, 545–554, 557, 563–564, 578–581
thermal nanofluid flow in double-tube heat exchanger, 590
thermophysical properties of, 578–584
viscosity of, 545–554, 557, 563–564, 581–584
Nano particles, 946
Nano-scale coatings, 1003
Nanowires, 1066, 1379
Naphthalene experiments, 1185–1186
Nappe, 174
Narrow-band materials, 1174
National Energy Code of Canada for Buildings (NECB), 814
National Institute of Standards and Technology, 23
Natural convection heat transfer, 588
Natural draft counterflow unit, 865
Natural gas, 1338
Navier–Stokes (NS), 994, 997
equations, 131, 1132–1133, 1141, 1443, 1496, 1499, 1507, 1509, 1513, 1528
NCGs, *see* Noncondensable gases
Nebulizers, 994
NECB, *see* National Energy Code of Canada for Buildings
Necroptosis, 667, 668
Negative thixotropy, 226
Net positive suction head (NPSH), 855, 857, 862
Net radiation method, 352
Network
analogy, 356, 366
energy-efficient, 1262
Network design and evolution, 888–892
heat exchanger networks (HENs), 875, 876*f*, 877*f*, 884–886, 888–891, 892*f*
integrating with other energy devices, 892–894
Neumann boundary conditions, 1448, 1512
Neumann solution, 398, 399
of two-phase Stefan problem, 400
Neutral velocity field, 1012
New-Newtonian flows, 223–244
classification of
Newtonian fluid behavior, 223–225
non-Newtonian fluids, 225–228
constitutive equation, 229–232
Newtonian-fluid mixtures, 581
Newtonian fluids, 223–225, 569, 1019
Newton–Raphson method, 1480
Newton's law of cooling, 258, 431
Newton's second law of motion, 154
N identical partially covered PVT-CPC active solar still
energy balance equations, 1402–1405
useful energy gain for, 1401–1402
Nikuradse equation, 240
NILM, *see* Nonintrusive load monitoring
Nissan drying phases of paper, 1288*f*
N-layer heat diffusion model, 1061
NMD, *see* Number median drop diameter
Nocturnal distillation, 1390
"No foul" approach, 617, 617*t*
Noise generation, fans and, 858
Noncircular ducts
laminar flow in, 239–240
turbulent flow in circular and, 240–243
Noncondensable gases (NCGs), 956, 965
Nonconductors, 345
Noncontinuum effects, 1004
Nondimensional governing equations, 1528–1529
Nondominated sorting genetic algorithm (NSGA), 775, 776*f*
Nonintrusive load monitoring (NILM), 730
Nonlinear equation solver, 1479–1481
Non-Newtonian fluids, 130, 315, 1019
classification of, 225–228
convection heat transfer in, 315–338
viscous time-independent, 321–322
Nonprogrammed cell, 667, 674
Nonrepeating thermal bridges, 812
Nonspherical particles
heat transfer to, 1007–1008
shapes, 1004–1005
Normal shock wave, 193–195
North America, energy performance standards in, 814
North American Energy Measurement and Verification Protocol, 746
No-slip boundary conditions, 1502, 1503*f*
Nozzle
designs, 992*f*
flow, 201, 1189, 1190–1192
flow meters, 1191*f*
NPSH, *see* Net positive suction head
NS, *see* Navier–Stokes
NSGA, *see* Nondominated sorting genetic algorithm
NTU, *see* Number of transfer units
Nucleate boiling, 371, 385
Nucleation, 394
Nukiyama–Tanasawa distributions, 987
Number density, 417
Number median drop diameter (NMD), 986
Number of transfer units (NTU), 461
NTU-effectiveness method, 605
ε-NTU method, 511*t*, 513*t*
Numerical discretization error, 1553
Numerical integration, 1487–1488
Nusselt numbers, 144, 327, 444, 469, 546, 559, 575, 584, 948, 1005, 1006, 1007, 1008, 1008*f*, 1013*f*, 1133, 1157, 1158*f*, 1219, 1228
average, 323, 334*t*
free convection average, 323*f*
fully turbulent heat transfer, 265
heat transfer, 389
laminar flows, 308
Newtonian, 315
Nu values, 563

O

Oblique shock
characteristics, 205*f*
waves, 204–206
on wedge, 204*f*
Occupational Safety and Hazard Administration (OSHA) guidelines, 721

Octane, combustion of liquid, 61
ODP, *see* Ozone depletion potential
Office building
  internal heat gain level for, 742*t*
  typical floor plan of audited, 740
Office equipment, 734
Offset strip fins, 525, 525*f*
Ohmic heating, 254, 1021, 1031, 1032
Ohnesorge number, 989
OHP, *see* Oscillating heat pipe
Oil-flooded screw compressor, 845*f*
Oil flooding, 832–833, 849
OIP, *see* Organic–inorganic perovskite
Oldham coupling, 831
ONB, *see* Onset of nucleate boiling
1-D fin efficiency, 517
One-dimensional integrals, 1487
One-step preparation, to preparing nanofluids, 575
Onset of nucleate boiling (ONB), 371
Opaque medium, 367
Opaque surfaces, radiative exchange between, 343–357
Open cavities, correlations for, 271–273
Open channel flow, 170–176
  critical flow, 170–172
  definition, 170
  gradually varied flow, 174–176
  hydraulic jump, 172–173
  uniform flow, 170
  weirs, 173–174
Open drive compressors, 824
Open feedwater heater, 21
OpenFOAM software, 1558
Open-source CFD codes/software, 1558
Open-source mathematical software, 1558
Open-type compressor, 919
Operating lines, 872
Optical phonons, 1046
Optical probe, 1214
Ordinary annuity, 104
Organic–inorganic perovskite (OIP), 1379
Organic phase changing materials, 782, 783*t*
Organ tissues, temperature dependence of, 637–638
Orifice(s)
  characteristics of, 1190*t*
  flow coefficients for concentric, 1191*f*
  flow through, 157
  geometry and pressure tap locations, 1190*f*
  plates, 1189
Orsat analyzer, 56
Oscillating heat pipe (OHP), 566, 567*f*
OSHA guidelines, *see* Occupational Safety and Hazard Administration guidelines
Osmophilic microorganisms, 1022
Oval shape orifices, 1159
Overall heat transfer coefficient, 262
Oxidation resistance, 1085
Ozone, 924
Ozone depletion potential (ODP), 924

