Second Edition

HANDBOOK OF PHOTOVOLTAIC SCIENCE AND ENGINEERING

Editors ANTONIO LUQUE • STEVEN HEGEDUS

WILEY

Handbook of Photovoltaic Science and Engineering

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

Handbook of Photovoltaic Science and Engineering

Second Edition

Edited by

Antonio Luque

Instituto de Energía Solar,
Universidad Politécnica de Madrid, Spain

and

Steven Hegedus

Institute of Energy Conversion,
University of Delaware, USA

WILEY

A John Wiley and Sons, Ltd., Publication

This edition first published 2011

First Edition published in 2003

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloguing-in-Publication Data

Handbook of photovoltaic science and engineering / edited by A Luque and S Hegedus. – 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-72169-8 (cloth)
1. Photovoltaic cells–Handbooks, manuals, etc. 2. Photovoltaic power generation–Handbooks, manuals, etc. I. Luque, A. (Antonio) II. Hegedus, Steven.
TK8322.H33 2010
621.31′244 – dc22

2010031107

A catalogue record for this book is available from the British Library.

Print ISBN: 978-0-470-72169-8
ePDF ISBN: 978-0-470-97466-7
oBook ISBN: 978-0-470-97470-4
ePub ISBN: 978-0-470-97612-8

Set in 9/11 Times by Laserwords Private Limited, Chennai, India.

Contents

About the Editors

Professor Antonio Luque was born in Malaga, Spain, in 1941. He is married with two children and five grandchildren. A full Professor at the Universidad Politécnica de Madrid since 1970, he currently serves at the Instituto de Energía Solar that he founded in 1979. There he has formed over 30 PhD Students and the research group he leads (Silicon and PV Fundamental Studies) is ranked first among the 199 consolidated research groups of his university.

In 1976 Professor Luque invented the bifacial cell and in 1981 he founded ISOFOTON; a solar cell company with a turnover of about 300 million dollars (2007). In 1997 he proposed the intermediate band solar cell (321 citations in WOK registered journals by September 2010). Today more than sixty research centers worldwide have published on this topic (WOK registered) with citation of his work.

The main focus of Professor Luque's present research is in further understanding and developing the intermediate band solar cell, but further to this he is involved in two major additional actions: the establishment (as founder, and CEO) of the silicon ultrapurification research company CENTESIL (owned by two universities and three corporations) to further reduce the costs of silicon solar cell; and the supervision as Chair of the Scientific International Committee of the new institute ISFOC for Concentrator Photovoltaic (CPV) systems, established under his plan to stimulate the introduction of the CPV technology worldwide. This institute has granted contracts (through the board he chairs) to seven companies (three from Spain, two from the USA, one from Germany and one from Taiwan) and over two MW of panels have already been installed at ISFOC using the new multijunction cell technology that has given cell efficiencies above 41%.

He has been honored by several important prizes and distinctions, including the membership to the Royal Academy of Engineering of Spain, the Honor membership of the Ioffe Institute in St. Petersburg and two Honoris Causa doctorates (Carlos III University of Madrid and Jaen University). He has also received three major Spanish National Prizes (two delivered by the King of Spain and one by the Crown Prince) on technology and environmental research as well as one from European Commission and one from the US IEEE-PV Conference, both on photovoltaics.

Dr. Steven Hegedus has been involved in solar cell research for 30 years. While earning a BS in Electrical Engineering/Applied Physics at Case Western Reserve University (1977) he worked on a solar hot water project. He worked on integrated circuit design and modeling at IBM Corp from 1977–1982, during which time he received a Masters in Electrical Engineering from Cornell, working on polycrystalline GaAs solar cells. In 1982 he joined the research staff of the Institute of Energy Conversion (IEC) at the University of Delaware (UD), the world's oldest photovoltaic research laboratory. He has worked on nearly all of the commercially relevant solar cell technologies. Areas of active research include optical enhancement and contacts to TCOs, high growth rate of PECVD nanocrystalline Si, thin film device analysis and characterization, a-Si/c-Si heterojunction processing, and stability under accelerated degradation conditions. While at the IEC, he got a Ph.D. in Electrical Engineering from UD. He has contracts with the US Department of Energy and several US companies, large and small, to assist their development of thin film and c-Si PV products. Dr. Hegedus has been lead author of nearly 50 papers in the field of solar cell device analysis, processing, reliability and measurements. He teaches a graduate class at UD in Solar Electric Systems. Dr. Hegedus is keenly aware of the impact of policy on solar energy commercialization and was appointed a Policy Fellow by UD's Center for Energy and Environmental Policy in 2006. He was the first resident of his town to install a rooftop PV system.

List of Contributors

Armin G. Aberle
Solar Energy Research Institute of Singapore (SERIS)
National University of Singapore (NUS)
4 Engineering Drive 3
Block E4-01-01
Singapore
117576
Singapore

Jesús Alonso
Departamento de I+D
ISOFOTON
C/Caleta de Velez, 52
Pol. Ind. Santa Teresa
29006 Malaga
Spain
Phone: +3495 224 3790
Fax: +3495 224 3449
email: j.alonso@isofoton.es

Ignacio Antón
Instituto de Energía Solar
Universidad Politécnica de Madrid
E.T.S.I Telecomunicatión
28040 Madrid
Spain

Ismael Guerrero Arias
DC Wafers
Ctra.
Madrid Km 320
24227 Valdelafuente
León, Spain

Sheila Bailey
NASA Glenn Research Center
Cleveland, OH
USA
Phone: +1 216 433 2228
Fax: +1 216 433 6106
email: Sheila.bailey@lerc.nasa.gov

Bruno Burger
Fraunhofer Institute for Solar Energy Systems ISE
Freiburg
Heidenhofstr. 2
79110 Freiburg
Germany

John Byrne
Center for Energy and Environmental Policy
University of Delaware
Newark
Delaware
19716
USA

Carlos del Cañizo
Instituto de Energía Solar
Universidad Politécnica de Madrid
E.T.S.I. Telecomunicación
28040 Madrid
Spain
Phone: +34 91 544 1060
Fax: +34 91 544 6341
email: canizo@ies-def.upm.es

Bruno Ceccaroli
Marche AS
P.O. Box 8309 Vaagsbygd
N-4676 Kristiansand
Norway
Phone: +47 38 08 58 81
Fax: +47 38 11 99 61
email: br-c@online.no

Alan E. Delahoy
New Millennium Solar Equipment Corp.
8 Marlen Drive
Robbinsville, NJ 08691
USA
Phone: +1 609 587 3000
Fax: +1 609 587 5355
email: a.delahoy@nmsec.com

Xunming Deng
Department of Physics and Astronomy
University of Toledo
Toledo, Ohio
USA
Phone: +1 419 530 4782
Fax: +1 419 530 2723
email: dengx@physics.utoledo.edu

Jaime Agredano
Instituto de Investigaciones Eléctricas
Gerencia de Energías No Convencionales
P.O. Box 475
Cuernavaca Morelos
62490 México
email: agredano@iie.org.mx

Keith Emery
NREL
1617 Cole Boulevard
Golden, CO 80401-3393
USA
Phone: +1 303 384 6632
Fax: +1 303 384 6604
email: keith_emery@nrel.gov

Arthur L. Endrös
Corporate R&D department
Siemens and Shell Solar GmbH
Siemens AG
Munich, Germany

Dieter Franke
Access e.V.
Aachen
Germany

D. J. Friedman
NREL
1617 Cole Boulevard
Golden, CO 80401-3393
USA

Jeffery L. Gray
Purdue University
School of Electrical and Computer Engineering
Electrical Engineering Building
465 Northwestern Ave.
West Lafayette
Indiana
47907-2035
USA
email:grayj@ecn.purdue.edu

Lalith Gunaratne
Solar Power & Light Co, Ltd
338 TB Jayah Mawatha
Colombo 10
Sri Lanka
Phone: +94 014 818395
Fax: +94 014 810824
email: laithq@sri.lanka.net

Sheyu Guo
Yiri Solartech (Suzhou) Co., Ltd.
Wujiang Hi-Tech Park
2358 Chang An Road, Wujiang City
Jiangsu Province, P. R. China 215200
Phone: +86 512 63970266
Fax: +86 512 63970278
email: sguo@yirisolartech.com

Christian Häßler
Central Research Physics
Bayer AG Krefeld
Germany
email: christian.haessler@bayerpolymers.com

Kohjiro Hara
Research Center for Photovoltaics (RCPV)
National Institute of Advanced Industrial Science and Technology (AIST)
Central 5
1-1-1 Higashi, Tsukuba, Ibaraki 305-8565, Japan
Phone: 29-861-4494
Fax: 29-861-6771
email: k-hara@aist.go.jp

Steven Hegedus
Institute of Energy Conversion
University of Delaware
Newark DE 19716
USA
email: ssh@udel.edu

Jorge M. Huacuz
Instituto de Investigaciones Eléctricas
Gerencia de Energías No Convencionales
P.O. Box 475
Cuernavaca Morelos
62490 México
email: jhuacuz@iie.org.mx

Raymond M. Hudson
BEW Engineering
2303 Camino Ramon
Suite 220
San Ramon CA 94583
USA
Phone: +1925 867 3330

Henk F. Kaan
ECN Energy Research Centre of the Netherlands
P.O. Box 1
1755 ZG Petten
The Netherlands

Juris P. Kalejs
RWE Schott Solar Inc.
4 Suburban Park Drive
Billerica, MA 01821 USA
Phone: 978-947-5993
Fax: 978-663-2868
email: jkalejs@asepv.com

Wolfgang Koch
Central Research, Physics (ZF-FPM), Photonic Materials
Chemicals-Bayer Solar, (CH-BS), Projects
Bayer AG
Geb.R82, PF111107
D-47812 Krefeld
Germany
Phone: +492151-883370
Fax: +492151-887503
email: wolfgang.koch.wk2@bayer-ag.de

Lado Kurdgelashvili
Center for Energy and Environmental Policy
University of Delaware
Newark
Delaware
19716
USA

Sarah Kurtz
NREL
1617 Cole Boulevard
Golden, CO 80401-3393
USA
Phone: +1 303 384 6475
Fax: +1 303 384 6531
email: sarah_kurtz@nrel.gov

Otto Lohne
Norwegian University of Science and Technology
Department of Materials Technology
N-7491 Trondheim
Norway
Phone: +47 73 59 27 94
Fax: +47 43 59 48 89
email: Otto.Lohne@sintef.no

Eduardo Lorenzo
Instituto de Energía Solar
Universidad Politécnica de Madrid
E.T.S.I. Telecomunicación
Ciudad Universitaria
28040 Madrid
Spain
Phone: +3491 366 7228
Fax: +3491 544 6341
email: lorenzo@ies-def.upm.es

Antonio Luque
Instituto de Energía Solar
Universidad Politécnica de Madrid
E.T.S.I. Telecomunicación
28040 Madrid
Spain
Phone: +34 91 336 7229
Fax: +34 91 544 6341
email: luque@ies-def.upm.es

Antonio Martí
Instituto de Energía Solar
Universidad Politécnica de Madrid
E.T.S.I. Telecomunicación
28040 Madrid
Spain
Phone: +34 91 544 1060
Fax: +34 91 544 6341
email: amarti@etsit.upm.es

Brian E. McCandless
Institute of Energy Conversion
University of Delaware
Newark, DE 19716
USA
Phone: +1 302 831 6240
Fax: +1 302 831 6226
email: bem@udel.edu

H. J. Möller
Institut für Experimentelle Physik
TU Bergakademie Freiberg
Silbermannstr.1
09599 Freiberg
Germany
Phone: +493731-392896
Fax: +493731-394314
email: moeller@physik.tu-freiberg.de

Shogo Mori
Department of Fine Materials Engineering
Faculty of Textile Science and Technology
Shinshu University
Ueda 386–8567
Japan

Hugh O'Neill
Center for Structural Molecular Biology
Chemical Sciences Division
Oak Ridge National Lab
Tennessee, USA

J. M. Olson
NREL
1617 Cole Boulevard
Golden, CO 80401-3393
USA

Ryne Raffaelle
National Center for Photovoltaics
National Renewable Energy Lab Golden
CO, USA

Anat Razon
BEW Engineering
2303 Camino Ramon
Suite 220
San Ramon CA 94583
USA
Phone: +1925 867 3330

Tjerk H. Reijenga
BEAR Architecten
Gravin Beatrixstraat 34
NL 2805 PJ Gouda
The Netherlands
Phone: +31 182 529 899
Fax: +31 182 582 599
email: Tjerk@bear.nl

Gabriel Sala
Instituto de Energia Solar
Universidad Politécnica
de Madrid
E.T.S.I Telecomunicatión
28040 Madrid
Spain

Dirk Uwe Sauer
Fraunhofer Institute for Solar Energy Systems
ISE
Heidenhofstrasse 2
D-79110 Freiburg
Germany
Phone: +49 761 4588 5219
Fax: +49 761 4588 9217
email: sauer@ise.fhg.de

Eric A. Schiff
Department of Physics
Syracuse University
Syracuse, New York 13244-1130
USA
http://physics.syr.edu/∼schiff

Jürgen Schmid
Fraunhofer Institute for Wind Energy and
Energy Systems Technology IWES, Kassel
Germany
Phone: +49 (0)5 61/72 94-3 45
Fax: +49 (0)5 61/72 94-3 00
email: jschmid@iset.uni-kassel.de

Heribert Schmidt
Fraunhofer Institute for Solar Energy Systems ISE
Freiburg
Heidenhofstr. 2
79110 Freiburg
Germany
Phone: +49 (0)7 61/45 88-52-26
Phone: +49 (0)7 61/45 88-92-26
email: heri@ise.fhg.de

Hugo Rodriguez San Segundo
DC Wafers
Ctra.
Madrid Km 320
24227 Valdelafuente
León, Spain

William N. Shafarman
Institute of Energy Conversion
University of Delaware
Newark, DE 19716
USA
Phone: 1 302 831 6215
Fax: 1 302 831 6226
email: wns@udel.edu

Susanne Siebentritt
University of Luxembourg
Laboratory for Photovoltaics
162a Avenue de la Faïencerie
L-1511 Luxembourg

James R. Sites
Department of Physics
Colorado State University
Fort Collins, CO 80523-1875
USA
Phone: +1 970 491 5850
Fax: +1 970 491 7947
email: sites@lamar.colostate.edu

Lars Stolt
Solibro Research AB
Vallvägen 5
75651 Uppsala
Sweden
Phone: +46 18 471 3039
Fax: +46 18 555 095
email: Lars.Stolt@angstrom.uu.se

Sam-Shajing Sun
Chemistry Department and PhD Program in Materials Science & Engineering
Norfolk State University
Virginia, USA

Ignacio Tobías
Instituto de Energía Solar
Universidad Politécnica de Madrid
ETSI Telecomunicación
28040 Madrid
Spain
Phone: +3491 5475700-282
Fax: +3491 5446341
email: Tobias@ies-def.upm.es

Timothy U. Townsend
BEW Engineering
2303 Camino Ramon
Suite 220
San Ramon CA 94583
USA
Phone: +1925 867 3330

Xavier Vallvé
Trama Tecno Ambiental
Avda. Meridiana, 153
planta baixa
08026 Barcelona
Spain

Per I. Widenborg
Formerly with School of Photovoltaic and Renewable Energy Engineering
University of New South Wales Sydney Australia
Now with Solar Energy
Research Institute of Singapore
National University of Singapore
Singapore

Charles M. Whitaker
BEW Engineering

Preface to the 2nd Edition

The first edition of the *Handbook of Photovoltaic Science and Engineering* was published in 2003. It described the results of 50 years of research, technology, product development, and applications of solar cells and modules. This included the first generation of terrestrial PV – crystalline Si wafers – the second generation of PV – thin films of amorphous Si, CdTe, or $CuInGaSe_2$ – and the third generation PV – organic dye-sensitized junctions mimicking photosynthesis or advanced very high efficiency theoretical concepts such as multiphoton and intermediate band solar cells, which had yet to be demonstrated in practice. It also included chapters on III–V based multijunctions (having the highest demonstrated efficiency) and concentrators. Applications of PV installed in outer space and on earth – from urban offices to rural villages – were described. Components of systems such as batteries and power conversion electronics such as inverters had their own chapters. Finally we included chapters on fundamental physics, measurements and characterization, and how to calculate the energy produced from a module installed anywhere for any configuration.

Almost coincident with this publication, interest in PV exploded. Sales and production increased over tenfold, from 600 MW of production in 2003 to 7300 MW in 2009. Growing interest in PV generated significant private and public investment, resulting in significant improvements in technology and applications. Much of this was driven by innovative national policies. Hundreds of companies, from brand new small start-ups to mature giant multinationals, tried to ride the surging wave of popular and technical interest in PV. Many of them bought copies of the first edition to help educate and inform their engineers, managers, analysts and investors. The PV field was maturing – companies were finally making profits, merging, scaling up production, and expanding. New technologies were finding their way into the marketplace.

A second edition was planned to represent these new developments. Ultimately, this second edition has benefited from the dose of reality of the past year's economic crisis. But it is a testament to the power of an idea whose time has come, that PV has continued to grow and prosper, one of the few industries which still increased its sales during the New Great Depression. In many countries, nurturing a PV industry has become a prominent strategy in economic recovery and job creation, in addition to being a potent weapon in the battle against global climate change.

What's new in this second edition? There are three completely new chapters, discussing the role of national energy policy in encouraging PV growth, transparent conductive oxides for thin film PV, and third-generation organic polymer-based devices. Five chapters have all new authors, giving a fresh view of crystalline Si wafer technology, second-generation thin film silicon cells, concentrating PV, power conditioning electronics, and off-grid and on-grid system design. All the other chapters have been significantly updated with new technical advances, state-of-the-art cell efficiencies, manufacturing status, and installation-related data.

The editors dedicate this book to all those who have worked so hard for over half a century to bring solar electricity to its present success, and to our colleagues present and future who must work even harder in the next half century to ensure that PV fulfills its potential as a widely available, carbon-free clean energy source.

The editors also owe tremendous debt to the authors of each chapter. Their long hours spent writing the best possible chapter covering their field of expertise, only to suffer a storm of editorial criticisms and corrections, has hopefully made this a high-quality publication of lasting value.

Finally we want to express our gratitude to our loved ones – Carmen, Ignacio, Sofia, (and their children), and Debbie, Jordan, Ariel – for many hours stolen from family life while working on this book.

Antonio Luque & Steven Hegedus
June 2010

1

Achievements and Challenges of Solar Electricity from Photovoltaics

Steven Hegedus[1] and Antonio Luque[2]

[1]*Institute of Energy Conversion, University of Delaware, USA*
[2]*Instituto de Energía Solar, Universidad Politécnica de Madrid, Spain*

1.1 THE BIG PICTURE

Congratulations! You are reading a book about a technology that has changed the way we think about energy. Photovoltaics (or PV) is an empowering technology that has shown that it can generate electricity for the human race for a wide range of applications, scales, climates, and geographic locations. Photovoltaics can bring electricity to a rural homemaker who lives 100 kilometers and 100 years away from the nearest electric grid connection in her country, thus allowing her family to have clean, electric lights instead of kerosene lamps, to listen to a radio, and to run a sewing machine for additional income. It can pump clean water from underground aquifers for drinking or watering crops or cattle. Or, photovoltaics can provide electricity to remote transmitter stations in the mountains, allowing better communication without building a road to deliver diesel fuel for its generator. It can allow a suburban or urban homeowner to produce some or all of their annual electricity, selling any excess solar electricity back into the grid. It can help a major electric utility in Los Angeles, Tokyo, or Madrid to meet its peak load on hot summer afternoons when air conditioners are working full time. Finally, photovoltaics has been powering satellites orbiting the Earth for 30 years or vehicles roving over the surface of Mars.

Every day the human race is more aware of the need for sustainable management of its Planet Earth. It upholds almost seven billion human beings of which one billion have adopted a

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

high-consumption lifestyle which is not sustainable. "High consumption" used to refer materials that could become scarce, but it increasingly refers to energy. Here the term energy, refers to "useful energy" (or exergy), that once used is degraded, typically to waste heat, and will be no longer be useful.

Here we are concerned with electrical energy which is a secondary form of energy. Fossil fuel (coal, petroleum and natural gas) combustion and nuclear fission are the primary processes which create heat to turn water into steam which rotate giant turbines which generate electricity. When the C–H bonds in fossil fuels are burned in the presence of air for heat, they produce CO_2, and H_2O. The latter waste is not a problem because there is already so much water in the seas and in the atmosphere. But CO_2 is a different story. Analysis of the air bubbles embedded in Antarctic ice layers provides information on the CO_2 concentration of the atmosphere in the last 150 000 years. This content shows an unprecedented growth in the last 300 years, coinciding with the beginning of the industrialization. This fact is linked by most scientists to global climate change, including global warming, sea level rise, more violent storms, and changes in rainfall. This will disrupt agriculture, disease control, and other human activities. Thus, a substantial fraction of our energy must be generated without any C emissions within the next 10–20 years, or else the Earth will become a dangerous experiment. Besides, fossil fuels cannot last forever. Supplies of petroleum and natural gas will both peak and then decrease within decades if not years, and coal within few centuries. We must develop large-scale alternatives to burning fossil fuels very soon.

Another primary energy source for electrical generation is radioactivity in the form of uranium, which when conveniently transformed, fuels nuclear plants through nuclear fission. Concerning the uranium, most of it consists of the 238 isotope, which is not "fissile" (not a nuclear fuel) and only about 0.7% is the 235 isotope, which is fissile. With the present technology, nuclear fuel will peak within decades. However, uranium 238 can be converted to an artificial fissile fuel by proper bombardment with neutrons. With this technology, not fully commercial today, it would be possible to have nuclear power (with unproven cost effectiveness) maybe for a millennium. Nuclear fusion, which is a totally different nuclear technology, could be practically inexhaustible, but its practical feasibility is very far from being proven.

While nuclear plants emit no CO_2, they are still inherently dangerous. Nuclear engineers and regulators take many precautions to ensure safe operation, and (excepting very few cases) the power plants function without catastrophic problems. But the storage of highly radioactive wastes, which must remain controlled for centuries, remains an unsolved issue worldwide, along with the possibility of diversion of nuclear fuel to making a bomb.

So the situation at the beginning of the 21st century is that the previous century's methods of generating our most useful form of energy, electricity, are recognized as unsustainable, due to either increasing CO_2 poisoning of the atmosphere or the increasing stockpile of radioactive waste with no safe storage. What about using existing energy more efficiently? This will be crucial for slowing and perhaps even reversing the increased CO_2 levels. Doing more with less energy or just doing less (considered unpopular with growth-oriented economic advocates) are certainly necessary to reduce our demand for energy. But a growing world population with a growing appetite for energy is difficult to reconcile with using less energy. Besides, there is a large group whose voices are often not heard in this discussion – namely, the one out of three human beings who lack any electricity at all.

In fact, access to and consumption of electricity is closely correlated with quality of life, up to a point. Figure 1.1 shows the human development Index (HDI) for over 60 countries, which includes over 90% of the Earth's population, versus the annual per capita electricity use (adapted from [1]). The HDI is compiled by the UN and calculated on the basis of life expectancy, educational achievement, and per capita gross domestic product. To improve the quality of life in many countries, as measured by their HDI, will require increasing their electricity consumption by factors of 10 or more, from a few hundred to a few thousand kilowatt-hours (kW h) per year.

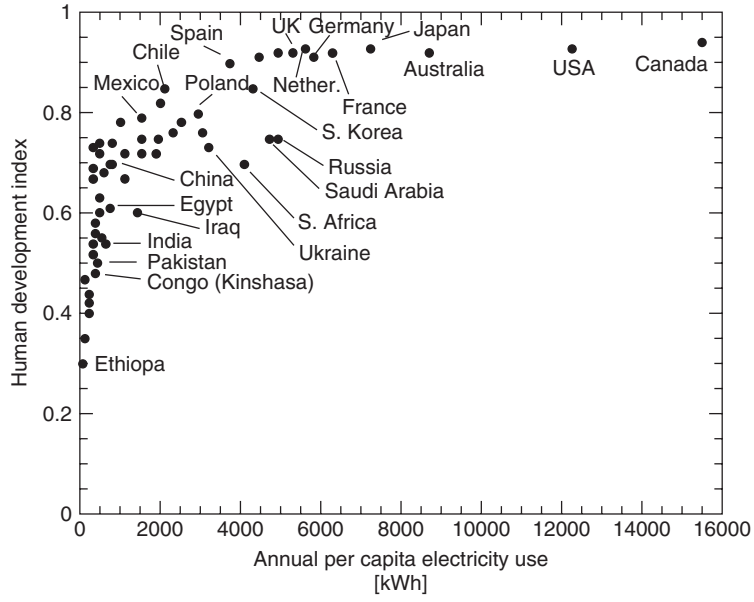


Figure 1.1 Human development index (HDI) versus per capita kW usage in year 2000 [1]

Adding two billion more inhabitants with increasing appetites for energy to the high-consumption pattern of today's one billion in the developed World, as would be expected from the development of China and India, would lead to unbearable stresses both in materials and energy. Barring their access (and that of others) to the wealth of the Western lifestyle is unfeasible, in addition to being unethical.

Renewable energies, and in particular solar energy, are the only clear solution to these issues. As matter of fact the amount of energy arriving on Earth from the Sun is gigantic: in the range of 10 000 times the current energy consumption of the human species. The ability of various forms of renewable energy to meet the "terawatt challenge" of providing world's present demand of 13 TW has been published [2]. We can also add geothermal energy (not renewable, properly speaking) and tidal energy, but they are insignificant in global terms, although locally, in some cases, their exploitation may be attractive.

Wind is generated by the solar energy (through the differential heating of the Earth in equatorial and polar regions). It has been calculated [3] that about 1% of the solar energy (10 times the global current consumption of energy) is converted into wind, but only a 4% of this is actually usable (but still 0.4 times the current consumption). It is estimated that with aggressive exploitation, land- and water-based wind generation is capable of providing about 10% of the world's expected energy demand [2]. Biomass converts solar energy in fuels but its efficiency is also very low, and its use for food has priority. Waves are caused by the wind and therefore a small fraction of the wind energy is passed to them. Sea currents, as winds also originate in the solar energy. The fraction passing to them is uncertain, but probably small. Finally, hydropower, produced by the transport of water from the sea to the land by means of solar energy, represents a tiny fraction of the total energy income, and the most promising sites are already in use. Summarizing, the direct exploitation of the solar energy is the real big energy resource [4].

Using photovoltaics with an efficiency of 10%, solar energy can be converted directly into enough electricity to provide 1000 times the current global consumption. Restricting solar collection

to the earth's solid surface (one quarter of the total surface area), we still have a potential of 250 times the current consumption. This means that using 0.4% of the land area could produce all the energy (electricity plus heat plus transportation) currently demanded. This fraction of land is much smaller than the one we use for agriculture.

Achieving the required strong penetration of solar energy is not trivial. In the rest of this chapter we shall present a description of the status PV and broadly outline some of the challenges for it to become a TW scale energy source. But let us advance arguments that are seldom spoken: (a) PV is technologically more mature than advanced nuclear fission or nuclear fusion technology, the two non-renewable CO_2-free energies permitting substantial increments of the global energy production; (b) even well-developed wind energy cannot match the amount of energy directly available from the sun; (c) biomass energy can expect further scientific development, but will probably not reach efficiency levels that will make of it a global alternative to solve the issues presented; (d) concentrating solar thermal power (CSP) could produce electricity in concurrence with PV. We think that PV has a bigger innovation potential and has also modularity properties (it operates at small or large scale) and lacks the geographic limitations of CSP which makes it a clear winner in this competition.

1.2 WHAT IS PHOTOVOLTAICS?

PV is the technology that generates direct current (DC) electrical power measured in watts (W) or kilowatts (kW) from semiconductors when they are illuminated by photons. As long as light is shining on the solar cell (the name for the individual PV element), it generates electrical power. When the light stops, the electricity stops. Solar cells never need recharging like a battery. Some have been in continuous outdoor operation on Earth or in space for over 30 years.

Table 1.1 lists some of the advantages and disadvantages of PV. Note, that they include both technical and nontechnical issues

Table 1.1 Advantages and disadvantages of photovoltaics

Advantages of photovoltaics

- Fuel source is vast, widely accessible and essentially infinite
- No emissions, combustion or radioactive waste (does not contribute perceptibly to global climate change or air/water pollution)
- Low operating costs (no fuel)
- No moving parts (no wear); theoretically everlasting
- Ambient temperature operation (no high-temperature corrosion or safety issues)
- High reliability of solar modules (manufacturers' guarantees over 30 years)
- Rather predictable annual output
- Modular (small or large increments)
- Can be integrated into new or existing building structures
- Can be very rapidly installed at nearly any point-of-use

Disadvantages of photovoltaics

- Fuel source is diffuse (sunlight is a relatively low-density energy)
- High initial (installed) costs
- Unpredictable hourly or daily output
- Lack of economical efficient energy storage

What is the physical basis of PV operation? Solar cells are typically made of semiconductor materials, which have weakly bonded electrons occupying a band of energy called the valence band. When energy exceeding a certain threshold, called the bandgap energy, is applied to a valence electron, the bonds are broken and the electron is somewhat "free" to move around in a new energy band called the conduction band where it can "conduct" electricity through the material[1]. Thus, the free electrons in the conduction band are separated from the valence band by the bandgap (measured in units of electron volts or eV). This energy needed to free the electron can be supplied by photons, which are particles of light.

Figure 1.2 shows the idealized relation between energy (vertical axis) and the spatial boundaries (horizontal axis). When the solar cell is exposed to sunlight of sufficient energy, the incident solar photons are absorbed by the atoms, breaking the bonds of valence electrons and pumping them up to higher energy in the conduction band. There, a specially made selective contact collects conduction-band electrons and drives these freed electrons to the external circuit. The electrons lose their energy by doing work in the external circuit such as pumping water, spinning a fan, powering a sewing machine motor, a light bulb, or a computer. They are restored to the solar cell by the return loop of the circuit via a second selective contact, which returns them to the valence band with the same energy that they started with. The movement of these electrons in the external circuit and contacts is called the *electric current*. The potential at which the electrons are delivered to the external world is less than the threshold energy that excited the electrons; that is, the bandgap. It is independent of the energy of the photon that created it (provided its energy is above the threshold). Thus, in a material with a 1 eV bandgap, electrons excited by a 2 eV (red) photon or by a 3 eV (blue) photon will both still have a potential voltage of slightly less than 1 V (i.e. both of the

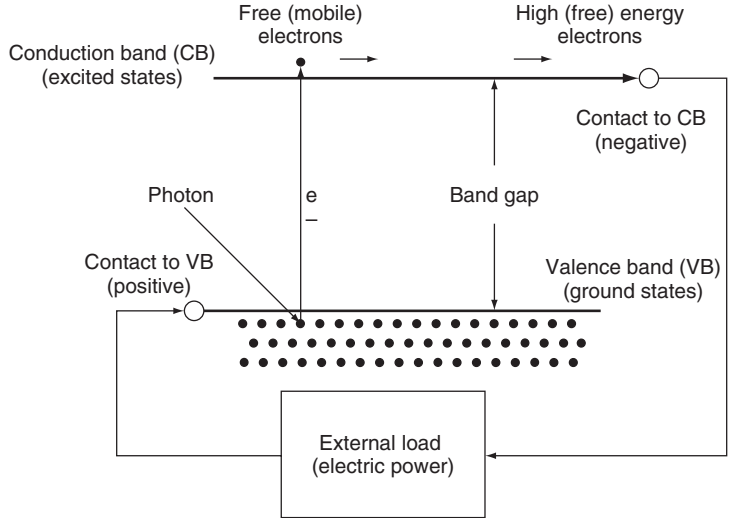


Figure 1.2 Schematic of a solar cell. Electrons are pumped by photons from the valence band to the conduction band. There they are extracted by a contact selective to the conduction band (an *n*-doped semiconductor) at a higher (free) energy and delivered to the outside world via wires, where they do some useful work, then are returned to the valence band at a lower (free) energy by a contact selective to the valence band (a *p*-type semiconductor)

[1] The bandgap energy or energy gap is a fundamental and unique parameter for each semiconductor material. To be a good absorber of solar energy on earth, a semiconductor should have a bandgap between about 1 and 2 eV. See figure 4.3.

electrons are delivered with an energy of about 1 eV). The electrical power produced is the product of the current times the voltage; that is, power is the number of free electrons times their electric charge times their voltage. Brighter sunlight causes more electrons to be freed resulting in more power generated.

Sunlight is a spectrum of photons distributed over a range of energy. Photons whose energy is greater than the bandgap energy (the threshold energy) can excite electrons from the valence to conduction band where they can exit the device and generate electrical power. Photons with energy less than the energy gap fail to excite free electrons. Instead, that energy travels through the solar cell and is absorbed at the rear as heat. Solar cells in direct sunlight can be somewhat warmer (20–30 °C) than the ambient air temperature. Thus, PV cells can produce electricity without operating at high temperature and without moving parts. These are the salient characteristics of PV that explain safe, simple and reliable operation.

At the heart of almost any solar cell is the *pn* junction. Modeling and understanding is very much simplified by using the *pn* junction concept. This *pn* junction results from the "doping" that produces conduction-band or valence-band selective contacts with one becoming the *n*-side (lots of negative charge), the other the *p*-side (lots of positive charge). The role of the *pn* junction and of the selective contacts will be explained in detail in Chapters 3 and 4. Here, *pn* junctions are mentioned because this term is often present when talking of solar cells, and is used occasionally in this chapter.

For practical applications, a certain number of solar cells are interconnected and encapsulated into units called PV modules, which is the product usually sold to the customer. They produce DC current that is typically transformed into the more useful AC current by an electronic device called *an inverter*. The inverter, the rechargeable batteries (when storage is needed), the mechanical structure to mount and aim the modules (when aiming is necessary or desired), and any other elements necessary to build a PV system are called the *balance of the system* (BOS). These BOS elements are presented in Chapters 19–21.

Most of the solar modules today in the market today are made of crystalline silicon (c-Si) solar cells (Chapters 5–7). About 10% are made of the so-called thin film solar cells (TFSC), comprising in reality a variety of technologies: amorphous silicon (a-Si, Chapter 12), copper indium gallium diselenide (CIGS, $Cu(InGa)S_2$, Chapter 13), cadmium telluride (CdTe, Chapter 14), and others (Chapter 11). Many think that thin film cells are more promising in reducing costs. There is also an incipient market of concentrator photovoltaics (CPV) where expensive and efficient multijunction (MJ) solar cells receive a high intensity of sunlight focused by concentrators made of lenses or mirrors (Chapters 8 and 10). The motivation of all these technologies is the same: to decrease the module costs compared with the dominant Si technology. Other options are under research and development, including organic solar cells (Chapters 15 and 16) and the new (or third) generation solar cells (Chapter 4).

1.2.1 Rating of PV Modules and Generators

A fuel-fired power generator is rated in watts (or kW or MW). This means that they are designed to operate producing this level of power continuously, as long as they have fuel, and will be able to dissipate the heat produced during its operation. If they are forced to operate at more than the rated power, they will use more fuel, suffer more wear and have a shorter lifetime. Some can be operated at lower power output, although with loss of efficiency, but many cannot be controlled at less-than-rated power.

PV modules, instead, are rated in watts of peak power (W_p). This is the power the module would deliver to a perfectly matched load when the module is illuminated with 1 kW/m^2 of

insolation (incident solar radiation) power of a certain standard spectrum (corresponding to bright sunlight) while the cell temperature is fixed at 25 °C. An array of modules is rated by summing up the watts peak of all the modules.

These "standard test conditions" or STC are universally applied to rate peak power output of a solar cell in a laboratory or a module out in the field, but rarely occur in real outdoor applications (see Chapter 18 for a complete discussion of testing conditions and Chapter 22 for real outdoor conditions). Generally, the irradiance (insolation power) is smaller and the temperature higher. Both factors reduce the power that can be delivered by the module to the matched load. In some cases the load is not so well matched (or the modules among themselves) reducing further the power. Thus while the output power is well defined under these STC, output power under real conditions varies considerably. While a 10 kW diesel generator produces 10 kW so long as it has diesel fuel, a 10 kW PV array will produce from perhaps 0–11 kW, depending on sunlight and temperature.

To enable useful predictions, the energy (not power) in kW h produced by the solar radiation falling in a generator in one year (or one month or one average day) is obtained by multiplying the rated power in kW_p times the number of "effective hours" of irradiance falling on the generator in one year (one month, one average day) times the performance ratio (PR), which accounts for losses above mentioned in real operation plus those in the wiring, the inverter (whose efficiency may be 0.90–0.97), etc. Time for maintenance is also included here. The PR in well-designed installations varies from 0.7 to 0.8 as discussed in Chapter 19, but may be even lower in warmer climates because the efficiency of the cell is reduced with the temperature.

What are the "effective" sun hours? Since the rating irradiance is 1 kW/m^2, the number of "effective" hours at the rating power is the number of kWh/m^2 falling on a plane with the same orientation of the PV generator. Thus, a typical mid-latitude location might receive a daily average of 4 kWh/m^2 of sunlight integrated over a period of 24 hours (including night time) on a horizontal surface, due to an incident power that ranged from 0 to 1 kW/m^2. This is equivalent to a constant incident solar power of 1 sun = 1 kW/m^2 for a period of only 4 hours, hence 4 'effective sun hours'. Locations such as Phoenix (United States), Madrid (Spain), Seoul (South Korea) or Hamburg (Germany) have respectively, 2373, 1679, 1387 and 1059 kW h/m^2 per year (or equivalently the same number of effective hours) for optimally oriented surfaces (facing south and tilted about 10° below the latitude). In these locations a PV plant of 1000 kW optimally oriented, with PR = 0.75 will produce 1 779 375; 1 2259 250, 1 040 250 and 793 857 kW h in one year. Table 1.2 shows, for four widely varying cities, the average daily input in solar irradiance, equivalent hours of full sunlight (at 1 kW/m^2), and average annual yield in kW h from each kW of installed PV, assuming a system performance ratio PR = 1. Once multiplied by the actual PR, this average yield is independent of the efficiency or area of the modules, thus demonstrating the simplicity of this method. These represent close to the entire range of sunlight conditions found where most people live. A world map with the effective hours on horizontal surface (kW h/m^2·year) is presented in Figure 1.3.

Table 1.2 Daily irradiance (kW/m^2), equivalent daily "sun hours" of 1 sun = 1 kW/m^2, and annual energy production per kW of installed PV (assuming PR = 1), all for optimum latitude tilt

City	Phoenix	Madrid	Seoul	Hamburg
Daily kW h/m^2	6.5	4.6	3.8	2.9
Daily "sun hours"	6.5	4.6	3.8	2.9
Annual kW h/kW	2372	1679	1387	1058

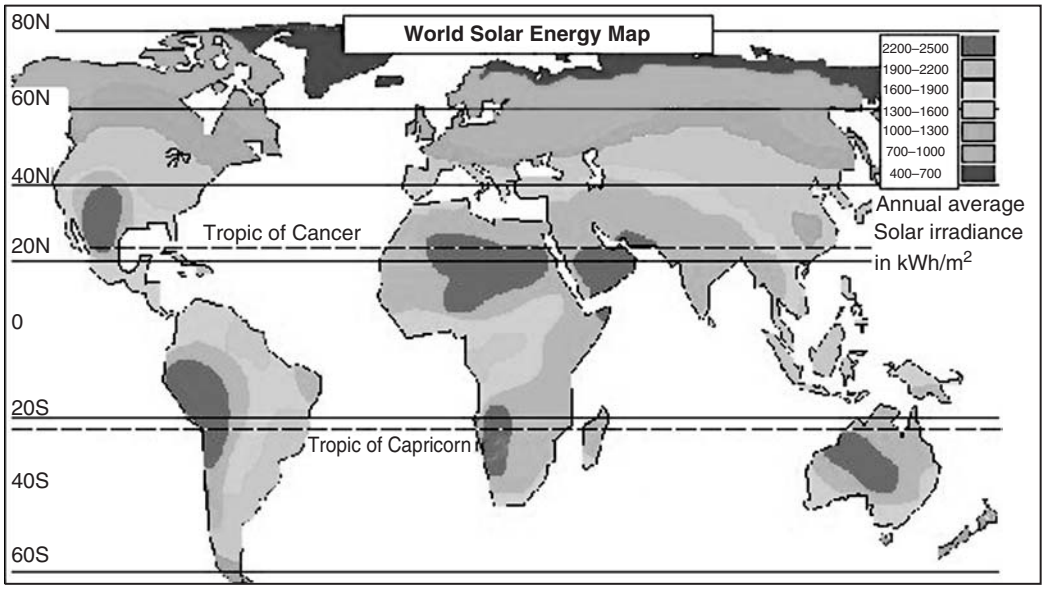


Figure 1.3 World distribution of the annual solar radiation (kW h/m^2) [obtained from www.rise.org .au/info/Applic/Array/image003.jpg] See Plate 1 for the colour figure

The rating of concentrator plants is still a subject of debate. Rating such a plant by summing the rating of the modules may be impossible as some concentrators do not have modules or they are too big for indoor measurements. However in other concentrators it might be applicable.

Chapter 22 contains much more detailed methods to calculate the incident sunlight and the PV module output as a function of location, time of day, month of year, etc. or various on-line calculators are available [5].

1.2.2 Collecting Sunlight: Tilt, Orientation, Tracking and Shading

Potential residential or commercial PV customers often worry "Does my roof have the right slope? Does my house have good solar exposure?" These are indeed important questions for fixed non-tracking arrays. Chapters 19 and 22 address these in more detail. The tilt angle to optimize yearly production for fixed non-tracking arrays is usually some few degrees below the local latitude (there is more insolation in summertime). However, many people are surprised to find that annual output is only weakly dependent on tilt, hence the slope of their roof. In fact, nearly any reasonable tilt is good, and even flat roofs are good for solar below 45° latitude. For example, at mid-latitudes, the difference in annual averaged effective hours varies by 10% as the tilt angle of the modules varies from horizontal (0°) to latitude tilt. Thus, for a home in Washington DC or Madrid or Seoul or Wellington, New Zealand, all at very roughly 40° latitude, the difference in annual effective hours between a horizontal flat roof (~4.4 effective hours per day) or a 40° tilted roof (~4.6 h/day) is 5%. The reason is that the sun's angle at that latitude varies from 27° to 72° between winter solstice to summer solstice at this latitude. In winter, a steeper roof will have more output than a shallow slope, and vice versa in summer, so the difference between flat and tilted averages out somewhat during the year.

What about orientation? For solar installations in the northern hemisphere, the optimum orientation for fixed non-tracking arrays is true south. But again, it is not very sensitive to minor deviations. An array oriented to the southeast will get more sunlight in the morning and less in the afternoon. Thus, for an array installed at 40 °N latitude with 40° tilt and oriented from 45° east or west of true south, the annual output will be only 6% less compared to the optimum true south orientation.

Or, you can install modules on movable supports that "track" the sun. They can track from E to W (oriented in long N–S linear arrays) called single-axis trackers. They can also be installed on special mounts that track the sun in both its daily E–W motion across the sky and its daily and seasonal variation in vertical height, called two-axis trackers. Single- and double-axis tracking generally increase the sunlight collected by 15–20% and 25–40%, respectively. They typically are only employed in large, utility-scale ground-mounted arrays. Of course, costs are higher than for fixed-mount arrays.

So, are there any limits to the location for the installation of an array such as on a roof or in a farmer's field? Yes! The array must not have much shadowing on it, at least not during the peak production hours from 9 am to 3 pm (solar time). The first obvious reason is that the shaded parts produce negligible energy because although PVs can operate with diffuse light, the amount of energy in this diffuse light is rather small. But there are other effects that are more insidious. Even a slight shadow, such as due to a thin pole or leafy tree, on a corner or edge of a module could dramatically reduce the output from the shadowed module and also from the entire array. This is because the modules are connected in series; restricting the flow of current in one cell will restrict the output of all other cells in that module and thus in all modules connected in it in series. But the use of bypass diodes in series strings reduces these losses to very acceptable values. This topic is further analyzed in Chapters 7 and 21. The shadow issue may present a significant limit in cities or towns with lots of trees or tall buildings. A proper preinstallation design will include a shading analysis. Some governments are considering "guaranteed solar access" laws to prevent a newly constructed building or neighbor's trees from shading another roof's array, but the legal problem is not trivial.

1.2.3 PV Module and System Costs and Forecasts

Although the important figure of merit for cost is \$/kW h, typically \$/W_P is used. Policy makers and consumers alike often ask "How much do PV modules cost?" Prices for the same module can differ from country to country. There are challenges of discussing a unique module price even within a single country such as Germany with a very mature and well-regulated PV market, educated consumers and high-volume installers. For example, using average module selling price data in Germany during 2009 [6], the factory gate price for c-Si modules was 2.34 €/W. Due to the excess inventory caused by the failure of the Spanish market in 2009 (a fact that will be explained later), the "market" price for c-Si modules was 16% lower. Market prices for less-efficient thin film a-Si and CdTe modules were about another 10% lower, approaching 1.50 €/W. Market prices for c-Si modules made in Asia were 19% below the average. This is consistent with a more detailed study showing 25% higher costs due to labor for a hypothetical 347 MW c-Si PV module factory in the US or Germany compared with China [7]. This range of module pricing in the most advanced PV market in the world indicates the difficulty of answering the question "how much does a module cost".

But what about the cost for complete systems? This is what really determines the price of solar electricity. We turn to a report analyzing installed costs of 52 000 PV systems (566 MW) installed in the US, mostly in California, from 1998 to 2008 [8]. The average price, before applying any incentives or state refunds, decreased from \$US 10.8/W to \$US 7.5/W, a 3.6% annual decrease.

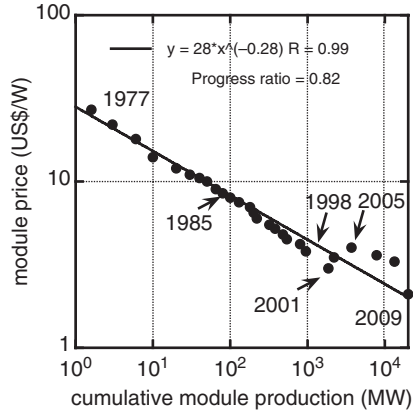


Figure 1.4 Experience curve for photovoltaics from 1976 until 2009. Straight line is fit indicating an experience factor of $1 - 2^{-0.28} = 0.18$ or equivalently a progress ratio $2^{-0.28} = 0.82$

As expected, prices decreased as the system size increased (2008 prices): \$US 9.2/W for small (2 kW) residential systems versus \$US 6.5/W for large (500–750 kW) commercial-scale system. Excluding any taxes, installed prices of residential systems in 2008 was \$US 6.1/W in Germany, \$US 6.9/W in Japan compared with \$US 7.9/W in the US. But prices decreased significantly in 2009 as this was being written.

Therefore, any discussion of module or system prices is complicated by numerous factors, including location, size of the system, discounts or incentives, and the PV technology. Furthermore, it is strongly time dependent. Nevertheless, analysts worldwide commonly assume some price in order to analyze trends and market influences, as in Chapter 2. A common method predict the cost evolution is the so-called learning curve that states the "price" (whatever definition of this is adopted) of the modules is reduced by a factor 2^n every time the cumulated production is doubled. Figure 1.4 shows a learning curve for PV modules based on their past prices. It suggests that to reach \$US 1/W at the present rate will require an order of magnitude increase in cumulative production.

1.3 PHOTOVOLTAICS TODAY

1.3.1 But First, Some PV History

The history of photovoltaics goes back to the nineteenth century. The first functional, intentionally made PV device was by Fritts [9] in 1883. He melted Se into a thin sheet on a metal substrate and pressed an Ag-leaf film as the top contact. It was nearly 30 cm^2 in area. He noted, "the current, if not wanted immediately, can be either stored where produced, in storage batteries, ... or transmitted a distance and there used." This man foresaw today's PV technology and applications over a hundred years ago. The modern era of photovoltaics started in 1954 when researchers at Bell Labs in the US accidentally discovered that *pn* junction diodes generated a voltage when the room lights were on. Within a year, they had produced a 6% efficient Si *pn* junction solar cell [10]. In the same year, the group at Wright Patterson Air Force Base in the US published results of a thin film heterojunction solar cell based on Cu_2S/CdS also having 6% efficiency [11]. A year later, a 6% GaAs *pn* junction solar cell was reported by RCA Lab in the US [12]. By 1960, several key papers by Prince [13], Loferski [14], Rappaport and Wysocki [15], Shockley (a

Nobel laureate) and Queisser [16], developed the fundamentals of *pn* junction solar cell operation, including the theoretical relation between bandgap, incident spectrum, temperature, thermodynamics, and efficiency. Thin films of CdTe were also producing cells with 6% efficiency [17]. By this time, the US space program was utilizing Si PV cells for powering satellites. Since space was still the primary application for photovoltaics, studies of radiation effects and more radiation-tolerant devices were made using Li-doped Si [18]. Similar achievements took place in the former USSR whose Sputnik II satellite in 1957, was already powered with silicon cells. In 1970, a group at the Ioffe Institute led by Alferov (a Nobel laureate), developed a heteroface GaAlAs/GaAs solar cell [19] which solved one of the main problems that affected GaAs devices and pointed the way to new device structures. GaAs cells were of interest due to their high efficiency and their resistance to the ionizing radiation in outer space. A significant improvement in performance occurring in 1973 was the "violet cell", having an improved short wavelength response leading to a 30% relative increase in efficiency over state-of-the-art Si cells [20]. GaAs heterostructure cells were also developed at IBM in the US having 13% efficiency [21]. Finally, in October 1973, the first world oil embargo was instituted by the Persian Gulf oil producers. This sent shock waves through the industrialized world. Several governments began programs to encourage solar energy, ushering in the modern age of photovoltaics and giving a new sense of urgency to research of photovoltaics for terrestrial applications.

An excellent history of the PV early times can be found in a book by John Perlin [22] or, more briefly, in Chapter 1 of the first edition of this book.

In the 1980s, the industry began to mature, as emphasis on manufacturing and costs grew. Manufacturing facilities for producing PV modules from Si wafer *pn* junction solar cells were built in the US, Japan, and Europe. New technologies began to move out of government, university and industrial laboratories, and into precommercialization or "pilot" line production. Companies attempting to scale up thin film PV technologies such as a-Si and $CuInSe_2$, which had achieved $>10\%$ efficiency for small area ($\sim 1\,cm^2$) devices made with carefully controlled laboratory-scale equipment, found that this was far more complicated than merely scaling the size of the equipment. Unfortunately, by the 1980s most large US semiconductor and oil companies gave up their R&D or pilot-scale efforts in the absence of large infusions of private or government support. One common result was the purchase of American companies and their technologies by foreign companies, displacing the center of the PV industrial activity from the US to Japan and Europe and later to China, currently the world's largest solar cell producer.

1.3.2 The PV Picture Today

In the last decade (1998–2008) the market of PV modules has multiplied by more than 20. The explosive growth transformed PV from a dream for environmentally conscious citizens to a reality that attracts investors eager to exploit this new Eldorado.

Who is making all the PV modules? Figure 1.5 shows where these modules have been produced. The US led the world in production during most of the 1990s (not shown) when Europe and Japan had relatively static manufacturing growth. Then in 1998, progressive and supportive government policies in Germany and in Japan resulted in substantial increases in their production. These policies were driven partly by a strong commitment to CO_2 reduction, as prescribed by the Kyoto Protocol, and partly to develop PV as an export.

But the big story in PV production since the first edition was published is the rapid rise of Chinese production since 2006. In 2003, none of the top ten manufacturers were from Asia. In 2008, three are from China and one from Taiwan. In 2009, China is expected to appear as the top manufacturing location.

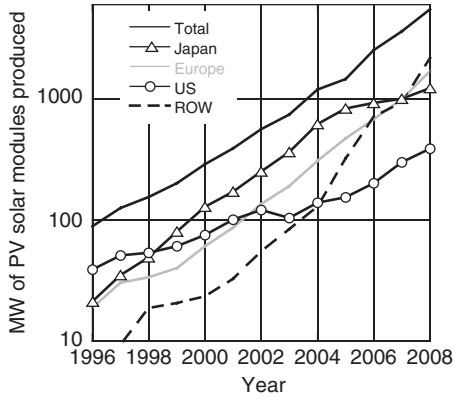


Figure 1.5 Production of PV modules by country or region (ROW = rest of world, mostly China and Taiwan, Europe is mainly Germany and Spain)

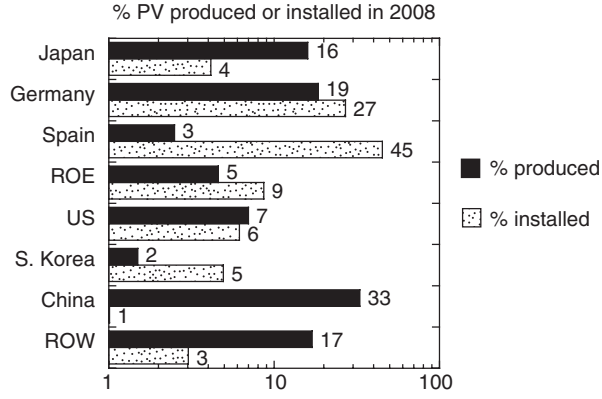


Figure 1.6 Percentage of PV cells/modules produced [23] and installed [24] in 2008 by country or region. Note the logarithmic scale. ROE = rest of Europe, ROW = rest of world (mostly Taiwan and India.). Total installed = 5500 MW; total produced = 7900

Where are the modules being installed? Figure 1.6 shows the geopolitical breakdown for 2008 including the top four producers – China, Germany, Japan, and the US – and the top four installers – Spain, Germany, US, and South Korea. Note that Spain installed 15 times more PV than they produced and China produced 30 times more than they installed. The US was nearly balanced in terms of imports and exports (6–7%). In 2008, Spain became the top destination for PV modules for the first time, overtaking Germany. Most of these modules were installed in large centralized plants >10 MW such as the one presented Figure 1.10. But the brief two years of Spanish leadership in PV installation, created by favorable feed-in tariff (FIT) legislation, ended in 2009 due to restrictive modifications of the law.

The difference between production and installation for 2008 has been variously quoted as 1500–2500 MW. This may represent a double counting of production caused by including both the cells produced in one factory and delivered to a second one for making into modules, and also counting the modules made in the second factory. Some believe that this discrepancy can be caused

by a surplus in inventory of unsold modules, but this is doubtful because the module prices were high throughout 2008 and the market experienced a shortage. Prices only decreased significantly in 2009, when the Spanish markets collapsed.

1.3.3 The Crucial Role of National Policies

The real origins of today's surging PV growth started in the mid 1990s when residential scale grid-connected applications in Europe and Japan began to grow rapidly, primarily owing to strong government support. Until then, the primary destination for a PV module was in an off-grid application, whether a rural home in developing country, a water pump for cattle or people, a vacation cabin in the mountains, or a radio transmission antenna. The relative change in dominance of the three main PV applications – off-grid, grid-connected residential and commercial, and large utility scale – is shown in Table 1.3. There are two types of incentive programs responsible for the success of grid-connected residential or commercial applications. One approach, pioneered in Japan and later copied by many US states, provides home or business owners a rebate from the government or their electric utility agency for 10–50% of the PV system cost. Then, their electric bill is determined by the utility using "net metering" where the customer pays only the net difference between what they used and what they generated. Thus, they get a reduction in the initial price of the system, and their PV electricity is valued at the same rate as their utility electricity. This initiated the new market of grid-connected residential and commercial buildings, PV's first big burst in growth, beginning around 1995. Interestingly, government support of photovoltaics in Japan has been decreasing while the market for PV homes has continued to show a good growth rate. The second approach, pioneered by Germany, paid the home or business owners for the electricity they feed into the public electric grid at a rate that is several times greater than the rate that they buy electricity from the grid. Additionally, German banks provided generous loans for purchasing the installation. But there is no government rebate to reduce initial cost. This has resulted in solar arrays being installed on German houses, barns, commercial roofs, government buildings, schools, dairy farms, abandoned airports, and parking garages – in short, any place they can face the sun and still be connected to the grid. This concept, called either a feed-in tariff (FIT) or production tax credit (PTC), has been implemented in Spain, the Netherlands, South Korea, Canada, recently in Japan and soon will be in a few municipalities or states in the US. The German FIT initiated the second great wave in grid-connected PV growth in the mid 2000s. Chapter 2 discusses these and other funding policies to promote PV, including solar renewable energy certificates (SRECs) and mandated renewable energy portfolios (REPs). And the Spanish FIT resulted in the explosive growth of utility scale PV projects in 2007–2008. While many people, especially PV engineers and scientists, might think otherwise, the rapid growth of grid-connected PV, hence all PV, is due more to innovations in policy than to advances in technology.

Table 1.3 Approximate percentages (±20% relative) of different types of application installed

Year	Off-grid (%)	Grid-connected residential or commercial (%)	Grid-connected utility scale (%)	Total installed each year (MW)
1996	95	5	<1	89
2000	60	40	1	288
2004	30	68	2	955
2008	10	35	55	5600

Table 1.3 shows very approximate percentages (±20% relative) of each type of application installed in that year. Off-grid includes single module rural homes, cabins, water pumping, diesel hybrids, and remote communication transmitters. Grid-connected residential and commercial is typically roof-mounted arrays <200 KW in 1996 and 2000, but maybe <1000 kW in 2004 and 2008. Anything larger defined as utility scale. Data and definitions vary from a variety of sources [25].

The importance of the Spanish market in 2007–2008 (3.4 GW_p in total) and its sudden collapse deserves some reflection. The subject is very well explained in Chapter 2, but we want to add here some additional thoughts. The FIT, issued by a Royal Decree, guaranteed a generous price (of about 43–46 €cents/kW h in 2008) over 25 years for the electricity privately produced and sold to the grid by a PV installation, finished and registered in that year. Three Royal Decrees were necessary to permit the market to expand. The first Royal Decree limited FIT payments to generators of less than 5 kW. Some entrepreneurs circumvented this limitation by gathering many small investors and making bigger plants, called solar farms, where each investor owed 5 kW. A second Royal Decree in 2004 lifted the power limitation per owner to 100 kW. The reaction of entrepreneurs was to register dozens of 100 kW installations to effectively create MW plants. Clearly there was higher profitability in larger installations. Finally in 2008 any restriction on size was removed, triggering a frantic activity to build big plants in Spain, so that by the end of 2008 they had built 40 of the 50 biggest plants in the world [26], totaling 2.6 GW in only one year, including the world's biggest PV installation at Olmedilla de Alarcón of 60 MW. No other energy technology can expand so quickly. A 1 GW nuclear plant with about five times more capacity factor (equivalent to 5 GW PV plants) would require at least 10 years to be built, so utility scale PV can be built five times faster.

The trend to build big plants has continued and by early 2010 there were 15 plants with power above 25 MW: 8 in Spain, 5 in Germany, 1 in Portugal and 1 in the US [26]. There are 1000 plants in the world of 1.3 MW or bigger totaling a power of 4.5 GW out of the total 14.7 GW installed. Thus, the Spanish FIT was responsible for the third wave of steep growth, namely that of utility scale projects after 2006.

Unfortunately, the unexpected success of this program resulted in overwhelming the funds which had been allocated, requiring a significant reduction in scale. The collapse of this market in 2009 reportedly resulted in the firing of 25 000 of the 75 000 people working in PV in Spain. It was caused by the unexpectedly strong and fast growth of the market and downturn in the world economy, neither of which was foreseen by the Government. The present regulation plans for a market of only 500 MW in 2009, divided between installations smaller than 20 KW (26.7 MW) and larger than 20 kW (241.3 MW) and the rest in ground installations. The offer for ground installations has largely exhausted the quota by several times (the excess being on a waiting list) while the quota is still open for roof-mounted plants. To date, the FIT regulations have been the most successful policy to expand PV. They explain why Europe (where the climatic conditions are not the best) leads the world in PV installation, with over 70% of the world's total cumulative installed PV power. Some countries are reluctant to adopt a FIT program, sometimes because they consider it an unacceptable intrusion into the market laws.Yet the FIT allows the competition between various PV technologies (Si wafer, thin films, concentrators), installation strategies (roof-top, ground mount, BIPV) and companies. A well-designed FIT must decrease with time (the time when the installation is commissioned, not the time years later when the kW h are sold) in a predictable way to force the industry to reduce their costs and hence their selling price.

1.3.4 Grid Parity: The Ultimate Goal for PV

The ultimate goal for PV is to reach grid parity without subsidies. The relevant parameter to evaluate the long-term cost of any energy source cost is the levelized cost of energy (LCOE). Calculation of LCOE ($/kW h) for PV is complicated; it includes the output of the system over its lifetime (efficiency, solar irradiance and temperature at that location, tilt angle, system losses,

annual degradation rate), cost to install and maintain the system (design, permits, BOS costs, site preparation, replacement cost of the inverter, batteries, repairs, profits), and cost of financing the system (discount rate, loans, inflation). Here, we present results from a model developed by NREL (US) called Solar Advisor Model (SAM) [27, 28]. A price in $/W is determined in six categories: module, inverter, BOS, installation, indirect (financing, design, permits, site preparation, and profits), and operation plus maintenance or O + M (replacing inverter, module cleaning, inspection) considering a given lifetime. The calculated value of LCOE can then be compared with the grid-supplied consumer or wholesale price to determine grid parity, cost–benefit for a given installation, etc. Table 1.4 shows the breakdown in costs and assumptions for four systems using SAM: 4.5 kW residential, 150 kW commercial, 12 MW single-axis tracking at tilt, and 12 MW two-axis tracking concentrator system. Parameters common to all are listed in the table. In the reference cited, each system is analyzed for 2005 (benchmark year using real data), and projected to 2011 and 2020 (using extrapolated data representing improved performance and costs). We list projected results for 2011. These results should not be taken as absolute values, but rather as relative comparisons.

According to this model, the residential system has the highest inverter, BOS, installation and LCOE costs. This derives from the higher design and fixed costs for the smaller systems, and less economy of scale. Note that the smaller inverters have lower lifetime, partly because they are supposed to lack professional inspection services. In reality, smaller systems like this would probably have higher module prices as well since volume purchases drive down costs. The 150 kW commercial system has the lowest LCOE costs, due mostly to the assumed lower installation costs. The two different 12 MW utility systems have very low inverter costs in common; most other parameters are quite different, yet they end with the same LCOE. They both have higher BOS costs, possibly due in part to having to purchase or rent land, unlike the other two rooftop applications. Module price is less than half of the total system price for the 4.5 kW residential, while it is more than half of the other types of systems. The LCOE ranges from $US 0.10 to 0.15/kW h, compared with the expected retail price for standard grid electricity of $US 0.07–0.12/kW h. An often neglected aspect of PV systems is the lifetime. Increasing the lifetime from 20 to 30 years decreases the LCOE by $US 0.023/kW h or 13%.

Table 1.4 LCOE analysis from NREL model SAM for four different systems, located in Phoenix (6.5 kW h/m^2·day) and projected for 2011. Common parameters: 96% inverter efficiency, 35 year lifetime, 1% degradation/year

Parameter (units)	Residential	Commercial	Utility 1-axis tracking	Utility concentrator
System rating	4.5 kW	150 kW	12 MW	12.5 MW
Module price ($/W)	2.20	2.20	2.20	3.00
Efficiency (%)	16	16	16	25
Inverter price ($/W)	0.69	0.51	0.35	0.35
Inverter life (yr)	10	15	15	15
BOS ($/W)	0.40	0.36	0.73	0.53
Installation ($/W)	0.57	0.17	0.16	0.33
Indirect ($/W)	1.14	0.76	0.46	0.09
O + M (% costs)	0.3	0.3	0.3	0.6
Installed price ($/W)	5.00	4.00	3.90	4.30
LCOE ($US/kW-hr)	0.15	0.10	0.12	0.12
Retail price of electricity ($US/kWh)	0.12 Residential	0.10 Commercial	0.07 Industrial	0.07 Industrial

Compared with the retail electric rates for the US shown at the bottom, the LCOE cost of a PV system in a very sunny location is very close to grid parity, in particular for the so-called commercial applications (roofs of factories or supermarkets). This is why in the US there is increasing interest in this type of system. Of course, such predictive calculations are only as valid as their assumptions, which do seem reasonable.

There is considerable uncertainty in future electric rates due to possible environmentally driven costs (carbon tax) and the need for substantial reinvestment in new transmission and grid control systems (which will probably be paid for by customers through higher rates). Thus, predicting the price of either PV or conventional electricity is problematic.

Figure 1.7 shows the dependence of the LCOE on available sunshine (irradiance in annual kW h/m^2) for module prices of \$US 1 or 2/W. The system was assumed to be a 2.5 kW residential array with a 30 year mortgage at 6% and 30 year lifetime with 1% annual degradation. The other fixed costs were taken from Table 1.4, and amounted to \$US 2.80/W. The module price of \$US 2/W is close to today's selling price, and \$US 1/W represents what many believe is achievable with reasonable advances in today's technologies in the near future. The symbols indicate the average residential electricity rate for various cities indexed by their annual irradiance. Already in 2009, PV electricity prices matched a few markets in the US and Europe that have relatively high-price electricity and/or high solar irradiation such as Italy, Hawaii, New York and California (San Diego, not shown), when determined as the LCOE. Similar analysis has been published using different assumptions (system cost of \$EU 3/W) but showing similar trends and conclusions [7]. In reality, the cost of conventional electricity is presently less than the PV LCOE for most cities but the gap is narrowing. The LCOE cost would be less for a larger system as in Table 1.4 due to lower installation costs. This graph shows that grid parity is a complex relation between geographic location, price of electricity, and price of PV system (related to size of PV array). Note that increasing the incident sunlight, hence module output energy, by a factor of two reduces the LCOE by the exact same amount, thus having significantly greater impact than decreasing the price by a factor of 2.

In order to reduce the LCOE, it is acknowledged that efficiency must increase and cost must decrease. But what is their relative contribution? Where should we put our efforts? Figure 1.8(a) presents the dependence of LCOE on module efficiency for module costs of \$US 1 and 2/W (again) for a 2.5 kW system in two US locations, sunny hot Phoenix or temperate coastal Philadelphia,

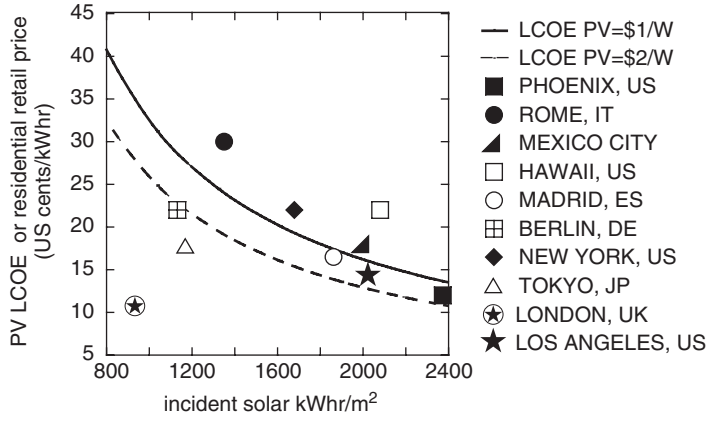


Figure 1.7 LCOE and cost of residential electricity versus average annual incident irradiance. LCOE calculated using SAM model parameters explained in the text assuming a 2.5 kW array with \$US 1 or 2/W module cost

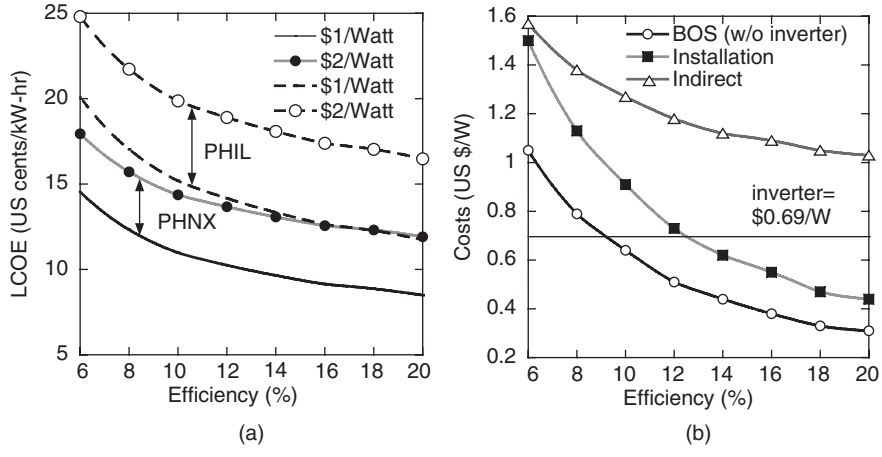


Figure 1.8 (a) Calculated impact of efficiency and module cost on LCOE for Phoenix (sunny, 2370 kW h/m^2/yr, solid lines) and Philadelphia (temperate, 1680 kW h/m^2/yr, dashed lines) for 2.5 kW systems at latitude tilt. (b) Fixed costs as function of efficiency. Indirect for $US 2/W module. A 30 year lifetime is assumed in all cases. Other assumptions are described in the text or caption for Figure 1.7

calculated using SAM with the same parameters as in Figure 1.7. (Philadelphia has comparable solar irradiance to Shanghai, Melbourne, New York or Madrid.)

As efficiency increases, the module area and hence area-related costs decrease. In this study, the number of 1 m^2 modules decreased from 41 to 12 as efficiency increased from 6 to 20%. The area-related costs were: BOS = \$64/m^2, installation \$91/m^2 and indirect \$180/m^2 as derived from Table 1.4 for the 2011 residential case. The 2.5 kW inverter cost \$1725 (\$0.69/W). For a factor of 2 decrease in module cost, the LCOE decreases by 12% (at 6% module efficiency) to 22% (at 20% module efficiency), independent of location. For a factor of 2 increase in efficiency (from 10 to 20%), the LCOE decreases by ~20%, independent of location. For a factor of two decrease in module cost (at 16% efficiency), the LCOE decreases by ~25%, independent of location. Thus both price and efficiency have comparable impact on LCOE. This weak correlation between either module cost and efficiency with LCOE might surprise some readers, but this just indicates the importance of area-related and fixed costs. The indirect costs are much greater than either the BOS or installation. This could be alleviated with lower interest rates for PV projects, and shorter mortgages. Figure 1.8(b) shows how the area related costs decrease with efficiency. Decreases in both LCOE and area related costs are steeper at low efficiency, becoming less sensitive at higher efficiency, suggesting the relatively greater impact of increasing the efficiency of low-efficiency thin film modules (rather than lowering their price) compared with high-efficiency Si modules. Together, these figures indicate the importance of efficiency and module cost as a driver to lower LCOE, but also the impact of reducing fixed or area-related costs as well, which are often neglected by research and development funding agencies. Finance charges must be minimized.

1.4 THE GREAT CHALLENGE

In this section, we discuss the requirements and limitations to very large-scale PV energy production in terms of the amount of land and raw materials needed, the environmental impact, the net energy balance, the reliability, and the readiness of manufacturing capacity.

First, what is size of the task? Estimating the world's demand for electricity in 2030 or 2050 is complicated by many assumptions. Will a larger fraction of our primary energy be shifted from chemical and thermal to electric (i.e. electric cars)? How large a role will efficiency and smart growth play in decreasing demand? How large a role will population and economic growth in developed versus developing countries play in increasing demand?

The 2007 Nobel Prize winning organization UN Intergovernmental Panel on Climate Change (UN-IPCC) [29] estimates the world will need the equivalent of 32 000 terawatt hours (1 TW h = 10^9 kW h = 10^{12} W h) of electrical energy by 2030, but efficiency improvements might reduce this to 22 000 TW h. Analysis by their Mitigation Working Group III on how this can be best accomplished to minimize the cost per ton of avoided CO_2 concludes that PV could only meet about ~1–2% of this demand (~150–300 TW h), limited largely by *cost* not technology or resource availability. This would be mostly to meet demand from rural electrification in developing countries. However, other bodies set higher targets for PV. The European Photovoltaic Industry Association (EPIA) predicts that PV could provide 12% of Europe's energy by 2020. The International Energy Agency anticipates the PV could provide over 11% of the world's electricity by 2050 [30]. California has a mandate to produce 33% of their electricity by renewable energy by 2020. Originally, PV was expected to contribute only about 10% relative (about 3.2% total) but due to recent PV price decreases,[2] the relative fraction of PV has been increased to 40% (about 15% total). Thus, Europe and California have similar expectations for PV, and are consistent with analysis of the ability of a well-regulated grid, without storage, to accept a variable energy source such as PV sets an upper limit of 10–20% by 2030 [31]. For instance, in Spain PV already provides about 2% of the annual demand, but if we add wind energy, the total intermittent energy is actually 14%.

How can we calculate how many GW (or more precisely GW_p) of installed PV would be required to produce a TW h of electricity?

In fuel power plants the concept of capacity factor (CF) is widely used. It is the ratio of the energy actually produced to the theoretical maximum energy that the plant could produce in one year (typically equal to the nameplate power rating times 8760 h). Thus, a 1 GW power plant operating at full capacity for half the year (or at half capacity for the entire year) has a CF = 50%, and would produce 1 GW × 8760 h/yr × 0.50 = 4380 GW h/yr. The CF is usually below unity not only because the plants might need to stop for maintenance (the case of nuclear plants) but also because the management of the electric grid requires some plants to idle for certain periods.

In a PV plant CF is calculated as the effective sun hours (see Section 1.2.1) times performance ratio divided by the total number of hours in the year (even though the sun only shines 50% of the year on any location on earth). The CF for PV generators ranges from 0.08 (Hamburg, fixed panels) to 0.26 (Albuquerque, sun tracking[3]), so we will use 0.15 as an average. Thus, to provide 300 TW h in 2030, we would need 230 GW of installed PV capacity, or an installation rate of 11.6 GW per year for the next 20 years. In fact, the world produced about 7 GW in 2009, and

[2] In 2008, with installed system costs of $US7/W or $US0.30/kW h, it was expected most of California's 3.2% PV demand would be large, centralized power plants. By 2009, system costs fell to $US3.70/W or $US0.17/kW h due to availability of lower-cost thin film modules. Combined with land use permitting and transmission access difficulties for the large PV power plants, distributed small-scale PV looked more attractive to provide much of the 15% demand (from Garrett Hering in Photon International December 2009, 12–14.)

[3] Sun tracking increases the CF. The number of equivalent hours in a sun-facing plane (called a two-axis tracker) increases by about 40% with respect to the fixed optimally oriented module. This is discussed further in Chapters 19 and 22.

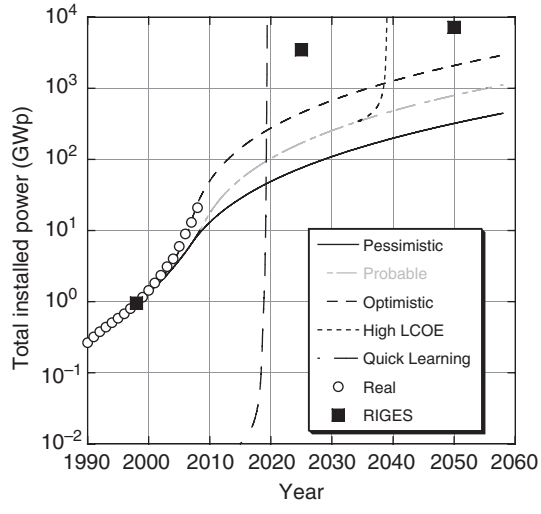


Figure 1.9 Forecast of the cumulative global PV installed power by year. (see [32] for details). "Real" represents the actual installed power (from P. Maycock, *Photovoltaic News*, 19(3) 1 (2000), more recent data from the *Photon International* yearly reports in the March issue). The dots labeled "RIGES" are the goals in Reference [33]

expects around 10–12 GW in 2010. So meeting the relatively low expectations of the UN-IPCC for PV will be easy, requiring no growth or scale-up at all.

In 2001 one of us published a paper [32] in which a differential equation for the evolution of the yearly market was formed by coupling the learning curve (Figure 1.4) and the demand elasticity (the relative reduction of price in a certain year divided by the relative increase of the yearly market). The price evolution was deduced as well as the cumulative sales for every year. For the initial forecasting period, the model parameters were extracted from past experience in yearly markets and prizes. The cumulative sales, approximately equating the total installed PV power, are represented in Figure 1.9 for several parameter choices.

The three curves: "optimistic" (developed countries are willing to spend 0.1% of their GDP on higher cost PV electricity), "probable" (spending 0.05%), and "pessimistic" (spending 0.025%) show a rapid increase followed by a slowing in growth as markets saturate compared with a given price of electricity. Note the striking coincidence between the actual PV installations up through 2008 (labeled "real" in the figure) and the "optimistic" forecast predicted in 2001. This coincidence gives credibility to the other implications predicted by this model. For 2030 this model predicted 1.6% of the total demand of 22 000 TW h forecasted by the IPCC under the hypothesis considered "probable" in 2001. This is not very different from the modest contribution of 1.4% assumed by the IPCC. But if we look at the real installations trend, which follows closely the "optimistic" forecast, the result is that by 2030 PV will be supplying the 4.4% of the demand. This is more than three times the prediction of the IPCC, but still modest.

So, let us assume that PV is to supply 12% (2,640 TW h) of the world's electricity by 2030, hence we would need to install on average 100 GW ($\cong 2640/(365 \times 24 \times 0.15 \times 20)$) per year (a nice round number) for the next 20 years. In fact a constant average annual production is completely unrealistic, but it serves to give the scale of the task. Some studies assume producing

20–30 GW/yr of PV until 2020 then ramping up an order of magnitude each decade until 2050 [34] while others see rapid initial growth followed by slowing as the installed capacity reaches its limit.

The paper which developed the model in Figure 1.9 stated in 2001:

> *First we predict several years of explosive market growth. . . . But this period cannot last too long. Not more than a decade. If it does, the capital involved will become excessive. Powerful voices will cease to consider PV as a curiosity and will question cost effectiveness. Other voices, not less powerful, will support PV. The equilibrium will determine subsequent growth. This equilibrium will induce a slower market growth, but at levels that are no longer negligible, at least in terms of business volume, but probably not enough for pollution abatement. Price decrease will continue, but slowly. For the next half-century they will be not competitive with common electricity unless some of the following facts happen.*
>
> (*a*) *Electricity prices rise.*
> (*b*) *Commercial schemes substantially reduce commercialization, installation and financing costs.*
> (*c*) *New inventions of lower initial costs or with more cost reducing potential appear.*

As the paper predicted, PV has become a substantial business (about €50 billion in 2008) but society has started to question its cost effectiveness, especially in view of the worldwide financial problems of 2009. For instance, there was the dramatic reduction of public support in Spain in 2009 discussed above, and less dramatic reduction in Germany in 2010, the world's two most active markets in 2008.

But let us examine now whether the conditions for competitiveness in the second part of the statement are being fulfilled or not. The model, intended to be simple, considers PV competing with only the wholesale price of the electricity, but if the retail price is taken into account instead, this would allow for a substantial growth within the first half of the century. The case is considered in the curve labeled "High LCOE". With this choice of parameters (consult the paper for details) a vertical asymptote appears around 2040 (assuming a module price of $\$1.25/W_p$). Since the actual trend is following the curve "optimistic", according to the model, the asymptote will appear before this date, maybe towards 2025. Thus, if a contract offered by the regional electric power provider is based on the retail price of electricity (essentially net metering), including the price paid to PV power plants, condition (a) would be automatically fulfilled.

The steep asymptotic increase in "High LCOE" is just an artifact of the model caused by the assumption that the electricity market is infinite. This is valid only if one knows when the LCOE "tipping point" is reached, not for studying the subsequent growth. In reality this asymptote will appear as a surge in the installations which will inevitably saturate and slow when sales to the new market sector are exhausted.

But, in addition, the model assumed that the cost of a PV installation was distributed in equal thirds, one for the module, one for the BOS and one for the commercialization, installation and financing. This is a reasonable assumption for distributed small home markets (prevalent when the model was developed) but the recent development of big power plants has drastically reduced the commercialization costs, leading to a situation where the plant cost is about twice (and not three times) the module's cost, according to the SAM model presented in Table 1.4. This fulfills condition (b) and brings the date of the surge in installations even closer.

Thus, what the model explains is what is already anticipated by many analysts, including the study presented in Figure 1.8 for a system price of \$US $2/W_p$. This price is already being offered in Spain for big ground-mounted PV power plants, and we think that expected surge will take place within the next five years in several countries (or US States) where the combination of insolation and the electricity retail price are high enough to reach grid parity with the retail electricity price.

Note that the retail electricity price includes the cost of generation as well as transmission, distribution and commercialization; therefore, it is only fairly applicable to homes and commercial PV plants. Utility-scale PV power plants should compete with the much lower wholesale price of generation. The application of the retail price to big plants is dependent on political decisions that waive the cost of the distribution for the big PV producers. Yet the application of FIT incentive suggests that this policy can be adopted in many countries.

Thus we have justified why the goal of meeting 12% of the world's electric demand by 2020 or 2030 with PV is not just wishful thinking, at least from the point of view of reaching reasonable prices.

But this is not enough. At the end of this century, many analysts expect solar energy to provide a large part of the demand, i.e. about 60% of the energy demand (not only electricity) [34]. This requires fulfilling condition (c), a technological breakthrough with a faster learning curve (curve labeled "quick learning"). It might happen at any time (2015 is assumed in the figure). It is possible that it is already happening. Consider the recent success of First Solar, a thin film manufacturer, who became the biggest solar cell producer in the world in 2009. Maybe the breakthrough will appear in 5–10 years with the concentrator systems intended to exploit the ultra-high-efficiency (over 40%) MJ solar cells. Or maybe, at an even later moment, the exploitation of novel concepts for a higher efficiency will permit this faster learning curve and therefore dominate the wholesale production of electricity. Certainly the so-called third-generation concepts aim at this. This is why we said that PV has more technological options than concentrating solar thermal electricity and we qualified PV as a winning option.

But this high penetration (>20%) will only be possible if some cheap and efficient procedure for electricity storage is developed. Advanced PV and storage have to be among the leading tasks for all this century for scientists and technologists.

Let us concentrate now on other constraints and challenges associated with our immediate 12% goal implying the installation of 100 GW_p per year, on average, during the next 20 years.

1.4.1 How Much Land Is Needed?

Estimates of *household* electricity usage from various sources for the US, Japan and Europe indicate about 5 kW h/person/day, or 20 kW h/day per family of four. With a CF of 0.15 this demand can be satisfied with a PV installation $20/(24 \times 0.15) = 5.5\ kW_p$. For a rated module efficiency (at STC) of 15% (150 W_p/m^2) this requires an area of $5500/150 = 37\ m^2$ of modules. There are many properly oriented roofs (with a lot of flexibility in the tilt and orientation, but generally facing southwards if in the northern hemisphere) with 37 m^2 of well-oriented roof available (see Chapter 23 for architectural integration). In fact, many roofs are larger, and many homes have sunny areas of this size around them, so it is possible for a family of four, with all the conveniences of a typical modern home, to provide all their power averaged over a year from PV modules on their house or on racks on their yard.

Let us assume that of the 100 GW we must fabricate each year 25 GW_p are grid-connected homes of 5 kW_p. This will imply installing 5 million residential generators per year. This is certainly a challenge, but is not impossible if we consider that the car industry produces more than 60 million cars per year. Neither is the availability of capital a problem, if the business is profitable.

Perhaps it is illustrative to know that the 125 more densely populated towns in the world [35], (whose density of population ranks from 26 650 people/km^2 in Mumbai to 1550 in Denver) with a total of 619 million inhabitants occupy an area of 124 000 km^2. With the same daily electricity consumption of 5 kW h per person and the same capacity factor of 0.15, the area of 15% efficient modules required is 129 000 km^2, practically the same as the total town area. Obviously in Mumbai

there will not be enough room for this, but in Denver it will be easy. There are many more towns less densely populated. We think this clarifies that space is not a limitation for generating a sizeable fraction of the electricity we use with PV, or for meeting our goal of 25 GW_p per year.

In the first edition we calculated how much land it would take to replace a 1000 MW coal or nuclear power plant. The answer was 60 km^2 (or 24 square miles). This is a square 8 km (or 5 miles) on a side. For the same electricity production, this is equivalent to the area for coal mining during a coal-powered plant's life cycle, if it is surface mining, or three times the area for a nuclear plant, counting the uranium mining area [36].

But building this size of PV plant is not a solution adopted by PV investors for the moment. Actually the 34 biggest plants built in the world comprise 1010 MW, occupying an area similar to the one calculated above. The size of these plants ranges between the 60 MW of the Olmedilla (Spain) plant to the 19.4 MW of the Helmeringen (Germany) plant. Figure 1.10 is a picture of the 10 MW plant at Jerez de los Caballeros (Spain) showing how PV is well integrated in the environment, permitting for instance, cattle raising. The PV "trees" are about the same size as the oak trees around, but collect solar energy about 100 times more effectively.

In any case we must not hide the fact that the annual electric energy (MW h) produced by the 1000 MW plant (or ensemble of plants) is less than the one powered by coal or nuclear energy because of the smaller capacity factor (0.15 in our example compared with at least 0.5 for the fuel plants).

Finally, how much land would be needed for PV to supply the 12% of worldwide demand (2 TW) we want for 2030. Others have analyzed very large-scale PV (VLS-PV) at seven extremely sunny, arid desert locations worldwide, roughly one per major geopolitical land area [37]. In those deserts, they would produce about 50% more energy per area than our typical, mid-latitude array analyzed above, so 100 MW requires approximately 1.7 km^2 of barren land. Scaling these areas from the VLS-PV study to our hypothetical 2 TW yields an area of 34 000 km^2, or about 5000 km^2 in each of the seven deserts. Certainly, an area of 70×70 km^2 (43×43 miles) could be found in these deserts where the installation of PV arrays would be accommodated without significant disruption of the natural surroundings. (We point out again that PV does not require water for its operation.)

Figure 1.10 The Jerez de los Caballeros 10 MW PV plant. Reproduced by permission of Guascor Solar SAW

But following the actual trend, the construction of 1000 plants per year of average size of 75 MW will lead to the 75 GW that, added to the 25 GW in buildings, will complete our goal of 100 GW per year. This will be a challenge, but not impossible if the business is profitable. Remember that in the last 2–3 years 1000 plants of average size 4.5 MW have been built (totaling 4.5 GW) and the biggest are approaching the 75 MW size.

1.4.2 Raw Materials Availability

Are there sufficient raw materials on Earth to make enough PV modules to provide a significant fraction of our energy in the future? This is an important question, because if the answer is "no", then PV will be ultimately relegated to a minor role.

The main material used today to make solar cells is silicon. Being the second most abundant material in the Earth's crust (after the oxygen), there is not a foreseeable shortage. High-purity quartzite (SiO_2) ore is used to produce Si today. The present production of Si is about 50 times more than needed for PV use and can be easily increased, so that producing 12% of the electricity with PV, which implies an increase of the cell production by at most, 15 times, can be easily supplied.

But for non-Si based cells, this is a very complex question, requiring decisions about how much of a given element is economically recoverable and at what rate (ton/yr or equivalently converted to GW/yr of PV). Several studies [38–40] are in general agreement that the potentially limiting materials for PV are Ag for contacts to Si cells (but Ag is not essential), In for Cu(InGa)Se_2, Te for CdTe, and Ge for Ge substrates commonly used for III–V concentrators cells or a-SiGe cells. In and Te metals are not actually mined, but are dilute by-products of other metal refining; i.e. Te from Cu ore; In and Ge from Zn ore and also from ashes in the combustion of coal. Conclusions differ significantly as do the assumptions, but generally these studies find that solar cells based on In or Te could provide a few percent of the world's future electricity without significant reductions in utilization (i.e. thickness), thus potentially meeting the entire, although trivial, requirement for PV assigned by the IPCC. Another view suggests that historical production rates of In and Te could easily provide 20 GW/yr [39], thus meeting about 20% of our hypothetical PV requirement of 100 GW/yr with CdTe and Cu(InGa)Se_2. This does not mean they should not be considered, since it is unlikely a single PV technology will dominate. Supplies of In and Te can enable multi-billion dollar annual PV industries[4]. Reducing the cell thickness, recycling of waste and expired modules, and increasing efficiency in manufacturing will reduce the total demand for a given raw material and extend its ability to meet these targets. These are all promising and active areas of research.

As for III–V/Ge concentrator devices, operating at 500–1000×, there are not major problems in producing 100 GW/year [41]. The main limitation comes from the Ge, and from plastic for lenses. The use of hybrid glass–silicone lenses, beginning to be common today, would remove the limitation on plastics (because they use a very thin layer of silicone). As for the Ge, it is probable that enough can be found in other sources, but it could be recycled or substituted by Si.

1.4.3 Is Photovoltaics a Clean Green Technology?

Would large-scale PV manufacturing and deployment degrade the environment, although in a different way from conventional energy production?

[4] There are also competing technological uses to consider. For example, Indium usage has increased substantially in recent years due to the need for transparent conductive indium tin oxide (ITO) layers in flat panel displays.

One of the most valuable characteristics of photovoltaics is its well-deserved image as an environmentally clean and "green" technology, resulting from the cleaner operation of a PV electricity generator compared with a fossil-fuel or nuclear-fired generator. But this must also extend to the manufacturing process itself as well as the recycling of discarded modules. Let it be stated at the beginning that the present Si-based PV technology which dominates the market has few environmental concerns and is considered totally safe to the public.

Hazards can be classified by whether they affect workers at a PV manufacturing plant, customers with photovoltaics on or near their homes, or members of the public who consume air and water near the PV plant. Very little risk is associated with the public or the PV owner or installer; the main risk being that of electric shock already existing with conventional electricity, but more severe because large areas can be electrically live. Adequate grounding is strongly recommended.

Safe handling procedures for some of the materials and processes are already well established from the integrated circuit or glass coating industries. But in the case of some unique materials and processes, safety procedures had to be developed by the PV industry. The PV Environmental Health Safety Assistance Center at Brookhaven National Laboratory in New York, USA provides worldwide leadership in risk analysis and safety recommendations for the PV industry [42]. An industry-focused PV Safety Working Group has been established in Europe [43].

Si modules, currently the most widely used, are totally free from the suspicion of releasing dangerous materials. Si module manufacturers have long used Pb-based solder to interconnect the wafers, as did the electronic chip manufacturers. This represents a minor risk that nevertheless is being further reduced through the use of Pb-free solders.

There has been considerable research into occupational and accidental exposure and risk analysis of one PV material in particular – thin films of CdTe – since Cd is a known carcinogen. The general conclusion is that CdTe thin film modules do not pose a risk to the public [44]. Much of the concern is misguided for two reasons: there are significant toxicological differences between elemental Cd (toxic) and compound CdTe (much less so), and because the CdTe is hermetically sealed between two sheets of glass. Those concerned about Cd sealed in CdTe modules should consider that most Cd in our environment is released directly into the atmosphere by combustion of coal and oil. Chapter 14 has more about Cd environmental issues. Even in the event of a house fire, studies have shown that roof-mounted PV modules do not release any potentially hazardous materials [45], and this is also true for those containing cadmium.

A related issue is what to do with PV modules at the end of their projected 25- to 30-year life. An excellent strategy is to recycle the modules. This solves two problems at once, namely, keeping potentially hazardous materials out of the environment and reducing the need for additional mining and/or refining of new materials. Semiconductor vendors have indicated a willingness to accept used modules, and to extract and purify the CdTe, CdS, or $Cu(InGa)Se_2$ for resale and reuse. As for the recycled Si, it is inherently purer than the Si used today as raw material.

Thus, we can say with confidence that PV is the cleanest and safest technology with which to generate electricity even at the GW production scale.

1.4.4 Energy Payback

Can PV modules produce much more energy over their lifetime than it took to make them; i.e. are they net energy producers? This oft-expressed concern is baseless. This concept is quantified by the "energy payback time " or EPT, which is how many years the PV system must operate to produce the energy required for its manufacture. After the payback time, all of the energy produced is truly new energy.

Many studies have concluded that EPT's have steadily decreased over the past decades, now estimated at 1.5 to 2.5 years for crystalline Si and 1 to 1.5 years for thin films [45, 46]. Thus, all the energy needed to produce our fictional 100 GW every year will be paid back to the grid within the next two years.

For crystalline Si, melting and forming the crystalline wafers is the major energy requirement. For thin films, where the semiconductor layers are 100 times thinner, and deposited at much lower temperature, their energy requirement is negligible. Instead, it is the energy embodied in the glass or stainless steel substrate, which is the major energy sink. The cosmetic Al frame around the module is responsible for a surprisingly large fraction of energy, and is being phased out. Although thin film modules have a shorter energy payback, they also have lower efficiency, which means a larger BOS is needed to support the larger number of modules. Thus, a larger amount of energy is embodied in the BOS for thin film photovoltaics compared to crystalline Si photovoltaics.

The case of concentrators is less studied, but again the use of semiconductor is reduced and the BOS becomes more important than even for the thin films because the concentrating structures are more massive. However, their efficiency is much higher. In summary, we can guess that their EPT will be similar to the case of thin films.

Recently, PV has been examined in terms of its potential for Carbon reduction [46–48]. The amount of CO_2 released during manufacture of PV systems is much less than the CO_2 avoided by the power it produces during its lifetime. PV technologies are responsible for about 30–50 g CO_2 emission per kW h produced over their life-cycle (all due to fossil fuel energy consumed during manufacturing) while coal-powered generating plants release about 20–30 times more. As the CO_2 load from our energy sources decreases, the amount of CO_2 per kW h of PV will decrease as well. Thus, PV is an excellent strategy to mitigate global climate change.

1.4.5 Reliability

Most modules seem to lose about 0.5–1% of their relative output annually for a variety of reasons. For the past few years, most modules have been sold with guarantees to maintain at least 90% of their rated output after 10 years and at least 80% after 20 or 25 years. They must pass rigorous reliability testing using accelerated extreme weather conditions. A recent study of >200 modules [49] produced in the early 1980s, before many of today's more rigorous standards and improved encapsulation methods and materials were available, found that after >20 years continuous operation outdoors, only 18% of them had lost more than 20% of their rated power (and they were only warranted for 10 years back then!). And remember, even if they lose 20% of their output after 20 years, they are still providing free electricity! Also, solar modules have been operating in space, a very harsh environment, for decades.

1.4.6 Dispatchability: Providing Energy on Demand

Could PV meet all of the world's needs today if we would just pass laws requiring photovoltaics and halting all fossil and nuclear plants?

The first technical problem faced would be the intermittent nature of the solar radiation, available only during the day and strongly reduced in overcast skies. Energy storage would solve this problem, but no cheap storage method appears on the horizon, although pumped hydro storage [50] is already used and compressed air [51], grid-tied electric vehicle batteries (when parked and plugged in) called V2G [52], and new battery technology (Chapter 20) are being actively explored.

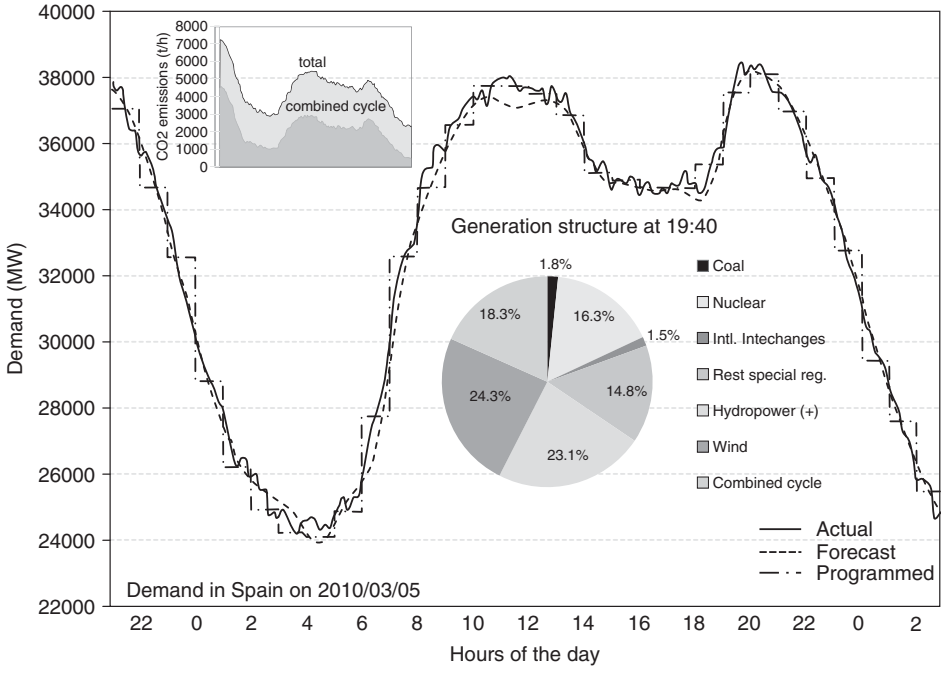


Figure 1.11 Dispatching of electricity in Spain over 24 hours on 2010-03-05; actual, forecast and programmed (dispatching plan). The wind energy resulted in 24.3% at the peak consumption time. PV is not independently registered. It is included in "Rest special regime" of 14.8% (the estimate for PV is about 2%). The high penetration of renewable is permitted by the high proportion of hydro/gas (combined cycle) plants which are easily dispatchable. Coal generation, very important in the past, has almost disappeared to allow for renewable introduction. Coal is the most CO_2-intensive source. Adapted with permission of Red Eléctrica de España

The problem of electricity storage is not specific to the intermittent production of electricity. It is a general need of the grid management. Actually, electricity demand is quite variable with the hour of the day and even seasonally, where differences of 40% are possible daily between its peak (maximum) load and its base (minimum) load. Thus, the electricity operators, who have a statistical knowledge of the demand behavior, plan the connection or disconnection of power plants to produce enough energy that approximately matches the demand as shown in Figure 1.11. The output of power plants is adjustable within certain limits. This process is called dispatching.