P

Packaged compressors, 843
Packing capability line, 873
Packing factor (PF), 1393
Pagers, 1111
Paints, temperature-sensitive, 1173–1174
Paper making, thermal aspects of
  composition/grades/properties, 1281–1284
  drawing of paper machine, 1275*f*
  dryer fabrics, 1292–1293
  drying of paper, 1285–1286
    global changes in paper structure during, 1297
    heat of desorption, 1298–1300
    physical properties, 1297–1302
    specific heat capacity, 1298
    thermal conductivity, 1300–1302
  global heat transfer coefficient on drying cylinder, 1289–1290
  hood/ventilation/heat recovery, 1295–1297
  organization of cylinder dryers, 1291–1292
  process
    cellulosic fibers and behavior during, 1277–1281
    finishing and converting operations, 1275–1277
    from pulp disintegration to drying and reeling, 1270–1275
    stages of drying process, 1286–1289
  steam and condensate systems, 1293–1295
Parallel path method, 809–810, 809*f*
Parallel plates, vertical, 323
Parallel pumps, selection of, 167*f*
Partial-array chip-scale package, 1085
Partial differential equations (PDEs), 1443, 1445, 1454
Partial pressures, 1306
  Dalton's law of, 419
Particle
  clouds, radiative properties of, 361–362
  cluster deposition, 552
  density, 1018
  gas convection, 388–392
  image velocimetry (PIV), 1209–1212
  tracking, 1208–1209
Particle image velocimetry (PIV), 1177, 1209, 1210*f*
Particle-source-in-cell method, 1005
Particle velocity distribution function, 1500–1501
Particulate fouling, 536, 609
Passive heat enhancement techniques, 933
Passive solar stills, 1389
  double slope passive solar still, 1420–1422
  modified single-slope solar still, 1414–1420
  nanofluids effect, 1422–1423
  thermal–electrical solar still with boosting reflector, 1423, 1424*f*
  wick-type solar still, 1423–1426, 1425*f*
Pasteurization, 1031–1032
Path relaxation, 892
Paulapuro's classification of paper, 1281
PBGA, *see* Power-plastic ball grid array
PCB, *see* Printed circuit boards
PCD, *see* Pulse combustion drying
PCM, *see* Phase change memory; Phase change materials
PDAs, *see* Personal digital assistants
PDEs, *see* Partial differential equations
PEC process, *see* Photoelectrochemical process
PEC, *see* Purchased equipment cost
Peclet number, 763
Peierls BTE, 1055
Pennes' model, 643
Penta-diagonal coefficient matrix, 1451

Perfusion, 649–650*f*
resolution, 648
temperature dependence of, 649–651
thermal measurements of, 648–651
Periodic boundary conditions, 1503
Perkins, Jacob, 957
Perkins, Ludlow Patton, 957
Perkins tube, 957
Perlite powder insulation, 720–721
Persistent bubble, 1145
Personal digital assistants (PDAs), 1111
Perturbation expansions, 401
PF, *see* Packing factor
pH, 1017–1018, 1022
Phase change materials (PCMs), in buildings, 749, 770–773, 1120
distributed and localized applications, 787–788, 788*f*
enhanced gypsum boards, 786, 787, 788*f*
examples of building applications, 792–793
phase change enthalpy values and transition temperatures for, 782*f*
properties of commonly used PCMs, 781–785
slurry systems, 792
thermal and energy performance analysis
laboratory evaluations of, 789–790
numerical tools for, 791–792
utilization
in building envelope and interior fabric, 785–787
in building space-conditioning and water heating systems, 789
Phase change memory (PCM), 1036
Phase-change processes, 368–413
boiling, 368–371
condensation, 380–385
hybrid methods, 412
improved temperature-based equivalent heat capacity method, 413
internal convective boiling, 376–379
level set methods, 405
melting and freezing, 393–413
numerical approaches, 404–413
phase field approach, 405
pool boiling, 370–375
Phase diagrams, 395–396
Phase/Doppler instruments, 987
PHE, *see* Plate heat exchangers
Phonon BTE, 1055
Phonon hotspot, 1058, 1071
Phonon–phonon scattering rates, 1057
Phonons, 1040
Phonons properties, in crystals, 1045–1049
Photoabsorption, 1375
Photodetectors, 1209
Photoelectrochemical (PEC) process, 1370
decoupling, 1380
physicochemical principles of, 1371–1375
Photon, 367
Photon Monte Carlo (PMC) method, 365
Photopotentials, 1380–1381
Photovoltaic thermal (PVT) collectors, 1390
Photovoltaic thermal-compound parabolic concentrator (PVT-CPC) collector, 1390, 1399, 1400*f*
PHPs, *see* Pulsating heat pipes
PHYSIBEL Voltra tool, 814
Physical properties, in thermal processing, 1018
PID controller, *see* Proportional–integral–derivative controller
Piezoelectric actuators, 1154
Pigging, 625
Pigments, 1273
Pilot paper machine
drying section of, 1274*f*
forming section of, 1272*f*
press section of, 1273*f*
Pinch technology
cost targeting, 884
cost optimization, 886–887
targeting for heat transfer area, 885–886
targeting for multiple utilities, 887–888
targeting for number of heat exchangers, 885
energy targeting
algebraic analysis, 881–884
golden rules of pinch analysis, 884
graphical analysis, 878–881
features of, 877–878
motivating example, 875–877
Pipe
ceramic fiber, 725
empirical representation for naturally rough, 152
fiberglass, 725
roughness, measure of, 155
solutions for flow and convection in smooth, 143–144
Pipe systems, hydraulics of, 154–169
basic computations, 154–155
equations, 154–155
fluid friction, 155–157
minor losses, 157–158
pipe design
external loads, 158–160
limiting velocities, 160
pressure class guidelines, 158
pump selection, 165–167
valve selection, 160–165
air valves, 165
check valves, 164
control valves, 160–163
Pitot probes, 1201–1202
Pitot–static probe, 1201*f*
PIV, *see* Particle image velocimetry
Planar heat flux (Schmidt–Boelter) gauge, 1176, 1178*f*
Planar synthetic jets, 1154
Planck's law, 341, 367, 1173
approximation of, 343
general behavior of, 341
Plane wall
steady diffusion across, 427–430
at steady state, 252–253
Plasma heat transfer
defining features of, 1004–1005
electron–ion transport effect on, 1004
multiparticle interactions, 1005
noncontinuum effects, 1004
nonspherical particle shapes, 1004–1005
variations in thermophysical properties, 1004
Plasma spheroidization, 1003
Plastic flammability standards, 1086
Plastics, Bingham, 229, 239, 331, 333*t*, 336*t*, 1020
Plate-and-frame PHE, *see* Gasketed PHE

Plate-fin exchanger
rating/sizing problem for crossflow, 529–535
Plate-fin heat exchangers, 502–503, 502–503*f*, 520
Plate heat exchangers (PHE), 522, 1032
nanofluids results in, 546, 560–566
Plate-type heat exchangers, 506–508
PLCs, *see* Programmable Logic Controllers
Plenum chamber, 441
PMC method, *see* Photon Monte Carlo method
Pneumatic conveying, 220–222
horizontal, 220*f*
vertical, 221*f*
Pneumatic dryers, 1321–1322, 1322*f*
Pneumatic robotics, 736
Poddar plot construction, 622
Point thermal bridges, 812, 813
Poisson equation, 1059, 1513
Poisson ratio, 1517
Polar coordinates, 137
Polar molecules, 368, 385
Polycarbonate, 1107–1108
Polyethylene oxide (Polyox), 243–244
Polymerization, 479
Polytropic exponent, 45
Polytropic processes, 45–46
Ponding/puddle regime, 1294
Pool boiling, 370–375, 385, 945–947
Popcorn effect, 678
Porosity, 1300
Porous catalysts, 432–434
Portable computers, heat transfer in, 1110*f*
Positive coefficient rule, 410
Positive-displacement pump, 853, 855–856
Post-dryout mist flow evaporation, 379
Postmortem necrosis, 674
Potential energy, 6, 21
Potential flow, 123–138, 136, 137*f*
Potential theory, 136
Powder insulation, 720*f*, 720*t*, 721*f*
perlite, 720–721
thermal conductivity, 720
Powder systems, 719–721
Power
absorption, 1030
cycles, 8
density, 1038*f*
dissipation, 1036, 1071–1073, 1095, 1107*f*
law, 317*t*, 1251
constitutive equation, 229–232
correlation, 1251
equation, 315
fluids, 320, 323
penetration, 1031
plants, 858
PowerFLOW software, 1558
Power-plastic ball grid array (PBGA), 1100
Prandtl number, 144, 268, 294, 440, 1181, 1185, 1186, 1219, 1228
Pray foam, 805
Precipitation fouling, 536
Predictive control, 1238
Pressure
atomizers, 988
class guidelines, 158
coefficient, 207
contours, 1520*f*
diffusion, 423
drag, 185–186, 188
drop, 210, 217
gradient, 181
effect of, 182–184
flow over flat plate with zero, 291–293
flows with, 283–284, 289–290
frictional, 217
zero, 183
path length, 357
as source term, 1465–1466
swirl atomizers, 991
Pressure drop, 479–480, 565, 603, 607–608
fin efficiency and extended surface efficiency, 513–522
heat transfer exchanger and, 509–522
performance, 527
Pressure–enthalpy diagram
for vapor compression refrigeration cycle, 917*f*
Pressure Poisson equation, 1510
Pressure probes, 1199
alignment effects, 1200
setup, 1199*f*
static pressure taps, 1201
streamline curvature, 1201
supersonic flow around, 1200*f*
turbulence effects, 1201
wall effects, 1201
Pressure ratio, 919
Printed circuit boards (PCB), 1081–1085, 1081*f*, 1084*f*
Printed circuit heat exchangers, 508
PRISM model, 746
Probe(s)
beam, 1062
design, 634–635
multicomponent, 1203–1204
Problem Table Algorithm (PTA), 881
Process
performance, restoring, 614
simulation software, 85
Product
changeover, 614
stream, 86
Productivity, of solar still, 1411–1412
Profile shaping, 188
Pro-flux flow, 763
Programmable Logic Controllers (PLCs), 168
Programmed cell death, 667
Property data, for reactive systems, 58–65
Property relations and data, 23–55
basic relations for pure substances, 23–27
Maxwell relations, 24–25
evaluating Δh, Δu, and Δs, 37–39
fundamental thermodynamic functions, 39
generalized charts for enthalpy, entropy, and fugacity, 50–52
ideal gas model, 43–45
ideal gas mixtures, 46–49
moist air, 49
polytropic processes, 45–46
multicomponent systems, 52–53
chemical potential, 54–56
ideal solution, 54–55

Property relations and data (*cont.*)
P-v-T relations, 27–37
compressibility charts, 32–34
equations of state, 34–35
gas mixtures, 35–37
graphical representation, 31–32
P-v-T surfaces, 29–31
specific heats and other properties, 25–27
thermodynamic data retrieval, 40–43
Proportional–integral–derivative (PID) controller, 775, 1225
disturbance rejection, 1241
step change in set point, 1239–1240
P rule, 90
Pseudoplastic fluids, 225, 230, 325*t*, 327*t*, 329–330*f*
Pseudoreduced specific volume, 32
PSI Therm tool, 814
Psychrometer, 443–444, 900
Psychrometric chart, 901–902, 902*f*, 1308
adiabatic mixing of two streams, 904–907
cooling and dehumidifying air, 902–903
heating and humidifying air, 903–904
Psychrometry, 1306
Psychrophilic microorganisms, 1022
Psychrotrophic microorganisms, 1022
PTA, *see* Problem Table Algorithm
Pulsating heat pipes (PHPs), 960–964, 960–961*f*
Pulsations, 824
Pulse combustion drying (PCD), 1328, 1329*f*
Pulverized coal, 361–362
Pump(s), 18
applications, 857
beam, 1062
data for preliminary selection of, 854*t*
positive-displacement, 853, 855–856
selection, 165–167
filling and flushing, 168
problem of, 166
for single pump, 166*f*
system monitoring and control, 168
velocity head, 853, 855
Pumping limitation, air-cooled heat exchangers, 496
Pumping power, 565, 566*f*
Pumps and fans, 853–862
centrifugal and other velocity head pumps, 854–855
fans, 858–861
positive-displacement pumps, 855–856
pump/flow considerations, 856–857
pumps, 853–854
Purchased equipment cost (PEC), 97*t*, 102
PVC pipe, guidelines for filling, 168
P–V diagram
of reciprocating compressor, 825, 826*f*
PVT collectors, *see* Photovoltaic thermal collectors
PVT-CPC water collector solar still, 1427–1429
PyFR software, 1558
Pyroptosis, 667, 668

Q

QFP, *see* Quad flat pack
Quad flat pack (QFP), 1085
Quadratic Upstream Interpolation for Convective Kinematics (QUICK) scheme, 1513
Quadrilateral elements, 1485
Quasi-enthalpy method, 412
Quasi-stationary approximation
for convective boundary condition, 402–403
for imposed flux, 402
for imposed temperature, 401–402
QUICK scheme, *see* Quadratic Upstream Interpolation for Convective Kinematics scheme

R

Radial-flow fans, 858
Radial Gaussian system, 1224
Radially split casings, 855
Radiation, 341–367
blackbody, 341–343
boundary condition, 1471
convective system, 1364, 1364*f*
devices, 1172, 1173
longwave, 910
nature of thermal, 341
radiative exchange between opaque surfaces
emissivity, 343
net radiation method, 352–357
radiative properties of surfaces, 343–347
view factors, 347–352
radiative exchange within participating media, 357–366
heat exchange in presence of participating medium, 362–364
radiative properties of molecular gases, 357–361
radiative properties of particle clouds, 361–362
resistance, 718
shields, 356
tempering, 1363–1364, 1363*f*
Radiation-absorbing pad, 1423
Radiative exchange
between opaque surfaces, 343–357
with participating media, 357–366
Radiative heat transfer, 1367
Radiative loss heat coefficient, 1396–1397
Radiative properties
of particle clouds, 361–362
of soot, 362
Radiative resistance, 356
Radiative transfer equation (RTE), 1358
solution methods, 364–365
Radio frequency, 1030
dryers, 1324–1325
Radiosity, 352, 367
Rankine cycle, 76
RANS models, *see* Reynolds averaged Navier–Stokes models
Rate limitation, air-cooled heat exchangers, 497
Rate of deformation
analysis of, 129
relationship between forces and, 129
Rate process description theory of, 659–662
Rayleigh breakup, 989*f*
Rayleigh curve, 197
Rayleigh number, 146, 439
Rayleigh–Plateau instability, 999, 1000
Rayleigh regime, 988
RDIMM, *see* Registered Dual In-Line Memory Module