The production of intermittent electricity is also predicted statistically and the output needed from fuel-based power plants can be planned accordingly. In Spain, about 14% of the yearly demand is intermittent or variable and this fraction may be much higher in specific windy or sunny moments. Baseload coal or nuclear power generators must be permanently connected and generally run at full power. As the wind/solar fraction grows (assuming no storage), on windy/sunny days when the intermittent generation alone can meet the demand, some of the wind/solar generation will have to be shut down, because conventional baseload plants cannot be disconnected or decreased quickly. A cheap storage, such as pumped hydro plants, would permit more penetration of intermittent renewable electricity. Either way, in any grid (even without intermittent renewable energy) there is always some idle generation capacity and the associated financial losses, but this can be increased by excessive intermittent generation. Adequate grid management would allow up to 35% of the electric production to be intermittent [33], even without sensible storage.

1.5 TRENDS IN TECHNOLOGY

Most of this book's chapters deal with technology. This section will give a broader perspective that the specific chapter authors cannot provide.

In 2008 almost 8 GW_p of modules were manufactured. The breakdown among the different technologies appears in Table 1.5 for 2003 and 2008. The crystalline Si (c-Si) modules dominated the market (87% in 2008) and are divided into multicrystalline (multi-Si), single- or monocrystalline (mono-Si), or ribbon silicon, depending on the type of Si wafer used. The thin film modules, a minority, but whose market share is expanding, are divided into a-Si, CdTe and CIS modules. The rest of the technological options are yet too immature to appear in the market breakdown.

As shown in Figure 1.8a, efficiencies are one of the most critical factors to reduce cost of electricity generated. "Champion" cell efficiencies are presented in Figure 1.12 for several technologies. Two technologies not yet commercially significantly appear in this figure (ironically having the highest and lowest efficiencies). Improvements in champion cell efficiency have slowed in the last decade, except for the III–V based multijunctions and CIGS. The efficiency of champion cells is typically 25–50% higher than the efficiency of the commercial products because the techniques used for making the highest efficiencies are seldom acceptable for cost effective manufacturing. This seems to be in contradiction with the statement that efficiency is the main driver of cost reduction, so we will qualify it as follows: the industry goal is to obtain the highest module efficiency compatible with a reasonably cheap, high-throughput, high-yield and reproducible process. Figure 6.18 in Chapter 6 shows a historic view of the evolution of the breakdown among technologies.

1.5.1 Crystalline Silicon Progress and Challenges

Table 1.5 show that c-Si, as either single or multicrystalline wafers or ribbons, was responsible for almost 90% of worldwide PV production. How did its dominance occur? First, Si cell technology has benefitted from the tremendous development of microelectronics that is also based on the same Si. While thin film cell researchers had to develop their own manufacturing equipment, Si

Table 1.5 Total MW of production and percentage of the three Si wafer and three thin film technologies in 2003 (when the first edition was published) and 2008

Technology	2003		2008	
	MW	%	MW	%
Multi-Si	429	57	3773	48
Mono-Si	242	32	3024	38
Ribbon-Si	33	4	118	1
a-Si	*34*	*4.5*	*403*	*5*
CdTe	*8*	*1*	*506*	*7*
Cu(InGa)Se$_2$	*4*	*0.5*	*79*	*1*
Total Crystalline Si	704	93	6915	87
Total Thin Films	*46*	*7*	*988*	*13*
total	750	100	7910	100

Data from *Photon International*, March 2009, p. 190

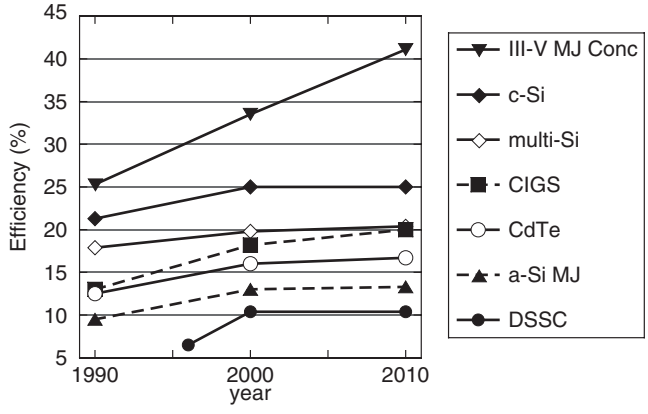


Figure 1.12 Best small-area (0.5–5 cm^2) efficiency for various cell technologies measured under standard laboratory test conditions as of 1990, 2000, 2010. MJ concentrators are double junctions before 1995, and triple junctions after. MJ a-Si represents stabilized efficiency after extended light soaking (see Chapter 12). Data for 2000 and 2010 from independently verified Solar Cell Efficiency Tables (*Progress in Photovoltaics*, John Wiley & Sons, Ltd, UK)

cell researchers could use that already developed for microelectronics, sometimes off-the-shelf and sometimes with some minor modifications. Second, the silicon bandgap, of 1.1 eV, is almost optimal to make a good solar converter (see Figure 4.3 in Chapter 4). Furthermore it is very abundant, clean and nontoxic. Finally, Si solar cells are very stable, even without encapsulation.

However, Si has mechanical limitations (it is brittle) and optical limitations (it absorbs sunlight weakly), requiring relatively thick cells. Therefore, some of the electrons pumped by the photons to the conduction band have to travel distances of the order of the thickness, to be extracted by the front face through the selective contact (the one-way valve of Figure 1.2 representing the *pn* junction). Consequently, a good material with high chemical purity and structural perfection is required to fight the natural tendency of the conduction-band electrons to return to the valence band. To avoid this loss process, called *recombination*, the electrons must be highly mobile, as they are in perfect silicon. Impurities and imperfections must be avoided as they can absorb the extra energy of the conduction-band electrons, thus eliminating the free electron.

Metallurgical grade (MG) Silicon is obtained by reduction of quarzite (SiO_2) with charcoal in an arc furnace. PV only uses about 2% of the world production of MG-Si. Then it is highly purified, commonly by a method developed by and named after the Siemens Company, consisting of the fractional distillation of chlorosilanes, which are obtained from the reaction of a chlorinated source with Si. Finally, chlorosilanes are reduced with hydrogen at high temperatures to produce hyperpure silicon, usually called s*emiconductor grade* (SG) s*ilicon* or just *polysilicon* which has many random grains of crystalline Si, typically of about 1 mm. Methods to produce polysilicon or a lower-cost form called solar grade Si are described in Chapters 5 and 6.

In the past, the polysilicon was produced by about half a dozen factories for the microelectronic manufacturers. Today PV is the biggest user of this polysilicon and more new factories are appearing now specializing in solar grade polysilicon. The definition of this grade is the subject of controversy as some manufacturers obtain it as enhanced MG-Si with reduced impurities, leading to a cheaper material, but unable to make high-efficiency cells; others want to keep the high purity, even with higher cost, necessary for the higher efficiency. More about this debate can be found in Chapters 5 and 7.

By melting and recrystallizing polysilicon, the wafer producers grow either single mono-Si crystal ingots by the Czochralski (Cz) technique or cast multi-Si blocks (grain size about 1 cm). Both methods yield large solid blocks which must be sliced into wafers (150–250 μm thick). Conventional Si cells are made by diffusing the junction and screen printing the contacts. Mono-Si wafers produce cell efficiencies of about 16–17% while the multi-Si wafers produce cells of about 13–15%. Mono-Si wafers are necessary to implement the unconventional and higher-efficiency structures such as the HIT and IBC cells, both with champion cell efficiencies of over 20% (see Chapter 7 for details).

Wafering the silicon blocks implies kerf losses and a sizeable proportion (40%) of the expensive polysilicon is lost in "sawdust". To avoid this, sheets of silicon can be grown as "ribbons". However the cell efficiency is not quite as high as multi-Si. The same few factories making Si ribbon also integrate it into cells.

Most crystalline solar cells are fabricated from wafers by the screen printing technology that is described in Chapter 7. There are some exceptions, such as the IBC and HIT cells also explained in that chapter. The bigger cell factories integrate the crystal growth and wafering processes.

The mono-Si technology comes directly from the microelectronics industry and was the first to be used for solar cells. The multi-Si technology was developed specifically for PV to avoid the high costs of the Cz growth process. But it has not been able to clearly dominate the market because of the slightly lower efficiency. There is considerable research to develop processing steps to reduce the efficiency gap between mono-Si and multi-Si. Concerning ribbon, besides the low efficiency, the growth per cm^2 is slower than for wafers, leading to higher capital costs. The production of fast ribbon growth without losing efficiency is the desired goal, but it is difficult, in part because the ingot formation is also a purifying step that is incompatible, for physical reasons, with fast ribbon growth (Chapter 6).

Once the cells are manufactured they are assembled into encapsulated modules, as described in Chapter 7. This is done, either in the cell factories or in module assembly factories that purchase cells from a variety cell factories. Thus nearly identical Si cells can be bought from a variety of suppliers and integrated into modules. This is an important origin for double counting the module production in market studies (because the cell manufacturers may count all their cells as MW_p produced and then the module manufacturer counts them again). This may cause discrepancies in the reports on manufacturing as well as in the breakdown between technologies (that are not very substantial).

A consortium of European Si PV companies and research groups (CrystalClear) has been collaborating on reducing the cost per watt of various Si PV cell technologies [53]. They established the cost structure for the current baseline standard multicrystalline Si module, called Basepower, averaging 2.1 €/W (2005 reference technology). Figure 1.13a presents the costs in terms of process step and Figure 1.13b is the cost in terms of manufacturing cost category.

The fraction of polysilicon cost (i.e. feedstock) is only ~14%, but during 2008 (in the explosion of the Spanish market) a shortage of polysilicon caused its price to rise occasionally in the spot market to about 10 times its normal price (of about US$50/kg). The combined shortages and high prices of Si modules gave an opportunity to less-mature technologies, such as TFSC and CPV to find a place in the market. Not everyone was prepared to profit from this, but we think that the extraordinary success of First Solar (the world's biggest company in 2009) with its thin film CdTe modules is partially due to this unique opportunity.

Increasing efficiency, followed at a distance by reducing the polysilicon cost, are the main single drivers of cost reduction in Si technology. In this respect the appearance of such technologies as the IBC of SunPower and HIT of Sanyo, both able to produce wafer-sized cells of more than 20%, are very promising. It is not really known if they are more cost effective than conventional screen printed cell technology (still largely dominating). But in 2008 they became the ninth and

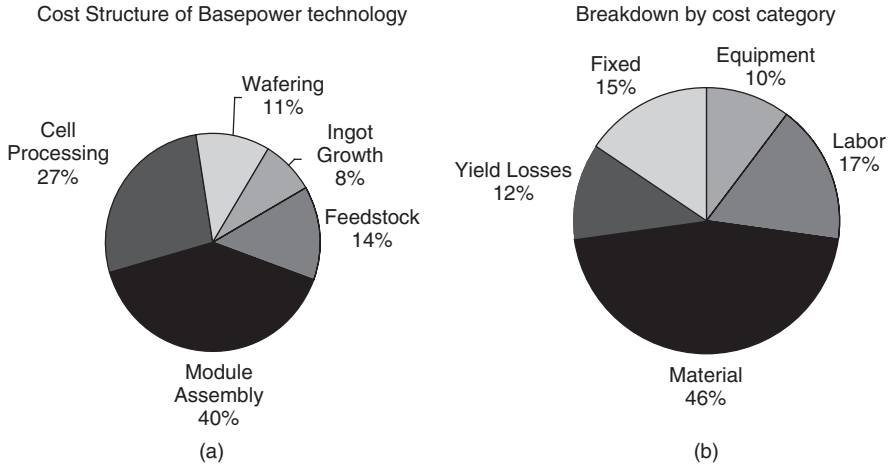


Figure 1.13 Breakdown of costs in the fabrication of a Si-wafer-based PV module. (a) The left-hand side is the percentage in terms of technical process steps; (b) the right-hand side is the percentage in terms of financial activities (adapted from del Cañizo *et al. Progress in Photovoltaics* 17, 199–209 (2009)). About half of the module assembly costs are materials

tenth largest solar cell manufacturing companies in the world, SunPower with 3% of world market and Sanyo with the 2.7%.

1.5.2 Thin Film Progress and Challenges

Why develop a totally different semiconductor technology for photovoltaics when Si is so well established? The simplest answer is in order to achieve lower cost and improved manufacturability at larger scales than could be envisioned for Si wafer-based modules.

The TFSC are based on materials that strongly absorb sunlight so that the cells can be very thin (1–3 micrometers). The electrons freed by the photons need to travel only this short distance inside the cell to the cell contacts (and from there to the external circuit to produce power). This reduces the demand for high purification and crystallinity of the material, one of the causes of the high cost of the Si cells. However, where the thin films have a real business advantage is that they are made directly into modules and not in cells. In other words, while Si cells are manufactured from wafers, then processed and assembled to form a module, in TFSC technology many cells are made and simultaneously formed as a module.

But there are disadvantages and in fact they have not yet dominated the market. We must understand why.

It was recognized almost as early as c-Si PV cells were developed in the 1950s that thin film semiconductors could make good solar cells. When fabricated into useful devices, they are so thin that they must be deposited on a foreign material, called a substrate, for mechanical support. This can be a glass and metal or a sheet of plastic, all of them of low cost (at least as compared with the self-supporting Si wafer). A framework for analyzing the material properties, device structures, device physics, and manufacturing issues unique to TFSC had to be developed since they differed considerably from Si wafers [54, 55]. Between 1981 and 82, four thin film technologies demonstrated the ability to cross the magical 10% efficiency barrier, thus becoming candidates for serious consideration: Cu_2S/CdS [56], a-Si [57], $CuInSe_2$/CdS [58], and CdTe/CdS [59]. Of

these four TFSC technologies, Cu_2S/CdS would soon be rejected for commercialization due to fundamental and fatal stability problems related to electrochemical decomposition [60]. In contrast, a-Si has a minor stability problem that, once stabilized, is predictable, reversible and seasonal, as discussed in Chapter 12. No fundamental stability problem has been found with Cu(InGa)Se_2 and CdTe modules, although they can develop unique degradation modes if not properly encapsulated. Consequently, significant industrial and government-sponsored research and resources have been directed worldwide at TFSC technology. This led to steady progress in champion cell efficiencies through the 1990s, as seen in Figure 1.12.

But the efficiency of TF modules is 25–50% lower than for Si modules which makes it difficult to translate the low cost per m^2 of TFSC modules to cost per W_p. This lower efficiency causes a higher area-related BOS cost when a TF array is installed, partially negating the natural cost advantage of TFSC, as discussed in connection with Figure 1.8b.

To obtain low manufacturing costs, TFSC plants must be operated at high volume throughput to offset the initial capital investment[5]. A detailed study of thin film module manufacturing options concluded that costs of current technologies would decrease 30–50% as the production facility increased from 25 to 200 MW per year [61].

The TFSC manufacturing process is designed such that they are deposited sequentially on moving substrates as in a continuous 'in-line' process or on many substrates at a time in a stationary batch process. This minimizes handling and facilitates automation, including laser scribing, to isolate and interconnect individual cells on the module, called monolithic integration. They are deposited at relatively low temperature (200–500 °C compared with ~800–1450 °C for the different main processes of c-Si). TFSC are either polycrystalline with small ~1 μm sized grains such as Cu(InGa)Se_2 or CdTe, amorphous like a-Si, or mixed amorphous/crystalline Si phases called nanocrystalline Si. The noncrystalline structure is a consequence of being deposited at temperatures too low and at rates too fast to allow perfect crystalline bond formation. TFSC typically consist of 5–10 different layers whose functions include reducing resistance, forming the *pn* junction, reducing reflection losses, and providing a robust layer for contacting and interconnection between cells. Some of the layers are only ~20 atoms thick (10 nm), yet they may be a meter wide! This requires excellent process control.

Besides the lower efficiencies (so far), TFSC have a much less-developed knowledge and technology base compared with c-Si, and their properties are more difficult to control. Consequently, under-capitalized companies have had to struggle to develop not only an understanding of the materials and devices, but also the equipment and processing to manufacture them. The thin film PV industry has had to develop the technologies all by itself with considerably less financial resources than the Si PV industry had. They were not able to adopt a mature technology from the Si electronics industry. These factors can lead a purchaser to hesitate to buy a product that is less mature, and not so cheap (because of the small manufacturing volume), and thus to prefer the standard Si wafer-based product.

What are the strengths and remaining challenges for the TFSC industry? We will review the salient characteristics of the three leaders: a-Si, Cu(InGa)Se_2/CdS, and CdTe/CdS.

Amorphous Si (Chapter 12) is deposited from hydride gases such as SiH_4 using plasma to decompose the gas. This is called *plasma-enhanced CVD* (PECVD) and allows for large areas to be coated rather uniformly and with excellent control, using the same technology as large-area flat panel displays. The a-Si film has 1–10% hydrogen bonded to the Si, and is often designated

[5] This is being proven by First Solar which has been manufacturing CdTe thin film modules for 10 years, achieving the lowest manufactured price per watt since 2008 (<1.00 $US/W), becoming the world leader in PV module production in 2009.

as *a-Si:H*. The H atoms passivate a large number of the defects resulting from the incomplete bonding of the Si atoms. The atomic structure lacks the long-range order of other crystalline or polycrystalline materials. This can be an advantage because the light absorption is increased with respect to c-Si. Films are typically deposited between 150 and 250 °C, the lowest temperature of any of the TFSC materials, allowing the use of lower-cost, low-temperature substrates. a-Si solar cells are deposited on glass, stainless steel foil, or plastic. The last two substrates are flexible allowing for "roll-to-roll" manufacturing where all the layers are deposited as the roll moves through their process zone. Nearly all a-Si modules contain multiple junction devices where two or three junctions are grown on top of each other. This allows for more efficient utilization of the sunlight. Increasingly, a nanocrystalline form of thin Si is being used as the low bandgap partner in making "micromorph" a-Si/nc-Si multijunction cells (Chapter 12, Section 12.5). The highest reported cell efficiency was 15% for a triple junction, which degraded to about 13% before stabilizing [62]. Micromorph modules 1.4 m^2 or larger are being reported with 8–10% stabilized efficiency, as discussed in Section 12.6, but standard products (not micromorph) are in the 5–7% range. The three major challenges for a-Si technology are: (1) to improve the standard module efficiency to 10–12%; (2) to minimize or eliminate the self-limited degradation which reduces efficiency by 2–3% (absolute); and (3) to increase the deposition rate of the layers and utilization of the gases, especially the nanocrystalline layer to allow faster, lower-cost manufacturing.

Polycrystalline layers of $Cu(InGa)Se_2$ (Chapter 13) alloys have produced the highest efficiency TFSC devices and modules. TFSCs based on $CuInSe_2$ (no Ga) achieved 12–15% efficiency, but were limited by the low bandgap. Alloying with Ga and/or S increases the bandgap and increases the efficiency of delivering the electrons to the circuit. While many deposition methods have been explored in the laboratory, there are two different processes under commercial development. Co-evaporation forms the alloy by simultaneous evaporation of the Cu, In, Ga, and Se from sources onto a heated substrate. The other process is called *selenization*, because layers of Cu, In, and Ga are deposited by a wide variety of methods onto a substrate, then heated in the presence of Se from a gas such as H_2Se or a Se vapor, thus contributing the fourth constituent of the alloy. A very active area of research is developing methods to incorporate these atoms and others into low-defect alloys to increase the bandgap even further.

Substrate temperatures typically reach 500–600 °C during some stage of the growth unless the substrate is a polymer, in which case 450 °C is the maximum. Substrates of Mo-coated glass are typically used although Mo-coated metal foils or plastic are in manufacturing. If sodium is not available from the substrate (diffusing from glass), it must be provided directly, either during or after deposition, to enhance electronic quality of the $Cu(InGa)Se_2$ and increase voltage. The $Cu(InGa)Se_2$ films are *p*-type, typically 1–3 μm thick and have crystallites or grains on the order of 1μm. The *pn* junction is formed by depositing an *n*-type layer of CdS, ZnO, or other new materials under development to replace the CdS (largely for "environmentally friendly" bragging rights). The highest reported cell efficiency is presently 20.0% [63] and several companies have limited manufacturing capacity (<20 MW) of modules with 10–13% efficiency. Transferring a high-efficiency small scale laboratory process on a stationary substrate to manufacturing on a large area moving substrate has proven more difficult for $Cu(InGa)Se_2$ than for a-Si or CdTe. The three major challenges for $Cu(InGa)Se_2$-related technology are: (1) to control the composition (Ga, S, Se, or Na) of the alloy through the film in a manufacturing environment on a moving substrate; (2) to find alternative junction partners to replace CdS; and (3) to find new alloys (with Ag, S, Te) or new deposition methods to give high-performance devices with higher-bandgap alloys.

Polycrystalline layers of CdTe (Chapter 14) have been investigated for photovoltaics since the 1970s. In contrast to limited process options for a-Si or $Cu(InGa)Se_2$, there are over 10 methods to deposit the CdTe films that have produced CdTe solar cells exceeding 10% efficiency. Four have reached precommercialization: spray pyrolysis (SP), electrodeposition (ED), vapor deposition (VD) and close-spaced sublimation (CSS). Some take place in liquid baths that are barely warm

~50 °C, with CdTe deposition rates of μm/h (ED) while others take place in vacuum systems at temperatures high enough to soften glass ~600 °C, with CdTe deposition rates of μm/min (CSS). There seem to be three critical steps, however, that all efficient CdTe solar cells require. First, they need a post-deposition anneal in the presence of Cl and O_2 at around 400 °C. This chemical/thermal treatment enlarges the grains, passivates the grain boundaries, and improves the electronic quality of the CdTe. Second, all CdTe layers need a surface treatment before applying a contact. This treatment can be a wet or dry process and prepares the CdTe surface by etching away unwanted oxides and leaving a Te-rich layer needed to make a low-resistance contact. Third, nearly all high- efficiency devices have a Cu-containing material somewhere in their CdTe contact process but again, there are many ways this can be achieved. Details of these three process steps tend to be very proprietary. Whichever process is used to deposit the CdTe, it has been found that the entire device process is highly coupled since processing steps strongly influence previous layers. This is partially due to the CdTe grain boundaries which act like paths for interdiffusion.

The *pn* junction is formed by first depositing an *n*-type layer of CdS on a glass substrate with a transparent conductive oxide contact layer (typically SnO_2) followed by the 2- to 8-μm-thick CdTe layer and appropriate chemical annealing. Once the solar cell is made, the CdTe films are slightly *p*-type with crystallites or grains of the order of 1 μm. The highest reported efficiency for a CdTe/CdS device is presently 16.5% [64] and modules are around 10–11%. Some CdTe modules have been in outdoor field-testing for over 10 years with negligible degradation. Of the three leading TFSC technologies, CdTe has surged into first place in terms of manufacturing capacity and lowest cost on the strength of a single company. The three key challenges are: (1) to better understand the various post-deposition optimizing treatments so they can be simplified and transferred into production; (2) to increase the output voltage commensurate with its bandgap; and (3) to maintain and evolve the safe and cost-effective Cd usage in the workplace, followed by recycling at the end of the module's life.

Technically astute investors know that other factors can be more important than efficiency in selecting a technology for development. This point is made obvious by examining the relative performance of the three major TFSC technologies – Cu(InGa)Se_2, CdTe, and a-Si – in Figure 1.12. Note that a-Si has always had the lowest efficiency. Yet, of the three, it was a-Si that was commercialized much earlier and more widely. This is partly because there has only been one generic deposition technology – PECVD – while CdTe and Cu(InGa)Se2 have a wide range of technologies, meaning each company must develop the unique process technology and equipment themselves. a-Si also had a stronger scientific research base, partly due to the other applications such as flat panel displays, which ensured that the relation between deposition conditions and fundamental material and device properties were well characterized, which comforted investors. In contrast, CdTe and Cu(InGa)Se_2 are "orphans" because they have no real application outside of photovoltaics. In 2008, Table 1.5 shows that a-Si accounted for about 4%, CdTe for 7%, and Cu(InGa)Se_2 still about 1%. Yet Cu(InGa)Se_2 has had the highest laboratory efficiency for two decades (Figure 1.12). This shows that translating research-grade champion cell performance into production modules coming off the production line day after day is a very challenging task. The phenomenal growth of CdTe is due to one company, First Solar, whose modules are among the lowest priced on the market. But it took them over 15 years of research and development with significant public and private investment to get there.

Conjecturing that the ideal PV technology would have some of the merits of c-Si (abundance, nontoxicity, stability) but be deposited as a thin film a few micrometers thick, several groups have tried to achieve the "best of both worlds" by developing thin films of multi-Si deposited on an inexpensive non-Si substrate. This is the subject of Chapter 11. At present, the best thin film multi-Si modules have the same efficiency ~10% as their $CuInGaSe_2$, CdTe, or a-Si based predecessors. This is partly because multi-Si thin-film photovoltaics also inherits some of the problems of both c-Si and thin films. In particular, passivation of grain boundaries and surfaces seems to be a major

problem, yet many of the well-established passivation methods from c-Si are not applicable to multi-Si thin films due to temperature limitations ($<600\,^{\circ}C$).

There are new thin-film technologies such as the solid–liquid junction dye-sensitized (Chapter 15) and polymer-based organic (Chapter 16) solar cells that operate on very different principles than an all-solid-state solar cell. Their main attraction is the potential for very low cost. However, these fascinating new technologies present many new challenges, including strong sensitivity to air and water vapor, hence the need for excellent encapsulation. There are some private investment efforts to scale them to manufacturing.

1.5.3 Concentrator Photovoltaics Progress and Challenges

Concentrator Photovoltaic Technology or CPV is based on separating the area for collection of sunlight from its conversion. The collection area is an optical element, mirror or lens, which casts the light into a much smaller area of solar cells. This allows using high-efficiency but more expensive solar cells since the area of cells is >100 times smaller than the light collection area. The design and operation of CPV is described in Chapter 10.

PV technologists were aware of this possibility from the beginning of PV development. In general CPV needs sun-tracking systems that makes it unsuitable for small-scale applications of PV. There were attempts in the 1980s when Si solar cells were still too expensive and the markets were still too small.

The interest in CPV has spread in the last five years when MJ III–V-based solar cells, discussed in Chapter 8, developed for space applications, started to approach 40% efficiency, which they have already surpassed, as shown in Figure 1.12. CPV can take advantage of these ultra-high-efficiency cells, which are inherently very expensive, because the total cell area is reduced under high concentration, thus mitigating their high cost. Concentration levels of $500\times$ (the cell is 500 times smaller than the optical aperture) are common with this technology and there is a trend to move towards the $1000\times$ which presently operates at slightly less efficiency. Projected costs of such concentrator systems might be very small [65]. Concentration helps to increase the efficiency which theoretically increases as the logarithm of the intensity, until it reaches too high a value of current density, leading to ohmic losses that reduce the efficiency. Thus concentrator cells must be specially designed to have very low series resistance.

However, there are disadvantages. First, CPV does not utilize diffuse radiation, thus losing ability to convert at least 10% of the global radiation even in the best climates, and often much more. Second, at $500–1000\times$, CPV requires tracking the position of the sun very accurately every minute of the day which adds cost and complexity to the installation. Non-imaging optics is a new scientific tool that may soften this requirement. Third, the optical elements reduce the overall efficiency. Nevertheless module efficiencies of 30% have been presented at conferences and are about to appear commercially. Finally, CPV modules must be able to dissipate a significant amount of heat, leading to complex construction and reliability issues.

Another temporary issue is the rating. An array is usually formed by modules mounted on a tracking system, but only when mounted can the efficiency of the array be determined because the tracking itself affects the efficiency. Even in a good tracking system it varies in about 5% from second to second because of small misalignments, so that defining the kW rating of an array, and how to measure it, is not yet decided. Under these conditions the bankability of a concentrator system is still difficult to assess or guarantee. Nevertheless, under the silicon module shortage in 2008 the Spanish company Guascor Solar, with license of the American Amonix, installed over 9 MW of concentrators, being so far the first in installed CPV.

With the short learning curve in CPV, competing with the cheap Si modules from China, and to some extent with the cheap CdTe modules from First Solar, is very difficult. It will be necessary to demonstrate good credibility in performance at low installed prices to see this technology widely deployed. The recently created Institute of CPV systems in Spain (ISFOC), has subsidized the installation of 3 MW from seven companies (three from Spain, two from the USA, one from Germany and one from Taiwan) to provide accredited performance data to help gain this credibility and to establish the rules for measuring CPV performance.

Some think that they will be able to make a low-concentration system with Si cells that will beat in prices the complex high-CPV systems and the flat panel systems alike. As matter of fact the Guascor Photon sales have been obtained using highly efficient (~25%) IBC-silicon cells, although they are moving now to MJ III–V cells.

So the result is that many companies, start-ups and established, are today involved in developing CPV options in the high and low concentration ranges. The next years will tell us about the success of these efforts.

1.5.4 Third-Generation Concepts

In 1961 Shockley and Queisser (SQ) published a paper [66] setting the thermodynamic efficiency limit[6] of a single-junction solar cell of about 40% under certain hypotheses that were thought to be fundamental and absolute. Cells based on principles that violate some of these hypotheses are called third- [67] or next-generation [68] solar cells.

The most studied (called by some revolutionary) third-generation solar cells [69] are: the Intermediate band solar cell [70] which discards the SQ hypothesis that photons below the bandgap are not absorbed, the multi exciton generation solar cell [71], which discards the SQ hypothesis that a photon can only pump one electron; and the hot carrier solar cell [72] which discards the QS hypothesis that the electrons are at the lattice temperature.

Some progress has been made in all three concepts since the first edition of this book, but a high-efficiency cell has not been produced with any of them. The 1st and 2nd generation solar cells used today have required decades to yield reasonably high efficiencies and a reliable manufacturing process. It will be the same with these new concepts. They will probably not be available to meet the 2030 challenge. Third-generation concepts may enable TFSC to operate with efficiencies above 20% or may permit achieving efficiencies of 50% in CPV cells permitting modules of 40%. To accomplish this, we will need detailed knowledge of these new fundamental concepts and more importantly how to integrate them into a functional device.

1.6 CONCLUSIONS

The human race is increasingly aware of the need for sustainable development. Solar energy is almost the only, and certainly the most developed way of producing energy for this sustainable development. This will be possible mainly through PV.

PV constitutes a new form of producing electric energy that is environmentally clean and very modular. It is highly appreciated by the public. It is unique for many applications of high social value such as providing electricity to people who lack it in remote areas. In recent years PV has experienced an unprecedented burst of growth. Today PV is a big business of around $US 50 billion

[6] This limit applies to the individual cells in a MJ stack, but not to the stack as a whole.

worldwide and growing at ~50% annually. PV-powered homes, commercial buildings, and power plants have been built around the world. It has been recently shown (in Spain) that PV electricity can be installed five times faster than nuclear power plants and that a penetration of intermittent electricity of around 15% can be handled by the electric grids with positive environmental effects.

Common forecasts predict that PV electricity will contribute only a few percent by 2030. We have shown that there is easily enough land, raw materials, safety protocols, capital, technological knowledge and social support to allow PV to provide over 12% of our electrical needs by 2030. And we have to be much more ambitious for the future because PV has to become the biggest supplier of electricity by the end of the century. This will require finding new ways of energy storage.

The present PV development has been possible by the public support, driven by public opinion, which has led to governments spending substantial money to subsidize PV. However, this has not been a waste. It has been an investment. Today PV electricity is very close to grid parity. Thanks to this we predict that PV will continue to grow at a fast pace towards the 12% goal by 2030. Still strong political support will be necessary. And the promise of significant growth in employment – due to raw materials processing, module manufacturing, installation, and non-PV system components – is becoming a major driving force behind that political support.

PV possesses a panoply of novel technologies that ensure a continuous advance in the reduction of costs throughout the whole century. In this PV is unique compared with other energy technologies. Nations that want to lead this irrepressible movement will have to support it by investing in R&D, industry and markets.

REFERENCES

1. Benka S, *Physics Today* **38**, 39 (2002); adapted from Pasternak A, *Lawrence Livermore Natl. Lab report UCRL-ID-140773* (October 2000).
2. Lewis N, *Material Research Society Bulletin* **32**, 808–820 (2007).
3. Gustavson M R, Limits to Wind Power Utilization. *Science* **204**, 13–17 (1979).
4. Zweibel K, Mason J, Fthenakis V, *Scientific American* **298**, 64–73 (2008).
5. The Effect of Tilt, Orientation, Array Size, etc Can be Effectively Determined for Many Locations Using the PV Watts On-Line Solar Calculator at www.nrel.gov/rredc/pvwatts/version1 .html.
6. Monthly Module Price Index, *Photon International* (November 2009), pp 84–87.
7. Fath P, Keller S, Winter P, Joos W, Herbst W, *Proceedings of the 34 IEEE PVSC*, Philadelphia, pp 002471-76 (2009).
8. Wiser R, Barbose G, Peterman C, Darghouth N, Tracking the Sun – II: Installed costs of PV in the US from 1998–2008, *US Department of Energy Lawrence Livermore Berkley Laboratory* (on-line) 2009. At http://eetd.lbl.gov/ea/emp/re-pubs.html.)
9. Fritts C, *Proceedings of the American Association for the Advancement of Science* **33**, 97 (1883).
10. Chapin D, Fuller C, Pearson G, *Journal of Applied Physics* **25**, 676–677 (1954).
11. Reynolds D, Leies G, Antes L, Marburger R, *Physical Review* **96**, 533–534 (1954).
12. Jenny D, Loferski J, Rappaport P, *Physical Review* **101**, 1208–1209 (1956).
13. Prince M, *Journal of Applied Physics* **26**, 534–540 (1955).
14. Loferski J, *Journal of Applied Physics* **27**, 777–784 (1956).
15. Wysocki J, Rappaport P, *Journal of Applied Physics* **31**, 571–578 (1960).
16. Shockley W, Queisser H, *Journal of Applied Physics* **32**, 510–519 (1961).
17. Cusano D, *Solid State Electronics* **6**, 217–232 (1963).
18. Wysocki J *et al.*, *Applied Physics Letters* **9**, 44–46 (1966).
19. Alferov ZhI, *Fizika i Tekhnika Poluprovodnikov* **4**, 2378 (1970).

20. Lindmayer J, Allsion J, *COMSAT Technical Review* **3**, 1–22 (1973).
21. Hovel H, Woodall J, *Proceedings of the 10th IEEE Photovoltaic Specialist Conference*, pp 25–30 (1973).
22. Perlin J, *From Space to Earth*. Ann Arbor, MI: Aatech Publications, 1999.
23. Annual Survey, *Photon International* 2009-3 (March 2009), pp 170–206.
24. *Photovoltaics International*, 2nd Quarter 2009, 160–162 www.pv-tech.org; also 2008 Annual Report (Table 3), IEA-PVPS www.iea-pvps.org.
25. Data for 1996, 2000, and 2004 from Maycock PV Market Update, published annually in *Renewable Energy World* August Issue (until 2007). Data for 2008 for grid and utility connected from IEA-PVPS Trends in *Photovoltaic Applications* http://www.iea-pvps.org/products/rep1_18.htm. Data for 2008 off-grid from several sources reporting 300-400 MW off-grid installations.
26. PV Resources http://www.pvresources.com.
27. US Department of Energy Solar Energy Technologies Mult-year Program Plan 2001-2011.
28. Gilman P, Blair N, Mehos M, Christensen C, Janzou S, Cameron C, *Solar Advisor Model User Guide for Version 2.0*, NREL Report No. TP-670-43704, 2008.
29. *UN IPCC Fourth Assessment Report: Climate Change 2007*, Working Group III – Mitigation of Climate Change, Section 4.4.3.3.
30. *IEA Technology Roadmap: Solar Photovoltaic Energy* (released May 11, 2010) at www.iea.org/papers/2010/pv_roadmap.pdf.
31. The first paper by these two authors shows that PV might provide 10–20% of the load of a traditional grid system while the second looks at what changes might be made to allow up to 50% penetration of PV: Denholm P, Margolis R, *Energy Policy* **35**, 2852–2861 (2007); Denholm P, Margolis R, *Energy Policy* **35**, 4424–4433 (2007).
32. Luque A, *Progress in Photovoltaics* **9**, 303–312 (2001).
33. Johansson T B, Kelly H, Reddy A K N, Williams R H, Burnham L, *Renewable Energy Sources for Fuel and Electricity*. Washington DC: Island Press, 1993.
34. German Advisory Council on Climate Change (WGBU) Energy in Transition, (2003), www.wgbu.de.
35. http://www.citymayors.com/statistics/largest-cities-density-125.html.
36. *Energy System Emissions and Material Requirements*, Meridian Corporation (Alexandria, VA) report prepared for the Deputy Assistant Secretary for Renewable Energy of the USA (1989).
37. Ito M, Kato K, Komoto K, Kichimi T, Sugihara H, Kurokawa K, *Proceedings of the 19th European PVSEC*, pp 2113–2116 (2004).
38. Andersonn B, *Progress in Photovoltaics* **16**, 61–76 (2000).
39. Feltrin A, Freundlich A, *Renewable Energy* **33**, 180–185 (2008).
40. PV FAQs from www.nrel.gov/ncpv.
41. Sala G, Luque A, Past Experiences and New Challenges in PV Concentrators. in: A Luque, V M Andreev (eds), *Concentrator Photovoltaics*, Berlin: Springer, 2007, pp 1–24.
42. National Photovoltaic Environmental Health and Safety Assistance Center at www.pv.bnl.gov
43. European PV Environmental Health and Safety Working Group http://www.iea-pvps.org/tasks/task12.htm.
44. Fthenakis V, Morris S, Moskowitz P, Morgan D, *Progress in Photovoltaics* **7**, 489–497 (1999); or www.nrel.gov/cdte.
45. Fthenakis V M, Fuhrmann M, Heiser J. Lanzirotti A, Fitts J, Wang W, *Progress in Photovoltaics* **13**, 713–723 (2005).
46. Fthenakis V, Alsema E, *Progress in Photovoltaics* **14**, 275–280 (2006).
47. Ito M, Kato K, Komoto K, Kichimi T, Kurokawa K, *Progress in Photovoltaics* **16**, 17–30 (2008).
48. Fthenakis V, Kim H, Alsema E, *Environmental Science and Technology* **42**, 2168–2174 (2008).
49. Skoczek A, Sample T, Dunlop E, *Progress in Photovoltaics* **17**, 227–240 (2009).
50. Denholm P, Kulcinski G, *Energy Conversion and Management* **45**, 2153–2172 (2004).

51. Fthenakis V, Mason J, Zweibel K, *Energy Policy* **37**, 387–389 (2009).
52. Kempton W, Tomić J, *Journal of Power Sources* **44**, 268–279 (2005).
53. del Cañizo C, del Coso G, Sinke W, *Progress in Photovoltaics* **17**, 199–209 (2009).
54. Barnett A, Rothwarf A, *IEEE Transactions of the Electron Devices* **27**, 615–630 (1980).
55. A truly pioneering classic text on photovoltaic devices is sadly out of print, and only available at a very high price from used book sellers: Fahrenbruch A, Bube R, *Fundamentals of Solar Cells*. New York: Academic Press, 1983.
56. Hall R, Birkmire R, Phillips J, Meakin J, *Applied Physics Letters* **38**, 925–926 (1981).
57. Catalano A *et al.*, *Proceedings of the 16th IEEE Photovoltaic Specialist Conference*, pp 1421–1422 (1982).
58. Mickelson R, Chen W, *Proceedings of the 16th IEEE Photovoltaic Specialist Conference*, pp 781–785 (1982).
59. Tyan Y, Perez-Albuerne E, *Proceedings of the 16th IEEE Photovoltaic Specialist Conference*, pp 794–799 (1982).
60. Phillips J, Birkmire R, Lasswell P, *Proceedings of the 16th IEEE Photovoltaic Specialist Conference*, pp 719–722 (1982).
61. Zweibel K, *The Terawatt Challenge*, NREL Technical Report NREL/TP-520-38350 (2005), p 25, available at www.nrel.gov.
62. Yan, B, Yue G, Owens J, Yang J, Guha S, *Proceedings of the 4th IEEE WCPEC*, Waikoloa, pp 1477–1482 (2006).
63. Repins I, Contreras M, Egaas B, Dehart C, Scharf J, Perkins C, To B, Noufi R, *Progress in Photovoltaics* **16**, 235–239 (2008).
64. Wu X *et al.*, *Proceedings of the 17th European PVSEC* Munich, pp 995–1000 (2001).
65. Yamaguchi M, Luque A, *IEEE Transactions on Electron Devices* **46**, 2139–2144 (1999).
66. Shockley W, Queisser H, *Journal of Applied Physics* **32**, 510–519 (1961).
67. Green M, *Third Generation Photovoltaics*. Berlin: Springer, 2003.
68. Martí A, and Luque A, (eds), *Next Generation Photovoltaics: High Efficiency through Full Spectrum Utilization*. Bristol: Institute of Physics Publishing, 2004.
69. Lewis L, Crabtree G, Nozik A, Wasielewski M, Alivisatos P, *Basic Research Needs for Solar Energy Utilization*. US Department of Energy, Office of Basic Science, 2005.
70. Luque A, and Martí A, *Physical Review Letters* **78**, 5014–5017 (1997).
71. Kolodinski S, Werner J, Wittchen T, Queisser H, *Applied Physics Letters* **63**, 2405–2407 (1993).
72. Ross R, Nozik A, *Journal of Applied Physics* **53**, 3813–3818 (1982).

2

The Role of Policy in PV Industry Growth: Past, Present and Future

John Byrne and Lado Kurdgelashvili

Center for Energy and Environmental Policy, University of Delaware, USA

2.1 INTRODUCTION

Recently, the photovoltaic (PV) industry has experienced phenomenal growth, with market demand expanding at an annual rate in excess of 40% [1, 2]. Technological improvements, increased economies of scale and manufacturing experience have allowed PV manufacturers to lower costs of production and, thereby, stimulate the market. But policy has been an equally important factor, and in some instances the most important driver of an industry boom (e.g. rapid growth in German and Spanish markets) that could rival global experience in computation and communications. Key policy instruments spurring PV's expansion include market and tax incentives (e.g. Feed-in tariffs, rebates and tax credits), regulations (e.g. renewable portfolio standards, new building codes requiring zero-energy capable operation, and solar energy mandates) and public research and development (R&D).

This chapter first provides a comprehensive review of policy strategies in key countries and markets and then offers a model to analyze and compare policy mechanisms to promote still wider and more rapid adoption of PV. With rising costs of conventional fuels and growing concern about carbon emissions from the energy sector, an even more rapid pace of PV's diffusion is likely to be needed if we are to address the challenge of sustainability [3–5].

2.1.1 Changing Climate in the Energy Industry

Over the course of the twentieth century, the energy sector became heavily dependent on fossil fuels and, recently, uranium for nuclear reactors. Although the share of fossil fuels (oil, natural gas and coal) in global energy supply has slightly decreased from 1980 levels (when they comprised 91% of worldwide commercial energy supply) by end of 2008 fossil fuels still supplied 87% of primary global energy (Figure 2.1). When nuclear energy is included, the conventional energy supply system is the source of 93% of current energy use. Due to high supply risks, volatile fuel

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

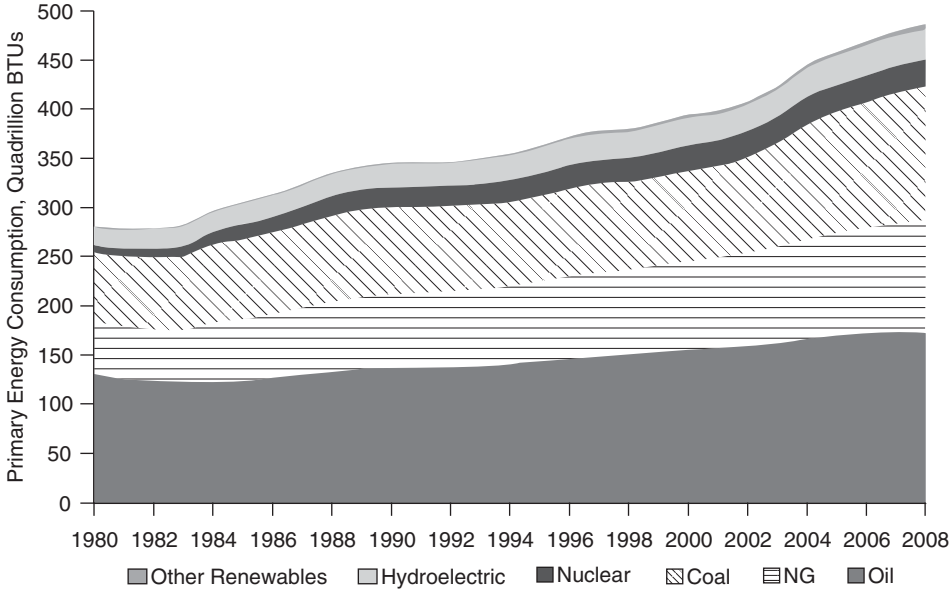


Figure 2.1 Global primary energy consumption. Data source [7]

prices and long-term environmental implications of fossil energy use, the existing energy supply structure is regarded by many as no longer viable [3, 5, 6].

For the last decade, energy prices for conventional power generation have significantly increased (Figure 2.2). Based on data from the US Energy Information Administration for 2000–2008, wholesale prices of residual fuel oil (No. 5 and 6 distillates) have increased by over 219% ($0.61/gallon to $1.94/gallon); the cost of natural gas used for electricity generation has increased by 113% (from $4.38/MCF to $9.35/MCF); the weighted average cost of uranium

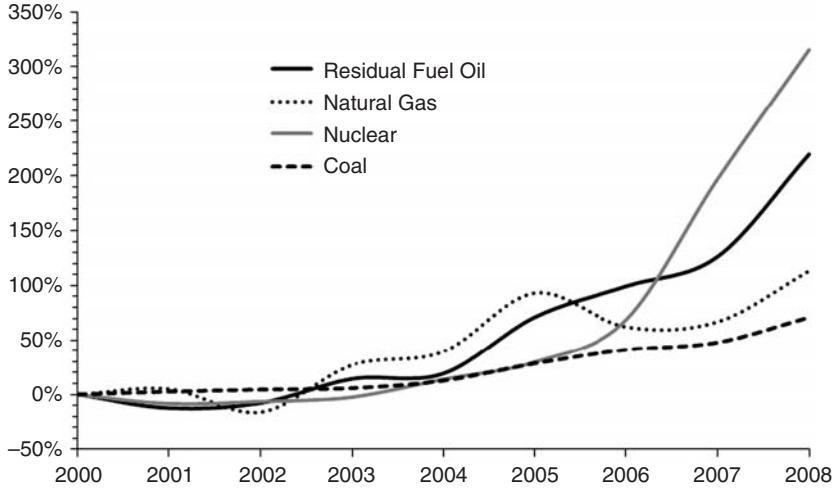


Figure 2.2 Energy price fluctuations for power generation in the US 2000–2008. Data source [9]

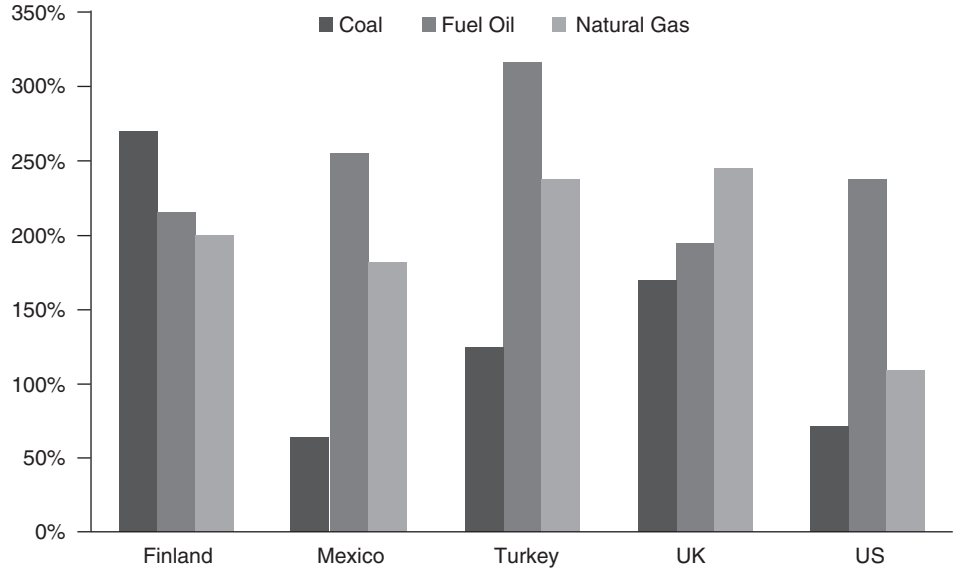


Figure 2.3 Percentage increase in energy price for power generation in selected countries, 2008 compared with 2000. Data source [10]

oxide (U_3O_8) used in nuclear power plants has increased by 316% ($11.04 to $45.88 per million pounds); and coal prices, despite being the least volatile, have also witnessed an increase of 72% ($27.5 to $47.4 per metric ton). Similar trends were also observed in other countries (Figure 2.3). In 2009, on the heels of the worst economic crisis since the 1920s, world oil prices fell, but only by about one-third of their peak 2008 value. Conventional fuel prices and downward pressure on material prices such as steel and copper resulting from the current economic crisis could reduce energy costs to end users in the short term, but the long-term trend is clear – conventional energy will cost more and, soon, much more [7, 8]. Higher conventional energy prices and the likelihood of significant fluctuations in the foreseeable future have made, and will continue to make, energy from PV power increasingly attractive.

Mounting concerns over global climate change and other environmental problems associated with conventional energy use are adding to the momentum to rethink the architecture of the energy sector. According to the IPCC [11], energy efficiency, changes in land use practices and wider adoption of renewable energy technologies are likely to be the principal tools to decarbonize the world economy (Figure 2.4). The baseline scenario in its 2007 assessment of mitigation options contains an IPCC forecast of 340 TW h of 2030 electricity generation, resulting in a decrease of 0.25 gigatons (GT) in CO_{2e} emissions. IEA's *World Energy Outlook* forecasts PV to provide 525 TW h or 2% of electricity demand under the 450 Scenario [7]. As discussed later in the chapter, these baseline scenarios can reasonably be surpassed, with as much as 25% of electricity needs supplied by PV, if the proper policy menu is embraced.

2.1.2 PV Markets

Over the last decade, the global PV industry has grown rapidly, faster than any renewable or non-renewable energy option. World PV annual cell production grew from 277 MW_p in 2000 to 6850 MW_p in 2008 (an annual average growth of more than 40%), reaching a cumulative worldwide

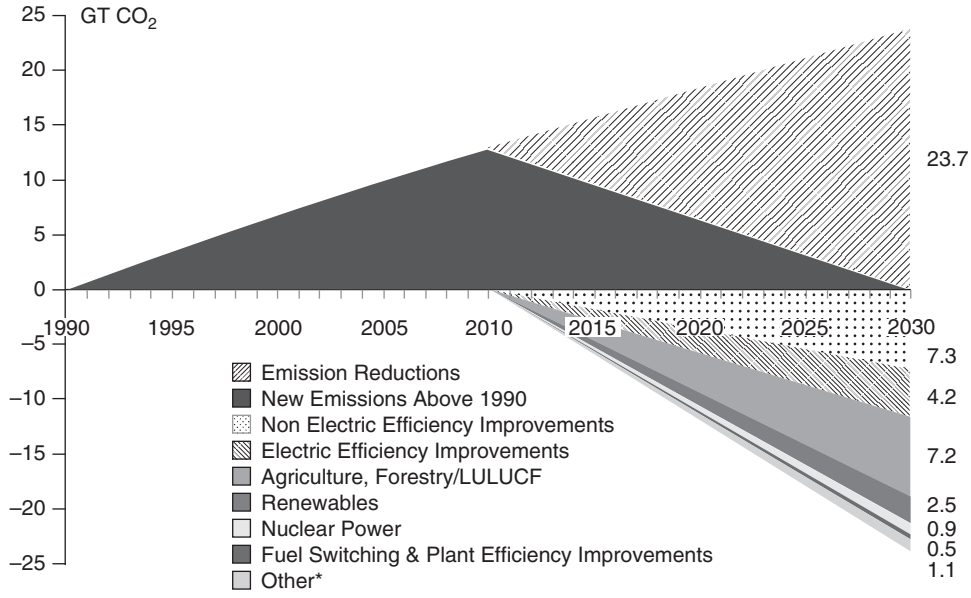


Figure 2.4 Potential GHG emissions avoided by 2030. * "Other" includes CO_2 capture and storage (0.4 GT) and improved waste management (0.7 GT). Data source: [11] (calculated by this chapter's authors based on information in the Fourth Assessment of WGIII)

PV cell shipment of over 19 GW_p at the end of 2008 (Figure 2.5). Japan, Germany and the US are traditional leaders in PV cell manufacturing. However, in recent years new players have emerged, especially China, which manufactured 2150 MW_p of PV in 2008 and became the largest PV producer in the world (followed by Germany (1510 MW_p) and Japan (1230 MW_p)).[1] The US has lost its standing as a manufacturer, producing only 430 MW_p in 2008, less than one-half of the output of Taiwan (865 MW_p). While Spain supplied less than 200 MW_p in 2008, its industry growth is expected to lead it past the US in the next few years [12, 13].

In 2008, annual new PV installations reached a record high of 5950 MW_p [17]. The major PV markets that have fueled PV demand and growth are Spain, Germany, the US, South Korea, Italy and Japan, which at the end of 2008 accounted for 41%, 31%, 6%, 5%, 4% and 4% of demand, respectively, while the "rest of the world" and the "rest of Europe" accounted for 5% and 4% respectively [17]. Although Germany is a leader in cumulative solar PV installations (5.3 GW_p), its annual installations of 1.86 GW_p is now second to Spain, with the latter installing an impressive 2.5 GW_p (up from 0.64 GW_p in 2007). Use of the technology in the US is much slower, reaching 0.4 GW_p (an increase from 0.2 GW_p in 2007) [12, 17]. At the end of 2008 it was estimated that cumulative global PV installations for power production reached 16.4 GW_p [2], with most of the installations occurring in industrialized countries (Figure 2.6).[2] While the cumulative installed capacity of PV reached a significant milestone in 2008, it is only a small fraction (0.4%) of the total global installed electric power generation capacity of about 4000 GW_p [18].

[1] Nearly all of the PV cells manufactured in China are exported to other markets.

[2] Discrepancies between cumulative PV cell shipments and installations can be attributed to delays in installation after shipment. This can be significant under a fast growing PV market. According to industry analysts, on average there is a delay of two quarters between a module being shipped and its connection to the grid [132]. At the beginning of 2009, the industry had started with over 2 GW_p of inventory [21].

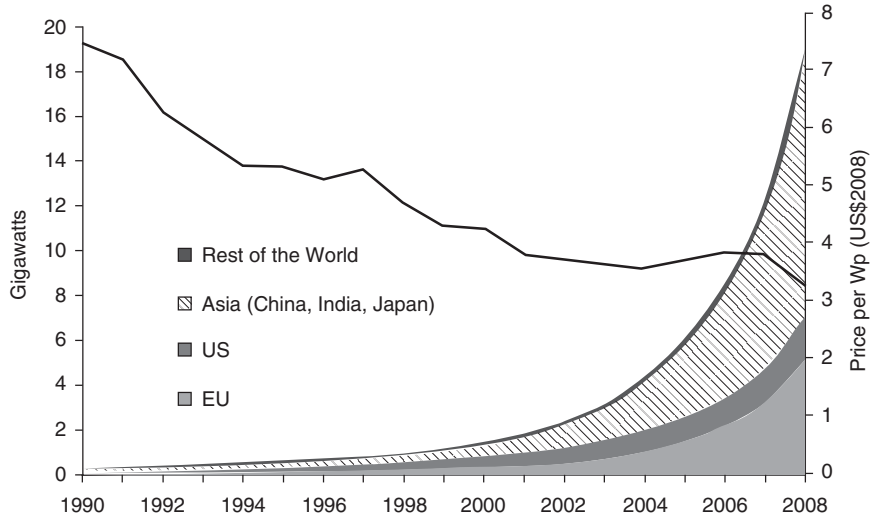


Figure 2.5 World cumulative photovoltaic shipments and retail prices (W_p) 1990–2008 (solid line = price per W_p). Data sources: [12, 14–16]

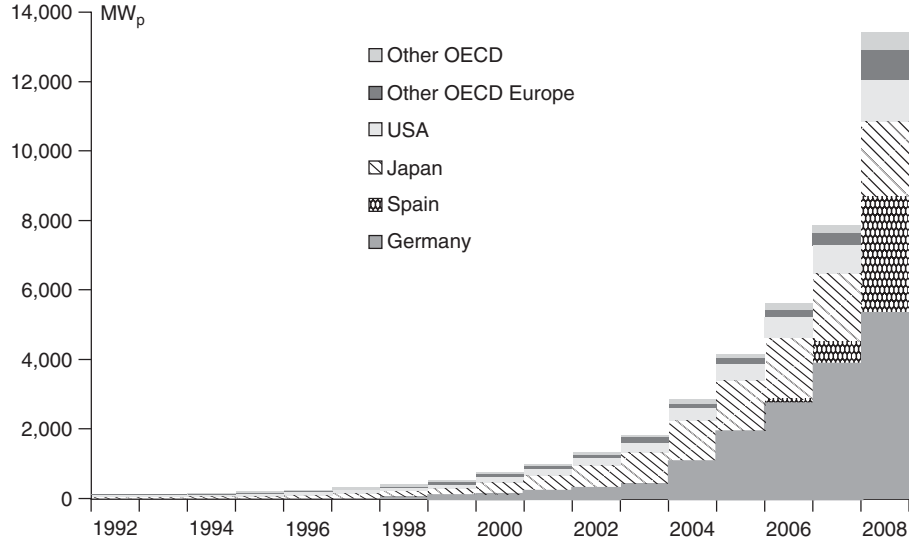


Figure 2.6 Cumulative photovoltaic installations in OECD countries 1992–2008. Data source: [12]

Until recently, upstream manufacturing improvements combined with downstream system integration experience served to drive down PV prices. However, since 2005 high demand for the technology led to a departure from this decade-long trend (Figure 2.5). A major contributor to the price increase was a shortage of polysilicon supply [19, 20]. Polysilicon historically sold at about $35/kg, but since 2004, the spike in demand caused by PV market growth led to spot market prices above $400/kg in 2008 [19, 21]. In response, a number of new polysilicon production lines were added around the world (particularly in China). As more suppliers enter the market, analysts

expect polysilicon contract prices to settle at around $70–80/kg [19]. Accordingly, module prices are expected to return to their historical downward trend by 2010.

2.2 POLICY REVIEW OF SELECTED COUNTRIES

Although significant cost reductions have been achieved, currently electricity produced from PV is not cost-competitive with conventional sources of generation.[3] National and local governments have supported PV deployment through a broad range of incentive, tax, regulatory and R&D instruments that include tax credits and exemptions, preferential interest rates and loan programs, direct incentives (e.g. performance-based incentives, capital subsidies), building code mandates, feed-in tariffs, renewable portfolio standards, voluntary green power programs, net metering, interconnection standards and "demonstration" or pilot projects [12, 22]. Policy leaders for PV deployment are Germany, Spain, Japan, South Korea and the U.S. Each country's national and, in some cases, local policies are reviewed below.[4]

2.2.1 Review of US Policies

In recent years, federal and state policies have facilitated strong demand for PV systems in the US. In 1998, the US had only 100 MW_p of installed PV capacity; ten years later, cumulative PV installations have reached 1.2 GW_p, of which 68% are grid-connected [12]. On the national level, the solar Investment tax credit (ITC) and modified accelerated cost recovery system (MACRS), a tax depreciation rule for PV and other capital equipment, have played key roles in reducing PV developer costs.

The ITC was first established in 1978 under the Energy Tax Act, which provided a tax credit of 15% of installed cost for solar energy installations. The Tax Reform Act of 1986 gradually reduced the ITC to 10%, and it remained at this level until 2005 [23]. The Tax Reform Act also introduced MACRS depreciation rules for commercial entities, allowing PV installations for the business sector to qualify for rapid 5-year tax depreciation. The Energy Policy Act of 2005 (EPAct 2005) increased the ITC to 30%. The Energy Improvement and Extension Act (EIEA) of 2008 removed the previous cap for residential installations (US $2000) and extended the 30% ITC through 2016 [24]. In 2009, the American Recovery and Reinvestment Act (ARRA) allowed commercial entities to receive cash grants from the US Treasury for PV installations occurring in 2009 and 2010. Cash grants provide incentives for those businesses which do not have high tax obligations to fully utilize benefits of the 30% federal tax credit. ARRA also provided a bonus depreciation benefit of 50% for projects implemented in 2009 [25]. For business applications, the ITC and MACRS can reduce initial PV project development costs by more than 50% [26].

In addition to federal policies supporting PV deployment in the US, an increasing number of states have used two instruments to improve PV marketability. One is a policy called a renewable portfolio standard (RPS), in which load-serving entities (LSEs) in the electricity sector must provide a fraction of their electricity supply from PV or distributed renewable energy technologies (which also includes PV). By the end of 2009, 29 states and the District of Columbia had broad RPS mandates. Importantly, 14 states and the District of Columbia (Washington,

[3] It should be noted that this comparison does not consider the relative subsidies for PV and its retail market competitors. When these are factored in, some analysts conclude that PV is very near to a market parity [19, 131]. If pollution and other external costs are included in the cost of conventional fuels, it is likely that PV is less expensive, at least over the long run [133].

[4] While China is the world's largest solar manufacturer, it mostly sells its production to overseas markets. This section focuses on policies to stimulate domestic use of PV and, for this reason, does not include China. Future editions of the *Handbook* will almost certainly need to profile China's domestic market which recently began to expand.

DC) had specific solar or distributed generation requirements for LSEs under their RPS laws. In addition, California, Oregon and Texas have created specific targets for distributed generation or PV unrelated to their RPS laws [25]. The second favored policy instrument among US states is net metering. Currently, 43 states and Washington, DC support net metering of PV electricity. Net metering allows customer-sited PV generators to offset electricity provided by the LSE with kW h supplied by their PV system [27, 25]. With net metering, electricity generated by PV is valued at the retail electricity price, providing additional incentive for PV deployment.

Historically, states and electric utilities have also supported PV deployment through rebate programs. In recent years however, support has begun to shift towards production or performance-based incentives. By early 2010, 29 states had production-based incentives embodied in utility obligations to purchase Renewable Energy Certificates (RECs). Sixteen of these states had solar electric sales mandates, which included either production-based incentives (viz. solar REC purchase obligations for utilities) or credits (applied to meet RPS mandates) [25]. Nevertheless two states, California and New Jersey, represented 67 and 9%, respectively, of the total US grid connected systems, and were the policy leaders in the country. Their approaches to market development have been successful, causing, for example, an increase in grid-connected PV installations between 2005 and 2008 of 208% (California) and 622% (New Jersey) [28–30]. The PV policies of these key states are reviewed below.

2.2.1.1 California

California has a long history of solar market development. In 1984, the Sacramento Municipal Utility District (SMUD) installed a 1 MW_p PV plant (PV1) – one of the first large-scale PV power plants in the world [31]. Over the past two decades, PV1 showed steady performance and it was gradually expanded, reaching 3.2 MW_p by 2004 [32]. In 1993, Pacific Gas and Electric Company (PG&E) installed a grid-connected 500 kW_p PV system (in Kerman) to serve peak power demand. Performance of the PV system demonstrated that PV output could reduce coincident utility load peaks [31–33]. More significantly, it demonstrated the value of PV to the utility in avoided costs that were comparable to the value of electrical energy itself [34].

In 1998 as part of California's electricity sector deregulation, financial incentives were created for renewable energy technologies under the California Energy Commission's Renewable Energy Program [35]. The initiative contained a special provision for "emerging renewables" which referred specifically to on-site generation technologies – primarily PV and small wind. From 1998 to 2004, the California Energy Commission's Emerging Renewables Program (CEC-ERP) offered rebates, which on average amounted to 40% of the installed price, to reduce (buy-down) the initial cost of the system. Beginning in 2005, the CEC-ERP offered participants the following options: (1) they could receive rebates amounting to 40% of installed costs; or (2) they could receive incentive payments based on actual system performance in the amount of 50 cents per kW h for three years [36, 37]. The CEC-ERP supported PV installations by customers of the three major investor-owned utilities (IOUs) serving the state PG&E, Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E).[5] Under the program (which ended in 2006), PV system size was limited to 30 kW_p. By the end of the program, 120 MW_p of grid-connected residential PV had been installed (this includes projects started under the program and completed in 2007 and 2008) [38].

[5] As with many jurisdictions in the US, IOUs are only one source of electricity supply. Customers may also receive power from so-called municipal or publicly owned utilities (utilities owned and operated by a governmental jurisdiction such as a city or incorporated region), electric cooperatives (suppliers owned by their customers which often are not subject to conventional utility regulation), and special federal authorities such as the Tennessee Valley Authority and the Bonneville Power Administration. IOUs serve approximately 97 million customers, while municipal utilities, cooperatives, special federal and state authorities together serve 40 million customers. Retail power marketers serve the remaining 6 million customers of the US [123].

In 2001 the California Public Utilities Commission created its Self-Generation Incentive Program (CPUC-SGIP). It was intended to complement the California Energy Commission's program[6] and provided incentives for PV installations exceeding 30 kW_p. By end of 2008, 135 MW_p of grid-connected PV systems were installed under this program [38].

These two policy initiatives were instrumental in promoting PV markets for IOU service areas. At the same time, a number of publicly owned utilities (POU) began developing policies to support PV installations within their service territories. Sacramento Municipal Utility District (SMUD) and Los Angeles Department of Water and Power (LADWP) were two major POUs pioneering PV use. As previously motioned, SMUD was one of the first public utilities in the world with a large-scale PV installation. Between 1998 and 2007, 11 MW_p of on-site PV was installed, utilizing a then-unique policy tool in which a SMUD citizen or business could elect to pay a higher electricity price and the utility would install, operate and maintain the system. This policy tool gave rise to a stream of new policies culminating in the property-assessed clean energy (or PACE) program in which electricity users voluntarily pledge an increase in their property tax assessments in order to retire the capital debt incurred by the installation of the PV system. This model is now being imitated across the US, and is the subject of national legislation [39, 40].

On August 20, 2004, California's governor announced the Million Homes Solar Plan which laid the groundwork for the 2009 Go Solar California campaign. Go Solar California aims to install an additional 3.0 GW_p of PV in the state within 10 years and is funded by ratepayers in the amount of $3.3 billion [41]. Go Solar California was launched in 2007 and includes two new solar incentive programs – the California Solar Initiative (CSI) and the New Solar Homes Partnership Program (NSHP). Residences that are served by publicly owned utilities (e.g. local municipal utilities) are not eligible for the CSI and NSHP programs. However, California now requires publicly owned utilities to offer an equivalent incentive program for their customers [42].

The CSI began in 2007 and led to the rapid installation of more than 130 MW_p in PV installations in one year (Figure 2.7). For 2007–2016, the CSI Program has a budget of $2.2 billion and a target of 1.75 GW_p of installations from the mainstream incentive program and an additional 190 MW_p from its low-income program (CPUC, 2008). Initially rebates stood at \$2.50/W_p for residential and commercial systems and \$3.25/W_p for government entities and the nonprofit sector. The incentive levels are scheduled to decline as the aggregate capacity of PV installations increases [25]. Installed and in-pipeline projects have already met 20% of the CSI target [38, 41].

The New Solar Homes Partnership Program provides funding for builders and developers who install PV systems on new, energy-efficient residential buildings that are served by investor owned utilities. NSHP is administered by the California Energy Commission. The program has a budget of $400 million and a goal of installing 400 MW_p of PV on new homes by 2016. This includes a 36 MW_p target for new low-income housing (California Energy Commission, [44]). Initially rebates range from \$2.50/W_p to \$3.50/W_p (for low income housing) and gradually decrease as PV installations increase [25].

The CSI and NSHP incentives are designed to stimulate rapid market demand while reducing incentive levels as the market for PV becomes viable. Depending on system size and customer choice, incentives are paid on a dollar-per-watt or cents-per-kilowatt-hour basis. The former is referred to as an expected performance-based buy-down (EPBB) incentive and the latter is called a performance-based incentive (PBI). EPBB is intended for residential and small commercial customers with systems less than 50 kW_p capacity. The incentive is in the form of a lump-sum, up-front payment. PBI is intended for large commercial, government and nonprofit customers. It is mandatory

[6] The CPUC has regulatory authority over the IOUs serving the state, while the California Energy Commission has responsibility for long-term energy policy and planning with special responsibilities for the promotion of energy efficiency, conservation and renewable energy (see www.energy.ca.gov/commission/index.html).

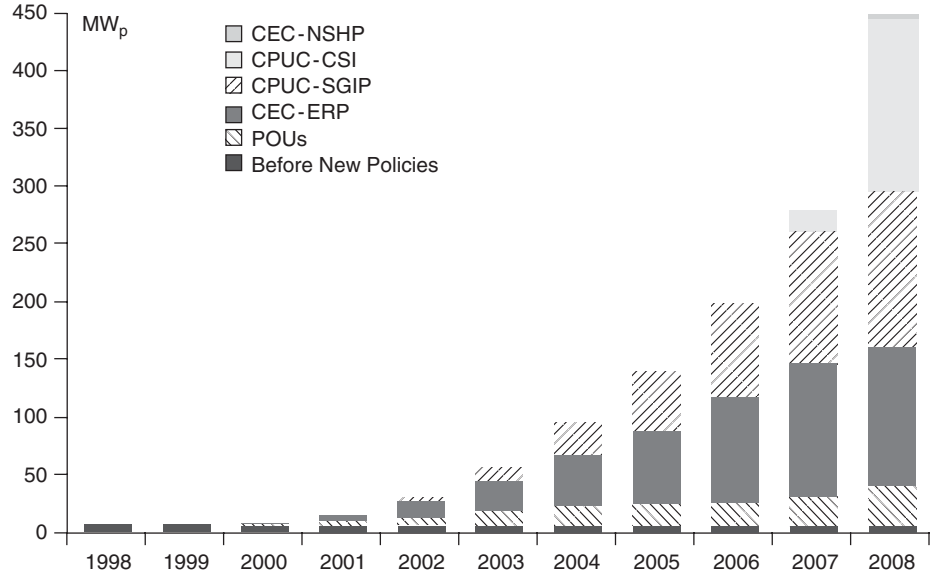


Figure 2.7 Cumulative grid-connected photovoltaic installations in California by policy, 1998–2008. Data source: [38]

for all systems greater than 50 kW_p capacity (systems less than 50 kW_p in size can opt in to PBI). The PBI program provides payments to PV users of 40–50 cents per kW h for five years. The incentive levels are scheduled to decline as the aggregate capacity of PV installations increases. Both incentives are performance based. In the case of EPBB, lump sum payments are predicated on a system's *expected* performance (factors include system AC rating, location, orientation and shading). This requires accurate and transparent predictive models. In the case of PBI, incentives are based on *actual* energy production and monthly payments are made during a 60-month period [42–44].

Energy policy innovation has been a hallmark of California for decades, and the promotion of PV use is no exception. The state hosts the largest PV capacity and greatest number of installations per capita of any US jurisdiction. Its policy tools have been widely adopted both in and beyond the country.

2.2.1.2 New Jersey

New Jersey has grown to be the second largest PV market in the US (by the end of June 2009, the state had 90 MW of installed PV systems). It was one of the first states to set specific targets for renewable energy sources in state electricity supply. In 1999 under the Electric Discount and Energy Competition Act (EDECA), a statewide renewable portfolio standard (RPS) was adopted and went into effect in 2001. A specific carve-out for PV was included which requires load-serving entities to procure 2.12% of electricity from PV by 2020. EDECA and the RPS laid the foundation for New Jersey's Clean Energy Program administered by the New Jersey Board of Public Utilities (BPU) [45].

From its launch in 2001, the Clean Energy Program (CEP) has directed significant funds for renewable energy development. Under CEP there were two initiatives supporting renewables: the Customer On-site Renewable Energy (CORE) strategy and the Renewable Energy Project Grants and Financing (REPGF) opportunity. CORE provided rebates for on-site renewable generation projects with less than 1 MW_p capacity. REPGF was to support development of so called Class 1 renewable

energy resources (which includes PV, solar thermal electric, wind, geothermal, fuel cells, landfill gas recovery and sustainable biomass) larger than 1 MW_p capacity for power generation [45].[7]

CORE has proved to be instrumental for PV market development in the state. Initially it provided rebates from \$3.75/W_p (for 100–500 kW_p systems) to \$5.50/W_p (for systems less than 10 kW_p). Later rebates were gradually reduced as installations increased and prices fell [46]. Under the program, 70 MW_p of PV have been installed [47] and an additional 50 MW_p of PV systems have been approved for rebates. In 2009, the Clean Energy Program redesigned its incentive program based on the success of CORE. A Renewable Energy Incentive Program (REIP) was created with lower rebates, but aggressive pricing for "solar renewable energy credits" (see below). Under REIP, a residential customer can receive a rebate of \$1.75/W_p for up to 10 kW_p of installed on-site PV if the customer agrees to receive a free energy audit (the rebate falls to \$1.55/W_p without an audit). Nonresidential customers can receive \$1.00/W_p rebates for up to 50 kW_p of installed PV [48].

The backbone of New Jersey's solar policy is now its Solar Renewable Energy Credits (SREC) initiative. In a significant departure from its previous incentives, up-front capital incentives are being phased out, replaced by an emphasis on performance-based production incentives. In fact, the state intends to terminate all rebates by 2012 [49]. SRECs are tradable certificates that represent the clean energy benefits of electricity generated from a solar electric system. Each time a PV system generates 1 MW h of electricity, an SREC is issued that can then be sold or traded separately from the power. New PV projects and projects already in the CORE program queue are eligible to participate in the SREC program. However, starting in 2009 customers were required to forgo rebates to participate in the SREC program.

New Jersey has also adopted an 8-year Solar Alternative Compliance Payment (SACP) schedule intended to enable project financing for large PV systems without up-front rebates. Utilities are required to pay an SACP of \$711 per MW h if they do not meet the state's Solar RPS through the purchase of SRECs. The SACP schedule gradually declines, reaching \$594 per MW h in 2016 [25–49]. The high SACP rate (the highest in the US) has led to high market prices for SRECs. In May 2009, for example, the weighted average price for SRECs was \$500/MW h, much higher than in previous years when prices hovered around \$240/MW h [50]. Only SRECs from PV installed within the state can be used by utilities to comply with New Jersey RPS [51].[8]

As a result, the SREC initiative has spurred rapid growth in the state's PV installation rate, outpacing the experience of the state's earlier and quite successful CORE program (Figure 2.8). When given the choice between up-front capital incentives (REIP) and production incentives (SRECs), customers have shown an overwhelming preference for the latter.[9] This policy innovation is now being actively considered in many jurisdictions throughout the country.

2.2.1.3 Other states

While California and New Jersey are acknowledged leaders of US solar policy innovation, several other states also qualify as pioneers in this area. Table 2.1 identifies ten American states with the highest per capita PV installation rates by the end of 2008. Importantly this group includes not only "sunny" locations or large markets, but also smaller states (e.g. Delaware with a population

[7] New Jersey has yet to build a PV project under the REPGF program.

[8] Across the US, policies regarding out-of-state SREC registration varies. At the beginning of 2010, New Jersey, and Maryland did not allow an out-of-state SREC registration, while Delaware, Ohio, Pennsylvania and the District of Columbia have accepted out-of-state SREC registrations [51].

[9] Customer support for the SREC approach may reflect the preference of solar project developers who reduce prices when SRECs are assigned to them. Because an SREC assignment for 8 years represents a predictable revenue stream, developers can sometimes find it easier to borrow needed capital from lending institutions.

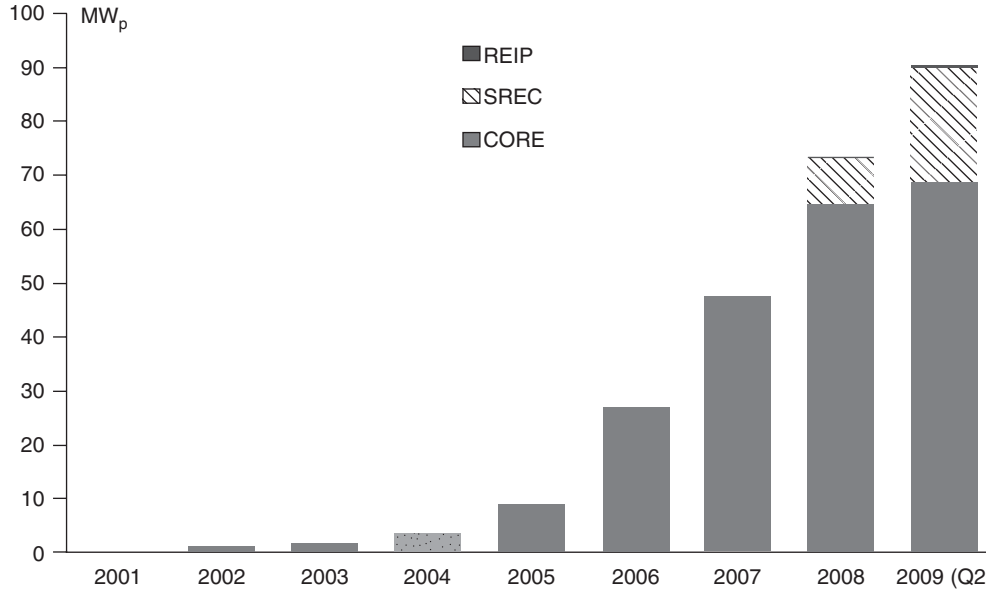


Figure 2.8 Cumulative grid-connected photovoltaic installations in New Jersey 2001–2009 (Q2). 2009 includes cumulative installations through second quarter (Q2) of 2009. Data sources: [52, 53]

Table 2.1 Top ten states by per capita capacity

	Per capita installed power in 2007 (W_p/person)	Per capita installed power 2008 (W_p/person)
California	9.1	14.6
Nevada	7.8	14.2
Hawaii	3.0	10.6
New Jersey	5.0	8.1
Colorado	3.1	7.7
Arizona	3.1	4.3
Connecticut	0.8	2.5
Delaware	1.4	2.2
Oregon	0.8	2.1
Vermont	1.2	1.8
National average	1.6	2.7

Data sources: [28, 29]

of approximately 900 000 residents) and those with above-average colder and cloudy days (Connecticut and Oregon). As the diversity of state leaders indicates, solar development is not necessarily driven by geographic, weather or insolation characteristics. Policy – both governmental and business – shapes how and how much PV is used. Four of states can be used to illustrate this point.

Nevada deploys PV on a per capita basis at the second highest rate in the country. The drivers for its performance are intersecting government and utility policy initiatives. The state was one of the earliest to create a carve-out provision in its RPS, requiring 1.5% of electric sales to come from PV by 2025. Having established an aggressive schedule for utility involvement in solar

energy development, the state organized a Task Force on Energy Conservation and Renewable Energy [25] to design utility programs to spur the market. A rebate program was launched in 2004 providing \$2.30/W_p of PV installed on residences and small businesses, and \$5.00/W_p for public buildings [54]. In addition, the state's largest utility agreed to buy SRECs for 20 years from the 14 MW_p PV plant at Nellis Air Force Base, making the project profitable. The state also took the initiative of supporting a 12.6 MW_p thin film solar farm in 2008 and its private backers – Sempra and First Solar – were able to obtain a long-term power purchasing agreement with the California utility Pacific Gas and Electric. The project showcases the benefits of a public–private partnership, with installed cost at \$3.20/W_p which some analysts have suggested is comparable to grid power [55].

Colorado also illustrates the importance of public–private partnerships. With only a moderate carve-out target of 0.8% of solar from PV by 2020, the state nonetheless hosts the third highest installation volume among the American states [28]. Its market has grown because utility programming has attracted investors. The state's largest utility, Xcel Energy, initiated a Solar Rewards Program that includes a rebate of \$2.00/W_p coupled with an SREC purchase agreement for 20 years at \$55 per MW h. Other utilities have followed suit, pushing Colorado's installed capacity above 38 MW_p by the end of 2009 [25, 56].

While Nevada and Colorado can count on large markets and good to excellent insolation, a state like Delaware has neither. Yet, its progress in promoting solar development is impressive. In 2005, the state adopted one of the most aggressive carve-out targets in the country, by which 2% of sales must come from PV by 2019, and created incentives which cover approximately one-third of installed costs [25]. By the end of 2009, installations already surpassed the RPS target for 2011 [57, 58]. But a major driver in Delaware has been its creation of the country's first Sustainable Energy Utility (SEU). With capital from proceeds of the state's participation in quarterly carbon allowance auctions under the Regional Greenhouse Gas Initiative [59], the SEU has become the largest SREC off-taker in the initial production years of the 10 MW_p SUN Park in the state's capital, Dover. The SUN Park will be completed in 2011 and will be one of the largest solar plants built on the American east coast. The SEU's participation is a key reason for the SUN Park's favorable economics. The SEU's program for SREC pricing has stimulated additional projects in Delaware, including a 2 MW_p rooftop application as part of a two-phase 6 MW_p PV installation at the state's largest university [60, 61] and another 2 MW_p distributed application on four campuses of the state's community college. Completion of these projects will catapult Delaware to a leader on a per capita basis from its present ranking as eighth (Table 2.1). The rapid growth in Delaware's sustainable energy market received national attention, with a recent story in the *New York Times* complimenting the SEU for its innovative policies [62].

Another small state – Vermont – is advancing an innovative strategy for PV market development. In 2005, its legislature created the Clean Energy Development Fund (CEDF) with authority to invest in PV and other clean energy options. The CEDF offers a wide range of financing from grants to loans, equity investments and direct incentives [63]. So far, the Fund has underwritten approximately 1 MW_p of PV installations [63]. Additionally, 1.7 MW_p was installed under Vermont's Small Scale Renewable Energy Incentive Program [64]. Vermont has also created its own tax credit for businesses investing in solar systems which covers 30% of initial capital outlays. Together with US tax credit of 30%, the state has dramatically lowered the up-front cost hurdle for investors [25].

Each state has crafted policies and programs that seek to take best advantage of the particular market and social assets of their jurisdictions in order to stimulate rapid growth in the utilization of PV. While some may worry that such policy diversity may create market confusion, state initiatives in the US have proven to date to be the incubators of policy innovation creating the country's fast expanding demand for solar energy. Indeed, researchers have shown that state policy innovation has nurtured a powerful civil society commitment to sustainable energy which is effectively challenging the country's traditional energy policies [65].

2.2.2 Europe

2.2.2.1 Germany

Germany has more than a 25-year policy history of promoting PV use. In 1983, with government support, the first 4 kW_p grid-connected PV system in Europe was installed on the roof of an occupied residence in Munich [66]. Yet, the German PV market was still in its infancy, accounting for only 1 MW_p in cumulative installations in 1989. The German 1000 Roofs Measurement and Analysis Program, introduced in 1990 as a pilot, spurred interest in the technology and led to more than 2000 roof-mounted systems with a capacity of 5.3 MW_p in just 5 years [67]. Performance of these systems was extensively monitored by government and university researchers, which led to significant technical and regulatory improvements. Several federal- and state-funded programs followed, which furnished capital subsidies per kW_p of installed PV ranging from 25 to 50% of initial investment costs [66]. These programs provided the initial government stimulus for funding PV systems.

At the end of 1990, Germany adopted the world's first feed-in tariff (FiT). Under the law, electric utilities were required to purchase electricity from PV systems at a price equal to at least 90% of retail electricity rates [68]. The first feed-in rate was not sufficient to spark significant development of PV. In contrast, wind, which initially shared the same tariff as PV, increased its penetration in electricity supply from 0% in 1990 to 1% in 1998 [18]. Nevertheless, the feed-in law, in combination with the country's 1000 Roofs Program and local grant initiatives, created a PV market of 54 MW_p of installed capacity by 1998 – a tenfold increase in just 5 years [69]. This experience would set in motion a policy regime that has arguably proved to be the most successful in the world.

In 1999, the national government initiated the 100 000 Roof Solar Energy Program, providing 10-year, zero-interest loans with the final installment (10% of the principal) being waived. By end of 1999, nearly 4000 systems with a total capacity of 10 MW_p were installed under the program [66]. In 2000, the government adopted the Renewable Energy Source Act (RESA), which increased the feed-in tariff for PV sixfold (from US$ 0.08 to US$ 0.50) and required utilities to sign minimum contracts of 20 years for a system's output [70, 71]. The high feed-in tariff, long contract length, and favorable financing through zero-interest loans created a rush to install PV projects throughout the country. During the first 4 months, more than 70 MW_p of PV projects sought government and utility support [66]. This was more than the existing 69 MW_p of installed capacity accumulated since 1983 [69]; in other words, in four months the new policy had created demand for PV that had taken 17 years to realize under the old approach. The capacity was much higher than anticipated (the original plan was to install 27 MW_p by 2000) and the government put a temporary moratorium on applications, dropped the waiver provision for the final installment of the loan, and increased the loan interest rate to 2%. It also increased the target for new PV system installations in the first year of the program from 27 to 50 MW_p, and shifted by one year the program's final target of 300 MW_p from 2004 to 2003.

The 100 000 Roofs Program met its goal by the end of 2003. To maintain growth in the PV market, the government then created the Solar Power Generation Program, administrated by the KfW Promotional Bank [72], to continue its low-interest financing incentive. In 2004, the Renewable Energy Source Act was again amended, setting the feed-in tariff still higher, to between €0.54 per kW h (for systems larger than 100 kW_p) and €0.57 per kW h (for systems smaller than 30 kW_p) for building-based applications; and for ground-mounted systems, the feed-in rate was set initially at €0.46 per kW h [73].[10] The law required a decline in the tariff paid to PV system owners of 5% per year for building-based systems and 6.5% per year for ground-mounted systems. This

[10] These rates in 2004 US dollars are equivalent to $0.66–$0.70 per kW h for building-based PV systems and $0.57 for ground-mount edPV systems (see http://www.bankofcanada.ca/en/rates/exchform.html).

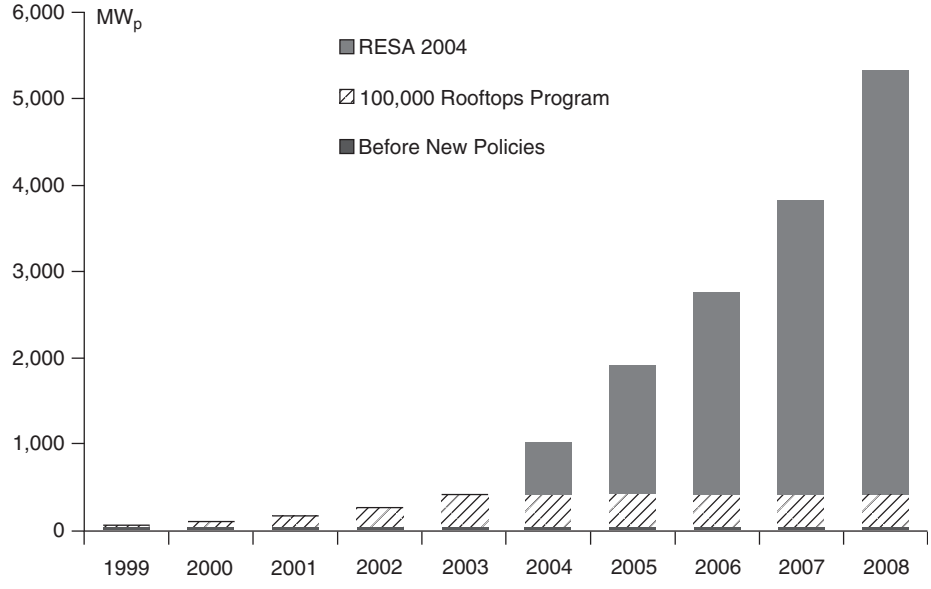


Figure 2.9 Cumulative grid-connected photovoltaic installations in Germany by policy, 1999–2008. Data sources: [12, 66]

reduction schedule applied through the end of 2008 (e.g. the 2008 feed-in rate paid to PV building system owners for applications smaller than 30 kW_p was €0.47.[11] The revised feed-in law created strong, sustained demand for PV in Germany and by the end of 2008 cumulative installed power reached 5.3 GW_p (Figure 2.9).

In 2009, the annual FiT schedule was again adjusted, with PV tariffs declining by 8% per year in place of the earlier 5% rate and the new rate applies to all systems less than 100 kW_p in size. For all ground-mounted applications and building-based systems greater than 100 kW_p, the tariff declines 10% per year. Beginning in 2011, the FiT is set to fall 9% per year for all systems. Additional adjustments may occur if actual annual PV installation rates grow faster than projections (Table 2.2). If the upper target is achieved early in a given year, the decrease in price will increase by 1%. Likewise, realization of the lower target will slow the FiT reduction rate by 1% [12, 74].

Table 2.2 German feed-in tariff reduction schedule: 2009–2011

	2009	2010	From 2011
Tariff reduction rate for small systems (<100 kW)	7% (<1000 MW) 8% (1000–1500 MW) 9% (>1500 MW)	7% (<1100 MW) 8% (1100–1700 MW) 9% (>1700 MW)	8% (<1200 MW) 9% (1200–1900 MW) 10% (>1900 MW)
Tariff reduction rate for large and ground systems (>100 kW)	9% (<1000 MW) 10% (1000–1500 MW) 11% (>1500 MW)	9% (<1100 MW) 10% (1100–1700 MW) 11% (>1700 MW)	8% (<1200 MW) 9% (1200–1900 MW) 10% (>1900 MW)

Data source: [12, 74].

[11] Due to the decline in the value of the US dollar, this rate was *higher*, nearly $0.74 per kW h, when valued in American currency.

Through a combination of low-interest financing and multi-year FiT pricing, Germany grew its market faster than any country had previously achieved and catapulted it to the top of the world's nations in installed PV capacity. This stunning achievement demonstrates the central role of policy in PV market development.[12]

2.2.2.2 Spain

Until 2004 Spain did not have a sizable PV market (2003 installed PV capacity stood at 12 MW_p [12]). In 2004, the country's law governing renewables [75] was amended establishing a new legal and financial framework for renewable energy applications. For PV systems with a capacity of less than 100 kW_p, a feed-in rate of 575% of the reference tariff for 25 years was created; any output after 25 years of operation would receive 460% of the reference tariff. The reference tariff was based on the national average electricity generation price.[13] For large systems (i.e. above 100 kW_p) the Spanish feed-in rate was set at 300%. The high FiT and the requirement that electric utilities purchase power from PV systems for a minimum of 25 years led to a rapid rise in PV installations.

By the end of 2004, PV installed capacity almost doubled, reaching 23 MW_p. In 2005 the Spanish government approved a new Renewable Energy Plan which established a national target of 400 MW_p PV installed by 2010 [76]. The announcement of the new Plan spurred even faster growth and by 2006 installed PV capacity had tripled from 48 MW_p in 2005 to 145 MW_p. Indeed, expansion of the Spanish market was so quick that the Plan target of 400 MW_p by 2010 was reached by the fall of 2007. The government promptly increased its target to 1200 MW_p [77]. In 2007 by

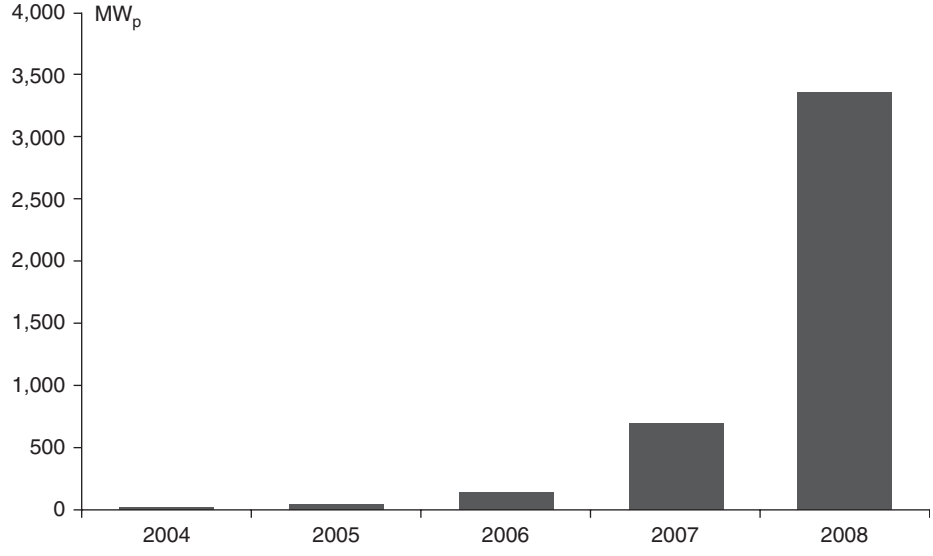


Figure 2.10 Cumulative grid-connected photovoltaic installations in Spain 2004–2008. Data source: [12]

[12] The importance of policy applies more generally to renewable energy as the German case confirms. The country has employed its FiT and financing approaches to wind, solar hot water and biomass markets with similarly impressive results, making Germany the leader in newly installed capacity in all of these markets [71].

[13] In 2004, the average price in Spain was 7.24 Euro cents per kW h or 9.0 US cents, making the PV FiT equal to nearly 52 US cents for 25 years.

Royal Decree 661, the government modified the FiT structure for PV systems: small systems (less than 100 kW_p) receive €0.44 per kW h; a new category of system size with a range of 0.1–10 MW_p, receives €0.42 per kW h; and large systems (10–50 MW_p) receive €0.23 per kW h. All contracts are for a minimum of 25 years [78].[14]

In 2008, Spain's PV market experienced even higher growth, increasing installed capacity fivefold, to 3.4 GW_p of installed capacity, second only to Germany [12]. The extremely rapid rate of installation within one year created significant financial pressure on ratepayers [77] and distorted world PV module pricing. In September 2008, Royal Decree 1578 set new feed-in tariffs for PV systems installed after September 29th of 2008. Under the new tariff structure, roof-mounted systems smaller than 20 kW receive €0.34 per kW h for 25 years. Roof-mounted systems larger than 20 kW_p and ground-mounted systems receive €0.32 per kW h.[15] The decree also capped the size of systems receiving FiT support at 2 MW_p for roof-mounted and 10 MW_p for ground-ounted systems. Total new installations receiving FiT support are limited to 500 MW_p in 2009, 502 MW_p in 2010 and 488 MW_p in 2011 [79] (Figure 2.10).

2.2.3 Asia

2.2.3.1 Japan

Japan has a long track record of supporting PV system deployment. In 1992, the government started the PV Field Test for Public and Other Facilities Program, which led to the installation of 4.9 MW_p of PV on public buildings such as schools, hospitals, clinics and government offices by 1997 when the program was ended. In 1997, the Japanese government through its Ministry of Economy, Trade and Industry (METI) initiated the Residential PV System Dissemination Program, managed by a government-created New Energy Foundation (NEF).[16] The program was instrumental in promoting rooftop PV technology in Japan. Initially, subsidies covering 50% of installed costs were provided to residential customers, with the subsidy rate declining as system costs fell [80]. The program was closed in 2005 after resulting in 932 MW_p of installed PV capacity. This was nearly two-thirds of the total installed capacity of Japan [81, 82]. The government concluded that the program was no longer needed because market mechanisms were sufficient to drive growth [81].

In 1998, Japan initiated a Field Test Project on Photovoltaic Power Generation Systems for Industrial and Other Applications, which led to a total of 18.1 MW_p of PV installations by the time it ended in 2002. A successor program titled the Field Test Project on New Photovoltaic Power Generation Technology has led to 62 MW_p, of industry-scale installed PV [83]. The purpose of the latest field testing program is to promote medium- and large-scale PV systems, with 50% of the system cost subsidized by the government.