Reaction
equilibrium, 64–66
product formation rates, kinetics models of, 659–660
velocity, overall, 660
Real time pricing (RTP), 729
Rear head, 476
Reciprocating compressors, 824–827, 919–920
Reciprocating retort, 1027
Reciprocity rule, 350
Recirculatory vortex, 1007
Rectangular orifices, 1159
Recuperator, 511
Redlich–Kwong equation, 35
Reference
junction, 1162, 1167
state schemes, 449
Refinery distillation unit, 617*f*
Refining, 1271
Reflectance, 367
Reflectivity, 344
Refractories, 344
Refrigerant(s)
ASHRAE classification of, 927
compressors, types of, 819
flammability and environmental characteristics of, 927*t*
performance characteristics of, 926*t*
physical properties of, 925*t*
replacements, CFC, 839
selection, 923–928
Refrigeration, 915
Refrigeration compressors
cross-sectional drawing of, 825*f*
performance and capacity range for positive displacement, 824*f*
Regenerative solar still, 1392
Regenerators, 504–506, 522
Regime of slugs and columns, 371
Registered Dual In-Line Memory Module (RDIMM), 1110*f*
Regression models, 746–747
Reinitialization equation, 1539
Relative cost difference, 92
Relative humidity (RH), 49, 900, 1283, 1307
Relaxation time, 1499
approximation, 1041
Repeating thermal bridges, 812
Required net positive suction head (NPSHR), 862
Resistance temperature detectors (RTD)
calibration and drift, 1172
electrical characteristics, 1170–1171
physical characteristics, 1170
steady-state self-heating, 1171–1172
thermal characteristics, 1171
Resistance thermometer, 1170, 1171*f*
Resolution of design, air-cooled heat exchangers, 498
Restricted dead state, 68
Restriction flow meters, 1188–1190, 1215
Retina, 681
Retorting sterilization, 1033
Retorts, heating equipment
classification based on continuity of process, 1025–1026
classification based on mixing, 1026–1027
Revenue-requirement method, 97*t*
Reversible adiabatic flow, 193
Reynolds averaged Navier–Stokes (RANS) models, 1545
Reynolds-averaging procedure, 994, 995
Reynolds number, 142, 158, 235, 271, 524, 529, 531, 546, 559, 560*f*, 936, 1013, 1139, 1154, 1157, 1158*f*, 1189, 1228, 1498, 1510
Reynolds transport theorem, 123–125
RH, *see* Relative humidity
Rheology, in preservation of foods, 1018–1020
Rheopexy, 226, 228*f*
Rigid-body rotation, 122
Rigid/flex interposer chip-scale package, 1085
Rimming regime, 1294
Rising bubble viscometers, 1212
RMS errors, 1259*t*
RNA transcription enzymes, 674
Robin boundary conditions, 1448
Rolling piston compressor, 824, 827–829, 827*f*
Rosin–Rammler distributions, 987
Rotameters, 1193–1194
Rotary atomizers, 991
Rotary dryers, 1322–1323, 1323*f*
Rotary kilns
axial velocity of solid particles, 1340–1341
basic solution methodology, 1349–1350
conduction heat transfer in refractory wall, 1346–1347
energy transport
in gas, 1348–1349
in solid, 1347–1348
firing of, 1338
food products drying in, 1342
gas convection, 1346
heat transfer in, 1338–1351
processes, 1341
modeling, 1341
phenomenon within solid bed
mixing in solid bed, 1340
transverse bed motion, 1339–1340
problem formulation, 1343–1349
salient features of, 1338
thermal radiation among hot gas, 1343–1346
Rotary regenerator, 504, 506
Rotary retort, 1027
Rotary screw compressor, 921–922
Rotary syphon, 1295
"Rotating" heat pipes, 967
Rotating vane rotary compressor, 828, 828*f*
Rotational viscometers, 549
Rothemuhle regenerator, 506
Roughing, 861
Rough surfaces, 933
RSI value, 750–751
RTD, *see* Resistance temperature detectors
RTE, *see* Radiative transfer equation
RTP, *see* Real time pricing
Runge–Kutta scheme, 1502, 1540
R-value (thermal resistance), 807–809

S

SAC105
elastic modulus of, 1091–1092*t*
ultimate tensile strength of, 1092–1093*t*

SAC305
elastic modulus of, 1093–1094*t*
ultimate tensile strength of, 1094*t*
Safety, in drying, 1333
Sage software, 1558
Saint Venant–Kirchhoff law, 1517
Same average cost, 90
Sanderson-Rocker Arm Motion (S-RAM) mechanism, 827
Satellite holes, 1160
Saturated air, 49
Saturated moist air specific enthalpy air, 899
Saturated moist air specific entropy, 899
Saturated moist air specific volume, 899
Saturated water vapor pressure, 899
Saturation
curve, 872
degree of, 900
pressure, 29
Sauter mean diameter (SMD), 987
SBDFA, *see* Solar distiller with a floating absorber
SCADA, *see* System Control and Data Acquisition System
Scattering time, 1042
SCF, *see* Supercritical fluid
Schmidt–Boelter gauges, 1176
Schmidt number, 424*t*, 425*t*, 440, 1229
Schrödinger wave equation, 1042
Scilab software, 1558
Scraped surface heat exchangers, 1029–1030, 1032
Scraping blades, 1032
Screw compressors, 832–834, 832–835
Scroll compressors, 829–832, 829*f*
gas flow in, 830*f*
Seasonal energy efficiency ratio (SEER), 832
Secondary fining, 1359
Secondary water, 930
Second law efficiency, thermodynamics, 928
Second law of thermodynamics, 8–12
Second mitochondria-derived activator of caspases (SMAC), 668
Sedimentation, 479
SEER, *see* Seasonal energy efficiency ratio
Segmented baffle, 1029
Self-heating effect, 1169
Self-heating, steady-state, 1171–1172
Semiconductor
devices, 1037*f*
electrons properties in, 1043–1045
thermal conductivity of, 1068*f*
thermal parameters, 1092–1102
Semiconductor–electrolyte interface
band bending in, 1372–1373, 1374
charge transfer processes at, 1373–1375
Semi discrete finite-element model, 1476–1477
Semi-hermetic compressors, 823, 823*f*, 829, 834
Semi-hermetic designs, 833
Semitransparent, 367
Sensible heat, 401
"Separate type gravity heat pipes," 957
Series–parallel path method, 810
Settling velocity, 545
Shape factors, 1343–1346, 1345–1346*f*
Shape stabilization, 782
Shape stabilized phase change materials, characteristics of, 793*t*
Shaping air, 990
Sharp-crested weir, 173
Sharp-interface, 1525
formulation, 1539
Sharp-interface level set method, 1540
Shear stress, 183, 622, 1019, 1019*f*, 1203
Shear viscosities, 549
Shell-and-tube exchanger, 624–625
Shell-and-tube heat exchangers, 461, 464–466*t*, 1028–1029, 1032
approximate design method, 480–484
construction features, 473–477
nomenclature of components, 471–473
principles of design, 477–480
TEMA nomenclature for, 474*f*
Shells, 474
Shell-side flow and heat transfer, effect of fouling on, 624–625
Shells-in-series arrangement, 623*f*
Sherwood number, 440
Shivering, 653
Shock waves
normal, 193–195
oblique, 204–206
Shortwave absorbtance, for common enclosure surface treatments, 1119*t*
Short-wavelength radiation model, 774
Shrinkage, 1282, 1297
Side loss heat coefficient, 1396
Silica aerogel, 720
Silicon, 1379
circuits, 1079
Similitude, 139–153
correlation of experimental data and theoretical values, 146–153
evaluation of *n*, 148–149
friction factor for commercial pipes for all conditions, 150–153
intermediate regimes, 149–150
pressure gradient in flow through packed bed of spheres, 150
dimensional analysis, 139–146
free convection from vertical isothermal plate, 144–146
fully developed flow of water through smooth round pipe, 140–143
fully developed forced convection in fully developed flow in round tube, 143–144
general principles, 139–140
Simplified Weinbaum–Jiji model, 646–647
Simpson method, 513
Simulating sprays, 994–997
Simulink environment, 771
Single-basin solar still, 1392
Single-cavity synthetic jet, 1160
Single component system, 54
Single-duct systems, 928
Single-field/fluid formulation, 1526, 1528
Single-layer breathing wall, 749, 760–762, 778
background, 762–763
measured and predicted thermal performance, 764–768
products and technologies, 763–764
Single-pass crossflow exchanger, 534

Single-payment compound-amount factor, 103*t*
Single-phase approach, for nanofluid flow simulations, 585
Single-phase forced convection, 936–945
  active and compound techniques for, 944–945
  performance evaluation criteria for, 942–944
Single-phase free convection, 935–936
Single-phase motors, 860
Single screw compressor (SSC), 832, 835*f*
Single slope solar still, 1388*f*, 1389, 1394, 1394*f*, 1402–1403, 1414–1420, 1415*f*, 1421*f*
Single-step method, to producing nanofluids, 547
Sinking fund factor, 1412
Sizing operations, paper making, 1275
Skin, 682
  burn threshold, 663*t*
  forming, 1318
  friction coefficient, 183*t*
Slack-wire platinum resistance thermometer, 1171*f*
Slip boundary conditions, 1502–1503
Slip fan, 1124
Slug calorimeter, 1176, 1177*f*
SMAC, *see* Second mitochondria-derived activator of caspases
Smagorinsky model, 1550–1552
Small artery model, 647–648
Small outline integrated circuit (SOIC), 1085
SMD, *see* Sauter mean diameter
Smoothened functions, 1538
SNB, *see* Statistical narrow band models
Snell's law, for refraction, 1052
Soft computing, 1217–1218
Software
  and computer codes, 1558
  process simulation, 85
SOIC, *see* Small outline integrated circuit
Solar collector systems, 347
Solar distillation system
  basic principle of, 1387–1389
  passive and active, 1386–1430, *see also* Solar still
Solar distiller with a floating absorber (SBDFA), 1420
Solar drying, 1306
Solar hydrogen
  challenges, 1380–1381
  efficiency metrics/photoreactor design/process design, aspects of, 1375–1378
  introduction, 1369–1371
  model identification and refinement, 1375
  physicochemical principles of PEC systems, 1371–1375
  realizable efficiencies, 1378–1379
  as renewable reductant, 1369–1382
Solar intensity, 1394
Solar irradiation, 353
Solar load, maximum, 1119
Solar radiation, 752, 787, 1387
Solar reflectivity, 753–754
Solar still
  active solar still
    ETC-integrated solar still, 1429
    FPC-integrated solar still, 1426–1427
    PVT-CPC water collector solar still, 1427–1429
  classification of, 1389–1390, 1390*f*
  climatic/operational/design parameters, 1429–1430
  defined, 1387
  distillation process, 1387
  efficiency of, 1407–1410
  electrical exergy efficiency, 1408–1409
  heat transfer in, 1394–1399
    external heat transfer, 1395–1396
    internal heat transfer, 1396–1397
    modified internal heat transfer, 1398–1399
    overall heat transfer, 1397–1398
  historical background, 1390–1394
  with inbuilt/seperate condenser, 1419*f*
  material requirement, 1389
  overall exergy efficiency, 1409
  overall thermal efficiency, 1409–1410
  passive solar stills
    double slope passive solar still, 1420–1422
    modified single-slope solar still, 1414–1420
    nanofluids effect, 1422–1423
    thermal–electrical solar still with boosting reflector, 1423
    wick-type solar still, 1423–1426
  performance parameters
    efficiency, 1407–1410
    energy matrices, 1410–1411
    energy output, 1405–1406
    enviroeconomic parameter, 1413–1414
    exergoeconomic parameter, 1412–1413
    exergy output, 1406–1407
    production cost of water and cost of electricity gain, 1414
    productivity, 1411–1412
  thermal circuit diagram, 1394, 1394*f*
  thermal efficiency, 1407–1408
  thermal exergy efficiency, 1408
  thermal modeling, 1399–1405
Solar to hydrogen (STH) efficiency, 1375, 1376*f*
Solder joint geometry, 1117
Solid(s)
  insulation systems, 721–724
  solubility and permeability of gases in, 428–429*t*
  transient diffusion in semi-infinite, 430
Solid bed
  mixing in solid bed, 1340
  phenomenon within, 1339
  transverse bed motion, 1339–1340
Solid density, 1018
Solidification, 393, 1524
  microscopic modeling of, 1530–1531, 1532*f*
Solid particles, axial velocity of, 1340–1341, 1340*f*
Solution curves, for gradually varied flow, 176*f*
Sommerfeld theory, 1042
Sondheimer model, 1044
Sonic limitation, of heat pipe, 968
Soot, radiative properties of, 362
Soret effect, 553
Sorption isotherms, 1280*f*, 1283–1284, 1309, 1309*f*
Sound, velocity of, 39, 192
Space-conditioning energy consumption, reduction of, 785
Spandrel panel, 806
Spark-ignition engine, 78
Specific enthalpy of dry air, 899
Specific entropy, 25
  of dry air, 899
Specific gravity, 1018
Specific heat, 1020
  capacity, 1298

Specific property, 3
Specific volume, 899
Spectral dependence, 344
Spectral models, 365–366
Spectral value, 367
Specularity, 1056
Specular reflection, 1502, 1503*f*
Speed of sound, 192
Sphere(s)
  eccentric, 275
  flows over, 295
  radius, dimensionless variables in, 328
Spherical particle, heat transfer to, 1005–1007
Spheroidization, plasma, 1003
Spill-return-type nozzle atomizer designs, 992
Spinodal curves, 369
Spinodal limit, 1140
Spiral-fin tubes, 939
Split-ring floating head, 477
Spores, 1021
Spray
  column, 1135
  devices, 991*t*
  drop
    number, 987
    size, 989
  dryers, 1315–1318
  modeling, 996*f*
Spurious invariants, 1498
SSC, *see* Single screw compressor
Stabilization, of nanofluids, 576–578
Stacked plate heat exchanger, 508
Stagnation point, 1007
Standard atmospheric conditions, 898
Standard chemical exergies, 69
Standard reference state, 58
Standard temperature and pressure (STP), 1193
Stanton number, 440
Staphylococcus aureus, 1022
STAR-CD software, 1558
Start-stop frequency, 847
Static cool roof, 757
Static retort, 1026–1027
Stationary compressors, 848
Stationary phase, of bacteria, 1021
Stationary syphon, 1295
Statistical narrow band models (SNB), 366
Steady state, long, cylindrical systems at, 253
Steam
  blanching, 1031
  and condensate systems, 1293–1295, 1294*f*
  function, 135–136
  turbines, 842
Steam-heated dryer cylinder, 1289*f*
Steam infusion heating system, 1032
Steam injection heating system, 1032
Stefan–Boltzmann constant, 342
Stefan–Boltzmann law, 1120
Stefan, Jozef, 393
Stefan number, 1148
Stefan problem, 393, 397–401
Stepless capacity control, 826
Steric stabilization, 577, 577*f*
Sterilization, 1033–1034
  aseptic processing, 1033–1034
  retorting, 1033
STH efficiency, *see* Solar to hydrogen efficiency
Stiffness matrix, 1475, 1482
Stirling cycle, 82
Stirling engine, 83
Stokes law, 545
Stokes number, 1155
STP, *see* Standard temperature and pressure
Straight-fin tubes, 939
Streaming/propagation step, 1496, 1497
Streams
  properties, 876*t*
Stress components, 225*f*
Strouhal number, 1155
Structured boiling surfaces, 946, 946*f*
Structured rough surfaces, 947
Studies of Thermal Injury, 659
SU2 software, 1558
Subgrid stress models, 1550–1552
Submerged objected, hydrostatic forces on, 116–119
Subsonic flow, 192
Suction, 935
Suction mufflers, 824
Suction stroke, in synthetic jet operation, 1153, 1153*f*
Summation rule, 350
Supercalendering, 1276
Supercooled liquids, 1352
Supercooling, 393
Supercritical fluid (SCF), 395
Superheated steam, 1342
  dryers, 1313–1315, 1314*f*
Supersonic flow, two-dimensional, 203–208
Supersonic wind tunnel, 201
Suppression, 624
Surface(s)
  diffusion, 426
  electrochemistry, 1375
  heat flux, uniform, 308–310
  radiative exchange between opaque, 343–352
  radiative properties of, 343–347
  roughness, 937–938
  scraping, 944
  solar absorptivity of selected, 347*t*
  states, on PEC processes, 1381
  temperature
    arbitrary, 288
    uniform, 287–288, 310
  temperature, 539, 611
  tracking, 1525
  vibration, 935, 945
Surface-tension devices, 935
Surrogate error, 1223
Surrogate organism, 1017, 1023
Swine left ventricle, temperature dependence of, 639
Swing check valves, 164
Swirl-flow devices, 934, 940–942, 943*f*, 947
Synthesis, of nanofluids, 575–576
Synthetic jets
  flow characteristics of, 1154–1157
  heat transfer characteristics, 1157–1161
  schematic of operation, 1153*f*
  stages of, 1155*f*
Syphon, 1295