The government has now folded its PV promotion efforts into a broad action plan to create a low-carbon society. Under the plan, 70% of new buildings are to have PV systems on their rooftops [84]. The plan sets aggressive goals for PV installation of 14 GW_p by 2020 and 53 GW_p by 2030 [12]. In addition to the national programs, up to 300 local governments have announced programs to support of PV installations. One of the largest programs was announced by the Tokyo Metropolitan Government, which is supporting the installation of about 1 GW_p of PV systems on homes and apartment buildings by 2010 [12. 84]. In addition, electric utilities have announced plans to build 30 centralized PV power plants with a total capacity of 140 MW_p by 2020 [12] (Figure 2.11).

[14] In 2007 US dollars, these FiT rates are $0.60, $0.57 and $0.31 per kW h.
[15] In 2008 US dollars, these FiT rates are $0.63 and $0.50 per kW h, respectively.
[16] The program began as the Residential PV System Monitor Program in 1994 [81].

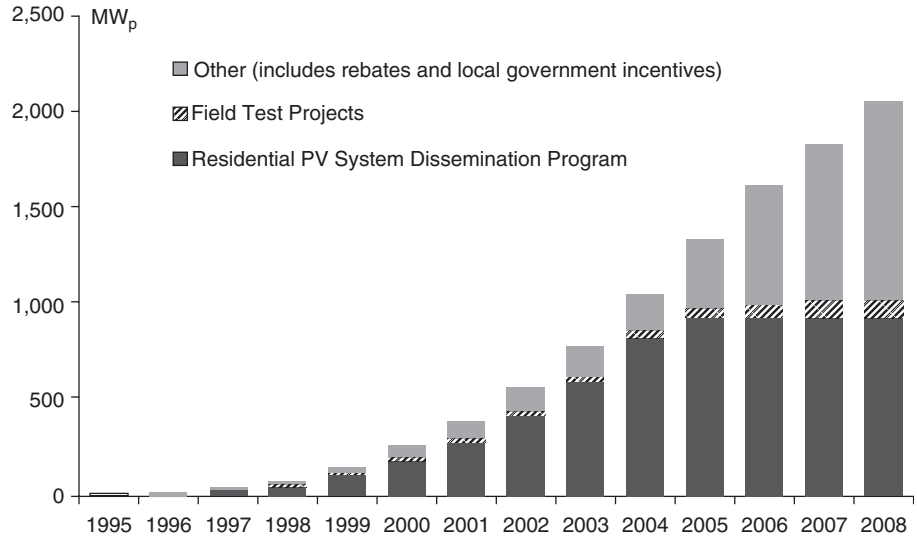


Figure 2.11 Cumulative grid-connected photovoltaic installations in Japan by policy, 1995–2008. Data sources: [12, 81, 83, 85–87]

2.2.3.2 South Korea

Since 1993, South Korea's Ministry of Commerce, Industry and Energy (MOCIE)[17] has been responsible for implementing PV demonstration and field test projects. However, results were meager – by 2004 less than 10 MW_p of PV systems were installed, half of which were off-grid applications (12). In December 2003, the Korean government announced a goal of meeting 5% of total energy consumption from renewable sources by 2011. For PV, a goal of 1.3 GW_p in cumulative installed capacity was set for 2012 and 4 GW_p by 2020 [88]. Following this announcement, growth in the PV market was significantly accelerated.

Successful programs include a rooftops initiative in which the government initially supported 50% of the installed cost of 1–3 kW_p PV installations in the residential sector, hoping to reach its target of 100 000 rooftops by 2012 [89]. In 2008, the government adopted a national plan to construct one million green homes and 200 green villages by 2020. For this task the government provides 60% of the initial PV system cost for single-family and private multi-family residences, and covers all initial costs for public multi-family rental buildings [12].

The government further supports PV development in the public sector through its General Deployment Program and Public Building Obligation Program. The Program serves schools, public facilities, and universities. Under this program PV systems from 5 to 200 kW_p are installed with government funds covering up to 60% of the installation cost. Under the program, new public buildings larger than 3000 m^2 must spend 5% of their total construction budget on renewable energy system installations [89, 89].

These programs targeting public buildings and residential sector have played important role, but the major driver for the recent significant expansion of PV installations in Korea is a newly adopted feed-in tariff policy. Feed-in tariffs are paid for 15–20 years at 700 Korean Won per kW h.[18] By the

[17] The name of the ministry was recently changed to the Ministry of the Knowledge Economy (MKE).
[18] In 2008 US dollars, this FiT rate is equivalent to $0.68 per kW h.

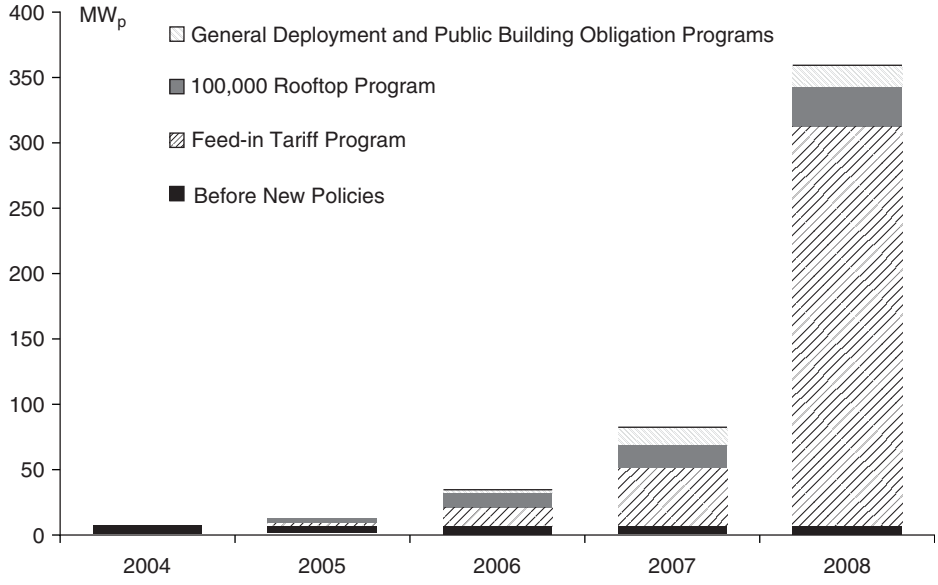


Figure 2.12 Cumulative grid-connected photovoltaic installations in Korea by policy, 2004–2008. Data sources: [12, 88–91]

end of 2008, approximately 300 MW_p of systems were installed under this scheme, 90% of which are larger than 100 kW_p in capacity, reflecting the very high percentage of residential, public and commercial buildings that are more than ten stories [12]. Beginning in 2012, the government intends to replace its feed-in tariff system with renewable portfolio standard scheme [90] (Figure 2.12).

2.3 POLICY IMPACT ON PV MARKET DEVELOPMENT

The US, Germany, Spain, Japan, and South Korea are the leading markets for grid-connected PV. By the end of 2008, the combined PV installations in these five countries stood at 12.4 GW_p, representing more than 90% of cumulative PV installations in the OECD bloc [12]. Review of PV deployment experience in these countries underscores the significance of policy in market development, market growth, and technology diffusion. It also enables us to identify policy effectiveness by the types of tools employed by this group.

Initial drivers of PV deployment in the US, Germany, Japan and South Korea were programs targeting small-scale installations in the residential sector. Germany and South Korea supported PV through their solar rooftop programs and Japan used its Residential PV System Dissemination Program to promote this sector's use of the technology. Likewise, California implemented its Emerging Renewables (CEC-ER) program, which supported PV installations in the residential sector. In order to achieve modest market start-up, these programs found that subsidies of 50% or higher of initial system costs were needed. Despite major differences in housing stock, these programs were able to launch small PV markets in the residential sector.

After initial success of PV deployment programs targeting small-scale residential applications, Japan, South Korea and the US supported programs with similar designs for their commercial

Table 2.3 Cumulative R&D spending in selected countries (in millions of US$ 2008)

	1975	1980	1985	1990	1995	2000	2005	2008
Japan	12	49	403	726	1029	1459	2095	2268
Germany	42	278	633	1006	1405	1680	1909	2073
US	8	812	1568	1872	2332	2739	3160	3469

Data sources: US data for 1991–2008 was provided by Robert Margolis from NREL. Data for Japan and Germany, for 1975–2008 and US for 1975–1990 was obtained from the IEA Database [92].

and public buildings sectors. Here again, market development was modest, but steady. An important contribution of these programs was the confidence they built in the technology.

The policy history of these countries shows a marked shift in the last 10 years away from investment incentives and toward production-based approaches. Feed-in rates and tradable solar renewable energy certificate (SREC) schemes are the preferred policy tools in this period. Germany, Spain and South Korea have implemented feed-in laws, which have triggered significant market growth. Amendments to national FiT schemes have ensured a wide array of market applications (commercial, industrial, and public as well as residential uses) and technology configurations (e.g. thin film as well as silicon; ground-mounted as well as roof-mounted). In the US, New Jersey has shown an ability to replicate the fast growth of FiT strategies with its vibrant SREC market approach, and California has used production-based Incentives to quickly grow its Go Solar Initiative. It can be expected that production based incentives will gain more prominence through SREC trading in the US and other countries (e.g. South Korea is planning to implement an RPS scheme in 2012), and through the use of solar carve-outs in national target-setting for renewable energy use.

Financial assistance of PV deployment projects is crucial for wider adoption of this technology. However, it is also important to have strong government support for research and development (R&D). Indeed, the governments of Germany, Japan and the US have provided significant R&D funding over the last 25 years (Table 2.3).

R&D expenditures have enabled countries to spur improvements in technology performance and, thereby, reduce user costs. As we will show in the following section, this factor is very important for the goal of building a policy strategy for a *long-term* sustainable PV market.

Direct government incentives, whether capital or performance based, combined with R&D funding have played a major role in PV cost reduction. An Important analytical question is to what extent direct government incentives, combined with R&D funding or other policy tools, such as carbon taxes, can impact future adoption of PV. In the following section, we describe the methodology for conducting such an analysis, and then we show concretely how different policy tools affect PV diffusion paths over the short and long term.

2.4 FUTURE PV MARKET GROWTH SCENARIOS

2.4.1 Diffusion Curves

The evolution of technology has been the focus of research for a long time. A prominent view holds that technology evolves in three phases: invention, innovation and diffusion [93]. Invention refers to the initial development of a scientifically or technologically innovative process or product while innovation refers to the point when the new product or process reaches the market. Diffusion

is the final stage in this evolution, and is the focus of this section. It refers to the process of dissemination through which successful innovations come to be widely available through the adoption by individuals and/or firms (Schumpeter, 1942, quoted in [93]). While the three stages of invention, innovation and diffusion are described sequentially, in actual practice there is a cyclical relationship between them. In this regard the role of feedback arising from diffusion, including second- or third-generation invention and innovation is important [93].

An interesting aspect of diffusion of innovation is the fact that not all potential buyers make the decision to invest in a product or process at the same time. Adopters can be categorized into five types based on personality and behavior, values and attitudes [94–96]. They include "innovators" who constitute about 2.5% of adopters; "early adopters" who constitute about 12.5–13.5% of adopters; "early majority" who constitute about 34–35% of adopters; "later majority" who constitute 34–35%; followed finally by the "laggards" who make up the remaining 15–16% of adopters [94–96].

Rogers [94] also proposes general characteristics for each adopter category, based on socioeconomic, personality and communications behavior. For instance, the "innovators" and "early adopters" tend to display characteristics such as more years of education and greater knowledge of the technology. This view has been modified by the argument that adopters who are "innovators" for one product, could be "laggards" for another product. This point underscores the importance of compatibility of a product, for instance photovoltaics, with the lifestyles, attitudes and values of potential adopters.

A useful addition to the diffusion of innovations theory is the idea of a "chasm" between the "early adopters" and the "early majority". The entry of the early majority in the market is critical to the commercial viability of a product or service. Unlike the "innovators" and "early adopters" this "early majority" is unlikely to take the long-term view and put up with inconveniences and product complexity. The incorporation of innovations, product enhancements and other "attractive" features, often based on feedback from early adopters is required to win this segment over [97].

Nature, markets and technologies experience growth patterns which are usually confined by some limits. These limits could be the size of the potential market, as in the case of technological innovations, or an ecosystem's carrying capacity, as in the case of animal and plant populations. The graphical representation of this type of growth resembles an S-shaped curve [98]. The diffusion of innovations, i.e. growth in the market for innovations, such as photovoltaics, computers or cellular phones, has also been found to follow an S-shaped or logistic growth curve [99, 100].

Logistic growth models have proven to be accurate tools for forecasting a wide range of phenomena, from human population growth (used by the Belgian mathematician Pierre Verhulst in 1838) to oil development [101, 102]. Often, technologies (e.g. computers or cell phones) grow exponentially during an initial phase. However, as a device eventually reaches saturation in the potential market, the rate of growth is seen to slow down and finally taper off. This methodology is commonly used to anticipate the entry of new technology [98, 103, 104], including new energy technologies [105, 106] (Figure 2.13).

A logistic growth curve, according to Laherrère [102] and Meyer *et al*. [98], can be represented by the following equation:

$$Q_t = \frac{U}{1 + e^{-b*(t-t_m)}} \tag{2.1}$$

where Q_t is the forecast variable (e.g., percentage of electricity supply by PV in a given year), U is the saturation (maximum) level for Q_t (e.g. maximum percentage of electricity supply assumed to feasibly come from PV), b is the slope term, reflecting an initial growth rate, t is a time variable (in years), and t_m represents the midpoint of the logistic curve.

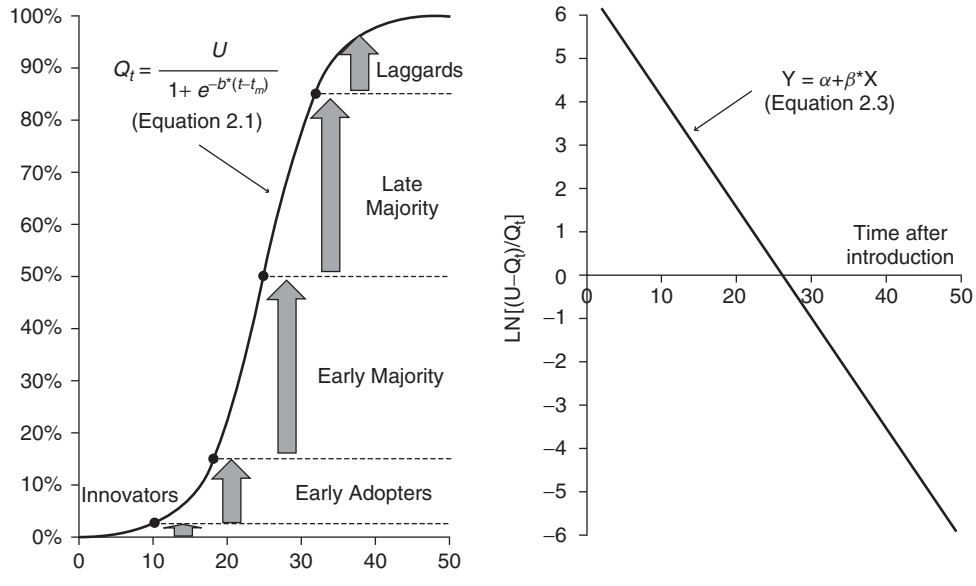


Figure 2.13 Representation of technology diffusion by an S-shaped curve and its linear form

Rearranging terms in Equation (2.1) and taking the logarithm of both sides gives:

$$\ln\left(\frac{U-Q_t}{Q_t}\right) = -b*t + b*t_m \tag{2.2}$$

Grouping variables, we then obtain:

$$Y = \ln\left(\frac{U-Q_t}{Q_t}\right)$$

This yields the familiar linear equation:

$$Y = \alpha + \beta * X \tag{2.3}$$

where $X = t$, parameter $\alpha = b^*t_m$ and parameter $\beta = -b$.

Applying statistical regression methods to Equation (2.3), the parameters α and β can be robustly estimated.

Noting that $t_m = -\alpha/\beta$ and $b = -\beta$, Equation (2.1) can be presented as:

$$Q_t = \frac{U}{1+e^{\beta\left(t+\frac{\alpha}{\beta}\right)}} \tag{2.4}$$

Equation (2.4) and the linear regression method used to estimate parameters in Equation (2.3) are consistent with the classic Fisher–Pry form of a logistic growth curve widely used to model technology diffusion [103]. In this way, a forecasting model can be built on available empirical experience to date for the technology of interest (PV, in this case). A key factor in the diffusion of new technology is cost-competiveness with its alternatives. At the initial phase, technology diffusion can be supported through government programs and incentives. However, for wide-scale adoption the technology should have an advantage over other alternatives and at least be cost-competitive.

The link between technology diffusion and technology cost trends can be characterized through experience curves described in the next section.

2.4.2 Experience Curves

Experience curves, also referred to as learning curves, describe the link between long-term cost trends and adoption rates for new technologies. In 1936, Wright [107] was the first to provide a mathematical representation of the experience curve [108, 109]. Since then experience curves have become a helpful tool for analysts to assess trends in the cost-competitiveness of different technologies [110–115].

Experience curves are typically used for long-term strategic rather than short-term tactical analysis. But in the formulation of competitive strategies, experience curves can be powerful instruments to model market development of innovations [114]. According to Neij [110], experience curves offer a means of projecting future cost trends based on past cost developments.

Figure 2.14 provides a schematic representation of an experience or a learning curve for PV on double-logarithmic scales. On the horizontal axis is the cumulative installation of PV systems; on the vertical axis is system price per W_p.[19] As cumulative installations of PV systems grow, so do the producers' and installers' experiences, leading to reductions in manufacturing and deployment costs. Mathematical representation of this relationship can be expressed as [112]:

$$C_t = C_o * \left(\frac{n_t}{n_o}\right)^{\beta} \tag{2.5}$$

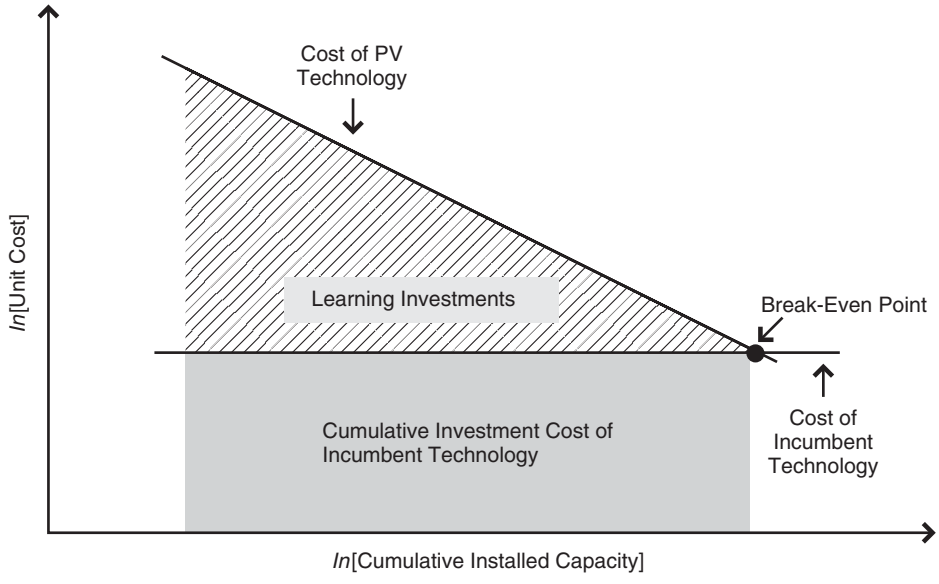


Figure 2.14 Schematic representation of learning curves and learning investments. Adapted from [116]

[19] System price is the installed cost of a system including the PV device, balance of system (e.g. inverters, wiring, and panel array structure) and labor and other installation costs, as well as rates of return to the manufacturers and installers.

where C_t represents the expected cost at an n_t cumulative production level at some time in the future. C_o is the known cost of the product or installation at the initial phase of product deployment. Typically, C_o is calculated when cumulative production $n_o = 1$ (e.g. 1 MW_p or 1 GW_p). The exponent β is very important in characterizing the rate of price decrease as discussed below. In logarithmic form, Equation (2.5) can be written as:

$$\ln(C_t) = \ln(C_o) + \beta * \ln(n_t) \tag{2.6}$$

This yields the familiar linear equation:

$$Y = \alpha + \beta * X \tag{2.7}$$

where $X = \ln(n_t)$ and parameter $\alpha = \ln(C_0)$.

If data points for the experience curve are known, the parameters for the underling Equation (2.7) can be obtained through linear regression analysis. Equation (2.6) shows that, in logarithmic form, the change in the cost per unit is directly proportional to change in cumulative output.

There are two important metrics devised to parameterize the information contained in an experience curve and to apply it for analysis: the progress ratio (PR) and the learning rate (LR).

Comparison of different experience curves can be made by determining the change in price, when cumulative production volume doubles. The corresponding change in price gives the progress ratio. Thus, if the cost per unit reduces to 0.80 of the original price by doubling the cumulative output, then the progress ratio of such a technology is 80%. The learning rate for a particular technology is derived from the progress ratio by subtracting it from 100%. Thus, if the progress ratio is 80%, the corresponding learning rate for the technology is 20%. Progress ratios and learning rates can be obtained through the following equations:

$$PR = 2^{\beta} \tag{2.8}$$

$$LR = 1 - PR \tag{2.9}$$

where β is the slope parameter that can be obtained through the regression (Equations 2.6 and 2.7).

It is important to provide policy support until a technology becomes cost-competitive with alternative sources. The point at which technology becomes cost-competitive is referred to as the break-even point (Figure 2.14). Experience curve analysis can show the level of investments required to make a technology market competitive. However, experience curves do not forecast when, in time, this break-even point would be reached. Even so, experience curves can be an effective tool for energy policy makers to set targets and implement measures to enable new technologies to become economically viable.

The level of required investments needed to reach a break-even price can be calculated by integration under the learning curve (Figure 2.14). Learning investments (LI) can be calculated as follows (adapted from Zwaan and Rabl [112]):

$$LI = \int_{n_c}^{n_b} (C_c - C_b) * dn = \frac{C_c}{\beta + 1} * \frac{n_b^{\beta+1} - n_c^{\beta+1}}{n_c^{\beta}} - (n_b - n_c) * C_b \tag{2.10}$$

where β is a slope parameter derived from equation (2.7), C_c is the technology's current cost, C_b is its cost at a break-even point, n_c is the current cumulative production level, and n_b is the cumulative production at a break-even point, which can be derived from the following equation:

$$n_b = n_c * \left(\frac{C_b}{C_c}\right)^{1/\beta} \tag{2.11}$$

Recently, several researchers have proposed to extend the simple formulation of learning curves (described above), also referred to as a single-factor learning curve, to two-factor learning curves (2FLC) [117–120]. The 2FLC model provides the added ability to measure the impact of research and development (R&D) activities on technology cost reduction. The two-factor learning curve can be expressed as:

$$\ln(C_t) = \ln(C_o) + \beta * \ln(n_t) + \gamma * \ln(K_t) \tag{2.12}$$

where K_t represents the stock of knowledge in time period t acquired due to past investments. The knowledge stock is defined as a function of past R&D investments that includes depreciation and time lag factors. The knowledge stock can be expressed as [118, 119]:

$$K_t = K_{t-1} * (1 - \rho) + ARD_{t-i} \tag{2.13}$$

where ρ is the annual knowledge stock depreciation rate, ARD is the annual expenditure in R&D, and $\acute{\imath}$ is the time lag between R&D investment and its effect. For PV technology, a typical value used for the annual knowledge stock depreciation rate ρ is 3%, and for the R&D time lag $\acute{\imath}$, researchers use two to three years [118, 121, 122].

The 2FLC results in two learning rates. The first is the learning-by-doing rate (LDR), representing experience gained through increasing scale of production and deployment and its impact on cost (analogous to an LR for single-factor learning curves). The second is the learning-by-searching rate (LSR), representing the impact of increased knowledge, obtained through R&D, on system cost. LDR and LSR can be represented as follows:

$$LDR = 1 - 2^{\beta} \tag{2.14}$$

$$LDR = 1 - 2^{\gamma} \tag{2.15}$$

where β is the learning-by-doing index and γ and is the learning-by-searching index.

The major problem with the 2FLC model is that its independent variables are highly correlated (i.e. high multicollinearity between cumulative installations and R&D knowledge stock). A common solution is to use a predefined value for the knowledge stock index (i.e. γ) and estimate the learning-by-doing index (i.e. β) by regression analysis. In our modeling, a value of to $\gamma = 0.154$ is assumed (based on [118, 120]). Our findings using a 2FLC model are reported below.

2.4.3 PV Diffusion in the US under Different Policy Scenarios

As discussed earlier, the US PV Market in recent years has experienced very rapid growth. PV cumulative installations have increased from 43.5 MW_p in 1992 to 1168 MW_p by the end of 2008. However, even with such rapid growth, PV's share in the US electricity supply in 2008 was only 0.04% [12, 123]. Nevertheless, if the current trend continues, the share of PV will likewise increase. Below, we model this process and estimate the rising share of US electricity supply from solar electric power under three policy scenarios: (1) a national carbon tax cap-and-trade policy; (2) a national renewable portfolio standard; and (3) an expanded national commitment to R&D to promote higher-efficiency, lower-cost PV modules.

2.4.3.1 Building a business-as-usual benchmark for US PV market development

For our diffusion analysis, under different policy scenarios, we have assumed 25% as a reasonable target for the maximum share of grid electricity supply provided by PV in the next few decades

(see, e.g. [124]). A standard diffusion model based on values obtained from Equation (2.2) above (i.e. values of

$$\left[\ln\left(\frac{U - Q_t}{Q_t}\right)\right]$$

plotted against a time variable) is used, which adopts empirical estimates from regression analysis of diffusion rates for the period 1990–2008 (Figure 2.15). From the figure, it is evident that the regression trend line has changed with PV diffusion accelerating after 2000. The β value (i.e. diffusion rate) for the initial period of 1992–2000 is −0.1172, but it more than doubles (−0.2543) for the later period of 2001–2008.

Between 2000 and 2008, PV's share of US electricity supply increased by approximately 30% per year. Based on this historical trend, we built a diffusion model to estimate future PV supply, when saturation is assumed to occur at 25% of total electricity generation.[20] Additionally, it was assumed that electricity generation in the US will increase by 1% per year. This assumption mirrors that of the US EIA for the period between 2010 and 2030 [8], and our analysis extended this assumption after 2030 until 25% saturation is reached under each policy scenario, when necessary.

Utilizing EIA projected total electricity supply for 2010–2030 and assuming a 1% growth rate thereafter, estimates of total and PV-sourced electricity generation were obtained (Figure 2.16 and Table 2.4). The projected path of PV supply in Figure 2.16 is treated as the business-as-usual scenario (BAU). The BAU projects PV capacity to increase from 1.8 GW_p in 2010 to 1076 GW_p in 2055; correspondingly PV's share of US electricity supply grows from 0.07% to 25% during the same period.

The BAU projections derived from our analysis are well within reach, particularly when we compare them with targets established by current state level RPS policies for PV and customer-sited

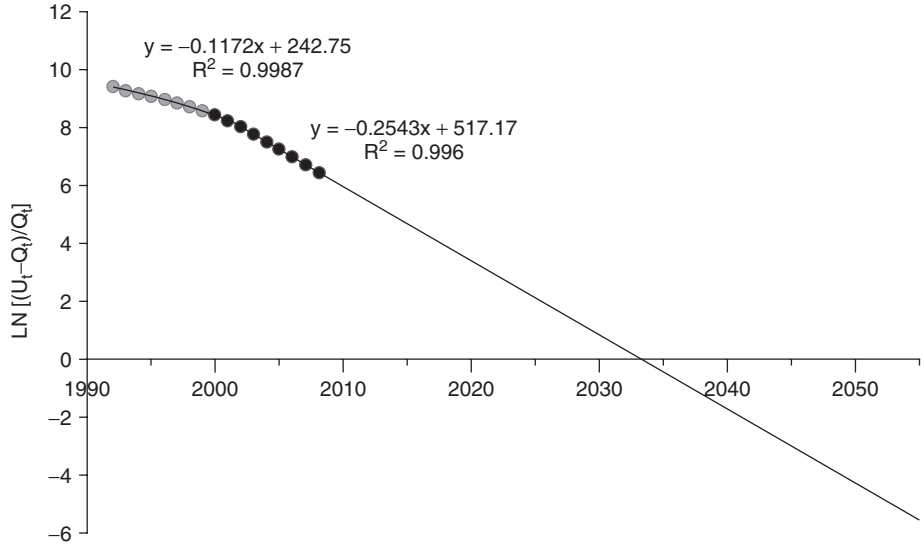


Figure 2.15 Regression analysis using a logistic growth model for US PV installation. Data sources: [12, 123]

[20] This saturation rate is based on research suggesting that the integration of intermittent resource into grid supply has a technical limit roughly at this rate (e.g. [124])

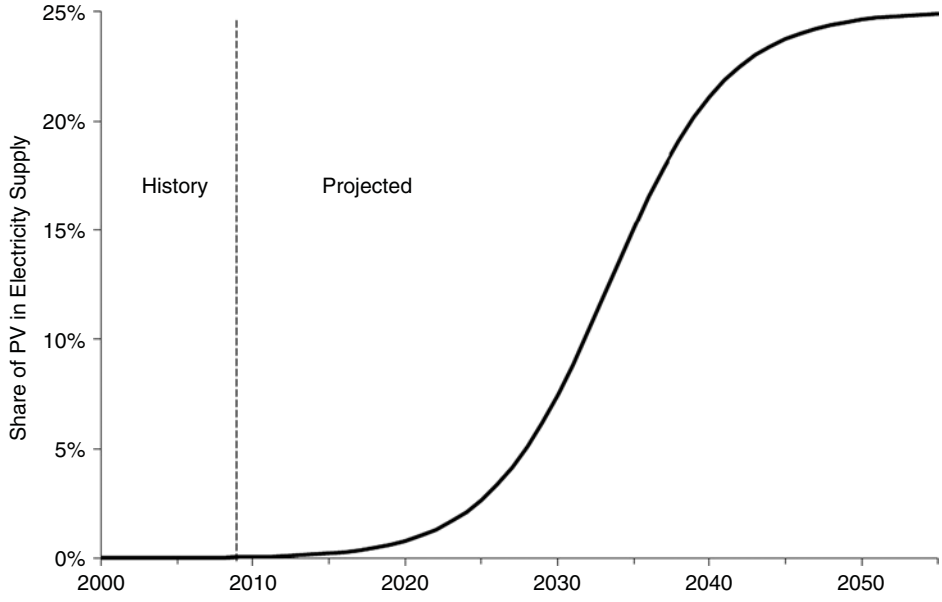


Figure 2.16 PV diffusion in the US under a business-as-usual scenario. This projection is based on the regression analysis shown in Figure 2.15, which assumes $U = 25\%$, and gives regression parameters of $\beta = -0.2543$, and $t_m = 2034$

Table 2.4 BAU projected PV generation and installed capacity

	US electricity generation* (billion kW h)	PV's share of US electricity supply (%)	PV generation (billion kW h)	PV capacity GW_p**
2010	4162	0.07	2.72	1.8
2015	4339	0.23	10.07	6.7
2020	4573	0.81	37.02	24.7
2025	4840	2.67	129.16	86.1
2030	5055	7.48	377.90	251.9
2035	5312	15.09	801.79	534.5
2040	5583	21.12	1179.09	786.1
2045	5868	23.78	1395.21	930.1
2050	6166	24.64	1519.96	1 013.3
2055	6482	24.90	1614.01	1 076.0

*2010–2030 numbers are from [8]; for 2030–2055, 1% of annual growth is assumed.
**Assumes annual energy production of 1500 kW h per kW_p for PV.

distributed energy sources. As noted earlier, 14 states and the District of Columbia (Washington, DC) have specific solar or distributed generation carve-outs in their RPS requirements. California, Oregon and Texas have specific targets for distributed generation or PV [25]. Applying current PV/DG RPS requirements to these states' electricity consumption (based on 1% annual increases in electricity demand), and using state-specific solar radiation, we determined the PV installation capacity needed to meet the RPS carve-outs (Table 2.5). Results show that states with PV/DG RPS will, in total, require 3.5 GW_p in 2015 and 11.8 GW_p in 2020 to meet their legislated targets. These

Table 2.5 State PV and distributed generation carve-out

	Solar carve-out				Projected installations (MW)			
	2010	2015	2020	2025	2010	2015	2020	2025**
AZ*	0.50%	1.50%	3.00%	4.50%	143	452	951	1499
CA							3500	3500
CO	0.20%	0.60%	0.80%	0.80%	70	219	308	323
DC	0.00%	0.20%	0.40%	0.40%	3	16	40	42
DE	0.00%	0.60%	2.00%	2.00%	2	54	204	214
IL	0.00%	0.60%	1.10%	1.50%	0	716	1317	1978
MD	0.00%	0.30%	1.50%	2.00%	11	115	724	1014
MO	0.00%	0.10%	0.20%	0.30%	0	66	138	218
NC	0.00%	0.10%	0.20%	0.20%	23	168	254	265
NH	0.00%	0.30%	0.30%	0.30%	4	30	32	34
NJ+	0.22%	965 GW h	2164 GW h	4610 GW h	137	742	1664	3545
NM	0.20%	0.60%	4.00%	4.00%	27	87	607	638
NV	0.60%	1.20%	1.30%	1.50%	131	276	319	381
NY*	0.10%	0.10%	0.10%	0.10%	108	176	185	195
OH	0.00%	0.20%	0.30%	0.50%	12	187	445	688
OR							20	20
PA	0.00%	0.10%	0.40%	0.50%	15	188	610	723
TX*							500	500
Total					686	3492	11 818	15 777

*Distributed generation (DG).
**Most of the RPS carve-outs have targets which must be reached by 2020.
+In January 2010, New Jersey changed its percentage-based solar target to a GW h target.
Data sources: [8, 25].
The analysis, based on EIA's most recent electricity forecast (8), assumes electricity consumption will grow 1% annually. The PV capacity required to meet solar or DG carve-outs is calculated using average daily solar radiation data from NREL [125], and a PV system performance ratio of 75%.

are nearly half (48–52%) of what is projected under our BAU scenario (i.e. 6.7 GW_p in 2015 and 24.7 GW_p in 2020). The selected 17 states (Table 2.5) and Washington DC represent one-half of US electricity demand (49%). Thus, if the remaining states follow similar policy initiatives and/or if the pioneering 18 US jurisdictions upgrade their targets while the bulk of the states without RPS rules adopt a policy strategy that approximates the goals of the early adopters, the BAU targets in Table 2.4 will be readily achievable. Moreover, in the BAU scenario, the projected PV share is 0.8% of total US electricity supply by 2020, which is not overly aggressive if we consider the current installed capacity of PV in other OECD countries.[21] Germany and Spain, for example, already reached this level of PV adoption in 2008 [12].

A key factor in the rate of PV deployment is the cost of PV electricity relative to other generating fuels and technologies. Historically, the levelized cost of electricity (LCOE) of conventional electricity supply sources has been significantly lower than PV, even after accounting for transmission and distribution costs.[22] For substantial penetration of grid-connected PV, the LCOE

[21] The BAU scenario presented here, assumes maximum level of PV share electricity supply at 25%. This level is projected to be reached in 2055.
[22] LCOE provides the means for economic evaluation and comparison of different electricity generation technologies. LCOE accounts for all costs over a technology's lifetime, including initial investment, operations and maintenance, cost of fuel, and cost of capital.

for PV production needs to decrease and/or non-PV LCOEs must increase. LCOE can be calculated as follows [112, 126, 127]:

$$LCOE = IC * \left(r_{O\&M} + \frac{r_{\text{int}}}{1 - (1 + r_{\text{int}})^{-n}} \right) \tag{2.16}$$

In this equation IC is initial capital cost including installation, $r_{O\&M}$ is the annual operation and maintenance cost (O&M) as a percentage of IC, r_{int} is the real interest rate, and n is the economic system lifetime in years. In this formulation, fuel costs appear in $r_{O\&M}$.

In 2009, the capacity-weighted average installed system cost of a PV system was \$6.80 per W_p [30]. According to a Deutsche Bank report, in the US, the financing rate associated with project development ranged between 6 and 8% [19]. For the annual inflation rate for 2010–2030, we utilized the EIA reference case [8], which assumes 2% per year for this timeframe, resulting in a real interest rate for PV of 5%.[23] Using a typical PV generation of 1500 kW h per kW_p for the US, with a project lifetime of 25 years, an installed cost of \$6.80 per W_p, annual O&M costs at 1% of initial installed cost, and a 30% federal tax credit, we estimated the LCOE for PV at 25.7 cents per kW h.[24] At the same time, the recent average weighted retail electricity price for commercial and residential customers in the US is 10.6 cents per kW h [123]. EIA projects no significant real price (adjusted for inflation) increases in their projections of US electricity retail prices [8]. Using this very conservative assumption, a \$2 per W_p PV system cost is required to be competitive with conventional grid-supplied electricity in the residential and commercial sectors. This \$2 per W_p system cost represents the PV break-even price.

The experience curve, as discussed above, offers a useful method for assessing cost-cutting knowledge gained through increasing scale of PV production and deployment. The cost reductions associated with this experience is expressed by a learning rate. Based on data obtained from NREL on the annual average installed system cost for 1998–2005 and cumulative installed capacity [30], an experience curve for the US PV system cost is derived. In the analysis, recent years affected by polysilicon shortages are excluded.[25] The resulting curve presented in double-logarithmic form is shown in Figure 2.17.

The slope parameter (β in Equations 2.5–2.8) for the experience curve displayed in Figure 2.17 has the value -0.214. Using Equations (2.8) and (2.9), the US PV system cost trend exhibits an 86.2% progress ratio and a 13.8% learning rate. Based on this learning rate, a break-even price of \$2 per W_p system cost will be reached when cumulative installations equal 280 GW_p. Under the BAU scenario, this level of PV installation in the US would be met by 2031.[26] As shown in Figure 2.17, before PV installations reach a break-even level, the cost differential between system cost and break-even price needs to be subsidized. The total subsidy required to reach the break-even point is illustrated as the shaded area, labeled learning investments.[27] Government

[23] 5% was obtained using 7% as a midpoint of Deutsche Bank's financing values, adjusted for 2% inflation (i.e. $[1.07/1.02 - 1] \approx 0.05$).

[24] This rate does not include other federal (e.g. MACRS) and local incentives. If MACRS depreciation rules are included, PV's LCOE in the BAU case falls to 14.5 cents per kW h. Inclusion of tax benefits is justified on the ground that all power plants and non-renewable fuels in the US have been subsidized by tax and other policy treatments. It should be noted that, typically, LCOEs for fossil fuel power plant include the MACRS tax benefit.

[25] During 2006–2008, silicon PV module prices actually increased the first time in 30 years. Then in 2009, they fell rapidly as new production capacity made it to market. For a discussion of this situation see Chapter 1, Section 1.2.3 in this *Handbook*.

[26] As Figure 2.16 reports, once a break-even price is reached in 2031, PV market share will grow and realize the target of 25% of total US electricity assumption in just 24 years (i.e. the rapid-growth interval of 2031–2055 in Figure 2.16).

[27] These investments are of the learning-by-doing variety.

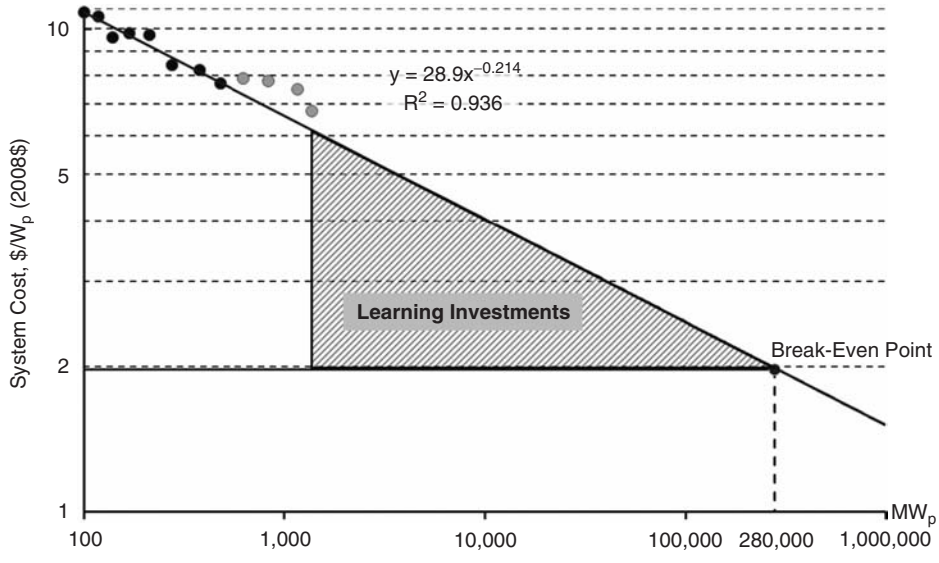


Figure 2.17 US PV system cost experience curve. Data sources: [30, 12]. The curve was derived based on system cost data for 1998–2005 and cumulative installations in the US Data from recent years affected by polysilicon shortages ware excluded

programs and incentives can increase the rate of PV production and deployment, driving PV system costs down the experience curve. The level of subsidy required each year to meet targets outlined under the BAU diffusion path in Figure 2.16 are calculated based on Equation (2.10).

The results of this calculation are displayed in Figure 2.18, and show that the level of subsidies to meet BAU annual installation targets need to increase between 2010 and 2025,

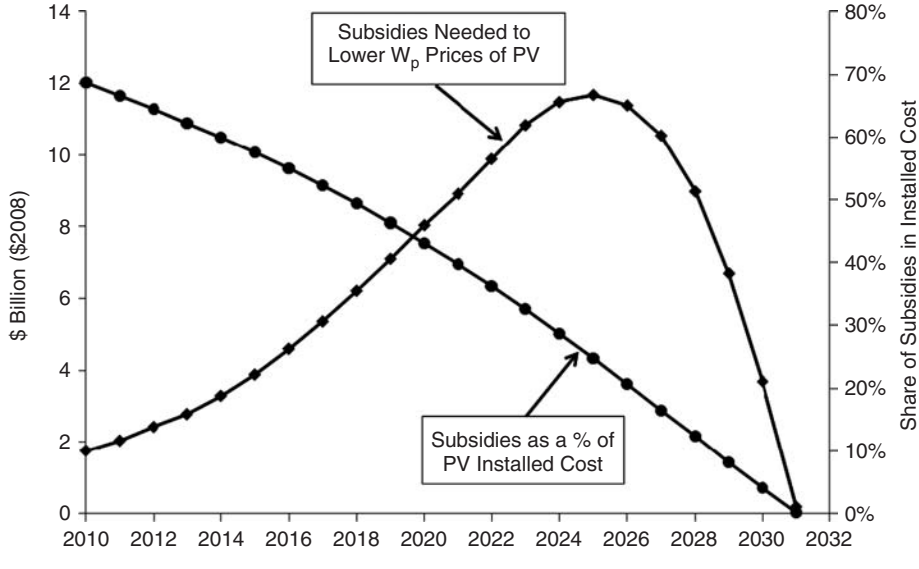


Figure 2.18 PV system cost subsidies required to follow BAU diffusion scenario

Table 2.6 Learning parameter estimates

	Index	Progress ratio (%)	Learning rate (%)
Learning-by-doing, estimated	−0.2028	86.9	13.1
Learning-by-searching, fixed	−0.1520	90.0	10.0

Annual R&D expenditure data for 1991–2009, used for these estimates was provided by Robert Margolis (NREL). Additional data for 1975–1990 was obtained from IEA Database [92]. The fixed learning-by-searching rate is taken from research Miketa and Schrattenholzer [118] and Kouvaritakis *et al*. [120].

with a subsequent decline thereafter. The share of subsidy in total PV investments gradually decreases from approximately 70% of total installed cost in 2010, until it reaches zero by 2031. This translates into a total learning investment of $142 billion (in 2008 dollars). However, this number does not account for cost reductions associated with research and development (R&D) expenditures. This is discussed below.

To estimate the impact of R&D, a two-factor learning curve (2FLC) is developed. Following the methodology described in the previous section (see Equations 2.12–2.15), parameters for the 2FLC model were obtained. The results are presented in Table 2.6. The estimated "learning-by-doing" parameters reflect cost reductions due to manufacturing and other experience gained through PV system deployment. "Learning-by-searching" indicators represent cost reductions associated with R&D expenditures. After factoring in the impact of R&D, the learning rate attributed to experience gained through PV system deployment and associated cost reductions is reduced from 13.8 to 13.1%.

In order to assess the level of annual R&D investments required to continue the same trend of cost reduction in Figure 2.17, the single-factor and two-factor learning curves must be harmonized so that system costs in Equations (2.6) and (2.12) are equal for the same cumulative installation level n_t. Our method to achieve this is now presented. Equations (2.17) and (2.18) modify the original forms in order to reflect different intercept and slope parameters for learning-by-doing and learning-by-searching.

$$\ln(C_t) = \ln(C_{o1}) + \beta_1 * \ln(n_t) \tag{2.17}$$

$$\ln(C_t) = \ln(C_{o2}) + \beta_2 * \ln(n_t) + \gamma * \ln(K_t) \tag{2.18}$$

After combining these equations and rearranging terms, the stock of knowledge variable K_t introduced in Equation (2.12) can be derived as:

$$K_t = n_t^{(\beta_1-\beta_2)/\gamma} * \exp\left(\frac{\ln(C_{o1}) - \ln(C_{o2})}{\gamma}\right) \tag{2.19}$$

After combining Equations (2.13) and (2.19), required annual R&D expenditures can be determined. An average of $122 million per year in R&D is needed to meet targets outlined under the BAU diffusion path. Thus, between 2010 and 2031, in addition to $142 billion required as learning investments, an estimated $2.7 billion in R&D is projected to be needed to reach the break-even point.

2.4.3.2 US PV policy scenarios

In addition to tax subsidies and R&D programs that have traditionally supported PV deployment, a new array of policy mechanisms in support of a US transition to a "green energy" economy are

being discussed. This section evaluates three policy tools to ascertain their likely impact on the diffusion of PV into the national electricity market.

2.4.3.2.1 Pricing carbon

The first policy tool we examine is a carbon tax or carbon cap-and-trade strategy.[28] Due to the dominance of fossil fuels in electricity generation, the introduction of a carbon pricing scheme will increase the cost of grid supplied electricity. This in itself will raise the necessary break-even price for PV, at which the technology is cost-competitive with grid-supplied electricity. An effective carbon tax at \$25 per ton of CO_2 will increase the break-even system cost by \$0.32/W_p and a carbon tax at \$50 per ton of CO_2 will increase the break-even system cost by \$0.64/W_p (see Figure 2.19).[29]

A carbon price of \$25 per ton is at the high end of the trading value in the EU [128]. The US is struggling to pass cap-and-trade legislature that would likely result in carbon prices less than \$25 per ton [129]. An additional scenario using a high price of \$50 per ton is included to capture what currently appears to be the outer reach of political possibility with regard to this tool.

Increasing the break-even cost will reduce the amount of cumulative PV system installations required to reach the break-even point. For the case of a \$25 per ton cost of CO_2 emissions, a break-even point will be reached at 139 GW_p, and for a \$50 per ton cost of CO_2 emissions a break-even point at 76 GW_p is needed (Figure 2.19).

The amount of required learning investments will be reduced from \$142.2 billion to \$47.0 billion under a \$50 per ton of CO_2 cost scenario and to \$79.5 billion under a \$25 per ton of CO_2

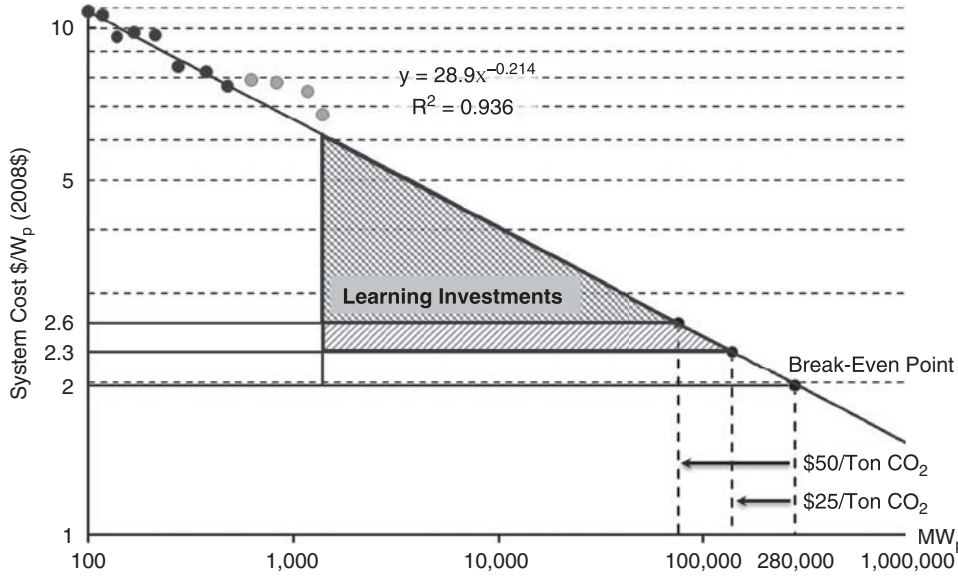


Figure 2.19 Impact of a carbon pricing policy on the break-even point for PV

[28] Although important differences exist between a carbon tax and a cap-and-trade strategy in terms of implementation, we focus here only on the effect on retail electricity prices. For this reason, we do not distinguish between the two in our analysis.

[29] For Figure 2.19, we assumed 0.6 tons of CO_2 is emitted per MW h of electricity generation, PV generation at 1,500 kW h per installed kW_p, and 25 year product life. A real 5% discount rate was assumed.

cost scenario. In this respect, if carbon taxes or cap-and-trade allowances collected from fossil fuel generators are directed to provide additional support for non-fossil fuel generation technologies, then PV deployment projects will not only gain from a higher break-even price, but would also benefit from increased public investment. For example, if PV generators, before reaching the break-even point, are paid \$50 or \$25 per ton for avoided CO_2 emissions, then additional monetary benefits to PV project development can be quantified at \$48 billion or \$44 billion, respectively.[30]

We can now estimate the impact of the two carbon pricing scenarios on PV diffusion. For a carbon price of \$25/ton, and including the price effect on conventional grid power and the effect of using the proceeds of carbon pricing to incentivize PV use, we project grid parity to occur in 2024 and 25% saturation to be reached in 2050. For \$50/ton, we project PV to reach grid parity in 2020 and to realize 25% saturation in 2045.

2.4.3.2.2 The impact of PV R&D

We now turn our attention to the impact of R&D policy on PV diffusion. What would be the likely impact of increasing public investment in R&D? To answer this question we analyzed the R&D scenario under which R&D investments are committed at the same level as learning-by-doing investments (i.e, rebates, tax credits and other project level incentives), shown in Figure 2.18. Increased investment in R&D facilitates faster decline of PV system costs. We modeled the following scenario: conventional policy incentives for PV development were decreased by approximately \$25 billion over 8-year period; instead \$25 billion was dedicated to public R&D; this increased R&D during 8-year period by roughly factor of ten – that is from \$2.7 billion to \$27.9 billion. As is demonstrated in Figure 2.20, increase of R&D spending tenfold raises the slope on the US

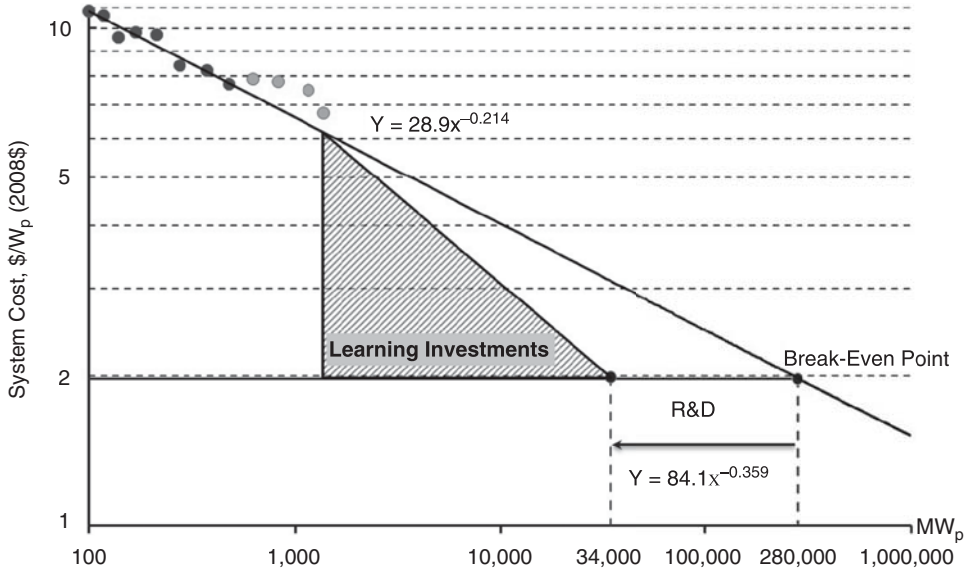


Figure 2.20 Impact of increased R&D expenditures on the break-even point for PV

[30] These numbers are based on the assumptions in footnote 29, and using the products of previously obtained numbers for carbon tax impact on break-even price and amount of additional cumulative PV installations required to reach a break-even point (i.e., \$0.64/W_p*[76 – 1.3], and \$0.32/W_p*[139 – 1.3]).

Table 2.7 Effect of R&D on total investments for reaching a PV break-even point, \$2 per W_p

	BAU scenario	Increased R&D scenario
R&D Investments (billion \$)	2.7	27.3
Learning-by-doing Investments (billion \$)*	142.2	27.3
Combined Investments (billion \$)	144.9	54.6

*Learning-by-doing investments equal the total policy incentives (rebates, tax credits, etc. see Figure 2.18)

experience curve from −0.214 to −0.359, thus, increasing the learning rate from 14 to 22% (see Equations 2.8 and 2.9).

Based on the results obtained, which are shown in Figure 20 and Table 2.7, it can be seen that Increases in R&D expenditures can significantly reduce the amount of combined (learning and R&D) investments needed to reach a break-even point.

Using these learning rates, we calculated that tenfold increase of US public R&D for PV would accelerate the year when grid parity would be reached to 2018 (from the BAU of 2031). It would also shorten the time to 25% saturation from a BAU of year 2055 to 2040.

2.4.3.2.3 Effect of a solar carve-out policy

Another policy option for rapid deployment of PV is the promotion of a solar carve-out under a national RPS, similar to those implemented in a number of US states (see Table 2.5). Under a Solar RPS requirement, energy suppliers need to create SRECs through their own power generation, or purchase them from third-party solar power providers. A well-structured RPS with sufficiently high noncompliance fees and a legal requirement for utilities to purchase SRECs through multi-year contract can result in rapid PV deployment similar to the German experience with feed-in tariffs. An RPS with a solar carve-out could create significant demand for Solar Renewable Energy Certificates (SRECs). By 2010, in states with established SREC markets, SRECs ranged from \$200 to \$700 per MW h [51]. For a national solar market, more conservative prices for SRECs were considered: SRECs are initially traded at \$200 per MW h (20 cents/kW h) in 2010 and gradually decrease to \$100 per MWh (10 cents/kW h) by 2020.

This scenario relies on an SREC pricing schedule that is below those in play in 14 pioneer states. Indeed, recent policy reforms in Delaware, Pennsylvania, Massachusetts, New Jersey and other states have led to SREC trading above \$200 per MW h [51]. In this sense, the carve-out scenario analyzed in Figure 2.21 is quite conservative. Under this scenario, the break-even point will be achieved at 66 GW_p of cumulative installations, and additional learning investments required to reach grid parity is \$30.5 billion (Figure 2.21).[31]

Under this conservative SREC pricing scenario, we project PV to reach grid parity by 2018 and to achieve a 25% market saturation by the year 2035.

2.4.3.3 Comparing policy impacts

Table 2.8 summarizes results of the policy scenarios discussed in this section. Under the BAU scenario, grid parity is reached in 2031 and it would require up to \$145 billion in cumulative combined learning investments (i.e. R&D and learning-by-doing). Of this amount, \$142.2 billion

[31] As with the carbon pricing and R&D scenarios, required learning investments needed to reach grid parity are assumed to be financed by rebates, incentives and other public and utility policy tools.

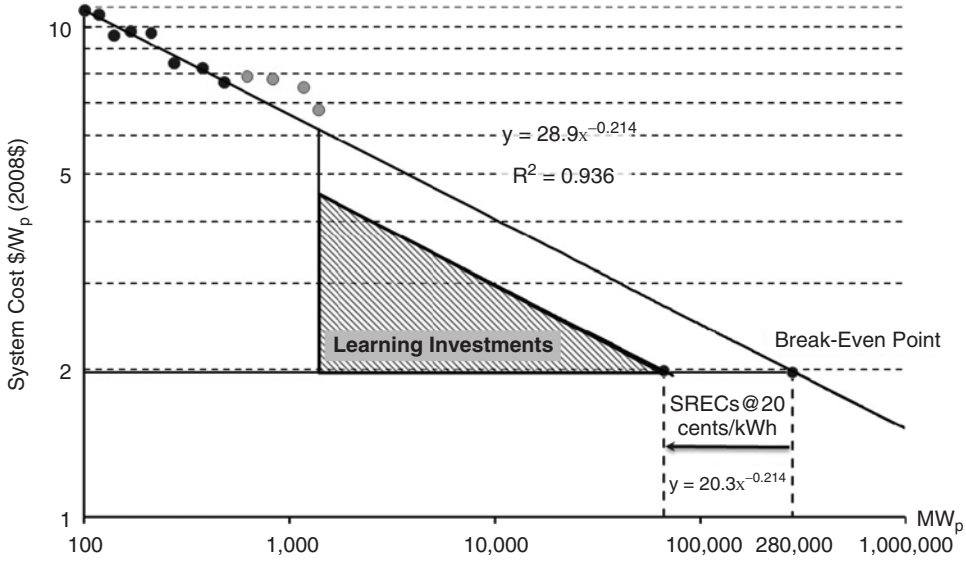


Figure 2.21 Impact of a national SREC requirement on the break-even point for PV

Table 2.8 Total investments required for reaching a PV break-even point, \$2 per W_p under different policy scenarios

	BAU scenario	CO_2 Price of \$25 per ton scenario	CO_2 Price of \$50 per ton scenario	Increased R&D scenario	SRECs at 20 cents/kW h scenario
R&D Investments (billion \$)	2.7	2.7	2.7	27.3	2.7
Learning-by-doing Investments (billion \$)	142.2	79.5	47.0	27.3	30.5
Combined Investments (billion \$)	144.9	81.2	49.7	54.6	33.2
Time to grid parity	2031	2024	2020	2018	2018

All scenarios assume learning-by-doing investments will be maintained at the schedule outlined in Figure 2.18 until break-even point is reached.

is policy incentives would be needed to reach grid parity in 2031. The introduction of carbon pricing (either through taxation or a cap-and-trade carbon market) would help non-carbon-based technologies, such as PV to compete with conventional fossil-based power generation technologies (natural gas turbines, coal plants, etc.) and would reduce the required learning-by-doing investments needed to reach grid parity. At a higher price per ton of CO_2 released, the cumulative volume of learning-by-doing investments needed to reach the break-even point on the learning curve would be lower (see Figure 2.19). Learning investments are reduced from \$142.2 billion to \$47.0 billion under a \$50 per ton of CO_2 emissions scenario and to \$79.5 billion under a \$25 per ton of CO_2 emissions scenario. Tenfold increase of public R&D would reduce the amount of learning-by-doing investments required to reach grid parity by \$115 billion to \$27.3 billion (the scenario assumes that same amount of cumulative investments are made in R&D). This dramatic reduction is mostly due to

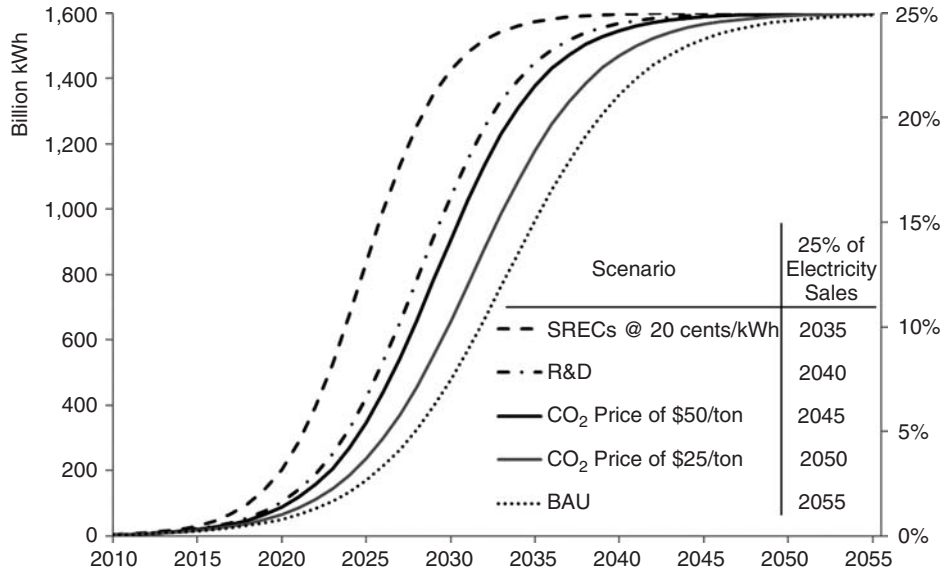


Figure 2.22 PV diffusion under different policy options in the US

the much earlier realization of grid parity, namely, 2018 for the R&D strategy compared with 2031 for the BAU strategy. Finally, the introduction of a national SREC market can reduce cumulative learning-by-doing investments to $30.5 billion and shorten the time to grid parity by 13 years (i.e. the BAU case of 2031 to the SREC case of 2018). Again, the savings from this scenario are very large – nearly $110 billion – and are attributable to the faster realization of the grid parity.

Diffusion curves under each of the policy scenarios investigated in this chapter were also constructed. The diffusion scenarios are benchmarked against the BAU scenario (see Figure 2.16). Based on this assumption, corresponding diffusion curves are constructed for each policy scenario. The results are presented in Figure 2.22, and show that a national RPS facilitates the fastest deployment of PV. A solar carve-out with SREC prices initially at 20 cents per kW h will shave nearly 20 years off the BAU time horizon of PV reaching 25% of the US electricity market share. That is, our BAU projection finds that 25% saturation will not occur until the year 2055, but a national SREC policy would hasten the achievement of that goal to 2035.

Increasing R&D funding for PV will also expedite the diffusion process. An expanded R&D policy can reduce the time needed to reach the 25% goal by almost 15 years. It may also rightly claim the benefit of strengthening the knowledge infrastructure to promote a green energy economy and, more broadly, a low-carbon future for the planet.

Finally, carbon taxes or cap-and-trade allowances have a noticeable impact, on PV diffusion, lessening the time for PV to reach the 25% goal by 5–10 years (compared with the BAU case). With regard to the effects of carbon pricing, three observations are in order. First, by comparison, the price of carbon modeled here is less than the SREC price options on a per kW h basis.[32] Thus, it is understandable why its impact is weaker. Of course, it is also important to observe that US jurisdictions have been able to enact laws leading to substantially higher SREC prices than the scenarios we have modeled. By contrast, no country or region has been able to sustain carbon

[32] A $25 per ton price of carbon emissions would increase the average retail electricity price in the US by about 1.5 cents per kW h, while a $50 per ton price would raise electricity prices by 3.0 cents per kW h.

pricing at or above $25 per ton. This leads to the third observation. Carbon pricing is perhaps the most complex and difficult policy approach to legislate because its effects are wide-ranging and will adversely affect the economics of some of the most powerful industries and companies in the world. Evaluated in this light, PV R&D and SREC have a contained policy reach and do not directly require raising the costs of fossil fuel competitors. Equally true, these tools only modestly affect the cost–benefit matrix that underlies the US and world economies. In this sense, their impacts can be strategically large, but cannot substitute for the systematic effect of a carbon pricing policy, even when its effective pricing of the pollutant is modest.

2.5 TOWARD A SUSTAINABLE FUTURE

In one of the boldest maps in print for our future, Herman Scheer's *The Solar Economy* [4] plots a worldwide shift from a non-renewable and unsustainable energy system to a renewable and, hopefully, sustainable future. Some may agree and others will disagree with assumptions behind the map or calculations he presents in support of its path. But one aspect of the map is indisputable: achieving a sustainable future requires solar energy to be at the center of the new economy.

Realizing this fact will not be easy, and almost certainly the world will disappoint Scheer with the slow pace and lack of vigor in its pursuit of a solar economy. Just because there is no other safe harbor for our future does not mean that humanity will avoid getting lost in its journey. That said, there are reasons and evidence for hope. In the last 15 years, policies have sprouted across the planet that have led to truly remarkable progress. Compared with where we were, a steady stream of policy incentives put into action – feed-in tariffs, solar carve-outs, renewable portfolio standards, community solar financing, and sustainable energy utilities – demonstrate a capacity for change that few would have predicted. In this regard, Scheer is more accurate about our ability to realize a different future than policy skeptics, or even policy moderates. The policy review in this chapter provides a detailed portrait of real change that has been gained, and underscores the potential for vastly larger improvement by using already invented tools which have been tested in very different national contexts, but with common results – a measurable advance in realizing a solar economy.

The policy analysis occupying the last half of the chapter offers a concrete projection of large-scale change built upon the use of already invented policies. Surely a global solar economy requires significant US participation, and the analysis plots a course of action for that country which would result in 25% of its electricity supply from solar generation in as little as 20 years. Achieving the goal will require public investment in PV R&D to sustain progress in the technology's performance and economic value while incentives are utilized to organize at-scale investment in solar energy. The experience of Germany, Spain and South Korea, among others, with the feed-in tariff tool (invented by Scheer)[33] indicates that policy can work to mobilize capital and can evolve markets that sustain significant PV penetration rates. Because national politics and economies differ, tools that aspire to the same outcome are also needed. The imitation of a feed-in tariff with a properly designed RPS, carve-out and SREC policy suite can promise the needed result and this chapter's analysis of the option shows that a transformed US electricity supply path is feasible.

It is also clear that systematic policy regime for realizing a solar economy will require comprehensive pricing of carbon uses and emissions. In the case of this tool, designs that initiate the process, even when carbon pricing is modest, have a demonstrable impact. In the longer run, this tool will guarantee that *all* energy decisions and investments reflect this factor. At the same time, a systematic policy regime *must have* specific initiatives such as FiT or SRECs to ensure rapid, concrete change. It would be a tragic error in policy design to fail to adopt the PV-focused policies modeled above.

[33] The 2009 Karl Boër Solar Medal was awarded to Hermann Scheer for his invention of the FiT [134].

The challenge of sustaining a future in the greenhouse remains large, daunting, and even unnerving [130]. The policy record is mixed and advances are slow. Still the interface of policy, technology and society is hopeful. Large change has happened before and the improvements of the last 15 years are promising.

REFERENCES

1. Jäger-Waldau, A, *PV Status Report 2009: Research, Solar Cell Production and Market Implementation of Photovoltaics*. Luxembourg: Office for Official Publications of the European Communities, 2009.
2. Renewable Energy Policy Network for the 21st Century, *Renewables Global Status Report 2009 Update*. Paris: REN21 Secretariat, 2009.
3. Byrne J, Toly N, Energy as a Social Project: Recovering a Discourse. in: J Byrne, N Toly, L Glover (eds) *Transforming Power: Energy, Environment, and Society in Conflict*. New Brunswick, NJ: Transaction Publishers, 2006, pp 1–32.
4. Scheer H, *The Solar Economy: Renewable Energy for a Sustainable Global Future*. London: Earthscan Publications Ltd., 2002.
5. Smil V, *Energy at the Crossroads: Global Perspectives and Uncertainties*. Cambridge, MA: The MIT Press, 2003.
6. Newman S, *The Final Energy Crisis*. London: Pluto Press, 2008.
7. International Energy Agency, *World Energy Outlook*. Paris: IEA Publications, 2009.
8. Energy Information Administration (U.S.), Annual Energy Outlook 2009 with Projections to 2030, Updated Annual Energy Outlook 2009 Reference Case with ARRA. [Online] 2009. [Cited: September 15, 2010.] http://www.eia.doe.gov/oiaf/servicerpt/stimulus/aeostim.html.
9. Energy Sources. [Online] 2010. [Cited: April 15, 2010.] http://www.eia.doe.gov.
10. International Energy Agency, *Electricity Information 2009*. Paris: IEA Publications, 2009.
11. Intergovernmental Panel on Climate Change, Fourth Assessment Report, Working Group III Report, Mitigation of Climate Change. [Online] 2007. [Cited: April 15, 2010.] http://www.ipcc.ch/ipccreports/ar4-wg3.htm.
12. International Energy Agency Photovoltaic Power Systems Program, Trends in Photovoltaic Applications: Survey Report of Selected IEA Countries between 1992-2008. [Online] 2009. [Cited: April 15, 2010.] http://www.iea-pvps.org/products/download/rep1_18.pdf.
13. Solarbuzz, World PV Industry Report Summary. [Online] 2009. [Cited: April 15, 2010.] http://www.solarbuzz.com./Marketbuzz2009-intro.htm.
14. Earth Policy Institute, Plan B 4.0 – Supporting Data for Chapters 4 and 5 – Solar. [Online] 2009. [Cited: April 15, 2010.] http://www.earth-policy.org/datacenter/pdf/book_pb4_ch4-5_solar_pdf.pdf.
15. Resources on Solar Energy. [Online] 2008. [Cited: April 15, 2010.] http://www.earthpolicy.org/Indicators/Solar/index.htm.
16. Prometheus Institute, *PV Manufacturing in the United States: Market Outlook, Incentives and Supply Chain Opportunities*. Boston, MA: Greentech Media Inc., 2009.
17. Marketbuzz, 2008 World PV Industry Report Highlights. [Online] 2009. [Cited: April 15, 2010.] http://www.solarbuzz.com/Marketbuzz2009-intro.htm.
18. Energy Information Administration (U.S.), International Energy Statistics. [Online] 2010a. [Cited: April 15, 2010.] http://tonto.eia.doe.gov/cfapps/ipdbproject/IEDIndex3.cfm.
19. O'Rourke S, Kim P, Polavarapu H, *Solar Photovoltaic Industry: Looking Through the Storm*. New York: Global Markets Research, Deutsch Bank Securities Inc., 2009.
20. Sawin J L, Another Sunny Year for Solar Power. *Worldwatch Institute*. [Online] 2008. [Cited: April 15, 2010.] http://www.worldwatch.org/node/5449#notes.
21. Mints P, 10 Years in the Sun: The Most Profitable Decade in PV History Draws to a Close. *Renewable Energy World Magazine* **13** (1), 40–45 (2010).

22. Goldman Sachs, *Alternative Energy: A Global Survey*. New York, NY: Goldman Sachs Global Markets Institute, 2007.
23. Solar Energy Industries Association, The Investment Tax Credit (ITC): SEIA's Top Legislative Priority, ITC Extended. [Online] 2008. [Cited: April 15, 2010.] http://www.seia.org/cs/solar_tax_policy.
24. Energy Information Administration (U.S.), Energy Improvement and Extension Act of 2008: Summary of Provisions. [Online] 2009c. [Cited: April 15, 2010.] http://www.eia.doe.gov/oiaf/aeo/otheranalysis/aeo_2009analysispapers/eiea.html.
25. Database of State Incentives for Renewables and Efficiency, Renewables Portfolio Standards with Solar/DG Provisions. [Online] 2010. [Cited: April 15, 2010.] http://www.dsireusa.org.
26. Bolinger M, Financing Non-Residential Photovoltaic Projects: Options and Implications. *Lawrence Berkeley National Laboratory*. [Online] 2009. [Cited: April 15, 2010.] http://eetd.lbl.gov/EAP/EMP/reports/lbnl-1410e.pdf.
27. International Renewable Energy Council, Net Metering Model Rules 2009 Edition. [Online] 2009. [Cited: April 15, 2010.] http://www.irecusa.org/NMmodel09.
28. Sherwood L, *U.S. Solar Market Trends* 2008. New York NY: International Renewable Energy Council, 2009.
29. *U.S. Solar Market Trends 2007*. New York: International Renewable Energy Council, 2008.
30. Wiser R, *et al.*, Tracking the Sun II: The installed Cost of Photovoltaics in the U.S. from 1990-2008. *National Renewable Energy Laboratory*. [Online] 2009. [Cited: April 15, 2010.] http://eetd.lbl.gov/ea/emp/reports/lbnl-2674e.pdf.
31. California Energy Commission, A Short History of Solar Energy and Solar Energy in California. [Online] 2009. [Cited: April 15, 2010.] http://www.gosolarcalifornia.ca.gov/solar101/history.html.
32. Pollution Engineering, Celebrating 20 Years of Solar Power. *Pollution Engineering*. Troy, Michigan: Business News Publishing Co., 2004. Vol. 36, p 36.
33. U.S. Department of Energy, Utility Applications Case Study: Power for a Utility Substation in California. [Online] 2009. [Cited: April 15, 2010.] http://www1.eere.energy.gov/solar/cs_ca_substation.html.
34. Hoff T, Shugar D S, The Value of Grid-Support Photovoltaics In Reducing Distribution System Losses. *IEEE Transactions on Energy Conversion* **10** (3), 569–576 (1995).
35. California Energy Commission, History of California's Renewable Energy Programs. [Online] 2009b. [Cited: April 15, 2010.] http://www.energy.ca.gov/renewables/history.html.
36. – . Emerging Renewables Program. [Online] 2010. [Cited: April 15, 2010.] http://www.energy.ca.gov/renewables/emerging_renewables/index.html.
37. – . Emerging Renewables Program: Fourth Edition. [Online] 2005. [Cited: April 15, 2010.] http://www.energy.ca.gov/2005publications/CEC-300-2005-001/CEC-300-2005-001-ED4F.PDF.
38. – . California Solar Photovoltaic Statistics & Data. [Online] 2009. [Cited: April 15, 2010.] http://www.energyalmanac.ca.gov/renewables/solar/pv.html.
39. Pace Financing, Property-Assessed Clean Energy (PACE) Financing Explained. [Online] 2010. [Cited: April 15, 2010.] http://pacefinancing.org.
40. Barnes J, *et al.* State Incentives and Policy Trends. [Online] 2009. [Cited: April 15, 2010.] http://irecusa.org/wp-content/uploads/2009/10/IREC-2009-Annual-ReportFinal.pdf.
41. California Public Utility Commission, California Solar initiative: Staff Progress Report, January 2009. [Online] 2009. [Cited: April 15, 2010.] http://www.energy.ca.gov/2009publications/CPUC-1000-2009-002/CPUC-1000-2009-002.PDF.
42. – . California Solar initiative: Staff Progress Report, January 2008. [Online] 2008. [Cited: April 15, 2010.] http://www.energy.ca.gov/2008publications/CPUC-1000-2008-002/CPUC-1000-2008-002.PDF.

43. California Energy Commission, The California Solar Initiative – CSI. [Online] 2009d. [Cited: April 15, 2010.] http://www.gosolarcalifornia.ca.gov/csi/index.html.
44. – . California Energy Commission's New Solar Homes Partnership. [Online] 2009. [Cited: April 15, 2010.] https://www.newsolarhomes.org.
45. Summit Blue Consulting, *Assessment of the New Jersey Renewable Energy Market*. Summit Blue Consulting. 2008. Volume 1. Submitted To: New JerseyBoard of Public Utilities Office of Clean Energy.
46. New Jersey Clean Energy Program, CORE Program Changes Chronology: Customer On-site Renewable Energy (CORE), Program Updated August, 2006. [Online] 2006. [Cited: April 15, 2010.] http://www.njcleanenergy.com/files/file/COREProgramUpdate081706.pdf.
47. – . NJ Renewable Energy Systems Installed. New Jersey Board of Public Utilities Office of Clean Energy. [Online] 2009. [Cited: April 15, 2010.] http://www.njcleanenergy.com/renewable-energy/program-updates/installation-summary.
48. – . Renewable Energy Incentive Program. [Online] 2009. [Cited: April 15, 2010.] http://www.njcleanenergy.com/renewable-energy/programs/renewable-energy-incentive-program.
49. – . New Jersey Approves Solar REC-Based Financing Program. [Online] 2007. [Cited: April 15, 2010.] http://www.njcleanenergy.com/files/file/SOLARTransitionFAQs121707%20fnl2(2).pdf.
50. – . SREC Pricing. [Online] 2010. [Cited: April 15, 2010.] www.njcleanenergy.com/renewable-energy/project-activity-reports/srec-pricing/srec-pricing.
51. SRECTrade, Solar Renewable Energy Certificates (SRECs). [Online] 2010. [Cited: April 15, 2010.] http://www.srectrade.com/background.php.
52. New Jersey Clean Energy Program, Customer On-Site Renewable Energy (CORE) Program. [Online] 2009. [Cited: April 15, 2010.] http://www.njcleanenergy.com/renewable-energy/home/home.
53. – . 2009 SREC Registration Process Status Reports. [Online] 2009. [Cited: April 15, 2010.] http://www.njcleanenergy.com/misc/renewable-energy/weekly-status-reports.
54. Nevada Energy, NV Energy Will Begin Accepting Applications For The Solar Incentive Program On April 21, 2010. *SolarGenerations*. [Online] 2010. [Cited: April 15, 2010.] http://www.nvenergy.com/renewablesenvironment/renewablegenerations/solargen/index.cfm.
55. Wang U, First Solar Reaches Grid-Parity Milestone, Says Report. *GreenTechSolar*. [Online] 2008. [Cited: April 15, 2010.] http://www.greentechmedia.com/articles/read/first-solar-reaches-grid-parity-milestone-says-report-5389/.
56. Xcel Energy, Solar Rewards. [Online] 2010. [Cited: April 15, 2010.] http://www.xcelenergy.com/Colorado/Residential/RenewableEnergy/Solar_Rewards/Pages/home.aspx.
57. Delaware Public Service Commission, Certified Eligible Energy Resources (Excel spreadsheet). [Online] 2010. [Cited: April 15, 2010.] http://depsc.delaware.gov/electric/rps_resources.xls.
58. Sustainable Energy Utility Task Force, The Sustainable Energy Utility: Delaware First. *Sustainable Energy Utility*. [Online] 2008. [Cited: April 15, 2010.] http://www.seu-de.org/docs/SEU_Final_Report.pdf.
59. Regional Greenhouse Gas Initiative, [Online] 2010. [Cited: April 15, 2010.] http://www.rggi.org/home.
60. Energize Delaware, Delaware Gets Clean Energy Jobs Boost. [Online] 2010. [Cited: April 15, 2015.] http://www.energizedelaware.org/sites/default/files/Dover%20Sun%20Park%20PR%281%29.pdf.
61. University of Delaware, Introducing the University of Delaware Climate Action Plan. [Online] 2009. [Cited: April 15, 2010.] http://www.udel.edu/sustainability/footprint/.
62. Rahim S, State and Local Governments Innovate to Cut Energy Waste. *The New York Times*. February 11, 2010.

63. Vermont Clean Energy Development Fund, Vermont Clean Energy Development Fund: 2008 Annual Report. *Vermont Department of Public Service*. [Online] 2009. [Cited: April 15, 2010.] http://publicservice.vermont.gov/energy/ee_cleanenergyfund.html.
64. Renewable Energy Resource Center, The Vermont Small Scale Renewable Energy Incentive Program. [Online] 2010. [Cited: April 15, 2010.] http://www.rerc-vt.org/incentives/index.htm.
65. Byrne J, *et al.* American Policy Conflict in the Greenhouse: Divergent Trends in Federal Regional, State, and Local Green Energy and Climate Change Policy. *Energy Policy*. **35** (9), 4555–4573 (2007).
66. Erge T, Hoffmann V U, Kiefer K, The German Experience with Grid-Connected PV-Systems. *Solar Energy*. **70** (6), 479–487 (2001).
67. Laukamp H, *et al.*, Reliability Issues in PV Systems-Experience and Improvements. *2nd World Solar Electric Buildings Conference*, Sydney, Australia, s.n., (2000).
68. Gipe P, The Original Electricity Feed Law in Germany. [Online] 2009. [Cited: April 15, 2009.] http://www.wind-works.org/FeedLaws/Germany/ARTsDE.html.
69. Wissing L, National Survey Report of PV Power Applications in Germany 2008. *International Energy Agency, Co-Operative Programme on Photovoltaic Power Systems*. [Online] 2, 2009. [Cited: April 15, 2010.] from http://www.iea-pvps.org/countries/download/nsr08/NSR%20Germany%202008.pdf.
70. Bolinger M, Wiser R, Support for PV in Japan and Germany. *Berkley Lab and Clean Energy Group*. [Online] 2002. [Cited: April 15, 2010.] http://eetd.lbl.gov/ea/EMP/cases/PV_in_Japan_Germany.pdf.
71. International Energy Agency, *Energy Policies of IEA Countries: Germany*. Paris: IEA Publications, 2007.
72. KfW Banking Group, Third Quarterly Report. [Online] 2005. [Cited: April 15, 2010.] http://www.kfw.de/DE_Home/Service/Download_Center/Finanzpublikationen/PDF_Dokumente_Berichte_etc./5_Quartalsberichte/3._Quartal/3_QB_2005_e.pdf.
73. Renewable Energy Sources Act, Erneuerbare-Energien-Gesetz EEG. [Online] 2004. [Cited: April 15, 2010.] http://www.iea.org/Textbase/pm/?mode=re&action=detail&id=1969.
74. Bundesverband Solarwirtschaft, EEG 2009 Important Changes and Feed-in Tariffs for Photovoltaics. [Online] 2009. [Cited: April 15, 2010.] http://en.solarwirtschaft.de/fileadmin/content_files/EEG_revision_EN_consol.pdf.
75. Royal Decree 436, Establishing The Methodology for the Updating and Systematisation of the Legal and Economic Regime for Electric Power Production in the Special Regime. [Online] 2004. [Cited: April 15, 2010.] http://onlinepact.org/fileadmin/user_upload/PACT/Laws/Spain_436_2004_english.pdf.
76. Gil J, Lucas H, Spain: New Plan for Renewable Energy. *Renewable Energy World*. **11** (2005).
77. Voosen P, Spain's Solar Market Crash Offers a Cautionary Tale About Feed-In Tariffs. *The New York Times*. August 18, 2009.
78. Held A, *et al.*, Feed-In Systems in Germany, Spain and Slovenia – A Comparison. *International Feed-in Cooperation*. [Online] 2007. [Cited: April 15, 2010.] http://www.feed-in-cooperation.org/wDefault_7/content/research/research.php.
79. Salas V, Status of PV Policy and Market in Spain. [Online] 2009. [Cited: April 15, 2010.] http://www.mbipv.net.my/dload/Spain.pdf.
80. Jäger-Waldau A, *PV Status Report 2003: Research, Solar Cell Production and Market Implementation in Japan, USA and the European Union*. Luxembourg: Office for Official Publications of the European Communities, 2003.
81. Ikki O, Matsubara K, National Survey Report of PV Power Applications in Japan 2006. *International Energy Agency Photovoltaic Power Systems Program*. [Online] 2007. [Cited: April 15, 2010.] http://iea-pvps.org/countries/download/nsr06/06jpnnsr.pdf.

82. Jäger-Waldau A, *PV Status Report 2006: Research, Solar Cell Production and Market Implementation of Photovoltaics*. Luxembourg: Office for Official Publications of the European Communities, 2006.
83. Ikki O, Matsubara K, National Survey Report of PV Power Applications in Japan 2007. *International Energy Agency Photovoltaic Power Systems Program*. [Online] 2008. [Cited: April 15, 2010.] http://iea-pvps.org/countries/download/nsr07/2007_NSR_Japan_080610.pdf.
84. Jäger-Waldau A, *PV Status Report 2008: Research, Solar Cell Production and Market Implementation of Photovoltaics*. Luxembourg: Office for Official Publications of the European Communities, 2008.
85. Ikki O, Kaizuka I, Overview of Urban Scale PV Projects in Japan. *IEA PVPS Task 10 Workshop*. [Online] 2005. [Cited: April 21, 2010.] from http://www.iea-pvps-task10.org/IMG/pdf/5-RTS-Corporation.pdf.
86. Ikki O, Tanaka Y, National Survey Report of PV Power Applications in Japan 2003. *International Energy Agency Photovoltaic Power Systems Program*. [Online] 2004. [Cited: April 15, 2010.] http://iea-pvps.org/countries/download/nsr03/jpn.pdf.
87. Green Gross International, 2008 Global Solar Report Cards: the Time Has Come to Harness the Sun. [Online] 2008. [Cited: April 15, 2010.] http://globalgreen.org/docs/publication-96-1.pdf.
88. Yoon K H, Kim D, Yoon K S, National Survey Report of PV Power Applications in Korea 2006. *International Energy Agency Photovoltaic Power Systems Program*. [Online] 2007. [Cited: April 15, 2010.] http://iea-pvps.org/countries/download/nsr06/06kornsr.pdf.
89. International Energy Agency Photovoltaic Power Systems Program, Trends in Photovoltaic Applications: Survey Report of Selected IEA Countries between 1992-2007. [Online] 2008. [Cited: April 15, 2010.] http://iea-pvps.org/products/download/rep1_17.pdf.
90. Yoon K H, Kim D, National Survey Report of PV Power Applications in Korea 2008. *International Energy Agency Photovoltaic Power Systems Program*. [Online] 2009. [Cited: April 15, 2010.] http://iea-pvps.org/countries/download/nsr08/NSR%20Korea%202008.pdf.
91. – . National Survey Report of PV Power Applications in Korea 2007. *International Energy Agency Photovoltaic Power Systems Program*. [Online] 2008. [Cited: April 15, 2010.] http://iea-pvps.org/countries/download/nsr07/2007_nsr_korea.pdf.
92. International Energy Agency, Energy Technology RD&D 2009 Edition Database. [Online] 2010. [Cited: April 15, 2010.] http://www.iea.org/stats/rd.asp.
93. Organisation for Economic Co-operation and Development, *Technology, Innovation, Development and Diffusion*. Paris: OECD/IEA, 2003.
94. Rogers E M, *Diffusion of Innovations*. New York, NY: Free Press, 1995.
95. International Energy Agency, *Creating Markets for Energy Technologies*. Paris: IEA Publications, 2003.
96. Jenkins N, *et al. Emerging Technologies, Energy Efficiency, Roles and Linkages*. San Francisco, CA: American Council for Energy Efficient Economy, 2004.
97. Faiers A, Neame C, Consumer Attitudes Towards Domestic Solar Power Systems. *Energy Policy*. **34** (14), 797–1906 (2006).
98. Meyer P S, Yung J W, Ausubel J H, A Primer on Logistic Growth and Substitution: The Mathematics of Loglet Lab Software. *Technology Forecasting and Social Change* **61** (3), 247–271 (1999).
99. Byrne J, *et al.* Beyond Oil: A Comparison of Projections of PV Generation and European and U.S. Domestic Oil Production. in: D Y Goswami (ed.), *Advances in Solar Energy: An Annual Review of Research and Development*. Sterling, VA: Earthscan, 2005, pp 35–69.
100. Byrne J, *et al.*, The Potential of Solar Electric Power for Meeting Future U.S. Energy Needs: a Comparison of Projections of Solar Electric energy Generation and Arctic National Wildlife Refuge Oil production. *Energy Policy* **32** (2), 289–297 (2004).
101. Hubbert M K, *Energy Resources: Report to the Committee on Natural Resources*. Washington, DC: National Academy of Science and National Resource Council, 1962.

102. Laherrère J H, The Hubbert Curve: its strength and Weaknesses. [Online] 2000. [Cited: April 15, 2010.] http://dieoff.org/page191.htm.
103. Fisher J C, Pry R H, A Simple Substitution Model of Technological Change. *Technological Forecasting and Social Change*. **3** (1), 75–88 (1971).
104. Woodall P, Untangling E-Conomics: A Survey of the New Economy. *The Economist*. September 2000, pp 23–29.
105. European Wind Energy Association, *Wind Force 10: A Blueprint to Achieve 10% of the World's Electricity from Wind Power by 2020*. London: EWEA, 1999.
106. Collantes G O, Incorporating Stakeholders' Perspectives into Models of New Technology Diffusion: The Case of Fuel-cell Vehicles. *Technological Forecasting and Social Change*. **74** (3), 267–280 (2007).
107. Wright T P, Factors Affecting the Cost of Airplanes. *Journal of Aeronautical Sciences* **3** (4), 122–128 (1936).
108. Duke R, Kammen D M, The Economics of Energy Market Transformation Programs. *The Energy Journal* **20** (4), 15–64 (1999).
109. Argote L, Epple D, Learning Curves in Manufacturing. *Science*. **247**, 920–924 (1990).
110. Neij L, Cost Development of Future Technologies for Power Generation – A Study Based On Experience Curve and Complementary Bottom-Up Assessments. *Energy Policy*. **36** (6), 2200–2211 (2008).
111. Poponi D, Byrne J, Hegedu S, Break-even Price Estimates for Residential PV Applications in OCED Countries with an Analysis of Prospective Cost Reductions. *Energy Studies Review* **14** (1), 104–117 (2006).
112. Zwaan B, Rabl A, Prospects for PV: A Learning Curve Analysis. *Solar Energy* **74** (1), 19–31 (2003).
113. Colpier U C, Cornland D, The Economics of the Combined Cycle Gas Turbine – an Experience Curve Analysis. *Energy Policy*. **30** (4), 309–316 (2002).
114. International Energy Agency, *Experience Curve for Energy Technology Policy*. Paris: IEA Publications, 2000.
115. Reis D A, Learning Curves in Food Services. *Journal of Operational Research Society* **42** (8), 623–629 (1991).
116. International Energy Agency, *Energy Technology Perspectives 2008: Scenarios & Strategies to 2050*. Paris: IEA Publications, 2008.
117. Berglund C, Söderholm P, Modeling Technical Change in Energy System Analysis: Analyzing the Introduction of Learning-by-doing in bottom-up Energy Models. *Energy Policy* **34** (12), 1344–1356 (2006).
118. Miketa A, Schrattenholzer L, Experiments with a Methodology to Model the Role of R&D Expenditure in Energy Technology Learning Processes. *Energy Policy* **32** (15), 1679–1692 (2004).
119. Barreto L, Kypreos S, Endogenizing R&D and Market Experience in the 'Bottom-up' Energy-Systems ERIS Model. *Technovation* **24** (8), 615–629 (2004).
120. Kouvaritakis N, Soria A, Isoard S, Modeling Energy Technology Dynamics: Methodology for Adaptive Expectations Model with Learning by Doing and Learning by Searching. *International Journal of Global Energy Issues*. **14** (1), 104–115 (2000).
121. Kobos P H, Reickson J D Drennen T E, Technological Learning and Renewable Energy Costs: Implications for U.S. Renewable Energy Policy. *Energy policy*. **34** (13), 1645–1658 (2006).
122. Watanabe C, Wakabayashi K, Miyazawa T, Industrial Dynamism and the Creation of a 'Virtuous Cycle' Between R&D, Market Growth and Price Reduction – the Case of Photovoltaic Power Generation (PV) Development in Japan. *Technovation*. **20** (6), 299–312 (2000).
123. Energy Information Administration (U.S.), Monthly Electric Utility Sales and Revenue Data: Utility Level Retail Sales of Electricity and Associated Revenue by End-Use Sector,

State, and Reporting Month. EIA-826, 1990-2009. [Online] 2010b. [Cited: April 15, 2010.] http://www.eia.doe.gov/cneaf/electricity/page/eia826.html.
124. Denholm P, Margolis R, Very Large-Scale Deployment of Grid-Connected Solar Photovoltaics in the United States: Challenges and Opportunities. *National Renewable Energy Laboratory*. [Online] 2006. [Cited: April 15, 2010.] http://www.nrel.gov/pv/pdfs/39683.pdf.
125. National Renewable Energy Laboratory, Photovoltaic Solar Resources of the United States. [Online] 2008. [Cited: April 15, 2010.] http://www.nrel.gov/gis/solar.html.
126. Masters G M, *Renewable and Efficient Electric Power Systems*. Hoboken, NJ: John Wiley & Sons, Inc., 2004.
127. Stoft S, *Power System Economics: Designing Markets for Electricity*. New York NY: John Wiley & Sons, Inc., 2002.
128. Ellerman D A, Joskow P L, The European Union's Emissions Trading System in Perspective. Prepared for the Pew Center on Global Climate Change. *Pew Center on Global Climate Change*. [Online] 2008. [Cited: April 15, 2010.] http://www.pewclimate.org/docUploads/EU-ETS-In-Perspective-Report.pdf.
129. Aldy J E, Pizer W A, The Competitiveness Impacts of Climate Change Mitigation Policies. *Pew Center on Global Climate Change*. [Online] 2009. [Cited: April 15, 2010.] http://www.pewclimate.org/docUploads/competitiveness-impacts-report.pdf.
130. Byrne J, Kurdgelashvili L, Hughes K, Undoing Atmospheric Harm: Civil Action to Shrink the Carbon Footprint. in: P Droege (ed.), *Urban Energy Transition: From Fossil Fuels to Renewable Power*. Oxford: Elsevier, 2008, pp 27–53.
131. Lazard, *Levelized Cost of Energy Analysis – Version 3.0*. 2009.
132. Sharma A, Wilkinson S, Guest Blog: 2009 PV Module Market–Installations and Shipments up, Revenues Down. [Online] 2010. [Cited: February 15, 2010.] http://www.pv-tech.org/editors_blog.
133. Owen A, Renewable Energy: Externality Costs as Market Barriers. *Energy Policy*. **34** (5), 634–642 (2006).
134. UDaily, [Online] March 17, 2009. [Cited: April 15, 2010.] http://www.udel.edu/udaily/2009/mar/boeraward031709.html.

3

The Physics of the Solar Cell

Jeffery L. Gray

Purdue University, West Lafayette, Indiana, USA

3.1 INTRODUCTION

Semiconductor solar cells are fundamentally quite simple devices. Semiconductors have the capacity to absorb light and to deliver a portion of the energy of the absorbed photons to carriers of electrical current – electrons and holes. A semiconductor diode separates and collects the carriers and conducts the generated electrical current preferentially in a specific direction. Thus, a solar cell is simply a semiconductor diode that has been carefully designed and constructed to efficiently absorb and convert light energy from the sun into electrical energy.

A simple conventional solar cell structure is depicted in Figure 3.1. Sunlight is incident from the top, on the front of the solar cell. A metallic grid forms one of the electrical contacts of the diode and allows light to fall on the semiconductor between the grid lines and thus be absorbed and converted into electrical energy. An antireflective layer between the grid lines increases the amount of light transmitted to the semiconductor. The semiconductor diode is fashioned when an n-type semiconductor and a p-type semiconductor are brought together to form a metallurgical junction. This is typically achieved through diffusion or implantation of specific impurities (dopants) or via a deposition process. The diode's other electrical contact is formed by a metallic layer on the back of the solar cell.

All electromagnetic radiation, including sunlight, can be viewed as being composed of particles called photons which carry specific amounts of energy determined by the spectral properties of their source. Photons also exhibit a wavelike character with the wavelength, λ, being related to the photon energy E_λ by

$$E_\lambda = \frac{hc}{\lambda} \tag{3.1}$$

where h is Plank's constant and c is the speed of light. Only photons with sufficient energy to create an electron–hole pair, that is, those with energy greater than the semiconductor bandgap

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

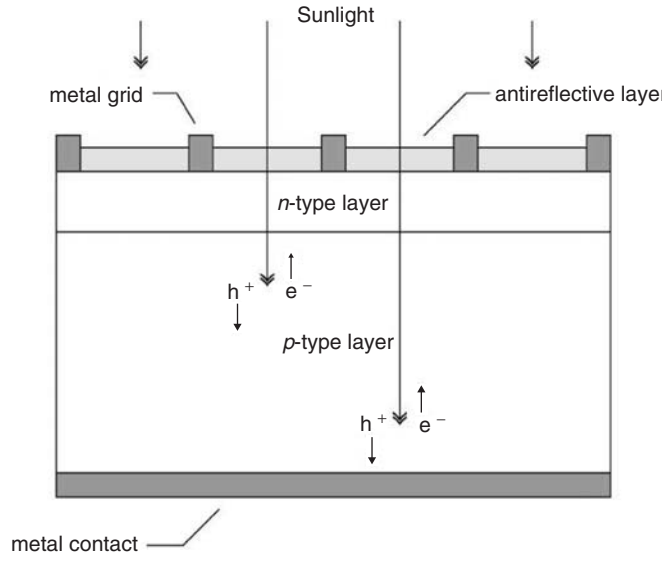


Figure 3.1 A schematic of a simple conventional solar cell. Creation of electron–hole pairs, e^- and h^+, respectively, is depicted

(E_G), will contribute to the energy conversion process. Thus, the spectral composition of sunlight is an important consideration in the design of efficient solar cells.