System boundary, 3
System Control and Data Acquisition System (SCADA), 168
System disturbance errors, 1181*f*

T

Tandem-PEC cell, 1373
Tank-type heat exchanger, 1032
Target flow meters, 1195
Target organism, 1022–1023
Taylor bubbles, 962, 963
Taylor series, 1445
TDTR, *see* Time-domain thermoreflectance
TEFRI program, 872
Telecommunication electronics, outdoor, 1118–1121
TEM, *see* Transmission electron microscopy
TEMA, *see* Tubular manufacturer's manufacturing association
TEMA *Standards* (2007), 473
Temperature(s), 899
  blood flow effect on, 641–643
  boundary layer, 284–286
  dependency, 1082–1084*t*
  dew-point, 900
  dry-bulb, 899
  effect, 512
  field plot, 619–620, 620*f*
  history, coupling to rate processes, 629
  limits, understanding, 1105–1107
  mean radiant, 908
  sensitive electrical parameters (TSEP), 1079, 1081*t*
  subzero, 684
  thermostat, 735
  uniform surface, 310, 312
Temperature aware task scheduling, 1038
Temperature distribution, in bi-metal plate, 1494*f*
Temperature–entropy diagram for Carnot refrigeration cycle, 916*f*
Temperature-sensitive electrical parameters (TSEPs), 1079
Tempering furnace
  heat transfer phenomena in, 1366–1367
  model, 1360, 1362*f*
Terminal slip velocity, 220
TES, *see* Thermal energy storage
Test chambers, 765*f*
  and breathing wall, 764–765
  energy balance model, 766*f*
Test rig, 555*f*
Tetrahedral elements, 1485–1486
Theoretical air, 57
Therapeutic heating, 656
Thermal and energy performance analysis
  laboratory evaluations of PCM, 789–790
  numerical tools for PCM, 791–792
Thermal anemometer, 1215
Thermal aspects of paper making, 1270–1302
Thermal blood vessel sealing, 678–680
Thermal boundary layers, 1004
Thermal breaks, 806
Thermal bridges
  in building structures, 800–816
  characteristic of building materials and construction subsystems, 803–804
  combined, 806
  definition and common consequences of, 801–802
  design problems, 808*t*
  durability consequences of, 806–807
  effect, 801, 801*f*
  energy performance standards focused on buildings, 814–815
  engineering methods for analysis, 807
  generated by building geometry and architectural details, 803
  isothermal planes method, 810
  linear thermal transmittance method, 811–813
  mitigation, methods of, 815–816
  modified zone method, 811
  numerical tools used in thermal performance analysis, 813–814
  parallel path method, 809–810
  R-value and *U*-factor, 807–809
  structural bridges due to construction, 804–806
  types of, 802–809
Thermal buoyancy, 771
Thermal coagulation, 675–676
Thermal comfort
  conditions, 907–908
  effect of thermal bridging, 802
Thermal conduction, in electronic microstructures and nanostructures, 1036–1074
Thermal conductivity
  coefficient, 1358
  combined, 718–719
  and conductance of materials/interfaces, 1065–1068
  in crystal, 1056*f*
  data
    for calcium silicate, 725, 725*t*
    for fibrous pipe and board insulation, 725*t*
    for perlite pipe insulations, 725*t*
    for specific tissues, 691–693
  data taken from ASTM standard specifications, 724*t*
  defined, 262
  and diffusivity, measurement, 631–635
  of food materials, 1020
  of materials, 716
  microelectromechanical-fabricated device to measuring, 1063, 1064*f*
  of myocardial tissue, 640*t*
  of nanofluids, 545–554, 557, 563–564, 578–581, 580*f*, 586
  of paper, 1300–1302, 1300*t*
  powder insulation, 720*f*, 721*f*
  ranges, 717*f*
  of silicon, 1090*f*
  thermal epoxies, 1104*t*
  thermal greases, 1103*t*
  of typical packaging materials, 1091*t*
Thermal coupling, 763
Thermal damage
  histologic markers of
    collagen shrinkage, 677–678
    lethal low-temperature thermal effects, 674–675
    nonlethal low-temperature tissue effects, 673–674
    thermal coagulation, 675–676
    thermally induced loss of birefringence, 676–677
    water-dominated effects, 678–680
  studies, 662–673
  in vivo red zone, 674–675

Thermal degradation, 479
Thermal diffusivity, 631–635, 1020, 1040
  as function of drop Peclet number, 1139*f*
Thermal discomfort risk, 775
Thermal dissipation, 1169*t*
Thermal drying, 1306
Thermal effectiveness, of synthetic jet heat sinks, 1161
Thermal efficiency, 1323
Thermal–electrical solar still, 1423, 1424*f*
Thermal energy, 1389
Thermal energy storage (TES), 739, 742
Thermal engineering, 1243–1245
Thermal enhancement
  configurations, 1112*t*
  techniques, 1113*f*
Thermal epoxies, and thermal conductivities, 1104*t*
Thermal expansion coefficient, isobaric, 324
Thermal fatigue reliability, 1114, 1114–1116*f*
Thermal greases, and thermal conductivities, 1103*t*
Thermal history experiments, transient, 682–683
Thermal–hydraulic analysis, 527
Thermal insulations, 716–726, 723–724, 732, 801, 803*f*, 806
  heat transfer in
    combined thermal conductivity, 718–719
    gas conduction and convection, 717–718
    radiation heat transfer, 718
    solid conduction, 716–717
  insulation application
    thermal insulation products and installation, 724–725
    vacuum maintenance, 725–726
  insulation systems, 719–724
  level, 754–755
Thermal interface materials, 1102–1105
Thermal mass flow meters, 1195–1197, 1196*f*
Thermal modeling, of solar stills, 1399–1405
Thermal nanofluid flow, in double-tube heat exchanger, 590
Thermal networks, 1246–1247
Thermal performance, 1101*f*
  of electronic packages, 1102*t*
  laboratory evaluations of PCM, 789–790
  numerical tools used in
    thermal bridging and construction details, 813–814
Thermal plasmas, 1003
Thermal processing system and preservation of foods
  heating equipment
    electromagnetic heating, 1030–1031
    Ohmic heating, 1031
    retorts, 1025–1027
    scraped surface heat exchanger, 1029–1030
    shell and tube, 1028–1029
    tubular heat exchanger, 1027–1028
  kinetics of reactions
    C value, 1025
    D value, 1023–1024
    F value, 1024
    process optimization, 1025
    z value, 1024
  microorganisms, 1021–1023
    classification based on gram staining, 1022
    classification based on oxygen, 1022
    classification based on pH/water activity/salt/osmotic pressure, 1022
    classification based on temperature, 1022
    surrogate organism, 1023
    target organism, 1022–1023
  process considerations
    blanching, 1031
    hot fill, 1032–1033
    pasteurization, 1031–1032
    sterilization, 1033–1034
  product considerations
    dielectric properties, 1020–1021
    electrical properties, 1021
    pH, 1017–1018
    physical properties, 1018
    rheological properties, 1018–1020
    thermal properties, 1020
    water activity, 1018
Thermal properties
  as function of temperature, 637
  as function of water and fat content, 639–641
  in thermal processing, 1020
Thermal radiation, 341
  among hot gas, 1343–1346
Thermal reservoir, 8
Thermal resistance, 566, 569*f*, 752*f*, 807–809, 1073
Thermal shorts, 803, 815–816
Thermal superconductor, *see* Heat pipes
Thermal systems, economic analysis of, 102–105
  calculation of product costs, 105
  estimation of total capital investment, 102
  principles of economic evaluation, 102–105
    annuities, 104
    cost escalation, 104
    levelization, 104–105
    parameters and assumptions, 105
Thermal tapes, 1103–1104
Thermal tempering, 1360, 1363
Thermal transmittance, 807–809
Thermistor, 1169*t*
  calibration, 1168
  resistance variation, 1169*t*
  thermal characteristics, 1168–1170
Thermocompressor, 1295
Thermocouples, 764, 1162–1164
  application characteristics of alloys, 1163*t*
  calibrations, 1168
  circuit, 1167
  E–T diagram of, 1166
  pairs, first names for, 1165
  substitute material extension wires for, 1164*t*
  theory, 1164–1168
  Type T, 1225
Thermodiffusion/thermophoresis, 553
Thermodynamic data retrieval, 40–43
Thermodynamic effectiveness, guidelines for improving, 84–85
Thermodynamic functions, 39
Thermodynamic limitation, air-cooled heat exchangers, 496–497
Thermodynamic optimization, 94
Thermodynamic properties of water vapor–air mixtures, 867*t*
Thermodynamic system, 79
Thermodynamic temperature, 4
Thermodynamic wet-bulb temperature, 899

Thermoeconomics, 67
Thermoforming, 1276
Thermomagnetic convection, 593
Thermometer
  constant volume gas, 4
  wet-bulb, 49
Thermometry techniques, 1060–1065
Thermophilic microorganisms, 1022
Thermophoresis, 545
Thermophysical data, of heat pipes, 977–978
Thermophysical properties, variations in, 1004
Thermopile bolometers, 1173
Thermoreflectance, 1060
Thermoregulation, physiological processes of human, 652
Thermoregulatory processes, 652–653
Thermostat temperatures, 735, 912
Thermosyphons, 973
THERM software, 814
Thin-airfoil theory, 206–208, 207*f*
Thin small outline package (TSOP), 1085
Third law of thermodynamics, 60
T-history method, 790
Thixotropic behavior, 227*f*
Thixotropic fluid, 1020
Three-dimensional elements, 1485–1487
Three-dimensional heat transfer programs, 814
Three-dimensional integrals, 1488
3-omega (3-ω) method, 1060, 1060f
Three-phase motors, 860
Threshold
  fouling, 618, 619
  function, 1222
  limit value (TLV), 928
Throttling devices, 18
Tick-off heuristics, 889
Time
  value, of money, 102
  variant models, 747
Time-dependent fluids, 225, 226*f*
  friction factor–Reynolds number behavior for, 242*f*
Time-domain thermoreflectance (TDTR), 1060, 1062*f*, 1063*f*
Time-independent fluids, 225
  types of, 228–229
Time integration algorithm, 1479, 1480*t*
Time variant models, 747
Tissue(s)
  alterations, 658
  death, 667, 674
  effects, subzero temperatures, 684
  representative electric properties of, 658–659
  thermal transport properties, 630–631
Titania, 1378–1379
Tollmein–Schlichting instabilities, 1132
Tomography, 1213, 1213*f*
Top loss heat coefficient, 1395–1396, 1398
Toroidal vortex flow, 1007
Torque, 160–161
Total capital investment, 102
Total emissivity, of reference state, 357
Total pressure, 1306
Total revenue requirement method, 105
TPM, *see* Transmission polarizing microscopy
TransAT software, 1558
Transfer number, 437
Transient diffusion, in plane slab, 430
Transient hot-wire method, for measuring thermal conductivity, 549
Transient nonlinear equation solver, 1481
Transient plane source techniques, for thermal conductivity, 549
Transient pressures, 162–163, 164
Transient puff topography, 597, 597*f*
Transition boiling, 372, 385
Transition flow, 524
Transmission electron microscopy (TEM), 547, 548*f*, 673
Transmission function, 1051
Transmission polarizing microscopy (TPM), 676
Transmissivity, 343, 367
Transmittance, 367
Transparent, 367
Transport equations, 1462
Trapezoidal rule, 868–869
Traverse, a probe, 1215
Treated surfaces, 933
Triangular elements, 1484
Tribological systems, 1247
Triple line, 29
Triple point of water, 5
Triple tube heat exchanger, 1028, 1028*f*
TRISCO tool, 814
Truncation error, 1446
TSEPs, *see* Temperature-sensitive electrical parameters
TSOP, *see* Thin small outline package
Tube(s)
  banks, 295–301
  coaxial, 311
  construction features, 473
  material, 539
  supports, 473
Tube-bank construction, in finned tube, 488
Tube-bank layouts, in finned tube, 487
Tube-fin heat exchangers, 503–504, 522
Tubular heat exchangers, 1032
  double tube, 1027–1028
  triple tube, 1028
Tubular manufacturer's manufacturing association (TEMA), 622
Tunnel dryers, 1312–1313, 1312*f*
Turbine(s), 18
  cascade, 1183
  flow meters, 1195, 1196*f*
  gas, 842, 988
  -generator, exergetic efficiency of, 74
  steam, 842
Turbomolecular pump, 861–862
Turbotal inserts, 622, 624*f*
Turbulence, 1545
  effects, 1201
  intensities, 1156
Turbulent boundary layer, 182, 286, 291, 449
Turbulent flows, 307–308, 312, 546, 584–585, 586
Turbulent regime, 151
Twin-fluid injectors, 992
  design, 993*f*
Twin screw compressor, 832–834, 833
Twin water bath method, 790
Twisted-tape inserts, 940–941, 942*f*, 947, 949

2D square lattice, 1500
Two-dimensional elements, 1484–1485
Two-dimensional integrals, 1487
Two-dimensional thermal analysis, 814
Two-phase approach, for nanofluid flow simulations, 585, 586–587
Two-phase flow, fired heaters and, 625–626
Two-step method, to producing nanofluids, 547
Two-step preparation, to preparing nanofluids, 575
Two-way radios, 1111