The sun has a surface temperature of approximately 5762 K and its radiation spectrum can be approximated by a black body radiator at that temperature. Emission of radiation from the sun, as with all black body radiators, is isotropic. However, the Earth's great distance from the sun (approximately 93 million miles or 150 million kilometers) means that only those photons emitted directly at the Earth contribute to the solar spectrum as observed from the Earth. Therefore, for most practical purposes, the light falling on the Earth can be thought of as parallel streams of photons. Just above the Earth's atmosphere, the radiation intensity, or solar constant, is about 1.353 kW/m^2 [1] and the spectral distribution is referred to as an *air mass zero* (AM0) radiation spectrum. The air mass is a measure of how absorption in the atmosphere affects the spectral content and intensity of the solar radiation reaching the Earth's surface. The air mass number is given by [1]

$$\text{Air mass} = \frac{1}{\cos\theta} \tag{3.2}$$

where θ is the angle of incidence ($\theta = 0$ when the sun is directly overhead). The air mass number is always greater than or equal to one at the Earth's surface.

A widely used standard for comparing solar cell performance is the AM1.5 ($\theta = 48.2^\circ$) spectrum normalized to a total power density of 1 kW/m^2. The spectral content of sunlight at the Earth's surface also has a diffuse (indirect) component due to scattering and reflection in the atmosphere and surrounding landscape, and can account for up to 20% of the light incident on a solar cell. The air mass number is therefore further defined by whether or not the measured spectrum includes the diffuse component. An AM1.5g (global) spectrum includes the diffuse component, while an AM1.5d (direct) does not. Black body ($T = 5762$ K), AM0, and AM1.5g radiation spectrums are shown in Figure 3.2. The air mass and solar radiation are described in more detail in Chapters 18 and 22.

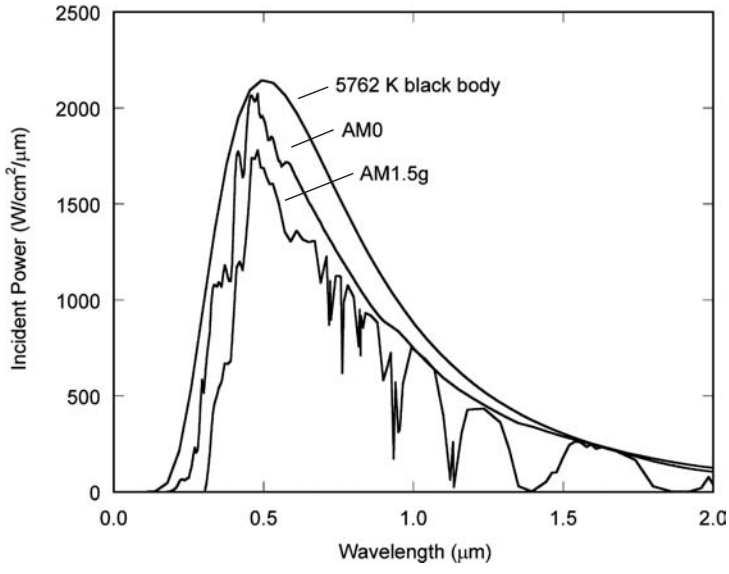


Figure 3.2 The radiation spectrum for a black body at 5780 K, an AM0 spectrum, and an AM1.5 global spectrum

The basic physical principles underlying the operation of solar cells are the subject of this chapter. First, a brief review of the fundamental properties of semiconductors is given that includes an overview of semiconductor band structure and carrier generation, recombination, and transport. Next, the electrostatic properties of the *pn*-junction diode are reviewed, followed by a description of the basic operating characteristics of the solar cell, including the derivation (based on the solution of the minority-carrier diffusion equation) of an expression for the current–voltage characteristic of an idealized solar cell. This is used to define the basic solar cell figures of merit, namely, the open-circuit voltage V_{OC}; the short-circuit current I_{SC}; the fill factor FF; the conversion efficiency η, and the collection efficiency η_C. Much of the discussion here will focus on how carrier recombination is the primary factor controlling solar cell performance. Finally, some additional topics relevant to solar cell operation, design and analysis are presented. These include the relationship between bandgap and efficiency, the solar cell spectral response, parasitic resistive effects, temperature effects, voltage-dependent collection, a brief introduction to some modern cell design concepts, and a brief overview of detailed numerical modeling of solar cells.

3.2 FUNDAMENTAL PROPERTIES OF SEMICONDUCTORS

An understanding of the operation of semiconductor solar cells requires familiarity with some basic concepts of solid-state physics. Here, an introduction is provided to the essential concepts needed to examine the physics of solar cells. More complete and rigorous treatments are available from a number of sources [2–6].

Solar cells can be fabricated from a number of semiconductor materials, most commonly silicon (Si) – crystalline, polycrystalline, and amorphous. Solar cells are also fabricated from other semiconductor materials such as GaAs, GaInP, Cu(InGa)Se_2, and CdTe, to name but a few. Solar cell materials are chosen largely on the basis of how well their absorption characteristics match the solar spectrum and upon their cost of fabrication. Silicon has been a common choice due to

Table 3.1 Abbreviated periodic table of the elements

I	II	III	IV	V	VI
		B	C	N	O
		Al	Si	P	S
Cu	Zn	Ga	Ge	As	Se
Ag	Cd	In	Sn	Sb	Te

the fact that its absorption characteristics are a fairly good match to the solar spectrum, and silicon fabrication technology is well developed as a result of its pervasiveness in the semiconductor electronics industry.

3.2.1 Crystal Structure

Electronic grade semiconductors are very pure crystalline materials. Their crystalline nature means that their atoms are aligned in a regular periodic array. This periodicity, coupled with the atomic properties of the component elements, is what gives semiconductors their very useful electronic properties. An abbreviated periodic table of the elements is given in Table 3.1.

Note that silicon is in column IV, meaning that it has four valence electrons – that is, four electrons that can be shared with neighboring atoms to form covalent bonds with those neighbors. In crystalline silicon, the atoms are arranged in a *diamond lattice* (carbon is also a column IV element) with tetrahedral bonding – four bonds from each atom where the angle between any two bonds is 109.5°. Perhaps surprisingly, this arrangement can be represented by two interpenetrating face-centered cubic (*fcc*) unit cells where the second *fcc* unit cell is shifted one-fourth of the distance along the body diagonal of the first *fcc* unit cell. The lattice constant, ℓ, is the length of the edges of the cubic unit cell. The entire lattice can be constructed by stacking these unit cells. A similar arrangement, the *zincblende* lattice, occurs in many binary III–V and II–VI semiconductors such as GaAs (a III–V compound) and CdTe (a II–VI compound). For example, in GaAs, one interpenetrating *fcc* unit cell is composed entirely of gallium atoms and the other entirely of arsenic atoms. Note that the average valency is four for each compound, so that there are four bonds to and from each atom with each covalent bond involving two valence electrons. Some properties of semiconductors are dependent on the orientation of the crystal lattice, and casting the crystal structure in terms of a cubic unit cell makes identifying the orientation easier by means of Miller indices.

3.2.2 Energy Band Structure

Of more consequence to the physics of solar cells, however, is how the periodic crystalline structure of the semiconductor establishes its electronic properties. An electron moving in a semiconductor material is analogous to a particle confined to a three-dimensional box that has a complex interior structure, due primarily to the potential fields surrounding the component atom's nucleus and tightly bound core electrons. The dynamic behavior of the electron can be established from the electron wavefunction, ψ, which is obtained by solving the time-independent Schrödinger equation

$$\nabla^2\psi + \frac{2m}{\hbar^2}[E - U(\vec{r})]\psi = 0 \tag{3.3}$$

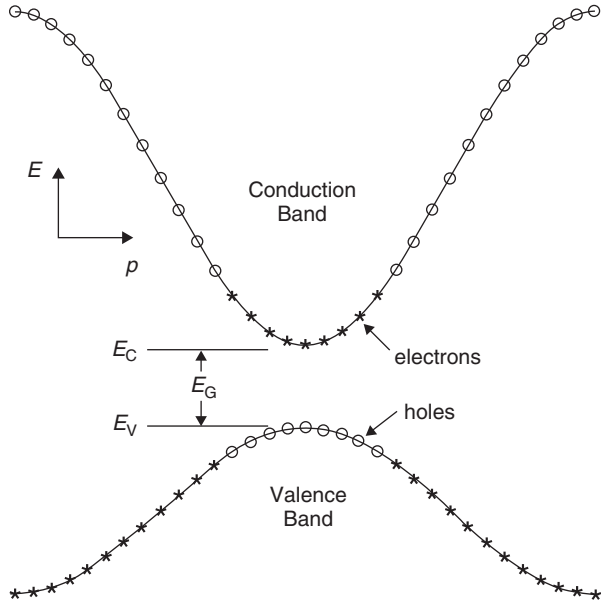


Figure 3.3 A simplified energy band diagram at $T > 0\,\mathrm{K}$ for a direct bandgap (E_G) semiconductor. Electrons near the maxima in valence band have been thermally excited to the empty states near the conduction-band minima, leaving behind holes. The excited electrons and remaining holes are the negative and positive mobile charges that give semiconductors their unique transport properties

where m is electron mass, $\hbar$ is the reduced Planck constant, E is the energy of the electron, and $U(\vec{r})$ is the periodic potential energy inside the semiconductor. Solving this quantum mechanical equation is beyond the scope of this work, but suffice it to say that the solution defines the band structure (the allowed electron energies and the relationship between the electron's energy and momentum) of the semiconductor and, amazingly, tells us that the quantum mechanically computed motion of the electron in the crystal is, to a good approximation, like that of an electron in free space if its mass, m, is replaced by an effective mass m^* in Newton's second law of motion. Newton's second law of motion, from classical mechanics, is

$$F = m^* a \tag{3.4}$$

where F is the applied force and a is the acceleration of the electron.

A simplified energy band structure is illustrated in Figure 3.3. The allowed electron energies are plotted against the crystal momentum, $p = \hbar k$, where k is the wave vector (represented here as a scalar for simplicity) corresponding to the wavefunction solutions of the Schrödinger equation. Only the energy bands of immediate interest are shown – energy bands below the valence band are presumed to be fully occupied by electrons and those above the conduction band are presumed to be empty. The electron effective mass is defined by the curvature of the band as

$$m^* \equiv \left[\frac{\mathrm{d}^2 E}{\mathrm{d}p^2}\right]^{-1} = \left[\frac{1}{\hbar^2}\frac{\mathrm{d}^2 E}{\mathrm{d}k^2}\right]^{-1}. \tag{3.5}$$

Near the top of the valence band, the effective mass is actually negative. Electrons ($*$) fill the states from bottom to top and the states near the top of the valence band are empty ($\circ$) due to some

electrons being thermally excited into the conduction band. These empty states can conveniently be regarded as positively charged carriers of current called *holes* with a positive effective mass. It is conceptually much easier to deal with a relatively few number of holes that have a positive effective mass since they will behave like classical positively charged particles.

Notice that the effective mass is not constant within each band. The top of the valence band and the bottom of the conduction band are approximately parabolic in shape and therefore the electron effective mass (m_n^*) near the bottom of the conduction band is a constant, as is the hole effective mass (m_p^*) near the top of the valence band. This is a very practical assumption that greatly simplifies the modeling of semiconductor devices such as solar cells.

When the minimum of the conduction band occurs at the same value of the crystal momentum as the maximum of the valence band, as it does in Figure 3.3, the semiconductor is a *direct bandgap* semiconductor. When they do not align, the semiconductor is said to be an *indirect bandgap* semiconductor. This is especially important when the absorption of light by a semiconductor is considered later in this chapter.

Even amorphous materials exhibit a similar band structure. Over short distances, the atoms are arranged in a periodic manner and an electron wavefunction can be defined. The wavefunctions from these small regions overlap in such a way that a *mobility gap* can be defined, with electrons above the mobility gap defining the conduction band and holes below the gap defining the valence band. Unlike crystalline materials, however, there are a large number of localized energy states within the mobility gap (band tails and dangling bonds) that complicate the analysis of devices fabricated from these materials. Amorphous silicon (a-Si) solar cells are discussed in Chapter 12.

3.2.3 Conduction-band and Valence-band Densities of State

Now that the dynamics of the electron motion in a semiconductor has been approximated by a negatively charged particle with mass m_n^* in the conduction band and by a positively charged particle with mass m_p^* in the valence band, it is possible to calculate the density of states in each band. This again involves solving the time-independent Schrödinger equation for the wavefunction of a particle in a box, but in this case the box is empty. All the complexities of the periodic potentials of the component atoms have been incorporated into the effective mass. The density of states in the conduction band is given by [3]

$$g_C(E) = \frac{m_n^* \sqrt{2m_n^*(E - E_C)}}{\pi^2 \hbar^3} \text{ cm}^{-3}\text{eV}^{-1} \tag{3.6}$$

while the density of states in the valence band is given by

$$g_V(E) = \frac{m_p^* \sqrt{2m_p^*(E_V - E)}}{\pi^2 \hbar^3} \text{ cm}^{-3}\,\text{eV}^{-1}. \tag{3.7}$$

3.2.4 Equilibrium Carrier Concentrations

When the semiconductor is in thermal equilibrium (i.e. at a uniform temperature with no external injection or generation of carriers), the Fermi function determines the ratio of filled states to available states at each energy and is given by

$$f(E) = \frac{1}{1 + e^{(E - E_F)/kT}} \tag{3.8}$$

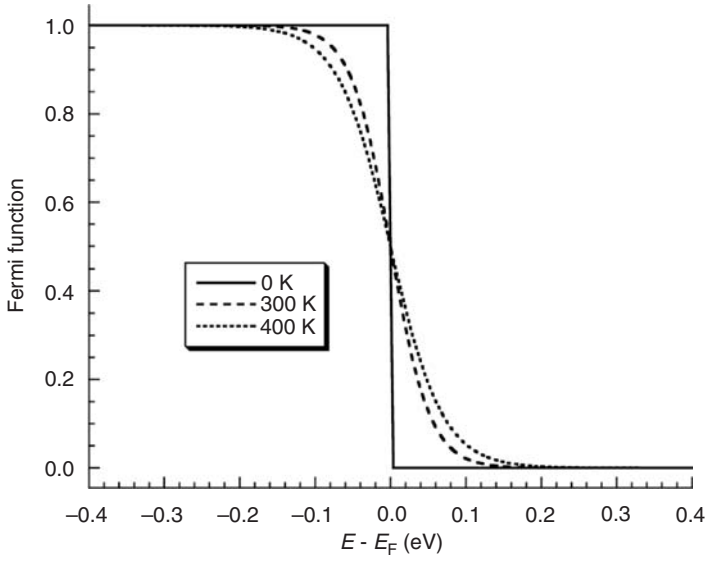


Figure 3.4 The Fermi function at various temperatures

where E_F is the Fermi energy, k is Boltzmann's constant, and T is the Kelvin temperature. As seen in Figure 3.4, the Fermi function is a strong function of temperature. At absolute zero, it is a step function and all the states below E_F are filled with electrons and all those above E_F are completely empty. As the temperature increases, thermal excitation will leave some states below E_F empty, and the corresponding number of states above E_F will be filled with the excited electrons.

The equilibrium electron and hole concentrations (number per cm^3) are therefore

$$n_o = \int_{E_C}^{\infty} g_C(E) f(E) \mathrm{d}E = \frac{2N_C}{\sqrt{\pi}} F_{1/2}((E_F - E_C)/kT) \tag{3.9}$$

$$p_o = \int_{-\infty}^{E_V} g_V(E)[1 - f(E)] \mathrm{d}E = \frac{2N_V}{\sqrt{\pi}} F_{1/2}((E_V - E_F)/kT) \tag{3.10}$$

where $F_{1/2}(\xi)$ is the Fermi–Dirac integral of order 1/2,

$$F_{1/2}(\xi) = \int_0^{\infty} \frac{\sqrt{\xi'}\mathrm{d}\xi'}{1 + \mathrm{e}^{\xi' - \xi}} \tag{3.11}$$

The conduction-band and valence-band effective densities of state (#/cm^3), N_C and N_V, respectively, are given by

$$N_C = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2} \tag{3.12}$$

and

$$N_V = 2\left(\frac{2\pi m_p^* kT}{h^2}\right)^{3/2}. \tag{3.13}$$

When the Fermi energy, E_F, is sufficiently far (>3 kT) from either band edge, the carrier concentrations can be well approximated (to within 2%) as [7]

$$n_o = N_C e^{(E_F - E_C)/kT} \tag{3.14}$$

and

$$p_o = N_V e^{(E_V - E_F)/kT}, \tag{3.15}$$

and the semiconductor is said to be *nondegenerate*. In nondegenerate semiconductors, the product of the equilibrium electron and hole concentrations is independent of the location of the Fermi energy and is just

$$p_o n_o = n_i^2 = N_C N_V e^{(E_V - E_C)/kT} = N_C N_V e^{-E_G/kT}. \tag{3.16}$$

In an undoped (intrinsic) semiconductor in thermal equilibrium, the number of electrons in the conduction band and the number of holes in the valence band are equal; $n_o = p_o = n_i$, where n_i is the intrinsic carrier concentration. The intrinsic carrier concentration can be computed from Equation (3.17), giving

$$n_i = \sqrt{N_C N_V} e^{(E_V - E_C)/2kT} = \sqrt{N_C N_V} e^{-E_G/2kT}. \tag{3.17}$$

The Fermi energy in an intrinsic semiconductor, $E_i = E_F$, is given by

$$E_i = \frac{E_V + E_C}{2} + \frac{kT}{2} \ln\left(\frac{N_V}{N_C}\right) \tag{3.18}$$

which is typically very close to the middle of the bandgap. The intrinsic carrier concentration is typically very small compared with the densities of states and typical doping densities ($n_i \approx 10^{10}\,\text{cm}^{-3}$ in Si) and intrinsic semiconductors behave very much like insulators; that is, they are not good conductors of electricity.

The number of electrons and holes in their respective bands, and hence the conductivity of the semiconductor, can be controlled through the introduction of specific impurities, or dopants, called *donors* and *acceptors*. For example, when semiconductor silicon is doped with phosphorus, one electron is donated to the conduction band for each atom of phosphorus introduced. From Table 3.1, it can be seen that phosphorous is in column V of the periodic table of elements and thus has five valence electrons. Four of these are used to satisfy the four covalent bonds of the silicon lattice and the fifth is available to fill an empty state in the conduction band. If silicon is doped with boron (valency of three, since it is in column III), each boron atom accepts an electron from the valence band, leaving behind a hole. All impurities introduce additional localized electronic states into the band structure, often within the forbidden band between E_C and E_V, as illustrated in Figure 3.5. If the energy of the state E_D introduced by a donor atom is sufficiently close to the conduction bandedge (within a few kT), there will be sufficient thermal energy to allow the extra electron to occupy a state in the conduction band. The donor state will then be positively charged (ionized) and must be considered when analyzing the electrostatics of the situation. Similarly, an acceptor atom will introduce a negatively charged (ionized) state at energy E_A. The controlled introduction of donor and acceptor impurities into a semiconductor allows the creation of the n-type (electrons are the primary carriers of electrical current) and p-type (holes are the primary carriers of electrical current) semiconductors, respectively. This is the basis for the construction

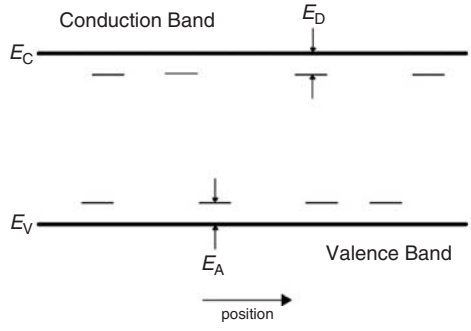


Figure 3.5 Donor and acceptor levels in a semiconductor. The nonuniform spatial distribution of these states reinforces the concept that these are localized states

of all semiconductor devices, including solar cells. The number of ionized donors and acceptors are given by [7]

$$N_{\mathrm{D}}^{+} = \frac{N_{\mathrm{D}}}{1 + g_{\mathrm{D}} \mathrm{e}^{(E_{\mathrm{F}} - E_{\mathrm{D}})/kT}} = \frac{N_{\mathrm{D}}}{1 + \mathrm{e}^{(E_{\mathrm{F}} - E'_{\mathrm{D}})/kT}} \tag{3.19}$$

and

$$N_{\mathrm{A}}^{-} = \frac{N_{\mathrm{A}}}{1 + g_{\mathrm{A}} \mathrm{e}^{(E_{\mathrm{A}} - E_{\mathrm{F}})/kT}} = \frac{N_{\mathrm{A}}}{1 + \mathrm{e}^{(E'_{\mathrm{A}} - E_{\mathrm{F}})/kT}} \tag{3.20}$$

where g_{D} and g_{A} are the donor and acceptor site degeneracy factors. Typically, $g_{\mathrm{D}} = 2$ and $g_{\mathrm{A}} = 4$. These factors are normally combined into the donor and the acceptor energies so that $E'_D = E_{\mathrm{D}} - kT \ln g_{\mathrm{D}}$ and $E'_A = E_{\mathrm{A}} + kT \ln g_{\mathrm{A}}$. Often, the donors and acceptors are assumed to be completely ionized so that $n_o \simeq N_D$ $n_o \simeq N_{\mathrm{D}}$ in n-type material and $p_o \simeq N_A$ in p-type material. The Fermi energy can then be written as

$$E_{\mathrm{F}} = E_{\mathrm{i}} + kT \ln \frac{N_{\mathrm{D}}}{n_{\mathrm{i}}} \tag{3.21}$$

in n-type material and as

$$E_{\mathrm{F}} = E_{\mathrm{i}} - kT \ln \frac{N_{\mathrm{A}}}{n_{\mathrm{i}}} \tag{3.22}$$

in p-type material.

When a very large concentration of dopants is introduced into the semiconductor, the dopants can no longer be thought of as a minor perturbation to the system. Their effect on the band structure must be considered. Typically, this so-called heavy doping effect manifests itself as a reduction in the bandgap, E_{G}, and thus an increase in the intrinsic carrier concentration, as can be seen from Equation (3.17). This bandgap narrowing (BGN) [8] is detrimental to solar cell performance and solar cells are typically designed to avoid this effect, though it may be a factor in the heavily doped regions near the solar cell contacts.

3.2.5 Light Absorption

The creation of electron–hole pairs via the absorption of sunlight is essential to the operation of solar cells. The excitation of an electron directly from the valence band (which leaves a hole behind)

to the conduction band is called *fundamental absorption*. Both the total energy and momentum of all particles involved in the absorption process must be conserved. Since the photon momentum, $p_\lambda = h/\lambda$, is very small compared with the range of the crystal momentum, $p = h/\ell$, the photon absorption process effectively conserves the momentum of the electron.[1] The absorption coefficient for a given photon energy, $h\nu$, is proportional to the probability, P_{12}, of the transition of an electron from the initial state E_1 to the final state E_2, the density of electrons in the initial state $g_V(E_1)$ and the density of available final states, and is then summed over all possible transitions between states where $E_2 - E_1 = h\nu$ [9],

$$\alpha(h\nu) \propto \sum P_{12} g_V(E_1) g_C(E_2), \tag{3.23}$$

assuming that all the valence-band states are full and all the conduction-band states are empty. Absorption results in creation of an electron–hole pair since a free electron excited into the conduction band leaves a free hole in the valence band.

In direct bandgap semiconductors, such as GaAs, GaInP, CdTe, and Cu(InGa)Se_2, the basic photon absorption process is illustrated in Figure 3.6. Both energy and momentum must be conserved in the transition. Every initial electron state with energy E_1 and crystal momentum p_1 in the valence band is associated with a final state in the conduction band at energy E_2 and crystal momentum p_2. Since the electron momentum is conserved, the crystal momentum of the final state is the same as the initial state, $p_1 \approx p_2 = p$.

Conservation of energy dictates that the energy of the absorbed photon is

$$h\nu = E_2 - E_1 \tag{3.24}$$

Since we have assumed parabolic bands,

$$E_V - E_1 = \frac{p^2}{2m_p^*} \tag{3.25}$$

and

$$E_2 - E_C = \frac{p^2}{2m_n^*} \tag{3.26}$$

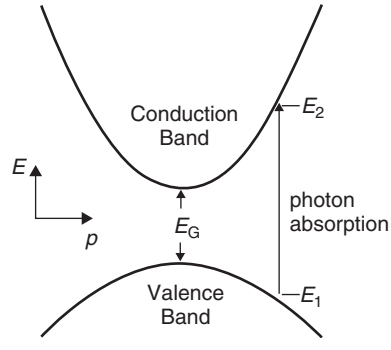


Figure 3.6 Photon absorption in a direct bandgap semiconductor for an incident photon with energy $h\nu = E_2 - E_1 > E_G$

[1] The wavelength of sunlight, λ, is of the order of a micrometer (10^{-4} cm), while the lattice constant is a few angstroms (10^{-8} cm). Thus, the crystal momentum is several orders of magnitude larger than the photon momentum.

Combining Equations (3.25), (3.26), and (3.27) yields

$$h\nu - E_G = \frac{p^2}{2}\left(\frac{1}{m_n^*} + \frac{1}{m_p^*}\right) \tag{3.27}$$

and the absorption coefficient for direct transitions is [9]

$$\alpha(h\nu) \approx A^*(h\nu - E_G)^{1/2}, \tag{3.28}$$

where A^* is a constant. In some semiconductor materials, quantum selection rules do not allow transitions at $p = 0$, but allow them for $p \neq 0$. In such cases [9]

$$\alpha(h\nu) \approx \frac{B^*}{h\nu}(h\nu - E_G)^{3/2}, \tag{3.29}$$

where B^* is a constant.

In indirect band gap semiconductors such as Si and Ge, where the valence-band maximum occurs at a different crystal momentum from that of the conduction-band minimum, conservation of electron momentum necessitates that the photon absorption process involve an additional particle. Phonons, the particle representation of lattice vibrations in the semiconductor, are suited to this process because they are low-energy particles with relatively high momentum. This is illustrated in Figure 3.7. Notice that light absorption is facilitated by either phonon absorption or phonon emission. The absorption coefficient, when there is phonon absorption, is given by

$$\alpha_a(h\nu) = \frac{A(h\nu - E_G + E_{ph})^2}{e^{E_{ph}/kT} - 1} \tag{3.30}$$

and by

$$\alpha_e(h\nu) = \frac{A(h\nu - E_G - E_{ph})^2}{1 - e^{-E_{ph}/kT}} \tag{3.31}$$

when a phonon is emitted [9]. Because both processes are possible,

$$\alpha(h\nu) = \alpha_a(h\nu) + \alpha_e(h\nu). \tag{3.32}$$

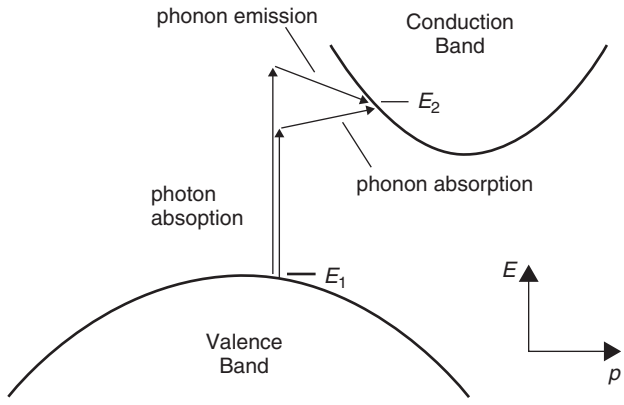


Figure 3.7 Photon absorption in an indirect bandgap semiconductor for a photon with energy $h\nu < E_2 - E_1$ and a photon with energy $h\nu > E_2 - E_1$. Energy and momentum in each case are conserved by the absorption and emission of a phonon, respectively

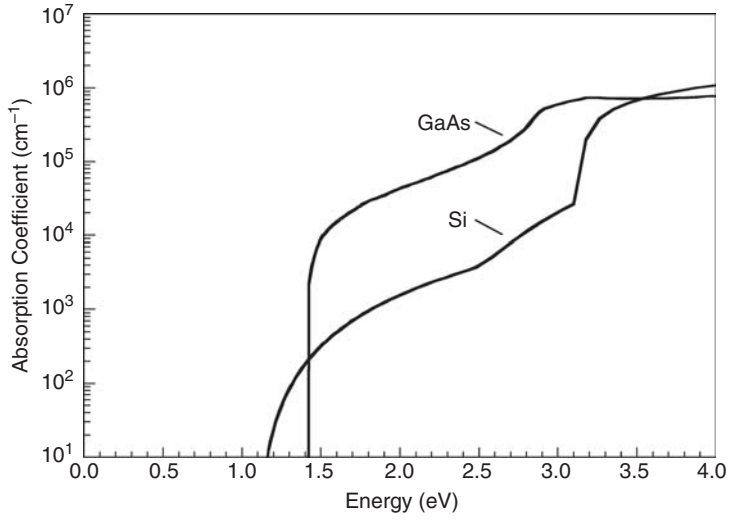


Figure 3.8 Absorption coefficient as a function of photon energy for Si (indirect bandgap) and GaAs (direct bandgap) at 300 K. Their bandgaps are 1.12 and 1.42 eV, respectively

Since both a phonon and an electron are needed to make the indirect gap absorption process possible, the absorption coefficient depends not only on the density of full initial electron states and empty final electron states but also on the availability of phonons (both emitted and absorbed) with the required momentum. Thus, compared with direct transitions, the absorption coefficient for indirect transitions is relatively small. As a result, light penetrates more deeply into indirect bandgap semiconductors than direct bandgap semiconductors. This is illustrated in Figure 3.8 for Si, an indirect bandgap semiconductor, and GaAs, a direct bandgap semiconductor. Similar spectra are shown for other semiconductors elsewhere in this handbook.

In both direct bandgap and indirect bandgap materials, a number of photon absorption processes are involved, though the mechanisms described above are the dominant ones. A direct transition, without phonon assistance, is possible in indirect bandgap materials if the photon energy is high enough (as seen in Figure 3.8 for Si at about 3.3 eV). Conversely, in direct bandgap materials, phonon-assisted absorption is also a possibility. Other mechanisms may also play a role in determining the optical absorption in semiconductors. These include absorption in the presence of an electric field (the Franz–Keldysh effect), absorption aided by localized states in the forbidden gap, and degeneracy effects when a significant number of states in the conduction band are not empty and/or when a significant number of state in the valence band are not full, as can happen in heavily doped materials (BGN) and under high-level injection (the Burstein–Moss shift). The net absorption coefficient is then the sum of the absorption coefficients due to all absorption processes or

$$\alpha(h\nu) = \sum_{\mathrm{i}} \alpha_{\mathrm{i}}(h\nu). \tag{3.33}$$

In practice, measured absorption coefficients or empirical expressions for the absorption coefficient are used in analysis and modeling. Chapter 17 has more details on extracting optical parameters from measurements and on the relation between optical and electric constants especially for thin film and conductive oxides, including heavily doped materials.

The rate of creation of electron–hole pairs (number of electron–hole pairs per cm^3 per second) as a function of position within a solar cell is

$$G(x) = (1 - s) \int_{\lambda} (1 - r(\lambda)) f(\lambda) \alpha(\lambda) e^{-\alpha x} \, d\lambda \tag{3.34}$$

where s is the grid-shadowing factor, $r(\lambda)$ is the reflectance, $\alpha(\lambda)$ is the absorption coefficient, and $f(\lambda)$ is the incident photon flux (number of photons incident per unit area per second per wavelength). The sunlight is assumed to be incident at $x = 0$. Here, the absorption coefficient has been cast in terms of the light's wavelength through the relationship $h\nu = hc/\lambda$. The photon flux, $f(\lambda)$, is obtained by dividing the incident power density at each wavelength by the photon energy.

Free-carrier absorption, in which electrons in the conduction band absorb the energy of a photon and move to an empty state higher in the conduction band (correspondingly for holes in the valence band), is typically only significant for photons with $E < E_G$ since the free-carrier absorption coefficient increases with increasing wavelength,

$$\alpha_{fc} \propto \lambda^{\gamma} \tag{3.35}$$

where $1.5 < \gamma < 3.5$ [9]. Thus, in single-junction solar cells, it does not affect the creation of electron–hole pairs and can be ignored (although free-carrier absorption can be exploited to probe the excess carrier concentrations in solar cells for the purpose of determining recombination parameters [10]). However, free-carrier absorption is a consideration in tandem solar cell systems in which a wide bandgap (E_{G1}) solar cell is stacked on top of a solar cell of smaller bandgap ($E_{G2} < E_{G1}$). Photons with energy too low to be absorbed in the top cell ($h\nu < E_{G1}$) will be transmitted to the bottom cell and be absorbed there (if $h\nu > E_{G2}$). Of course, more solar cells can be stacked as long as $E_{G1} > E_{G2} > E_{G3} \ldots$, and so on. The number of photons transmitted to the next cell in the stack will be reduced by whatever amount of free-carrier absorption occurs. This loss can be avoided by splitting the incident spectrum and directing the matched portion of the spectrum to each component solar cell of a multijuction system [11]. Multijunction solar cells are discussed more completely in Chapters 8 and 12.

3.2.6 Recombination

When a semiconductor is taken out of thermal equilibrium, for instance by illumination and/or the injection of current, the concentrations of electrons (n) and holes (p) tend to relax back toward their equilibrium values through a process called *recombination* in which an electron falls from the conduction band to the valence band, thereby eliminating a valence-band hole. There are several recombination mechanisms important to the operation of solar cells – recombination through traps (defects) in the forbidden gap, radiative (band-to-band) recombination, and Auger recombination – that will be discussed here. These three processes are illustrated in Figure 3.9.

The net recombination rate per unit volume per second through a single level trap (SLT) located at energy $E = E_T$ within the forbidden gap, also commonly referred to as *Shockley–Read–Hall recombination*, is given by [12]

$$R_{SLT} = \frac{pn - n_i^2}{\tau_{SLT,n}(p + n_i e^{(E_i - E_T)/kT}) + \tau_{SLT,p}(n + n_i e^{(E_T - E_i)/kT})}. \tag{3.36}$$

The carrier lifetimes are given by

$$\tau_{SLT} = \frac{1}{\sigma v_{th} N_T} \tag{3.37}$$

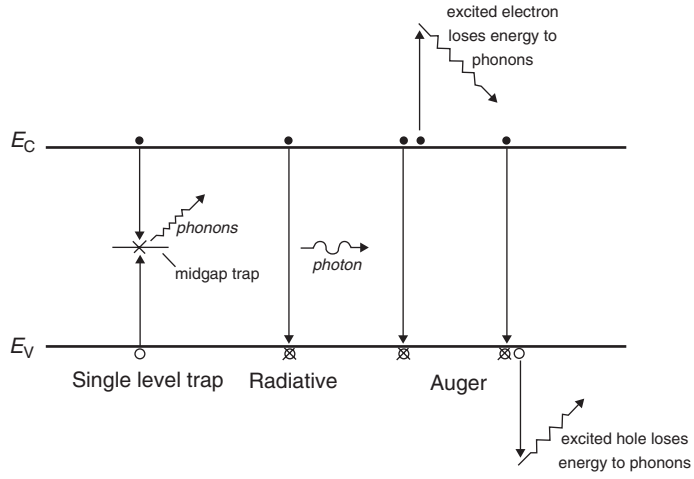


Figure 3.9 Recombination processes in semiconductors

where σ is the capture cross-section (σ_n for electrons and σ_p for holes), v_{th} is the thermal velocity of the carriers, and N_{T} is the concentration of traps. The capture cross-section can be thought of as the size of the target presented to a carrier traveling through the semiconductor at velocity v_{th}. Small lifetimes correspond to high rates of recombination. If a trap presents a large target to the carrier, the recombination rate will be high (low carrier lifetime). When the velocity of the carrier is high, it has more opportunity within a given time period to encounter a trap and the carrier lifetime is low. Finally, the probability of interaction with a trap increases as the concentration of traps increases and the carrier lifetime is therefore inversely proportional to the trap concentration.

Some reasonable assumptions allow Equation (3.36) to be simplified. If the material is p-type ($p \approx p_o \gg n_o$), in low injection ($n_o \leq n \ll p_o$), and the trap energy is near the middle of the forbidden gap ($E_{\text{T}} \approx E_{\text{i}}$), the recombination rate can be written as

$$R_{\text{SLT}} \approx \frac{n - n_o}{\tau_{\text{SLT},n}}. \tag{3.38}$$

Notice that the recombination rate is solely dependent on the minority carrier. This is reasonable since there are far fewer minority carriers than majority carriers and one of each is necessary for there to be recombination.

If high-injection conditions prevail ($p \approx n \gg p_o, n_o$),

$$R_{\text{SLT}} \approx \frac{n}{\tau_{\text{SLT},p} + \tau_{\text{SLT},n}} \approx \frac{p}{\tau_{\text{SLT},p} + \tau_{\text{SLT},n}}. \tag{3.39}$$

In this case, the effective recombination lifetime is the sum of the two carrier lifetimes. While the recombination rate is high due to the large number of excess holes and electrons, the carrier lifetime is actually longer than in the case of low injection. This can be of significance in the base region of solar cells, especially concentrator cells (solar cells illuminated with concentrated sunlight), since the base is the least doped layer.

Radiative (band-to-band) recombination is simply the inverse of the optical generation process and is much more efficient in direct bandgap semiconductors than in indirect bandgap

semiconductors. When radiative recombination occurs, the energy of the electron is given to an emitted photon – this is how semiconductor lasers and light emitting diodes (LEDs) operate. In an indirect bandgap material, some of that energy is shared with a phonon. The net recombination rate due to radiative processes is given as

$$R_\lambda = \mathrm{B}(pn - n_\mathrm{i}^2) \tag{3.40}$$

If we have an n-type ($n \approx n_o \gg p_o$) semiconductor in low injection ($p_o \leq p \ll n_o$), the net radiative recombination rate can be written in terms of an effective lifetime, $\tau_{\lambda,p}$,

$$R_\lambda \approx \frac{p - p_o}{\tau_{\lambda,p}} \tag{3.41}$$

where

$$\tau_{\lambda,p} = \frac{1}{n_o \mathrm{B}}. \tag{3.42}$$

A similar expression can be written for p-type semiconductors. If high-injection conditions prevail ($p \approx n \gg p_o, n_o$), then

$$R_\lambda \approx Bp^2 \approx Bn^2. \tag{3.43}$$

Since photons with energies near that of the bandgap are emitted during this recombination process, it is possible for these photons to be reabsorbed before exiting the semiconductor. A well-designed direct bandgap solar cell can take advantage of this *photon recycling* and increase the effective lifetime [13].

Auger recombination is somewhat similar to radiative recombination, except that the energy of transition is given to another carrier (in either the conduction band or the valence band), as shown in Figure 3.9. This electron (or hole) then relaxes thermally (releasing its excess energy and momentum to phonons). Just as radiative recombination is the inverse process to optical absorption, Auger recombination is the inverse process to *impact ionization*, where an energetic electron collides with a crystal atom, breaking the bond and creating an electron–hole pair. The net recombination rate due to Auger processes is

$$R_\mathrm{Auger} = (C_n n + C_p p)(pn - n_\mathrm{i}^2) \tag{3.44}$$

In an n-type material in low injection (and assuming C_n and C_p are of comparable magnitudes), the net Auger recombination rate becomes

$$R_\mathrm{Auger} \approx \frac{p - p_o}{\tau_{\mathrm{Auger},p}} \tag{3.45}$$

with

$$\tau_{\mathrm{Auger},p} = \frac{1}{C_n n_o^2}. \tag{3.46}$$

A similar expression can be derived for minority electron lifetime in p-type material. If high-injection conditions prevail ($p \approx n \gg p_o, n_o$), then

$$R_\mathrm{Auger} \approx (C_n + C_p)p^3 \approx (C_n + C_p)n^3 \tag{3.47}$$

While the SLT recombination rate can be minimized by reducing the density of single-level traps and the radiative recombination rate can be minimized via photon recycling, the Auger recombination rate is a fundamental property of the semiconductor.

Each of these recombination processes occurs in parallel. And, there can be multiple and/or distributed traps[2] in the forbidden gap – in which case the net recombination is a sum of the contributions of each trap ($\sum_{\text{traps } i} R_{\text{SLT},i}$). Thus, the total recombination rate is the sum of rates due to each process

$$R = \left[\sum_{\text{traps } i} R_{\text{SLT},i} \right] + R_{\lambda} + R_{\text{Auger}}. \tag{3.48}$$

An effective minority-carrier lifetime for a doped material in low-level injection is given as

$$\frac{1}{\tau} = \left[\sum_{\text{traps } i} \frac{1}{\tau_{\text{SLT},i}} \right] + \frac{1}{\tau_{\lambda}} + \frac{1}{\tau_{\text{Auger}}}. \tag{3.49}$$

The distribution of traps in the energy gap for semiconductor materials can be influenced by the specific growth or processing conditions, impurities, and crystallographic defects.

Interfaces between two dissimilar materials, such as those that occur at the front surface of a solar cell, have a high concentration of defects due to the abrupt termination of the crystal lattice. These manifest themselves as a continuum of traps (*surface states)* within the forbidden gap at the surface and electrons and holes can recombine through them just as with bulk traps. These surface states are illustrated in Figure 3.10. Rather than giving a recombination rate per unit volume per second, surface states give a recombination rate per unit area per second. A general expression for surface recombination is [12]

$$R_{\text{S}} = \int_{E_{\text{V}}}^{E_{\text{C}}} \frac{pn - n_{\text{i}}^2}{(p + n_{\text{i}} e^{(E_{\text{i}} - E_{\text{t}})/kT})/s_n(E_{\text{t}}) + (n + n_{\text{i}} e^{(E_{\text{t}} - E_{\text{i}})/kT})/s_p(E_{\text{t}})} D_{\Pi}(E_{\text{t}})\, \text{d}E_{\text{t}} \tag{3.50}$$

where E_{t} is the trap energy, $D_{\Pi}(E_{\text{t}})$ is the surface state (the concentration of traps is probably varies with trap energy), and $s_n(E_{\text{t}})$ and $s_p(E_{\text{t}})$ are surface recombination velocities, analogous to the carrier lifetimes for bulk traps. The surface recombination rate is generally written, for simplicity, as [12]

$$R_{\text{S}} = S_p(p - p_o) \tag{3.51}$$

in n-type material and as

$$R_{\text{S}} = S_n(n - n_o) \tag{3.52}$$

in p-type material. S_p and S_n are effective surface recombination velocities. It should be mentioned that these effective recombination velocities are not necessarily constants independent of carrier concentration, though they are commonly treated as such.

[2] It is unlikely that more than one trap will be involved in a single recombination event since the traps are spatially separated.

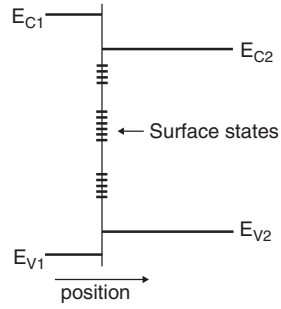


Figure 3.10 Illustration of surface states at a semiconductor surface or interface between dissimilar materials such as a semiconductor and an insulator (i.e., antireflective coating), two different semiconductors (heterojunction) or a metal and a semiconductor (Schottky contact)

3.2.7 Carrier Transport

As has already been established, electrons and holes in a semiconductor behave much like a free particle of the same electronic charge with effective masses of m_n^* and m_p^*, respectively. Thus, they are subject to the classical processes of drift and diffusion. Drift is a charged particle's response to an applied electric field. When an electric field is applied across a uniformly doped semiconductor, the bands bend upward in the direction of the applied electric field. Electrons in the conduction band, being negatively charged, move in the opposite direction to the applied field and holes in the valence band, being positively charged, move in the same direction as the applied field (Figure 3.11) – in other words, electrons *sink* and holes *float*. This is a useful conceptual tool for analyzing the motion of holes and electrons in semiconductor devices.

With nothing to impede their motion, the holes and electrons would continue to accelerate without bound. However, the semiconductor crystal is full of objects with which the carriers collide and are scattered. These objects include the component atoms of the crystal, dopant ions, crystal defects, and even other electrons and holes. On a microscopic scale, their motion is much like that of a ball in pinball machine, the carriers are constantly bouncing (scattering) off objects in the crystal,

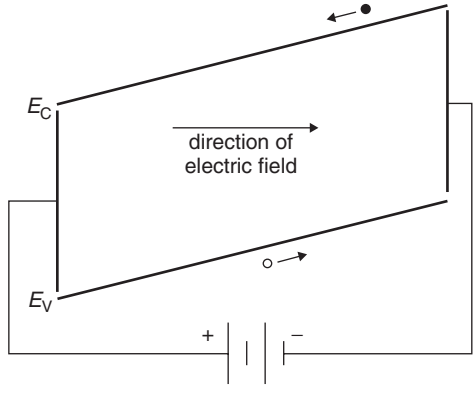


Figure 3.11 Illustration of the concept of drift in a semiconductor. Note that electrons and holes move in opposite directions. The electric field can be created by the internal built-in potential of the junction or by an externally applied bias voltage

but generally moving in the direction prescribed by the applied electric field, $\vec{E} = -\nabla\phi$, where ϕ is the electrostatic potential. The net effect is that the carriers appear to move, on a macroscopic scale, at a constant velocity, v_{d}, the drift velocity. The drift velocity is directly proportional to the electric field

$$|\vec{v}_d| = \left|\mu\vec{E}\right| = |\mu\nabla\phi| \tag{3.53}$$

where μ is the carrier mobility. The carrier mobility is generally independent of the electric field strength unless the field is very strong, a situation not typically encountered in solar cells. The drift current densities for holes and electrons can be written as

$$\vec{J}_p^{\mathrm{drift}} = qp\vec{v}_{\mathrm{d},p} = q\mu_p p\vec{E} = -q\mu_p p\nabla\phi \tag{3.54}$$

and

$$\vec{J}_n^{\mathrm{drift}} = -qn\vec{v}_{\mathrm{d},n} = q\mu_n n\vec{E} = -q\mu_n n\nabla\phi. \tag{3.55}$$

The most significant scattering mechanisms in solar cells are lattice (phonon) and ionized impurity scattering. These component mobilities can be written as

$$\mu_{\mathrm{L}} = C_{\mathrm{L}}T^{-3/2} \tag{3.56}$$

for lattice scattering and as

$$\mu_{\mathrm{I}} = \frac{C_{\mathrm{I}}T^{3/2}}{N_{\mathrm{D}}^{+} + N_{\mathrm{A}}^{-}} \tag{3.57}$$

for ionized impurity scattering. These can then be combined using Mathiessen's rule to give the carrier mobility [14]

$$\frac{1}{\mu} = \frac{1}{\mu_{\mathrm{L}}} + \frac{1}{\mu_{\mathrm{I}}}. \tag{3.58}$$

This is a first-order approximation that neglects the velocity dependencies of the scattering mechanisms. These two types of mobility can be distinguished experimentally by their different dependencies on temperature and doping. A better approximation is [14]

$$\mu = \mu_{\mathrm{L}}\left[1 + \left(\frac{6\mu_{\mathrm{L}}}{\mu_{\mathrm{I}}}\right)\left(\mathrm{Ci}\left(\frac{6\mu_{\mathrm{L}}}{\mu_{\mathrm{I}}}\right)\cos\left(\frac{6\mu_{\mathrm{L}}}{\mu_{\mathrm{I}}}\right) + \left[\mathrm{Si}\left(\frac{6\mu_{\mathrm{L}}}{\mu_{\mathrm{I}}}\right) - \frac{\pi}{2}\right]\sin\left(\frac{6\mu_{\mathrm{L}}}{\mu_{\mathrm{I}}}\right)\right)\right], \tag{3.59}$$

where Ci and Si (not to be confused with the abbreviation for silicon) are the cosine and sine integrals, respectively.

When modeling solar cells, it is more convenient to use measured data or empirical formulas. Carrier mobilities in Si at 300 K are well approximated by [14]

$$\mu_n = 92 + \frac{1268}{1 + \left(\dfrac{N_{\mathrm{D}}^{+} + N_{\mathrm{A}}^{-}}{1.3 \times 10^{17}}\right)^{0.91}}\mathrm{cm}^2/\mathrm{V} - \mathrm{s} \tag{3.60}$$

$$\mu_p = 54.3 + \frac{406.9}{1 + \left(\dfrac{N_{\mathrm{D}}^{+} + N_{\mathrm{A}}^{-}}{2.35 \times 10^{17}}\right)^{0.88}}\mathrm{cm}^2/\mathrm{V} - \mathrm{s} \tag{3.61}$$

and are plotted in Figure 3.12. At low impurity levels, the mobility is governed by intrinsic lattice scattering, while at high levels the mobility is governed by ionized impurity scattering.

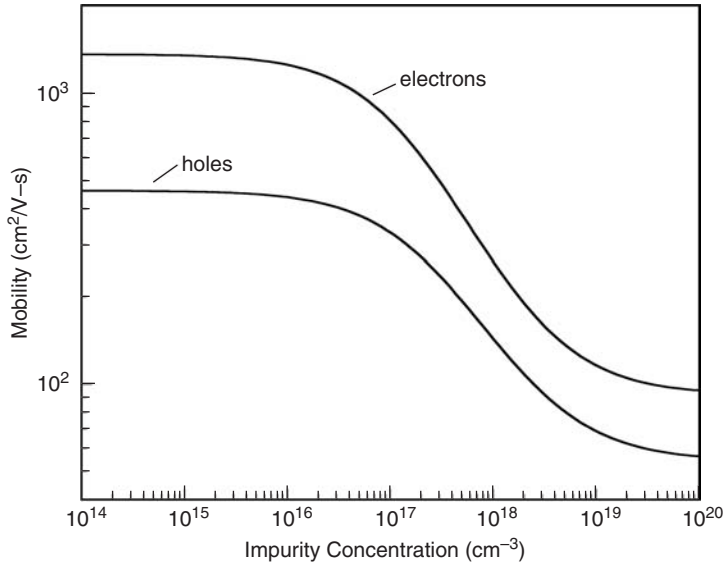


Figure 3.12 Electron and hole mobilities in silicon for $T = 300\,\mathrm{K}$

Electrons and holes in semiconductors tend, as a result of their random thermal motion, to move (diffuse) from regions of high concentration to regions of low concentration. Much like the way the air in a balloon is distributed evenly within the volume of the balloon, carriers, in the absence of any external forces, will also tend to distribute themselves evenly within the semiconductor. This process is called *diffusion* and the diffusion current densities are given by

$$\vec{J}_p^{\,\mathrm{diff}} = -qD_p \nabla p \tag{3.62}$$

$$\vec{J}_n^{\,\mathrm{diff}} = qD_n \nabla n \tag{3.63}$$

where D_p and D_n are the hole and electron diffusion coefficients, respectively. Note that the currents are driven by the gradient of the carrier densities.

In thermal equilibrium, there can be no net hole current and no net electron current – in other words, the drift and diffusion currents must exactly balance. In nondegenerate materials, this leads to the Einstein relationship

$$\frac{D}{\mu} = \frac{kT}{q} \tag{3.64}$$

and allows the diffusion coefficient to be directly computed from the mobility. Generalized forms of the Einstein relationship, valid for degenerate materials, are

$$\frac{D_n}{\mu_n} = \frac{1}{q} n \left[\frac{\mathrm{d}n}{\mathrm{d}E_\mathrm{F}}\right]^{-1} \tag{3.65}$$

and

$$\frac{D_p}{\mu_p} = -\frac{1}{q} p \left[\frac{\mathrm{d}p}{\mathrm{d}E_\mathrm{F}}\right]^{-1}. \tag{3.66}$$

The diffusion coefficient actually increases when degeneracy effects come into play.

The total hole and electron currents (vector quantities) are the sum of their drift and diffusion components

$$\vec{J}_p = \vec{J}_p^{\text{drift}} + \vec{J}_p^{\text{diff}} = q\mu_p p\vec{E} - qD_p\nabla p = -q\mu_p p\nabla\phi - qD_p\nabla p \tag{3.67}$$

$$\vec{J}_n = \vec{J}_n^{\text{drift}} + \vec{J}_n^{\text{diff}} = q\mu_n n\vec{E} + qD_n\nabla n = -q\mu_n n\nabla\phi + qD_n\nabla n \tag{3.68}$$

The total current is then

$$\vec{J} = \vec{J}_p + \vec{J}_n + \vec{J}_{\text{disp}} \tag{3.69}$$

where $\vec{J}_{\text{disp}}$ is the *displacement current* given by

$$\vec{J}_{\text{disp}} = \frac{\partial\vec{D}}{\partial t}. \tag{3.70}$$

$\vec{D} = \varepsilon\vec{E}$ is the dielectric displacement field, where ε is the electric permittivity of the semiconductor. The displacement current can be neglected in solar cells since they are static (dc) devices.

3.2.8 Semiconductor Equations

The operation of most semiconductor devices, including solar cells, can be described by the so-called semiconductor device equations, first described by Van Roosbroeck in 1950 [15]. A generalized form of these equations is given here.[3]

$$\nabla\cdot\varepsilon\vec{E} = q(p - n + N) \tag{3.71}$$

This is a form of Poisson's equation, where N is the net charge due to dopants and other trapped charges. The hole and electron continuity equations are

$$\nabla\cdot\vec{J}_p = q\left(G - R_p - \frac{\partial p}{\partial t}\right) \tag{3.72}$$

$$\nabla\cdot\vec{J}_n = q\left(R_n - G + \frac{\partial n}{\partial t}\right) \tag{3.73}$$

where G is the optical generation rate of electron–hole pairs. Thermal generation is included in R_p and R_n. The hole and electron current densities are given by (Equations 3.67 and 3.68)

$$\vec{J}_p = -q\mu_p p\nabla(\phi - \phi_p) - kT\mu_p\nabla p \tag{3.74}$$

and

$$\vec{J}_n = -q\mu_n n\nabla(\varphi + \varphi_n) + kT\mu_n\nabla n. \tag{3.75}$$

Two new terms, ϕ_p and ϕ_n, have been introduced here. These are the so-called band parameters that account for degeneracy and a spatially varying bandgap (heterostructure solar cells) and electron affinity [17]. These terms were ignored in the preceding discussion and can usually be ignored in nondegenerate homostructure solar cells.

[3] In some photovoltaic materials such as GaInN, polarization is important and Poisson's equation becomes $\nabla\cdot(\varepsilon\vec{E} + \vec{P}) = q(p - n + N)$, where $\vec{P}$ is the polarization [16].

The intent here is to derive a simple analytic expression for the current–voltage characteristic of a solar cell, and so some simplifications are in order. It should be noted, however, that a complete description of the operation of solar cells can be obtained by solving the full set of coupled partial differential equations, Equations (3.71–3.75). The numerical solution of these equations is briefly addressed later in this chapter.

3.2.9 Minority-carrier Diffusion Equation

In a uniformly doped semiconductor, the bandgap and electric permittivity are independent of position. Since the doping is uniform, the carrier mobilities and diffusion coefficients are also independent of position. As we are mainly interested in the steady-state operation of the solar cell, the semiconductor equations reduce to

$$\frac{\mathrm{d}\vec{E}}{\mathrm{d}x} = \frac{q}{\varepsilon}(p - n + N_{\mathrm{D}} - N_{\mathrm{A}}) \tag{3.76}$$

$$q\mu_p \frac{\mathrm{d}}{\mathrm{d}x}(p\vec{E}) - qD_p \frac{\mathrm{d}^2 p}{\mathrm{d}x^2} = q(G - R) \tag{3.77}$$

and

$$q\mu_n \frac{\mathrm{d}}{\mathrm{d}x}(n\vec{E}) + qD_n \frac{\mathrm{d}^2 n}{\mathrm{d}x^2} = q(R - G) \tag{3.78}$$

In regions sufficiently far from the pn-junction of the solar cell (quasi-neutral regions), the electric field is very small. When considering the minority carrier (holes in the n-type material and electrons in the p-type material) and low-level injection ($\Delta p = \Delta n \ll N_{\mathrm{D}}, N_{\mathrm{A}}$), the drift current can be neglected with respect to the diffusion current. Under low-level injection, R simplifies to

$$R = \frac{n_P - n_{Po}}{\tau_n} = \frac{\Delta n_P}{\tau_n} \tag{3.79}$$

in the p-type region and to

$$R = \frac{p_N - p_{No}}{\tau_p} = \frac{\Delta p_N}{\tau_p} \tag{3.80}$$

in the n-type region. Δp_N and Δn_P are the excess minority-carrier concentrations. The minority-carrier lifetimes, τ_n and τ_p, are given by Equation (3.49). For clarity, the capitalized subscripts, P and N, are used to indicate quantities in p-type and n-type regions, respectively, when it may not be otherwise apparent. Lowercase subscripts, p and n, refer to quantities associated with minority holes and electrons, respectively. For example, Δn_P is the minority electron concentration in the p-type material.

Thus, Equations (3.77) and (3.78) each reduce to what is commonly referred to as the *minority-carrier diffusion equation*. It can be written as

$$D_p \frac{\mathrm{d}^2 \Delta p_N}{\mathrm{d}x^2} - \frac{\Delta p_N}{\tau_p} = -G(x) \tag{3.81}$$

in n-type material and as

$$D_n \frac{\mathrm{d}^2 \Delta n_P}{\mathrm{d}x^2} - \frac{\Delta n_P}{\tau_n} = -G(x) \tag{3.82}$$

in p-type material. The minority-carrier diffusion equation is often used to analyze the operation of semiconductor devices, including solar cells, and will be used in this way later in this chapter.

3.2.10 *pn*-junction Diode Electrostatics

Where an n-type semiconductor comes into contact with a p-type semiconductor, a pn-junction is formed. In thermal equilibrium there is no net current flow and by definition the Fermi energy must be independent of position. Since there is a concentration difference of holes and electrons between the two types of semiconductors, holes diffuse from the p-type region into the n-type region and, similarly, electrons from the n-type material diffuse into the p-type region. As the carriers diffuse, the charged impurities (ionized acceptors in the p-type material and ionized donors in the n-type material) are uncovered – that is, they are no longer screened by the majority carrier. As these impurity charges are uncovered, an electric field (or electrostatic potential difference) is produced, which counteracts the diffusion of the holes and electrons. In thermal equilibrium, the diffusion and drift currents for each carrier type exactly balance, so there is no net current flow. The transition region between the n-type and the p-type semiconductors is called the *space-charge region*. It is also often called the *depletion region*, since it is effectively depleted of both holes and electrons. Assuming that the p-type and the n-type regions are sufficiently thick, the regions on either side of the depletion region are essentially charge-neutral (often termed *quasi-neutral*). The electrostatic potential difference resulting from the junction formation is called the *built-in voltage*, V_{bi}. It arises from the electric field created by the exposure of the positive and the negative space charge in the depletion region.

The electrostatics of this situation (assuming a single acceptor and a single donor level) are governed by Poisson's equation

$$\nabla^2\phi = \frac{q}{\varepsilon}(n_o - p_o + N_{\text{A}}^- - N_{\text{D}}^+) \tag{3.83}$$

where ϕ is the electrostatic potential, q is magnitude of the electron charge, ε is the electric permittivity of the semiconductor, p_o is the equilibrium hole concentration, n_o is the equilibrium electron concentration, N_{A}^- is the ionized acceptor concentration, and N_D^+ is the ionized donor concentration. Equation (3.83) is a restatement of Equation (3.71) for the given conditions.

This equation is easily solved numerically; however, an approximate analytic solution for an abrupt pn-junction can be obtained that lends physical insight into the formation of the space-charge region. Figure 3.13 depicts a simple one-dimensional (1D) pn-junction solar cell (diode), with the metallurgical junction at $x = 0$, which is uniformly doped, with a doping density of N_{D} on the n-type side and of N_{A} on the p-type side. For simplicity, it is assumed that the each side is nondegenerately doped and that the dopants are fully ionized. In this example, the n-type side is assumed to be more heavily doped (n^+) than the p-type side.

Within the depletion region, defined by $-x_N < x < x_P$, it can be assumed that p_o and n_o are both negligible compared to $|N_{\text{A}} - N_{\text{D}}|$ so that Equation (3.83) can be simplified to

$$\begin{aligned} \nabla^2\phi &= -\frac{q}{\varepsilon}N_{\text{D}}, \text{ for } -x_N < x < 0 \text{ and} \\ \nabla^2\phi &= \frac{q}{\varepsilon}N_{\text{A}}, \quad \text{for } 0 < x < x_P \end{aligned} \tag{3.84}$$

Outside the depletion region, charge neutrality is assumed and

$$\nabla^2\phi = 0, \text{ for } x \le -x_N \text{ and } x \ge x_P. \tag{3.85}$$

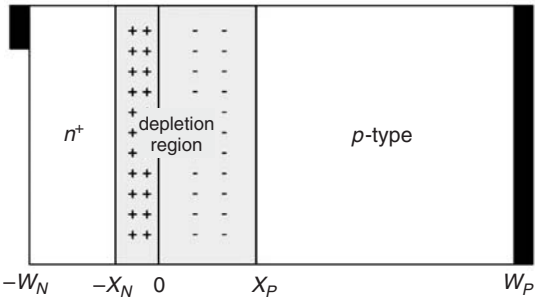


Figure 3.13 Simple solar cell structure used to analyze the operation of a solar cell. Free carriers have diffused across the junction ($x = 0$) leaving a space-charge or depletion region practically devoid of any free or mobile charges. The fixed charges in the depletion region are due to ionized donors on the n-side and ionized acceptors on the p-side

This is commonly referred to as the *depletion approximation*. The regions on either side of the depletion regions are the quasi-neutral regions.

The electrostatic potential difference across the junction is the built-in voltage, V_{bi}, and can be obtained by integrating the electric field, $\vec{E} = -\nabla\phi$.

$$\int_{-x_N}^{x_P} \vec{E}\,\mathrm{d}x = -\int_{-x_N}^{x_P} \frac{\mathrm{d}\phi}{\mathrm{d}x}\mathrm{d}x = -\int_{V(-x_N)}^{V(x_P)} \mathrm{d}\phi = \phi(-x_N) - \phi(x_P) = V_{\text{bi}} \tag{3.86}$$

Solving Equations (3.84) and (3.85) and defining $\phi(x_P) = 0$, gives

$$\phi(x) = \begin{cases} V_{\text{bi}}, & x \le -x_N \\ V_{\text{bi}} - \dfrac{qN_{\text{D}}}{2\varepsilon}(x + x_N)^2, & -x_N < x \le 0 \\ \dfrac{qN_{\text{A}}}{2\varepsilon}(x - x_P)^2, & 0 \le x < x_P \\ 0, & x \ge x_P \end{cases} \tag{3.87}$$

The electrostatic potential must be continuous at $x = 0$. Therefore, from Equation (3.87),

$$V_{\text{bi}} - \frac{qN_{\text{D}}}{2\varepsilon}x_N^2 = \frac{qN_{\text{A}}}{2\varepsilon}x_P^2 \tag{3.88}$$

In the absence of any interface charge at the metallurgical junction, the electric field is also continuous at this point (really, it is the displacement field, $\vec{D} = \varepsilon\vec{E}$, that is continuous, but in this example, ε is independent of position), and

$$x_N N_{\text{D}} = x_P N_{\text{A}} \tag{3.89}$$

This is simply a statement that the total charge in either side of the depletion region exactly balance each other and therefore the depletion region extends furthest into the more lightly doped side.

Solving Equations (3.88) and (3.89) for the depletion width, W_D, gives[4]

$$W_D = x_N + x_P = \sqrt{\frac{2\varepsilon}{q}\left(\frac{N_A + N_D}{N_A N_D}\right) V_{bi}}. \tag{3.90}$$

Under nonequilibrium conditions, the electrostatic potential difference across the junction is modified by the applied voltage V which is zero in thermal equilibrium. As a consequence, the depletion width is dependent on the applied voltage,

$$W_D(V) = x_N + x_P = \sqrt{\frac{2\varepsilon}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)(V_{bi} - V)}. \tag{3.91}$$

As previously stated, the built-in voltage, V_{bi}, can be calculated by noting that, under thermal equilibrium, the net hole and electron currents are zero. The hole current density is

$$\vec{J}_p = q\mu_p p_o \vec{E} - qD_p \nabla p = 0. \tag{3.92}$$

Thus, in 1D and utilizing the Einstein relationship, the electric field can be written as

$$\vec{E} = \frac{kT}{q}\frac{1}{p_o}\frac{dp_o}{dx} \tag{3.93}$$

Rewriting Equation (3.86) and substituting Equation (3.93) yields

$$V_{bi} = \int_{-x_N}^{x_P} E\vec{d}x = \int_{-x_N}^{x_P} \frac{kT}{q}\frac{1}{p_o}\frac{dp_o}{dx}dx = \frac{kT}{q}\int_{p_o(-x_N)}^{p_o(x_P)} \frac{dp_o}{p_o} = \frac{kT}{q}\ln\left[\frac{p_o(x_P)}{p_o(-x_N)}\right] \tag{3.94}$$

Since we have assumed nondegeneracy, $p_o(x_P) = N_A$ and $p_o(-x_N) = n_i^2/N_D$. Therefore,

$$V_{bi} = \frac{kT}{q}\ln\left[\frac{N_D N_A}{n_i^2}\right]. \tag{3.95}$$

Figure 3.14 shows the equilibrium energy band diagram (a), electric field (b), and charge density (c) for a simple abrupt *pn*-junction silicon diode in the vicinity of the depletion region. The conduction band edge is given by $E_C(x) = E_0 - q\phi(x) - \chi$, the valence band edge by $E_V(x) = E_C(x) - E_G$, and the intrinsic energy by Equation (3.18). E_0, defined as the vacuum energy, serves as a convenient reference point and is universally constant with position. An electron at the vacuum energy is, by definition, completely free of influence from all external forces. The electron affinity χ is the minimum energy needed to free an electron from the bottom of the conduction band and take it to the vacuum level. The electric field is a result of the uncovered ionized donors and acceptors, and opposes the diffusion of electrons and holes in the quasi-neutral regions. The charge density plot illustrates the balance of charge between the two sides

[4] A somewhat more rigorous treatment of equation 3.89 would yield a factor of $2kT/q$ which is ~50 mV at 300 K, or

$$W_D = \sqrt{\frac{2\varepsilon}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)(V_{bi} - 2kT/q)} \quad [3].$$

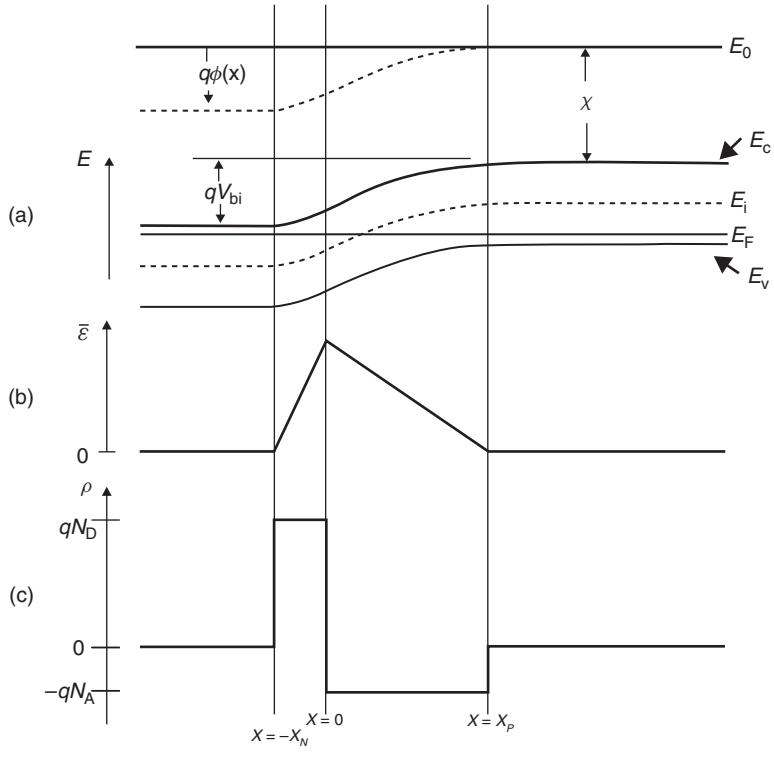


Figure 3.14 Equilibrium conditions in a solar cell: (a) energy bands; (b) electric field; (c) charge density

of the depletion region. In heterostructures, both the bandgap and the electron affinity are position-dependent – making the calculation of the junction electrostatics and energy band diagram more complex, as discussed in Section 3.4.8.

3.2.11 Summary

The fundamental physical principles relevant to solar cell operation have been reviewed and the basic solar cell structure has now been established (Figures 3.1 and 3.13). A solar cell is simply a *pn*-junction diode consisting of two quasi-neutral regions on either side of a depletion region with an electrical contact made to each quasi-neutral region. Typically, the more heavily doped quasi-neutral region is called the *emitter* (the *n*-type region in Figure 3.13) and the more lightly doped region is called the *base* (the *p*-type region in Figure 3.13). The base region is also often referred to as the *absorber region* since the emitter region is usually very thin and most of the light absorption occurs in the base. This basic structure will now serve as the basis for deriving the fundamental operating characteristics of the solar cell.

3.3 SOLAR CELL FUNDAMENTALS

The basic current–voltage characteristic of the solar cell can be derived by solving the minority-carrier diffusion equation with appropriate boundary conditions.

3.3.1 Solar Cell Boundary Conditions

In Figure 3.13, at $x = -W_N$, the usual assumption is that the front contact can be treated as an ideal ohmic contact, i.e.

$$\Delta p(-W_N) = 0. \tag{3.96}$$

However, since the front contact is usually a grid with metal contacting the semiconductor on only a small percentage of the front surface, modeling the front surface with an effective surface recombination velocity is more realistic. This effective recombination velocity models the combined effects of the ohmic contact and the antireflective passivation layer (SiO_2 in silicon solar cells). In this case, the boundary condition at $x = -W_N$ is

$$\frac{d\Delta p}{dx} = \frac{S_{F,\text{eff}}}{D_p}\Delta p(-W_N) \tag{3.97}$$

where $S_{F,\text{eff}}$ is the effective front surface recombination velocity. As $S_{F,\text{eff}} \to \infty$, $\Delta p \to 0$, and the boundary condition given by Equation (3.97) reduces to that of an ideal ohmic contact (Equation 3.96). In reality, $S_{F,\text{eff}}$ depends upon a number of parameters and is bias dependent. This will be discussed in more detail later.

The back contact can also be treated as an ideal ohmic contact, so that

$$\Delta n(W_P) = 0. \tag{3.98}$$

However, solar cells are often fabricated with a *back-surface field* (BSF), a thin, more heavily doped region at the back of the base region. An even more effective BSF can be created by inserting a wider bandgap semiconductor material at the back of the solar cell (a heterojunction). The BSF keeps minority carriers away from the back ohmic contact and increases their chances of being collected and it can be modeled by an effective, and relatively low, surface recombination velocity. This boundary condition is then

$$\left.\frac{d\Delta n}{dx}\right|_{x=W_P} = -\frac{S_{BSF}}{D_n}\Delta n(W_P), \tag{3.99}$$

where S_{BSF} is the effective surface recombination velocity at the BSF.

All that remains now is to determine suitable boundary conditions at $x = -x_N$ and $x = x_P$. These boundary conditions are commonly referred to as the *law of the junction*.

Under equilibrium conditions, zero applied voltage and no illumination, the Fermi energy, E_F, is constant with position. When a bias voltage is applied, it is convenient to introduce the concept of quasi-Fermi energies. It was shown earlier that the equilibrium carrier concentrations could be related to the Fermi energy (Equations 3.14 and 3.15). Under nonequilibrium conditions, similar relationships hold. Assuming the semiconductor is nondegenerate,

$$p = n_i e^{(E_i - F_P)/kT} \tag{3.100}$$

and

$$n = n_i e^{(F_N - E_i)/kT} \tag{3.101}$$

It is evident that, under equilibrium conditions, $F_P = F_N = E_F$. Under nonequilibrium conditions, assuming that the majority carrier concentrations at the contacts retain their equilibrium

values, the applied voltage can be written as

$$qV = F_N(-W_N) - F_P(W_P) \tag{3.102}$$

Since, in low-level injection, the majority carrier concentrations are essentially constant throughout the quasi-neutral regions, that is, $p_P(x_P \leq x \leq W_P) = N_A$ and $n_N(-W_N \leq x \leq -x_N) = N_D$, $F_N(-W_N) = F_N(-x_N)$ and $F_P(W_P) = F_P(x_P)$. Then, assuming that both the quasi-Fermi energies remain constant inside the depletion region,

$$qV = F_N(x) - F_P(x) \tag{3.103}$$

for $-x_N \leq x \leq x_P$, that is, everywhere inside the depletion region. Using Equations (3.100) and (3.101), this leads directly to the *law of the junction*, the boundary conditions used at the edges of the depletion region,

$$p_N(-x_N) = \frac{n_i^2}{N_D} e^{qV/kT} \tag{3.104}$$

and

$$n_P(x_P) = \frac{n_i^2}{N_A} e^{qV/kT}. \tag{3.105}$$

3.3.2 Generation Rate

For light incident at the front of the solar cell, $x = -W_N$, the optical generation rate takes the form (see Equation 3.34)

$$G(x) = (1-s)\int_\lambda (1 - r(\lambda)) f(\lambda)\alpha(\lambda) e^{-\alpha(x+W_N)}\, d\lambda. \tag{3.106}$$

Essentially, only photons with $\lambda \leq hc/E_G$ contribute to the generation rate.

3.3.3 Solution of the Minority-carrier Diffusion Equation

Using the boundary conditions defined by Equations (3.97), (3.99), (3.104), and (3.105) and the generation rate given by Equation (3.106), the solution to the minority-carrier diffusion equation, Equations (3.81) and (3.82), is easily shown to be

$$\Delta p_N(x) = A_N \sinh[(x + x_N)/L_p] + B_N \cosh[(x + x_N)/L_p] + \Delta p'_N(x) \tag{3.107}$$

in the n-type region and

$$\Delta n_P(x) = A_P \sinh[(x - x_P)/L_n] + B_P \cosh[(x - x_P)/L_n] + \Delta n'_P(x) \tag{3.108}$$

in the p-type region. The particular solutions, $\Delta p'_N(x)$ and $\Delta n'_P(x)$, due to $G(x)$ are given by

$$\Delta p'_N(x) = -(1-s)\int_\lambda \frac{\tau_p}{(L_p^2\alpha^2 - 1)}[1 - r(\lambda)] f(\lambda)\alpha(\lambda) e^{-\alpha(x+W_N)}\, d\lambda \tag{3.109}$$

and

$$\Delta n'_P(x) = -(1-s)\int_\lambda \frac{\tau_n}{(L_n^2\alpha^2 - 1)}[1 - r(\lambda)] f(\lambda)\alpha(\lambda) e^{-\alpha(x+W_N)}\, d\lambda. \tag{3.110}$$

Using the boundary conditions set above, A_N, B_N, A_P, and B_P in Equations (3.107) and (3.108) are readily solved for and are needed to obtain the diode current–voltage ($I-V$) characteristics.

3.3.4 Derivation of the Solar Cell $I-V$ Characteristic

The minority-carrier current densities in the quasi-neutral regions are just the diffusion currents, because the electric field is negligible. Using the active sign convention for the current (since a solar cell is typically thought of as a battery) gives

$$\vec{J}_{p,N}(x) = -q D_p \frac{\mathrm{d}\Delta p_N}{\mathrm{d}x} \tag{3.111}$$

and

$$\vec{J}_{n,P}(x) = q D_n \frac{\mathrm{d}\Delta n_P}{\mathrm{d}x} \tag{3.112}$$

The total current is given by

$$I = A[J_p(x) + J_n(x)] \tag{3.113}$$

and is true everywhere within the solar cell (A is the area of the solar cell). Equations (3.111) and (3.112) give only the hole current in the n-type region and the electron current in the p-type region, not both at the same point. However, integrating Equation (3.73), the electron continuity equation, over the depletion region, gives

$$\int_{-x_N}^{x_P} \frac{\mathrm{d}\vec{J}_n \mathrm{d}x}{\mathrm{d}x} \mathrm{d}x = \vec{J}_n(x_P) - \vec{J}_n(-x_N) = q \int_{-x_N}^{x_P} [R(x) - G(x)] \mathrm{d}x \tag{3.114}$$

$G(x)$ is easily integrated and the integral of the recombination rate can be approximated by assuming that the recombination rate is constant within the depletion region and is $R(x_m)$ where x_m is the point at which $p_\mathrm{D}(x_m) = n_\mathrm{D}(x_m)$ and corresponds to the maximum recombination rate in the depletion region. If recombination via a midgap single level trap is assumed, then, from Equations (3.36), (3.100), (3.101), and (3.103), the recombination rate in the depletion region is

$$R_\mathrm{D} = \frac{p_\mathrm{D} n_\mathrm{D} - n_\mathrm{i}^2}{\tau_n(p_\mathrm{D} + n_\mathrm{i}) + \tau_p(n_\mathrm{D} + n_\mathrm{i})} = \frac{n_\mathrm{D}^2 - n_\mathrm{i}^2}{(\tau_n + \tau_p)(n_\mathrm{D} + n_\mathrm{i})} = \frac{n_\mathrm{D} - n_\mathrm{i}}{(\tau_n + \tau_p)} = \frac{n_\mathrm{i}(\mathrm{e}^{qV/2kT} - 1)}{\tau_\mathrm{D}} \tag{3.115}$$

where τ_D is the effective lifetime in the depletion region. From Equation (3.114), $\vec{J}_n(-x_N)$, the majority carrier current at $x = -x_N$, can now be written as

$$\begin{aligned} \vec{J}_n(-x_N) &= \vec{J}_n(x_P) + q \int_{-x_N}^{x_P} G(x)\,\mathrm{d}x - q \int_{-x_N}^{x_P} R_\mathrm{D}\,\mathrm{d}x \\ &= \vec{J}_n(x_P) + q(1-s) \int_\lambda [1 - r(\lambda)] f(\lambda) \mathrm{e}[^{-\alpha(W_N - x_N)} - \mathrm{e}^{-\alpha(W_N + x_P)}]\,\mathrm{d}\lambda \\ &\quad - q \frac{W_\mathrm{D} n_\mathrm{i}}{\tau_\mathrm{D}} (\mathrm{e}^{qV/2kT} - 1) \end{aligned} \tag{3.116}$$

where $W_\mathrm{D} = x_P + x_N$. Substituting into Equation (3.113), the total current is now

$$I = A\left[J_p(-x_N) + J_n(x_P) + J_\mathrm{D} - q \frac{W_\mathrm{D} n_\mathrm{i}}{\tau_\mathrm{D}} (\mathrm{e}^{qV/2kT} - 1) \right] \tag{3.117}$$

where

$$J_{\mathrm{D}} = q(1-s)\int_{\lambda}[1-r(\lambda)]f(\lambda)(\mathrm{e}^{-\alpha(W_N-x_N)}-\mathrm{e}^{-\alpha(W_N+x_P)})\,\mathrm{d}\lambda \tag{3.118}$$

is the generation current density from the depletion region and A is the area of the solar cell. The last term of Equation (3.117) represents recombination in the space-charge region.

The solutions to the minority-carrier diffusion equation, Equations (3.107) and (3.108), can be used to evaluate the minority-carrier current densities, Equations (3.111) and (3.112). These can then be substituted into Equation (3.117), which, with some algebraic manipulation, yields the solar cell current–voltage characteristic

$$I = I_{\mathrm{SC}} - I_{o1}(\mathrm{e}^{qV/kT}-1) - I_{o2}(\mathrm{e}^{qV/2kT}-1). \tag{3.119}$$

where I_{SC} is the short-circuit current and is the sum of the contributions from each of the three regions: the n-type region ($I_{\mathrm{SC}N}$), the depletion region ($I_{\mathrm{SCD}} = AJ_{\mathrm{D}}$), and the p-type region ($I_{\mathrm{SC}P}$)

$$I_{\mathrm{SC}} = I_{\mathrm{SCN}} + I_{\mathrm{SCD}} + I_{\mathrm{SCP}} \tag{3.120}$$

where

$$I_{\mathrm{SC}N} = qAD_p\left[\frac{\Delta p'(-x_N)T_{p1} - S_{\mathrm{F,eff}}\Delta p'(-W_N) + D_p\left.\frac{\mathrm{d}\Delta p'}{\mathrm{d}x}\right|_{x=-W_N}}{L_pT_{p2}} - \left.\frac{\mathrm{d}\Delta p'}{\mathrm{d}x}\right|_{x=-x_N}\right] \tag{3.121}$$

with

$$T_{p1} = D_p/L_p\sinh[(W_N-x_N/L_p] + S_{\mathrm{F,eff}}\cosh[(W_N-x_N/L_p] \tag{3.122}$$

$$T_{p2} = D_p/L_p\cosh[(W_N-x_N/L_p] + S_{\mathrm{F,eff}}\sinh[(W_N-x_N/L_p] \tag{3.123}$$

and

$$I_{\mathrm{SC}P} = qAD_n\left[\frac{\Delta n'(x_P)T_{n1} - S_{\mathrm{BSF}}\Delta n'(W_P) - D_n\left.\frac{\mathrm{d}\Delta n'}{\mathrm{d}x}\right|_{x=W_P}}{L_nT_{n2}} + \left.\frac{\mathrm{d}\Delta n'}{\mathrm{d}x}\right|_{x=x_P}\right] \tag{3.124}$$

with

$$T_{n1} = D_n/L_n\sinh[(W_P-x_P)/L_n] + S_{\mathrm{BSF}}\cosh[(W_P-x_P)/L_n] \tag{3.125}$$

$$T_{n2} = D_n/L_n\cosh[(W_P-x_P)/L_n] + S_{\mathrm{BSF}}\sinh[(W_P-x_P)/L_n] \tag{3.126}$$

I_{o1} is the dark saturation current due to recombination in the quasi-neutral regions,

$$I_{o1} = I_{o1,p} + I_{o1,n} \tag{3.127}$$

with

$$I_{o1,p} = qA\frac{n_{\mathrm{i}}^2}{N_{\mathrm{D}}}\frac{D_p}{L_p}\left\{\frac{D_p/L_p\sinh[(W_N-x_N)/L_p] + S_{\mathrm{F,eff}}\cosh[(W_N-x_N/L_p]}{D_p/L_p\cosh[(W_N-x_N)/L_p] + S_{\mathrm{F,eff}}\sinh[(W_N-x_N/L_p]}\right\} \tag{3.128}$$

and

$$I_{o1,n} = qA\frac{n_i^2}{N_A}\frac{D_n}{L_n}\left\{\frac{D_n/L_n \sinh[(W_P - x_P)/L_n] + S_{BSF}\cosh[(W_P - x_P/L_n]}{D_n/L_n \cosh[(W_P - x_P)/L_n] + S_{BSF}\sinh[(W_P - x_P/L_n]}\right\} \quad (3.129)$$

These are very general expressions for the dark saturation current and reduce to more familiar forms when appropriate assumptions are made, as will be seen later.

I_{o2} is the dark saturation current due to recombination in the space-charge region,

$$I_{o2} = q\,A\frac{W_D n_i}{\tau_D} \quad (3.130)$$

and is bias-dependent since the depletion width, W_D, is a function of the applied voltage (Equation 3.91).

3.3.5 Interpreting the Solar Cell $I-V$ Characteristic

Equation (3.119), repeated here, is a general expression for the current produced by a solar cell.

$$I = I_{SC} - I_{o1}(e^{qV/kT} - 1) - I_{o2}(e^{qV/2kT} - 1) \quad (3.131)$$

The short-circuit current and dark saturation currents are given by rather complex expressions (Equations 3.120, 3.127, 3.128, 3.129, and 3.130) that depend on the solar cell structure, material properties, and the operating conditions. A full understanding of solar cell operation requires detailed examination of these terms. However, much can be learned about solar cell operation by examining the basic form of Equation (3.131). From a circuit perspective, it is apparent that a solar cell can be modeled by an ideal current source I_{SC} in parallel with two diodes – one with an ideality factor of 1 and the other with an ideality factor of 2, as shown in Figure 3.15. Note that the direction of the current source is such that it serves to forward-bias the diodes.

The current–voltage ($I-V$) characteristic of a typical silicon solar cell is plotted in Figure 3.16 for the parameter values given in Table 3.2. Note that it is the minority-carrier properties which determine the solar cell behavior, as indicated by Equations (3.119–3.129). For simplicity, the dark current due to the depletion region (diode 2) has been ignored (a reasonable and common assumption for a good solar cell, especially at larger forward biases). It illustrates several important figures of merit for solar cells – the short-circuit current, the open-circuit voltage,

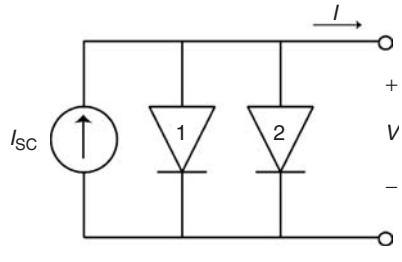


Figure 3.15 Simple solar cell circuit model. Diode 1 represents the recombination current in the quasi-neutral regions ($\propto e^{qV/kT}$), while diode 2 represents recombination in the depletion region ($\propto e^{qV/2kT}$)

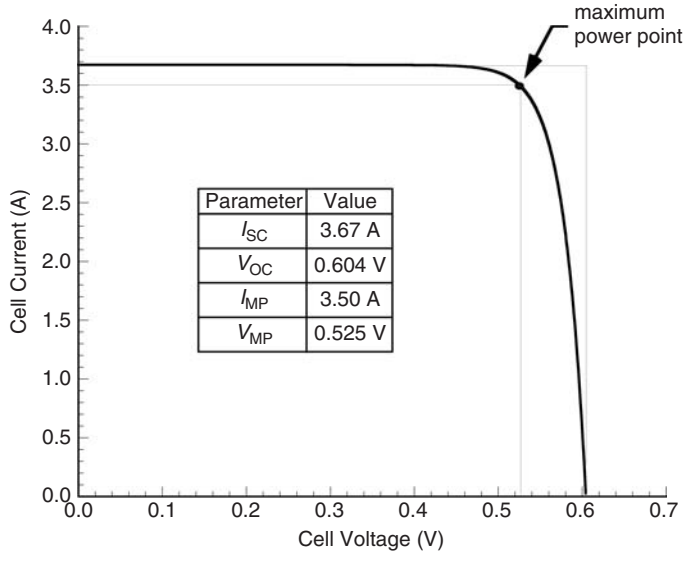


Figure 3.16 Current–voltage characteristic calculated for the silicon solar cell defined by Table 3.2 (area $A = 100\ \text{cm}^2$)

Table 3.2 Si solar cell model parameters

Parameter	n-type Si emitter	p-type Si base
Thickness	$W_N = 0.35\,\mu\text{m}$	$W_P = 300\,\mu\text{m}$
Doping density	$N_\text{D} = 1 \times 10^{20}\ \text{cm}^{-3}$	$N_\text{A} = 1 \times 10^{15}\ \text{cm}^{-3}$
Surface recombination	$D_\text{p} = 1.5\ \text{cm}^{-2}/\text{V s}$	$D_\text{n} = 35\ \text{cm}^{-2}/\text{V s}$
Minority-carrier diffusivity	$S_{\text{F,eff}} = 3 \times 10^4\ \text{cm/s}$	$S_\text{BSF} = 100\ \text{cm/s}$
Minority-carrier lifetime	$\tau_p = 1\,\mu\text{s}$	$\tau_n = 350\,\mu\text{s}$
Minority-carrier diffusion length	$L_p = 12\,\mu\text{m}$	$L_n = 1100\,\mu\text{m}$

and the fill factor. At small applied voltages, the diode current is negligible and the current is just the short-circuit current, I_SC, as can be seen when V is set to zero in Equation (3.131). When the applied voltage is high enough so that the diode current (recombination current) becomes significant, the solar cell current drops quickly.

Table 3.2 shows the huge asymmetry between the n-emitter and the p-base in a typical solar cell. The emitter is ~1000 times thinner, 10 000 times more heavily doped, and its diffusion length is ~100 times shorter than the corresponding quantities in the base.

At open-circuit ($I = 0$), all the light-generated current I_SC is flowing through diode 1 (diode ignored, as assumed above), so the open-circuit voltage can be written as

$$V_\text{OC} = \frac{kT}{q} \ln \frac{I_\text{SC} + I_{o1}}{I_{o1}} \approx \frac{kT}{q} \ln \frac{I_\text{SC}}{I_{o1}}, \tag{3.132}$$

where $I_\text{SC} \gg I_{o1}$.

Of particular interest is the point on the $I-V$ curve where the power produced is at a maximum. This is referred to as the *maximum power point* with $V = V_{MP}$ and $I = I_{MP}$. As seen in Figure 3.16, this point defines a rectangle whose area, given by $P_{MP} = V_{MP}I_{MP}$, is the largest rectangle for any point on the $I-V$ curve. The maximum power point is found by solving

$$\left.\frac{\partial P}{\partial V}\right|_{V=V_{MP}} = \left.\frac{\partial (IV)}{\partial V}\right|_{V=V_{MP}} = \left.\left[I + V\frac{\partial I}{\partial V}\right]\right|_{V=V_{MP}} = 0 \tag{3.133}$$

for $V = V_{MP}$. The current at the maximum power point, I_{MP}, is then found by evaluating Equation (3.131) at $V = V_{MP}$.

The rectangle-defined by V_{OC} and I_{SC} provides a convenient reference for describing the maximum power point. The fill factor, FF, is a measure of the *squareness* of the $I-V$ characteristic and is always less than one. It is the ratio of the areas of the two rectangles shown in Figure 3.16 or

$$FF = \frac{V_{MP}I_{MP}}{V_{OC}I_{SC}} = \frac{P_{MP}}{V_{OC}I_{SC}}. \tag{3.134}$$

Arguably, the most important figure of merit for a solar cell is its power conversion efficiency, η, which is defined as

$$\eta = \frac{P_{MP}}{P_{in}} = \frac{FFV_{OC}I_{SC}}{P_{in}} \tag{3.135}$$

The incident power P_{in} is determined by the properties of the light spectrum incident upon the solar cell. Further information regarding experimental determination of these parameters appears in Chapter 18.

Another important figure of merit is the collection efficiency, which can be defined relative to both optical and recombination losses as an *external* collection efficiency

$$\eta_C^{ext} = \frac{I_{SC}}{I_{inc}} \tag{3.136}$$

where

$$I_{inc} = qA\int_{\lambda<\lambda_G} f(\lambda)\,d\lambda \tag{3.137}$$

is the maximum possible photocurrent that would result if all photons with $E > E_G$ ($\lambda < \lambda_G = hc/E_G$) created electron–hole pairs that were collected. The collection efficiency can also be defined with respect to recombination losses as the internal collection efficiency

$$\eta_C^{int} = \frac{I_{SC}}{I_{gen}} \tag{3.138}$$

where

$$I_{gen} = qA(1-s)\int_{\lambda<\lambda_G} [1 - r(\lambda)f(\lambda)(1 - e^{-\alpha(W_N+W_P)})\,d\lambda \tag{3.139}$$

is the light-generated current. This represents what the short-circuit current would be if every photon that is absorbed is collected and contributes to the short-circuit current. $I_{gen} = I_{inc}$ when there is no grid shadowing, no reflective losses, and the solar cell has infinite optical thickness.

3.3.6 Properties of Efficient Solar Cells

Using these figures of merit, the properties of a good (efficient) solar cell can be ascertained. From Equation (3.135), it is clear that an efficient solar cell will have a high short-circuit current I_{SC}, a high open-circuit voltage, V_{OC} and a fill factor FF as close as possible to 1. A more detailed understanding of what influences the solar cell efficiency can be obtained by rewriting the efficiency as [18]

$$\eta = \frac{P_{\max}}{P_{in}} = \eta_{ideal}\eta_{photon}FF\eta_V\eta_C^{\text{int}}, \tag{3.140}$$

where FF and η_C^{int} have been previously defined (Equations 3.134 and 3.138, respectively) and η_{ideal}, η_{photon}, and η_V are defined below.

Assuming the maximum energy that can be extracted from an absorbed photon is E_G, the ideal efficiency can be expressed as

$$\eta_{ideal}(E_G) = \frac{\frac{1}{q}E_G I_{\text{inc}}}{P_{\text{in}}} = \frac{E_G}{(P_{\text{in}}/A)}\int_{\lambda<\lambda_G} f(\lambda)\,d\lambda. \tag{3.141}$$

Since only photons with $h\nu > E_G$ can create electron–hole pairs and contribute to the output power of the solar cell, it is clear that the bandgap determines how well the solar cell is coupled to the solar spectrum. A simple analysis can be performed to predict the ideal efficiency. This is plotted in Figure 3.17 for an AM1.5 global spectrum and shows a maximum efficiency of 48% at about $E_G = 1.1$ eV, close to the bandgap of silicon, although bandgaps between 1.0 and 1.6 eV have comparable ideal efficiencies.

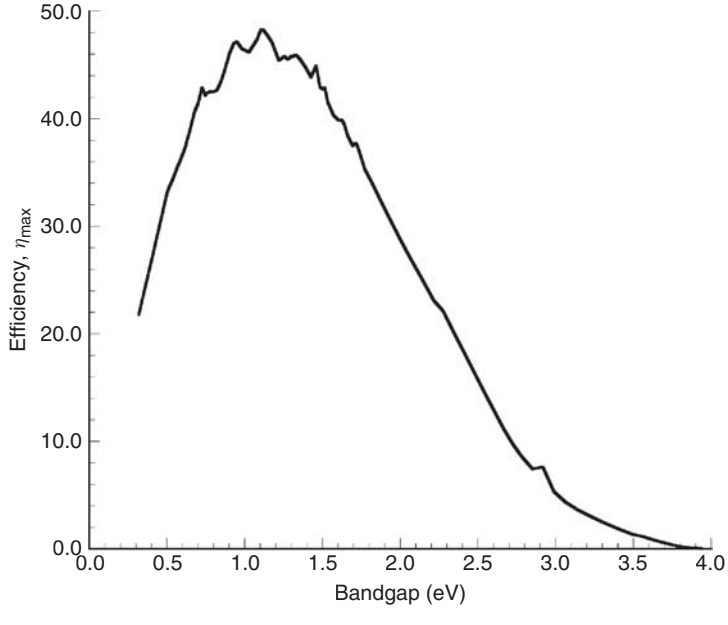


Figure 3.17 Ideal efficiency as a function of semiconductor band gap for an AM1.5 global spectrum

Of course, this assumes that $V_{OC} = \frac{1}{q}E_G$ and $FF = 1$, which are obvious exaggerations. Perfect light trapping is also assumed so that $I_{SC} = I_{inc}$, but that is a more realistic prospect. However, this quantity does serve to set an upper bound on the efficiency of a single-junction solar cell. Of course, multijuction photovoltaic systems will have a higher ideal efficiency. More complete analyses of the theoretical limits of solar cells are given elsewhere [19–21] and are also discussed in Chapter 4 of this handbook.

The photon efficiency η_{photon} accounts for photons that are reflected, transmitted through, or otherwise not absorbed in the solar cell and can be written as

$$\eta_{photon} = \frac{I_{gen}}{I_{inc}} = \frac{\eta_C^{ext}}{\eta_C^{\text{int}}}. \tag{3.142}$$

To maximize $\eta_{photon}(\eta_{photon} \to 1$ when $I_{\text{gen}} \to I_{\text{inc}}$ or, equivalently, $\eta_C^{ext} \to \eta_C^{\text{int}})$, the solar cell should be designed with a minimum amount of grid shadowing s, minimum reflectance $r(\lambda)$, and be optically thick enough such that nearly all the photons with $E > E_{\text{G}}$ are absorbed.

A transcendental relationship between V_{OC} and V_{MP} can be obtained from the solution of Equation (3.131) for the single-diode model, from which the following semi-empirical expression for the fill factor can be extracted [22]

$$FF = \frac{V_{\text{OC}} - \dfrac{kT}{q}\ln[qV_{\text{OC}}/kT + 0.72]}{V_{\text{OC}} + kT/q}. \tag{3.143}$$

It can be seen that FF is a weak function of the open-circuit voltage, increasing slowly as the open-circuit voltage increases. This expression neglects any series and shunt resistances which tend to degrade the fill factor, as will be discussed later in this chapter.

The voltage efficiency η_V is the ratio of the open-circuit voltage to the bandgap voltage

$$\eta_V = \frac{V_{OC}}{\frac{1}{q}E_G}. \tag{3.144}$$

Empirically, the best solar cells have an open-circuit voltage approximately 0.4 V less than the bandgap voltage (no solar concentration). For silicon, this gives $\eta_V = 0.643$. It is clearly desirable to have the open-circuit voltage approach the bandgap voltage and this is one of the challenges in the development of next generation solar cells. At open-circuit, since there is no flow of carriers out of the devices, every electron–hole pair must recombine. It is the rate of this recombination, or the reverse saturation current, that constrains the open-circuit voltage. The open-circuit voltage (Equation 3.132)

$$V_{\text{OC}} \approx \frac{kT}{q}\ln\frac{I_{\text{SC}}}{I_{o1}} \tag{3.145}$$

is logarithmically proportional to the short-circuit current and to the reciprocal of the reverse saturation current I_{o1}. Therefore, reducing the saturation current will increase the open-circuit voltage. From Equations (3.128) and (3.129), it is obvious that $I_{o1} \to 0$ as $\tau \to \infty$ and $S \to 0$.

The final term in Equation (3.140) is for the internal collection efficiency, which was defined previously in Equation (3.138), $I_{\text{SC}} = \eta_{\text{C}}^{\text{int}} I_{\text{gen}}$ is dependent on the recombination in the solar cell and will approach 1 as $\tau \to \infty$ and $S \to 0$. Voltage-dependent collection can compensate for low effective carrier lifetimes [23] in achieving a higher short-circuit current, but the open-circuit voltage and fill factor do not benefit from this effect. This effect is briefly discussed later in this chapter.

From this discussion, it can be seen that the design of an efficient solar cell has several key goals:

1. Selection of a semiconductor material with a bandgap well matched to the solar spectrum, i.e. maximizing η_{ideal}.
2. Minimizing optical losses such as grid shadowing, reflectance, and absorption in the optical components, as well as maximizing the optical thickness of the solar cell, thereby maximizing η_{photon}.
3. Minimizing series and shunt resistances in the cell and its connections, thereby maximizing the fill factor, FF.
4. Minimization of the bulk and surface recombination rates, maximizing η_V and hence the open-circuit voltage.
5. Minimization of the bulk and surface recombination rates will also maximize the internal collection efficiency η_C^{int} and hence the short-circuit current.

Simultaneous achievement of all of these goals will result in a very efficient solar cell. For a silicon solar cell with $V_{\text{OC}} = 0.72$ V, the predicted efficiency from Equation (3.140) (assuming $\eta_{photon} = \eta_C^{\text{int}} = 1$ under AM1.5 global illumination is 26.2% – just slightly higher than the best reported silicon solar cells [24].

It is evident that, despite the apparent complexity of the expressions describing the fundamental operation of solar cells, the basic operating principles are easy to understand, as illustrated by the above discussion.

3.3.7 Lifetime and Surface Recombination Effects

The solar cell characteristics previously derived (Equations 3.119–3.132) allow examination of the dependence of the solar cell performance on specific sources of recombination. Figure 3.18 shows how the base minority-carrier lifetime affects V_{OC}, I_{SC}, and the FF. Unless otherwise stated, the parameters of Table 3.2 are used to compute the solar cell performance. Short lifetimes mean that the diffusion length in the base is much less than the base thickness and carriers created deeper than about one diffusion length in the base are unlikely to be collected. When this is true ($L_n \ll W_P$), the contribution to the dark saturation current in the base (Equation 3.129) becomes

$$I_{o1,n} = qA\frac{n_{\text{i}}^2}{N_{\text{A}}}\frac{D_n}{L_n} \tag{3.146}$$

and is commonly referred to as the *long-base approximation*. In this case, the BSF has no effect on the dark saturation current. On the other hand, when the base minority-carrier lifetime is long ($L_n \gg W_P$), the carriers readily come in contact with the BSF and the dark saturation current is a strong function of S_{BSF}

$$I_{o1,n} = qA\frac{n_{\text{i}}^2}{N_{\text{A}}}\frac{D_n}{(W_P - X_P)}\frac{S_{\text{BSF}}}{S_{\text{BSF}} + D_n/(W_P - x_P)} \tag{3.147}$$

When S_{BSF} is very large (i.e. no BSF), this reduces to the more familiar short-base approximation

$$I_{o1,n} = qA\frac{n_{\text{i}}^2}{N_{\text{A}}}\frac{D_n}{(W_P - x_P)}. \tag{3.148}$$

Figure 3.19 shows how S_{BSF} affects V_{OC}, I_{SC}, and FF. Notice that the break point in the curves occurs when $S_{\text{BSF}} \approx D_n/W_P = 1000$ cm/s, as can be inferred from Equation (3.147).

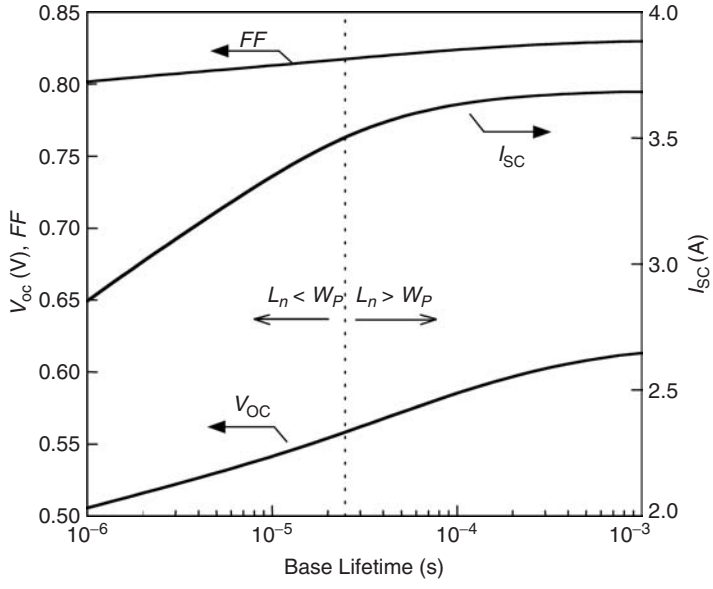


Figure 3.18 Effect of base lifetime on solar cell performance for the solar cell parameters in Table 3.2. The minority-carrier diffusion length ($L_n = \sqrt{D_n \tau_n}$) is equal to the base thickness (W_P) when $\tau_n = 25.7\,\mu s$

Front surface recombination for solar cells with contact grids on the front of the device is really an average over the front surface area of the relatively low surface recombination velocity between the grid lines and the very high surface recombination velocity of the ohmic contact. An expression for the effective front surface recombination velocity is given by [25]

$$S_{F,\text{eff}} = \frac{(1-s)S_F \bar{G}_N \tau_p \left(\cosh \frac{W_N}{L_p} - 1\right) + p_o(e^{qV/A_o kT} - 1)\left[s\frac{D_p}{L_p}\frac{\cosh \frac{W_N}{L_p}}{\sinh \frac{W_N}{L_p}} + S_F\right]}{(1-s)\left[p_o(e^{qV/A_o kT} - 1) + \bar{G}_N \tau_p \left(\cosh \frac{W_N}{L_p} - 1\right)\right]} \tag{3.149}$$

where S_F is the surface recombination velocity between the grid lines and $\bar{G}$ is the average generation rate in the emitter region. It is obvious that $S_{F,\text{eff}}$ is dependent upon the solar cell operation point. This is better seen in Table 3.3 where some special cases are illustrated (assuming $L_p \gg W_N$).

3.4 ADDITIONAL TOPICS

3.4.1 Spectral Response

The spectral response, $SR(\lambda)$, of a solar cell permits an examination of how photons of different wavelengths (energy) contribute to the short-circuit current. Just as the collection efficiency can be

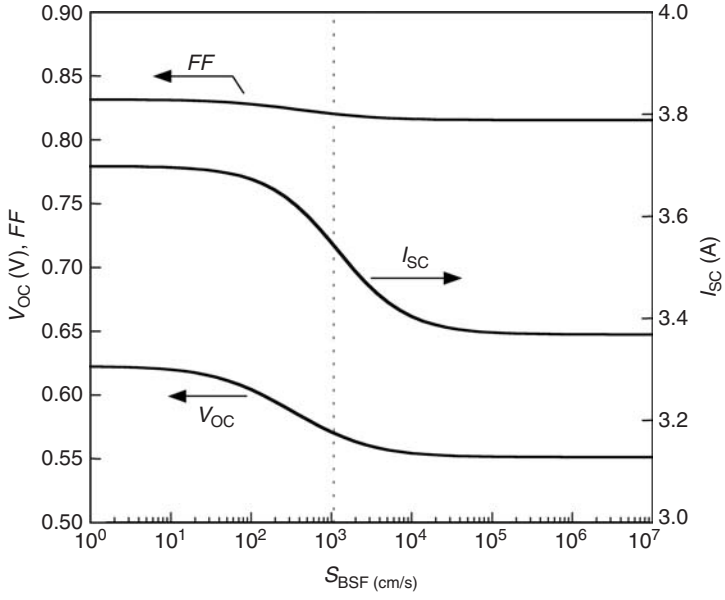


Figure 3.19 Effect of the back-surface field recombination velocity on solar cell performance. All other parameters are from Table 3.2

Table 3.3 Special cases of $S_{F,eff}$

No grid ($s = 0$)	$S_{F,eff} = S_F$
Full grid ($s = 1$)	$S_{F,eff} \to \infty$
Dark ($\overline{G} = 0$)	$S_{F,eff} = \dfrac{S_F + s D_p/W_N}{1 - s}$
Short-circuit (V = 0)	$S_{F,eff} = S_F$
V large ($\approx V_{OC}$)	$S_{F,eff} = \dfrac{S_F + s D_p/W_N}{1 - s}$

measured as either an external or internal collection efficiency, so can the spectral response. The spectral response is defined as the *short-circuit current* $I_{SC}(\lambda)$, resulting from a single wavelength of light normalized by the maximum possible current. The external spectral response is defined as

$$SR_{ext} = \frac{I_{SC}(\lambda)}{qAf(\lambda)} \tag{3.150}$$

and the internal spectral response as

$$SR_{int} = \frac{I_{SC}(\lambda)}{qA(1-s)(1-r(\lambda))f(\lambda)(e^{-\alpha(\lambda)W_{opt}} - 1)}, \tag{3.151}$$

where W_{opt} is the optical thickness of the solar cell (technically, also a function of wavelength). Experimentally, the external spectral response is measured. The internal spectral response is

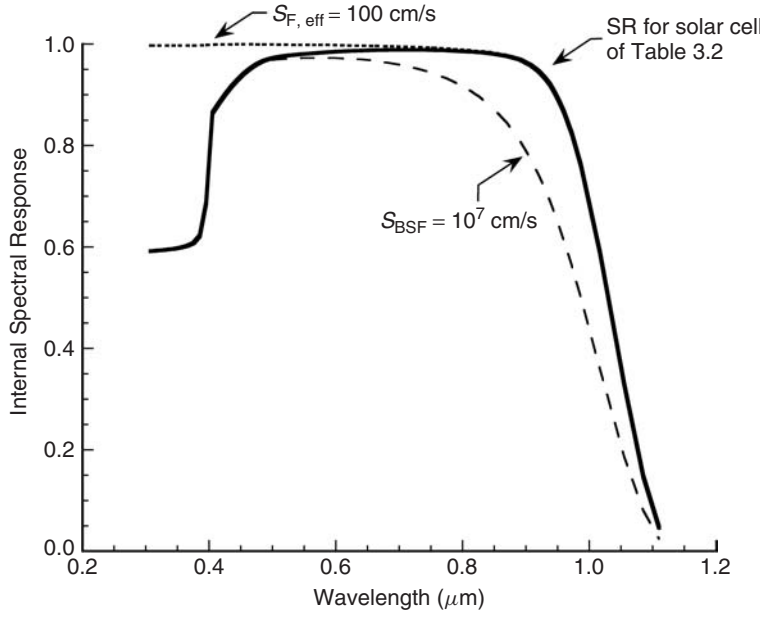


Figure 3.20 Internal spectral response of the silicon solar cell defined in Table 3.2

determined from it, along with the knowledge of the grid shadowing, reflectance, and optical thickness. W_{opt} can be greater than the cell thickness if light-trapping methods are used. Such methods include textured surfaces [26] and back-surface reflectors [27] and are discussed in Chapters 11 and 12. The short-circuit current can be written in terms of the external spectral response as

$$I_{SC} = \int_{\lambda} SR_{ext}(\lambda) f(\lambda)\, d\lambda. \tag{3.152}$$

The internal spectral response gives an indication of which sources of recombination are affecting the cell performance. This is demonstrated in Figure 3.20 where the internal spectral response of the silicon solar cell described by the parameters of Table 3.2 is shown. Also shown is the spectral response when $S_{F,eff} = 100$ cm/s (a well-passivated front surface) and the spectral response when $S_{BSF} = 1 \times 10^7$ cm/s (in effect, no BSF). The short-wavelength response improves dramatically when the front surface is passivated since the absorption coefficient is highest for short-wavelength (high-energy) photons. Conversely, removing the BSF makes it more likely that electrons created deep within the base region of the solar cell (those created by the long-wavelength, low-energy photons) will recombine at the back contact and therefore, the long-wavelength response is dramatically reduced.

3.4.2 Parasitic Resistance Effects

Equation (3.143) neglects the parasitic series and shunt resistances typically associated with real solar cells. Incorporating these resistances into the circuit model of Figure 3.15, as shown in

Figure 3.21, yields

$$I = I'_{SC} - I_{o1}(e^{q(V+IR_S)/kT} - 1) - I_{o2}(e^{q(V+IR_S)/2kT} - 1) - \frac{(V + IR_S)}{R_{Sh}} \tag{3.153}$$

where I'_{SC} is the short-circuit current when there are no parasitic resistances. The effect of these parasitic resistances on the $I-V$ characteristic is shown in Figures 3.22 and 3.23. As can also be seen in Equation (3.153), the shunt resistance R_{Sh} has no effect on the short-circuit current, but reduces the open-circuit voltage. Conversely, the series resistance R_S has no effect on the open-circuit voltage, but reduces the short-circuit current. Sources of series resistance include the metal contacts, particularly the front grid, and the transverse flow of current in the solar cell emitter to the front grid.

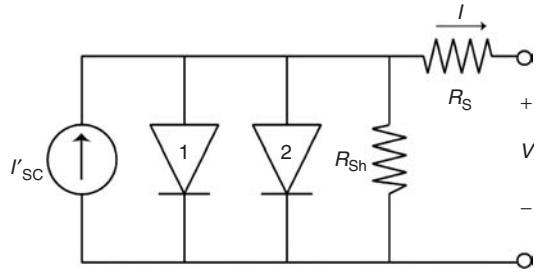


Figure 3.21 Solar cell circuit model including the parasitic series and shunt resistances

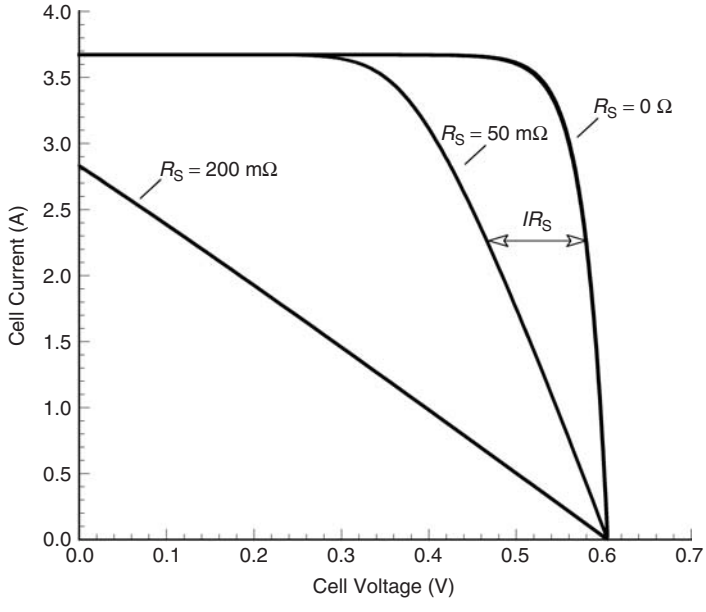


Figure 3.22 Effect of series resistance on the current–voltage characteristic of a solar cell ($R_{Sh} \rightarrow \infty$)

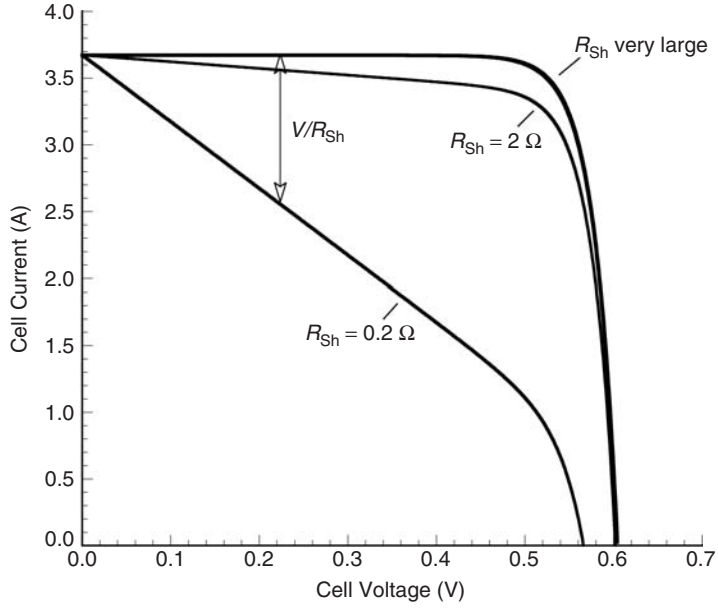


Figure 3.23 Effect of shunt resistance on the current–voltage characteristic of a solar cell ($R_S = 0$)

It is often more convenient to rewrite Equation (1.153) as

$$I = I'_{SC} - I_o(e^{q(V+IR_S)/A_okT} - 1) - \frac{(V + IR_S)}{R_{Sh}} \tag{3.154}$$

where A_o is the diode ideality (quality) factor and typically has a value between 1 and 2, with $A_o \approx 1$ for diode dominated by recombination in the quasi-neutral regions and $A_o \to 2$ when recombination in the depletion region dominates. In solar cells where the recombination in each region is comparable, A_o is somewhere in between. At short-circuit, Equation (3.154) becomes

$$I_{SC} = I'_{SC} - I_o(e^{qI_{SC}R_S/A_okT} - 1) - I_{SC}R_S/R_{Sh} \tag{3.155}$$

and at open-circuit, it becomes

$$0 = I'_{SC} - I_o(e^{V_{OC}/A_okT} - 1) - V_{OC}/R_{Sh}. \tag{3.156}$$

When $\log(I_{SC})$ is plotted versus V_{OC} (where I_{SC} and V_{OC} are obtained over a range of illumination intensities), there is typically a regime where neither the series nor shunt resistances are important, as illustrated in Figure 3.24. The slope of this line will yield the diode ideality factor A_o, while the y-intercept will give I_o. In the regime where only series resistance is important, Equations (3.155) and (3.156) can be combined to give

$$I_{SC}R_S = \frac{A_okT}{q} \ln\left[\frac{I_oe^{qV_{OC}/A_okT} - I_{SC}}{I_o}\right] \tag{3.157}$$

and a plot of I_{SC} versus $\log[I_oe^{qV_{OC}/A_okT} - I_{SC}]$ will then permit R_S to be extracted from the slope of this line. Similarly, in the regime where only R_{Sh} is important, Equations (3.155) and (3.156)

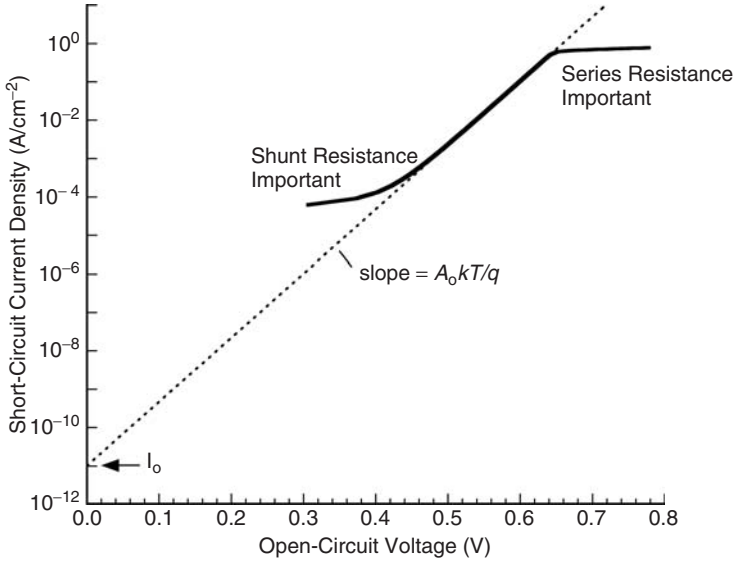


Figure 3.24 Short-circuit current versus open-circuit voltage plot illustrating the effects of series and shunt resistances

can be combined to give

$$\frac{V_{\mathrm{OC}}}{R_{\mathrm{Sh}}} = I_{\mathrm{SC}} - I_o \mathrm{e}^{qV_{\mathrm{OC}}/A_o kT} \tag{3.158}$$

and R_{Sh} can be determined from the slope of the line given by plotting V_{OC} versus $[I_{\mathrm{SC}} - I_o \mathrm{e}^{qV_{\mathrm{OC}}/A_o kT}]$. If the series and shunt resistances are such that there is no regime where they can be neglected, the parameters can, with patience, be extracted through the process of trial and error.

3.4.3 Temperature Effects

From Equations (3.128) and (3.129), it is apparent that

$$I_{o1,n},\, I_{o1,p} \propto n_{\mathrm{i}}^2 \tag{3.159}$$

and from Equation (3.130) that

$$I_{o2} \propto n_{\mathrm{i}}. \tag{3.160}$$

An increase in the intrinsic carrier concentration increases the dark saturation (recombination) current and results in a decrease in the open-circuit voltage, as can be seen from Equation (3.145). The dark saturation current contains other temperature-dependent terms (D, τ, and S), but the temperature dependence of the intrinsic carrier concentration dominates. The intrinsic carrier concentration is given by Equation (3.17), which when combined with Equations (3.12) and (3.13) yields

$$n_{\mathrm{i}} = 2(m_n^* m_p^*)^{3/4} \left(\frac{2\pi kT}{h^2}\right)^{3/2} \mathrm{e}^{-E_{\mathrm{G}}/2kT}. \tag{3.161}$$

The effective masses are generally taken to be weak functions of temperature. The bandgap decreases with temperature, and its temperature dependence is well modeled by

$$E_{\mathrm{G}}(T) = E_{\mathrm{G}}(0) - \frac{aT^2}{T+\beta}. \tag{3.162}$$

where α and β are constants specific to each semiconductor. It is clear that as the temperature increases, n_{i} increases, and thus recombination increases, and cell performance is impaired. Bandgap narrowing, referred to earlier, is a reduction in bandgap due to high doping and also serves to increase n_{i} and impair solar cell performance.

The open-circuit current expression, Equation (3.145), can be rearranged and the temperature dependence explicitly included to give

$$I_{\mathrm{SC}} \approx I_{o1}\mathrm{e}^{qV_{\mathrm{OC}}/kT} \approx BT^{\zeta}\mathrm{e}^{-E_{\mathrm{G}}(0)/kT}\mathrm{e}^{qV_{\mathrm{OC}}/kT} \tag{3.163}$$

where B is a temperature-independent constant and $T^{\zeta}\mathrm{e}^{-E_{\mathrm{G}}(0)/kT}$ accounts for the temperature dependence of the saturation current. The short-circuit current is relatively unaffected by temperature under typical operating conditions, so by differentiating with respect to T, the temperature dependence of the open-circuit voltage can be expressed as [22]

$$\frac{\mathrm{d}V_{\mathrm{OC}}}{\mathrm{d}T} = -\frac{\dfrac{1}{q}E_{\mathrm{G}}(0) - V_{\mathrm{OC}} + \zeta\dfrac{kT}{q}}{T} \tag{3.164}$$

which for silicon at 300 K corresponds to about $-2.3\,\mathrm{mV/^{\circ}C}$. Equation (3.163) can be rearranged as follows:

$$V_{\mathrm{OC}}(T) = \frac{1}{q}E_{\mathrm{G}}(0) - \frac{kT}{q}\ln\left(\frac{BT^{\zeta}}{I_{\mathrm{SC}}}\right). \tag{3.165}$$

V_{OC} varies roughly linearly with temperature and an extrapolation of V_{OC} to $T = 0$ is the bandgap voltage since $\lim_{T\to 0}[T\ln T] = 0$.

The reason this is so important is that modules typically operate at 20–40 °C above ambient, depending on the module design and sunlight intensity. Typical Si modules have a negative temperature coefficient of power output of -0.4 to -0.5% relative per °C, largely due to the temperature dependence of V_{OC} as indicated in Equation 3.165. The effect of temperature on module performance is discussed further in Chapters 18 and 19.

3.4.4 Concentrator Solar Cells

Operating solar cells under concentrated illumination offers two main advantages. The first is that since fewer solar cells are required to collect the sunlight falling on a given area, their cost of manufacture can be higher than that for cells designed for unconcentrated illumination, and they are therefore presumably of higher quality (efficiency). The second is that operation under concentrated illumination offers an advantage in the solar cell efficiency. If sunlight is concentrated by a factor of X (X suns illumination), the short-circuit at that concentration is

$$I_{\mathrm{SC}}^{\mathrm{Xsuns}} = \mathrm{X}I_{\mathrm{SC}}^{\mathrm{1sun}}. \tag{3.166}$$

This is assuming that the semiconductor parameters are unaffected by the illumination level and that the cell temperature is the same at both levels of illumination – not necessarily valid assumptions,

especially at very large X. However, these assumptions will allow the demonstration of the potential efficiency of concentrator solar cells. Substituting Equation (3.166) into Equation (3.135) gives

$$\eta = \frac{FF^{\text{Xsuns}} V_{\text{OC}}^{\text{Xsuns}} I_{\text{SC}}^{\text{Xsuns}}}{P_{\text{in}}^{\text{Xsuns}}} = \frac{FF^{\text{Xsuns}} V_{\text{OC}}^{\text{Xsuns}} \text{X} I_{\text{SC}}^{\text{1sun}}}{\text{X} P_{\text{in}}^{\text{1sun}}} = \frac{FF^{\text{Xsuns}} V_{\text{OC}}^{\text{Xsuns}} I_{\text{SC}}^{\text{1sun}}}{P_{\text{in}}^{\text{1sun}}} \tag{3.167}$$

From Equation (3.132),

$$V_{\text{OC}}^{\text{Xsuns}} = V_{\text{OC}}^{\text{1sun}} + \frac{kT}{q} \ln \text{X}. \tag{3.168}$$

FF is a function of V_{OC} (Equation 3.143), so

$$\eta^{\text{Xsuns}} = \eta^{\text{1sun}} \left(\frac{FF^{\text{Xsuns}}}{FF^{\text{1sun}}} \right) \left(1 + \frac{\dfrac{kT}{q} \ln \text{X}}{V_{\text{OC}}^{\text{1sun}}} \right). \tag{3.169}$$

Both factors multiplying the 1 sun efficiency increase as the illumination concentration increases. Therefore, the efficiency of concentrator cells increases as the illumination concentration increases. For a silicon solar cell with $V_{\text{OC}}^{\text{1sun}} = 0.72\text{V}$, the efficiency at 1000 suns can potentially be more than 25% higher than its 1 sun value.

Of course, there are many obstacles to achieving this. Concentrator cells must be cooled, since an increase in operating temperature reduces V_{OC}, and hence the cell efficiency. In real devices, the FF^{Xsuns} eventually decreases with increasing solar cell current due to parasitic series resistance. Concentrator solar cells are discussed in more detail in Chapter 10.

3.4.5 High-level Injection

In high-level injection, the excess carrier concentrations greatly exceed the doping in the base region, so $\Delta p \approx \Delta n \approx n \approx p$ if the carriers are moving generally in the same direction. This occurs with back-contact solar cells, such as the silicon point-contact solar cell [28], which is illustrated in Figure 3.25. Since both electrical contacts are on the back, there is no grid shadowing. These cells are typically used in concentrator application and high-level injection conditions pervade. Assuming high-level injection, a simple analysis is possible.

Returning to Equations (3.77) and (3.78), it can be seen that in high-level injection, the electric field can be eliminated (it is not necessarily zero), resulting in the ambipolar diffusion equation

$$D_{\text{a}} \frac{\text{d}^2 p}{\text{d}x^2} - \frac{p}{\tau_n + \tau_p} = -G(x), \tag{3.170}$$

where the ambipolar diffusion coefficient is given by

$$D_{\text{a}} = \frac{2 D_n D_p}{D_n + D_p}. \tag{3.171}$$

In silicon, where $D_n / D_p \approx 3$ over a wide range of doping, the ambipolar diffusion coefficient is $D_{\text{a}} \approx 3/2 D_p \approx 1/2 D_n$ and, if we also assume $\tau_p \approx \tau_n$, the ambipolar diffusion length is

$$L_{\text{a}} \approx \sqrt{3} L_p \approx L_n. \tag{3.172}$$

Thus, the increased high-injection lifetime (see Equation 3.40) offsets the reduced ambipolar diffusion coefficient.

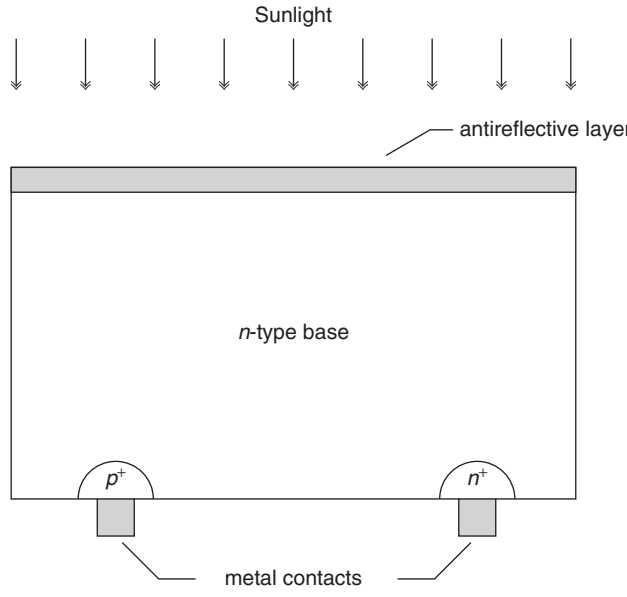


Figure 3.25 Schematic of a back-contact solar cell

It is crucial that the front surface of a back-contacted cell be well passivated, so we will assume that $S_F = 0$. We will further assume that optical generation is uniform throughout the base region. At open-circuit, with these assumptions, $d^2p/dx^2 = 0$ and therefore

$$V_{OC} = \frac{2kT}{q} \ln\left[\frac{G(\tau_n + \tau_p)}{n_i}\right]. \tag{3.173}$$

The short-circuit current (with $p \simeq 0$ at the back of the cell) is

$$I_{SC} = qAL_aG \sinh(W_B/L_a) \tag{3.174}$$

which, when $L_a \gg W_B$, becomes

$$I_{SC} = qAW_BG. \tag{3.175}$$

3.4.6 *p-i-n* Solar Cells and Voltage-dependent Collection

The *p-i-n* solar cell takes advantage of the fact that in many semiconductor materials, especially direct bandgap semiconductors (i.e. large absorption coefficient), most of the electron–hole pairs are created very near the surface. If an intrinsic (undoped) layer is placed between the (very thin) n and p regions, the depletion region thickness is the most significant fraction of the total solar cell thickness, as illustrated in Figure 3.26. Carrier collection is now aided by the electric field in the depletion region, which helps offset the low lifetimes in some materials, such as amorphous silicon (see Chapter 12). The $I-V$ characteristic of a *p-i-n* solar cell can be described with minor modifications to the previously derived expressions. The most significant modification is to Equation (3.130) where the depletion width is now written as

$$W_D = \chi_N + W_I + \chi_P \tag{3.176}$$

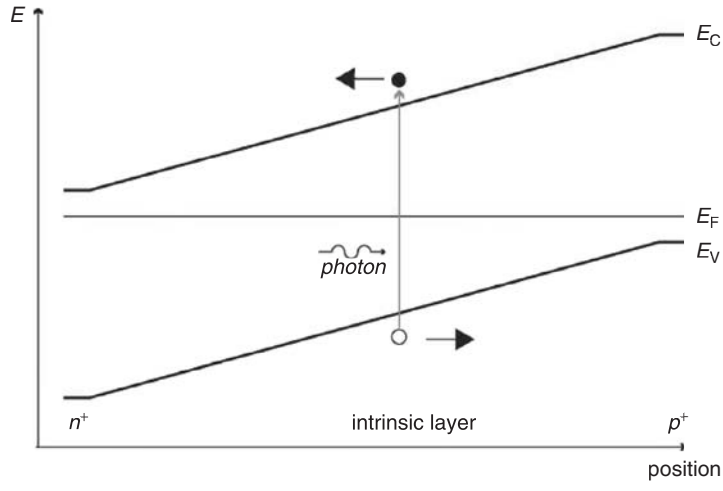


Figure 3.26 Band diagram of a *p-i-n* solar cell illustrating field-enhanced collection

where W_I is the thickness of the intrinsic layer. Since χ_N and χ_P are very thin, short-base approximations are in order for Equations (3.128) and (3.129). Also, there is no BSF ($S_{BSF} \rightarrow \infty$).

As mentioned above, the electric field in the depletion region of *p-i-n* solar cells aids in the collection of carriers. This is referred to as voltage-dependent collection (VDC) and is an important effect in any solar cell in which there is significant photogeneration in the depletion region, as is the case in most thin film solar cells. At the maximum power point and at open-circuit, the electric field in the depletion region is lower than at short-circuit, which leads to a lower *FF* and V_{OC} than might otherwise be expected. An excellent analysis of VDC in CdTe/CdS solar cells is given in [23].

3.4.7 Heterojunction Solar Cells

Reducing the recombination losses in the emitter will improve the efficiency of the solar cell. This can be accomplished by reducing the junction area [19]. Another way is by using a wider bandgap material for the emitter of the solar cell, as shown in Figure 3.27 for an n^+p solar cell. Ideally, minority-carriers in each region are collected by the junction to become majority-carriers in the opposite region. Majority-carriers injected into the opposite region that become minority-carriers are a source of recombination and reduce the efficiency of the solar cell.

The larger barrier presented to holes by the wider bandgap emitter of a heterojunction solar cell substantially reduces the number of holes injected into the emitter, thus reducing recombination in the emitter and improving the efficiency of the solar cell. The arrows in Figure 3.27 illustrate this point. This can also be seen analytically from Equation (3.128), the emitter component of the dark saturation current, which is reproduced below with $n_i = n_{i,emitter}$

$$I_{o1,p} = qA\frac{n_{i,\text{emitter}}^2}{N_D}\frac{D_p}{L_p}\left\{\frac{D_p/L_p \sinh[(W_N - x_N)/L_p] + S_{F,\text{eff}}\cosh[(W_N - x_N/L_p]}{D_p/L_p \cosh[(W_N - x_N)/L_p] + S_{F,\text{eff}}\sinh[(W_N - x_N/L_p]}\right\} \quad (3.177)$$

Recall from Equation (3.17) that

$$n_{i,\text{emitter}} = \sqrt{N_C N_V}e^{-E_{G,\text{emitter}}/2kT}. \quad (3.178)$$

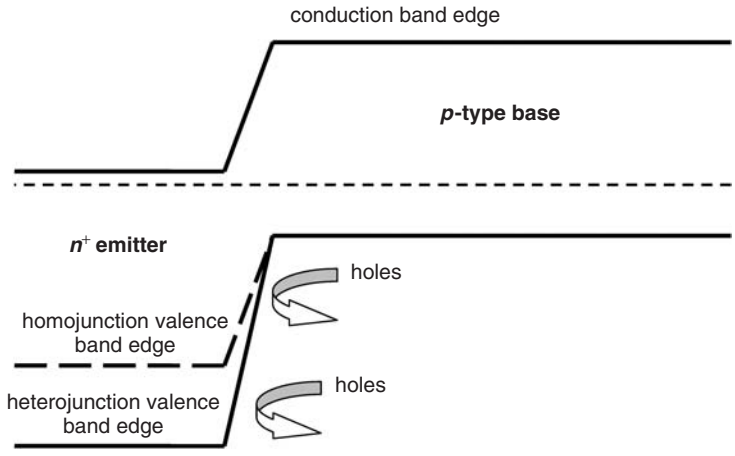


Figure 3.27 Band diagram of a heterojunction solar cell

The intrinsic carrier concentration $n_{i,emitter}$ is much smaller for a wider bandgap emitter, thus reducing the emitter component of the dark saturation current and therefore the net recombination in the emitter region.

An additional advantage of a heterojunction solar cell is that incident photons with energy less than the emitter bandgap energy (and higher than the base region bandgap) will be absorbed in the base region of the solar cell rather than near the front surface where recombination can be high (as would be the case in a homojunction). Heterojunction solar cells are discussed in more detail in Chapters 8, 13, and 14.

3.4.8 Detailed Numerical Modeling

While analytic solutions such as those discussed thus far in this chapter provide an intuitive understanding of solar cell performance and are therefore very important, they are limited in their accuracy due to the many simplifying assumptions that must be made in order to obtain them. It is rather straightforward to solve the semiconductor equations numerically without the need to make so many simplifications. Several computer codes have been written that solve the semiconductor equations for the explicit purpose of modeling solar cells: PC-1D [29], AMPS [30], ADEPT [31], and its predecessors [32, 33], for example.

The basic design of these computer programs is very similar. The semiconductor equations (three coupled nonlinear partial differential equations) are cast in a normalized form [34] to simplify the calculations. Finite difference or finite element methods are then used to discretize the equations on a mesh (grid), resulting in a set of three coupled nonlinear difference equations. Using appropriately discretized boundary conditions, these equations are solved iteratively using a generalized Newton method to obtain the carrier concentrations and electric potential at each mesh point. Each Newton iteration involves the solution of a very large matrix equation of order $3N$, where N is the number of mesh points. One-dimensional simulations typically utilize on the order of 1000 mesh points, so the matrix is 3000×3000. In 2D, the minimum mesh is typically at least 100×100, so $N = 10^4$ and the matrix is of order 3×10^4 and contains 9×10^8 elements. Fortunately, the matrices are sparse and can be solved using considerably less computer memory than one would expect at first glance.

Numerical simulation allows analysis of solar cell designs and operating conditions for which simple analytic expressions are inadequate. The necessity of ignoring the spatial variation of parameters is eliminated and more accurate representations of the solar cell are possible. In particular, nonuniform doping, heterojunction solar cells (the bandgap varies spatially), amorphous silicon solar cells (complex trapping/recombination mechanisms, field-assisted collection), and concentrator solar cells (high-level injection, 2D/3D effects) can all be modeled with more precision.

3.5 SUMMARY

It has been the objective of this chapter to give the reader a basic understanding of the physical principles that underlie the operation of solar cells. Toward that end, the fundamental physical characteristics of solar cell materials that permit the conversion of light into electricity have been reviewed. These characteristics include the ability of semiconductors to absorb photons by conferring that energy to carriers of electrical current and the ability of semiconductor materials to conduct electricity.

The basic operating principles of the solar cell (a carefully designed *pn*-junction diode) were derived from the (simplified) equations describing the dynamics of holes and electrons in semiconductors. This led to the definition of the solar cell figures of merit – the open-circuit voltage (V_{OC}), the short-circuit current (I_{SC}), the fill factor (FF), and the cell efficiency (η). The two key factors determining solar cell efficiency – electron–hole pair generation and recombination – were identified and discussed. In particular, the need to minimize all sources of recombination in the solar cells was demonstrated through examples.

The importance of matching the bandgap of the solar cell material to the solar spectrum was also discussed and it was shown that silicon, with a bandgap of 1.12 eV, is a reasonably good match to the solar spectrum. The effects of parasitic resistances and temperature on solar cell performance were examined and, finally, some advanced cell concepts were briefly introduced. Many of these topics will be expanded upon in the following chapters of this handbook.

REFERENCES

1. Green M, *Solar Cells: Operating Principles, Technology, and System Applications*, Chap. 1, Prentice Hall, Englewood Cliffs, NJ, 1–12 (1982).
2. Pierret R, in Pierret R, Neudeck G (Eds), *Modular Series on Solid State Devices, Volume VI: Advanced Semiconductor Fundamentals*, Addison-Wesley, Reading, MA (1987).
3. Sze S, *Physics of Semiconductor Devices*, 2nd Edition, John Wiley & Sons, Inc., New York, NY (1981).
4. Böer K, *Survey of Semiconductor Physics: Electrons and Other Particles in Bulk Semiconductors*, Van Nostrand Reinhold, New York, NY (1990).
5. Shur M, *Physics of Semiconductor Devices*, Prentice Hall, Englewood Cliffs, NJ (1990).
6. Singh J, *Physics of Semiconductors and Their Heterostructures*, McGraw-Hill, New York, NY (1993).
7. Pierret R, *Semiconductor Device Fundamentals*, Chap. 2, Addison-Wesley, Reading, MA, 23–74 (1996).
8. Slotboom J, De Graff H, *Solid-State Electron.* **19**, 857–862 (1976).
9. Pankove J, *Optical Processes in Semiconductors*, Chap. 3, Dover Publications, New York, NY, 34–81 (1971).
10. Sanii F, Giles F, Schwartz R, Gray J, *Solid-State Electron.* **35**, 311–317 (1992).
11. Barnett A, Honsberg C, Kirkpatrick D, *et al*., *Proc. 4th World Conference on Photovoltaic Energy Conversion*, 2560–2564 (2006).

12. Pierret R, in Pierret R, Neudeck G (Eds), *Modular Series on Solid State Devices, Volume VI: Advanced Semiconductor Fundamentals*, Chap. 5, Addison-Wesley, Reading, MA, 139–179 (1987).
13. Dubin S, Gray J, *IEEE Trans. Electron Devices* **41**, 239–245 (1994).
14. Pierret R, in Pierret R, Neudeck G (Eds), *Modular Series on Solid State Devices, Volume VI: Advanced Semiconductor Fundamentals*, Chap. 6, Addison-Wesley, Reading, MA (1987).
15. Van Roosbroeck W, *Bell Syst. Tech. J.* **29**, 560–607 (1950).
16. Sacconi F, Di Carlo A, Lugli P, Morkoc H, *IEEE Trans. Electron Devices* **48**, 450–457 (2001).
17. Lundstrom M, Schulke R, *IEEE Trans. Electron Devices* **30**, 1151–1159 (1983).
18. Gray J, Haas A, Wilcox J, Schwartz R, *Proc. 33rd IEEE Photovoltaic Specialist Conf.*, 1–6 (2008).
19. Gray J, Schwartz R, *Proc. 18th IEEE Photovoltaic Specialist Conf.*, 568–572 (1985).
20. Mathers C, *J. Appl. Phys.* **48**, 3181, 3182 (1977).
21. Shockley W, Queisser H, *J. Appl. Phys.* **32**, 510 (1961).
22. Green M, *Solar Cells: Operating Principles, Technology, and System Applications*, Chap. 5, Prentice Hall, Englewood Cliffs, NJ, 85–102 (1982).
23. Hegedus S, Desai D, Thompson C, *Prog Photovolt: Res. Applic.* **15**, 587–602 (2007).
24. Green M, Emery K, King D, Hisikawa Y, Warta W, *Prog Photovolt: Res. Applic.* **14**, 45–51 (2006).
25. Gray J, *Two-Dimensional Modeling of Silicon Solar Cells*, Ph.D. thesis, Purdue University, West Lafayette, IN (1982).
26. Baraona C, Brandhorst Jr. H, *Proc. 11th IEEE Photovoltaic Specialist Conf.*, 44–48 (1975).
27. Chai A, *Proc. 14th IEEE Photovoltaic Specialist Conf.*, 156–160 (1980).
28. Sinton R, Kwark Y, Swanson R, *Proc. 18th IEEE Photovoltaic Specialist Conf.*, 61–64 (1984).
29. Rover D, Basore P, Thorson G, *Proc. 18th IEEE Photovoltaic Specialist Conf.*, 703–709 (1984).
30. Rubinelli F *et al*., *Proc. 22nd IEEE Photovoltaic Specialist Conf.*, 1405–1408 (1991).
31. Gray J, *Proc. 22nd IEEE Photovoltaic Specialist Conf.*, 436–438 (1991).
32. Lundstrom M, *Numerical Analysis of Silicon Solar Cells*, Ph.D. thesis, Purdue University, West Lafayette, IN (1980).
33. Gray J, *IEEE Trans. Electron Devices* **36**, 906–912 (1989).
34. Snowden C, *Introduction to Semiconductor Device Modeling*, Chap. 2, World Scientific, Singapore, 14–36 (1986).

4

Theoretical Limits of Photovoltaic Conversion and New-generation Solar Cells

Antonio Luque and Antonio Martí

Instituto de Energía Solar, Universidad Politécnica de Madrid, Spain

4.1 INTRODUCTION

Efficiency is an important matter in the photovoltaic (PV) conversion of solar energy because the sun is a source of power whose density is not very low, so it gives some expectations on the feasibility of its generalised cost-effective use in electric power production. However, this density is not so high as to render this task easy. After over a third of a century of attempting it, cost still does not allow a generalised use of this conversion technology.

Efficiency forecasts have been carried out from the very beginning of PV conversion to guide the research activity. In solar cells the efficiency is strongly related to the generation of electron–hole pairs caused by the light, and their recombination before being delivered to the external circuit at a certain voltage. This recombination is due to a large variety of mechanisms and cannot be easily linked to the material used to make the cell. Nevertheless, already in 1975 Lofersky [1] had established an empirical link that allowed him to predict which materials were most promising for solar cell fabrication.

In 1960, Shockley and Queisser [2] pointed out that the ultimate recombination mechanisms – impossible to avoid – was just the detailed balance counterpart of the generation mechanisms. This allowed them to determine the maximum efficiency to be expected from a solar cell. This efficiency limit (40.7% for the photon spectrum approximated by a black body at 6000 K) is not too high because solar cells make rather ineffective use of the sun's photons. Many of them are not absorbed, and the energy of many of the absorbed ones is only poorly exploited.

Revisiting the topic of efficiency limits is pertinent because today there are renewed attempts to invent and develop novel concepts in solar cells – sometimes known as third-generation

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

cells [3] – aimed towards obtaining a higher efficiency. The oldest of them [4], and well established today, is the use of multijunction solar cells, to be studied in detail in Chapter 9 of this book. However, novel attempts are favoured by the general advancement of science and technology that puts new tools in the inventor's hands, such as nanotechnology, not available just a decade ago.

Conventional solar cells are semiconductor devices in which an interaction between electrons and holes is produced by absorption of light photons in order to produce electrical work. Unlike solar thermal converters, they can operate at room temperature.[1] In this chapter we shall start by presenting some of the irreversible thermodynamic background regarding the photon–electron interaction. Special attention is paid to the conditions for complying with the second law of the thermodynamics as a guide for new device inventors.

This thermodynamic approach will give us the efficiency that a certain solar converter can attain under certain ideal conditions. This will be applied not only to present solar cells, but also to a number of proposed new converters, some of them actually tried experimentally and some not.

The solar converter theory developed in this chapter looks very different from the one presented in Chapter 3 and is complementary to it. There the solar cell will be analysed in a solid-state physics context, on the basis of conventional photo-excitation of an electron from valence to conduction band, followed by electric carrier transport and recombination. In this chapter all materials will be considered ideal. Entropy-producing mechanisms will be reduced to those inherent to the concept being studied and all other mechanisms (surface and bulk nonradiative recombination, nonideal contacts) will be ignored. In this way, the efficiencies given in this chapter are to be considered as upper bounds of the solar cells studied both in Chapter 3 and the remaining chapters of this book. The lower values found there are not necessarily due to poor technology. They may be fundamental when linked to the actual materials and processes used to manufacture real solar cells. However, they are not fundamental in the sense that other materials and processes could, in principle, be found where different materials or process limitations will apply, which perhaps might be less restrictive.

Compared with the first edition, this newly revised chapter contains more discussion of the experimental results and attempts at implementing these concepts, although to keep the size unchanged some theoretical developments have been removed.

4.2 THERMODYNAMIC BACKGROUND

4.2.1 Basic Relationships

Thermodynamics defines a state function for a system in equilibrium. This function can be the entropy, the internal energy, the canonical grand potential, the enthalpy and so on, all of them mutually equivalent, containing the same information of the system and related to each other through the Legendre transformation [5]. The variables used to describe the macroscopic state of a system are divided into *extensive* (volume U, number of particles N, entropy S, internal energy E, etc.) and *intensive* (pressure P, electrochemical potential μ, temperature T, etc.) variables. A number of relationships may be written with them. For instance,

$$\mathrm{d}E = T\,\mathrm{d}S - P\,\mathrm{d}U + \mu\,\mathrm{d}N \tag{4.1}$$

For our purposes it is convenient to choose the grand potential Ω as our preferred state function. It is the state function that describes the state of a system when the electrochemical

[1] Solar cells usually operate at some 40–60 °C, but there is nothing theoretical against improving cooling to reach a cell temperature as close to the ambient one as desired.

potential, the volume and the temperature are chosen as independent variables. It is also the Legendre transform of the energy with respect to the temperature and the electrochemical potential:

$$\Omega \doteq E - TS - \mu N = -PU \tag{4.2}$$

Its definition is given by the equal-by-definition sign ($\doteq$). The right-hand equality represents a property proven with generality [6].

Other thermodynamic variables that characterise the system in equilibrium can be obtained from the grand potential as derivatives. Hence, the number of particles, entropy and pressure of the system under consideration can be obtained as

$$N = -\left.\frac{\partial \Omega}{\partial \mu}\right|_{U,T}; \quad S = -\left.\frac{\partial \Omega}{\partial T}\right|_{U,\mu}; \quad P = -\left.\frac{\partial \Omega}{\partial U}\right|_{\mu,T} \tag{4.3}$$

For the description of systems in nonequilibrium, the system under study is assumed to be subdivided into small subsystems, each one comprising an elementary volume in the phase space[2] (x, y, z, v_x, v_y, v_z) and having a size sufficient to allow the definition of thermodynamic magnitudes in it. Within such volumes, the subsystems are assumed to be in equilibrium. Thus, thermodynamic magnitudes are dependent on the position $\boldsymbol{r} \doteq (x, y, z)$ of the elementary volume and the velocity of motion $\boldsymbol{v} \doteq (v_x, v_y, v_z)$ of the elementary bodies or particles in it.

To describe a system in nonequilibrium, it is necessary to introduce the concept of *thermodynamic current densities* [7], $\boldsymbol{j}_x$. They are related to the extensive variables X and are defined for those elementary bodies with velocity $\boldsymbol{v}$ at the point $\boldsymbol{r}$ as follows:

$$\boldsymbol{j}_x(\boldsymbol{r}, \boldsymbol{v}) = x(\boldsymbol{r}, \boldsymbol{v})\boldsymbol{v} \tag{4.4}$$

where $x = X/U$ is the contribution to the extensive variable X, per unit of volume U at the point $\boldsymbol{r}$ of the elementary bodies with velocity $\boldsymbol{v}$.

Equation (4.2) can be applied to the thermodynamic current densities, allowing us to write

$$\boldsymbol{j}_\omega = \boldsymbol{j}_e - T(\boldsymbol{r}, \boldsymbol{v})\boldsymbol{j}_s - \mu(\boldsymbol{r}, \boldsymbol{v})\boldsymbol{j}_n = -P(\boldsymbol{r}, \boldsymbol{v})\boldsymbol{v} \tag{4.5}$$

For the thermodynamic current densities, we can write the following continuity equations [8]:

$$g \doteq \frac{\partial n}{\partial t} + \nabla \cdot \boldsymbol{j}_n \tag{4.6}$$

$$\upsilon \doteq \frac{\partial e}{\partial t} + \nabla \cdot \boldsymbol{j}_e \tag{4.7}$$

$$\sigma \doteq \frac{\partial s}{\partial t} + \nabla \cdot \boldsymbol{j}_s \tag{4.8}$$

where g, υ and σ are formally defined as the number of particles, energy and entropy generation rates per unity of volume. The symbol "$\nabla\cdot$" is the divergence operator.[3]

[2] x, y and z are the spatial coordinates giving the position of the particle and v_x, v_y and v_z are the coordinates of its velocity.

[3] The linear unbounded operator "$\nabla\cdot$" applied to the vector $\boldsymbol{A} \doteq (A_x, A_y, A_z)$ is defined as $\nabla \cdot \boldsymbol{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$.

4.2.2 The Two Laws of Thermodynamics

Equations (4.7) and (4.8) have close links with the laws of thermodynamics. A certain elementary subsystem or body can draw energy from or release energy to another body close to it, but the first law of thermodynamics states that the sum of the energies generated at all the i elementary bodies at a given position $\boldsymbol{r}$ must be zero, that is,

$$\sum_i \upsilon(\boldsymbol{r}, \boldsymbol{v}_i) = 0 \tag{4.9}$$

In the same way, the entropy generated by an elementary body can possibly be negative, but the second law of thermodynamics, as stated by Prigogine [8], determines that the sum of all the entropy generated by all the bodies, σ_{irr}, must be non-negative everywhere.

$$\sum_i \sigma(\boldsymbol{r}, \boldsymbol{v}_i) = \sigma_{\text{irr}}(\boldsymbol{r}) \geq 0 \tag{4.10}$$

4.2.3 Local Entropy Production

It is illustrative to have a look at the sources for entropy production. Using Equation (4.2) in the basic thermodynamic relationship of Equation (4.1), we obtain an interesting relationship between thermodynamic variables per unit volume:

$$\mathrm{d}s = \frac{1}{T}\,\mathrm{d}e - \frac{\mu}{T}\,\mathrm{d}n \tag{4.11}$$

If this relationship and Equation (4.5) are substituted in Equation (4.8), we find that

$$\sigma = \frac{1}{T}\frac{\partial e}{\partial t} - \frac{\mu}{T}\frac{\partial n}{\partial t} + \nabla \cdot \left(\frac{1}{T}\boldsymbol{j}_e - \frac{1}{T}\boldsymbol{j}_\omega - \frac{\mu}{T}\boldsymbol{j}_n\right) \tag{4.12}$$

Introducing Equations (4.6) and (4.7) in (4.12) and after some mathematical handling we obtain that [9]

$$\sigma = \frac{1}{T}\upsilon + \boldsymbol{j}_e \nabla\frac{1}{T} - \frac{\mu}{T}g - \boldsymbol{j}_n \nabla\frac{\mu}{T} + \nabla \cdot \left(-\frac{1}{T}\boldsymbol{j}_\omega\right) \tag{4.13}$$

where "∇" is the gradient operator.[4] This is an important equation allowing us to identify the possible sources of entropy generation in a given subsystem. It contains terms involving energy generation (from another subsystem: υ) and transfer (from the surroundings: $\boldsymbol{j}_e \nabla 1/T$), free energy ($\mu g$) generation, Joule effect ($\boldsymbol{j}_n \nabla \mu$) and existence of a gradient pressure ($\nabla \boldsymbol{j}_\omega / T$). This equation is very important and will be used to prove the thermodynamic consistency of solar cells.

4.2.4 An Integral View

Fluxes, $\dot{X}$, of the thermodynamic currents, $\boldsymbol{j}_x$, will be frequently used in this chapter. In this text they will be also called *thermodynamic variable rates*. By definition, the following relationship exists:

$$\dot{X} \doteq \int_A \sum_i \boldsymbol{j}_x \cdot \mathrm{d}\boldsymbol{A} \tag{4.14}$$

[4] The gradient operator, "∇" (without the dot) is a linear unbounded operator whose action on a scalar field $f(x, y, z)$ is defined as

$$\nabla f = \frac{\partial f}{\partial x}\boldsymbol{e}_1 + \frac{\partial f}{\partial y}\boldsymbol{e}_2 + \frac{\partial f}{\partial z}\boldsymbol{e}_3$$

where $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ are the orthonormal basis vectors of the Cartesian framework being used as reference.

where the sum refers to the different subsystems with different velocities to be found at a given position. A is the surface through which the flux is calculated. Actually, $\boldsymbol{j}_x \cdot \mathrm{d}\mathbf{A}$ represents the scalar product of the current density vector $\boldsymbol{j}_x$ and the oriented surface element $\mathrm{d}\boldsymbol{A}$ (orientation is arbitrary and if a relevant volume exists the orientation selected leads to the definition of *escaping* or *entering* rates).

4.2.5 Thermodynamic Functions of Radiation

The number of photons in a given mode of radiation is given [10] by the well-known Bose–Einstein factor $f_{\mathrm{BE}} = \{\exp[(\varepsilon - \mu)/kT] - 1\}^{-1}$, which through Equation (4.3) is related to the grand canonical potential $\Omega = kT \ln\{\exp[(\mu - \varepsilon)/kT] - 1\}$. In these equations, most of the symbols have been defined earlier: ε is the photon energy in the mode and k is the Boltzmann constant. The corresponding thermodynamic current densities for these photons are

$$\boldsymbol{j}_n = f_{\mathrm{BE}}\mathbf{c}/(Un_{ref}); \quad \boldsymbol{j}_e = \varepsilon f_{\mathrm{BE}}\mathbf{c}/(Un_{ref}); \quad \boldsymbol{j}_\omega = \Omega\mathbf{c}/(Un_{ref}) \tag{4.15}$$

where $\mathbf{c}$ is the velocity of light (a vector since it includes its direction) in vacuum and n_{ref} is the index of refraction of the medium in which the photons propagate, which is assumed to be independent of the direction of propagation. Thus, $\mathbf{c}/n_{ref}$ is the velocity of the photons in the medium.

The number of photon modes with energy between ε and $\varepsilon + \mathrm{d}\varepsilon$ is $8\pi U n_{ref}^3 \varepsilon^2/(h^3c^3)\,\mathrm{d}\varepsilon$. When the modes with energies $\varepsilon_m < \varepsilon < \varepsilon_M$ are taken into account, the total grand canonical potential of the photons, Ω_{ph}, associated with these modes is the sum of the contributions from each mode and can be written as

$$\Omega_{\mathrm{ph}}(U, T, \mu) = \frac{8\pi U n_{ref}^3}{h^3c^3} \int_{\varepsilon_m}^{\varepsilon_M} \varepsilon^2 kT \ln(1 - \mathrm{e}^{(\mu-\varepsilon)/kT})\,\mathrm{d}\varepsilon \tag{4.16}$$

where h is the Planck's constant.

Photons do not interact among themselves, so that temperatures and chemical potentials can be different for each mode. This means that they can be a function of the energy and of the direction of propagation. In the non-equilibrium case they can also be a function of the position. If we only take into account the photons' propagation in a small solid angle $\mathrm{d}\varpi$, the grand potential in equation (4.16) must be multiplied by $\mathrm{d}\varpi/4\pi$. The same coefficient affects other thermodynamic variables of the radiation.

The flux of a thermodynamic variable X of the radiation through a surface A with a solid angle ϖ is then given by

$$\dot{X} = \int_A \sum_i \boldsymbol{j}_x \cdot \mathrm{d}\boldsymbol{A} = \int_{A,\varpi} \frac{1}{4\pi}\frac{X}{U}\frac{c}{n_{ref}} \cos\theta \mathrm{d}\varpi\,\mathrm{d}A = \int_H \frac{1}{4\pi}\frac{X}{U}\frac{c}{n_{ref}^3}\,\mathrm{d}H \tag{4.17}$$

where the angle θ is defined in Figure 4.1. In this case, the sum of equation (4.14) has been substituted by the integration over solid angles. In many cases the integration will be extended to a restricted domain of solid angles. It is, in particular, the case of the photons when they come from a remote source such as the sun.

The differential variable $\mathrm{d}H = n_{ref}^2 \cos\theta \mathrm{d}\varpi\,\mathrm{d}A$, or its integral on a certain domain (at each position of A it must include the solid angle ϖ containing photons), is the so-called multilinear Lagrange invariant [11]. It is invariant for any optical system [12]. For instance, at the entry aperture of a solar concentrator (think of a simple lens), the bundle of rays has a narrow angular dispersion at its entry since all the rays come from the sun within a narrow cone. Then, they are collected

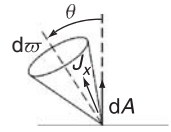


Figure 4.1 Diagram to show the flux of a thermodynamic variable across a surface element d***A***

across the whole entry aperture. The invariance for H indicates that it must take the same value at the entry aperture and at the receiver, or even at any intermediate surface that the bundle may cross. If no ray is turned back, all the rays will be present at the receiver. However, if this receiver is smaller than the entry aperture, the angular spread with which the rays illuminate the receiver has to be bigger than the angular spread that they have at the entry. In this way, H becomes a sort of measure of a bundle of rays, similar to its four-dimensional area with two spatial dimensions (in dA) and two angular dimensions (in dϖ). Thus, we may talk of the H_{sr} of a certain bundle of rays linking the sun with a certain receiver.

Besides *Lagrange invariant*, this invariant may have other names. In treatises of thermal transfer it is called *vision* or *view factor*, but Welford and Winston [13] have recovered for this invariant the old name given by Poincaré that in our opinion accurately reflects its properties. They refer to it as *étendue* (extension) of a bundle of rays. We shall adopt in this chapter this name as a shortened denomination for this *multilinear Lagrange invariant*.

When the solid angle of illumination consists of the total hemisphere, $H = n_{ref}^2 \pi A$, where A is the area of the surface traversed by the photons. However, in the absence of optical elements, the photons from the sun reach the converter located on the Earth within a narrow cone of rays. In this case, $H = \pi A \sin^2 \theta_s$, where θ_s is the sun's semi-angle of vision (taking into account the sun's radius and its distance to Earth), which is equal [14] to 0.265°. No index of refraction is used in this case because the photons come from a vacuum.

Fluxes of some thermodynamic variables and some formulas of interest, which can be obtained from the above principles, are collected in Table 4.1.

4.2.6 Thermodynamic Functions of Electrons

For electrons, the number of particles in a quantum state with energy ε is given by the Fermi–Dirac function $f_{\mathrm{FD}} = \{\exp[(\varepsilon - \varepsilon_F)/kT] + 1\}^{-1}$ where the electrochemical potential (the chemical potential including the potential energy due to electric fields) of the electrons is usually called the *Fermi level* ε_F. Unlike photons, electrons are in continuous elastic interaction with the very abundant phonons that randomise their direction so that, unlike photons, their thermodynamic properties are seldom direction dependent (an exception being ballistic electrons). They are also in strong inelastic interaction among themselves, and with some phonons so that finding different temperatures for electrons is rather difficult. In fact, once a monochromatic light pulse is shone on the semiconductor, the electrons or holes thermalise to a uniform internal electron temperature in less than 100 fs. Cooling to the lattice temperature will take approximately 10–20 ps [15]. However, electrons in semiconductors are found in two bands separated by a large energy gap in which electron states cannot be found. The consequence of this is that, in nonequilibrium, different electrochemical potentials, ε_{Fc} and ε_{Fv} (also called *quasi-Fermi levels*), can exist for the electrons in the conduction band (CB) and the valence band (VB), respectively. Sometimes we prefer to refer to the electrochemical potential of the holes (empty states at the VB), which is then equal to $-\varepsilon_{Fv}$. Once the excitation is suppressed, it can even take milliseconds before the two quasi-Fermi levels become a single one, as it is in the case of single-crystal high-quality silicon.

Table 4.1 Several thermodynamic fluxes for photons with energies between ε_m and ε_M distributed according to a black body law determined by a temperature T and chemical potential μ. The last line involves the calculations for a full-energy spectrum ($\varepsilon_m = 0$ and $\varepsilon_M = \infty$). One of the results (the one containing T^4) constitutes the Stefan–Boltzmann law

$$\dot{\Omega}(T, \mu, \varepsilon_m, \varepsilon_M, H) = kT\frac{2H}{h^3c^2}\int_{\varepsilon_m}^{\varepsilon_M} \ln(1 - \mathrm{e}^{(\mu-\varepsilon)/kT})\varepsilon^2 \mathrm{d}\varepsilon = \int_{\varepsilon_m}^{\varepsilon_M} \dot{\omega}(T, \mu, \varepsilon, H)\,\mathrm{d}\varepsilon \quad \text{(I-1)}$$

$$\dot{N} = (T, \mu, \varepsilon_m, \varepsilon_M, H) = \frac{2H}{h^3c^2}\int_{\varepsilon_m}^{\varepsilon_M} \frac{\varepsilon^2 d\varepsilon}{e^{(\varepsilon-\mu)/kT} - 1} = \int_{\varepsilon_m}^{\varepsilon_M} \dot{n}(T, \mu, \varepsilon, H)\,\mathrm{d}\varepsilon \quad \text{(I-2)}$$

$$\dot{E}(T, \mu, \varepsilon_m, \varepsilon_M, H) = \frac{2H}{h^3c^2}\int_{\varepsilon_m}^{\varepsilon_M} \frac{\varepsilon^3 d\varepsilon}{e^{(\varepsilon-\mu)/kT} - 1} = \int_{\varepsilon_m}^{\varepsilon_M} \dot{e}(T, \mu, \varepsilon, H)\,\mathrm{d}\varepsilon \quad \text{(I-3)}$$

$$\dot{S} = \frac{\dot{E} - \mu\dot{N} - \dot{\Omega}}{T}; \quad \dot{F} \doteq \dot{E} - T\dot{S} = \mu\dot{N} + \dot{\Omega} \quad \text{(I-4)}$$

$$\dot{E}(T, 0, 0, \infty, H) = (H/\pi)\sigma_{\mathrm{SB}}T^4; \dot{S}(T, 0, 0, \infty, H) = (4H/3\pi)\sigma_{\mathrm{SB}}T^3;$$

$$\sigma_{\mathrm{SB}} = 5.67 \times 10^{-8}\,\mathrm{Wm^{-2}\,K^{-4}} \quad \text{(I-5)}$$

4.3 PHOTOVOLTAIC CONVERTERS

4.3.1 The Balance Equation of a PV Converter

In 1960, Shockley and Queisser (SQ) published an important paper [2] in which the efficiency upper limit of a solar cell was presented. In this article it was pointed out, for the first time in solar cells, that the generation due to light absorption has a detailed balance counterpart, which is the radiative recombination. This SQ efficiency limit, detailed below, occurs in ideal solar cells that are the archetype of the current solar cells. Such cells (Figure 4.2) are made up of a semiconductor with a VB and a CB, which is more energetic and separated from the VB by the band gap, ε_g. Each band is able to develop separate quasi-Fermi levels, ε_{Fc} for the CB and ε_{Fv} for the VB, to describe the carrier concentration in the respective bands. In the ideal SQ cell the mobility of the carriers is infinite, and since the electron and hole currents are proportional to the quasi-Fermi level gradients times the mobility, it follows that the quasi-Fermi levels are constant. Contact to the CB is made by depositing a metal on an n^+-doped semiconductor. The carriers going through this contact are mainly electrons due to the small hole density. The few holes passing through this contact are accounted for as surface recombination, assumed zero in the ideal case. Similarly, contact with the VB is made with a metal deposited on a p-doped semiconductor. The metal Fermi levels ε_{F+} and ε_{F-} are levelled, respectively, to the hole and the electron quasi-Fermi levels ε_{Fv} and ε_{Fc} at each interface. In equilibrium, the two quasi-Fermi levels become just one.

The voltage V appearing between the two electrodes is given by the splitting of the quasi-Fermi levels, or, more precisely, by the difference in the quasi-Fermi levels of majority carriers

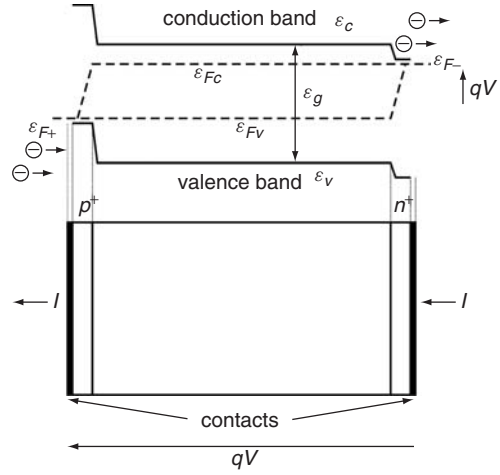


Figure 4.2 Band diagram of an ideal solar cell under illumination with its contacts

at the ohmic contact interfaces. With constant quasi-Fermi levels and ideal contacts, the split is simply

$$qV = \varepsilon_{Fc} - \varepsilon_{Fv} \tag{4.18}$$

Photons are absorbed by pumping electrons from the VB to the CB through the process known as electron–hole pair generation. However, as required by the detailed balance, the opposite mechanism is also produced so that a CB electron can decay to the VB and emit a photon, leading to what is called a *radiative recombination process*, responsible for luminescence. In fact, many of such luminescent photons whose energy is slightly above the band gap are reabsorbed, leading to generation of new electron–hole pairs. Only the recombination processes leading to the effective emission of a photon out of the semiconductor produce a net recombination. Taking into account that the luminescent photons are emitted isotropically, only those photons emitted near the cell faces, at distances in the range of or smaller than the inverse of the absorption coefficient, and directed towards the cell faces with small angles (those reaching the surface with angles higher than the limit angle will be reflected back) have chances to actually leave the semiconductor, and thus to contribute to the net radiative recombination. This is generally referred to as photon recycling [16], and including it in solar cell modelling is difficult and not yet very common. Note that radiative recombination is negligible in c-Si due to the indirect bandgap, and in thin films due to the higher density of localized midgap defect states.

In the ideal SQ cell any nonradiative recombination mechanism, which is an entropy-producing mechanism, is assumed to be absent.

The difference between the electrons pumped to the CB by external photon absorption and those falling again into the VB and effectively emitting a luminescent photon equals the current extracted from the cell. This can be presented in an equation form as

$$I/q = \dot{N}_s - \dot{N}_r = \int_{\varepsilon_g}^{\infty} (\dot{n}_s - \dot{n}_r)\, \mathrm{d}\varepsilon \tag{4.19}$$

where $\varepsilon_g = \varepsilon_c - \varepsilon_v$ is the semiconductor band gap and $\dot{N}_s$ and $\dot{N}_r$ are the photon fluxes entering or leaving the solar cell, respectively, through any surface. When the cell is properly contacted, this

current is constituted by the electrons that leave the CB through the highly doped n-contact. In a similar balance, in the VB, I/q are also the electrons that enter the VB through the highly doped p-contact. Note that the sign of the current is the opposite to that of the flow of the electrons.

Using the nomenclature in Table 4.1, the terms in Equation (4.19) for unit-area cells are $\dot{N}_s = a\dot{N}(T_s, 0, \varepsilon_g, \infty, \pi \sin^2\theta_s)$ for the cell facing the sun directly or $\dot{N}_s = a\dot{N}(T_s, 0, \varepsilon_g, \infty, \pi)$ for the cell under full concentration and $\dot{N}_r = \xi\dot{N}(T_a, qV, \varepsilon_g, \infty, \pi)$ where a and ξ are the absorptivity and emissivity of the cell. T_s is the sun temperature and T_a is the room temperature. Full concentration means using a concentrator without losses that is able to provide isotropic illumination; this is the highest illuminating power flux from a given source. The conservation of the étendue requires this concentrator to have a concentration C fulfilling the equation $C\pi \sin^2\theta_s = \pi n_{ref}^2$, that is, $C = 46050 n_{ref}^2$. This concentration is indeed unrealistic, but it does lead to the highest efficiency. Furthermore, it can be proven [17] that when the quasi-Fermi level split is uniform in the semiconductor bulk, then $a = \xi$. We shall assume – from now on in this chapter – that the solar cell is thick enough and perfectly coated with antireflection layers so as to fully absorb any photon with energy above the band gap energy so that $a = \xi = 1$ for these photons. Assumption of 100% optical conversion is not so unrealistic, the best cells have only 1–3% optical losses for above-bandgap light.

The assumption $\dot{N}_r = \dot{N}(T_a, qV, \varepsilon_g, \infty, \pi)$ states that the temperature associated with the emitted photons is the room temperature T_a. This is natural because the cell is at this temperature. However, it also states that the chemical potential of the radiation emitted, μ_{ph}, is not zero, but

$$\mu_{\text{ph}} = \varepsilon_{Fc} - \varepsilon_{Fv} = qV \tag{4.20}$$

This is so because the radiation is due to the recombination of electron–hole pairs, each one with a different electrochemical potential or quasi-Fermi level. A plausibility argument for admitting $\mu_{\text{ph}} = \varepsilon_{Fc} - \varepsilon_{Fv}$ is to consider that photons and electron–hole pairs are produced through the reversible equation (i.e. not producing entropy): electron + hole $\leftrightarrow$ photon. Equation (4.20) would then result as a consequence of equating the chemical potentials before and after the reaction. Equation (4.20) can be also proven by solving the continuity equation for photons within the cell bulk [17, 18].

When the exponential of the Bose–Einstein function is much higher than 1, the recombination term in Equation (4.19) for full concentration can be written as

$$\begin{aligned}\dot{N}_r &= \frac{2\pi}{h^3c^2}\int_{\varepsilon_g}^{\infty} \varepsilon^2 \exp\left(\frac{qV-\varepsilon}{kT_a}\right) \mathrm{d}\varepsilon \\ &= \frac{2\pi kT}{h^3c^2}[\mathbf{4}(kT)^2 + 2\varepsilon_g kT + \varepsilon_g^2] \exp\left(\frac{qV-\varepsilon_g}{kT_a}\right)\end{aligned} \tag{4.21}$$

This equation is therefore valid for $\varepsilon_g - qV \gg kT_a$. Within this approximation, the current–voltage characteristic of the solar cell takes its conventional *single* exponential appearance. In fact, this equation, with the appropriate factor $\sin^2\theta_s$, is accurate in all the ranges of interest of the current–voltage characteristic of ideal cells under unconcentrated sunlight.

The SQ solar cell can reach an efficiency limit given by

$$\eta = \frac{\{qV[\dot{N}_s - \dot{N}_r(qV)]\}_{\max}}{\sigma_{\text{SB}} T_s^4} \tag{4.22}$$

where the maximum is calculated by optimising V. This efficiency limit was first obtained by Shockley and Queisser [2] (for unconcentrated light) and is plotted in Figure 4.3 for several illumination spectra as a function of the bandgap.

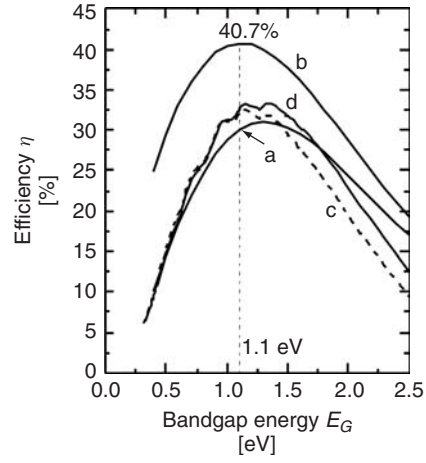


Figure 4.3 SQ efficiency limit for an ideal solar cell versus bandgap energy for unconcentrated black body illumination, for full concentrated illumination and for illumination under the terrestrial sun spectrum: (a) unconcentrated 6000 K black body radiation (1595.9 Wm^{-2}); (b) full concentrated 6000 K black body radiation (7349.0×10^4 Wm^{-2}); (c) unconcentrated AM1.5-Direct [19] (767.2 Wm^{-2}); (d) AM1.5 Global [19] (962.5 Wm^{-2})

Outside the atmosphere the sun is seen quite accurately as a black body whose spectrum corresponds to a temperature of 5758 K [20]. To stress the idealistic approach of this chapter, we do not take this value in most of our calculations, but rather 6000 K for the sun temperature and 300 K for room temperature.

It must be pointed out that the limiting efficiency obtained for full concentration can also be obtained at lower concentrations if the étendue of the escaping photons is made equal to that of the incoming photons [17]. This can be achieved by locating the cell in a cavity [21] that limits the angle of the escaping photons.

It is interesting to note that the SQ analysis limit does not make any reference to semiconductor *pn*-junctions. William Shockley, who first devised the *pn*-junction operation [22], was also the first in implicitly recognising [2] its secondary role in solar cells. In fact, a *pn*-junction is not a fundamental constituent of a solar cell. What seems to be fundamental in a solar cell is the existence of two gases of electrons with different quasi-Fermi levels (electrochemical potentials) and the existence of selective contacts [23] that are traversed by each one of these two gases. The importance of the role of the existence of these selective contacts has not been sufficiently recognised [24]. This is achieved today with *n*- and *p*-doped semiconductor regions, not necessarily forming layers, as in the point contact solar cell [25], but in the future it might be achieved otherwise, maybe leading to substantial advancements in PV technology. The role of the semiconductor, of which the cell is made, is to provide the two gases of electrons that may have different quasi-Fermi levels owing to the gap energy separation that makes the recombination difficult.

So far, for unconcentrated light, the most efficient single-junction solar cell, made of GaAs, has achieved an efficiency of 25.9% [26] of AM1.5G spectrum. This is only 21% (relative) below the highest theoretical efficiency of 32.8% for the GaAs bandgap, of 1.42 eV, for this spectrum [27]. The theoretical maximum almost corresponds to the GaAs bandgap. However, most cells are manufactured so that the radiation is also emitted towards the cell substrate located in the rear face of the cell, and little radiation, if any, turns back to the active cell. The consequence of this is that the étendue of the emitted radiation, which is π for a single face radiating to the air, is

enlarged. The étendue per unity of area is then $\pi + \pi n_{ref}^2$. The term πn_{ref}^2, absent in the ideal case, proceeds from the emission of photons towards the substrate of the cell, which has a refraction index n_{ref}. This reduces the limiting efficiency of the GaAs solar cell from 32.8 to 30.7%. Taking this into account, the efficiency of this best experimental cell is only 15.6% below the achievable efficiency of this cell, only on the basis of radiative recombination. Some substantial increase in efficiency might then be achieved by putting a reflector at the rear side of the active layers of the cells [28] (not behind the substrate!). This would require the use of thin GaAs solar cells [29] or the fabrication of Bragg reflectors [30] (a stack of thin semiconductor layers of alternating refraction indices) underneath the active layers. Bragg reflectors have been investigated for enhancing the absorption of the incoming light in very thin cells, but the reduction of the luminescent emission towards the substrate might be an additional motivation.

4.3.2 The Monochromatic Cell

It is very instructive to consider an ideal cell under monochromatic illumination. When speaking of monochromatic illumination, we mean that, in fact, the cell is illuminated by photons within a narrow interval of energy $\Delta\varepsilon$ around the central energy ε. The monochromatic cell must also prevent the luminescent radiation of energy outside the range $\Delta\varepsilon$ from escaping from the converter.

For building this device an ideal concentrator [31] can be used that collects the rays from the solar disc, with an angular acceptance of just θ_s, with a filter on the entry aperture, letting the aforementioned monochromatic illumination pass through. This concentrator is able to produce isotropic illumination at the receiver. By a reversal of the ray directions, the rays issuing from the cell in any direction are to be found also at the entry aperture with directions within the cone of semi-angle θ_s. Those with the proper energy will escape and be emitted towards the sun. The rest will be reflected back into the cell where they will be recycled. Thus, under ideal conditions no photon will escape with energy outside the filter energy and, furthermore, the photons escaping will be sent directly back to the sun with the same étendue of the incoming bundle H_{sr}.

The current in the monochromatic cell, ΔI, will then be given by

$$\begin{aligned} \Delta I/q &\equiv i(\varepsilon, V)\Delta\varepsilon/q = (\dot{n}_s - \dot{n}_r)\Delta\varepsilon \\ &= \frac{2H_{sr}}{h^3c^2}\left[\frac{\varepsilon^2\Delta\varepsilon}{\exp\left(\dfrac{\varepsilon}{kT_s}\right)-1} - \frac{\varepsilon^2\Delta\varepsilon}{\exp\left(\dfrac{\varepsilon - qV}{kT_a}\right)-1}\right] \end{aligned} \tag{4.23}$$

This equation allows for defining an equivalent cell temperature T_r,

$$\frac{\varepsilon}{kT_r} = \frac{\varepsilon - qV}{kT_a} \Rightarrow qV = \varepsilon\left(1 - \frac{T_a}{T_r}\right) \tag{4.24}$$

so that the power produced by this cell, $\Delta\dot{W}$, is

$$\begin{aligned} \Delta\dot{W} &= \frac{2H_{sr}}{h^3c^2}\left[\frac{\varepsilon^3\Delta\varepsilon}{\exp\left(\dfrac{\varepsilon}{kT_s}\right)-1} - \frac{\varepsilon^3\Delta\varepsilon}{\exp\left(\dfrac{\varepsilon}{kT_r}\right)-1}\right]\left(1 - \frac{T_a}{T_r}\right) \\ &= (\dot{e}_s - \dot{e}_r)\Delta\varepsilon\left(1 - \frac{T_a}{T_r}\right) \end{aligned} \tag{4.25}$$

We realise that the work extracted from the monochromatic cell is the same as that extracted from a Carnot engine fed with a heat rate from the hot reservoir $\Delta\dot{q} = (\dot{e}_s - \dot{e}_r)\Delta\varepsilon$. However, this similarity does not hold under a non-monochromatic illumination because, for a given voltage, the equivalent cell temperature would depend on the photon energy ε being unable to define a single equivalent temperature for the whole spectrum. Note that the equivalent cell temperatures corresponding to short-circuit and open-circuit conditions are T_a and T_s, respectively.

To calculate the efficiency, $\Delta\dot{W}$ in Equation (4.25) must be divided by the appropriate denominator. We could divide by the black body incident energy $\sigma_{\mathrm{SB}} T_s^4$, but this would be unfair because the unused energy that is reflected by the entry aperture could be deflected with an optical device and used in other solar converters. We can divide by $\dot{e}_s \Delta\varepsilon$, the rate of power received at the cell, thus obtaining the *monochromatic* efficiency, η_{mc}, given by

$$\eta_{\mathrm{mc}} = \left.\frac{q(\dot{n}_s - \dot{n}_r)V}{\dot{e}_s}\right|_{\max} = \left.\left(1 - \frac{\dot{e}_r}{\dot{e}_s}\right)\left(1 - \frac{T_a}{T_r}\right)\right|_{\max} \tag{4.26}$$

that is represented in Figure 4.4 as a function of the energy ε.

Alternatively, we could have used the standard definition of efficiency used in thermodynamics [32, 33] to compute the efficiency of the monochromatic cell. In this context, we put in the denominator the energy really wasted in the conversion process, that is, $(\dot{e}_s - \dot{e}_r)\Delta\varepsilon$. Actually, the energy $\dot{e}_r \Delta\varepsilon$ is returned to the sun, perhaps for later use (slowing down, for example, the sun's energy loss process!). This leads to the *thermodynamic* efficiency:

$$\eta_{\mathrm{th}} = \left(1 - \frac{T_a}{T_r}\right) \tag{4.27}$$

This efficiency is the same as the Carnot efficiency obtained by a reversible engine operating between an absorber at temperature T_r and the ambient temperature, and suggests that an ideal solar cell may work reversibly, without entropy generation. Its maximum, for $\Delta\dot{W} \geq 0$, is obtained for $T_r = T_s$, although unfortunately it occurs when $\Delta\dot{W} = 0$, that is, when negligible work is done by the cell (actually, none). However, this is a general characteristic of reversible engines, which yield the Carnot efficiency only at the expense of producing negligible power.

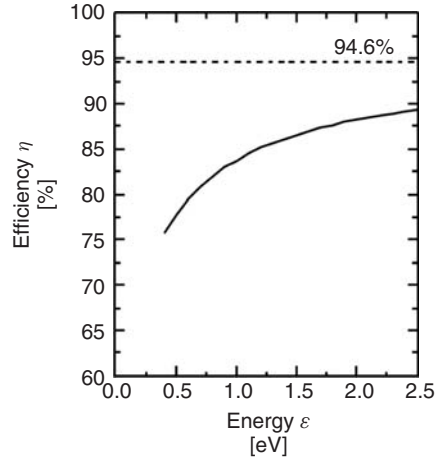


Figure 4.4 Monochromatic cell efficiency versus photon energy

4.3.3 Thermodynamic Consistency of the Shockley–Queisser Photovoltaic Cell

Electrons and photons are the main particles interacting in a solar cell [9]. However, other interactions occur as well. In general, looking at Equation (4.13), the generation of entropy for each kind of particle (electrons, photons and others) is written as

$$\sigma = \sum_i \left[\frac{1}{T} \upsilon + \boldsymbol{j}_e \nabla \frac{1}{T} - \frac{\mu}{T} g - \boldsymbol{j}_n \nabla \frac{\mu}{T} + \nabla \cdot \left(-\frac{1}{T} \boldsymbol{j}_\omega \right) \right] \tag{4.28}$$

where the summation extends to the different states of the particle.

First we analyse the generation of entropy by electrons, σ_{ele}. Except for the case of ballistic electrons, the pressure – appearing inside $\boldsymbol{j}_\omega$ as shown in Equation (4.5) – is very quickly equilibrated; it is the same for $+\boldsymbol{v}$ and for $-\boldsymbol{v}$ as a result of the frequent elastic collisions, mostly with the ubiquitous acoustic phonons. Thus, the term in $\boldsymbol{j}_\omega$ for electrons disappears when the sum is extended to all the states in each energy. These collisions also cause the temperature and the free energy to become the same for all directions, at least for a given energy. Furthermore, in conventional solar cells, the temperature of the electrons is the same for any energy and is equal to the lattice temperature T_a. Also, the electrochemical potential is the same for all the electrons within the same band at a given $\boldsymbol{r}$.

In a real cell, the energy flow $\boldsymbol{j}_e$ goes from the high-temperature regions towards the low-temperature ones $[\nabla(1/T) \geq 0]$ so that the term in Equation (4.28) involving $\boldsymbol{j}_e$ produces positive entropy. However, in the ideal cell, the lattice temperature, which is also that of the electrons, is constant and the term involving $\nabla(1/T)$ disappears. The electron flow, $\boldsymbol{j}_n$, opposes the gradient of electrochemical potential, thus also producing positive entropy for constant temperature. Furthermore, in the SQ [2] ideal cell, mobility is infinite and, therefore, its conduction and valence band electrochemical potentials or quasi-Fermi levels (ε_{Fc}, ε_{Fv}) are constant throughout the whole solar cell and their gradients are also zero. Therefore, all the gradients in Equation (4.28) disappear and the electron contribution to entropy generation, σ_{ele}, is given by

$$\sigma_{\text{ele}} = \sum_{i-\text{ele}} \left[\frac{1}{T_a} \upsilon_{i-\text{ele}} - \frac{\varepsilon_{Fc(v)}}{T_a} g_{i-\text{ele}} \right] \tag{4.29}$$

The quasi-Fermi level to be used in this case is ε_{Fc} or ε_{Fv} depending on the band to which the electronic state i-ele belongs. This is represented by $\varepsilon_{Fc(v)}$.

Other interactions may occur in the cell involving other particles besides the electrons and the photons. In principle they are phonons. Phonons are very abundant; much more than CB electrons in a nondegenerate semiconductor. Their density of modes at room temperature is three times the density of atoms and each mode is populated with some few phonons on average.

We shall assume that in these interactions the bodies involved (labelled as *others*) also have a direction-independent pressure (elastic collisions among phonons due to anharmonic interaction are very frequent) and that they are also at the lattice temperature. Furthermore, they are assumed to have zero chemical potential (the high density of phonons and their frequent anharmonic interaction makes it unlikely for them to be out of equilibrium). Using these assumptions, the contribution to the irreversible entropy generation rate from these *other* particles becomes

$$\sigma_{\text{others}} = \sum_{i-\text{others}} \left[\frac{1}{T_a} \upsilon_{i-\text{others}} \right] \tag{4.30}$$

For the case of the photons, the situation is rather different. As said before, photons do not interact with each other and, therefore, they are essentially ballistic. Their thermodynamic intensive variables may experience variations with the photon energy and also with their direction of propagation. In fact, they come from the sun in a few directions only, and only in these directions do they exert a pressure. The direct consequence is that a nonvanishing current of grand potential exists for the photons. (It vanishes in gases of photons that are confined and in thermal equilibrium with the confining walls. Using this condition for photon beams from the sun, as is sometimes done, is not correct.)

N_{ph} being the number of photons in a mode corresponding to a certain ray moving inside the semiconductor, their evolution along a given ray path corresponding to a radiation mode is given by [18]

$$N_{\mathrm{ph}}(\varsigma) = f_{\mathrm{BE}}(T, qV)\left\lfloor 1 - \mathrm{e}^{-\alpha\varsigma}\right\rfloor + N_{\mathrm{ph}}(0)\mathrm{e}^{-\alpha\varsigma} \tag{4.31}$$

where ς is the length coordinate along the ray, f_{BE} is the Bose–Einstein factor for luminescent photons whose chemical potential is the separation between the conduction and the valence band electron quasi-Fermi levels – in this case equalling the cell voltage V (times q) – and α is the absorption coefficient. Equation (4.31) shows a nonhomogeneous profile for N_{ph} contributed to by luminescent photons that increase with ς (first term on the right-hand side) and externally fed photons that decrease when the ray proceeds across the semiconductor (second term), and these photons are absorbed. $N_{\mathrm{ph}}(0) = f_{\mathrm{BE}}(T_s, 0)$ is usually taken in solar cells that correspond to illumination by free (i.e. with zero chemical potential) radiation at the sun temperature T_s.

In general, the photons in a mode of energy ε are considered as a macroscopic body [10] for which temperature and chemical potential can be defined. However, thermodynamically, they can be arbitrarily characterised by a chemical potential μ and a temperature T as long as $(\varepsilon - \mu)/T$ takes the same value. For example, the incident solar photons may be considered at the solar temperature T_s with zero chemical potential or, alternatively, at room temperature T_a with an energy variable chemical potential $\mu_s = \varepsilon(1 - T_a/T_s)$. This property has already been used in the study of the monochromatic cell.

Indeed, this arbitrary choice of T and μ does not affect the entropy production. This becomes evident if we rewrite Equation (4.13) in the case of photons as

$$\sigma_{\mathrm{ph}} = \sum_{i-\mathrm{ph}}\left[\frac{\varepsilon - \mu}{T}g + \boldsymbol{j}_n\nabla\left(\frac{\varepsilon - \mu}{T}\right) + \nabla\cdot\left(-\frac{1}{T}\boldsymbol{j}_\omega\right)\right] \tag{4.32}$$

where we have made use of Equation (4.15) and of the fact that $\upsilon = \varepsilon g$. This equation depends explicitly only on $(\varepsilon - \mu)/T$. This is less evident in the term in $\boldsymbol{j}_\omega/T$, but as discussed in the context of Equation (4.15), $\boldsymbol{j}_\omega$ is proportional to T and thus $\boldsymbol{j}_\omega/T$ only depends on $(\varepsilon - \mu)/T$. Note, however, that $\boldsymbol{j}_\omega$ is affected by the specific choice of T.

For simplicity, we shall consider the photons at room temperature and we shall calculate their chemical potential, μ_{ph}, from setting the equality

$$N_{\mathrm{ph}}(\varsigma, \varepsilon) = \frac{\varepsilon^2}{\exp\left[\dfrac{\varepsilon - \mu_{\mathrm{ph}}(\varsigma, \varepsilon)}{kT_a}\right] - 1} \tag{4.33}$$

However, we might have chosen to leave $\mu_{\mathrm{ph}} = 0$, and then the effect of the absorption of light when passing through the semiconductor could have been described as a cooling down of the photons (if N_{ph} actually decreases with ς).

With the room-temperature luminescent photon description, the production of entropy by photons is given by

$$\sigma_{\mathrm{ph}} = \sum_{i-\mathrm{ph}} \left[\frac{\upsilon_{i-\mathrm{ph}}}{T_a} - \frac{\mu_{i-\mathrm{ph}} g_{i-\mathrm{ph}}}{T_a} - \frac{\boldsymbol{j}_{n,i-\mathrm{ph}} \nabla \mu_{i-\mathrm{ph}}}{T_a} - \frac{\nabla \boldsymbol{j}_{\omega,i-\mathrm{ph}}}{T_a} \right] \tag{4.34}$$

However, using Equations (4.15) and (4.3),

$$\nabla \boldsymbol{j}_{\omega} = \frac{c}{U n_{ref}} \frac{\mathrm{d}\Omega}{\mathrm{d}\mu} \nabla \mu = -\frac{c}{U n_r} f_{\mathrm{BE}} \nabla \mu = -\boldsymbol{j}_n \nabla \mu \tag{4.35}$$

and

$$\sigma_{\mathrm{ph}} = \sum_{i-\mathrm{ph}} \left[-\frac{\upsilon_{i-\mathrm{ph}}}{T_a} + \frac{\mu_{i-\mathrm{ph}} g_{i-\mathrm{ph}}}{T_a} \right] \tag{4.36}$$

where

$$g = (c/U n_{ref}) \alpha f_{\mathrm{BE}}(T_a, qV) \mathrm{e}^{-\alpha \varsigma} - (c/U n_{ref}) \alpha f_{\mathrm{BE}}(T_s, 0) \mathrm{e}^{-\alpha \varsigma}; \quad \upsilon = \varepsilon g \tag{4.37}$$

The total irreversible entropy production is obtained by adding Equations (4.29), (4.30) and (4.36), taking into account Equation (4.37). Now, the terms in energy generation, all at the same temperature, must balance out by the first law of thermodynamics. The net absorption of photons corresponds to an electron transfer (positive and negative generations) between states gaining an electrochemical potential qV, so the terms $\mu g/T$ corresponding to electrons and photons cancel each other out. No additional generations are assumed to take place. Thus the total irreversible entropy generation rate can be written as

$$\sigma_{\mathrm{irr}} = \frac{c\alpha}{U n_{ref}} \sum_{i-\mathrm{ph}} \frac{(\mu_{i-\mathrm{ph}} - qV)[f_{\mathrm{BE}}(T_s, 0) \mathrm{e}^{-\alpha \varsigma} - f_{\mathrm{BE}}(T_a, qV) \mathrm{e}^{-\alpha \varsigma}]}{T_a} \tag{4.38}$$

For a given mode the second factor balances out when $f_{\mathrm{BE}}(T_s, 0) = f_{\mathrm{BE}}(T_a, qV)$, and so does the irreversible entropy generation. In this case, $N_{\mathrm{ph}} = f_{\mathrm{BE}}(T_a, qV)$ is constant along the ray and $\mu_{i-\mathrm{ph}} = qV$ is constant at all points. The irreversible entropy generation rate is then zero everywhere. If $f_{\mathrm{BE}}(T_s, 0) > f_{\mathrm{BE}}(T_a, qV)$, then $N_{\mathrm{ph}} > f_{\mathrm{BE}}(T_a, qV)$ and $\mu_{i-\mathrm{ph}} > qV$, so that both factors are positive and so is the product. If $f_{BA}(T_s, 0) < f_{BA}(T_a, qV)$, then $N_{\mathrm{ph}} < f_{BA}(T_a, qV)$ and $\mu_{i-\mathrm{ph}} < qV$. In this case both factors are negative and the product is positive.

Thus, we have proven that every mode contributes non-negatively to the entropy. We can then state that the SQ cell produces non-negative entropy and, in this sense, it complies with the second law of thermodynamics.

For less idealised cases, as we have mentioned above, the existence of quasi-Fermi levels or temperature gradients generally produces additional positive entropy. Nonradiative net recombination of electrons from the conduction to the valence band also produces positive entropy. However, net generation would contribute to the production of negative entropy and, therefore, it may incur in violation of the second law of thermodynamics if no other mechanism contributing to the creation of positive entropy exists. So the inclusion of imaginative carrier generation rates in novel device proposals, without counterparts, must be considered with some caution.

4.3.4 Entropy Production in the Whole Shockley–Queisser Solar Cell

The preceding approach is applicable to regions in which the physical properties of the system are continuous and differentiable, but not to abrupt interfaces. For testing the compliance with the second law, the continuity equations must be integrated in such cases by choosing volumes surrounding the suspected interfaces. This integral approach can also be extended to the whole converter to check for any violation of the second law and also to calculate the whole entropy production of a device. Note, however, that, if we accept the Prigogine formulation of the second law of thermodynamics (that requires *local* non-decrease of the entropy and not only global), the integral approach, if not complemented with the local approach, is valid for proving thermodynamic inconsistency, but not for proving consistency, which has to be proven at every point.

In the integral analysis we follow steps similar to those used in the local analysis. In particular, the first law of thermodynamics is applied by integrating Equation (4.7) and using the first law expressed by Equation (4.9). Then, we obtain, for the stationary case

$$0 = \int_A \sum_i \boldsymbol{j}_{e,i} \cdot \mathrm{d}\boldsymbol{A} = +\dot{E}_r - \dot{E}_s + \dot{E}_{mo} - \dot{E}_{mi} + \dot{E}_{\text{others}} \tag{4.39}$$

where $\dot{E}_s$ and $\dot{E}_r$ are the radiation energies entering or escaping from the converter, $\dot{E}_{mi}$ and $\dot{E}_{mo}$ are the energies of the electrons entering the VB and leaving the CB, respectively, and $\dot{E}_{\text{others}}$ is the net flow of energy leaving the semiconductor as a result of other mechanisms.