U

U-factor (thermal transmittance), 807–809
Ultimate tensile strength (UTS), 1092–1093*t*
Ultrapasteurization, 1032
Ultrasonic flow meter, 1197*f*
Ultrasound, 656
Ultrasound dryers, 1326–1327, 1327*f*
Undercooling, 394
Uniform flow, 170
Uniform-series present-worth factor, 103*t*, 104
Unimolecular process
  activation and denaturalization, 659*f*
  descriptions, 660–662
Unit cost, 74
Universal acceleration transform, 1117
Universal gas constant, 5, 32
Unsteady term, 1462
Upstream disturbances, 164
Utility cost analysis, 729
Utility data analysis, 730, 740
UTS, *see* Ultimate tensile strength
U-tube manometer, 116*f*
U-value, 762

V

Vacuum
  brazed PHE, 508
  maintenance, 725–726
  multifoil insulation, 722
  pumps, 861*f*
Vacuum freeze-drying, 1327–1328, 1328*f*
Vacuum insulation panels (VIPs), 802
Vacuum maintenance materials, 726*t*
Valence band maximum (VBM), 1372
Valve(s)
  air, 165
  flow through, 157
  selection, 160–165
Van der Waals forces, 550, 575
Vane types, used on centrifugal fan, 859*f*
Vapor and gas power cycles, 76–82
  Carnot, Ericsson, and Stirling cycles, 82–83
  Otto, diesel, and dual cycles, 78–81
  Rankine and Brayton cycles, 76–78
Vapor compression (VC) cycle, 916–918, 917*f*
Vapor cycles, modifications to conventional, 819–822
  liquid flooded compression, 820
  vapor injected compression, 820–822
Vapor injected compression, 820–822, 821*f*
Vaporization, adiabatic, 442
Vapor–liquid flows, 209
Vapor power systems, 76
Vapor-space condensation, 948–949
Variable air volume (VAV), 735, 744, 928–929, 930
Variable conductance heat pipes (VCHPs), 965–966, 966*f*
Variable reflectivity coating roofs (VRCs), 749, 778
  annual energy costs impact, 756–758
  background, 753
  cooling thermal loads impact, 754–756
  cost–benefit analysis of, 758–760
  heating thermal loads impact, 754
  thermal analysis, 753–754
Vasculature-based models, 645–648
VAV, *see* Variable air volume
VBM, *see* Valence band maximum
VC cycle, *see* Vapor compression cycle
VCHPs, *see* Variable conductance heat pipes
Velocity(ies)
  boundary conditions, 1503
  definitions of, 421–422
  error, 1180
  head, 853
  head pumps, 855
  potential, 134
  profiles, developing, 306*f*
  of sound, 39
  vectors, 1520*f*
Ventilated cavity wall, 771, 772*f*
Ventilation, 1296, 1296*f*
Ventilation system, mechanical, 737
Venting, 1033
Venturi meters, 1192
  characteristics of, 1190*t*
VERA2D-84 program, 872
Vermiculite, 719, 720
Vertically split casings, 855
Very large-scale integrated (VLSI), 1036
Vibratory atomizers, 994
Vibrofluidized system, 1321
View factors, 347–352, 367
VIPs, *see* Vacuum insulation panels
Virial coefficients, 34
Viscoelastic fluids, 130, 322–323
Viscoelasticity, effect of, 243–244
Viscoplastic models, 232
Viscosity
  apparent, 229
Viscosity, 1018, 1353, 1354*f*
  apparent, 225
  measurements, 1212
  of nanofluids, 581–584, 583*f*, 586
Viscous boundary layer, 1011
Viscous dissipation, 284–285, 289
Viscous forces, 545
Viscous shear, 183
Viscous/vapor pressure limitation, of heat pipe, 967–968
VLSI, *see* Very large-scale integrated
VOF method, *see* Volume of fluid method
Void fractions, 214–217, 388
Volume averaging, principles of, 1526–1528
Volume conservation equation, 1528
Volume median diameter, 986
Volume of fluid (VOF) method, 405, 999, 1000*f*
  advection, 1533
  application, 1536–1537

equation, 1534
numerical methodology, 1534–1536
solution algorithm for two-phase flow simulation, 1536
Volume-tracking methods, 1525
Volumetric flow rate, 125, 906
Volute, in centrifugal pump, 854
Vortex circulation inside confined liquid Taylor plugs, 962
Vortex-shedding frequency, 1155
Vortex shedding meters, 1195
Vorticity field, 1518, 1518*f*
VRCs, *see* Variable reflectivity coating roofs

W

WALE viscosity model, *see* Wall adaptive local Eddy viscosity model
Walk-through audit, 728
Wall adaptive local Eddy (WALE) viscosity model, 1552
Wall-bounded flows
large eddy simulation (LES) of, 1545–1556
Wall–floor intersections, 805, 805*f*
Wall-modeled LES model, 1549, 1555
Wall modeling, 1553–1554
Wall-resolved LES model, 1549, 1554
Warme und feuchte instationar (WUFI), 791
Washcoat, 432, 433
Water
activity, 1018
–air interface, 419*f*
blanching, 1031
characteristics using fixtures, 739*t*
loading, 869
management, indoor, 738
production cost of, 1414
secondary, 930
thermodynamic properties of saturated, 898
triple point of, 5
-using fixtures, characteristics of, 739
Water heating systems
utilization of PCMs in, 789
Water retention value (WRV), 1281
Water splitting PEC system, 1371–1372
Water vapor pressure, 899
Wave, growth rate, 989
Weber number, 988
Wedge angle, 290
Weibull distributions, 1114, 1116*f*
Weighted essentially non-oscillatory (WENO) scheme, 1513, 1540
Weighted sum of gray gases (WSGG), 366
Weinbaum–Jiji model, 645–646
Weir
broad-crested, 173
coefficient, 174
sharp-crested, 173
Welded PHEs, 507
WENO scheme, *see* Weighted essentially non-oscillatory scheme
Wet-bulb psychrometer, 443–444
Wet-bulb temperature, 900, 1307
Wet-bulb thermometer, 49
Wet flows, 962
Wet materials, properties of, 1308–1309
Wettability, 962
Wick structures, 955, 967
Wick-type solar distillation, 1391
Wick-type solar still, 1423–1426, 1425*f*
Wide band, 1174
Wide-range atomizer designs, 992
Wiedemann–Franz Law, 1043, 1053
Wilson, Carlos, 1390
Window cavity, heat loss across, 278–279
Windows, replacement of, 733
Window-wall perimeters, 806
Wind tunnel
role of, 189–190
supersonic, 201
tests, 842
Wire mesh sensors, 1214
"Wire-to-water" efficiencies, for centrifugal pumps, 856*f*
Wissler model, 653*f*
of human thermoregulation, 653–656
Wood fiber, 1277, 1278*f*
affinity with water, 1279
Working fluid, 76
Work-rate term, 131
WRV, *see* Water retention value
WSGG, *see* Weighted sum of gray gases
WUFI, *see Warme und feuchte instationar*

X

Xerophilic microorganisms, 1022

Y

Yeasts, 1021
Yield stress, 1020
Young modulus, 1517–1518

Z

Zero flow, 858, 1209
Zero-net-mass-flux (ZNMF), 1153
Zero pressure gradient, 283–284, 291–293
Zero viscosity, 179
ZNMF, *see* Zero-net-mass-flux
z value, 1024