Taking Equation (4.5) into account, the fact that no chemical potential is associated with *other* particles and processes, and the annihilation of the grand canonical potential current density, the term corresponding to the other elements is more conventionally written as

$$\dot{E}_{\text{others}} = T_a \dot{S}_{\text{others}} \doteq \dot{Q} \tag{4.40}$$

where $\dot{Q}$ is defined as the rate of heat leaving the converter.

The second law of thermodynamics, expressed by Equations (4.8) and (4.10), integrated into the whole volume of the converter can be written, for the stationary case, as

$$\dot{S}_{\text{irr}} = \int_U \sigma_{\text{irr}} \,\mathrm{d}U = \int_A \sum_i \boldsymbol{j}_{s,i} \cdot \mathrm{d}\boldsymbol{A} = \dot{S}_r - \dot{S}_s + \dot{S}_{mo} - \dot{S}_{mi} + \dot{S}_{\text{others}} \tag{4.41}$$

The irreversible rate of entropy production is obtained by the elimination of the terms with subscript *others* by multiplying Equation (4.41) by T_a, subtracting Equation (4.39) from the result and taking into account Equation (4.40). In this way we obtain

$$T_a \dot{S}_{\text{irr}} = (\dot{E}_s - T_a \dot{S}_s) - (\dot{E}_r - T_a \dot{S}_r) + (\dot{E}_{mi} - T_a \dot{S}_{mi}) - (\dot{E}_{mo} - T_a \dot{S}_{mo}) \tag{4.42}$$

From Equation I-4 (Table 4.1) and considering the annihilation of the grand canonical potential flow for electrons, $(\dot{E}_{mi} - T_a \dot{S}_{mi}) = \varepsilon_{Fv} \dot{N}_{mi}$ and $(\dot{E}_{mo} - T_a \dot{S}_{mo}) = \varepsilon_{Fc} \dot{N}_{mo}$, so that we can write

$$T_a \dot{S}_{\text{irr}} = \varepsilon_{Fv} \dot{N}_{mi} - \varepsilon_{Fc} \dot{N}_{mo} + (\dot{E}_s - T_a \dot{S}_s) - (\dot{E}_r - T_a \dot{S}_r) \tag{4.43}$$

Taking into account that $\dot{N}_{mi} = \dot{N}_{mo} = I/q$ and $\varepsilon_{Fc} - \varepsilon_{Fv} = qV$, Equation (4.43) is now rewritten as

$$T_a \dot{S}_{\text{irr}} = -\dot{W} + (\dot{E}_s - T_a \dot{S}_s) - (\dot{E}_r - T_a \dot{S}_r) \tag{4.44}$$

Here, this equation has been derived from the *local* model. However, it can also be obtained with more generality from a classical formulation of the second law of thermodynamics [34]. It is valid for ideal as well as for non-ideal devices. The values of the thermodynamic variables to be used in Equation (4.44) are given in Table 4.1. The power $\dot{W}$ – which in other cases will be that of the converter under study – corresponds in this case to the power of an SQ ideal solar cell and is given by the product of Equations (4.18) and (4.19).

We have already discussed the basic ambiguity for the thermodynamic description of any radiation concerning the choice of the temperature and the chemical potential. A useful corollary is derived from this fact [34]. If the power rate produced by a radiation converter depends on the radiation only through its rate of incident energy or number of photons, then, any radiation received or emitted by the converter can be changed into a luminescent radiation at room temperature T_a, and with chemical potential μ_x, without affecting the rate of power and of irreversible entropy produced. The chemical potential μ_x of this equivalent luminescent radiation is linked to the thermodynamic parameters of the original radiation, T_{rad} and μ_{rad}, through the equation

$$\frac{\varepsilon - \mu_{\text{rad}}}{kT_{\text{rad}}} = \frac{\varepsilon - \mu_x}{kT_a} \Rightarrow \mu_x = \varepsilon\left(1 - \frac{T_a}{T_{\text{rad}}}\right) + \mu_{\text{rad}}\frac{T_a}{T_{\text{rad}}} \tag{4.45}$$

Note that, in general, μ_x is a function of the photon energy, ε, as it may also be T_{rad} and μ_{rad}.

The proof of the theorem is rather simple and is brought in here because it uses relationships of instrumental value. It is straightforward to see that

$$\dot{n}_x = \dot{n}_{\text{rad}}; \quad \dot{e}_x = \varepsilon\dot{n}_x = \dot{e}_{\text{rad}}; \quad \dot{\omega}_x/T_a = \dot{\omega}_{\text{rad}}/T_{\text{rad}}; \quad \dot{e}_{\text{rad}} - T\dot{s}_{\text{rad}} = \mu_x\dot{n}_x + \dot{\omega}_x \tag{4.46}$$

where, again, the suffix *rad* labels the thermodynamic variables of the original radiation.

The equality of the energy and the number of photons of the original and equivalent radiation proves that the power production is unchanged. Furthermore, the last relationship can be introduced in Equation (4.44) to prove that the calculation of the entropy production rate also remains unchanged.

Taking all this into account, the application of Equation (4.44) to the SQ solar cell is rather simple. Using the SQ cell model for the power and using the room-temperature equivalent radiation (of chemical potential μ_s) of the solar radiation, we obtain

$$\begin{aligned} T_a\dot{S}_{\text{irr}} &= -\int_{\varepsilon_g}^{\infty} qV[\dot{n}(T_a,\mu_s) - \dot{n}(T_a,qV)]\,\mathrm{d}\varepsilon + \int_{\varepsilon_g}^{\infty}[\mu_s\dot{n}(T_a,\mu_s) + \dot{\omega}(T_a,\mu_s)]\,\mathrm{d}\varepsilon \\ &\quad - \int_{\varepsilon_g}^{\infty}[qV\dot{n}(T_a,qV) + \dot{\omega}(T_a,qV)]\,\mathrm{d}\varepsilon \\ &= \int_{\varepsilon_g}^{\infty}[\dot{\omega}(T_a,\mu_s) - \dot{\omega}(T_a,qV)]\,\mathrm{d}\varepsilon + \int_{\varepsilon_g}^{\infty}[(\mu_s - qV)\dot{n}(T_a,\mu_s)]\,\mathrm{d}\varepsilon \end{aligned} \tag{4.47}$$

The integrand is zero for $qV = \mu_s$, but as μ_s varies with ε it is not zero simultaneously for all ε. To prove that it is positive, we differentiate with respect to qV. Using $\mathrm{d}\dot{\omega}/\mathrm{d}\mu = -\dot{n}$,

$$\frac{\mathrm{d}(T_a\dot{S}_{\text{irr}})}{\mathrm{d}qV} = -\int_{\varepsilon_g}^{\infty}[\dot{n}(T_a,\mu_s) - \dot{n}(T_a,qV)] = I/q \tag{4.48}$$

Thus, for each energy, the minimum of the integrand also appears for $qV = \mu_s(\varepsilon)$. Since this minimum is zero, the integrand is non-negative for any ε and so is the integral, proving also the thermodynamic consistence of this cell from the integral perspective. Furthermore, the minimum of the entropy, which for the non-monochromatic cell is not zero, occurs for open-circuit conditions. However, the ideal monochromatic cell reaches zero entropy production, and therefore reversible operation, under open-circuit conditions, and this is why this cell may reach the Carnot efficiency, as discussed in Section 4.3.2.

4.4 THE TECHNICAL EFFICIENCY LIMIT FOR SOLAR CONVERTERS

We have seen that with the technical definition of efficiency – in whose denominator the radiation returned to the sun is not considered – the Carnot efficiency cannot be reached with a PV converter. What then is the technical efficiency upper limit for solar converters?

We can consider that this limit would be achieved if we could build a converter producing zero entropy [35]. In this case, the power that this converter can produce, $\dot{W}_{\lim}$, can be obtained from Equation (4.44) by setting the irreversible entropy-generation term to zero. Substituting the terms involving emitted radiation by their room-temperature luminescent equivalent, Equation (4.44) becomes

$$\dot{W}_{\lim} = \int_{\varepsilon_g}^{\infty} \{[\dot{e}_s - T_a \dot{s}_s] - [\mu_x(\varepsilon)\dot{n}_x + \dot{\omega}_x]\}\, \mathrm{d}\varepsilon = \int_{\varepsilon_g}^{\infty} \dot{w}_{\lim}(\varepsilon, \mu_x)\, \mathrm{d}\varepsilon \tag{4.49}$$

The integrand should now be maximised [34] with respect to μ_x. For this, we calculate the derivative

$$\frac{\mathrm{d}\dot{w}_{\lim}}{\mathrm{d}\mu_x} = -\dot{n}_x - \frac{\mathrm{d}\dot{\omega}_x}{\mathrm{d}\mu_x} - \mu_x \frac{\mathrm{d}\dot{n}_x}{\mathrm{d}\mu_x} = -\mu_x \frac{\mathrm{d}\dot{n}_x}{\mathrm{d}\mu_x} \tag{4.50}$$

where we have used that the derivative of the grand potential with respect to the chemical potential is the number of particles with a change of sign. This equation shows that the maximum is achieved if $\mu_x = 0$ for any ε, or, in other words, if the radiation emitted is a room-temperature thermal radiation. This is the radiation emitted by all the bodies in thermal equilibrium with the ambient. However, the same result will be also achieved if the emitted radiation is any radiation whose room-temperature luminescent equivalent is a room-temperature thermal radiation.

Now, we can determine this efficiency, according to Landsberg [35], as

$$\begin{aligned} \eta &= \frac{\left(\frac{H_{sr}}{\pi}\right)\left[\left(\sigma T_s^4 - \frac{4}{3}\sigma T_a T_s^3\right) - \left(\sigma T_a^4 - \frac{4}{3}\sigma T_a^4\right)\right]}{(H_{sr}/\pi)\sigma T_s^4} \\ &= 1 - \frac{4}{3}\left(\frac{T_a}{T_s}\right) + \frac{1}{3}\left(\frac{T_a}{T_s}\right)^4 \end{aligned} \tag{4.51}$$

which for $T_s = 6000$ K and $T = 300$ K gives 93.33% instead of 95% of the Carnot efficiency.

No ideal device is known that is able to reach this efficiency. Ideal solar thermal converters, not considered in this chapter, have a limiting efficiency of 85.4% [36, 37] and therefore do not reach this limit. Other high-efficiency ideal devices considered in this chapter do not reach it either. We do not know whether the Landsberg efficiency is out of reach. At least it is certainly an upper limit of the technical efficiency of any solar converter.

4.5 VERY-HIGH-EFFICIENCY CONCEPTS

4.5.1 Multijunction Solar Cells

A conceptually straightforward way of overcoming the fundamental limitation of the SQ cell, already pointed out by SQ, is the use of several solar cells of different bandgaps to convert photons of different energies. A simple configuration to achieve this is by just stacking the cells so that the upper cell has the highest bandgap, and lets the photons with energy less than its bandgap pass through towards the inner cells (Figure 4.5). The last cell in the stack is the one with the narrowest bandgap. Between cells we put low-energy pass filters so that the reflection threshold of each filter is the bandgap of the cell situated above. This prevents the luminescent photons from being emitted for energies different from those with which the photons from the sun are received in each cell. In this configuration, every cell has its own load circuit, and therefore, is biased at a different voltage. It has been shown [33] that a configuration without back reflectors leads to a lower efficiency if the number of cells is finite.

The detailed balance equations explained above can be used to obtain the limit efficiency of the stack. The maximum efficiency is obtained with an infinite number of solar cells, each one biased at its own voltage $V(\varepsilon)$ and illuminated with monochromatic radiation. The efficiency of this cell is given by

$$\eta = \frac{\int_0^\infty \eta_{\mathrm{mc}}(\varepsilon)\dot{e}_s \, \mathrm{d}\varepsilon}{\int_0^\infty \dot{e}_s \, \mathrm{d}\varepsilon} = \frac{1}{\sigma_{\mathrm{SB}} T_s^4} \int_0^\infty \eta_{\mathrm{mc}}(\varepsilon)\dot{e}_s \, \mathrm{d}\varepsilon = \frac{1}{\sigma_{\mathrm{SB}} T_s^4} \int_0^\infty i(\varepsilon, V)V|_{\max} \, \mathrm{d}\varepsilon \tag{4.52}$$

where $\eta_{\mathrm{mc}}(\varepsilon)$ is the monochromatic cell efficiency given by Equation (4.26) and $i(\varepsilon,V)$ was defined in Equation (4.23). For $T_s = 6000$ K and $T_a = 300$ K, the sun and ambient temperature, respectively, this efficiency is [36] 86.8%. This is the highest efficiency limit of known ideal converters.

Multijunction cells emit room-temperature luminescent radiation. This radiation presents, however, a variable chemical potential $\mu(\varepsilon) = qV(\varepsilon)$ and therefore it is not a radiation with zero chemical potential (free radiation). In addition, the entropy produced by this array is positive since the entropy produced by each one of the monochromatic cells forming the stack is positive. None of the conditions for reaching the Landsberg efficiency (zero entropy generation rate and emission

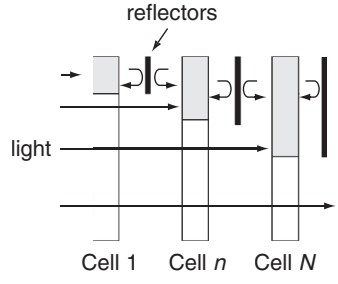


Figure 4.5 Stack of solar cells ordered from left to right in decreasing band gap ($E_{g1} > E_{g2} > E_{g3}$). (Reprinted with permission from Reference [33])

of free radiation at room temperature) is then fulfilled and, therefore, mutijunction cells do not reach this upper bound.

It is highly desirable from a manufacturing economy and reliability perspective to obtain monolithic stacks of solar cells, that is, on the same chip. In this case, the series connection of all the cells in the stack is the most compact solution. Chapter 9 will deal with this case extensively. If the cells are series-connected, a limitation appears that the same current must go through all the cells. The total voltage obtained from the stack is the sum of the voltages in all the cells. The case has been studied in detail [38, 39]. This limitation establishes a link between the voltage in each cell and the current which reduces the value of the maximum achievable efficiency when the number of cells is finite. Our interest now is to determine the top efficiency achievable in this case when the number of cells is increased to infinity. Surprisingly enough, it is found that the solution is also given by Equation (4.52) and therefore, the limiting efficiency of a set of cells that is series-connected is also 86.8%.

A lot of experimental work has been done on this subject. In many of the cell technologies the use of multijunction cells is being used, or at least investigated, as a means of increasing efficiency. To date (September 2010), the highest efficiency of 41.6%, has been obtained by Spectrolab (USA) in 2009, using a monolithic two-terminal lattice matched triple junction cell of GaInP/GaInAs/Ge operating at a concentration factor of 364 suns, that is, at about 36.4 Wcm^{-2} measured under the ASTM G-173-03 direct beam AM1.5 spectrum at a cell temperature of 25 °C. The several cells are series-connected through tunnel junctions, and the final structure comprises about twenty layers of different materials all grown on a Ge single crystal substrate.

Further efficiency improvement could possibly be achieved if the filters mentioned in this section were implemented [39].

4.5.2 Thermophotovoltaic and Thermophotonic Converters

Thermophotovoltaic (TPV) converters are devices in which a solar cell converts the radiation emitted by a heated body into electricity. This emitter may be heated, for example, by the ignition of a fuel. However, in our context, we are more interested in solar TPV converters in which the sun is the source of energy that heats an absorber at temperature T_r, which then emits radiation towards the PV device. Figure 4.6 draws the ideal converter for this situation.

An interesting feature of the ideal TPV converter is that the radiation emitted by the solar cells is sent back to the absorber assisting in keeping it hot. To make the cell area different from the radiator area (so far cells in this chapter have all been considered to be of unit area), the absorber in Figure 4.7 radiates into a reflecting cavity containing the cell. The reflective surfaces in the cavity and the mirrors in the concentrator are assumed to be free of light absorption.

In the ideal case [40], the radiator emits two bundles of rays, one of étendue H_{rc} (emitted by the right side of the absorber in the Figure 4.6a), which is sent to the cell after reflections in the cavity walls, with energy E_r' without losses, and another (unavoidable) one H_{rs}, with energy E_r, which is sent back to the sun (by the left face of the absorber) after suffering reflections in the left-side concentrator. Ideally, the radiator illuminates no other absorbing element or dark part of the sky: all the light emitted by the radiator's left side is sent back to the sun by the concentrator. Rays emitted by the radiator into the cavity may return to the radiator again without touching the cell, but since no energy is transferred, such rays are not taken into account.

On the other hand, the radiator is illuminated by a bundle of rays, coming from the sun, of étendue $H_{sr} = H_{rs}$ and by the radiation emitted by the cell itself, of étendue $H_{cr} = H_{rc}$. Also,

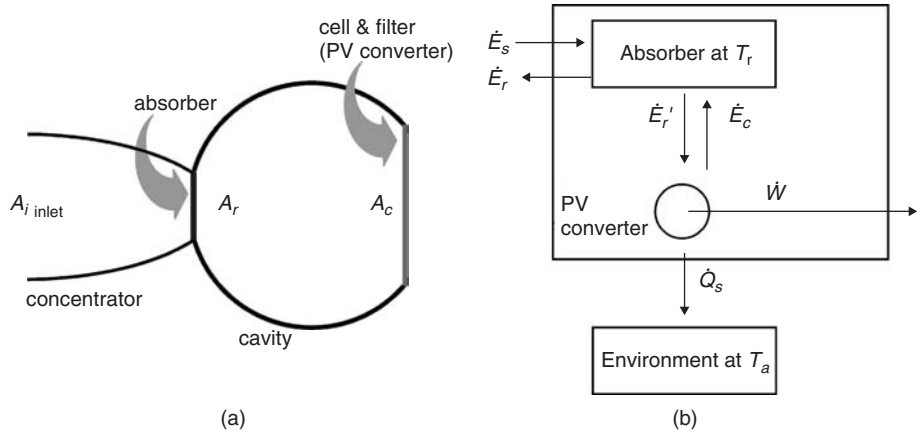


Figure 4.6 (a) Schematic of an ideal TPV converter with the elements inserted in loss-free reflecting cavities; (b) illustration of the thermodynamic fluxes involved. In the monochromatic case, $\dot{E}_s \equiv \dot{E}(T_s, 0, 0, \infty, H_{rs})$, $\dot{E}_r \equiv \dot{E}(T_r, 0, 0, \infty, H_{rs})$, $\dot{E}_r' \equiv \dot{e}(\varepsilon, T_r, 0, H_{rc})\Delta\varepsilon$, $\dot{E}_c \equiv \dot{e}(\varepsilon, T_a, qV, H_{rc})\Delta\varepsilon$

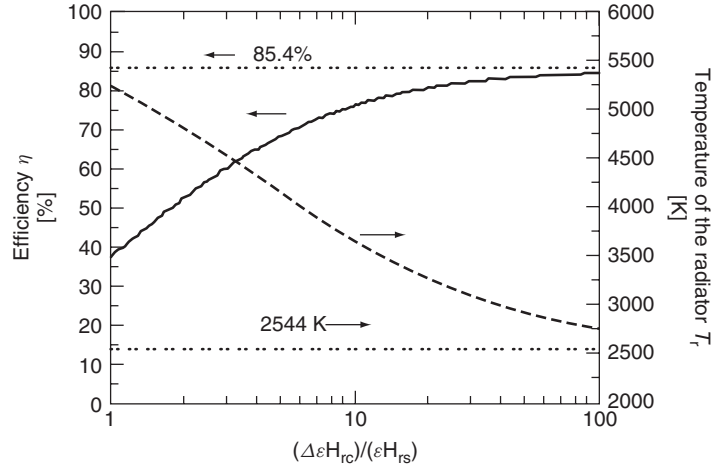


Figure 4.7 TPV ideal converter efficiency versus $H_{rc}\Delta\varepsilon/H_{rs}\varepsilon$. The energy ε and the cell voltage V are optimized

the cell may emit some radiation into the cavity, which returns to the cell again. This radiation is therefore not accounted for as an energy loss in the cell. In addition, we shall assume that the cell is coated with an ideal filter that allows only photons with energy ε and within a bandwidth $\Delta\varepsilon$ to pass through, while the others are totally reflected (no matter the direction). In this situation the energy balance in the radiator becomes

$$\dot{E}(T_s, 0, 0, \infty, H_{rs}) + \dot{e}(\varepsilon, T_a, qV, H_{rc})\Delta\varepsilon = \dot{e}(\varepsilon, T_r, 0, H_{rc})\Delta\varepsilon + \dot{E}(T_r, 0, 0, \infty, H_{rs}) \quad (4.53)$$

where the first equation member is the net rate of energy received by the radiator and the second member is the energy emitted.

Using $\dot{e}(\varepsilon, T_r, 0, H_{rc})\Delta\varepsilon - \dot{e}(\varepsilon, T_a, qV, H_{rc})\Delta\varepsilon = \varepsilon\Delta i/q = \varepsilon\Delta\dot{w}/(qV)$, where Δi is the current extracted from the monochromatic cell and $\Delta\dot{w}$ is the electric power delivered, it is obtained that

$$\frac{\varepsilon\Delta\dot{w}}{qV} = \frac{H_{rs}\sigma_{SB}}{\pi}(T_s^4 - T_r^4) \Leftrightarrow \frac{\boldsymbol{\varepsilon}^2\Delta i}{qH_{rc}\Delta\varepsilon} = \frac{H_{rs}\varepsilon}{H_{rc}\Delta\varepsilon}\frac{\sigma_{SB}}{\pi}(T_s^4 - T_r^4) \tag{4.54}$$

This equation can be used to determine the operation temperature of the radiator, T_r, as a function of the voltage V, the sun temperature T_s, the energy ε and the dimensionless parameter $H_{rc}\Delta\varepsilon/H_{rs}\varepsilon$. Notice that the left-hand side of Equation (4.54) is independent of the cell étendue and of the filter bandwidth (notice that $\Delta i \propto H_{rc}\Delta\varepsilon$).

Dividing by $H_{rs}\sigma_{SB}T_s^4/\pi$, the solar input power, allows expressing the efficiency of the TPV converter as

$$\eta = \left(1 - \frac{T_r^4}{T_s^4}\right)\left(\frac{qV}{\varepsilon}\right) = \left(1 - \frac{T_r^4}{T_s^4}\right)\left(1 - \frac{T_a}{T_c}\right) \tag{4.55}$$

where T_c is the equivalent cell temperature as defined by Equation (4.24). As presented in Figure 4.7, this efficiency is a monotonically increasing function of $H_{rc}\Delta\varepsilon/H_{rs}\varepsilon$. For $(H_{rc}\Delta\varepsilon/H_{rs}\varepsilon) \to \infty$, $\Delta I \to 0$ and $T_r \to T_c$. In this case an optimal efficiency [41] for $T_s = 6000$ K and $T_a = 300$ K is found to be 85.4% and is obtained for a temperature $T_c = T_r = 2544$ K. This is exactly the same as the optimum temperature of an ideal solar thermal converter feeding a Carnot engine. In reality, the ideal monochromatic solar cell is a way of constructing the Carnot engine. This efficiency is below the Landsberg efficiency (93.33%) and slightly below the one of an infinite stack of solar cells (86.8%).

It is worth noting that the condition $(H_{rc}\Delta\varepsilon/H_{rs}\varepsilon) \to \infty$ requires that $H_{rc} \gg H_{rs}$, and for this condition to be achieved, the cell area must be very large compared with the radiator area. This compensates the narrow energy range in which the cell absorbs. This is why a mirrored cavity must be used in this case.

A recent concept for solar conversion has been proposed [3, 42] with the name of thermophotonic (TPH) converter. In this concept, a solar cell converts the luminescent radiation emitted by a heated light-emitting diode (LED) into electricity. As in the TPV device, the LED can be heated with a fuel, but in our context it is heated as well with radiation absorbed from the sun. To emit luminescent radiation, the LED absorbs electric power, in addition to the power delivered from the photons that illuminate the absorber. This power is to be subtracted from the electric power converted by the solar cell. For a detailed analysis, please go to reference [42].

In the ideal case, and for vanishing conversion rates, the efficiency of solar systems using TPH converters is very high. They are strictly equivalent to TPV converters and have the same limiting efficiency. When the conversion of high-power fluxes, expressed as a solar concentration factor, is considered, TPH offers the possibility of high efficiency over 40% in the 100–1000 suns range and at moderate emitter temperatures of about 300 °C. In this respect it is more attractive than TPV. However, the achievement of practical high-efficiency, high-concentration TPH converters, besides the difficulty of fabricating a semiconductor device operating at high temperature, faces very difficult fundamental problems. The external quantum efficiency (photons emitted to electrons received) of the LED should be near unity, otherwise the TPH efficiency drops very abruptly.

4.5.3 Multi-exciton Generation Solar Cells

One of the drawbacks that limit the efficiency of single-junction solar cells is the energy wasted from each photon that is absorbed because it is not converted into electrical power. After the observation

of quantum efficiencies (ratio of electrons delivered to photons received) slightly above unity for high energy photons, Werner, Kolodinski, Brendel and Queisser [43, 44] have proposed a cell in which each photon may generate more than one electron–hole pair. Let us examine the implications of this without discussing a specific physical mechanism. Admitting that every photon may create $M(\varepsilon)$ electron–hole pairs, where M(e) ≥ 1, the current extracted from the device would be given by

$$I/q = \int_{\varepsilon_g}^{\infty} [M(\varepsilon)\dot{n}_s - M(\varepsilon)\dot{n}_r(T, \mu)]\, \mathrm{d}\varepsilon \tag{4.56}$$

In this equation, ε_g is the energy threshold for photon absorption and the factor M in the generation term is our initial assumption. The same term must appear in the recombination term to fulfil the detailed balance: if the sun temperature is brought to the ambient temperature, the current will be zero when $\mu = 0$, only if the factor M also appears in the recombination term. For the moment we are saying nothing about the chemical potential μ of the photons emitted.

The power delivered $\dot{W}$ is given by

$$\dot{W} = \int_{\varepsilon_g}^{\infty} qV[M\dot{n}_s - M\dot{n}_r(T, \mu)]\, \mathrm{d}\varepsilon \tag{4.57}$$

Let us consider a monochromatic cell and calculate the irreversible entropy generation rate $\dot{S}_{\mathrm{irr}}$ in the whole device. With the aid of the general Equation (4.44) and Equation I-4 in Table 4.1, it is given by

$$T_a \dot{S}_{\mathrm{irr}}/\Delta\varepsilon = (\mu_x \dot{n}_x + \dot{\omega}_x) - (\mu\dot{n}_r + \dot{\omega}_r) - qV(M\dot{n}_x - M\dot{n}_r) \tag{4.58}$$

where the source of photons has been substituted by its equivalent room-temperature luminescent radiation characterised by the chemical potential μ_x and the ambient temperature T_a. The open-circuit conditions are achieved when $\mu_{\mathrm{OC}} = \mu_x$. For this value the entropy rate is zero since then $\dot{n}_x = \dot{n}_r$ and $\dot{\omega}_x = \dot{\omega}_r$.

Let us calculate the derivative of the irreversible entropy generation rate (Equation 4.58) with respect to μ and particularise it for the open-circuit value of μ. Considering in what follows V as an unknown function of μ only, and independent of the way of obtaining the excitation (which is the case for infinite mobility) and using the fundamental relationship $\partial\dot{\omega}_r/\partial\mu = -\dot{n}_r$, the result is

$$\left[\frac{\mathrm{d}(T_a \dot{S}_{\mathrm{irr}}/\Delta\varepsilon)}{\mathrm{d}\mu}\right]_{\mu\mathrm{oc}} = (qMV_{\mathrm{OC}} - \mu_{\mathrm{OC}})\left[\frac{\mathrm{d}\dot{n}_r}{\mathrm{d}\mu}\right]_{\mu\mathrm{oc}} \tag{4.59}$$

This derivative is only zero if $qMV_{\mathrm{OC}} = \mu_{\mathrm{OC}}$. Since $\mu_{\mathrm{OC}} = \mu_x$ can take any value by changing the source adequately, we obtain the result $qMV = \mu$. Any other value would produce a negative rate of entropy generation in the vicinity of the open circuit, contrary to the second law of thermodynamics. This is a demonstration, based on the second law of thermodynamics, of the relationship between the chemical potential of the photons and the voltage (or electron and hole quasi-Fermi level split).

Internal quantum efficiency (IQE) greater than one was first found in Si [45, 46] for visible photons of high-energy and UV photons. However, the IQE was only slightly above unity, in the range of 1.3 at 280 Å. The effect was attributed to impact ionization, a mechanism in which the electron or the hole created by the high-energy photon, instead of thermalising by scattering with phonons, by means of impact processes transfers its high energy to a valence-band electron that results pumped into the conduction band. This mechanism has a detailed balance counterpart

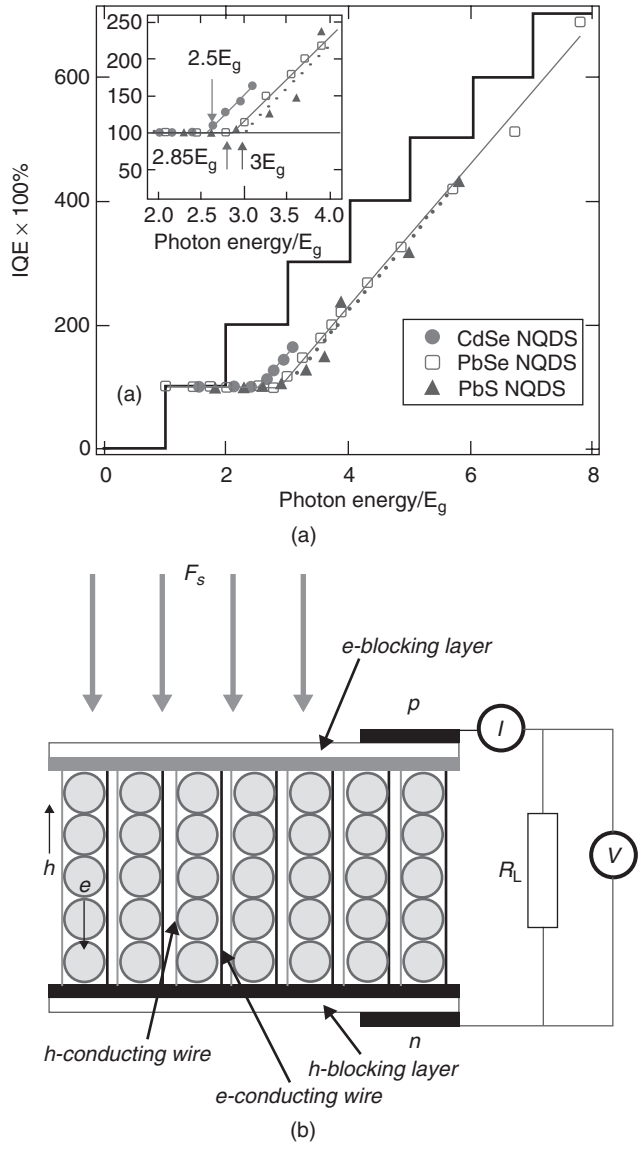


Figure 4.8 (a) Ideal staircase-like internal quantum efficiency (IQE) derived from energy conservation and deduced form measurements in CdSe, PbSe, and PbS nanocrystal quantum dots (NQDs) versus $h\nu$/Eg along with linear fits with a slope of about 1/Eg and carrier multiplication (CM) thresholds of 2.5Eg, 2.85Eg, and 3Eg, respectively (see an expanded view of IQE plots in the inset); (b) a schematic representation of the ideal NQD solar cell, in which each of the dots is in direct electrical contact with both electron- and hole-conducting wires that deliver charges to n and p electrodes, respectively. Short-circuiting of the device is prevented by electron- and hole-blocking layers. RL denotes an external load resistor. F_s is the incident solar fluence (reproduced with permission from Reference [47])

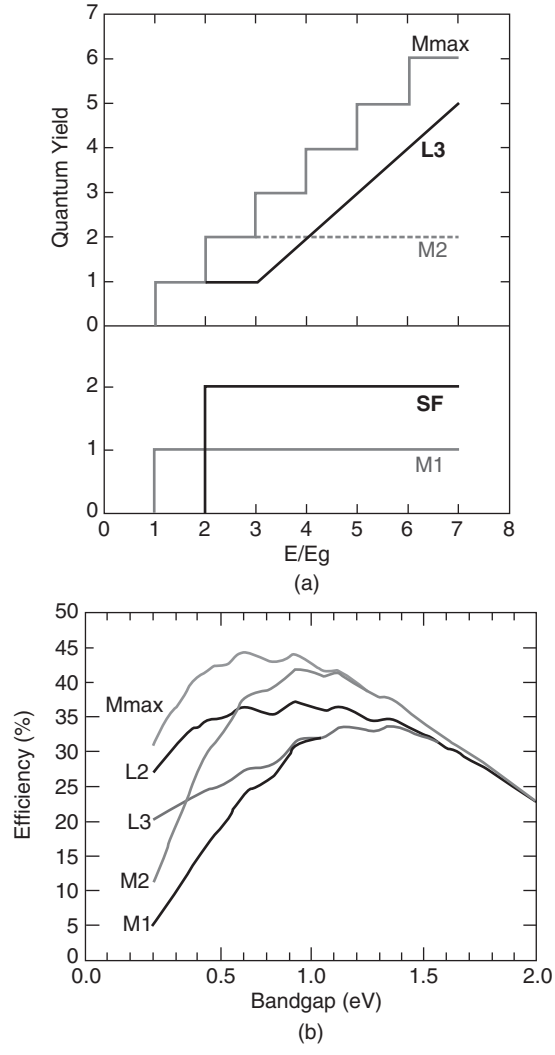


Figure 4.9 Efficiency calculations in MEG solar cells with different internal quantum efficiency (quantum yield) models characterising several types of absorbers. (a) Quantum efficiency models used in the calculations for the following absorber types: multiple-exciton generation (MEG) absorbers M1 and M2 presenting the termodynamic 'staircase', but with *M* limited to 1, 2, etc. and *M*max (full staircase); the (so-called) singlet fission (SF) absorber; a linear IQE absorber (L3) with threshold energy 3Eg and slope 1. (b) PV conversion efficiency at one sun for devices with the following absorber types: MEG absorber with multiplications M1 (this is an ordinary cell) M2, and Mmax, and a linear IQE absorber with threshold energy 3Eg and slope 1 (L3) and threshold energy 2Eg and slope 1 (L2). The L3 curve represents the maximum theoretical efficiency that could be obtained assuming a quantum yield similar to that supposedly measured experimentally in PbSe quantum dots (reproduced with permission from Reference [50])

that is Auger recombination, in which the energy recovered in the recombination is transferred to an electron or a hole, which thus acquires a high kinetic energy.

More recently this effect, also called multi-exciton generation (MEG), has been announced in quantum dots, which are semiconductor nanocrystals of several nanometers acting as artificial atoms with discrete energy levels appearing in the semiconductor bandgap and up to seven excitons (bound electron–holes) have been found generated from a single high-energy photon of 7.8 times the energy gap in PbSe quantum dots (QDs) [48], as presented in Figure 4.8a. Other workers have also measured the MEG effect [49] in other QD materials. However, the way of making a cell with the NQD material, schematically sketched in Figure 4.8b, is unclear (so far, most of the QDs studied for this purpose are in liquid suspension). Nozik and coworkers have calculated the detailed balance efficiency of such cells [50] taking into account the limiting and measured energy dependence of the quantum efficiency, as shown in Figure 4.9a. The MEG cells might present an important efficiency gain, shown in Figure 4.9b, for low bandgaps but note, looking at the efficiency in the cases L2 (whose quantum efficiency is not represented in Figure 4.9a, but is easily inferred from the explanations in the caption) and L3 (closer to experimental data), the important impact that the threshold energy has on it. This is such that the case M2, with a limited $M = 2$, exceeds in potential efficiency the seven-exciton NQD material for practical bandgaps.

However, a controversy has recently arisen because several of the published results have not been reproduced [51] and even the authors of reference [48] have reported that they cannot reproduce their results [52]. To date, all reports relating to the observation of MEG in QDs have been based on several types of spectroscopic measurements (transient absorption) where the MEG quantum efficiency is determined by analysing the optically induced formation of multiple excitons. Those are all indirect methods of measuring MEG. The most direct and definitive method for determining how many electron–hole pairs are created per absorbed photon would be to measure the IQE; i.e. count the electrons in the photocurrent measured in an external circuit connected to a solar cell consisting of QDs where light is absorbed only in the QDs. If the measured quantum efficiency for the photocurrent is greater than 1.0, then there can be no doubt whatsoever that MEG is occurring within the QDs. Such device has not so far (by 2009) been produced [53].

Additionally, some authors [54] have recently presented experimental and theoretical results proving that the carrier multiplication occurs more easily, for a given energy photon, in bulk PbS and PbSe than in QD's made of these materials due to the reduced density of states in the QD material. However, in QDs, the MEG mechanism would take place at a lower photon energy/bandgap energy ratio.

A related concept is the placement of a down-converter [55] on top of an ordinary solar cell. This converts one high-energy photon into two lower-energy ones, with energy closer to the cell bandgap; these down-converted photons illuminate the cell located immediately below. So far the absorption losses in this material have not been able to compensate the multiplication factor of existing down-converters.

4.5.4 Intermediate Band Solar Cell

One of the major causes of efficiency reduction in single-junction solar cells is the transparency of the semiconductor to sub-bandgap photons; i.e. the inability to absorb photons with an energy less than the fundamental bandgap. These photons carry the 36% of the solar energy, which is lost, in the world record 25.9% [26] efficient GaAs cell. The inclusion of an intermediate band (IB) may greatly increase the efficiency by allowing the formation of electron-hole pairs from sub-band gap photons [56]. We show in Figure 4.10 a band diagram of the photon absorption and emission in this

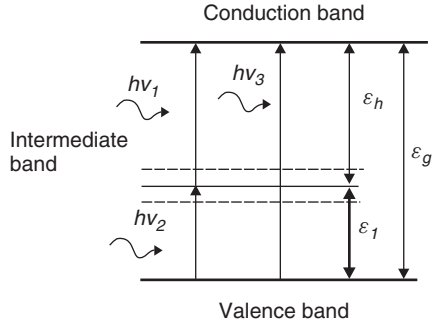


Figure 4.10 The band diagram of an intermediate band solar cell

intermediate band material. Photons are absorbed not only by pumping electrons from the VB to the CB as in a traditional solar cell (photons with energy $h\nu_3$) but also by transitions from the VB to the IB (photons with energy $h\nu_2$) and from the IB to the CB (photons with energy $h\nu_1$). In total, two low-energy photons are used to pump an electron from the valence band to the conduction band, passing through the intermediate band. This certainly increases the cell current.

The three absorption mechanisms detailed above are effective if the IB is a band partially filled with electrons. In this way, there are empty states in it to accommodate the electrons from the VB, and there are electrons to sustain a strong pumping to the CB. The detailed balance imposes photon emissions that are opposite to each one of the three absorption mechanisms.

The cell is contacted as shown in Figure 4.11 [57]. The electrons are extracted from the VB and returned to the CB using two layers of n and p ordinary semiconductors. The voltage is the sum of the two semiconductor junction voltages occurring at both sides of the IB material.

In this IB material we admit that there are three quasi-Fermi levels, one for each band. The cell voltage is q times the splitting of the CB and VB quasi-Fermi levels.

In general, there is an energy threshold for each one of the absorption mechanisms described above. However, the ideal structure is the one in which the upper energy of a photon that can be absorbed in certain mechanisms – involving, for example, transitions from the VB to the IB – is the threshold of the next one – for example, transitions from the IB to the CB. More specifically, calling ε_g the energy interval between the CB and the VB, ε_l the interval between the Fermi level in the intermediate band and the valence band and $\varepsilon_h = \varepsilon_g - \varepsilon_l$, and assuming that $\varepsilon_l < \varepsilon_h$, we consider that all the photons in the interval $(\varepsilon_l, \varepsilon_h)$ are absorbed by transitions from the VB to the IB, all the photons in the interval $(\varepsilon_h, \varepsilon_g)$ are absorbed by transitions from the IB to the CB and all the photons in the interval (ε_g, ∞) are absorbed by transitions from the VB to the CB.

Under such conditions the equations that rule the current–voltage characteristic including contribution to the current I from all three band transitions of the cell are

$$I/q = \left\lfloor \dot{N}(T_s, 0, \varepsilon_g, \infty, \pi) - \dot{N}(T, \mu_{CV}, \varepsilon_g, \infty, \pi) \right\rfloor + [\dot{N}(T_s, 0, \varepsilon_h, \varepsilon_g, \pi) - \dot{N}(T, \mu_{CI}, \varepsilon_h, \varepsilon_g, \pi)] \quad (4.60)$$

$$\dot{N}(T_s, 0, \varepsilon_l, \varepsilon_h, \pi) - \dot{N}(T, \mu_{IV}, \varepsilon_l, \varepsilon_h, \pi) = \dot{N}(T_s, 0, \varepsilon_h, \varepsilon_g, \pi) - \dot{N}(T, \mu_{CI}, \varepsilon_h, \varepsilon_g, \pi) \quad (4.61)$$

with μ_{XY} being the photon chemical potentials equal to the separation of quasi-Fermi levels between band X and band Y. Equation (4.60) states the balance of electrons in the CB (addition of those that are pumped from the IB and the VB by means of the absorption of the corresponding photon less those that recombine radiatively). Equation (4.61) states a similar balance equation for the electrons

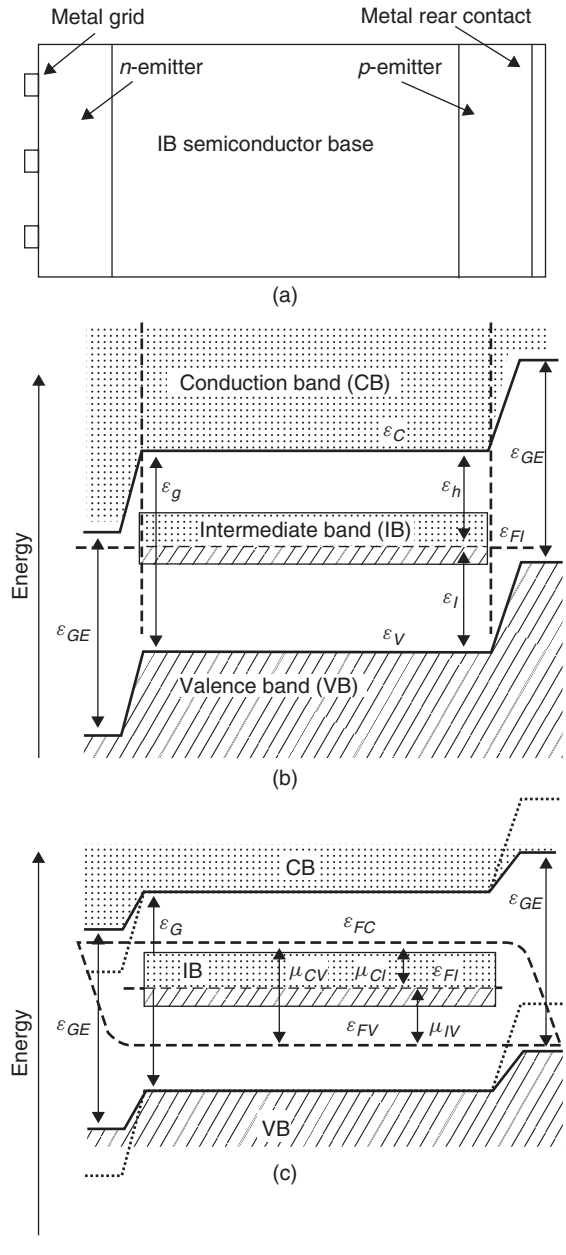


Figure 4.11 (a) Structure of the IB solar cell; (b) band diagram in equilibrium; (c) band diagram of forward-bias conditions (reproduced with permission from Reference [57])

in the IB and takes into account that no current is extracted from the IB. Meanwhile, the voltage is just as expected of a single bandgap solar cell not absorbing sub-bandgap photons, that is:

$$qV = \mu_{CV} = \mu_{CI} + \mu_{IV} \tag{4.62}$$

By eliminating μ_{CI} and μ_{IV} from the last three equations, we obtain the current–voltage characteristic of the cell. Efficiencies for different values of ε_l and ε_g are plotted in Figure 4.12.

The maximum efficiency of 63.2% is achieved for a cell of gap 1.95 eV with the IB Fermi level located at 0.71 eV from one of the bands. This efficiency is higher than the one corresponding to two series-connected ideal cells in tandem, of 54.5% (for bandgaps of 0.8 and 1.54 eV). A detailed analysis of this cell operation can be found in references [56–58], including the effect of overlapping of absorption coefficients in the transitions between different bands. In this respect, it can be shown that overlapping absorption coefficients can also lead to maximum efficiency as long as, for each of the energy intervals involved in the operation of the IB solar cell, the strength of the absorption coefficient associated with it is much higher than the others for the overlapping energies. Engineering the partial filling of the IB could be used to achieve this tailoring of the absorption coefficients. In practice one set of absorption coefficients may be weaker than the others. For this case, the use of light management structures have been proposed in order to increase the absorptivity of the weakest transitions [59, 60].

The generalization of the concept to more than two intermediate band gaps (multiband solar cells) has been studied in [61].

In principle, deep impurity levels could be used to create intermediate bands [62]. However deep levels are known to be a major cause of nonradiative recombination. The mechanism by which this recombination is produced has been extensively studied and, to an extent it is still an open

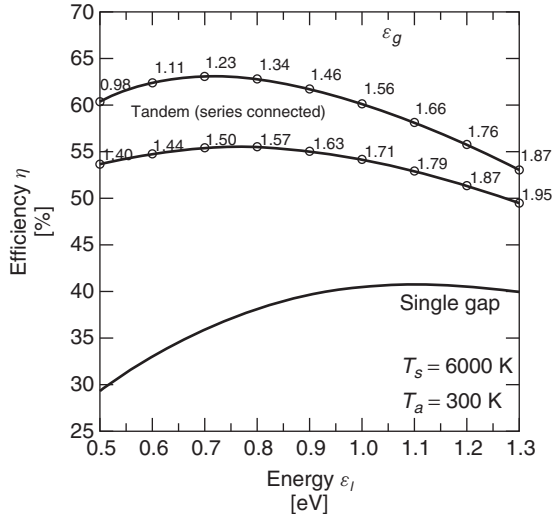


Figure 4.12 Limiting efficiency for the IB solar cell under maximum concentrated sunlight. For comparison, the limiting efficiency of single-gap solar cells (in this case, $\varepsilon_l \equiv \varepsilon_g$) and a tandem of two cells that are series-connected are also shown (in this case, ε_l labels the lowest bandgap of the cells and the figures on the curve correspond to the highest of the bandgaps) (reproduced with permission from Reference [56])

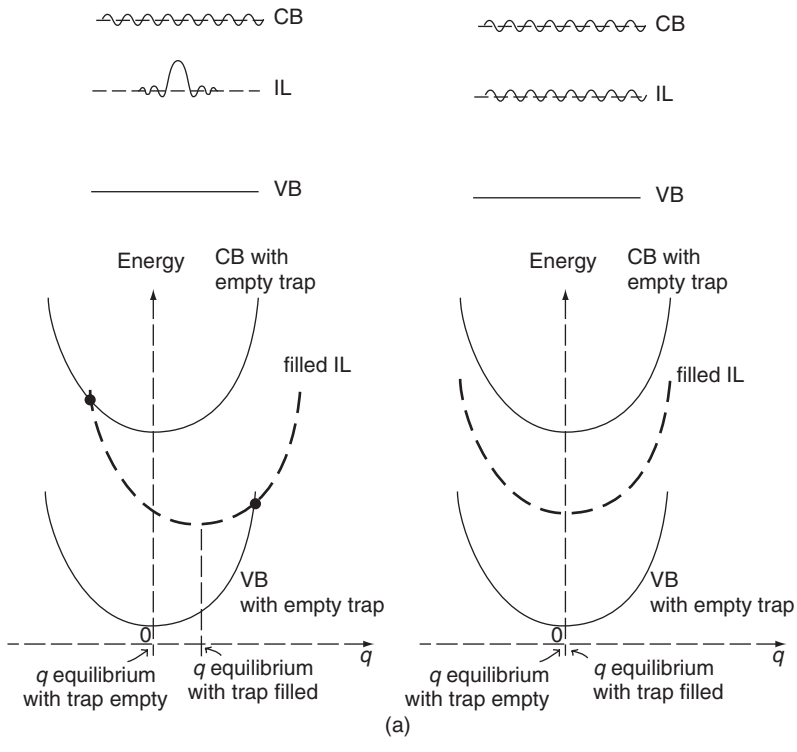


Figure 4.13 (a) Configuration diagram illustrating the potential energy of the nuclear equation (including electrons and nuclei) per electron along the line of maximum potential slope parameterized by q. When the impurity density is low the transition occurs between localised and delocalised states as schematically represented above. When it is high, the impurity states are also delocalised and the transition by the MPE mechanism cannot happen (reproduced with permission from Reference [62]). (b) Si wafer measured lifetime for several levels of Ti ion implantation followed by LPM steps. The higher implant doses, the longer lifetime (reproduced with permission from Reference [64])

question. However, we think that in many cases it is due to the multi-phonon emission (MPE) mechanism proposed by Lang and Henry [63]. In this mechanism the recombination starts with a transition between a CB state, represented by an extended wavefunction, and a localised impurity state. If the transition takes place, the change in the charge density near the impurity atom renders it strongly out of equilibrium with the surrounding lattice. Because of this, its energy is high enough as to enable its fundamental electronic state to enter into the CB state so that the transition becomes possible. A strong vibration takes place that is damped by the sequential (but not simultaneous) emission of many phonons. The second part of the recombination involves the transition from the impurity band localised state to the VB extended state. Again here a similar mechanism takes place.

We have proposed that a high density of impurities should lead to delocalised impurity states and therefore to no significant charge transfer in the transition. Since the nonequilibrium situation is not produced, the recombination is suppressed. Recent evidence of reduced recombination in Si that was very heavily doped with Ti (above 10^{21} cm^{-3}) by ion implant and then recrystallized by pulsed laser melting (PLM) supports the proposed model [64]. It is important to realize that this goes against the common belief that increasing the deep-level impurity density would increase the recombination monotonically, but at the same time explains why in semiconductors the recombination can be low (CB and VB states are all delocalised).

Extensive band calculations by Wahnon and coworkers at the Universidad Politécnica de Madrid (see for instance references [65, 66]) result in a number of possible half-filled IB materials among III–V, II–VI, and I–III–VI_2 semiconductors, not all of them thermodynamically stable. Some of them have been synthesized and light absorption evidence of the IB existence has been clearly produced in $V_{0.25}$ $In_{1.75}S_3$ by Conesa and coworkers at the CSIC Instituto de Catálisis [67].

Prior to this, IB materials had been detected based on the band anti-crossing mechanism produced in highly mismatched alloys. In particular the intermediate band has been detected by photoreflectance spectroscopy in $Zn_{0.88}Mn_{0.12}Te_{0.987}O_{0.013}$ [68] and in $GaN_xAs_{1-x-y}P_y$ alloys with $y > 0.3$ [69]. IB cells have been manufactured with O doped ZnTe showing higher photocurrent and efficiency than the ZnTe cell made for reference, combined with small reduction of voltage [70]. Unfortunately the efficiency of the IB cell and the ordinary cell is in both cases below 1%.

IB behaviour has also been detected in heavily Ti doped Si through Hall effect measurements at different temperatures [71].

The use of the confined levels of QDs has also been proposed to make IB cells [72]. The concept is represented in Figure 4.14. As the IB derives from states originated in the CB they are naturally empty. Donor doping of the barrier material can provide a half-filled IB which is thought to be required for efficient IB solar cell performance [73]. Cells were first made in 2004 growing ten layers of InAs QD's in GaAs [74] by MBE in the Stranski Kastranov mode. A very small current increase was observed in the expected energy region (below 1.41 eV) but, due to the few layers grown, was insufficient to compensate the voltage reduction produced. Nevertheless, the cells experimentally verified the two photon absorption mechanism [75] and the three quasi-Fermi levels [76].

Several groups have now prepared InAs/GaAs QD IB solar cells and the highest efficiency so far achieved is of 18.3% [77] although the GaAs cell used for reference has higher efficiency. The problems associated with present IB solar cells are discussed in detail in reference [78].

Up-converters, located behind a bifacial (active for light on both its sides) ordinary cell can receive two photons with energy below the cell bandgap and emit a higher energy photon that might subsequently be absorbed by the cell [79]. The efficiency of present up-converters is small in spite of the considerable progress already achieved [80].

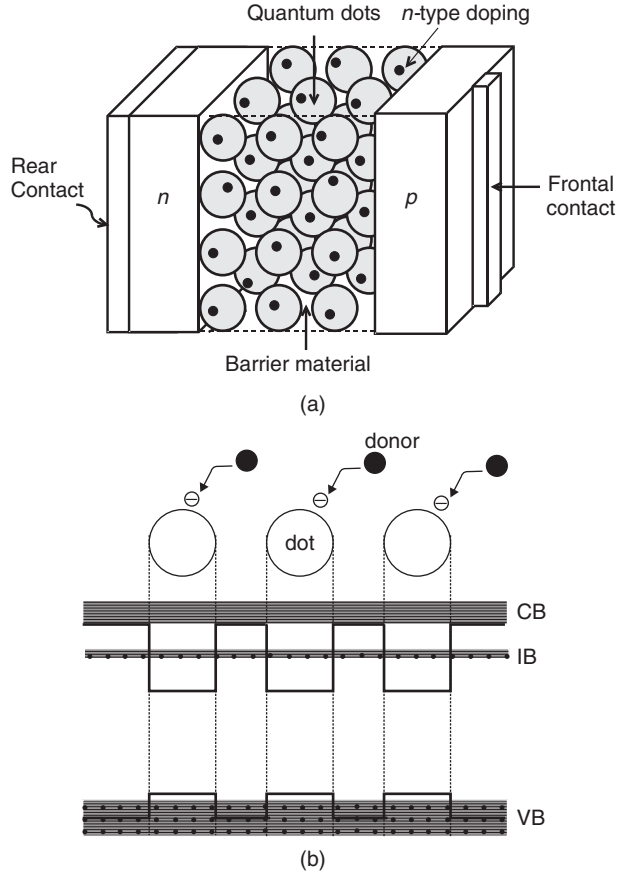


Figure 4.14 IB formation in QD arrays. (a) schematics; (b) IB region bands. Donor impurities in the barrier material partially fill the IB

Up-converter material might certainly be coupled to thin film solar cells and also to TiO_2 crystals [80], injecting the electrons there to form an IB dye-sensitised solar cell [81].

4.5.5 Hot Electron Solar Cells

Hot electrons, or hot carriers in general, are electrons within a certain band (or conduction electrons in the CB and holes in the VB) which are not in equilibrium with the surrounding lattice. In Section 4.2.6, the interactions of electrons among themselves and with phonons were described. Of course, they also interact with photons. In the hot carrier solar cell [82, 83], the electrons receive energy from the solar photons, but the inelastic interaction with phonons is small; absent in the ideal case. However, elastic interactions with the phonons are very frequent and make the electron thermodynamic functions direction independent. They are however considered energy dependent.

As in the case of photons, for every energy they are characterized by a temperature and a chemical potential; one of them can be chosen arbitrarily thus determining the other. For the moment we shall assume that the temperature is that of the lattice.

The interaction between electrons, which for the moment we assume dominates, can be described by reversible reactions such as $e_1^- + e_2^- \leftrightarrow e_3^- + e_4^-$. In this equation transference of energy between electrons is produced by couples so that $\hat{\varepsilon}_1 + \hat{\varepsilon}_2 = \hat{\varepsilon}_3 + \hat{\varepsilon}_4$ where $\hat{\varepsilon}$ is electron (not photon) energy. In addition, at constant temperature, the condition of reversibility also establishes that the electrochemical potentials of the electrons are related by $\varepsilon_F(\hat{\varepsilon}_1) + \varepsilon_F(\hat{\varepsilon}_2) = \varepsilon_F(\hat{\varepsilon}_3) + \varepsilon_F(\hat{\varepsilon}_4)$. Therefore, the electrochemical potential of the electrons is a linear function of the energy in the form

$$\varepsilon_F(\hat{\varepsilon}) = \beta\hat{\varepsilon} + \varepsilon_{F0} \tag{4.63}$$

Since the Fermi function for electrons is written as

$$\frac{1}{\exp\left[\dfrac{\hat{\varepsilon} - \beta\hat{\varepsilon} - \varepsilon_{F0}}{kT_a}\right] + 1} = \frac{1}{\exp\left[\dfrac{\hat{\varepsilon} - \varepsilon_{F0}/(1-\beta)}{kT_a/(1-\beta)}\right] + 1} \tag{4.64}$$

the electron distribution can be regarded equivalently as a distribution at the lattice temperature T_a, but characterised by a varying electrochemical potential given by Equation (4.64) or as a hot carrier distribution with a constant electrochemical potential given by $\mu_{hc} = \varepsilon_{F0}/(1-\beta)$ and a *hot carrier temperature* $T_{hc} = T_a/(1-\beta)$.

If besides the direction-randomising elastic interaction with phonons an inelastic additional interaction is introduced through the reaction $e_1^- + \text{phonon} \leftrightarrow e_2^-$, a condition is set to the chemical potentials of the species interaction. The additional anharmonic interaction of the very abundant phonons among themselves forces them to have a chemical potential of zero because they are freely created and annihilated, their number is not fixed, and they certainly aren't affected by the presence of the few electrons. Therefore, $\varepsilon_F(\hat{\varepsilon}_1) = \varepsilon_F(\hat{\varepsilon}_2)$. However, $\hat{\varepsilon}_1 \neq \hat{\varepsilon}_2$ but $\hat{\varepsilon}_2 = \hat{\varepsilon}_1 + \varepsilon$, where ε is the phonon energy and, $\varepsilon_F(\hat{\varepsilon}_1) = \varepsilon_F(\hat{\varepsilon}_2)$ is only fulfilled if $\beta = 0$ and consequently all the electrons have the same electrochemical potential ε_{F0} and are at the lattice temperature, discarding any possibility of hot carriers. But inelastic electron–phonon interaction is considered totally suppressed in the ideal hot carrier cell to be analysed later.

If the interaction with photons is considered now, $e_1^- + \text{photon} \leftrightarrow e_2^-$, the equilibrium of the reaction is represented by $\varepsilon_F(\hat{\varepsilon}_1) + \mu_{\text{ph}} = \varepsilon_F(\hat{\varepsilon}_2)$ where μ_{ph} is the photon chemical potential (it can be n-zero because they do not interact among themselves). Taking into account Equation (4.64), it is obtained that $\mu_{\text{ph}} = \beta(\hat{\varepsilon}_2 - \hat{\varepsilon}_1) = \beta\varepsilon$ where ε is the energy of the photon involved. The result is that we have an energy-dependent photon chemical potential μ_{ph}. The Bose function that describes the occupation probability of the photon energy level ε becomes

$$\frac{1}{\exp\left[\dfrac{\varepsilon - \beta\varepsilon}{kT_a}\right] - 1} = \frac{1}{\exp\left[\dfrac{\varepsilon}{kT_a/(1-\beta)}\right] - 1} \tag{4.65}$$

which can be seen as free radiation (zero chemical potential) at the hot carrier temperature.

It has been pointed out [84, 85] that to manufacture a solar cell, we should be careful with the contacts because they might not allow for the maintenance of hot carriers inside the absorbing material. Special contacts must be able to 'cool' the electrons from the 'hot' temperature T_{hc} to the contact temperature T_a reversibly by changing their electrochemical potential from ε_{F0} to the electron electrochemical potential at the contacts (Fermi level at the metals) ε_{F+} and ε_{F-}. In reference [83], these special contacts are devised as selective membranes (Figure 4.15) that

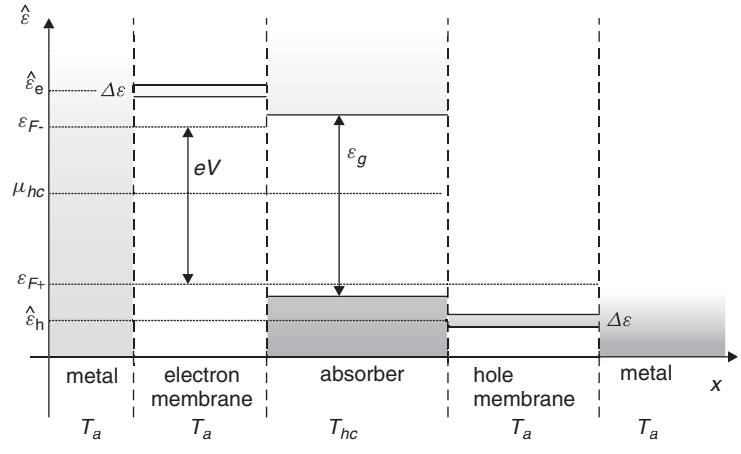


Figure 4.15 Band structure of a hot electron solar cell showing contacting scheme by means of selective membranes. (Reprinted with permission from Reference [83])

only allow electrons with energy centred around ε_e (left contact) and ε_h (right contact) to pass through. The reversible change in temperature and electrochemical potential of the electrons at the membranes is obtained by setting

$$\begin{aligned}\frac{\hat{\varepsilon}_e - \mu_{hc}}{kT_{hc}} = \frac{\hat{\varepsilon}_e - \varepsilon_{F-}}{kT_a} \Leftrightarrow \varepsilon_{F-} = \hat{\varepsilon}_e\left(1 - \frac{T_a}{T_{hc}}\right) + \mu_{hc}\frac{T_a}{T_{hc}} \\ \frac{\hat{\varepsilon}_h - \mu_{hc}}{kT_{hc}} = \frac{\hat{\varepsilon}_h - \varepsilon_{F+}}{kT_a} \Leftrightarrow \varepsilon_{F+} = \hat{\varepsilon}_h\left(1 - \frac{T_a}{T_{hc}}\right) + \mu_{hc}\frac{T_a}{T_{hc}}\end{aligned} \tag{4.66}$$

where $T_{hc} = T_a/(1-\beta)$. The cell voltage will therefore be given by

$$qV = \varepsilon_{F-} - \varepsilon_{F+} = (\hat{\varepsilon}_e - \hat{\varepsilon}_h)\left(1 - \frac{T_a}{T_{hc}}\right) \tag{4.67}$$

The current extracted from the cell is determined by the rate at which electron–hole pairs of energy $\hat{\varepsilon}_e - \hat{\varepsilon}_h$ can be withdrawn from the cell. Since no energy is lost as heat, the energy balance equation (first law) leads to

$$I(\hat{\varepsilon}_e - \hat{\varepsilon}_h)/q = \dot{E}(T_s, 0, \varepsilon_g, \infty, H_s) - \dot{E}(T_{hc}, 0, \varepsilon_g, \infty, H_r) \tag{4.68}$$

and the power extracted from the cell can be finally computed as

$$\dot{W} = IV = [\dot{E}(T_s, 0, \varepsilon_g, \infty, H_s) - \dot{E}(T_{hc}, 0, \varepsilon_g, \infty, H_r)]\left(1 - \frac{T_a}{T_{hc}}\right) \tag{4.69}$$

which is independent of the carrier-extracting energies of the contacts. In other words, a large separation of the extracting energies leads to high voltage and low current and *vice versa*. Note that T_{hc} is a parameter for Equation (4.68) and for Equation (4.69). By elimination, we obtain the $W(V)$ and from it the derived $I-V$ curve. The cell efficiency is obtained from the maximum of $W(V)$ or $W(T_{hc})$.

For $\varepsilon_g \to 0$ the limit efficiency becomes

$$\eta = \left(1 - \frac{T_{hc}^4}{T_s^4}\right)\left(1 - \frac{T_a}{T_{hc}}\right) \tag{4.70}$$

just as in the TPV converters, leading to a limiting efficiency of 85.4%.

Note that we have only introduced the band gap (ε_g) in Equation (4.68) to limit the photon absorption and emission. If a bandgap exists the transference among bands should be modelled and a band particle balance would represent an additional link. These points have been addressed in reference [85].

The monoenergetic membrane for electron and hole transfer to the contacting metals might be an insulator with an impurity band [86], but the nature of the phonon-insulated absorber is yet a subject of speculation. Inelastic electron photon coupling has been calculated [87] in a set of nonpolar and polar semiconductors and the conclusion is that Si seems to be the best choice, but perhaps only marginally attractive. It has been suggested [88] that, since thermal coupling is produced with optical phonons, some materials like InN with large phononic bandgap [89, 90] might have the required poor thermal coupling between optical and acoustic phonons. This behaviour has been also related to the phonon bottleneck in low-dimensional structures. This concept is attractive and, in addition, most probably Si, Ge and other classic materials have too strong a thermal coupling so that novel solutions have to be sought. In any case the decoupling of the optical and acoustical phonon temperatures relies in the low thermal conductivity attributed to the optical phonons. This property has to be examined because in materials with phononic bandgap the dispersion curve of the optical phonons can be considered horizontal only in a first-order study and optical phonons have small, but not zero group velocity so that the thermal conductivity may not be negligible.

4.6 CONCLUSIONS

In this chapter we have provided a thermodynamic basis that allows evaluation of the thermodynamic consistency of classical and newly proposed solar cells. Also, we have assessed the efficiency limit of several PV concepts.

As deduced by Shockley and Queisser, the upper limit reachable for single-junction solar cells illuminated by a black body source of photons at 6000 K is 40.7%, assuming the cell temperature at 300 K. This value is rather low if we take into account that the Carnot limiting efficiency for a reversible engine operating between hot and cold heat reservoirs at 6000 K and 300 K, respectively, is 95%. High-efficiency devices that can ideally surpass the Shockley–Queisser efficiency have been called *third-generation PV converters*. Thus, the following question arises: could we invent a solar converter that exhibits this Carnot efficiency?

The answer is negative and the reason for it lies in the definition of efficiency. In the definition of the Carnot efficiency, the denominator is the power *consumed*, that is, the power *arriving* at the converter less the power *leaving* the converter owing to the radiation that is emitted. In the conventional definition of the efficiency for solar converters, the term entering in the denominator is the power *arriving* at the converter, and because it is higher than the *consumed* power this leads to a lower efficiency. With this definition, the higher achievable efficiency is the Landsberg efficiency of 93.33%. However, even this, cannot be reached with any known ideal solar converter.

A very high efficiency of 85.4% can ideally be reached with several devices such as the TPV converter, constituted by an ideal solar cell and a black body absorber and the TPH converter, conceived with an ideal solar cell and a LED that also plays the role of an absorber, or even a hot

carrier solar cell. To reach this efficiency they all must emit radiation with zero chemical potential (free radiation) at 2544 K. This efficiency is also the limit for solar thermal devices.

However, this is not the highest efficiency that can be reached in solar converters. Efficiencies of up to 86.8% can be achieved using an array of solar cells of different bandgap, either series-connected or independently connected.

Except for the TPV and the TPH cases, all solar cells operate at ambient temperature. This is a highly desirable feature. The TPH concept, which allows for higher efficiencies at lower temperatures than the TPV concept, may bring some advances. However, for such devices to be practical requires an almost ideal external quantum efficiency of the LED and the ability of working at high temperatures, both requirements being very difficult to achieve.

MJ solar cells have already reached 41% efficiency and most probably 45% will be possible at concentration of about 1000 suns. This concentration is necessary to render these concepts cost effective [91, 92].

QD structures have proven to be appropriate for MEG solar cells. Although subject to some controversy, more than one electron–hole pairs per photon have been found experimentally in nanostructured material and beyond doubt in bulk material. The extraction of the current from this cell is yet unsolved although some ideas have been advanced.

Experimental work on IB solar cells – whose upper limit efficiency is 63.2% – is now strong and some promising results have started to appear. Several IB bulk materials have been found and solar cells with rather promising results have been fabricated with QDs. It is conceivable that IBSCs may once substitute for triple-junction devices with perhaps less complexity. They may also be combined in tandem with more ordinary or IB cells.

Finally, calculations prove that Si and Ge might be adequate materials for hot carrier solar cells. Low-dimensionality semiconductors have also been proposed as possible candidate materials. Concerning the selective contacts preliminary work has been done on selective contacts using impurity bands in high-bandgap semiconductors.

In summary, new solar cell principles have been presented that can improve the efficiency of solar cells in general. For instance the IB principle can be used to improve thin film or even dye-sensitised solar cells. On the other hand they may be combined in MJ stacks to reach higher efficiencies with lower complexities. In our opinion, very high efficiencies, certainly above 50%, will be achieved in actual PV devices. The use of very-high-concentration elements, in the range of 1000 or more, will be necessary to make these very expensive cells cost-effective. Wide acceptance angle concentrators with this concentration factor have already been developed [93] that seem very suitable for mounting in low-cost tracking structures [93].

REFERENCES

1. Lofersky J, *Postepy-Fizkyi* **26**, 535–560 (1975).
2. Shockley W, Queisser H, *J. Appl. Phys*. **32**, 510–519 (1961).
3. Green M, *Prog. Photovolt*. **9**, 123–135 (2001).
4. Jackson ED, *Trans. Conf. on the Use of Solar Energy*, Tucson, 1955, University of Arizona Press, Tucson, vol. 5, pp 122–126, 1958.
5. Callen H, *Thermodynamics*, John Wiley & Sons, Inc., New York (1981).
6. Landau L, Lifchitz E, *Physique Statistique*, Chap. I §24, Mir, Moscou (1967).
7. Badescu R, *Equilibrium and Nonequilibrium Statistical Mechanics*, John Wiley & Sons, Inc., New York (1975).

8. Kondepudi D, Prigogine I, *Modern Thermodynamics*, John Wiley & Sons, Ltd, Chichester (1999).
9. Luque A, Martí A, Cuadra L, *Physica E*. **14**, 107–114, (2002).
10. Landau L, Lifchitz E, *Physique Statistique*, Chap. V §52, Mir, Moscou (1967).
11. Welford W, Winston R, *The Optics of Nonimaging Concentrators*, Appendix I, Academic Press, New York, NY (1978).
12. Landau L, Lifchitz E, *Mécanique*, Chap. VII §46 La Paix, Moscou, (prior to 1965).
13. Welford W, Winston R, *The Optics of Non-imaging Concentrators*, Chapter 2, §2.7. Academic Press, New York (1978).
14. Luque A., *Solar Cells and Optics for Photovoltaic Concentration*. Chap. 13, §13.1. Adam Hilger, Bristol (1989).
15. Nozik A, *Annu. Rev. Phys. Chem*. **52**, 193–231 (2001).
16. Martí A, Balenzategui J, Reyna R, *J. Appl. Phys*. **82**, 4067–4075 (1997).
17. Araújo G, Martí A, *Sol. Energy Mater. Sol. Cells*, **31**, 213–240 (1994).
18. Luque A, Martí A, *Phys. Rev. Lett*. **78**, 5014–5017 (1997).
19. Hulstrom R, Bird R, Riordan C, *Sol. Cells* **15**, 365–391 (1985).
20. De Vos A, *Endoreversible Thermodynamics of Solar Energy Conversion*, Chap. 2 §2.1, Oxford University, Oxford (1992).
21. Miñano J, *Optical Confinement in Photovoltaics*, in Luque A, Araújo G (eds), *Physical Limitations to Photovoltaic Energy Conversion*, pp 50–83, Adam Hilger, Bristol (1990).
22. Shockley W, *Bell Syst. Tech*. **28**, 435–489 (1949).
23. Würfel P, *Physica E* **14**, 18–26 (2002).
24. Würfel P, *Physics of Solar Cells: From Basic Principles to Advanced Concepts*. John Wiley & Sons, Ltd, Chichester (2009).
25. Sinton R, Kwark Y, Gan J, Swanson R, *IEEE Electron. Dev. Lett*. **EDL7**, 567–569 (1986).
26. M. A. Green, K. Emery, Y. Hishikawa, and W. Warta, Prog. *Photovolt: Res. Appl*. **17**, 85–94 (2009).
27. Araújo G, *Limits to Efficiency of Single and Multiple Bandgap Solar Cells*, in Luque A, Araújo G (eds), *Physical Limitations to Photovoltaic Energy Conversion*, pp 119–133, Adam Hilger, Bristol (1990).
28. Araújo G, Martí A, *IEEE Trans. Elec. Dev*. **37**, 1402–1405 (1998).
29. Gale R, King B, Fan J, *Proc. 19th IEEE PSC*, 293–295, IEEE, New York (1987).
30. Tobin S, Vernon S, Sanfacon M, Mastrovito A, *Proc. 22nd IEEE PSC*, 147–152, IEEE, New York (1991).
31. Miñano J, *J. Opt. Soc. Am. A* **3**, 1345–1353 (1986).
32. Parrot J, in Luque A, Araújo G, (eds), *Physical Limitations to Photovoltaic Energy Conversion*, Adam Hilger, Bristol (1990).
33. Martí A, Araújo G, *Sol. Energy Mater. Sol. Cells* **43**, 203–222 (1996).
34. Luque A, Martí A, *Phys. Rev. B* **55**, 6994–6999 (1997).
35. Landsberg P, Tonge G, *J. Appl. Phys*. **51**, R1–20 (1980).
36. De Vos A, Pauwels H, *Appl. Phys*. **25**, 119–125 (1981).
37. Luque A, Martí A, *Sol. Energy Mater. Sol. Cells* **58**, 147–165 (1999).
38. Brown A, Green M, *Prog. Photovolt.: Res. Appl*. **10**, 299–307 (2002).
39. Tobías I, Luque A, *Prog. Photovolt.: Res. Appl*. **10**, 323–329 (2002).
40. Luque A, *Coupling Light to Solar Cells*, in Prince M (ed.), *Advances in Solar Energy*, Vol. 8, ASES, Boulder, CO (1993).
41. Castañs M, *Revista Geofísica* **35**, 227–239 (1976).
42. Tobias I, Luque A, *IEEE Trans Electron Dev*. **49**, 2024–2030 (2002).
43. Werner J, Kolodinski S, Queisser H, *Phys. Rev. Lett*. **72**, 3851–3854 (1994).
44. Werner J, Brendel R, Queisser H, *Appl. Phys. Lett*. **67**, 1028–3014 (1995).
45. Kolodinski S, Werner J, Queisser H, *Appl. Phys. Lett*. **63**, 2405–2407 (1993).

46. Kolodinski S, Werner J, Queisser H, *Sol. Energy Mater. Sol. Cells* **33**, 275–285 (1994).
47. Klimov V I, *Applied Physics Letters* **89**, 123118 (2006).
48. Schaller R D, Sykora M, Pietryga JM, Klimov V I, *Nano Letters* **6**, 424–429 (2006).
49. Beard M C, Knutsen K P, Yu P R, Luther J M, Song Q, Metzger W K, Ellingson R J, Nozik A J, *Nano Letters* **7**, 2506–2512 (2007).
50. Hanna M C, Nozik A J, *J. Appl. Phys*. **100**, 074510–8 (2006).
51. Nair G, Geyer S M, Chang L Y, and Bawendi M G, Carrier multiplication yields in PbS and PbSe nanocrystals measured by transient photoluminescence, *Physical Review B*, vol. 78, p. 10, (2008).
52. McGuire J A, Joo J, Pietryga J M, Schaller R D, and Klimov V I, New Aspects of Carrier Multiplication in Semiconductor Nanocrystals, *Accounts of Chemical Research.* **41**, 1810–1819, (2008).
53. Luther J M, Law M, Beard M C, Song Q, Reese M O, Ellingson R J, and Nozik A J, Schottky Solar Cells Based on Colloidal Nanocrystal Films, *Nano Letters*, **8**, pp. 3488–3492, (2008).
54. Pijpers J J H, Ulbricht R, Tielrooij K J, Osherov A, Golan Y, Delerue C, Allan G, and Bonn M, *Nature Physics*, **5**, 811–814, (2009).
55. Trupke T, Green M, Würfel P, *J. of Appl. Phys*. **92**, 4117–4122 (2002).
56. Luque A, Martí A, *Phys. Rev. Lett*. **78**, 5014–5017 (1997).
57. Luque A, Martí A, *Prog. Photovolt.: Res. Appl*. **9**, 73–86 (2001).
58. Cuadra L, Marti A, and Luque A, *IEEE Trans. Electron Dev*. **51**, 1002–1007 (2004).
59. Luque A, Marti A, Mendes M J, and Tobias I, *Journal of Applied Physics*, **104**, 113118, (2008).
60. Tobias I, Luque A, and Marti A, *Journal of Applied Physics*, **104**, 034502, (2008).
61. Green M, *Prog. Photovolt.: Res. Appl*. **9**, 137–144 (2001).
62. Luque A, Martí A, Antolín E, Tablero C, *Physica B* **382**, 320–327 (2006).
63. Lang D V, Henry C H, *Phys. Rev. Lett*. **35**, 1525–1528 (1975).
64. Antolin E, Marti A, Olea J, Pastor D, Gonzalez-Diaz G, Martil I, and Luque A, *Appl. Phys. Lett*. **94**, 042115 (2009).
65. Wahnón P, Tablero C, *Phys. Rev. B* **65**, 155115 (2002).
66. Palacios P, Aguilera I, Sanchez K, Conesa J C, Wahnón P, *Phys. Rev. Lett*. **101**, 046403 (2008).
67. Lucena R, Aguilera I, 72alácios P, Wahnon P, Conesa J C, *Chem. Mat*. **20**, 5125 (2008).
68. Yu K M, Walukiewicz W, Wu J, Shan W, Beeman J W, Scarpulla M A, Dubon O D, Becla P, *Phys. Rev. Lett*. **91**, 246403 (2003).
69. Yu K M, *et al*., *Appl. Phys. Lett*. **88**, 092110 (2006).
70. Wang W, Lin A S, and Phillips J D, *Applied Physics Letters*, **95**, 011103, (2009).
71. G. Gonzalez-Diaz, J. Olea, I. Martil, D. Pastor, A. Marti, E. Antolin, A. Luque, *Solar Energy Materials and Solar Cells*, **93**, 1668–1673, (2009).
72. Martí A, Cuadra L, Luque A, *Proc. 28th IEEE Photovoltaic Specialist Conf*., pp 940–943, IEEE, New York (2000).
73. Martí A, Cuadra L, Luque A, *IEEE Trans. Elec. Dev*. **48**, 2394–2399 (2001).
74. Luque A., Martí A., Stanley C., López N., Cuadra L., Zhou D., Mc-Kee A., *J. Appl. Phys*. **96**, 903–909 (2004).
75. Marti A, Antolin E, Stanley C R, Farmer C D, Lopez N, Diaz P, Canovas E, Linares P G, Luque A, *Physical Review Letters*, **97**, 247701–4, (2006).
76. Luque A, Marti A, Lopez N, Antolin E, Canovas E, Stanley C, Farmer C, Caballero L J, Cuadra L, Balenzategui J L, *Applied Physics Letters*, **87**, 083505, (2005).
77. S. A. Blokhin, A. V. Sakharov, A. M. Nadtochy, A. S. Pauysov, M. V. Maximov, N. N. Ledentsov, A. R. Kovsh, S. S. Mikhrin, V. M. Lantratov, S. A. Mintairov, N. A. Kaluzhniy, M. Z. Shvarts, AlGaAs/GaAs Photovoltaic Cells with an Array of InGaAs QDs, *Semiconductors*, **43**, 514–518, (2009).
78. Luque A, Marti A, *Advanced Materials*, DOI 10.1002/adma.200902388, (2009).
79. Trupke T, Green M, Würfel P, *J. Appl. Phys*. **52**, 1668–1674 (2002).

80. Ekins-Daukes N J, Schmidt T W, *Appl. Phys. Lett*. **93**, 063507 (2008).
81. Grätzel M, *J. Photochem. Photobiol. C: Photochem. Reviews* **4**, 145–153 (2003).
82. Ross R, Nozik A, *J. Appl. Phys*. **53**, 3813–3818 (1982).
83. Wurfel P, *Sol. Energy Mater. Sol. Cells* **46**, 43–52 (1997).
84. O'Dwyer M F, Humphrey T E, Lewis R A, and Zhang C, *Microelectronics J*. **39**, 656–659 (2008).
85. Wurfel P, Brown A S, Humphrey T E, and Green M A, *Progr. Photov*. **13**, 277–285 (2005).
86. Conibeer G, *et al*., *Solar* **511**, 654–662 (2006).
87. Luque A, Martí, *Solar Energy Materials and Solar Cells*, **94**, 287–296 (2010).
88. Conibeer G J, Guillemoles J F, König D *et al*., *21st European Photov. Sol. En. Conf*. WIP, Dresden, (2006), p. 90.
89. Davydov V Y, Emtsev V V, Goncharuk I N, *et al*. *Appl. Phys. Lett*. **75**, 3297–3299 (1999).
90. Bungaro C, Rapcewicz K, and Bernholc J, *Phys. Rev. B* **61**, 6720–6725 (2000).
91. Yamaguchi M, Luque A, *IEEE Trans. Electron. Dev*. **46**, 2139–2144 (1999).
92. Algora C, Rey-Stolle I, García I, Galiana B, Baudrit M, and González JR, ASME J. Sol. En. *Eng*., **129**, 336 (2007).
93. Luque A, Andreev V M (eds), *Concentrator Photovoltaics*, Springer, Berlin, (2007).
94. Luque-Heredia I, Martin C, Mananes M T, Moreno J M, Auger J, Bodin V, Alonso J, Diaz V, Sala G, in *Proc. 3rd World Conf. on Photovolt*., Vol. 1, p. 857 (2003).

5

Solar Grade Silicon Feedstock

Bruno Ceccaroli[1] and Otto Lohne[2]

[1]*Marche AS, Kristiansand, Norway,* [2]*Norwegian University of Science and Technology (NTNU), Trondheim, Norway*

5.1 INTRODUCTION

The Photovoltaic (*PV*) industry is still in its infancy and at the moment it is very difficult to predict which technical, economical and social patterns its deployment will follow before reaching maturity[1]. However, if photovoltaics are to become a major energy source in the future, it is appropriate to question which materials and which natural elements are critical to secure the long-term sustainability of this energy source. This is particularly valid for the semiconductor materials whose bandgap has to perform the efficient conversion of sunlight to electricity. The history of photovoltaics (since the 1950s) reveals an intense activity of research and development, embracing a broad range of disciplines and leading to a healthy multitude of innovations. Organic versus inorganic semiconductors, intrinsic versus extrinsic semiconductors, homojunctions versus heterojunctions and amorphous versus crystalline structures are a few dilemmas that new research achievements steadily bring to the scientific and industrial community[2]. It will take years, perhaps decades, before scientific and industrial companies are able to solve the challenges and answer the questions addressed above.

Up to now, the dominant semiconductor material converting light to electricity is elemental silicon. Two main classes of silicon must be distinguished: amorphous and crystalline. Crystalline cells are either single[3] or multicrystalline. Within each group of technology several variants may be distinguished[4]. The most recent market surveys[5] issued by commercial consultants, governmental organisations and agencies clearly confirm the dominance of silicon-based,

[1] In this chapter we will indifferently designate by *solar cells, photovoltaic(s)* and the abbreviation *PV* all scientific, industrial and commercial activities related to conversion of light to electricity.
[2] These aspects are covered in other chapters in this handbook.
[3] In this chapter we use indifferently the terms single crystalline and monocrystalline.
[4] The elaboration of the cells by different silicon-based technologies and their characteristics are described in the Chapter 7 of this handbook.
[5] This chapter was written in the last quarter of 2008 and first quarter of 2009. Proofreading was executed in September-October 2010 without inserting new comments although recent market development might have called for such.

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

and particularly crystalline silicon-based technologies. Analysing annual growth by technology and material in 2003, we could write in the first edition (Chapter 5) of the handbook that multicrystalline silicon took the lion's share of the growth. In the first half of the 1990s multicrystalline sawn wafers accounted for just half of the single-crystal shipments. In 1998 both technologies were equivalent in size. However, during the first decade of the new century single-crystal technology has regained a strong position being in 2007 at par with multicrystalline, whereas amorphous silicon cells have regressed from 12% of total cell production in 1999 to 5.2% in 2007. It is worth noting that non-silicon technologies have remarkably grown from 1% in 1999 to 5.2% in 2007, mainly due to the success of CdTe cells. To keep the numbers in the right perspective we must mention that in the same period the cell production has increased from 200 MW in 1999 to 4.3 GW in 2007, which means an average annual growth near 50%. There is no doubt that with 90% share crystalline silicon technologies must be credited this extraordinary growth achievement [1] (see also Section 5.5 and Table 5.7).

In the first edition we predicted that crystalline silicon is and will remain for at least a decade the workhorse of this growing market. Long-term visionary forecasts predict that by 2050, 30 000 TW h PV electricity will be generated annually worldwide. This will require an installed PV output capacity totalling approximately 15 million metric tons of solar grade silicon feedstock, assuming that silicon remains dominant and that cell efficiency and material yields have steadily improved (see also Section 5.5 for trends of specific consumption g/W of silicon). To build up such a capacity over fifty years will represent an annual production of 300 000 metric tons solar grade silicon feedstock. The annual present consumption of pure silicon for photovoltaics was then in year 2000 approximately a hundredth part of that (i.e. 4000 metric tons). In 2007 the consumption of silicon by the photovoltaic industry is in the range of 30 000 metric tons, whereas the prognoses for demand and supply in 2012 concur to 120–160 000 metric tons approaching the average value of 300 000 indicated in the first edition. Again to keep these numbers in a correct perspective it must be mentioned that in 2007 more than 2 million metric tons of metallurgical grade silicon were produced for all purposes, the solar cell application representing still no more than 2.5%. However, since 2003–2004 the demand in solar grade silicon has exceeded the offer. This new situation has created a situation of material shortage, which in turn caused sharp price increases along the value chain. Naturally and fortunately it also stimulated the search for numerous improvements to save silicon and inspired a plethora of initiatives to produce more silicon and to explore new routes to solar grade silicon. At the time of writing, the issue of material feasibility and availability remains a challenge in spite of the sudden slow-down in the world economy caused by the 2008 financial crisis. The present chapter is therefore dedicated to silicon, its extraction, purification and availability by current and future practice.

5.2 SILICON

Silicon (Si) is the second member of group IVA in the periodic system of elements. It never occurs free in nature, but in combination with oxygen, forming oxides and silicates. Most of the Earth's crust is made up of silica and miscellaneous silicates associated with aluminium, magnesium and other elements. Silicon constitutes about 26% of the Earth's crust and is the second most abundant element in weight, oxygen being the largest.

5.2.1 Physical Properties of Silicon Relevant to Photovoltaics

Silicon is a semiconductor with a bandgap E_g of 1.12 eV at 25 °C. At atmospheric pressure, silicon crystallises into a diamond cubic structure, which converts into a body-centred lattice when subjected to ca 15 GPa. Under some circumstances, slow-growing faces of silicon are {111} but in epitaxial films and polysilicon deposition {111} is the fastest growth direction. Vapour deposition

below 500 °C results in amorphous silicon. If reheated above this temperature, crystallisation will occur.

Unlike most compounds and elements, silicon contracts when melting or expands when solidifying.

Impurities incorporated in the silicon lattice during the crystal growth or during the post-treatment (diffusion, implantation etc.) ionise at low temperatures, thus providing either free electrons or holes. Impurities from group VA replace a Si atom in the atomic lattice to supply electrons and are called *n-dopants* or *donors*, whereas elements from group IIIA substitute for a Si atom to supply holes and are called *p-dopants* or *acceptors* (see Section 5.6.3). Phosphorus and boron represent these groups and are used in PV processing to control the semiconductor properties (doping levels) of silicon. In semiconductor sciences and technologies impurity concentrations are expressed in atoms of impurity per cubic centimetre of the host material (silicon). In silicon semiconductor devices, these vary from 10^{14} to 10^{20} atoms per cm^3 (for 5×10^{22} atoms/cm^3 of Si see Table 5.1) and can be directly measured by analytical instruments. Another way used by physicists to express the impurity concentration is the ratio of the atoms of impurity to the atoms of the host material (silicon) i.e. *parts per million of atoms ppm(a)* for one (1) atom of impurity in 1 million atoms of silicon and by extension *part per billion ppb(a), part per trillion ppt(a)* etc. Production of silicon is however the profession of chemists and metallurgists whose chemical analyses are expressed in the weight ratio of the impurity to the host material, i.e. 1% designing 1g of impurity in 100 g of silicon, 1 *ppm(w)* 1 g of impurity in 1 metric ton of silicon and by extension *1 ppb(w)* etc[6]. An indirect measure of impurity concentration is the *minority-carrier lifetime*. This is the time that elapses before a free electron in the lattice recombines with a hole. The transition metals, Fe, Cr, Ni, degrade the minority-carrier lifetime and the solar cell performances. High-purity silicon crystals with metal content less than 10 ppb(w) may have minority-carrier lifetime values as high as 10 000 µs. Semiconductor wafers with phosphorus and boron dopants have values from 50 to 300 µs. Solar cells require minority-carrier lifetime value of at least 25 µs.

The relatively high refractive index limits the optical applications of silicon. The absorption/transmission properties in the 0.4–1.5 µm spectral wavelength range are important in the performance of PV cells and photoconductive devices. In PV applications antireflective layers applied to silicon are commonly used.

Silicon, even when alloyed with small quantities of impurities, is brittle. Shaping silicon for PV applications requires sawing and grinding. Microelectronic applications require polishing. These mechanical operations are very similar to those applied to glasses. Various thermal and mechanical properties are reported in Table 5.1. For more details the reader is invited to consult the references [2–6] used by the authors in writing this chapter.

Table 5.1 Thermal and mechanical properties of silicon

Property	Value
Atomic weight	28.085
Atomic density (atoms/cm^3)	5.0×10^{22}
Melting point (°C)	1414
Boiling point (°C)	3270
Density (g/cm^3 at 25 °C)	2.329
Heat of fusion (kJ/g)	1.8
Heat of vaporisation at MP (kJ/g)	16
Volume of contraction on melting (%)	9.5

[6] In the present chapter we use the different expressions according to the context.

5.2.2 Chemical Properties Relevant to Photovoltaics

Silicon is stable in the tetravalent state and has a strong affinity for oxygen, forming stable oxides and silicates, the only natural occurrences known for silicon. Artificially isolated elemental silicon immediately oxidises, forming a thin protective film of silica of less than 100 Å, which prevents further oxidation. Oxygen plays an important role in silicon semiconductor devices, for instance, in manufacturing metal oxide semiconductor (MOS) transistors.

Silicon and carbon (group IVA) form a strong Si–C bond and stable products. Silicon carbide is artificially synthesised in several allomorphic structures, finding various applications in photovoltaics and electronics. Primary uses are the abrasive properties of SiC for wafering silicon crystals[7] and the emerging applications of SiC semiconductors. The strong Si–C bond is also the origin of the rich organosilicon chemistry, encompassing numerous polysiloxanes (commonly named *silicones*) and organosilanes in which organic radicals are attached to silicon atoms through a covalent Si–C bond.

The tetravalence and the similarity of silicon and carbon are illustrated in the ability of silicon to form bonds with itself, Si–Si, and to form polymers, for example, $-(SiH_2)_p-$, $-(SiF_2)_p-$, comparable to hydrocarbons and fluorocarbons, although the length of the chains remains modest in the case of the polysilanes.

Silicon forms hydrides; monosilane (SiH_4) is a key chemical compound for the production of amorphous silicon and the purification of silicon to semiconductor grade (see Section 5.4.2 and 5.4.3 later in this chapter).

The chemical reactivity of silicon with chlorine is also extremely important. Alkyl- and arylchlorosilanes are the necessary intermediates to build the polysiloxane chains (silicones). Trichlorosilane and tetrachlorosilane,[8] because they are volatile at low temperature and can be decomposed to elemental silicon at high temperature, are both the intermediates and by-products of the purification processes upgrading metallurgical grade silicon to semiconductor purity (see Section 5.4 later in this chapter). Other chlorosilanes (e.g. dichlorosilane SiH_2Cl_2) are also used in chemical vapour deposition applications. The halogen atom is easily substituted by a hydroxyl group, –OH, through hydrolysis. Such a hydroxyl group tends to react with other functional groups by exchanging the hydrogen atom. This is the basis of a rich surface chemistry.

Fluorosilanes or fluorosilicates are used by one company as a precursor of monosilane to further produce semiconductor or solar grade (see Section 5.4.3 later in this chapter). Silicon bromides and iodides are also considered as volatile intermediates which can be highly purified by distillation to produce solar grade silicon (see Tables 5.11 and 5.12).

Silicon and germanium (group IVA) are isomorphous and mutually soluble in all proportions.

Tin and lead, also elements of group IVA, do not react with silicon and are not miscible in silicon, which is mentioned as a remarkable curiosity.

For more details on the chemical properties the reader is invited to consult the references [5–9] used by the authors in writing this chapter.

5.2.3 Health, Safety and Environmental Factors

The surface of elemental silicon is oxidised and is relatively inert and is considered as nontoxic. Hazard risks with elemental silicon are high when silicon occurs as a fine powder in the presence of

[7] See Chapter 6.
[8] Indifferently also called *silicon tetrachloride* in this chapter.

an ignition source. Damaging and fatal explosions have been reported by the silicon industry. The raw material, from which silicon is made, quartz or quartzite, is one of the sources of silicosis. Most of the hazards are related when quartz/quartzite is quarried, exposure taking place during drilling, crushing, loading and bulk handling. Several protection methods must be applied to the quarries and to the metallurgical plants to prevent silicon dust explosion and silicosis. Volatile silanes such as monosilane and chlorosilanes are extremely reactive in the presence of oxygen, water or moisture. They are also classified as hazardous chemical substances whose handling requires special care. Saturated long-chain silanes, polysiloxanes as well as amorphous silica are known to be chemically inert and not toxic. Because of that they are widely used in pharmacy, the food industry and cosmetics.

The production of metallurgical silicon and electronic grade silicon has an environmental impact through energy consumption, associated with climatic and polluting gases, principally CO_2, NO_x and SO_2. However, it must be noticed that the corresponding nuisances and energy consumption involved in manufacturing and installing PV systems are 'paid back' by the same system in the form of emission-free 'green' electricity already one to two years (four to five years in the first edition of the handbook) after being taken into service of a guaranteed service duration of more than twenty years [10–12]. More quantified examples of environmental and energy 'payback' are reported from Europe, Japan and Australia and commented in [13].

5.2.4 History and Applications of Silicon

Since antiquity, silicon has been of great importance to mankind. However, the first applications were based on naturally occurring forms of silicon, for instance, flint (*silex*, *silicis* in Latin), a variety of quartz used from the Stone Age to the Neolithic Era to make tools, weapons and later potteries. Glass made of silicate dated back to 12 000 B.C. Elemental silicon was prepared for the first time in 1824 by Berzelius*, passing silicon tetrachloride over heated potassium. Silicon tetrachloride could be prepared by chlorinating silicate/silica. The first crystalline silicon was made accidentally in 1854 by Sainte-Claire Deville†, working on aluminium electrolysis. The first preparation of silicon/silicon rich alloys in an electric arc furnace was performed by Moissan‡ in 1895 and the industrial production was by Bozel and Rathenau§ independently from 1897 to 1898. Around the same time Acheson¶ also discovered silicon carbide accidentally while trying to make artificial diamond. Silicon alloys, particularly ferroalloys, have from the end of the nineteenth century played an important role in the production of steel. *Silicon metal* (silicon content higher than 96% according to definitions outlined by trade organisations) was not current until the Second World War.[9] Three major applications have since the Second World War greatly stimulated the production and purification

* Berzelius, Jöns, Jacob (1779–1848): Swedish physician and chemist. He is considered as one of the most important scientist who ever lived. He discovered among others *the law of constant proportions* which provided evidence of the *atomic theory* of John Dalton. Beside silicon he is credited with identifying selenium, thorium and cerium.

† Sainte-Claire Deville, Henri (1818–1881): French chemist, one of the most influent French scientists of his time, mainly retained by history for inventing the first industrial process to aluminum (1854).

‡ Moissan, Henri (1852–1907): French scientist, professor at la Sorbonne (Paris), most famous for isolating fluorine for which he was awarded the Nobel prize for chemistry (1906). In 1900 he invented an arc furnace capable to reach 3500 °C opening the road to the discovery of numerous elements and compounds including silicon metal and ferroalloys.

§ Rathenau, Walther (1867–1922): German chemist, industry leader and politician. President to the German energy group AEG, he organized war economics during World War I. Charismatic politician he was both beloved and hated. He became in 1922 the foreign affairs minister to the Weimar Republic and as such he signed the Rapallo treaty. Shortly after, he was assassinated by two nationalist activists.

¶ Acheson, Edward, Goodrich (1856–1931): US inventor, he worked first for Edison. He is most famous for inventing carborundum (silicon carbide SiC) and synthetic graphite as well as other abrasives and lubricants.

[9] We will indifferently designate in this chapter this commercial grade of silicon either as *metallurgical grade silicon* or *silicon metal*. The former expression referred to the first historical use of this product in the alloy

Table 5.2 Chemical characteristics of commercial silicon metal as used by aluminium and chemical industries. All data are in ppm(w)

Element	O	Fe	Al	Ca	C	Mg	Ti	Mn	V	B	P
Low	100	300	300	20	50	5	100	10	1	5	5
High	5000	25 000	5000	2000	1500	200	1000	300	300	70	100
Element	Cu	Cr	Ni	Zr	Mo						
Low	5	5	10	5	1						
High	100	150	100	300	10						

of silicon, that is, aluminium, silicones and solid-state electronics. Silicon carbide has also found a broad range of applications taking advantage of its hardness and chemical noble character. More recently SiC has found applications in electronics because of its excellent semiconductor properties, and has tended to become a strategic material for cutting silicon ingots and boules into thin wafers.

At the beginning of the new 2000 millennium, approximately one million metric tons per year (slightly more) of metallurgical grade silicon were produced and sold in the world market. This is a relatively small amount compared with the multimillion metric ton markets of crude iron, steel, aluminium or even ferroalloys. However, it is a relatively fast growing segment (8% per year over the 10 past years) by comparison with other mature products from the metallurgical industry. In 2007 the global output has grown to approximately two million tons. Industrial location of silicon metal production has been guided by the vicinity of rich and pure quartz deposits and/or the availability of abundant electrical power. Leading producing countries are China, the United States of America, Brazil, Norway and France. In spite of some recent mergers, the industry remains fragmented. One may find three dozens of companies outside and inside China producing and marketing silicon metal, most of the plants having an annual output of 20 000–60 000 metric tons (see Section 5.3 for production of silicon metal). The chemical characteristics of commercial metallurgical grade silicon are indicated in Table 5.2.

The silicon metal market is traditionally divided into two main subgroups, that is, the aluminium and the chemical segments, each consuming respectively 60 and 40% of the worldwide output. There are, however, some differences in the characteristics requested by each. Until recently (the turn of the millennium) semiconductor and solar applications were considered a part of the chemical segment. But since 2003 the demand for hyper-pure silicon has exploded, boosted by the demand from the solar market. Several traditional metallurgical grade silicon producers are now deploying considerable efforts at their plants to develop purification methods bypassing the need in a chemical purification through a volatile compound. These producers are putting more efforts into this trendy product development than in capacity expansion, whose profitability seems to them rather questionable.

5.2.4.1 Applications in aluminium

In the aluminium industry, silicon is added to molten aluminium in which it is dissolved. A simple eutectic composition occurs at 12.6% silicon in aluminium. This has important consequences for industrial applications in the aluminium industry. Silicon is used in order to improve the viscosity, the fluidity of liquid aluminium and the mechanical properties of commercial alloys. The iron, calcium and phosphorus content in silicon are particularly critical for such applications.

industry, particularly aluminium, which remains so far the largest consumer of metallurgical grade silicon. The latter expression does not refer to silicon's physical properties. Silicon is not a metal, but a metalloid. It rather refers to its historical production method by metallurgists and its metallic shiny appearance.

There are two important groups of aluminium alloys in which silicon is one of the main alloying elements.

5.2.4.1.1 Casting alloys

By adding silicon to the melt, the fluidity is improved. Aluminium alloys near the eutectic composition are therefore used in thin-walled castings. Typical concentrations are 7–12%. If a few tenths of a percent of magnesium is added, the alloys may be age-hardened and thereby nearly double their yield strength.

To counteract the formation of large needle-shaped silicon particles, the alloys are normally modified with sodium, strontium or phosphorus.

The alloys present good corrosion properties.

5.2.4.1.2 Wrought alloys

AlMgSi alloys are widely used as medium-strength structural alloys. Typical silicon content is 0.5–1.0%. These alloys have good hot-working properties and are age-hardening. These alloys are therefore well suited for extrusion of profiles, which by heating at 150–200 °C are given their final strength.

The alloys present good corrosion and welding properties. Typical markets are building and transport industries. The aluminium consumes 50–60% of the worldwide metallurgical grade silicon.

5.2.4.2 Applications in chemistry

5.2.4.2.1 Silicones

Since the discovery of the direct synthesis (see Reaction 5.1) of dimethyldichlorosilane, during the Second World War independently by Rochow and Müller, the silicones industry has developed to become a strong and growing chemical business consuming about 35–40% of worldwide silicon metal output or 700 000 tons in 2008 versus 400 000 metric tons in 2000.

$$Si(s) + 2CH_3Cl(g) \rightarrow (CH_3)_2SiCl_2 \text{ (Cu catalyst, 260 to 370 °C)} \quad (5.1)$$

$$(CH_3)_2SiCl_2 + 2H_2O \rightarrow (CH_3)_2Si(OH_2) + 2HCl \quad (5.2)$$

$$n(CH_3)Si(OH)_2 \rightarrow [(CH_3)_2SiO]_n + nH_2O \quad (5.3)$$

The direct synthesis (Reaction 5.1) is industrially performed in a fluidised bed reactor, requiring small particles or powder of silicon (20–300 μm). The reaction is exothermic and needs to be activated with copper catalysts as well as promoters Zn, Sn, P and others. Whereas Fe does not seem to play an important role, Ca and Al have been shown to take an active part in the overall reaction [7–9].

5.2.4.2.2 Synthetic silica

Varieties of synthetic silica such as pyrogenic silica (also called *fumed silica*) or silica ingots as feedstock to optical fibres are industrially prepared by burning silicon tetrachloride:

$$SiCl_4(g) + 2H_2(g) + O_2(g) \rightarrow SiO_2(s) + 4HCl(g) \quad (5.4)$$

Silicon tetrachloride may be prepared by chlorination of natural silica. However, industrially, silicon tetrachloride is produced by reacting chlorine with metallurgical grade silicon in a direct

synthesis, performed either in a fluidised bed or a fixed bed reactor:

$$Si(s) + 2Cl_2(g) \rightarrow SiCl_4(g) \quad (5.5)$$

$$Si(s) + 4HCl(g) \rightarrow SiCl_4(g) + 2H_2(g) \quad (5.6)$$

The present fumed silica market is about 120 000–150 000 metric tons, which in turn consumes around 60 000–75 000 metric tons of metallurgical silicon (in 2007–2008). It is noted that the main market for fumed silica is as additives in silicone rubbers, used to increase the mechanical strength and the elasticity of these elastomers.

5.2.4.2.3 Functional silanes

This generic term covers a broad range of products built on silane molecules, in which an atom of hydrogen or chlorine is substituted with an organic radical bearing a functional group, for example, amine, acid, ester, alcohol and so on.

$$X_{4-(n+p+q)}SiR_nR'_pR''_q$$

represents a general formula of functional silanes, in which R_n, R_p, R_q (are organic radicals and X is a halide, generally Cl or H. There exists a multitude of functional silanes. One of their major applications is as coupling agent between inorganic and organic compounds, for example, inorganic fillers (glass, silica, clays, etc.) in organic matrices (epoxy, polyester, etc.) Orthoethylsilicate or tetraethoxysilane, $Si(OC_2H_5)_4$, is an important chemical molecule for the glass, ceramic, foundry and paint industries.

The total consumption of silicon metal to functional silanes and orthosilicates may be estimated at 10 000–20 000 metric tons per year (in 2007–2008).

5.2.4.3 Semiconductor and photovoltaic silicon

Silicon is by far the most important and popular semiconductor material since the emergence of solid-state electronics in the late 1950s and early 1960s. Ultra-pure silicon (commercially called *polysilicon*[10]) with adequate semiconductor properties is industrially prepared through the distillation and the thermal decomposition or chemical reduction of volatile silicon compounds, for example, trichlorosilane, $SiHCl_3$, and monosilane, SiH_4. These operations are performed in large chemical plants, which for synergy reasons are sometimes incorporated in plants producing other silicon-based compounds, such as those described in Section 5.2.4.2. Although the ultimate application in the case of polysilicon is in the semiconductor industry, this particular process is, from a silicon raw material perspective, counted among the chemical applications of silicon metal.

The current production of polysilicon in 2000 was approximately 20 000 metric tons, whereas the installed worldwide capacity was estimated around 25 000 metric tons. Except for downgraded material (a couple of thousands tons), all polysilicon was then used exclusively in semiconductor applications. Compared with the other applications of silicon (aluminium: Section 5.2.4.1, silicones and silica: Section 5.2.4.2), the use of polysilicon in terms of volume remains rather modest. In 2007 it was estimated at 40 000 metric tons and in 2008 at around 55 000 tons. However, it is a high-value product. For example, the silicon value is multiplied by a factor of 30–50 through upgrading metallurgical grade silicon to polysilicon. It is also a fast growing application of silicon with an annual growth rate of 5–7% for semiconductor and 30–50% for the solar market. Boosted by the growth of the photovoltaic industry, the demand for polysilicon has since 2003 exceeded the supply when solar cells could no longer be produced only out of rejects from semiconductor

[10] This term has background in the semiconductor industry, referring to its polycrystalline structure as opposed to the single crystals made from polysilicon through the Czochralski and float zone techniques.

materials. Producers and consumers of silicon for solar applications were late to understand the necessity to form an alliance and secure expansion capacity of polysilicon. This resulted in a shortage of material, which was detrimental to all parties, including the semiconductor industry. After 2003 significant polysilicon expansion projects have been launched, but it takes a minimum of three years to build a green field plant. As pointed out earlier in this chapter the photovoltaic market has grown dramatically (30–50% from year to year). The solar sector has now surpassed the semiconductor sector for the consumption of silicon. The vast majority of the expansion projects are driven by the demand from the solar sector, leaving in the shadow the specific needs of the electronic industry. Although new types of solar grade silicon emerge, as for instance solar grade silicon purified through metallurgical processes, the dominant feedstock to solar cells is currently in 2008 and for the years to come virgin polysilicon made through traditional processes developed for the semiconductor industry.[11] Therefore, a more detailed description of these processes will be given in Sections 5.3, 5.4 and 5.7 later in the present chapter.

5.2.4.4 Other applications

There are a few other applications of silicon in various fields such as explosives (silicon powder), refractories and advanced ceramics (silicon nitride and carbide). These applications presently do not account for more than 1% of the worldwide silicon metal output.

Because of their anticipated excellent mechanical and chemical resistant properties, alloys rich in silicon have a bright future. They may be prepared by powder metallurgy, mixing and sintering silicon powder with metallic powders (e.g. Cu, Al, Ti, Co, V etc.) [4].

The present chapter does not review the applications of silicon, such as glasses, refractories, ceramics and ferroalloys, in which silicon is usually added to the production process as natural silicate, quartz, quartzite or other silicon alloys.

5.3 PRODUCTION OF SILICON METAL/METALLURGICAL GRADE SILICON

5.3.1 The Carbothermic Reduction of Silica

Metallurgical grade silicon or silicon metal, with a minimum of 96% and typical purity of 98.5% Si, is produced in submerged electric arc furnaces (for purity, see Table 5.2). In principle, this process is much the same as it was at the beginning of the twentieth century, when it was first developed for ferrosilicon and other alloys. However, practical execution has greatly improved, with larger furnaces, more efficient material handling and improved control of the operations. This has led to a continuous decrease of the specific energy consumption concomitant to higher degrees of raw material utilisation.

The furnace consists essentially of a crucible filled with quartz and carbon materials. Silicon is freed by the carbothermic reduction of silica according to the overall reaction:

$$SiO_2(s) + 2C(s) = Si(l) + 2CO(g) \tag{5.7}$$

[11] We designate by *solar grade silicon* a silicon material from which commercial solar cells can be produced. We will indifferently use the terms *solar grade, PV, photovoltaic grade silicon*. We designate indifferently by *electronic* or *semiconductor grade silicon* a purer silicon material, from which solid state electronic devices can be manufactured. The term *feedstock* or *silicon feedstock* has been historically used in the industry to designate the silicon raw material. This emphasises the dominance of silicon. In this chapter we will use the term *feedstock* in this sense.

Contrary to what is often claimed in popular articles or reviews, silica sand is currently not used for this purpose. Lumpy quartz (e.g. 10–100 mm) with appropriate purity and thermal resistance is preferred. Carbon raw material generally consists of metallurgical grade coal as well as woodchips and/or charcoal and coke. The metallurgical coal is co-produced with coal used for crude steel production. As a rule this coal needs to be washed in order to remove most of the ash containing unsuitable impurities. Raw materials, both quartz and carbon, are selected in order to achieve high product quality (silicon and silica fumes), to maximise furnace performances and to minimise the environmental damages (i.e. SO_2 and NO_x emissions). The raw material reactivity and the consistency of the mix of raw materials in the charge, for instance its porosity, are extremely important factors in achieving good furnace performance in terms of high material yield, lower power consumption and good product quality.

The raw material mix or charge is heated by means of an intense electric arc sustained between the tip of three submerged electrodes and the electrical ground of the furnace. Although important exceptions exist, the current practice is to run this process in a three-phase current, open and rotating furnace at a working electrical load normally between 10 and 30 MW, depending on the size of the furnace. The tendency is to increase the furnace size and the electrical load in order to achieve higher output and productivity (recent furnace constructions allow load as high as 45 MW. They remain however rare exceptions).

Electrodes are also made of carbon. Originally, expensive graphite electrodes were used. They were displaced by pre-baked electrodes, which in turn tend to be replaced by more cost-efficient self-baking electrodes. The electrode technology is an important aspect to the present development of this industry: half a dozen electrode types ranging from pre-baked to self-baking electrodes of Søderberg type are currently used or are in the process of development.

Liquid silicon metal is tapped from the bottom of the furnace, and the thoroughly mixed raw materials are charged on the top. The reaction co-product, carbon monoxide CO(g), is further oxidised to carbon dioxide CO_2(g) in open furnaces and released into the atmosphere. In open furnaces, side-reactions leading to the formation of silica fumes play an important role for the overall economics of the process:

$$\mathrm{Si(l)} + \tfrac{1}{2}\mathrm{O_2(g)} = \mathrm{SiO(g)} \tag{5.8}$$

$$\mathrm{SiO(g)} + \tfrac{1}{2}\mathrm{O_2(g)} = \mathrm{SiO_2(g)} \tag{5.9}$$

The silica fumes, which consist mainly of very fine particles of amorphous silica, less than 1 μm, are passed through filter cloths installed in large bag-house systems adjacent to the furnaces. The collected amorphous finely divided silica finds valuable applications as additives in concrete and refractory. Depending on the quality of the raw materials used and the operational strategy and skills, the silicon yield as metallurgical silicon ranges from 80 to 90%, the balance resulting in silica fume.

Reactions (5.7–5.9) are a simplification of the complex system. Several main principles can be understood from a more detailed description of the chemistry. There are two important intermediate compounds: the gaseous silicon monoxide SiO(g) as already mentioned in Reaction (5.8) and the solid silicon carbide SiC(s). To interpret the chemistry occurring in the furnace, it is convenient to conceptually split the furnace reaction inner space into an inner hot zone and an outer cooler zone. Liquid silicon is produced in the inner zone, where the dominant chemistry is described by the reactions:

$$\mathrm{2SiO_2(l) + SiC(s) = 3SiO(g) + CO(g)} \tag{5.10}$$

$$\mathrm{SiO(g) + SiC(s) = 2Si(l) + CO(g)} \tag{5.11}$$

The temperature in the inner zone is in the range 1900–2100 °C, allowing a high proportion of SiO(g) in this zone, which is absolutely indispensable for further reduction according to Reaction (5.11).

In the outer zone, where temperature is below 1900 °C, SiO(g) and CO(g) convected away from the inner zone meet and react with free carbon. Consequently, silicon carbide SiC(s) and condensation products of Si(l) in a matrix of SiO_2(s,l) are formed as the partial pressure of SiO(g) drops:

$$SiO(g) + 2C(s) = SiC(s) + CO(g) \quad (5.12)$$

$$2SiO(g) = Si(l) + SiO_2(g) \quad (5.13)$$

A schematic description of the furnace is given in Figure 5.1.

The high-temperature nature of this process implies operation as continuous as possible. Raw materials are therefore fed in small batches at frequent intervals, and are judiciously distributed on the top of the charge. Liquid silicon is continuously, or at frequent intervals, drained out from the bottom of the furnace, whereas gas exhaust and fumes are constantly passing through the filter to clean the fumes and recapture the silica.

Liquid crude silicon contains 1–4% impurities, depending on the raw materials and the type of electrodes. The main impurities are: iron Fe: 0.2–3%, aluminium Al: 0.4–1%, calcium Ca: 0.2–1%, titanium Ti: 0.01–0.1%, carbon C: 0.1–0.15%, oxygen O: 0.01–0.05%. The other main impurities are metals, particularly transition elements V, Cr, Mn, Co, Ni, Cu, Zr, Mo present in tens to hundreds ppm(w) each. Boron B and phosphorous P are always present in concentration 10–100 ppm(w) each.

5.3.2 Ladle Refining

Most of the applications of silicon as described above in Section 5.2.4 require further refining. The crude silicon is therefore tapped as liquid in large ladles (containing up to ten tons of silicon) and treated when still liquid with oxidative gas and slag-forming additives, mainly silica sand (SiO_2) and lime/limestone ($CaO/CaCO_3$). Other chemicals such as dolomite (CaO–MgO), calcium fluoride (CaF_2) and others are used, depending on plant practice and customer requirements. Elements less noble than silicon such as Al, Ca and Mg are oxidised and the degree of refining is determined by distribution equilibria (Reactions 5.14–5.17), where the (parentheses) refer to components dissolved in a slag phase and the underscored symbols refer to dissolved elements in liquid silicon:

$$4\underline{Al} + 3(SiO_2) = 3Si(l) + 2(Al_2O_3) \quad (5.14)$$

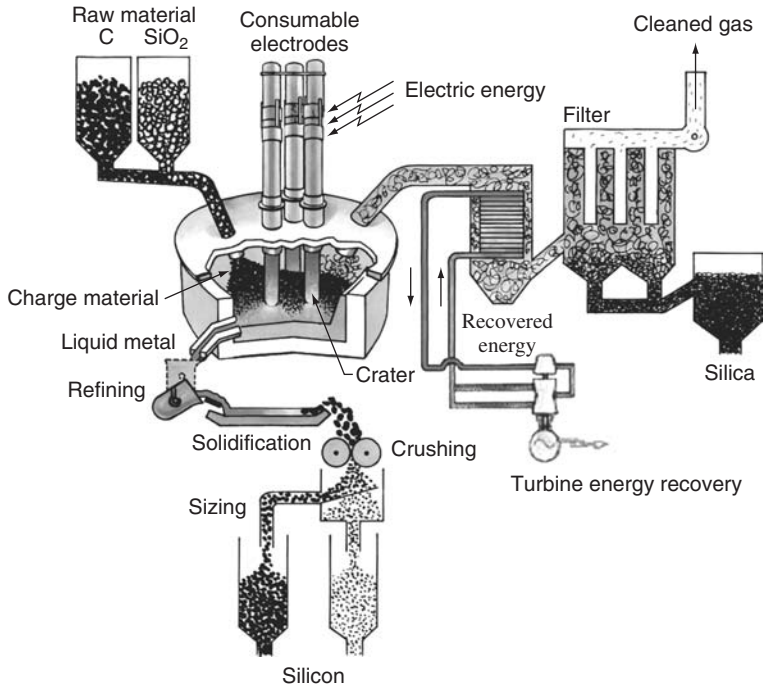
$$2\underline{Ca} + (SiO_2) = Si(l) + 2(CaO) \quad (5.15)$$

$$2\underline{Mg} + (SiO_2) = Si(l) + 2(MgO) \quad (5.16)$$

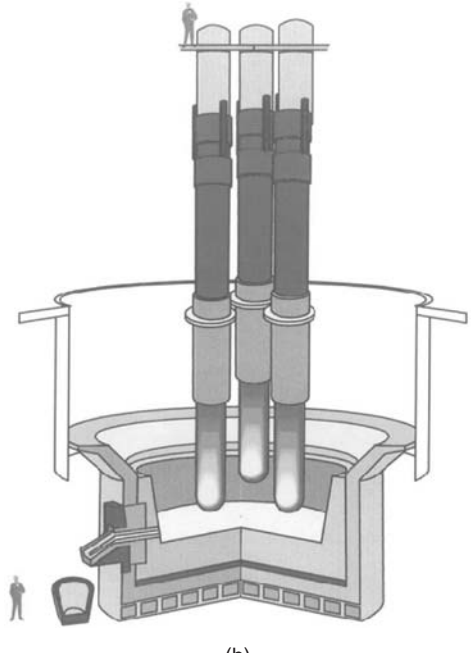
$$Si(l) + O_2 = (SiO_2) \quad (5.17)$$

Theoretically it is possible to remove Al and Ca to very low levels, but in practice this is prevented by the large heat losses occurring during this operation. The temperature drops to 1700–1500 °C, and to avoid freezing of the melt, some of the silica needed for slag formation is provided by direct oxidation of Si(l) with added oxygen (oxygen is added to heat silicon to keep it liquid through the exothermic oxidation of silicon). A disadvantage of this operation is also a partial oxidation of silicon, resulting in expensive material losses.

After completion of oxidative refining in the ladle, the slag, which contains part of the impurities, is removed mechanically or by gravity, and liquid silicon is poured into a casting mould.



(a)



(b)

Figure 5.1 (a, b) Schematic representation of a furnace for production of metallurgical grade silicon. Reproduced from Schei A, Tuset J, Tveit H, *Production of High Silicon Alloys*, Tapir forlag, Trondheim (1998) [14] with permission of Halvard Tveit

The slag-forming additives influence the slag density and viscosity, hence the practical separation of slag and the ultimate purity of the poured silicon. For instance, a high CaO content will lead to a low viscosity slag, which will sink to the bottom of the ladle, while CaF_2 will increase the viscosity. Sufficiently different properties in density and viscosity of both the slag and the molten silicon are required to achieve a good separation. Many studies and practical on-site developments have been devoted to this step of the process [14].

Carbon is present in crude liquid silicon mainly as dissolved C and suspended SiC particles. The fraction of SiC increases as the temperature is lowered; SiC particles are then efficiently captured by the slag phase and thus are removed from liquid silicon during the ladle treatment and the subsequent pouring. SiC is removed simply by mechanical separation, precipitated particles sticking to the walls of the ladle and the other devices containing the liquid silicon [15]. Dissolved carbon in the range 80–100 ppm(w) in best cases will finally remain in the purified alloy of metallurgical silicon.

Oxygen is present in liquid silicon mainly as dissolved O, but when solidified one may find in the silicon matrix oxide particles mainly from the slag.

The use of these refining principles to prepare solar grade silicon will be further discussed later in Section 5.7.2.

5.3.3 Casting and Crushing

The refined melt is poured from the ladle into a cast iron mould or onto a bed of silicon fines. The casting should preferably be removed from the mould while 'bleeding', that is, not fully solidified. After solidification in standard industrial conditions, metallurgical grade silicon is multicrystalline. The individual Si grains vary in size typically from 1 mm close to the iron mould wall to up to more than 100 mm in the centre section if cast on a bed of silicon fines [8]. The impurities are generally located at the Si grain boundaries as silicides and intermetallic compounds, but may also be incorporated in the Si grains if solidification has been sufficiently rapid [16]. Oxides and carbides are found as inclusions located at the grain boundaries and to a lesser degree inside the Si grains.

For use in customers' processes, solidified silicon needs to be further crushed down to small lumps up to 100 mm. This is performed in jaw crushers and roll crushers, since at room temperature metallurgical grade silicon is hard and brittle. This operation generally generates a significant amount of fines, which are undesirable because they may be contaminated by impurities and are difficult to handle during further transport and handling. Therefore, fines are removed after the primary crushing. The dominant fracture mode was found by Forwald *et al*. [4, 17] to be transgranular. For chemical applications, silicon lumps need to be further reduced to small powder particles of a few tens to a few hundreds of micrometres. This is carried out in industrial equipment such as ball mills.

Alternative methods based on rapid cooling have been developed to increase the homogeneity of the solidified structure of silicon through an even distribution of the impurities and intermetallic phases. Granulation in water, resulting in small granules of a few millimeters and thus avoiding casting and subsequent coarse crushing, has become a standard practice for several producers [18–20]. In an earlier attempt to avoid casting, crushing and milling, gas atomisation was tested by producers and users, but was not further industrialised for economic reasons [21, 22].

5.3.4 Purity of Commercial Silicon Metal

Purity characterizing silicon metal, as produced and refined along the above principles and as marketed to the million tons aluminium and chemical market, is reported in Table 5.2. The table

indicates for each impurity a rather broad range, but specifications by application and by customer are much tighter. This is particularly true for iron, one of the most important impurities in silicon metal; only a few applications tolerate concentration higher than 0.35% for this element.

The purity of the commercial grades in Table 5.2 is not high enough for direct use in the photovoltaic industry. Further refining is necessary. The applications to the manufacture of solar grade silicon of the refining principles described above in this Section 5.3 will be further discussed later in Sections 5.7 and 5.7.2. Before doing so, we need to keep in mind that oxidative refining in the ladle affects aluminium, calcium, magnesium, carbon and boron, but not iron and the other transition elements. Casting and crushing have some effect on these latter impurities as they segregate at the grain boundaries and can partly be removed by screening and magnetic separation. Chemical wet leaching and washing are also used as refining techniques with respect to these elements [23].

5.3.5 Economics

The carbothermic reduction of quartz in the submerged arc furnace consumes large amounts of energy and raw materials. Best industrial furnace performances are 10–11 MWh per metric ton of silicon metal and 90% silicon yield. Therefore, the silicon metal economics are extremely sensitive to both availability and price of electrical power and raw materials such as quartz and coal. There are wide differences in the economical factors influencing the cost from region to region and from producer to producer. It is not our purpose to make a detailed economic analysis. We will limit us to indicate general trends useful to a good understanding of this chapter. Table 5.3 reports aggregated statistical values for western producers in 1993 [24] and both western and Chinese producers in 2006 [25].

In spite of an apparent relative stability of the cost structure for the western producers as shown in Table 5.3, the industry has gone through dramatic developments (sharp increase of power and raw material price) and improvements (electrode technology, labour reduction) at the individual producer level.

Although the growth is relatively high for this type of traditional metallurgical product (8% annual growth over the last decade) western producers have not expanded their manufacturing capacity. In a few cases they have converted ferrosilicon furnaces to silicon metal as the technology and the equipment are quite the same. Capacity expansion, either on green- or brown-field sites requires affordable long-term contracts on electrical power, but these are now more difficult to renew or acquire than in the past. A green plant with a typical annual capacity of 40-60 000 tons

Table 5.3 Production cost structure of silicon metal by western (1993, 2006) and Chinese producers (2006)

Cost Factors	Western producers	Western producers	Chinese producers
	Year 1993 [24]	Year 2006 [25]	Year 2006 [25]
Reduction materials (as coal)	20%	22%	17%
Quartz	9%	8%	5%
Electrodes	12%	9%	9%
Electric power	21%	28%	46%
Supplies and equipment	16%	17%	14%
Labour	17%	11%	4%
Transport (to customers)	5%	5%	5%

takes three years to build and requires at least 200 million US$, a capital expenditure the western producers are not willing to risk until they see prices above 2.5 US$/kg. This is well (40–50%) above the average price they enjoyed during the last decade, and this, in spite of a sharp price increase in 2006–2008, following all raw material indexes until the international 2008 financial crisis preceded the 2009 recession. Therefore, the capacity expansion has, during the past decade, come only from China, which has grown to be the largest producing and exporting country, with half of the worldwide capacity.

Instead of expanding their capacity, western producers prefer to focus on other and new businesses. Solar grade silicon through the metallurgical route is offering such opportunities. The most renowned producers and a few others are now allocating all their R&D and technical resources on solar projects (see Section 5.7.2.4).

5.4 PRODUCTION OF POLYSILICON/SILICON OF ELECTRONIC AND PHOTOVOLTAIC GRADE

Beside silicon metal, the other well-established commercial grade of silicon is the hyper-pure material for the semiconductor industry, also, as mentioned earlier, called polysilicon. Impurities in the ppb(a)–ppt(a) range are required for polysilicon supplied to the semiconductor industry. The ultra-high purity is needed to ensure exacting semiconductor properties in the grown silicon crystals. This is achieved first by the preparation of a volatile silicon hydride and its purification, generally using fractional distillation. This is followed by the decomposition of this hydride to hyper-pure elemental silicon by reductive pyrolysis or chemical vapour deposition. The preparation of the volatile Si compound involves external reactants and its decomposition generates by-products, which need to be recycled. The various polysilicon routes must therefore control four successive steps. All have a strong impact on the overall feasibility and economics of the suitable polysilicon product:

1. preparation/synthesis of the volatile silicon hydride;
2. purification;
3. decomposition to elemental silicon;
4. recycling of by-products.

Many processes to produce polysilicon have been tested, patented and a few operated for many years. Only three large commercial processes are currently active.

The most popular process is based on the thermal decomposition of trichlorosilane at 1100 °C on a heated silicon rod or filament placed inside a deposition chamber. This process, which was developed in the late 1950s, is commonly referred to as the *Siemens process* with reference to the company that carried out its early development.

$$2SiHCl_3 = SiCl_4 + Si + 2HCl \tag{5.18}$$

In 2001 this process and its variants still accounted for at least 60% of the worldwide production of polysilicon.

In a more recent process (early 1980s) developed by *Union Carbide Chemicals* in the United States of America and by *Komatsu Electronic Materials* in Japan, the trichlorosilane has been replaced by monosilane SiH_4, but the principle of decomposition on a heated silicon rod inside a closed deposition chamber is maintained.

$$SiH_4 = Si + 2H_2 \tag{5.19}$$

This process, presently run by the company *Renewable Energy Corporation Silicon, Inc*., the successor of *Advanced Silicon Materials, LLC* in the United States has gained significant market acceptance during the past 25 years.

Finally, in the third process, also making use of monosilane SiH_4, the heated silicon rod in the closed reaction chamber has been replaced by a fluidised bed of heated silicon particles. The particles act as seeds on which SiH_4 is continuously decomposed to larger granules of hyper-pure silicon. Unlike the two first mentioned this process is a continuous one. It is known as the *Ethyl Corporation process*, after the name of the US chemical company that developed it also in the 1980s to 1990s. This process is presently run by the US corporation *MEMC* in Pasadena, Texas, the indirect successor of Ethyl Corporation. At the time of writing it is also introduced by Renewable Energy Corporation at its Moses Lake plant in Washington (USA). Both companies have, however, their own proprietary technology both to make silane and to decompose it in fluidised bed reactors.

The respective features, advantages and disadvantages of these different routes are described in the following sections.

5.4.1 The Siemens Process: Chlorosilanes and Hot Filament

A schematic overview of the process is given in Figure 5.2.

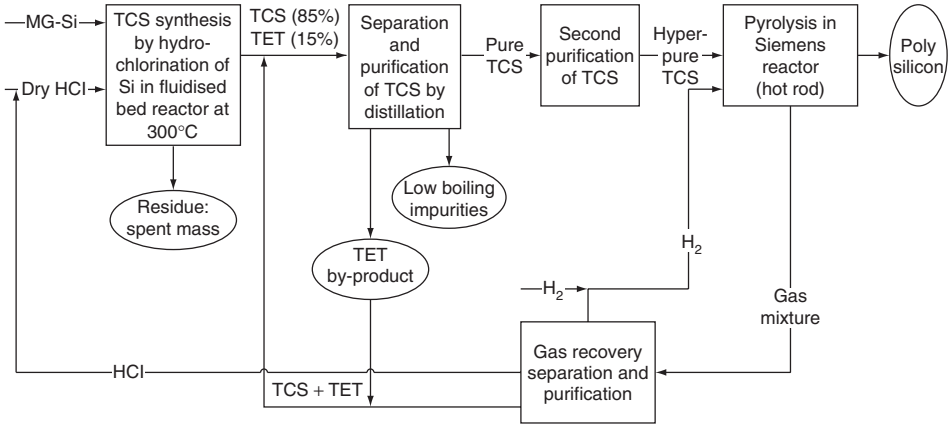
Trichlorosilane $HSiCl_3$ is prepared by hydrochlorination of metallurgical grade silicon in a fluidised bed reactor:

$$Si(s) + 3HCl = HSiCl_3 + H_2 \tag{5.20}$$

This reaction occurs at 300–350 °C, normally without a catalyst. A competing reaction is

$$Si(s) + 4HCl = SiCl_4 + 2H_2 \tag{5.21}$$

contributing to the formation of unsuitable tetrachlorosilane in molar proportion of 10–20%.



MG-Si: Metallurgical Grade Silicon
TCS: Trichlorosilane
TET: Tetrachlorosilane

Figure 5.2 Schematic representation of the Siemens process

Trichlorosilane is chosen because of its high deposition rate, its low boiling point (31.8 °C) and its comparatively high volatility and hence the ease of purification with respect to boron and phosphorus down to the ppb level. The boiling point of other silanes frequently found with trichlorosilane are as follows: SiH_4 (−112 °C), SiH_2Cl_2 (8.6 °C) and $SiCl_4$ (57.6 °C). The suitable trichlorosilane undergoes a double purification through fractional distillation, the first step removing the heaviest components resulting from the direct synthesis and the second step eliminating the components lighter than trichlorosilane, also called *volatiles*.

High-purity $SiHCl_3$ is then vaporised, diluted with high-purity hydrogen and introduced into the deposition reactors. The gas is decomposed onto the surface of heated silicon seed rods, electrically heated to about 1100 °C, growing large rods of hyperpure silicon.

The main reactions in the deposition chamber are:

$$2SiHCl_3 = SiH_2Cl_2 + SiCl_4 \tag{5.22}$$

$$SiH_2Cl_2 = Si + 2HCl \tag{5.23}$$

$$H_2 + HSiCl_3 = Si + 3HCl \tag{5.24}$$

$$HCl + HSiCl_3 = SiCl_4 + H_2 \tag{5.25}$$

The stream of reaction by-products, which leaves the reactor, contains H_2, HCl, $HSiCl_3$, $SiCl_4$ and H_2SiCl_2.

A schematic representation of the Siemens reactor is given in Figure 5.3.

The Siemens process is highly energy consuming, a major part of the energy being dispersed and lost. To avoid deposition on the inner surfaces of the reaction chamber, this has to be cooled. Originally, the decomposition chamber consisted of a quartz bell jar containing one single inverted U-shaped silicon seed rod. A major advancement in polysilicon production was the utilisation of metal bell jars in place of the quartz bell jars. Quartz bell jars could not be produced in large diameters and were susceptible to breakage. The development of steel bell jars made it possible to accommodate 30 or more inverted U-rods in each reactor. This dramatically increased the productivity while decreasing the energy consumption per kilogram of produced polysilicon.

As Reactions and equilibria (5.22–5.25) show, the deposition process generates by-products. Unfortunately, for each mole of Si converted to polysilicon, 3–4 moles are converted to $SiCl_4$, binding large amounts of chlorine and valuable silicon. The main industrial application of tetrachlorosilane is as a source material to produce pyrogenic (also called *fumed*) silica as described above in this chapter. The fumed silica market grows at a much slower rate (5% per year) than the polysilicon industry, boosted by the solar demand. Moreover, a significant portion of fumed silica is also produced by burning derivatised by-products from the silicones industry. In the early stages of the polysilicon industry, the fumed silica business could absorb the excess of silicon tetrachloride generated by the Siemens process. This explains the arrangements between polysilicon and fumed silica producers all around the world. With polysilicon production growing much faster than silicones and fumed silica, the question was whether to eliminate or recycle the tetrachlorosilane. This became an environmental and economic necessity. The concept of recycling on site the by-product back to the valuable starting material to form a closed-loop production process is generally an ideal preferred solution today. There are two basic chemical processes applicable to reconvert $SiCl_4$ to $SiHCl_3$:

1. The high temperature reduction of silicon tetrachloride with hydrogen.

$$SiCl_4 + H_2 = SiHCl_3 + HCl \tag{5.26}$$

At about 1000 °C, a 1:1 molar mixture of $SiCl_4$ and H_2 produces approximately 20–25% molar $SiHCl_3$ in the gaseous mixture. This process requires a fair amount of electrical energy, but

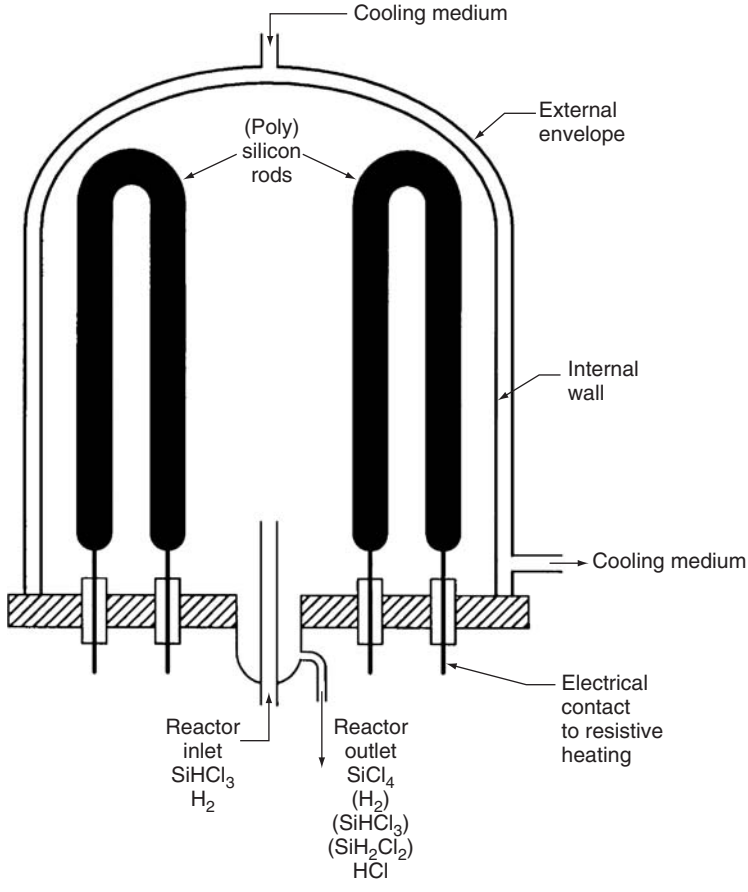


Figure 5.3 Schematic representation of the traditional Siemens reactor

has a distinct advantage that the trichlorosilane produced is of very high quality because both reactants, silicon tetrachloride and hydrogen, are basically electronic grade when produced by Reactions (5.22) and (5.25).

2. The hydrogenation of silicon tetrachloride in a mass bed of metallurgical silicon.

$$3SiCl_4 + 2H_2 + Si = 4SiHCl_3 \tag{5.27}$$

This hydrogenation reaction can produce approximately 20–30% trichlorosilane at 500 °C, 35 atm with a 1:1 ratio of $SiCl_4$ to H_2 in one pass through a mass bed of metallurgical grade silicon in a fluidised bed reactor.

In spite of its widespread and dominant position in the industry, the Siemens process as described above suffers from the following disadvantages:

- High energy consumption, over 90% of the input power is lost to the cold walls of the reactor.
- Two power supplies and preheating of the seed rods are normally required because the high-resistivity (~230 000 Ω cm) seed rods require very high power supplies and high initial power

rates to heat the rods. Therefore, a separate power supply for quartz lamps or graphite rod induction heating is used to preheat seed rods to about 400 °C (~0.1 Ω cm). Lower power electrical supplies can then be used to provide continued heating and control.

- Electrical contacts to seed rods are made of graphite, which is a source of contamination.
- Power failure (especially when starting the process) causes run abortion.
- Hot spot formation and filament burn out may occur.
- Problems arise owing to gas inclusions and to nonuniform deposition at the joints.
- Gas flows and electrical power have to be adjusted during the process to obtain optimal deposition rate.
- The process is operated batchwise.
- Large amounts of by-products need to be handled or recycled.

More recent developed processes have attempted to overcome some of these disadvantages.

5.4.2 The Union Carbide and Komatsu Process: Monosilane and Hot Filament

Research on this process was initiated in 1976 after the international oil crisis. The US Government funded several projects with the objective of finding a route to inexpensive solar grade polysilicon. The Union Carbide process for silane production, with fluidised bed production of polysilicon, was selected for further funding. When the funding was not provided for political reasons, Union Carbide decided to use the silane technology for the production of semiconductor grade polysilicon. Silane deposition technology for polysilicon rods was licensed from Komatsu Electronic Materials, Japan, which also became Union Carbide's main polysilicon customer. In 1990, the business was sold to Komatsu, which became *Advanced Silicon Materials Inc*. (*ASiMI*). By 1998, two large industrial plants were constructed in the United States with the nominal capacity of 5500 metric tons polysilicon, thus ranking ASiMI as number three worldwide among the polysilicon producers. In 2002 ASiMI formed the joint venture company *Solar Grade Silicon LLC* with Renewable Energy Corporation (*REC*) of Norway to run the oldest of both plants (Moses Lake, Washington) exclusively for the solar market and to develop the deposition technology in fluidised bed reactors. In 2005 REC acquired from Komatsu the total polysilicon business and technology of ASiMI. REC has then expanded the capacity of both plants and maintains a ranking of number two or three (volume-wise) among the polysilicon producers in the world. A schematic overview of the process is given in Figure 5.4.

The main process steps are as follows:

The hydrogenation of tetrachlorosilane through a mass bed of silicon metal is carried out in a fluidised bed reactor as already described by Reaction (5.27).

The trichlorosilane is separated by distillation while the unreacted tetrachlorosilane is recycled back to the hydrogenation reactor.

The purified trichlorosilane is then redistributed in two separate steps through fixed-bed columns filled with quarternary ammonium ion exchange resins acting as catalyst to both redistribution equations:

$$2HSiCl_3 = H_2SiCl_2 + SiCl_4 \tag{5.28}$$

$$3H_2SiCl_2 = SiH_4 + 2HSiCl_3 \tag{5.29}$$

Products of Reactions (5.28) and (5.29) are separated by distillation.

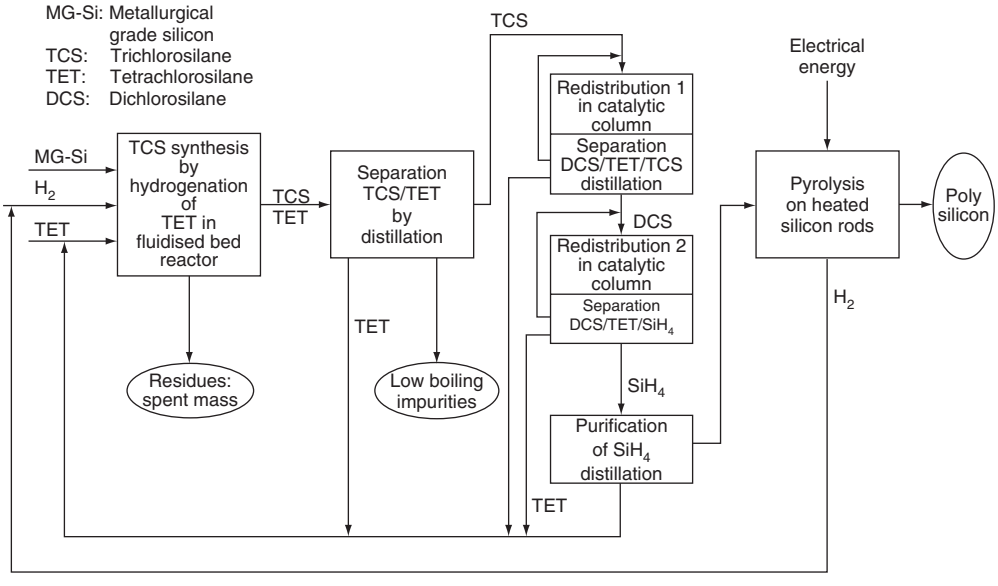


Figure 5.4 A schematic representation of the Union Carbide Polysilicon process

Tetrachlorosilane and trichlorosilane are recycled to the hydrogenation (5.27) reactor and the first redistribution step (5.28), respectively. Silane is further purified by distillation and then pyrolysed to produce polysilicon onto heated silicon seed rods mounted in a metal bell-jar reactor:

$$SiH_4 = 2H_2 + Si \tag{5.30}$$

With hydrogen and chlorine recycled, the only raw material requirement is metallurgical grade silicon in granular form designed for fluidisation. Because Reactions (5.27–5.29) yield low portions of suitable products and because distillation has to take place after each of them, the intermediates tri- and tetrachlorosilane are recycled and purified many times before conversion to silane. This results in extremely high purity for the silane and the subsequent polysilicon. This is operated as a closed-loop process.

Other advantages of using SiH_4 are that the pyrolysis may be operated at significantly lower temperature, e.g. 800 °C, the decomposition is complete, conversion efficiency is higher and no corrosive compounds are formed. Uniform, large-diameter, long, dense, void-free cylindrical rods of polysilicon produced this way are particularly suitable for single-crystal manufacturing by the floating zone (FZ) method.

The disadvantage of the monosilane-based process is the higher cost of the volatile molecule, since additional steps are requested to convert trichlorosilane to monosilane. Moreover, the recycling of unsuitable chlorosilanes is compulsory when choosing this option because the redistribution equations yield only a small percentage of the suitable silane.

Also silicon powder formation resulting from homogeneous decomposition of silane to silicon must be avoided by cooling the inner part of the reactor chamber, concentrating the heat on the rods. Efficient cooling results, however, in additional heat losses and higher energy consumption.

5.4.3 The Ethyl Corporation Process: Silane and Fluidised Bed Reactor

This process was developed by the US company Ethyl Corporation at the same time, in similar conditions and political context as the above-described Union Carbide process. Although managed differently, the outcome of both projects was similar in the sense that both shifted their focus from solar grade polysilicon and ended up as new commercial polysilicon processes serving the electronic industry. The Ethyl Corporation process is, by comparison with the Siemens and the Union Carbide processes, revolutionary in all aspects except the concept of purifying and decomposing a volatile silicon compound by pyrolysis.

The first radical change was the choice not to use metallurgical grade silicon as the primary raw material for silane. The idea was to make use of alkaline fluorosilicate (M_2SiF_6, M being an alkaline element), which is a waste by-product of the huge fertiliser industry. Tens of thousands of tonnes of such fluorosilicates every year are available. This is potentially a very low-cost starting material. Silicon tetrafluoride SiF_4 is sublimated by heating the fluorosilicates. Silicon tetrafluoride is then hydrogenated to monosilane by metal hydrides such as lithium aluminium hydride or sodium aluminium hydride.

$$2H_2 + M + Al = AlMH_4, \text{ M being Na or Li} \tag{5.31}$$

$$SiF_4 + AlMH_4 = SiH_4 + AlMF_4 \tag{5.32}$$

$AlMF_4$ is believed to find application in the aluminium industry, making it a valuable saleable product.

After distillation, monosilane SiH_4 is thermally decomposed to polysilicon as described by Reaction (5.30). However, to realise this process, Ethyl Corporation introduced a second radical change, not using static silicon seed rods in a bell-jar reactor, but dynamic silicon seed spheres or 'beads' in a fluidised bed sustained by a gas stream of silane and hydrogen. A schematic representation of a fluidised bed reactor is given in Figure 5.5.

The fluidised bed reactor offers some significant advantages compared with the bell-jar reactor. Most of the shortcomings identified for the Siemens process are then eliminated. The energy losses and hence the energy consumption are considerably reduced because the decomposition operates at a lower temperature and because the requirement to cool the bell jar does not apply. It is a major advantage as specific energy consumption for decomposition can be reduced by 80%. Another advantage is that large reactors may be constructed and operated continuously, reducing further the capital and operating costs. The product coming out of the reactor is ready for use: no post-treatment, no crushing are necessary, thus avoiding significant work load and possible contamination.

The end products are small granules of polysilicon that present some advantages when automation in manufacturing is appropriate or when continuous feeding is requested in the customer process.

A disadvantage of the process is the generation of powder due to homogeneous decomposition of SiH_4 in the free reactor space and the possible hydrogen absorption into the polysilicon deposition layer. Because of high specific surface and risk of contact between the granules and the equipment wall, some contamination is hardly unavoidable. Because of that fluidised bed granules have not gained full acceptance in the semiconductor market and are rather considered a better material for photovoltaics.

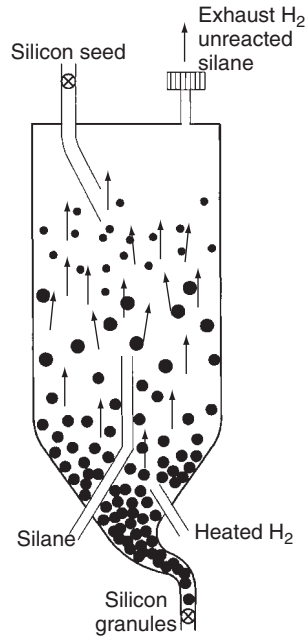


Figure 5.5 A schematic representation of a fluidised bed reactor for polysilicon production

5.4.4 Economics and Business

Until 2003–2004 the polysilicon business was entirely dedicated to the semiconductor industry, the solar sector was considered as a convenient outlet for rejects and downgraded material. The formation of *Solar Grade Silicon Inc* (see Section 5.4.2 above) reopening an idle polysilicon plant with the sole purpose of feeding the growing photovoltaic market was first to change the perception on the industry. In 2003–2004 a persistent shortage of silicon feedstock became an undeniable reality which opened the eyes of many players and observers. Encouraged by sharp price increases and attractive long term sales contracts, the producers responded with capacity expansion at their existing sites (de-bottlenecking green- and brown-field new lines or plants). In 2006 the polysilicon portion consumed by the solar market surpassed for the first time the portion going to semiconductor market. Now the three largest producers (together accounting for more than half of the global output) appear more dedicated to the solar than the semiconductor market. The global polysilicon output in 2007–2008 is around 50 000 metric tons. There is no official listed price for polysilicon. We suggest for the recent past period that prices have widely fluctuated between 50 US$/kg (contract price) to several hundreds US$/kg (spot or secondary market price). We further suggest that the average price paid by the solar industry during the past five years was around 65 US$/kg. We may also for the future expect ample fluctuations between 20 and 100 US$/kg. The polysilicon business is a multibillion US$ business, supporting two multibillion markets: electronics and photovoltaics.

The first characteristic of the polysilicon industry is to be extremely capital intensive. New plants or lines have capacity from 2 000 to 10 000 metric tons, the specific capital requirement being 130–200 US$/kg.

The second characteristic is that the decomposition process requires fairly large amounts of energy. In the early stages of the industry, consumption of 350 kW h per kg polysilicon was not unusual. Efficiencies improved to 150–160 kW h per kg, and advanced processes are now about

100 kW h per kg. The fluidised bed process is even more efficient. A long-term contract on electrical power is therefore a prerequisite to any expansion project. This is a serious challenge for a rapid global expansion of polysilicon able to respond to the solar demand.

The long-established members among the polysilicon producers are concentrated in the United States (five plants, four producers), the European Union (two plants, two producers) and Japan (three plants, two producers). High margins, long-term sales contracts and favourable outlook for the deployment of solar energy have attracted new comers into this business (e.g. chemical and metallurgical companies, solar companies having a strong foothold in the downstream part of the value chain, investors, etc). Several new players have already successfully entered into the market in Japan, China, Korea and several projects are in the completion phase in Europe, USA, China, Korea and Russia.

Until years, technology was proprietary and well protected within the confines of each long-established producer. This protective strategy seems not able to resist the pressure created by the recent extraordinary demand. Now several technology suppliers and engineering companies are offering turnkey projects to potential new entrants. So far the technology offered is limited to chlorosilane synthesis, chlorosilane recycling and polysilicon deposition from trichlorosilane in Siemens type (hot filament) reactors. All new projects incorporate the recycling of chlorosilane by-products.

Because of capital and energy consumption polysilicon remains an expensive product. It costs 20–30 times more to produce than silicon metal. The mass deployment of photovoltaic energy calls for drastic cost and energy reduction. Both the long-established players and the newcomers spend considerable R&D and technical resources to develop new and innovative processes to address this challenge. These will be presented later in Section 5.7.1.

5.5 CURRENT SILICON FEEDSTOCK TO SOLAR CELLS

We have seen in Sections (5.3) and (5.4) that two main commercial grades of silicon have coexisted for a long period (1960 onwards): on one hand silicon metal available in millions of tons at a fluctuating price of 1–4 US$/kg, but containing percentage quantities of impurities; on the other hand polysilicon developed for semiconductor manufacture, containing impurities at the ppb level, but available in tens of thousands of tons at a fluctuating contract price of 20–100 US$/kg.

Chemical purity of typical silicon metal is given in Table 5.2. For some price premium, a few silicon producers can upgrade it to higher purity by metallurgical treatment, as tentatively indicated in Table 5.4. For semiconductor grade silicon, impurity levels are the ppb level and resistivity within 1000–30 000 Ω cm.

Table 5.4 Metallurgical silicon and polysilicon best purity

Impurity	Upgraded silicon metal, including one-directional solidification	Solar grade polysilicon	Electronic grade polysilicon
Total metallic impurities (ppm(w))	<1	<0.05	<0.001
Donor/phosphorous (ppm(a))	<5	<0.005	<0.0005
Acceptor/boron (ppm(a))	<5	<0.0005	<0.0001
Carbon (ppm(w))	<50	<5	<0.1
Oxygen (ppm(w))	<100	<5	<1

Table 5.5 Specification of chemical impurities in lowest-grade silicon currently purchased to produce multicrystalline silicon wafers

Impurity	Specification
Fe, Al, Ca, Ti, metallic impurities	Less than 0.1 ppm(w) each
C	Less than 4 ppm(w)
O	Less than 5 ppm(w)
B	Less than 0.3 ppm(w)
P	Less than 0.1 ppm(w)

Table 5.6 Electrical specification of lowest-grade silicon currently purchased to produce multicrystalline silicon wafers

Property	Specification
Resistivity	Higher than 1 Ω cm, *p*-type
Minority-carrier lifetime	Higher than 25 μs

The minimum purity required for growing multicrystalline silicon ingots for wafering is given in Tables 5.5 and 5.6. We believe these guidelines have been observed by the ingot makers in the most recent period. However, we have observed that current praxis is to select the material by classes of resistivity more than by chemical purity control.

From Tables 5.4–5.6 we can see that the impurity content in the best grade of silicon metal, particularly boron and phosphorous, prevents the exclusive use of such material to make solar cells. However, purity requirements are not as stringent for solar cells as for semiconductors and therefore, polysilicon appears too good and unnecessarily expensive.

In year 2000 the photovoltaic industry consumed around 4000 metric tons of silicon, five years later the consumption was up to 17 000 tons and in 2008 it is estimated close to 40 000 tons. Here it must be noticed that significant progress has been made to reduce the specific consumption of silicon related to the cell electrical output (g Si/W). Progress has been achieved by reducing the thickness of the wafers (from 320 to 180 micrometres) and by overall increase of the cell conversion efficiency. Between 2000 and 2008 the specific consumption decreased from 15–17 g/W to 7–9 g/W, making thousands of tons of silicon available for manufacturing more cells.

Until 2000, in order to reduce prices, the industry selected its silicon raw material from various second-grade classifications of semiconductor silicon, less expensive than prime-grade polysilicon. We have assumed that this feedstock has cost the PV industry, on average, one-third of what prime-grade polysilicon would have cost. This feedstock consisted of a mix of rejects from crystal growth (*tops and tails*, ingots from aborted runs, loss structure ingot sections, crucible leavings also called *potscrap*) and rejects from polysilicon manufacture of many sorts (these were reviewed in details in the first edition of this handbook, Chapter 5). Rejects from crystal growth and from polysilicon plants in 2000 accounted in total for 2500–3000 metric tons at most. The balance was necessarily covered by standard semiconductor grade polysilicon, either prime-grade or a grade produced on purpose for photovoltaics. Around 2000 polysilicon capacity exceeded the demand by some 5000 metric tons. It was therefore attractive for some producers to offer a solar grade in order to optimise their overall costs by operating their plants at higher capacity. It was also attractive for ingot makers to rapidly increase their capacity and dilute impurities from the other feedstock sources. This rather lucky situation gave the consumers, ingot and cell makers the opportunity to learn about polysilicon. In the same period cell efficiency increased noticeably in all regions.

However, the correlation between improved efficiency and more extensive use of virgin polysilicon as feedstock is to our best knowledge not clearly proven, as many other improving factors were introduced during this same period. Around 2003–2004, a sudden but persistent shortage of polysilicon occurred due to the strong demand both from semiconductor and solar sectors. That triggered, at the producers, some expansions as well as development programs on new routes, both through metallurgical and chemical refining processes, aiming at the best suitable solar grade silicon with respect to quality, purity and cost. These developments are not yet at their term. They will be further described and commented in Section 5.7. Meanwhile, some silicon metal producers or refiners are already offering upgraded versions of silicon metal. Table 5.4 gives a tentative purity characterization of such upgraded silicon metal, which we learned about through recent communications [26, 27].

The shortage of silicon prevailing, many ingot and cell makers have been more than willing to test and evaluate these emerging materials. Currently ingot and wafer producers use a silicon mix in which the bulk part is virgin polysilicon, manufactured mainly in solar dedicated plants. Rejects from crystal growth and polysilicon manufacture continue to be used, but in much lower proportion as their volume has not grown as fast as the demand. Upgraded silicon metal is gradually increased in the mix although not by all ingot producers. Such a mix is more adequate for multicrystalline ingot manufacturing. Single-crystal growers are more careful with their requirement because of operational (nonhomogeneity and risk of structure loss) and commercial reasons (monocrystalline wafers targeting the high-efficiency cell market).

Table 5.7 gives an update of the respective share of the main commercial cell technologies used in terrestrial application. Only 5.2% are non-silicon based and outside the scope of this chapter, whereas 5.2% are silicon thin film technologies, the balance (close to 90%) being crystalline silicon technologies (single, multi- and ribbon).

Monosilane (SiH_4) is the source of silicon required for the deposition of silicon amorphous and microcrystalline thin film in a glow discharge or low-temperature plasma. Silane (monosilane) is mass-produced by the Union Carbide (REC Silicon) and Ethyl Corporation (MEMC) methods, which are described in Sections 5.4.2 and 5.4.3 above. The global annual output capacity is approximately 10 000 metric tons (in 2007), including minor volumes produced in Japan and Korea, and is in rapid expansion. The major part of it is used to produce polysilicon on site (United States), the balance being sold through distributors, for example industrial gas companies, to a vast group of customers. The worldwide silane market for silicon films in the semiconductor, photovoltaic, glass and ceramic industries is about 3 000 tons (in 2007). Applications include passivating and semiconducting layers for integrated circuits particular for flat screens, epitaxial films, architectural glass coatings, special ceramics, surface treatment and amorphous silicon. Silane is available in quantities and purity exceeding the needs of the photovoltaic market. Quantity and cost of silicon is of less importance for silicon thin film since the specific consumption per watt output is 50–100 times less than that for bulk crystalline silicon cells, that is, 100–400 mg/W versus 7–9 g/W as mentioned above.

Table 5.7 Solar modules shipment by technology [1]

Technology	2000 (MW)	2000 (%)	2007 (MW)	2007 (%)
Single-crystalline silicon wafers	107	37.4	1800	42.2
Multicrystalline silicon wafers	138	48.2	1940	45.2
Ribbon/multicrystalline silicon film	12	4.3	100	2.3
Amorphous silicon/microcrystalline silicon	28	9.6	220	5.2
Noncrystalline silicon	1.5	0.5	220	5.2
Total terrestrial PV shipments	287	100	4280	100

5.6 REQUIREMENTS OF SILICON FOR CRYSTALLINE SOLAR CELLS

Impurity concentration, element by element or by combination of some of them, should be the critical criteria for smart selection and mix composition of the silicon feedstock. The impact of a specific impurity in silicon on the final performances is not fully understood. Until now raw material selection and mix composition have been left to semi-empirical rules and individual experience or feeling. At most of the ingot producers the only criterion applied is resistivity measurement. In this section we want to investigate with a scientific approach, combining theory, experimental data and modelling the structural and chemical limitations imposed on silicon to reach optimal performances for crystalline solar cells. After a short description of the valid principles of crystallisation, we will examine the effect of critical impurities and the effect of structural imperfections.

5.6.1 Directional Solidification

When silicon solidifies, a homogeneous melt will normally not result in a homogeneous solid. Concentration gradients will appear as shown in Figures 5.6 and 5.7.

A schematic phase diagram of the silicon corner is shown in Figure 5.6. When cooling a melt of composition X_0, it starts to solidify at a temperature T_0. At a lower temperature, T_1, a fraction has solidified. At equilibrium the fractional compositions of the solid and liquid are X_s and X_l, respectively. The equilibrium distribution coefficient k_0 at T_1 is defined as the ratio of the equilibrium solid and liquid compositions:

$$k_0 = X_s/X_1 \tag{5.33}$$

(at *equilibrium* both the solid and the liquid have homogeneous compositions X_s and X_l, respectively).

During normal solidification the solid and the liquid are not in equilibrium. This principle can be used to purify the material. If: (1) the solid–liquid interface is flat; (2) the melt has a uniform composition; and (3) negligible diffusion takes place in the solid, a composition profile like the one in Figure 5.6 is obtained when equilibrium prevails at the interface.

A horizontal cylinder of liquid alloy of initially uniform composition X_0 is cooled at the left end (see Figure 5.6b). Let a small amount of solid form so that the solid–liquid interface is located at position 1. When a small volume has solidified, the composition in that volume has dropped from X_0 to k_0X_0. A mass of solute proportional to the cross-hatched area to the left of position 1 has been removed from the solid and rejected into the remaining liquid. This will increase the liquid composition to a level above X_0. If we consider that local equilibrium prevails at the solid–liquid interface during solidification, the liquid and solid compositions at the interface are tied together by the equation $X_s = k_0X_1$. When the composition of the liquid is raised, the solid composition must also rise as solidification proceeds. When the solid–liquid interface has moved to position 2, the solid composition will have gradually increased, as shown in Figure 5.6(b). The normal solidification equation for a rod of length L is expressed as

$$\mathrm{X_s(Z) = k_o X_o\{1 - Z/L\}^{ko-1}} \tag{5.34}$$

It is shown in Figure 5.7 that k_0 will have a marked effect on the distribution of impurities in solid silicon. Values of k_0 for different elements are given in Table 5.8 [28, 29]. For impurities having a low value of k_0, for example for iron Fe, $k_{0,\mathrm{Fe}} = 8 \times 10^{-6}$, the solidification process has a large purifying effect: only one Fe atom out of about 100 000 in the melt will enter the solid when

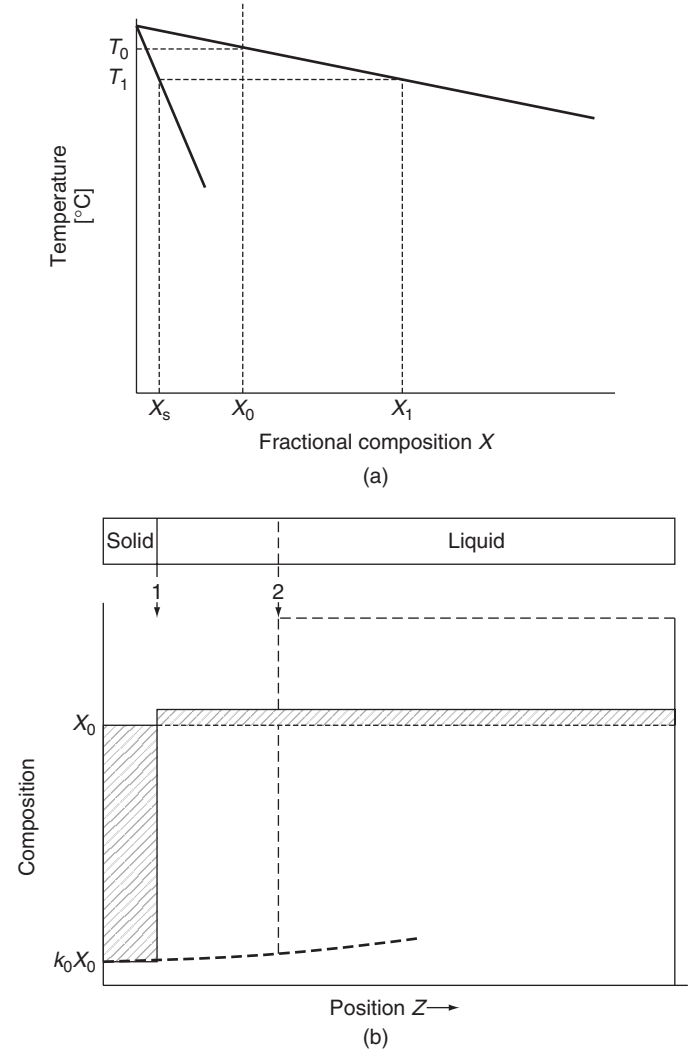


Figure 5.6 (a) Left-hand side of a phase diagram with $k_0 < 1$; (b) composition development in solid and liquid along a rod directionally solidified from the left end under 'normal freezing'

solidification starts. For elements in which k_0 is near to 1, no marked change in the concentration of impurities will appear between the melt and the solid silicon.

During the solidification process, the impurity concentration at the interface varies with the fraction of melt solidified and k_0. If the solute rejected into the boundary region is not transported immediately into the melt in front of the interface, solute will build up at the boundary. As the solute builds up, however, its concentration gradient across the boundary layer becomes steeper and the rate of transport through the boundary layer by diffusion increases until a balance is obtained between the solute being rejected into and out of the boundary layer. At this point the ratio $(X_1)_{\text{interface}}/(X_1)_{\text{bulk}}$ becomes constant. The effective distribution function k_{eff} is defined as

$$k_{\text{eff}} = X_{\text{s,interface}}/X_{1,\text{bulk}} \tag{5.35}$$

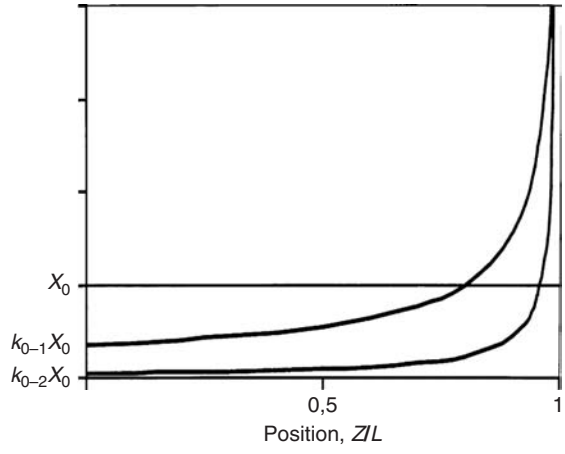


Figure 5.7 Composition profile after the entire rod of Figure 5.6 has solidified, illustrating normal solidification at two different values of k_0, $k_{0-1} = 0.35$ (upper curve) and $k_{0-2} = 0.05$

Table 5.8 Distribution coefficients, k_0, for some elements in silicon at the melting point [14, 28–30]

Element	k_0	Element	k_0
B	0.75	Ti	3.6×10^{-4}
Al	0.002	Cr	1.1×10^{-5}
Ga	0.008	Mn	1×10^{-5}
N	7×10^{-4}	Fe	8×10^{-6}
P	0.35	Co	8×10^{-6}
As	0.3	Ni	8×10^{-6}
C	0.07	Cu	4×10^{-4}
O	0.85	Zn	1×10^{-5}

and tells a great deal about liquid mixing in front of the interface. Poor mixing in the melt results in $k_{\text{eff}} = 1$. Good mixing (no build-up) means $k_{\text{eff}} = k_0$.

Burton *et al*. [31] have derived an expression for k_{eff} when a rotating crystal is pulled from the melt. At a growth rate f of the crystal:

$$k_{\text{eff}} = k_0/[k_0 + (1 - k_0)]\exp^{-\Delta} \tag{5.36}$$

where

$$\Delta = f\delta/D_1 \text{ and } \delta = 1.6D_1^{1/3}\nu^{1/6}\omega^{-1/2} \tag{5.37}$$

where δ is the distance from the growing interface to where the concentration in the melt is uniformly equal to $X_{1,\text{bulk}}$, D_1 is the diffusion coefficient of the solute, ν is the cinematic viscosity and ω the crystal rotation rate.

Kodera [32] has applied this to obtain δ/D_1 and D_1 values for melts doped with various impurities at different rotating rates, as shown in Table 5.9. For Czochralski CZ-grown silicon where the pull speed is around 1 mm/min, Δ is small and k_{eff} is close to k_0 [28].

Table 5.9 Values of δ/D_1 and D_1 for some elements [32]

Impurity element	Rotation rate [rpm]	δ/D_1 [s/cm]	Diffusion coefficient, D_1 [cm^2/s]
B	10	170 ± 19	$(2.4 \pm 0.7) \times 10^{-4}$
	60	84 ± 37	$(2.4 \pm 0.7) \times 10^{-4}$
Al	10	86 ± 34	$(7.0 \pm 3.1) \times 10^{-4}$
	60	40 ± 17	$(7.0 \pm 3.1) \times 10^{-4}$
P	5	127 ± 36	$(5.1 \pm 1.7) \times 10^{-4}$
	55	60 ± 19	$(5.1 \pm 1.7) \times 10^{-4}$
As	5	190 ± 53	$(3.3 \pm 0.9) \times 10^{-4}$
	55	79 ± 16	$(3.3 \pm 0.9) \times 10^{-4}$

Kvande *et al*. [33] estimated $k_{\text{eff}} = 2 \times 10^{-5}$ for iron when solidifying multicrystalline silicon intentionally doped using the standard Bridgman technique. Iron was present both as solid solution and precipitates.

The validity of the assumption of no diffusion in the solid will not be met at a slow solidification rate. Back-diffusion of Fe in a pilot-scale casting experiment has been reported to influence the iron concentration to a depth of 17 mm from the top of a 110-mm-high ingot [33].

The ability to purify molten silicon by solidification will therefore vary for the different impurities and with the solidification parameters. In addition, impurities diffuse from the crucible into the solid, the amount depending on temperature, time and purity of coating. Selection of the optimal casting process parameters and cooling/heating conditions are complicated, and improvements will depend on detailed experiments and the ability to model the total process.

5.6.2 Effect of Crystal Imperfections

It is important to differentiate between impurities in solid solution and those existing as precipitates. The negative effect of impurities on solar cell efficiency present as solid solution may to a large extent be reduced by gettering[12]. Impurities in precipitates at grain boundaries and dislocations show less effect of gettering and may act as impurity sources, leaking into the grains during heat treatments.

Measurements of the local electronic properties in individual grains and near grain boundaries show that the properties may vary from grain to grain. The influence of the boundary plane on interface properties such as solute segregation, energy and kinetics vary with the misorientation and how well the grains fit together. Twins have a low effect on carrier lifetime, while subgrain boundaries have a large negative effect [34]. It is hoped that further research on such topics will clarify relations between grain structure and the electronic properties in multicrystalline silicon.

The density of dislocations influences the lifetime of the minority carriers. Experiments show good correlation between areas having a high dislocation density and a short carrier lifetime. To control the dislocation density during casting has therefore become an important task. The dislocation movements and multiplication, and their interaction with impurity atoms present in solid solution or existing as complexes during the production processing steps challenges our ability to model complex relationships.

[12] For gettering, see Chapter 7.

5.6.3 Effect of Various Impurities

It is well known that impurity atoms have a strong effect on the efficiency of silicon as a photovoltaic material. It is also known that the effect of impurities can be changed by heat treatments and by exposing the material to gettering.

The impurity atoms may appear as solid solutions, as pairs with other elements, for example FeB, BO or as larger aggregates/precipitates with silicon and/or other elements, for example Fe_2Si, SiC to mention a few. This depends upon the temperature, the concentration and the density of the imperfections (dislocations, grain boundaries). If the temperature or the (chemical) surroundings are altered, it will take some time before a new equilibrium is established. The time to reach equilibrium may depend on parameters such as temperature, cooling/heating rate, and chemical composition, grain size, dislocation density and others. When comparing results from literature values in which the specifications of relevant parameters are not defined, it is likely that differences may appear.

Most of the impurities in silicon used for PV cells exist at very low concentrations. Since measurements of trace quantities are difficult, much of the progress has occurred when new and better instrumentation has become available. Over the years many review articles and books dealing with the effect of impurities in crystalline silicon have been published. The interested reader is therefore encouraged to look up some of the references for a wider treatment of the subject [29, 35–40]. The following will therefore be an attempt to briefly summarise and update the general knowledge of the subject.

The maximum solid solubilities, $X_s(\max)$, of impurities in silicon are related to the distribution coefficients at the melting point according to an empirical relation found by Fischler [41]:

$$X_s(\max) = 0.1k_0 \tag{5.38}$$

or in units of atoms/cm^3:

$$C(\max) = 5.2 \times 10^{21} k_0 \tag{5.39}$$

Although deviations have been found for nitrogen, carbon and oxygen [42], the relation may be useful.

5.6.3.1 Atoms from groups IIIA (B, Al, Ga...) or VA (N, P, As, Sb...)

These atoms act as substitutional impurities in silicon. At a site where a group VA impurity (e.g. phosphorus) has replaced a silicon atom, four *d*-electrons are bound to the Si neighbours, while the fifth electron is weakly tied to the group VA atom. The fifth electron is not completely free to move, but it is easily activated to the conduction band. Group VA atoms are therefore donor atoms.

In an analogous way, group IIIA atoms (e.g. boron) do not have enough valence electrons to satisfy the four covalent neighbour bonds. This gives rise to holes, weakly tied to the group IIIA atoms. These impurities therefore create energy levels for electrons in the forbidden gap just above the valence band edge. Group IIIA atoms are therefore acceptors.

To be able to control the level of doping, the concentration of unwanted elements from groups IIIA and VA should be well below the concentration of the doping element. That becomes a challenge when using recycled material or non-virgin polysilicon feedstock (see Section 5.5). If the concentration of phosphorus is too high in a melt for it to be doped by boron, a cross-over may appear in the solid where the material changes from a *p*-type material to an *n*-material. This will reduce the raw material utilization yield. This effect may be partly counteracted by increasing

the level of boron doping, so-called *compensation*. However, increasing the concentration of boron makes the solar cells more prone to light-induced degradation (LID), the effect showing a strong correlation with boron and oxygen concentration.[13]

Our theoretical understanding of the properties of impurities in silicon other than those from groups IIIA and VA is less well developed. However, as new instrumentation has become available, a lot of experimental results of high scientific standard have been published.

5.6.3.2 Carbon

Carbon is a common substitutional impurity. Like silicon, it has four valence electrons and is therefore electrically neutral. The carbon atom is smaller than the silicon atom and may therefore be involved in precipitation of species expanding the lattice such as silicon oxide.

The solid solubility limit is $C_S = 3.5 \times 10^{17}$ atom/cm^3 at the melting point or $C_S(T) = 4 \times 10^{24} \exp(-2.3\,\text{eV}/kT)$ atoms/cm^3, where k is the Boltzmann's constant [43, 44].

In metallurgical grade silicon (see Tables 5.2 and 5.4) carbon is present at levels above the solid solubility limit and SiC precipitates are commonly present. Coarse SiC particles present in ingots for solar cells may cause problems for wire sawing and shunting in solar cells [45]. In electronic grade (see Table 5.4) the concentration of carbon is low. But in contact with the crucible and the carbon-rich atmosphere (graphite), it is difficult to avoid contamination of the melt.

As a substitutional element, carbon diffuses rather fast [$D = 1.9 \exp(-3\,\text{eV}/kT)$ cm^2/s], but much slower than interstitial impurities.

5.6.3.3 Oxygen

The subject of oxygen atoms in silicon has been studied for many years. Since results from experiments performed under different thermal histories have varied, the discussion goes on. Oxygen atoms in solid solution are electrically inactive and predominantly enter interstitial sites. The equilibrium value of the solid solubility at the melting point is generally accepted to be about 1×10^{18} atom/cm^3 [14]. However, the results from measurements of solid solubility as a function of temperature differ markedly. The value of k_0 may therefore vary. In a work by Itoh and Nozaki [39] an important point was made that the concentration of oxygen at equilibrium at a given temperature was determined by the dominant stable oxide present at that temperature. Silica, SiO_2, appears to be the thermodynamically stable component of the Si–O system. However, SiO_2 occurs in a number of crystalline (allomorphic) phases. Although there has been some discussion of the stability of the various phases and the effect of impurities, quartz is the stable phase up to 870 °C. Above that temperature tridymite is reported [46] to be the stable phase up to temperatures above the melting point of Si. Cristobalite is the stable phase above 1470 °C and up to 1725 °C where it melts. The transformation of quartz to tridymite may take hours, even at temperatures up to 1400 °C. Jackson [30] has calculated a eutectic (Si) at 0.003 at% O that is 0.04 K below the melting point of Si. The data used on solubility of O in liquid silicon are a bit uncertain. Schei *et al*. [14] have shown literature results where the solubility of O in liquid silicon near the melting point varies from about 25 to 40 ppm(w) and have advised the reader to take Jacksons result as 'tentative'. However, using Jackson's results gives $k_0 < 1$, and may be close to 0.85.

On cooling, supersaturation is easily obtained and oxygen precipitates at a rate that depends on oxygen content, temperature, and time at temperature and nucleation sites. Small silicon oxide

[13] This is further explained in Chapter 6

particles are frequently found at grain boundaries [47, 48]. Carbon is also found to influence the precipitation of silicon oxides, possibly because the carbon atoms reduce the expansion of the lattice when silicon oxide is growing.

Oxygen has a high diffusion coefficient, $D = 0.13 \exp(-2.53/kT)$ [49]. By heat treatments, the distribution of oxide particles may change since particles may get dissolved when heated and grow when cooled. Oxygen is found to change the effect of other impurities, a process known as internal gettering. This is a well-known effect, being used in the integrated circuit manufacturing industry where near-surface properties are important.

One important oxygen source is the crucible. Silicon is normally melted in high-purity fused quartz (SiO_2) crucibles (single crystals) or quartz crucibles coated with high-purity silicon nitride (Si_3N_4) (multicrystalline ingots). If holes appear in the coating, quartz will easily dissolve in the melt and raise the level of oxygen in the solid.

As mentioned above in Section 5.6.3.1 oxygen may effect light-induced degradation (LID) in combination with boron.

The variety of effects observed with oxygen in silicon will certainly stimulate further research for a long time.

5.6.3.4 Transition metals

Properties of transition metal impurities in silicon have been reported in a large number of articles, reviews and books [35–37, 40, 50–54]. The transition metals are presented by the symbols 3d, 4d and 5d, which specify the outer electron configuration of a neutral atom. Most of the metals forming deep energy levels ('midway' between the conduction band and the valence band) in silicon belong to this group and have therefore a large influence on the lifetime of minority carriers in silicon.

The main impurities found in silicon belong to the 3d transition metals (Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu). In general, these are present as interstitial impurities.

The *diffusivity* increases with increasing atomic number in the 3d row (see Figure 5.8), with Ni and Cu having the largest diffusion coefficients known in silicon.

The concentration of vacancies is very low in silicon. The high diffusion coefficients of 3d elements can therefore only be explained by interstitial diffusion mechanisms independent of vacancies.

Solid solubility diagrams for 3d transition elements are shown in Figures 5.9 and 5.10. The metals have a steep temperature-dependent solid solution limit that makes them easily supersaturated during cooling. Therefore, they frequently form complexes/precipitates at dislocations, grain boundaries or other crystal defects. It is worth noting the retrograde behaviour which may have a great influence on the development of microstructures during casting of ingots [55].

The solubility of 3d elements at room temperature is very low. However, some 3d elements are mobile, even at room temperature. So the atoms with the highest mobility (Co, Ni, Cu) will therefore be out of the solid solution during or just after the cooling. The 3d metal atoms with a low diffusivity may remain in the solid solution at interstitial sites for a much longer time after cooling. This may depend upon crystal perfection, which determines the diffusion length to reach sinks such as dislocations, grain boundaries and precipitates.

The solid solubility in silicon of the 3d elements (transition metals) is reported in Figure 5.10.

A severe deterioration of the electrical properties is expected from solid solution impurities capturing the charge carriers. The low injection level minority-carrier lifetime, τ_0, is inversely

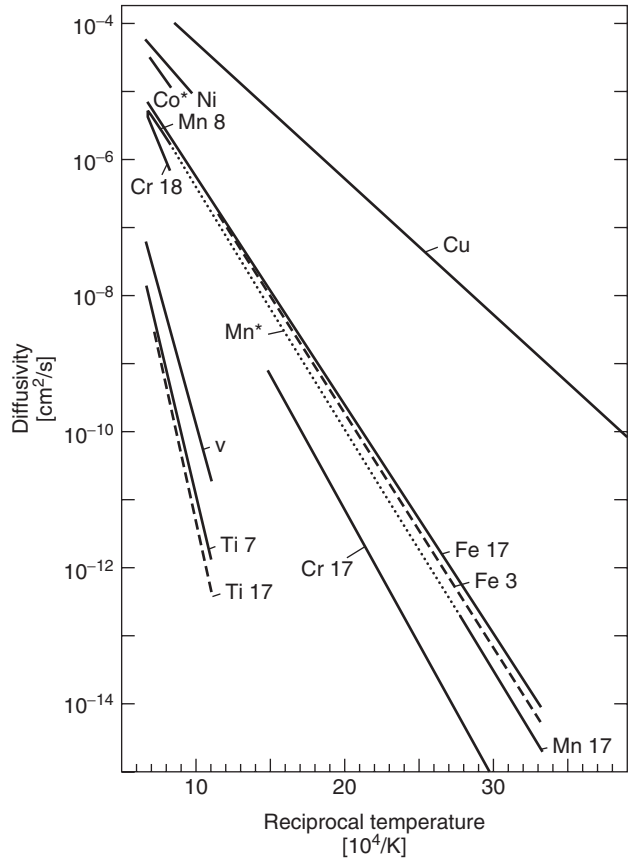


Figure 5.8 Diffusivities of 3d transition elements. Reproduced from *Metal Impurities in Silicon-Device Fabrication*, Graff K, 29, 2000, © Springer-Verlag GmbH & Co. KG

proportional to the impurity concentration, N (1/cm^3) [35]:

$$\tau_0 = (\sigma v N)^{-1} \tag{5.40}$$

where v is now the thermal velocity, which is the average speed of the electrons as they randomly collide with atoms, impurities or other defects, and σ, having the units of cm^2, represents the impurity atoms effective cross-section for the capture of a minority carrier.

Here the carrier capture cross-section for electrons, σ_e (cm^2), must be inserted for p-type silicon and σ_h for holes in n-type silicon.

The capture cross-sections for different transition metals can differ by several orders of magnitude. As a consequence, the carrier lifetime of a silicon sample can even be determined by an impurity of minor concentration if this is a 'lifetime killer' with a high minority-carrier capture cross-section. Therefore, the tolerable impurity concentration for acceptable lifetime values depends upon the chemical nature of the respective impurity, and its carrier capture cross-section for electrons in p-type silicon and for holes in n-type silicon. Both parameters can differ by orders of magnitude, and consequently the acceptable concentration for a defined impurity can be quite different in p- and n-type silicon (Figures 5.11 and 5.12; [37]).

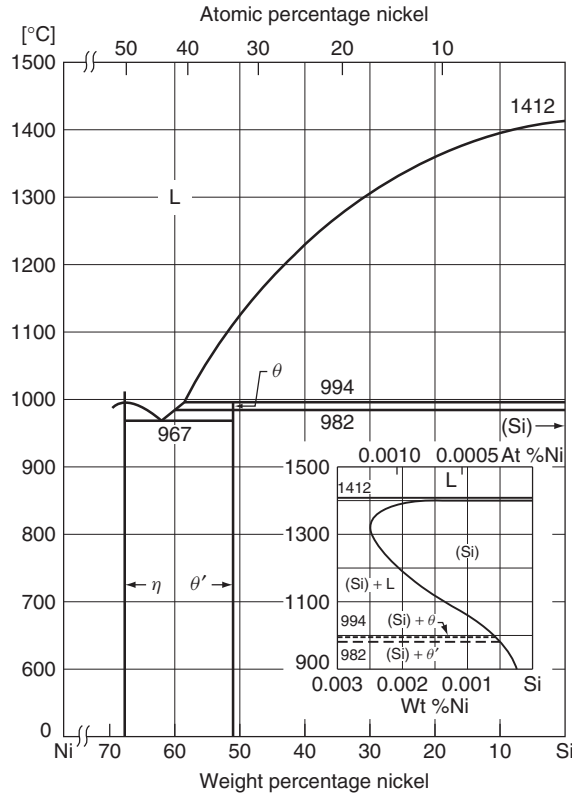


Figure 5.9 The Ni-Si phase diagram. The shape of the solidus line on the Si side is shown in the insert. It is seen that the solid solubility of Ni in Si is low, that it increases at temperatures above the peritectic point (at 994 °C) and reaches a maximum at about 1300 °C *(retrograde solubility)*. Redrawn from ASM Handbook, Vol. 8, *Metallography, Structures and Phase Diagrams* 8th Edition, ASM International, Materials Park, Ni-Si Phase Diagram, p. 325

Copper and nickel have high diffusivities and low capture cross-sections. These elements will rapidly enter a low solid solution level after cooling to room temperature and may therefore be expected to have less effect on lifetime than elements with lower diffusivities (Fe, Ti) and high capture cross-sections.

5.6.3.5 Precipitates

Except for copper, which forms Cu_3Si particles when cooled, the other 3d metals form $MeSi_2$ precipitates (Me designating a metal). Investigations of crystallographic structure, morphology and composition of several transition metal precipitates in silicon identify crystalline silicides, $FeSi_2$, $CoSi_2$, $NiSi_2$, as compounds that influence the minority-carrier diffusion length. The morphology and density depend upon temperature and time at temperature. The precipitates are mainly present at grain boundaries and dislocations [48, 51–54] Elements having a high diffusivity, such as Ni and Cu, precipitate easily. This reduces the number of atoms in solid solution and may therefore change the electric properties. Kittler and coworkers [38] studied the precipitation and coarsening of $NiSi_2$ and the effect of precipitates upon the minority-carrier diffusion length in *n*-Si single crystals (float zone FZ material). Scanning electron microscopy (SEM) using the method of electron

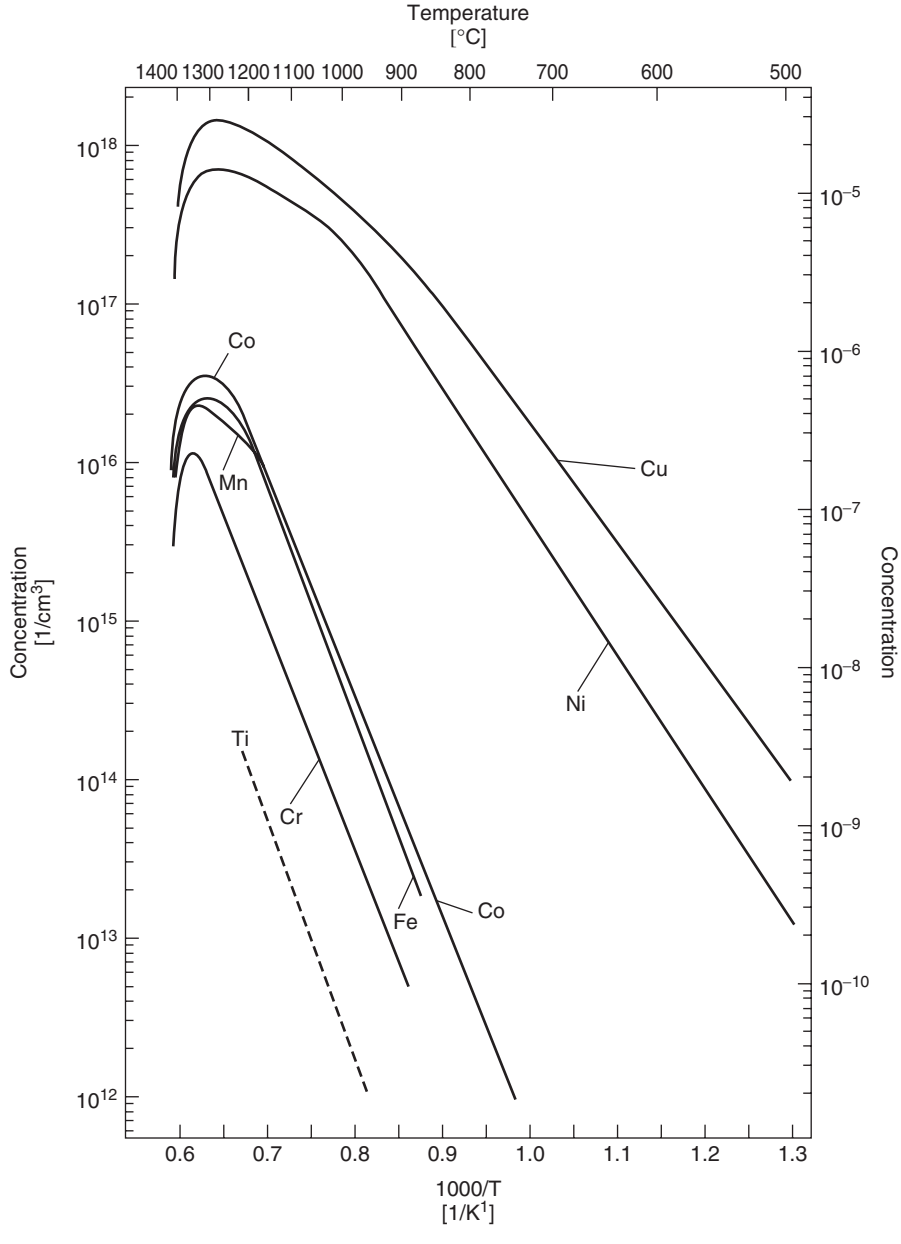


Figure 5.10 Solid solubility of 3d transition elements in silicon. Sc is still missing, but is assumed to be a little below the values of Ti. Adapted from *J. Appl. Phys*. Weber E, **A30**, 1–22 1983, © Springer-Verlag GmbH & Co. KG. The values for Ti are taken from [56]

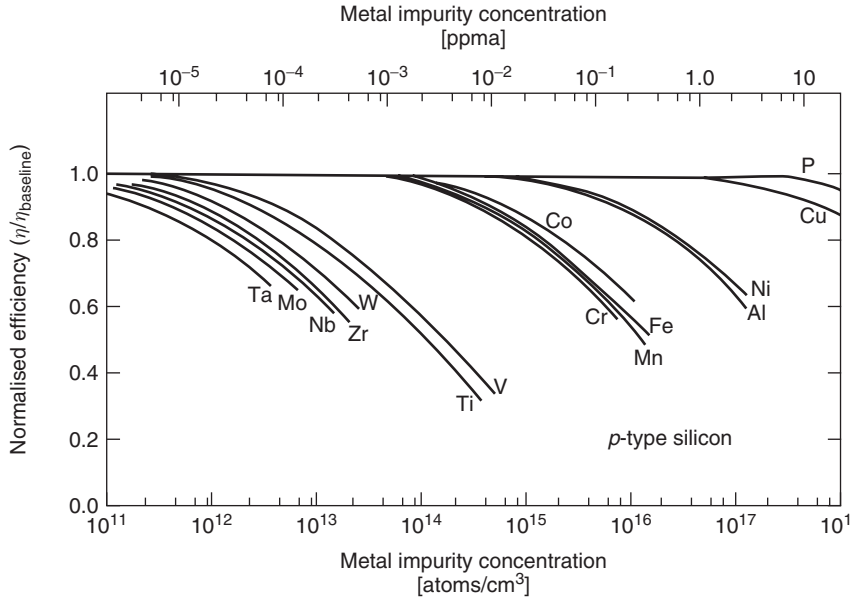


Figure 5.11 Solar cell efficiency versus impurity concentration for 4 Ω cm *p*-base devices. Reproduced from Davis Jr. J *et al., IEEE Trans. Electron Devices* © 1980 IEEE [37]

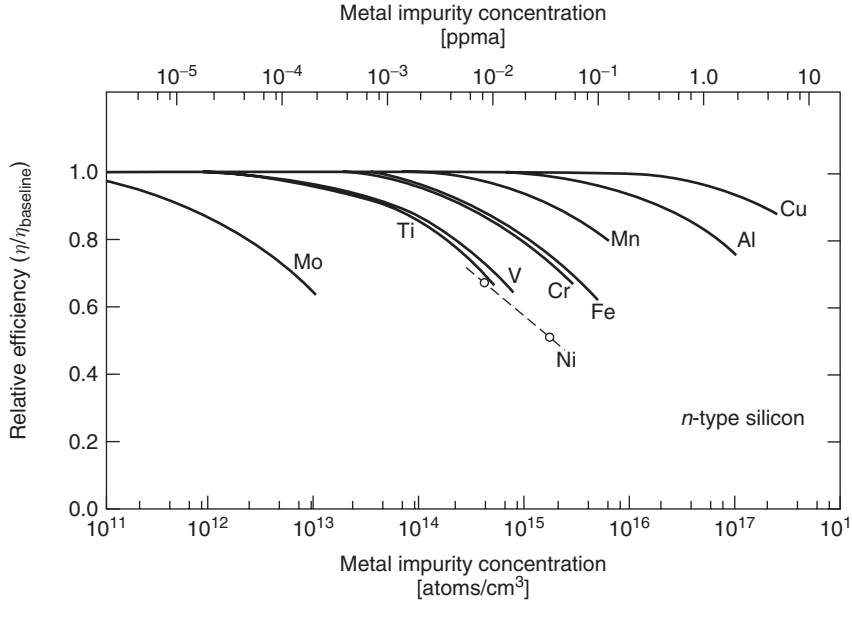


Figure 5.12 Solar cell efficiency versus impurity concentration for 1.5 Ω cm *n*-base devices. Reproduced from Davis Jr. J *et al., IEEE Trans. Electron Devices* © 1980 IEEE [37]

beam-induced current (EBIC) revealed that $NiSi_2$ precipitates were efficient recombination centres and that the minority-carrier diffusion length, L_D, was related to the precipitate density, N_P, by

$$L_D = 0.7 \cdot N_P^{-1/3} \tag{5.41}$$

This relationship revealed that the diffusion length depends only on the density of precipitates and not on the concentration of impurities. Therefore, a suitable temperature processing can increase the diffusion length L_D. This occurs if during this process large precipitates grow at the expense of the smaller ones, increasing the free distance between the precipitates. This ripening process of precipitates is observed repeatedly during heat treatments.

5.7 ROUTES TO SOLAR GRADE SILICON

Until now in this chapter we have presented the two commercial classes of silicon, their production methods, their usage and their economics. We have seen that their respective ranges of impurity do not overlap and that the gap between both ranges is quite large (several orders of magnitude). Furthermore, experimental data, theoretical considerations and modelling lead to the conclusion that the level of purity required to make efficient solar cells for most elements falls in this gap. Combining this scientific approach with empirical knowledge from commercial praxis we see now tentative initiatives to establish standard specifications and recommendations on raw material mix and technology strategy. Table 5.10 reports such tentative initiative by the European research consortium *Crystal Clear* at its workshop in late 2008 [27].

Table 5.10 Tentative solar grade silicon chemical specification. All data are in ppm(w) except numbers followed by (a), indicating ppm(a) [27]

	Feedstock A	Feedstock B	Feedstock C
B	0.05	0.45	1.5
P	0.1(a)	0.6	4(a)
Al	0.05(a)	5(a)	5(a)
Fe	0.05	5	
Cu	0.01	1	All four Fe, Cu, Ni,
Ni	0.01	1	Cr: 5
Cr	0.05	1	
Ti(*)	0.005	0.05	0.05
Na	0.01(a)	0.01(a)	
K	0.01(a)	0.01(a)	All Na, K: 0.01(a)
Zn		2	
Ca			
C	5	30 (multi) 1 (mono)	
O	5 (multi), 1 (mono)	20 (multi)	
Donnor/acceptor Compensation in mix to ingot	No	Yes	Yes
Extra defect engineering in cell process	No	No	Yes

These technical guidelines beside cost and volume constraints are likely to be useful in designing new process routes specific to solar grade silicon. Such an attempt is not a new challenge. It was first seriously and massively addressed in the mid-1970s after the first international oil crisis. Considerable R&D efforts were made in the years 1975–1985, particularly in the United States under the guidance of the Department of Energy (DoE) and Japan (NEDO). During and at the end of this productive period, numerous specific articles, exhaustive review studies and symposium proceedings devoted to these topics were published [57–59]. With the severe and persistent shortage of silicon occurring in 2003–2004 most of these historic studies are being revisited by academic and industrial researchers. The following gives a summary and update, although not exhaustive, of these efforts. For the sake of simplification all approaches are classified within four subgroups.

1. The first subcategory deals with the processes going through reduction/decomposition of a purified volatile silicon compound. That includes the simplification of the existing polysilicon processes. Other volatile compounds than those currently used and other ways to reduce silicon are also envisaged in this subgroup. The challenge here is to achieve sufficient output at acceptable cost.
2. Upgrading the purity of metallurgical silicon is a technique with the potential to produce millions of tons of silicon as we have demonstrated earlier in this chapter. This is reviewed in the second subcategory. The goal of and the challenge for this subgroup are to achieve adequate purity at acceptable cost.
3. Other methods, e.g. electrolysis, metallothermic reduction of silicon are summarised in a third category of methods.
4. The two latter subcategories of processes may hardly reach the suitable metallic purity and in most cases a directional solidification step must be inserted in the overall process. Crystallisation is known to be an effective purification technique. Crystallisation is also a necessary step to produce the silicon ingot and wafers. The question is where the crystallisation steps in the manufacturing chain to the solar cell can advantageously be incorporated. This is reviewed in the fourth category of methods.

5.7.1 Further Polysilicon Process Development and New Processes Involving Volatile Silicon Compounds

The Union Carbide and Ethyl Corporation polysilicon processes (Sections 5.4.2 and 5.4.3) resulted from attempts to make the Siemens process more economical. Project goals were to meet a price target of 10 US$/kg, set in 1975 by the US Department of Energy (DoE). The Siemens process was fully developed around 1960, but polysilicon for electronic devices and other purposes had been produced since the 1940s. The Siemens process was the first design of a rational industrial process, which gained international recognition through rapid and broad licensing. At least a dozen processes had been developed prior to the Siemens process and coexisted with the Siemens process until the beginning of the 1970s. Purity and high production rates based on semiconductor industry demand were the primary criteria for the process design. Capital expenditure and energy consumption were considered secondary criteria. While large production quantities of ultra-pure polysilicon were available at acceptable prices for semiconductors, this price was too high for the development of low-cost solar systems. For solar cells, a process with reduced costs, reduced energy consumption and increased production rates, with purity levels not as important, is needed. Since several processes had been developed in the past, a review and re-evaluation of the silicon chemistry of these processes were suggested (see Tables 5.11 and 5.12).

To make the process cost effective, the economics of both the production of the volatile silicon compound and its decomposition or reduction to elemental silicon had to be reviewed.

Table 5.11 Historical processes to manufacture polysilicon (according to Strategies Unlimited [60])

Companies	Volatile silicon source	Reduction agent	Reactor type
Dupont	$SiCl_4$	Zinc	Inside-quartz tube
Bell labs	$SiCl_4$	Hydrogen	Ta-filament
Union Carbide	$SiHCl_3$	Hydrogen	Inside-quartz tube
Int'l Telephone	$SiCl_4$	Sodium hydride NaH	Ta-filament
Mallinckrodt	SiI_4	Hydrogen	Inside-quartz tube
Transitron	SiH_4	Decomposition	Inside-quartz tube
Texas Instruments	$SiCl_4$	Hydrogen	Outside-quartz tube
Foot Mineral	SiI_4	Decomposition	Si-filament
Chisso	$SiCl_4$	Hydrogen	Inside-quartz tube
Siemens	$SiHCl_3$	Hydrogen	Si-filament
Komatsu	$SiH_4/SiHCl_3$	Decomposition	Si-filament
Motorola	$SiHCl_3$	Hydrogen	W-filament
Phoenix Materials	$SiCl_4$	Hydrogen	Si-filament
Texas Instruments	$SiHCl_3$	Hydrogen	Fluidised bed spheres

Table 5.12 Polysilicon research projects stimulated by the objectives of low-cost solar cells programs after 1975 [60]

Companies/groups	Volatile silicon source	Reduction agent	Reactor type
Aerochem Res. Lab.	$SiCl_4$	Sodium	Free space
Eagle Picher/General Atomic/Allied	SiH_4	Decomposition	Fluidised bed
Battelle	$SiCl_4$	Zinc	Fluidised bed
Hemlock	SiH_2Cl_2	Hydrogen	Si-filament/rod
Union Carbide	SiH_4	Decomposition	Free space
Union Carbide	SiH_4	Decomposition	Fluidised bed
Ethyl Corp	SiH_4	Decomposition	Fluidised bed
Motorola	Si_nF_{2n+2}	Decomposition	
NEDO	$SiHCl_3$	Hydrogen	Fluidised bed
Rhône–Poulenc	SiH_4	Decomposition	
Schumacher	$SiHBr_3$	Hydrogen	Fluidised bed
SRI International	SiF_4	Sodium	Solid separation
Westinghouse	$SiCl_4$	Sodium	Free space
Bayer	$SiCl_4$	Aluminium	Melt
Bayer	SiH_4	Decomposition	Fluidised bed
Wacker	$SiHCl_3$	Hydrogen	Fluidised bed

Compared with the historical processes, the research projects did not discover radically new silicon compounds. Apparently silicon halides $SiX_{4-n}H_n$ could not be avoided. However, innovative methods were discovered to produce these silicon compounds. Metallurgical grade silicon remained in most cases the starting point, but silicon fluoride, obtained as co-product from phosphates leaching (fertilisers) and direct chlorination of natural silica to produce tetrachlorosilane, were envisaged as serious challenges. The research abandoned completely the concept of heated filament or seed rod

that was clearly perceived as too expensive. Texas Instruments [57] had demonstrated the benefit of the fluidised bed technology, that is:

- larger throughput
- lower energy consumption
- continuous operation
- lower capital expenditure.

Although the quality was not acceptable for making microelectronics devices, this was not a drawback for the development of the solar cell market.

In addition to the fluidised bed, a free space reactor implying spontaneous seedless formation of solid silicon particles through homogeneous decomposition of silane has also been envisaged among others by Union Carbide [57].

We emphasise again that the list in Tables 5.11 and 5.12 is not exhaustive, but gives a good survey on envisaged and partially explored routes. The silicon shortage crisis of 2003–2004 contributed to revitalise and accelerate a multitude of research programs on this topic. Surprisingly most, if not all, of these 'novel' concepts are more or less revisiting those listed in the historic Tables 5.11 and 5.12. Because of higher demand on volumes, side technology development (e.g. in crystallisation techniques) and also more relaxed specifications, several of these concepts are perceived more viable today than thirty years ago. They attract renewed attention from researchers, industry, and investors. Currently we see the main developments as follows.

The long-established polysilicon companies (at least the four majors) are expanding their capacities based on existing proprietary technologies, adding some process modifications outside the core technology to reduce cost and taking advantage of the relaxed specification of photovoltaic as compared with semiconductor (higher growth rate, simpler crushing, no post-treatment and less product analysis).

In parallel with these expansion programs these companies are developing and implementing new deposition technologies to reduce energy consumption. Most successful so far is the fluidised bed technology, long practised by MEMC (Ethyl Corporation process; see Section 5.4.3), under implementation at REC Silicon's new plant (USA), under development at Wacker (Germany) and possibly also at Hemlock (USA). Tokuyama (Japan) is developing the vapour to liquid deposition (VLD) reactor, in which trichlorosilane is decomposed and deposited on a liquid film of silicon. The advantages of this method are higher deposition rate, higher yield of deposition and less by-products. The disadvantages are the higher operating temperature and contamination of silicon by the reactor lining (carbon).

New entrants are purchasing turnkey projects based on classical Siemens technology, but incorporating by-products recycling and process simplification allowed by the relaxed specifications of solar grade silicon. Volume and time to market are their first priority. Developing new and less costly processes is a second priority, kept in mind for the next wave of expansion at the new entrants.

Beside the producers we see a long list of candidates aspiring to enter the market with new technologies.

Joint Solar Silicon, a joint venture between Solar World (successor of Bayer Solar) and Evonik (a rename of Degussa) is developing the free space reactor using monosilane to generate silicon powder which can then be mechanically compressed to pellets. According to the company the technology is undergoing up-scaling at the Evonik site of Rheinfelden (Germany).

New fluidised bed reactor concepts are subject to studies with monosilane (SiH_4), trichlorosilane ($SiHCl_3$), tribromosilane ($SiHBr_3$) and iodosilane.

The reduction of tetrachlorosilane ($SiCl_4$) by metals is retained by at least two companies in Europe and Japan, both using zinc (Zn). The challenge is to close the material loop of both chlorine and zinc.

The fairly large abundance of cheap fluorosilicates co-produced with fertiliser phosphates still seems attractive as a source of volatile silicon tetrafluoride (SiF_4), which after purification can directly be reduced to polysilicon by a metal, e.g. sodium (Na) or via monosilane as in the Ethyl Corporation method (see Section 5.4.3).

All these developments are promising to secure significant volumes of pure silicon for the manufacture of the more demanding high-efficiency crystalline cells and for the deployment of the solar electricity market.

5.7.2 Upgrading Purity of the Metallurgical Silicon Route

The processes used in the elaboration of metallurgical grade silicon for the aluminium and chemical industries proved that most of the metallic impurities could, by repetition and combination of methods, be decreased to a rather low level, for example 100 ppm(w) (see Table 5.13). Since metals could be further decreased by a post-directional solidification, it was rational to explore the full potential of this expected low-cost route.

5.7.2.1 Use of pure raw materials and pure lining

The use of pure linings in the furnaces and the intermediate vessels containing the liquid silicon is a first approach. It was investigated by several companies and research consortia such as Dow-Corning, Elkem/Exxon, Siemens, NEDO, to mention the most accomplished [61–68]. Quartz or silica sands selected for their original purity were chemically treated by leaching or precipitated from waterglass solutions. Purified carbon black was used as the reduction material in most of these approaches. The morphology of sand or precipitated silica and carbon black are quite different from, respectively, lumpy crystalline quartz and bulk carbon, charcoal or woodchips currently used in the carbothermic reduction. New technologies, in particular, to agglomerate the powder-like raw materials had to be developed. Pellets consolidated by means of organic non-contaminating binders (e.g. sucrose) or rice hulls, perceived as potential high-volume, low-cost raw materials since they contain both silica and carbon, are worth mentioning. Extruded pellets of cocked rice hulls, with sucrose as a binder, appeared to be reactive in the submerged arc furnace technology. Using these special raw materials, boron and phosphorus were found in the range of 1–4 ppm(w). In standard metallurgical grade, B and P range between 7 and 50 ppm(w), typical average values being around 25 ppm. The boron level achieved with rice hulls could be as low as 1 ppm(w). However, phosphorus was as high as 40 ppm(w) and would require an additional and specific treatment to make use of this source of silicon in solar cells.

Table 5.13 Typical best results achieved with the carbothermic reduction using high-purity raw materials at the end of the eighties

Impurity	B	P	Al	Fe	Ti
Minimum content ppm(w)	2	1	100	100	10
Maximum content ppm(w)	4	3	300	200	20

5.7.2.2 Post-treatment by chemical leaching

Metallic impurities segregate at the grain boundaries, forming intermetallic phases with silicon, basically made of silicides and silicates. Hydrometallurgical upgrading of silicon by acid treatment and leaching has a long history. The *Silgrain* process, a trademark of Elkem AS, is currently (in 2008) producing around 30 000 metric tons of metallurgical grade silicon by refining ferrosilicon (90% Si) in an acidic liquor of ferric chloride. The process is particularly efficient in removing iron and transition elements because these have very low segregation coefficients in silicon. The dissolution of iron and the transition elements in the leaching liquor is eased and enhanced by the formation of stable intermetallic phases with Al, Ca or isostructural elements, for example, Sr, Ba, Ga or a lanthanide. The process has been further developed during the past 20 to 30 years with solar grade silicon as an objective [69–71].

Surface treatment or leaching by hydrofluoric acid or by HCl/HF mixtures is also well known to wash out residual impurities present as oxides and silicates. Grinding silicon prior to leaching to increase the surface exposure is a way to enhance the purification [72, 73]. Acid leaching has therefore been envisaged by several groups and companies, for example Wacker and Elkem, as a post-treatment purification after the carbothermic reduction [69–73].

However, it must be emphasised that this method cannot in practice refine silicon below the dissolution limit characteristic of each element at the solidification composition. For the major impurities Fe, Ca and Al present in metallurgical grade silicon the limit of dissolution, as obtained from classical refining in a ladle, exceeds by far the ppm level. Effective post-treatment leaching of metallurgical grade silicon has made it possible to achieve less than 10 ppm(w) for these major elements and less than 1 ppm(w) for minor transition elements. One pass of directional solidification was then envisaged as sufficient to achieve the requested purity for solar cells.

The acid treatment or leaching methods is not effective in removing impurities such as boron, carbon and oxygen. Adding Ca to the silicon alloy prior to acid treatment, however, proved that P could be reduced by a factor of 5 down to a concentration less than 5 ppm(w), probably because P is dissolved in calciumsilicide [70, 71]. Adding barium to the alloy also proved some effect on boron removal by acid leaching [72, 73].

5.7.2.3 Post-treatment by extraction metallurgy in ladle

Post-treatments of liquid silicon are common practices in refining metallurgical grade silicon for applications in the aluminium and chemical industries. The objective is then to adjust Al, Ca and possibly C to a suitable concentration of some hundreds or thousands of ppm(w). In that respect the reader may with great benefit consult the comprehensive handbook by Schei *et al*. [14]. As described above, crystallisation and leaching are efficient means of removing chemical elements with high ability to segregate from liquid to solid silicon, that is, Fe, and most of the metallic transition elements. Extraction metallurgy from a liquid phase of silicon, either liquid–liquid, liquid–solid or liquid–gas, has received considerable attention in order to remove the critical elements P, B and C. When silicon is kept liquid, it is possible to displace the equilibrium between both phases present and gradually remove the unsuitable impurity, by continuously extracting it.

Impurity (liquid silicon) = Impurity (liquid slag)	*Ksi/ls*
Impurity (liquid silicon) = Impurity (solid slag)	*Ksi/ss*
Impurity (liquid silicon) = Impurity (gas)	*Ksi/g*

Particular attention has been paid to boron and phosphorus because these elements are the major *p*- and *n*-type dopants of silicon and because they coexist in metallurgical grade silicon in concentration one or two orders of magnitude too high for solar cell applications.

Boron has nearly the same affinity towards oxygen as silicon. Boron forms gaseous suboxides, BO being analogous to SiO, and its stable oxide B_2O_3 behaving similarly to SiO_2 in the presence of alkaline earths at slag-forming temperatures. Therefore, there are good reasons to expect the removal of boron as an oxide constituent of the extracting slag or as a gaseous suboxide at elevated temperature. Both theoretical possibilities have been experimentally verified.

Since the work published by Theuerer in 1956 [74], it has been known that liquid silicon becomes purified with respect to boron when brought in contact with a gas mixture of $Ar–H_2–H_2O$. The sole role of H_2 and H_2O assisting the extraction has been emphasised by several authors such as Khattak *et al*. [75–78], whereas Amouroux, Morvan *et al*. of the University of Paris [79–82] and the Japanese group at Kawasaki Steel/NEDO [83–87] have underlined the benefit of using an oxidative plasma in the presence of moisture and hydrogen. Amouroux, Morvan *et al*. have also shown that boron elimination was enhanced when fluoride, for instance in the form of CaF_2, was injected into the plasma gas.

Experiments in removing B by slag extraction have been done by several companies and groups, including Kema Nord, Wacker, Elkem and NEDO/Kawasaki. Schei [88] has described a counter-flow solid–liquid reactor to remove boron by fractional extraction in a semi-continuous process in a patent assigned to Elkem AS of Norway.

It has been demonstrated that phosphorus could be evaporated from a silicon melt under vacuum conditions [89, 90]. Miki *et al*. [91] have explained the thermodynamics of this process by reactions involving mono- and diatomic phosphorus in the gas phase:

$$P_2(g) = 2\underline{P}(1\%\ \text{inSi(l)}) \tag{5.42}$$

$$P(g) = \underline{P}(1\%\ \text{inSi(l)}) \tag{5.43}$$

Underscored symbols refer to a dissolved element in liquid silicon as already defined in Section 5.3.2. With a segregation coefficient of 0.35 (see Table 5.8) P can also be refined by repeated crystallisation.

Silicon produced by carbothermic reduction is, so to speak, supersaturated in SiC when tapped from the furnace, and may contain as much as 1000–1500 ppm(w) C. As this silicon is cooled down to the solidification temperature, the majority of this carbon precipitates out as SiC particles leaving around 50 to 60 ppm(w) in liquid silicon. Carbon removal from liquid silicon is therefore a two-step operation:

1. the removal of precipitated SiC as close to the solidification temperature as possible;
2. the removal of dissolved $\underline{C}$ by oxidation to CO(g).

As already mentioned, SiC particles become effectively captured by the slag phase during oxidative refining in which the main objective may be to remove Al and Ca as industrially practised today or in a similar operation with the purpose of removing B or P. Depending on the temperature and the degree of slag/molten silicon intermixing, this treatment may give a product with 80–100 ppm(w) C. Other methods, which have been applied and proved effective at a temperature closer to the solidification temperature, are filtration, centrifugation or settling combined with slow cooling. Several studies have provided valuable contributions suggesting several methods, for example, settling in combination with directional solidification [92], filtration combined with oxidation [93], oxidative plasma [85–87] and decarburisation by inert gas purging or under vacuum [15, 93]. Klevan [15] has developed a mathematical model, which describes the kinetics

of decarburisation when inert gas purging is applied. Mechanical removal is, however, not efficient enough to affect the substituted carbon. Stronger methods able to displace the equilibria, such as oxidative plasma and vacuum vaporisation, are believed to be more powerful techniques.

It is worth noting that these types of operation can all be carried out in carbon-lined ladles. The several steps dealing with the removal of dissolved carbon $\underline{C}$ from liquid silicon Si(l), however, has to be carried out in the absence of C and at a highest possible temperature in order to optimise the equilibrium and the kinetic conditions for the reaction:

$$\underline{C} + \tfrac{1}{2}O_2 = CO(g) \tag{5.44}$$

A parallel reaction, which affects the silicon yield, unfortunately, also takes place:

$$Si(l) + \tfrac{1}{2}O_2 = SiO(g) \tag{5.45) = (5.8}$$

The value of solid solubility of carbon in silicon is approximately 10 ppm(a), corresponding to the homogeneous distribution of substituted carbon atoms on silicon lattice sites. This type of carbon impurity is detected by infrared spectroscopy. Higher concentrations of carbon result in SiC precipitates of different size and morphology. This type of carbon is detected and analysed by combustion methods or secondary ion mass spectroscopy (SIMS).

5.7.2.4 Challenges and achievements

The improved carbothermic reduction of silica, followed by leaching and directional solidification, has the capability to produce silicon with sub-ppm concentration of all metallic impurities. Controlling the nonmetallic elements, particularly B, P, C and O, at the level required for solar cell applications remains the major challenge for the metallurgical route. Unfortunately, there is no universal method for reducing these critical elements simultaneously. As a consequence, several refining steps are necessary, with the risk of reducing the silicon yield. Another risk is the recontamination by impurities from the reactants and the lining during handling and treatment of liquid silicon. Extensive studies were done on this route in the early 1980s, but efforts slowed down considerably for almost 20 years before being revitalized after 2003. The Japanese program headed by NEDO is to our knowledge the earliest accomplished project representing this route (Table 5.14; [86]). On the basis of this project's results, Kawasaki (renamed JFE Steel Corporation) has built a small commercial plant of 400 metric ton capacity, but the economical feasibility of the route long-term remains uncertain. JFE has admitted higher cost than suitable and might run the process only as long as the shortage prevails. In the NEDO process, metallurgical grade silicon is purified through:

1. melting of silicon by electron beam and evaporation under vacuum;
2. first directional solidification;
3. remelting of silicon and gas treatment ($O_2 + H_2O$) assisted by a plasma torch;
4. second directional solidification.

Table 5.14 Solar grade as obtained through upgrading metallurgical grade silicon by the NEDO method according to [86]

B [ppm(w)]	P [ppm(w)]	Al [ppm(w)]	Fe [ppm(w)]	Ti [ppm(w)]	O [ppm(w)]	C [ppm(w)]	Resistivity [ohm cm]	Lifetime [μs]
0.04–0.10	0.03–0.14	<0.01	<0.05	<0.01	<6	<5	0.8–1.2	>7.7

Other producers of silicon metal (e.g. CBCC/Dow Corning in Brazil, Silicium Becancour/Timminco Solar in Quebec, Elkem Solar in Norway, Ferroatlantica/Photosil in Spain/France, Globe Metallurgical in the USA and Fesil/Sunergy in Norway) are producing or about to start production of upgraded metallurgical silicon for solar customers, applying a succession of refining steps comprising:

1. judicious selection of raw materials to the carbothermic reduction of silica, aiming at concentration of boron and phosphorous as low as possible;
2. pyrometallurgical refining in a ladle, or a holding furnace, with or without a plasma or other similar equipment, for two or three phase extraction (liquid, solid, gas): the effect is to eliminate the elements less noble than silicon; these steps are also designed to reduce the most difficult boron and in some cases phosphorous as well;
3. directional solidification to reduce metallic impurities.

Process details are available only through patent publications and it is of course uncertain which patent details are effectively applied in the industrial processes. The producers remain very secretive about their recipes and about the accurate level of purity achieved. We think that the performances so far are in accordance with the specifications given in Table 5.10 (feedstocks B and C). However, these methods have the potential to reach the same level of purity as indicated in Table 5.14, the question remains at what cost.

5.7.3 Other Methods

Besides the large avenues described above in Sections 5.7.1 and 5.7.2, one finds some other routes. The following processes are worth noting:

1. The aluminothermic reduction of silica

$$3SiO_2 + 4Al = 3Si + 2Al_2O_3 \tag{5.46}$$

   explored by Wacker [58]. We have not noticed new recent interest for this process.

2. The electrolytic transfer of silicon from an anode made partly of metallurgical grade silicon alloyed with Cu in (Cu_3Si-Si) to a graphite cathode through a liquid electrolyte made of $KF:LiF:SiK_2F_6$ as described by Olson *et al*. [94–96].

$$Si(\text{metallurgical grade}) \rightarrow Si(4+) + 4e(-) \rightarrow Si(\text{refined}) \tag{5.47}$$

The electrochemical methods have also been revisited by several academic groups and companies, in our opinion none of them being close to industrialisation.

5.7.4 Crystallisation

As mentioned in the introduction to Section 5.7, none of the processes in Sections 5.7.2 or 5.7.3 have the capability to reach a suitable metal concentration without inserting a directional solidification, i.e. a crystallisation step. The efficiency of the crystallisation processes may be predicted by the segregation or the distribution coefficient of each impurity element (see Section 5.6.1) between the solid and the liquid silicon phases. Published data (see Table 5.8) clearly show that the elements belonging to group IIIA (B, Al, Ga) and group VA (P, As) have such distribution coefficient values making difficult to separate them from the silicon. These elements are the

electronic dopants for silicon and their concentration must be closely controlled during all steps of the manufacturing process. Their chemical and physical behaviour in interaction with silicon are two aspects of their close neighbourhood to this element. Crystallisation as such cannot be satisfactory without prior treatments, specifically to remove these elements, particularly P and B.

Crystallisation from a silicon-melt. Various methods have been claimed to be useful for refining silicon by directional solidification, resulting in large oriented crystals. All these various crystallisation methods are also used in manufacturing the silicon photovoltaic devices, and they are described in Chapter 6 of this handbook. Other valuable references are to be found in [58, 59].

Crystallisation from an aluminium melt. Silicon forms a unique eutectic phase at the concentration of 12.6% (w) in aluminium. Several companies (Union Carbide, Alcoa and Wacker) have tried to take advantage of this property in order to purify silicon. By cooling a hypereutectic composition, pure silicon phases crystallise and precipitate from the aluminium melt. After solidification of the hypereutectic alloy, the aluminium matrix can be dissolved by leaching, releasing the pure silicon crystals [58, 97–100]. Published data indicated that the aluminium content in the resulting silicon was as high as 300 ppm(w). The segregation coefficient of aluminium in silicon limits further separation. A further reduction of aluminium, by directional solidification of the silicon thus obtained, adds significant cost just for this element. Other metallurgical routes to eliminate Al at 1 ppm(w) level or less need to go through silicon melting and subsequent extraction with the double risk of recontamination and heavy material losses. We have noticed a renewed interest for this method.

5.8 CONCLUSIONS

At the time the second edition of the handbook is written (2008), silicon is the dominant PV material with nearly 95% of cells and modules based on it. Crystalline technologies account for nearly 90%. Although new thin film materials are coming, we continue predicting that silicon will remain a predominant material in the foreseeable future, at least until the third edition of the handbook. The crystalline technologies require significant consumption of pure silicon. The industry first selected silicon raw materials among the residuals from single-crystal pulling and second-grade polysilicon. When the first edition of the handbook was edited (2003), these sources appeared not sufficient to cover the rapidly growing demand from the industry. Polysilicon produced specifically for solar applications and prime-grade polysilicon started to be extensively used. This had the negative consequence of increasing the cost of solar cells, but it apparently did not slow down the extraordinary growth of the PV industry during the last decade (near 50% annually in average). In 2003–2004 a persistent shortage of silicon feedstock arose. This stimulated capacity expansions at the producers as well as a multitude of research programs on new technologies, partly introducing new ideas, but mainly revitalising old concepts formulated 30–35 years ago after the two first oil crises. Two main directions still retain most of the attention of scientists and technologists: the simplification of chemical purification routes such as the polysilicon processes and the further development of pyrometallurgical treatments of liquid silicon. The crystallisation processes applied to silicon to produce the PV devices play a role in purifying and shaping silicon and should affect to a larger extent the development and the final definition of solar grade silicon. Capital cost is perceived as a significant part of the overall cost and adds to the barriers that need to be overcome. Quality and cost requirements have imposed strict limits to solar grade silicon. However, the technical requirements or specifications to silicon raw material are still not well defined. With the shortage of feedstock and the emergence of new solar silicon processes more effort is nowadays put into understanding the role of individual, as well as classes of impurities, through experimental and theoretical studies. We have no doubt that the period to come until the third edition of the

handbook will bring considerable and certainly surprising new results on this fundamental topic, along with the commercial maturation of new emerging raw material and cell processes.

REFERENCES

1. Hirshman W P, Hering G, Schmela M *Photon International*. **3**, 140–174 (2008).
2. Murphy G, Brown R, *Silicon in Bulletin 675 Mineral Facts and Problems*, Bureau of Mines, US Department of Interior, Washington DC (1985).
3. Kerkhove D, *Silicon Production Technology*, Technische Universiteit Delft, Netherlands (1994).
4. Forwald K, *Dissertation NTNU*, Norway, MI-47 (1997).
5. Dosaj V, *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th edn, Vol. **21**, 1104–1122, John Wiley & Sons, Inc., New York (1997).
6. *Properties of Crystalline Silicon*, EMIS Datareviews series No. 20. INSPEC, Robert Hull, University of Virginia, USA.
7. Moretto H, Schulze M, Wagner G, *Ullmann's Encyclopaedia Ind. Chem*. **A24**, 57–93 (1993).
8. Rong H, *Dissertation NTNU*, Norway, IUK-67 (1992).
9. Sørheim H, *Dissertation NTNU*, Norway, IUK-74 (1994).
10. Alsema E, de Wild-Scholten M, *Proc.22nd European Photov. Solar Energy Conf*., Milan, Italy (2007).
11. Rangei M, Frankl P, Alsema E, de Wild-Scholten M, Fthenakis U, Kim H, *Proceedings AIST Symposium Expectation and Advanced Technologies in Renewable Energy*, Chiba, Japan (2007).
12. Alsema E, de Wild-Scholten M, *Proceedings MRS Fall Meeting*, Boston, USA (2007).
13. Watt M, *Added Values of Photovoltaic Power Systems*, Report IEA-PVPS T1-09 (2001).
14. Schei A, Tuset J, Tveit H, *Production of High Silicon Alloys*, Tapir forlag, Norway (1998).
15. Klevan O, *Dissertation NTNU*, Norway, MI-167 (1997).
16. Anglezio J, Servant C, *J. Mater. Res*. **5**, 1894–1899 (1990).
17. Forwald K, Schüssler G, *Proc. Silicon for the Chemical Industry I*, pp 39–46, Norway (1992).
18. Nygaard L, Brekken H, *Proc. Silicon for the Chemical Industry II*, pp 61–67, Norway (1994).
19. Brekken H, Nygaard L, Andresen B, *Proc. Silicon for Chemical Industry III*, pp 33–45 Norway (1996).
20. Pachaly B, *Proc. Silicon for Chemical Industry II*, pp 55–60, Norway (1994).
21. Forwald K, Soerli Oe, Schüssler G, EP 0 372 918 B1 (1989).
22. Schulze M, Licht E, *Proc. Silicon for the Chemical Industry I*, pp 131, Norway (1992).
23. Aas H, *The Met. Soc. AIME, TMS Paper Selection* **A71-47**, 651–667 (1971).
24. Boardwine C *et al*., Progress in Organic Chemistry, *Proc. Int. Symp. on Organosilicon Chemistry X (1993)*, pp 555–569, Gordon & Breach Science Publishers, Amsterdam (1995).
25. De Linde J, 2207 *Int. Forum Annual Conference of Chinese Industry*, Shenyang (2007).
26. Photon International, *5th Silicon Conference*, Munich (2008).
27. Crystal Clear, *Workshop on Arriving at Well Founded SOG Silicon Feedstock Specifications*, Amsterdam (2008).
28. Kobayashi S, in Hull R (ed.), *Properties of Crystalline Silicon*, pp 6–22, University of Virginia, USA (1999).
29. Trumbore F, *Bell Syst. Tech. J. (USA)* **39**, 205 (1960).
30. Jackson K A, *Bulletin on Alloy Phase Diagrams*, **9**(5), 548–549 (1988).
31. Burton J, Kolb E, Slichter W, Struthers J, *J. Chem. Phys*. **21**, 1991–1996 (1953).
32. Kodera H, *Japan J. Appl. Phys*. **2**, 212–219 (1963).
33. Kvande R,. Geerligs L J, Coletti G, Arnberg L, Di Sabatino M, Øvrelid E, Swanson C C, *J.Appl.Phys*. **104**, 064905 (2008).
34. Stokkan G, Riepe S, Lohne O, Warta W, *J. Appl. Physics*, **101**, 053515.

35. Graff K, *Metal Impurities in Silicon-Device Fabrication*, 29, Springer, Berlin (2000).
36. Istratov A, Hieslmair H, Weber E, *Appl. Phys*. **A70**, 489–534 (2000).
37. Davis Jr. J *et al*., *IEEE Trans. Electron Devices* **Ed-27**, 677–687 (1980).
38. Kittler M, Lärz J, Seifert W, *Appl. Phys. Lett*. **58**, 911–913 (1991).
39. Itoh Y, Nozaki T, *Japan J. Appl. Phys*. **24**, 279–284 (1985).
40. Weber E, *J. Appl. Phys*. **A30**, 1–22 (1983).
41. Fischler S, *J. Appl. Phys*. **33**, 1615 (1962).
42. Jaccodine R, Pearce C, in Bullis W, Kimerling L (Eds), *Defects in Silicon*, 115–119, The Electrochemical Society, Pennington, NJ (1983).
43. Nozaki T, Yatsurgui Y, Akiyama N, Endo Y, Makida Y, Behaviour of Light Impurity Elements in the Prod. of Semicond. Silicon, *J. Radioanal. Chem*. **19**, 109–128 (1974).
44. Bean A, Newmon R, *J. Phys. Chem. Solids* **32**, 1211–1219 (1971).
45. Al Rifai M H, Breitenstein O, Rakotoniaina J P, Werner M, Kaminski A, Le Quang L, *Proc.9th European Photov.Solar Energy Conf*., Paris, June 7–11, 632–635 (2004).
46. Wriedt H A, *Bulletin of Alloy Phase Diagrams*, **11** (1), 43–61 (1990).
47. Møller H, Long L, Riedel S, Rinio M, Yang D, Werner M, *7th Workshop on The Role of Impurities and Defects in Silicon Device Process*., NREL, Vail, Colorado, 41–50 (1997).
48. Nordmark H, Di Sabatino M, Acciarri M, Libal J, Binetti S, Øvrelid E J, Walmsley J C and Holmestad R, *Proc. 33rd IEEE PhotovoltaicSpecialists.Conf*., San Diego, USA (2008).
49. Mikkelsen Jr. J, *Mater. Res. Soc. Symp. Proc. (USA)* **59**, 19 (1998).
50. Solberg J K, Nes E, *Acta Crystallogr*, **A 34**, 684 (1978).
51. Ryoo K, Drosd R and Wood W, *J.Appl.Phys*. **63**, 4440 (1988).
52. Heuer M, Buonassisi T, Istratov A A, Pickett M D, *J.Appl.Phys*. **101**, 123510 (2007).
53. Rakotoniaina J P, Breitenstein O, Werner M, Al Rifai M H, Buonassisi T, Pickett M D, Ghosh M, Muller A, Quang N L, *Proc. 20th European Photov.Solar Energy Conf. and Exh*., Barcelona (Spain) p. 773 (2005).
54. Nordmark H, Di Sabatino M, Øvrelid E J, WalmsleyJ C and Holmestad R, *Proc.22nd European Photov. Solar Energy Conf*., Milan, p. 1710 (2007).
55. Buonassisi T, Heuer M, Istratov A A, Pickett M D, Marcus M A,. Lai B, Cai Z, Heald S M, Weber E R, *Acta materialia*, **55**, 6119 (2007).
56. Hocine S, Mathiot D, *Mater. Sci. Forum* **38-41**, 725 (1989).
57. Lutwack G, *Proc. of the Flat-Plate Solar Array Project Workshop on Low-Cost Polysilicon for Terrestrial Photovoltaic Solar-Cell Applications*, Contract DOE/JPL-1012-122, JPL Publication 86-11 (1986).
58. Dietl J, *Silicon for Photovoltaics*, Vol. 2, 285–352, North Holland, Amsterdam (1987).
59. Lanier F, Ang T, *Photovoltaic Engineering Handbook*, pp 3–17, Adam Hilger, Bristol (1990).
60. *Silicon Industry Vol. 2 – Technology Assessment*, Strategies Unlimited, Mountain View, CA 94040, USA (1983).
61. Hunt L, Dosaj V, Final Report, Contract DOE/JPL 954 559-78/7 (1979).
62. Hunt L, Dismukes J, Amick J, *Proc. Symp. Materials and Processing Technologies for Photovoltaics*, 106, The Electrochemical Society, Pennington, NJ (1983).
63. Amick J *et al*., *Proc. Symp. Materials and Processing Technologies for Photovoltaics*, 67, The Electrochemical Society, Pennington, NJ (1983).
64. Amick J *et al*., *Proc. 5th EC Photovoltaic Solar Energy Conf*., 336, Kovouri, Athens (1983).
65. Aulich H, Eisenrith K, Urbach H, Grabmaier J, *Proc. 3rd Symp. Materials and Processing Technologies for Photovoltaics*, 177 The Electrochemical Society. Pennington, NJ (1982).
66. Aulich H *et al*., *Proc. 4th EC Photovoltaic Solar Energy Conf*., 868 (Stresa, 1982).
67. Aulich H, *Proc. 5th EC Photovoltaic Solar Energy Conf*., 936 Kovouri, Athens (1983).
68. Yoshiyagawa M *et al*., Production of SOG-Si by Carbothermic Reduction of High Purity Silica, Presented at *Silicon for Solar Cells Workshop* (Schliersee, 1981).
69. Aas H, Kolflaath J, *US Patent 3,809,548* (1974).

70. Halvorsen G, *US Patent 4 539 194* (1985).
71. Ceccaroli B, Friestad K, Norwegian Patent Application WO 01/42 136 (2000).
72. Pizzini S, *Sol. Energy Mater*. **6**, 253 (1982).
73. Dietl J, *Sol. Cells* **10**, 145 (1983).
74. Theuerer H, *J. Met*. **8**, 1316 (1956).
75. Khattak C, Schmid F, Hunt L, *Proc. Symp. Electronic Properties of Polycrystalline or Impure Semiconductors and Novel Silicon Growth Methods*, pp 223–232, Saint Louis, MO,The Electrochemical Society, Pennington, NJ (1980).
76. Khattak C, Schmid F, *Proc. Symp. Materials and Processing Technologies for Photovoltaics*, pp 478–489, The Electrochemical Society. Pennington, NJ (1983).
77. Khattak C, Schmid F, *Silicon Processing for Photovoltaics II*, pp 153–183, Elsevier Science Publishers B.V., Amsterdam (1987).
78. Schmid F, Khattak C, *US Patent 5,972,107* (1999).
79. Amouroux J, Morvan D, *High Temp. Chem. Processes* **1**, 537–560 (1992).
80. Cazard-Juvernat I, Bartagnon O, Erin J, *High Temp. Chem. Processes* **3**, 459–466 (1994).
81. Combes R, Morvan D, Picard G, Amouroux J, *J. Phys. III France* **3**, 921–943 (1993).
82. Erin J, Morvan D, Amouroux J, *J. Phys. III France* **5**, 585–604 (1995).
83. Suzuki K, Sakaguchi K, Takano K, Sano N, *J. Jpn. Met*. **54**, 168–172 (1990).
84. Suzuki K, Kumagai T, Sano N, *ISJN Int*. **32**, 630–634 (1992).
85. Baba H *et al*., *Proc. 13th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 390–394 (Nice, 1995).
86. Nakamura N *et al*., *Proc. 2nd World Conf. on Photovoltaic Solar Energy Conversion* (Vienna, 1998).
87. Nakamura N *et al*., *EP Patent 0 855 367 A1* (1998).
88. Schei A, *US Patent 5,788,945* (1998).
89. Suzuki K, Sakaguchi K, Nakagiri T, Sano N, *J. Jpn. Inst. Met*. **54**, 61 (1990).
90. Ikeda T, Maeda M, *ISIJ Int*. **32**, 635–642 (1992).
91. Miki T, Morita K, Sano N, *Met. Mater. Trans*. **27B** 937–941 (1996).
92. Aulich H, Schulze F, Urbach H, Lerchenberger A, *Proc. Flat-Plate Solar Array Project Workshop on Low-Cost Polysilicon for Terrestrial Photovoltaic Solar-Cell Applications*, Contract DOE/JPL-1012-122, JPL Publication 86-11, pp 267–278 (1986).
93. Sakaguchi K, Maeda M, *Met. Trans. B* **23B**, 423–427.
94. Olson J, Carleton K, Kibbler A, *Proc. 16th IEEE Photovoltaic Specialists Conf*., pp 123–127 (1982).
95. Olson J, Carleton K, *J. Electrochem. Soc*. **128**, 2698 (1981).
96. Carleton K, Olson J, Kibbler A, *J. Electrochem. Soc*. **130**, 782–786 (1983).
97. Dawless R, *US Patent 4,246,249* (1981).
98. Smith F, Dawless R, *Electrochem. Soc*. **81-2**, 1147 (1981).
99. Kotval P, Strock H, *US Patent 4,195,067* (1980).
100. Hanoka J, Strock H, Kotval P, *J. Appl. Phys*. **52**, 5829–5832 (1981).

6

Bulk Crystal Growth and Wafering for PV

Hugo Rodriguez[1], Ismael Guerrero[1], Wolfgang Koch[2], Arthur L. Endrös[3], Dieter Franke[4], Christian Häßler[2], Juris P. Kalejs[5] and H. J. Möller[6]

[1]DC Wafers, Leon, Spain, [2]Bayer AG, Krefeld, Germany, [3]Siemens and Shell Solar GmbH, Munich, Germany, [4]Access e.V., Aachen, Germany, [5]RWE Schott Solar, Massachusetts, USA, [6]TU Bergakademie Freiberg, Freiberg, Germany

6.1 INTRODUCTION

The workhorse of the photovoltaics industry is silicon. Approximately nine of every ten kilowatts of solar modules produced come from crystalline silicon modules, around 50% from multicrystalline silicon material (Figure 6.1).

Silicon solar cells first were made about 50 years ago from Czochralski (Cz)-pulled monocrystals with technology adapted from the microelectronics industry. Subsequently, world-record cell efficiencies have been achieved at a very high cost on a laboratory scale with float zone (FZ) monocrystals. These cells reach almost 25% efficiency, while top efficiency industrial cells reach nowadays almost 22%.

Cost pressures have forced the development of multicrystalline material solidification processes for production of very large silicon ingots (blocks), which reached typical sizes of 450 kg in 2009. A good theoretical understanding of growth processes, together with numerical simulations of the entire process down to a microstructural defect level in the crystal today has resulted in the economical production of high-quality material. Sawing of silicon crystal into the thin wafers required for the best performance in solar cells wastes about 40–50% of expensive, pure silicon feedstock and is very costly. One alternative is crystalline silicon foil, that is, ribbon production processes, and these are now in various stages of R&D and commercialization. Another alternative, the so-called kerf-free, or no-waste alternatives, based on peeling wafers from the ingot rather than sawing, are slowly crawling into R&D and first pilot demonstrations.

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

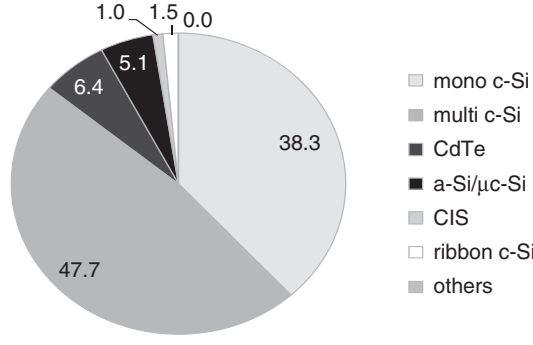


Figure 6.1 Photovoltaic technology share in 2009. Almost nine of every ten photovoltaic modules are made of crystalline silicon [1]

6.2 BULK MONOCRYSTALLINE MATERIAL

The dominant issue of the photovoltaic industry is to fabricate solar cells in large volumes that are both highly efficient and cost-effective. An overall industrial goal is to significantly lower the costs per watt. The dominant absorber material used today for the majority of the commercially produced solar cells is the Czochralski-grown crystalline silicon (c-Si) in monocrystalline and block-cast material in multicrystalline form (mc-Si). Up to now, a lot of effort has been undertaken to increase the electrical efficiency of industrial solar cells reproducibly towards and even above 20% [2], whereas much higher efficiencies have been claimed and reached [3, 4]. Unfortunately, efficiency improvements are often reached only with the help of cost-intensive process steps so that most steps cannot be directly implemented into industrial products, but have to be re-engineered for sufficiently low cost. Hence, there still remains a significant efficiency gap between monocrystalline laboratory cells with efficiencies above 24% and cost-effective, commercial Cz solar cells that are presently produced and sold in high volume at approximately 16–17% efficiency.

While some years ago the cost of a module was driven almost equally by the cost of the wafer (33%), the cell process (33%) and the module making (33%), this well-known ratio has changed significantly – both for single crystal silicon (sc-Si) and mc-Si. After 2004, in the recent years of polysilicon shortage, in most products the wafer has sometimes contributed more than 50% of the module cost, whereas the cell process and the module process contributed to the rest with similar portions of ~25%. The main reason for this is on one hand a steady cost reduction in the cell and module processes and on the other hand a significant increase in feedstock price, together with an almost unchanged wafer thickness of 250–350 µm in production. Today, where polysilicon is no longer scarce and wafer thicknesses have gone down to under 200 µm, the distribution is as follows: 14% feedstock, 20% wafer, 26% cell and 40% module [5].

Although a bit less stringent, wafers are still an important cost contribution. One way to meet today's demand of lower wafer cost is to: (1) reduce the cost of crystal growth by improvement of productivity and material consumption at constant wafer quality; (2) reduce the cost of the wire-sawing process; and (3) cut even thinner wafers. After several years of continuous reduction from some 300 µm, Si solar cells have a present wafer thickness of 180–200 µm. With current machines, further reduction is still possible, but wafers below 150–170 µm will be difficult to achieve at industrial scale. The main issue such thin wafers must face now is mechanical stability. Maintaining the breakage rate at low levels has increasingly become a headache for wafer, cell and module manufacturers, as well as wafer handling equipment suppliers. Theoretically, for full light absorption, a thickness of only 60–100 µm has been calculated to be the physical

optimum thickness for silicon solar cells [6]. In this thickness regime the maximum theoretical efficiency for c-Si solar cells can be reached. Also in this optimum thickness regime, mechanical stability is less important, as wafers become less fragile and more flexible. This is true for multicrystalline silicon sooner than for monocrystalline, already at some 110 μm [7]. However, for such thicknesses manufacturing processes have to be adapted, redesigned or newly developed to avoid bending and breaking of ultra-thin wafers. Also, ultra-thin wafers may be more flexible, but obviously less strong than thicker wafers, which must be taken into account. With reduced wafer thickness, there is also an increased necessity for surface passivation. Since this cannot be done without adding to the cost, any added process step has to add adequate efficiency to remain cost-effective. Other important issues for the efficiency of the 'valuable' wafer are improvements in antireflection (AR) coating, grid shadowing, 'blue' response of the emitter and volume passivation. All of these aspects are being developed parallel to the thickness reduction in recent years.

The demand for high-quality polysilicon feedstock in the world market grew quickly, not only in the microelectronic, but also in the photovoltaic industry. In 1980 the worldwide production of single-crystal silicon amounted to approximately 2000 metric tons per year. This number was equivalent to ~100 000 silicon crystals every year by either the Czochralski technique (80%) or the float zone technique (20%). The PV industry used both the high-quality tops and tails of microelectronics crystals for less than 5$/kg to fit the feedstock demand and the depreciated Cz pullers of the 'big brother' microelectronics industry. Since the microelectronics industry was and is still driven by continuously increasing ingot diameters, the 'small' Cz machines became available for the PV industry at interesting prices. During the last expansion phase of the microelectronic industry (1993–1999), the PV industry had to struggle with a severe shortage of affordable feedstock. To reach the production volume, even pot scrap Si material had to be used. New and demanding techniques were developed in a hurry to separate the Si from the quartz crucible parts and to preselect and pre-clean this material. Also, fine-grain material had to be made usable. The situation in the first decade of the new millennium is similar, but from a very different perspective. At the year 2000, predictions for the annual world requirement for solar quality silicon in 2010 were estimated at 8000–10 000 metric tons. But the growth of PV has outstripped even the most optimistic predictions, on average by 35% year-on-year, with years such as 2004 and 2008 above 65%. A new silicon shortage, this time due to the PV growth, occurred in 2004, again skyrocketing raw material prices. Ever since, polysilicon producers have increased capacity. Global polysilicon consumption reached 65 000 metric tons in 2008, 100 000 tons in 2009, and is expected to reach 170 000 in 2010, with already a bigger portion dedicated for PV than for microelectronics (55 versus 45% as early as 2006) [8]. The severe raw material shortage has also promoted a dedicated Si feedstock supply only for PV, which has been a necessity in recent years. Alternative sources such as upgraded metallurgical grade silicon, or less strict Siemens material, as well as other approaches, have attracted much attention and produced many good results lately.

It can no longer be denied that the growing of silicon crystals has matured from an art into a scientific business. In today's PV business, some of the bigger companies convert more than 20 tons of silicon per day into solar grade Cz crystals and solar cells. Since PV has different main requirements for crystal growing from the microelectronic industry, the focus of machine and process development differ.

6.2.1 Cz Growth of Single-crystal Silicon

Solar cells made out of Czochralski (Cz)-grown crystals and wafers play – together with multicrystalline cells – a dominant part in today's PV industry. This is due to the following advantages.

Cz crystals can be grown from a wide variety of differently shaped and doped feedstock material. This enables the PV industry to buy cost-effective feedstock silicon with sufficient quality, even on spot markets. Since the feedstock is molten in a crucible, the shape, the grain size and the resistivity of the different feedstock materials can be mixed for the required specifications, although a given feedstock alone would fail. However, special care must be taken to avoid any macroscopic particles (SiO_2, SiC) that would not be dissolved in the melt, especially when pot scrap material is used.

The Cz process acts as a purification step with respect to lifetime-limiting elements. The effective distribution coefficients of the most dominant lifetime-limiting metals (Fe, Ni, Au, Ti, Pt, Cr) are in the range of 10^{-5} or below. Together with appropriate gettering steps during cell processing, highly efficient commercial solar cells can even be made out of ingots grown from low-grade pot scrap material. The last decade has been very productive regarding defect and impurity engineering. Many efforts have been dedicated to investigating the real effect of defects and impurities inside the silicon matrix [9]. Surprising results have been obtained as to the high levels of metal impurities that can be allowed in silicon for PV applications [10]. Many studies have been made, specially for iron, where silicon has been intentionally contaminated with contents as high as 10^{15} atoms/cm^3 [11]. Unlike the conventional microelectronics belief that concentrations above 10^{12} atoms/cm^3 would destroy minority-carrier lifetime, such high concentrations still achieve good cell efficiencies. All these results indicate that some or many of the assumptions inherited from the microelectronics may be less strict for PV. Impurity and defect engineering is therefore a new field in photovoltaics that is bringing interesting results and should be followed with attention.

The Cz process itself acts as a quality control step since proper crystallisation, that is, dislocation-free growth of an ingot, can only take place in a well-defined process window. The homogeneity of a well-grown solar grade Cz ingot for PV application is excellent with respect to the bandwidth of electronic and structural properties, whereas mc-Si block casting produces specifications with higher variances in most parameters. Cell processes with Cz-Si can therefore use high-efficiency processes with smaller process windows that require well-defined starting material.

Cz technology is mature and cost-effective. Equipment and processes for semi-automated growing of crystals are commercially available so that several Cz pullers can be run by a single operator. Owing to the robust construction of the machines, many Cz growers more than 20 years of age are still in production.

The ingot can be pulled in a defined <100> orientation. This is a big economic advantage since the solar cell process can use this crystallographic property to homogeneously texture solar cells with a very cost-effective wet chemical etching step. By anisotropic etching, a surface structure with random pyramids is built that couples the incoming light very effectively into the solar cell. This effect, together with the usually higher diffusion length of Cz crystals gives rise to the increased efficiency of Cz-Si solar cells compared with similarly processed mc-Si cells.

There exists a high potential for increasing the net pulling speed, that is, the productivity of a puller by a clever design of the hot zone, by sophisticated recharging concepts of Si in the hot crucible and by tuning the growth recipe to the optimum pull speed. Here the PV industry is in the novel position that it can neglect most specifications that are required in the microelectronics industry. For instance, most of the Cz growers nowadays already have a recharging system in their pullers [12, 13].

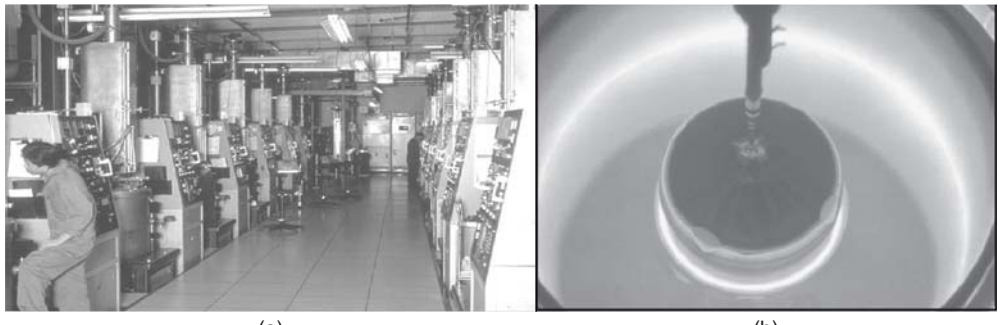
This is possible since the PV specifications are strongly reduced in the number of required parameters in contrast to microelectronic material. A PV specification ‘simply’ focuses on the maximum productivity, a minority-carrier diffusion length of the material that should exceed the cell thickness and a usually *p*-type doping that leads to a specific resistivity between 0.3 and 10 Ω cm, depending on the fabricated solar cell type.

One of the main disadvantages of Cz crystallisation of silicon is the fact that square cells are best suited to build a highly efficient solar module, whereas Cz ingots have a round cross-section. In order to use both the crystal and the module area in the best manner, the ingots are usually cut into a pseudo-square cross-section before they are cut into wafers. Additionally, the tops and tails of the ingots cannot be used for wafer production. The cropped and slapped materials, that is, tops and tails and so on, are then fed back into the growth process again.

The equipment and the basic principle for Cz pulling is shown in Figure 6.2. The Cz equipment consists of a vacuum chamber in which feedstock material, that is, polycrystalline silicon pieces or residues from single crystals, is melted in a crucible and a seed crystal is first dipped into the melt. Then the seed is slowly withdrawn vertically to the melt surface, whereby the liquid crystallises at the seed. High-vacuum conditions can be used as long as the melt weight is small (<1–2 kg), but with larger melts (nowadays usual sizes exceed 100 kg, patented equipment can reach even 200 kg [14]) only pulling under an argon inert gas stream is practicable. Owing to the reduced argon consumption, the argon pressure is set in the 5–50 mbar regime in the PV industry, whereas in the microelectronic industry, atmospheric pressure is also used.

After the silicon is completely molten, the temperature of the melt is stabilised to achieve the required temperature to lower the seed into the melt. The temperature must be chosen so that the seed is not growing in diameter (melt too cold) or decreasing in diameter (melt too hot). In PV the seed is usually <100>-oriented, is monocrystalline and is pulled upwards to grow a 'crystal neck'. Since dislocations propagate on (111) planes that are oblique in an <100>-oriented crystal, the dislocations grow out of the crystal neck after a couple of centimetres so that the rest of the crystal grows dislocation-free, even if the growth was started from a dislocated seed. The dislocation-free state of the grown crystal manifests itself in the development of 'ridges' on the crystal surface. If this state is achieved, the diameter of the crystal can be enlarged by slower pulling until it reaches the desired value. The transition region from the seed node to the cylindrical part of the crystal has more or less the shape of a cone and is therefore called the 'seed cone'. This cone can be pulled differently, either flat or steep.

Shortly before the desired diameter is reached, the pulling velocity is raised to the specific value at which the crystal grows with the required diameter. Owing to the seed rotation, the crystal cross-section is mostly circular. In general, the pulling velocity during the growth of the cylindrical part is not kept constant, but is reduced towards the bottom end of the crystal. This is mainly caused by the increasing heat radiation from the crucible wall as the melt level sinks. The heat removal of the crystallisation thus becomes more difficult and more time is needed to grow a certain length of



(a) (b)

Figure 6.2 Cz pullers in a PV production environment (a) and growing Cz crystal (b) in a quartz crucible

the crystal. Standard pull speeds in the body range from 0.5 to 1.2 mm/min. The diameter of the crystal in PV is often chosen between 200 and 300 mm [15]. This is due to the short-circuit current of big solar cells where values of 6 A per cell are exceeded. It is difficult to provide a proper contacting scheme in screen print technology that can handle such high currents in the front contacts without high series-resistance losses. With even larger cell sizes, this effect becomes more problematic.

To complete the crystal growth free of dislocations, the crystal diameter has to be reduced gradually to a small size, whereby an end cone develops. For this purpose, the pulling speed is raised and the crystal diameter is decreased. If the diameter is small enough, the crystal can be separated from the melt without a dislocation forming in the cylindrical part of the crystal. The withdrawal of the crystal from the residual melt can be done with a rather high velocity, but not too fast, because thermal shock would cause plastic deformation called 'slip' in the lower part of the crystal. The final crystal length is dependent on the crucible charge and varies between 200 and 400 cm [16].

Nowadays, the seed crystals used for Cz crystal growth are usually dislocation-free. However, each time the seed crystal is dipped into the melt, dislocations are generated by the temperature shock and by surface tension effects between the melt and the crystal. Normally these dislocations are propagated, or move into the growing crystal, particularly if the crystals have large diameters. The movement of dislocations is affected by cooling strain and faulty crystal growth.

The strain that occurs as a result of different cooling rates between the inner and the outer parts of the crystal is probably the main reason for the dislocation movement in the case of large crystals. At the usual <100> crystal orientation, no (111) lattice plane, that is, no main glide plane, extends parallel to the crystal axis. All (111) glide planes are oblique to the crystal axis and as a result all dislocations that move only on one glide plane are conducted out of the crystal at some time. For movement in the pulling direction, the dislocations have to move downwards in a zigzag motion using at least two of the four different (111) glide planes. Dislocation-free crystal growth is relatively stable, even for large crystal diameters, in spite of the higher cooling strain. This is so because it is difficult for a perfect crystal to generate a first dislocation. However, if a first dislocation has been formed, it can multiply and move into the crystal. In this way, numerous dislocations are generated and spread out into the crystal until the strain becomes too low for further movement of dislocations. Therefore, if a dislocation-free growing crystal is disturbed at one point, the whole cross-section and a considerable part of the already-grown good crystal are inundated with backward-moving slip dislocations. The length of the slip-dislocated area is approximately equal to the diameter. Wafers and cells that show these types of dislocations can easily be hydrogen-passivated.

After losing the dislocation-free state, the crystal continues to grow with a high density of dislocations that are usually arranged irregularly. They are partly grown-in to the crystal and partly generated later by strain-induced processes. At high temperatures, climb processes take place that distort the dislocation array even more. This further increases the irregular shape and the distribution of the dislocations. In contrast to simple 'slip' dislocations, these 'grown-in' dislocations cannot be passivated well by a hydrogen-passivation step later in the solar cell processing. With crystal diameters above 30 mm, the monocrystalline, but dislocated, growth is not stable and in most cases changes to polycrystalline growth because of the tendency of a Si crystal to form twins in the presence of strain and dislocations. These twins also multiply and form higher-order twins and thus rapidly form a polycrystal. This fine-grained poly-Si material is not usable for solar cell production. Known causes for the generation of the critical first dislocation are either solid particles in the melt that move to the solidification front, gas bubbles that are trapped at the solidification front, impurities that exceed the solubility limit in the melt, vibrations of crystals and melt, thermal shocks or a too high cooling strain.

The growth of the seed cone is the most critical stage in the pulling of the Cz crystal. For productivity reasons very flat seed cones are preferable since the time needed to make the cone is not productive. However, the probability of introducing dislocations in the seed cone is lowest for tapered cones, although this means an increase in the pulling time by 15–25% for the same body length and additional loss in usable material. The loss of time and material gets worse for larger ingot diameters.

Owing to the reaction between the liquid Si and the quartz crucible, the crucible is of considerable importance to the growth. The silica of the crucible supplies considerable amounts of oxygen to the melt and, owing to the high purity of the silica, only small amounts of other impurities. However, the crucible tends to dissolve after a long time so that the risk for particles in the melt from the crucible is increased with increased pulling time. The oxygen of the melt adds up to 10^{18} oxygen atoms/cm^3 to the growing crystal, whereas carbon is usually $<10^{17}$/cm^3 and has only little impact on the solar cell performance. Oxygen effects such as thermal donors and precipitates can be well controlled in Cz cell processing.

6.3 BULK MULTICRYSTALLINE SILICON

Multicrystalline silicon besides monocrystalline silicon represents the basis of today's photovoltaic technology. Multicrystalline silicon offers advantages over monocrystalline silicon with respect to manufacturing costs and feedstock tolerance at, however, slightly reduced efficiencies. Another inherent advantage of multicrystalline silicon is the rectangular or square wafer shape, yielding a better utilisation of the module area in comparison to the mostly round or pseudo-square monocrystalline wafers. The efficiencies of multicrystalline silicon solar cells are affected by recombination-active impurity atoms and extended defects such as grain boundaries and dislocations. A key issue in achieving high solar cell efficiencies is a perfect temperature profile of both ingot fabrication and solar cell processing in order to control the number and the electrical activity of extended defects. Moreover, the implementation of hydrogen-passivation steps in solar cell processing turned out to be of particular importance for multicrystalline silicon. With the introduction of modern hydrogen-passivation steps by silicon nitride layer deposition, the efficiencies of industrial multicrystalline silicon solar cells were boosted to the 14–15% efficiency range and consequently market shares were continuously shifted towards multicrystalline silicon as the standard material of photovoltaics. Nowadays, these industrial efficiencies reach 16%, and defect engineering is becoming more important in ingot growth and solar cell processing.

6.3.1 Ingot Fabrication

Two different fabrication technologies for multicrystalline silicon, the Bridgman and the block-casting process (Figures 6.3 and 6.4) are employed. In both processes the solidification of high-quality multicrystalline silicon ingots with weights of 450 kg, dimensions of up to 90 × 90 cm and heights of more than 30 cm have been successfully realised. While the Bridgman technology is a quite commonly used technique, the only two companies mainly employing the casting technology are Kyocera (Japan) and Deutsche Solar GmbH (Germany) [17, 18].

The main difference between both techniques is that for the melting and crystallisation process only one crucible (Bridgman) is used, whereas for the crystallisation process a second crucible (block-casting) is used.

In the case of the Bridgman process, a silicon nitride (Si_3N_4)-coated quartz crucible is usually employed for melting of the silicon raw material and subsequent solidification of the multicrystalline ingot. The Si_3N_4 coating thereby serves as an anti-sticking layer, preventing

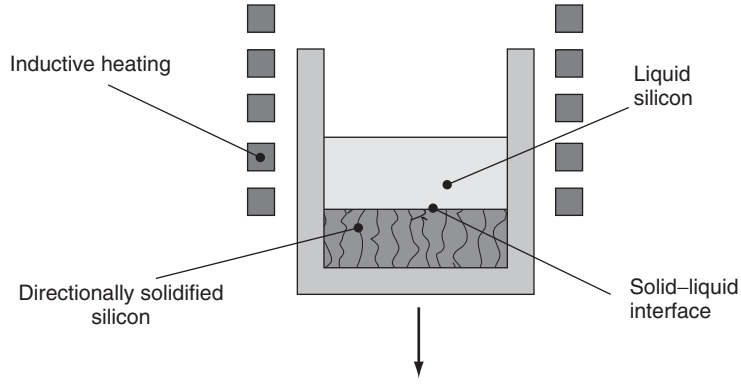


Figure 6.3 Conventional Bridgman technique that is still mainly used for the fabrication of multicrystalline ingots. Both melting and crystallisation of the silicon is performed in a Si_3N_4-coated quartz crucible. Crystallisation is realised by either slowly lowering the liquid silicon-containing crucible out of the inductively heated hot zone of the process chamber, or raising this zone out of the crucible

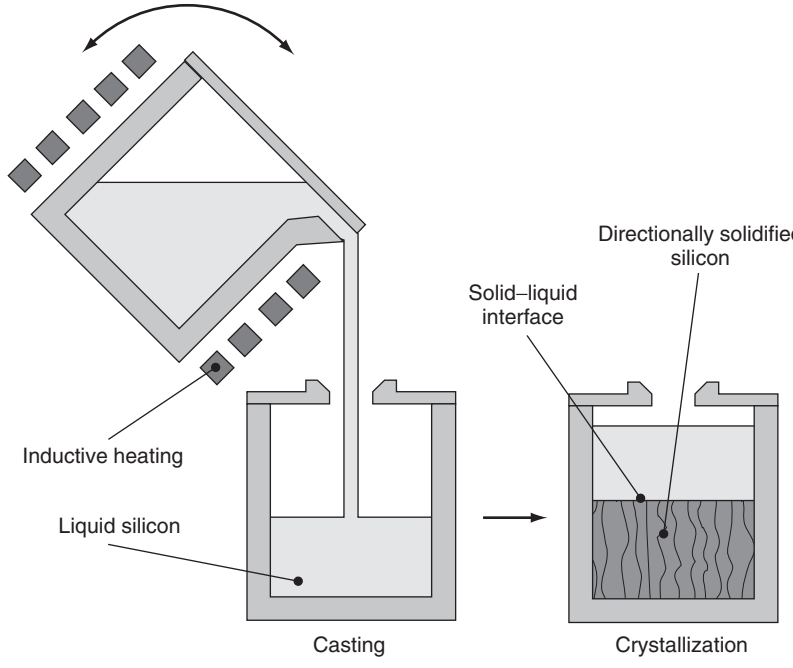


Figure 6.4 Block-casting process for the fabrication of multicrystalline silicon. After melting the silicon in a quartz pot, the silicon is poured into a second quartz crucible with a Si_3N_4 coating. The heating elements of the crystallisation crucible are not shown in the figure. In comparison with the Bridgman technique (Figure 6.3), shorter crystallisation and cooling times can be realised by employing a more variable heater system

the adhesion of the silicon ingot to the quartz crucible walls. Concerning the block-casting process, the melting is performed in a quartz crucible without a coating, whereas – after pouring the molten silicon into a second crucible – for the crystallisation also a Si_3N_4-coated one is used.

Usually, in both production technologies, crystallisation starts at the bottom of the crucible by reducing the temperature below the melting temperature (1410 °C) of silicon. Within the Bridgman process the temperature reduction is achieved by simply lowering the liquid silicon-containing crucible out of the hot area of the crystallisation furnace, or the opposite: elevating the hot area out of the standing crucible. Contrarily, the temperature control during the block-casting process is achieved by a corresponding adjustment of the heaters, whereas the crucible itself is not moved during solidification.

After solidification starts in the bottom region, the crystallisation front, that is, the liquid–solid interphase, moves in a vertical direction upwards through the crystallisation crucible. This so-called directional solidification results in a columnar crystal growth and consequently adjacent wafers fabricated out of the ingots show nearly identical defect structures (grain boundaries and dislocations).

Common crystallisation speeds used for the Bridgman technology are in a range of about 1.5 cm/h (corresponding to a weight of approximately 25 kg/h for large ingots). With regard to the increase in crystallisation speed, that is, productivity, mainly cooling of the already crystallised fraction of the ingot has to be taken into account. Too high process speeds cause large thermal gradients within the solidified silicon that may result in cracks or even destruction of the ingot. For the block-casting technology, however, owing to the more versatile and sophisticated heater system, considerably higher crystallisation speeds can be achieved [18].

6.3.2 Doping

Standard multicrystalline silicon is a boron-doped *p*-type material with a specific electrical resistivity of about 1.5Ω cm, which corresponds to a boron concentration of about $1 \times 10^{16}/cm^3$. The specific resistivity is adjusted in a way such that optimal solar cell performance is guaranteed. Naturally, the boron concentration can be varied according to the requirements of specific solar cell processes. Specific resistivities in a range of 0.1–5 Ω cm have been used for solar cell fabrication so far. The boron concentration is normally adjusted by adding a choice of different boron sources to the silicon raw material prior to the melting of the silicon: the equivalent amount of B_2O_3; the equivalent amount of boron powder; or the equivalent amount of heavily doped Cz-grown silicon. When using alternative doping elements like gallium (*p*-type) or phosphorus (*n*-type), the segregation coefficient governing the resistivity decrease with increasing block height has to be considered. With a segregation coefficient of 0.8, boron is nearly always the optimal doping element, giving only a small resistivity change over the silicon ingot (Figure 6.5), whereas gallium and phosphorus with segregation coefficients of 0.008 and 0.35, respectively, are less favourable.

For phosphorus as an *n*-type dopant, the additional disadvantage is encountered of a lower minority charge carrier (i.e. holes) mobility and a more complicated solar cell process, using, for example, higher process temperatures for boron instead of phosphorous diffusion. However, there is a series of advantages which could render *n*-type material an attractive new feedstock source for photovoltaics: it offers an increased supply of reject silicon from the microelectronics industry; it is also immune to the well-known boron–oxygen defect that can affect boron-doped *p*-type silicon. Another potential advantage of *n*-type mc-Si is its possible greater tolerance of some important metal impurities, which is still under debate [19, 20]. In any case, there is ample evidence for very high carrier lifetimes in *n*-type mc-Si [21]. The last decade has been fruitful in improving process

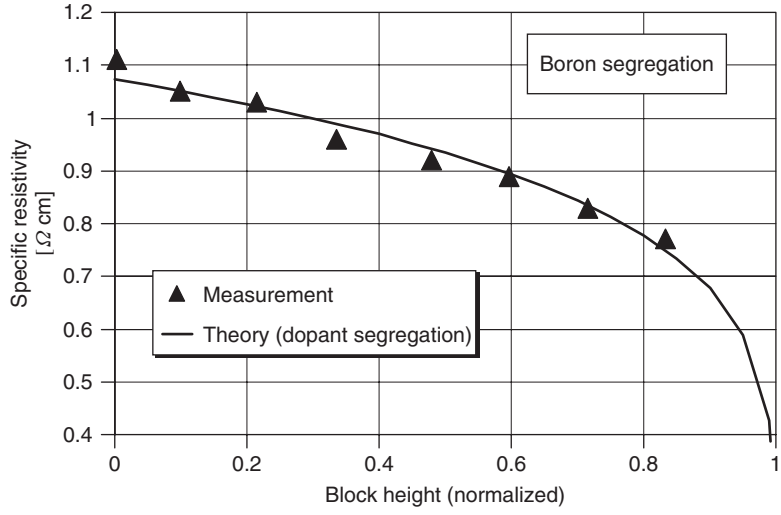


Figure 6.5 Decrease in the specific resistivity of *p*-type multicrystalline silicon due to segregation of the doping element boron

steps with *n*-type material, reaching efficiencies of above 18%, with promising results for cheap industrialisation in the coming years [22–24].

6.3.3 Crystal Defects

The main crystal defects in multicrystalline silicon are grain boundaries and dislocations. Concerning the attainable efficiencies of solar cells, not only the concentration of these defects but also their electrical activity is considered as crucial.

With respect to the grain boundaries and the grain size, basically smaller grains are observed at the beginning of the crystallisation process in the ingot bottom part. With increasing block height, individual grains prevail at the expense of surrounding grains and thus give rise to an increase in the mean grain size. This increase of grain size, however, depends on the crystallisation speed (Figure 6.6). A higher crystallisation speed also means higher temperature gradients and thus an increased probability for the formation of crystal seeds in the melt that in turn lead to a limitation of the grain size. This is also the reason for faster crystallised block-cast material usually exhibiting smaller grains than conventional Bridgman-type multicrystalline silicon.

On the other hand, the grain sizes achieved with modern block-cast material are still large enough not to degrade solar cell efficiencies, provided that the electrical activity of the grain boundaries is low enough. Grain boundaries and dislocations, if electrically charged, effectively attract minority charge carriers and consequently represent highly active recombination centres for photo-generated charge carriers. The electrical activity of grain boundaries and dislocations is determined by their impurity decoration (specifically by transition metals) and strongly increases with higher impurity concentrations. Recently it has also been observed that the grain boundary type has an influence in Fe gettering, and therefore on electrical activity: low-Σ grain boundaries have a lower gettering ability than high-Σ or random grain boundaries, therefore leading to lower electrical activity [25]. Also small-angle (SA) grain boundaries show more electrical activity than other grain boundary types.

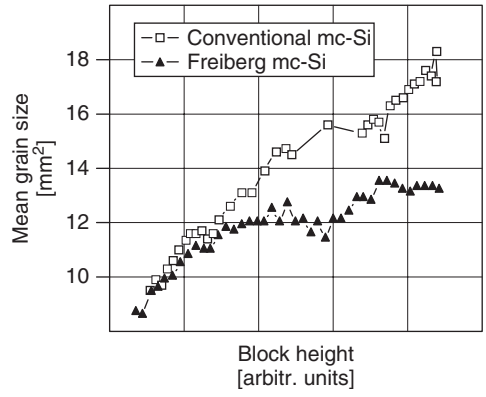


Figure 6.6 Mean grain size as a function of the block height for conventional Bridgman-type multicrystalline silicon (mc-Si) and the faster crystallised block-cast material from the Freiberg production facility of Deutsche Solar GmbH. Reduced crystallisation times lead to slightly lower grain sizes of the Freiberg mc-Si

Moreover, it was discovered that the shape of the crystallisation front during solidification also has considerable influence on the grain boundary activity [26]. Preserving a strictly planar solidification front clearly leads to a reduced grain boundary activity. Also because in modern high-throughput production processes a nearly perfectly planar solidification front is maintained throughout the crystallisation process, grain boundaries show only weak electrical activities and therefore are generally considered as less important for solar cell efficiencies.

Crystal dislocations, however, turned out to be the most efficiency-relevant defects in multicrystalline silicon for solar cells. The dislocation density that is experimentally accessible by counting micrometer-sized etch pits after appropriate chemical etching steps shows a nearly perfect correlation to the wafer lifetime and diffusion length (Figure 6.7) that are closely linked to

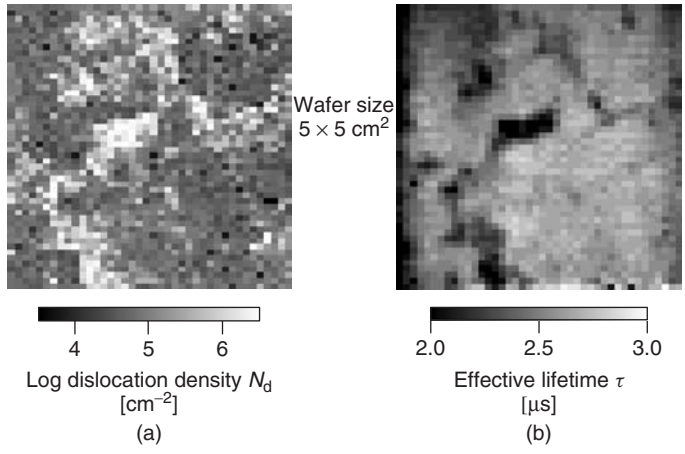


Figure 6.7 Topographies of the dislocation density N_d (a) and the effective lifetime τ_{eff} (b) of a typical multicrystalline silicon wafer, showing the excellent correlation of both parameters. The effective lifetime measurements were conducted out without any surface passivation and thus are limited to less than approximately 3 µs by surface recombination processes

solar cell performance. Dislocations are induced and multiplied by thermal stress that originates from temperature inhomogeneities during crystallisation and cooling of the ingot. The reduction of these temperature variations while maintaining high process speed is therefore considered one of the most important issues for the further improvement of multicrystalline silicon. Alternatively, in the so-called dislocation-engineering there is a new trend of reducing crystal dislocations by high-temperature annealing at the wafer stage, and very promising results have been obtained lately with high reduction in dislocation density [27]. However, some issues such as lifetime preservation during this annealing remain unresolved.

An optimal process scenario for the production of multicrystalline silicon from both the crystal defect and the productivity point of view starts with a small crystallisation speed and minimal temperature gradients in order to secure a low-defect-density bottom region of the ingot. After that, crystallisation speed should be largely increased for productivity reasons while keeping the solidification front planar and thermal gradients within the solidified silicon low.

6.3.4 Impurities

Despite boron being the standard dopant and thus an intentionally introduced impurity, even higher impurity concentrations in multicrystalline silicon are observed for both oxygen and carbon.

The interstitial oxygen concentration in multicrystalline silicon is affected by two processes, which are oxygen incorporation via the quartz crucible during melting and oxygen loss through evaporation of SiO, that is, the evaporating gaseous silicon monoxide that is stable at high temperatures only.

Because the segregation coefficient >1, the oxygen content decreases with increasing block height. Typical concentrations of the interstitial oxygen content of Bridgman and block-cast material are given in Table 6.1. Obviously, although the silicon melt never stays in direct contact with the quartz crucible, lower oxygen concentrations with Bridgman-type material compared with silicon from the block-casting process are not feasible. It can therefore be concluded that there also has to exist an oxygen release from the Si_3N_4 coating (containing some percentage of oxygen) into the silicon melt during the Bridgman process. In addition, a much more rapid decrease of the oxygen concentration with increasing block height is observed for block-cast material, which is attributed to the lower ambient pressure and enhanced gas exchange normally employed during this process.

Although oxygen residing on interstitial lattice sites is not electrically active, recombination-active oxygen complexes such as thermal donors [28–30], new donors [31, 32] and oxygen

Table 6.1 Typical concentrations of interstitial oxygen [O_i] for block-cast material from the Freiberg production plant of Deutsche Solar GmbH and for typical material coming from a Bridgman process. For the determination of the oxygen concentration by Fourier transform infrared spectroscopy (FTIR), a conversion factor of $2.45 \times 10^{17}/cm^2$ was used

Ingot position	Interstitial oxygen concentration [O_i] [$10^{17}/cm^3$]	
	Block-casting process	Bridgman process
Bottom	6.5	6
Middle	0.9	3.5
Top	0.5	2

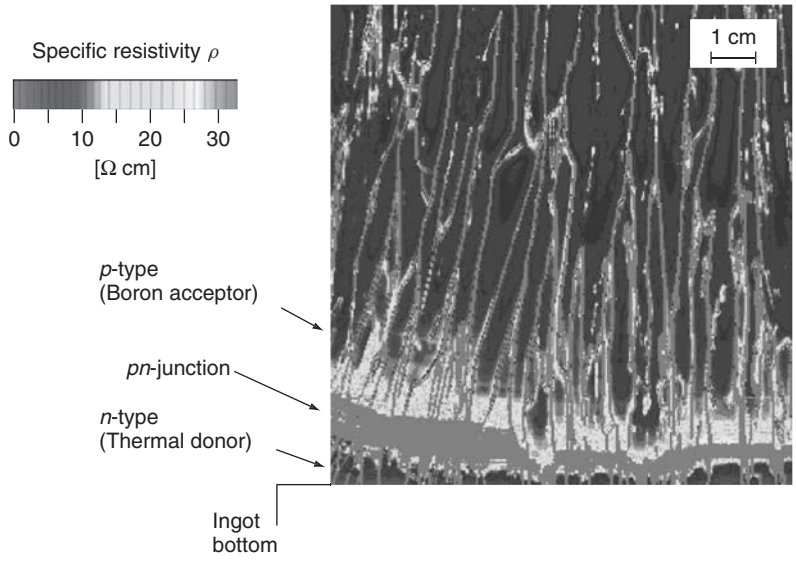


Figure 6.8 High-resolution map of the specific resistivity (van der Pouw measurement technique) of a vertically cut wafer from a special high-resistivity *p*-type multicrystalline silicon test ingot. Owing to the increased oxygen content, the formation of thermal oxygen donors changes the conductivity to *n*-type in the bottom part. The *pn*-junction can be identified by a marked increase in the specific resistivity

precipitates may be formed after annealing steps, specifically in the high oxygen concentration bottom part of the ingot (see an example of the donor activity in Figure 6.8).

Specifically, thermal donors turned out to be mainly responsible for a broad low-lifetime region with a width of 4–5 cm in the bottom part of Bridgman-type ingots [33]. Owing to the instability of the thermal donors in high-temperature steps during solar cell processing, these low lifetimes, however, do not lead to low efficiencies. It is worth noting that the width of this low-lifetime region in the bottom part of the ingots is largely reduced for material from the block-casting process. The most likely explanation for this is the shorter process times that consequently give less time for the formation of oxygen complexes out of interstitial oxygen atoms.

Similar to metals, oxygen segregation at grain boundaries and dislocations enhances the recombination strength of these extended defects. Oxygen precipitates may also getter metal impurities during crystallisation, which are released afterwards during solar cell processing as highly recombination-active point defects.

Generally, the manifold involvement of oxygen in efficiency-relevant microscopic processes makes the reduction of the oxygen incorporation into multicrystalline silicon one of the most important targets of material improvement efforts.

Like oxygen, carbon in multicrystalline silicon appears in concentrations considerably higher than those of the boron dopant concentrations. Typical concentrations of substitutional carbon are in the range of $2–6 \times 10^{17}/\text{cm}^3$, generally increasing with increasing block height. The incorporation of carbon into the silicon melt takes place via gaseous CO formation inside the crystallisation chamber by SiO chemically reacting with the graphite heaters. The main problem that is caused by an increased carbon concentration is the formation of SiC crystals (often associated with oxygen

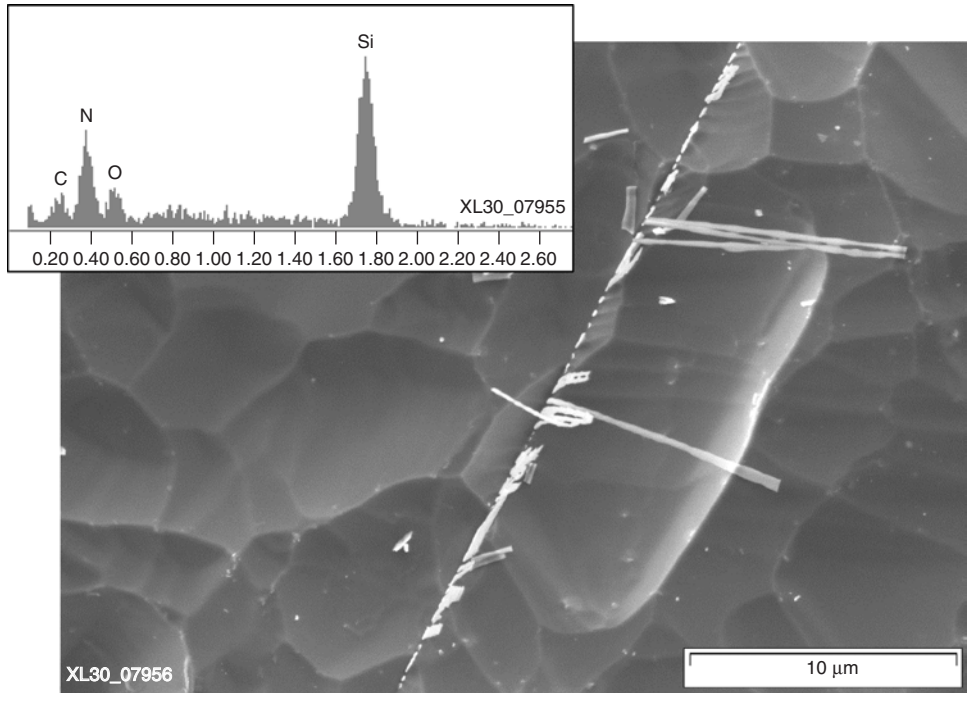


Figure 6.9 SEM (scanning electron microscope) image of a heavily shunted solar cell region. The microscopic investigations reveal needle-shaped structures containing silicon, nitrogen, oxygen and carbon. The shunting mechanism is assumed to be due to electrically conductive SiC that short-circuits the solar cell *pn*-junction. Si_3N_4 inclusions may have a similar effect

and nitrogen, see Figure 6.9) within the silicon material. SiC, representing an electrically conductive semiconductor material, effectively shorts the solar cell *pn*-junction, thereby leading to drastically reduced efficiencies. The problem of SiC formation, however, usually occurs only in the uppermost region of the ingot that is anyway rejected because of segregation of metallic impurities. Inclusions can also be caused by a high nitrogen concentration, precipitating in Si_3N_4 particles. Its electrical activity and consequences may be similar to that of SiC inclusions [34].

Despite oxygen and carbon being present in much higher concentrations, transition metals such as iron or titanium are considered as much more important with regard to solar cell efficiencies, with the exception of the outer edges (width 5–10 mm) of an ingot where in-diffusion from the Si_3N_4 coating may occur and the top segregation region metal. Point defects in high-quality multicrystalline silicon are present in concentration levels below the detection limit of deep-level transient spectroscopy (DLTS) measurements, that is, below approximately $10^{12}/cm^3$. The importance of metal impurities for multicrystalline silicon solar cells is, however, based on metal impurities controlling the activity of extended defects, specifically that of crystal dislocations.

It has been anticipated that metallic impurities are, for example, responsible for the observed systematic changes of the lifetime of multicrystalline silicon wafers after high-temperature steps in the range 800–1000 °C (e.g. the phosphorus diffusion step for fabrication of the solar cell *pn*-junction). The wafer lifetime quite commonly decreases in annealing steps above 900 °C, where this decrease is even more significant at enhanced cooling speeds after the anneal. The proposed mechanism behind this lifetime degradation is a release of metal atoms from extended

defects such as dislocations into the wafer bulk material and a subsequent quenching of the metal atoms as highly recombinative point defects.

Another hint of an extensive interaction between extended defects and metal impurities is given in Figure 6.10, showing the theoretical segregation profile of iron in an intentionally contaminated multicrystalline ingot (mean iron concentration: $7.9 \times 10^{17}/cm^3$) compared with the experimental one. A reduced experimentally determined segregation effectiveness can be clearly identified, most probably caused by iron segregation into extended defects competing with the segregation process in the liquid silicon phase during crystallisation.

In order to prevent such defect–metal interaction processes leading to enhanced recombination activity, a very effective segregation of metallic impurities into the ingot top region has to be assured. This segregation effectiveness, however, decreases with both increasing crystallisation speed and increasing concentration of extended defects.

This in turn verifies the importance of a properly adjusted and controlled crystallisation speed. In order to assure an effective impurity segregation for high-quality multicrystalline silicon, specifically in regions with increased defect densities (e.g. ingot bottom part), solidification at a low crystallisation speed is essential.

In parallel to improving segregation during crystallization, defect and impurity engineering during crystal growth and cell processing can also be applied to reduce the electrical activity of impurities. Much research has been done lately, not only trying to reduce iron content in silicon, but also transforming it into an electrically inactive form: it has been shown that iron co-precipitates with other metals are less harmful than interstitial iron in silicon [35]. Also, it has been recently shown that slower cool-down steps after P-diffusion have an impact on lifetime, and iron co-precipitation is assumed to be the case [36]. With all these new findings, defect and impurity engineering remains a very promising field to further improve cell efficiency.

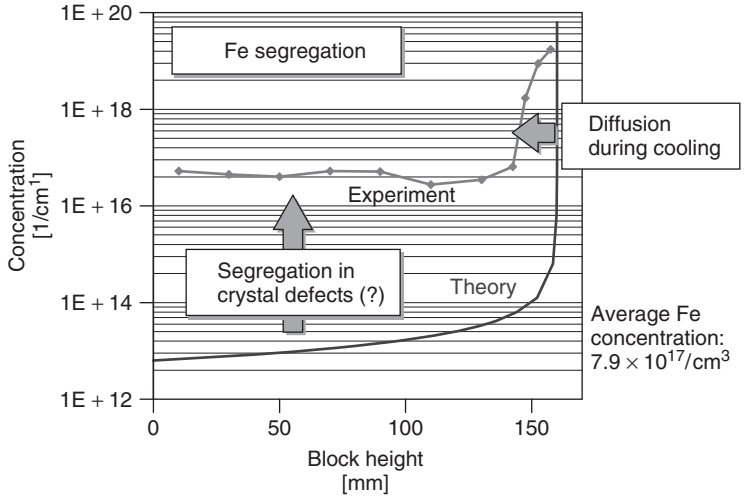


Figure 6.10 Experimentally determined iron concentration of an intentionally contaminated multicrystalline silicon test ingot (mean Fe concentration: $7.9 \times 10^{17}/cm^3$). The experimental data is given as a function of the block height and in comparison with the theoretically expected profile. The much higher than theoretically predicted concentration in the bottom and middle part of the ingot is attributed to segregation, not only into the silicon melt, but also into extended crystal defects during solidification

6.4 WAFERING

Between 80 and 90% of the current solar cell production requires the cutting of large silicon crystals [37]. Multicrystalline ingots grown by the Bridgman or gradient freeze technique now reach cross-sections of more than 80 × 80 cm and weigh over 400 kg; monocrystalline Cz crystals have diameters of up to 20 cm today. While in the last few years the cost of solar cell processing and module fabrication could be reduced considerably, the sawing costs remain high.

Figure 6.11 shows that the sawing costs are a substantial part (29%) of the wafer production cost and thus contribute considerably to the total module cost. Since the sawing of the crystals is connected with high material losses (about 40%), ribbon growth techniques or the thin film technology, which avoid the sawing step, have a high potential for developing cheaper solar cells, although these technologies have other problems involved that increase their cost and make them less attractive when polysilicon prices are low. Both technologies still have to overcome serious difficulties and their full development will probably take another 5–10 years. The present task is therefore to optimise the sawing technique for further cost reduction in mass production.

At the beginning of the PV industry, the available sawing technology of the microelectronic industry was used. The ingots were mainly cut by inner diameter (ID) saws. This technology is, however, relatively slow and not economical for mass production [38]. It was therefore gradually replaced by the multi-wire slicing technology [39]. The advantages are the higher throughput of about 8000–10 000 wafers per day and per machine, a smaller kerf loss of 100–180 µm and almost no restrictions on the size of the ingots. Currently, wafers between 180 and 220 µm are usually cut, but a wafer thickness down to about 100 µm can be achieved by the technique in the laboratory [40]. Since the technology is relatively new and still under development, most wafer manufacturers have to optimise the sawing process by their own experience. The sawing process depends on several variable parameters as will be described next, which makes it difficult to optimise the process in view of throughput, material losses, reduction of supply materials and wafer quality. Basic knowledge about the microscopic details of the sawing process is required in order to slice crystals in a controlled way. In the following section, the principles of the sawing process will be described as far as they are understood today.

6.4.1 Multi-wire Wafering Technique

After crystal growth the silicon ingots are cut in a first step, either by band saws or by wire saws into columns with a cross-section that is determined by the final wafer size. Standard sizes are about 15 × 15 cm, but larger wafers sizes up to 21 × 21 cm are being studied for their use in the solar cell

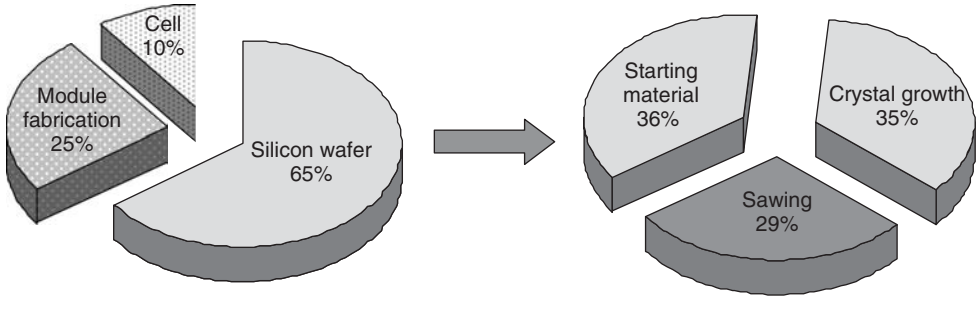


Figure 6.11 Cost distribution for modules and silicon wafers

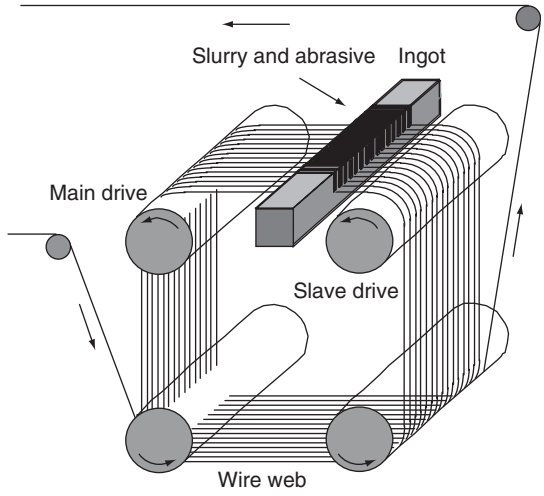


Figure 6.12 Schematic diagram depicting the principle of the multi-wire sawing technique

technology. The columns are glued to a substrate holder and placed in a multi-wire saw that slices them into the final wafers. The principle of the multi-wire technology is depicted in Figure 6.12. A single wire is fed from a supply spool through a pulley and tension control unit to the two to four wire guides that are grooved with a decreasing pitch due to the wire degradation during the cut. Multiple strands of a wire net (known as web) are formed by winding the wire on the wire guides through the 1000–1200 parallel grooves. A take-up spool collects the used wire. The wire is pulled by the torque exerted by the main drive and slave as shown in the figure. The tension on the wire is maintained by the feedback control unit at a prescribed value. The silicon column on the holder is pushed against the moving wire web and sliced into hundreds of wafers at the same time. The wire either moves in one direction or oscillates back and forth. Solar cell wafers are mainly cut by a wire that is moving in one direction, whereas wafers for the microelectronic industry are cut by oscillating wires. Cutting in one direction allows higher wire speeds between 5 and 20 m/s, but yields less planar surfaces. Smoother and more even surfaces are obtained by oscillating sawing. Depending on the cutting speed (the so-called table speed, or speed at which silicon bricks move through the wires, an average of 0.3 mm/min), the wires have a length between 250 and 550 km in order to cut up to four columns in one run. The wire material is usually stainless steel.

Cutting is achieved by an abrasive slurry, which is supplied through nozzles over the wire web and carried by the wire into the sawing channel. The slurry consists of a suspension of hard grinding particles. Today, silicon carbide (SiC) is the most commonly used abrasive. SiC is very expensive and accounts for 25–35% of the total slicing cost. The volume fraction of solid SiC particles can vary between 20 and 60%, and the mean grain size between 5 and 15 µm. For polishing smaller grains sizes below 1 µm are used. The main purpose of the slurry is to transport the abrasive particles to the sawing channels and to the crystal surface. It also has to keep the particles apart and must prevent their agglomeration. The entry of the slurry is a result of the interaction between the wire and the highly viscous slurry. Usually, only a small amount of slurry enters the cutting zone. The two important factors here are the viscosity and the wire speed, but to understand the fluid mechanical problems that are involved a complex physical modelling is required. First attempts of a description have been reported recently [40–43].

Most of the commercial slurries are based on polyethylene glycol (PEG) as the carrying fluid and SiC as the abrasive element. The properties of the slurry are critical for a good quality

cut and parameters such as granulometry distribution, humidity and viscosity are closely monitored by most of the wafer manufacturers. Both SiC and PEG are recycled up to 80% nowadays through different technologies, this being very important both from an economic point of view since very large quantities of slurry are used on every cut, and from the environmental perspective.

Material is continuously removed through the interaction of the SiC particles below the moving wire and the silicon surface. The abrasive action of SiC depends on many factors such as wire speed, force between wire and crystal, solid fraction of SiC in the suspension, viscosity of the suspension, size distribution and shape of the SiC particles. The viscosity of the slurry depends on the temperature and the solid fraction of particles. Since temperature rises as a result of the cutting process, the suspension has to be actively cooled and the temperature controlled during sawing. The viscosity of the slurry also changes because of the continuous abrasion of silicon and iron from the wire. This gradually degrades the abrasive action and the slurry has to be replaced or mixed with new slurry after some time.

The kerf loss and surface quality are determined by the wire diameter, the SiC particle size distribution and the wire transverse vibrations. The amplitude of vibration is mainly sensitive to the wire tension, but it also depends on the damping effect of the slurry. Increasing the tension will reduce the amplitude of vibration, hence the kerf loss [44]. Typical wire diameters are 120 to 140 μm. With the mean size of active particles of 5 to 15 μm, this yields kerf losses around 150–200 μm per wafer.

The objective of efficient sawing is to slice with a high throughput, with a minimum loss of slurry and silicon and with a high quality of the resulting wafers. Since many parameters can be changed, the optimisation of sawing becomes a difficult task, which is nowadays mainly carried out by the wafer manufacturers. They are mostly guided by experience. In the following section, the main results of investigations are summarised, which describe the current understanding of the microscopic details of the wire sawing and yield some guidelines to optimise the process.

6.4.2 Microscopic Process of Wafering

Figure 6.13 shows schematically a cross-section of the wire in the cutting zone. The space between the wire and the crystal surface is filled with slurry and SiC particles. The pressure of the wire on the particles varies along the contact area. The forces are maximal directly below the wire and decrease towards the side faces. Because of the transverse vibrations, the wire also exerts a sideward force, which determines the surface quality of the sliced wafers. The interaction between the abrasive SiC particles and the silicon crystal yields a distinct damage pattern on the surface that

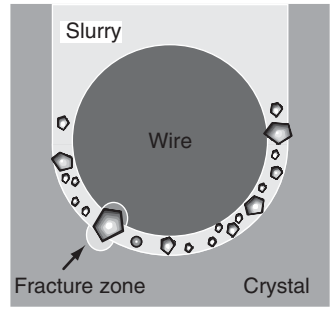


Figure 6.13 Cross-section of wire, slurry with abrasive and crystal in the cutting zone

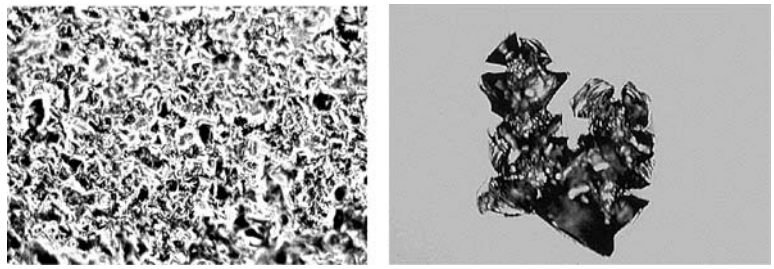


Figure 6.14 (a) Optical micrograph of the surface of an as-cut silicon wafer and (b) several micro-indentations on a polished silicon surface

can be analysed by microscopic techniques. A typical surface structure as seen under an optical microscope is shown in Figure 6.14. Similar structures are obtained along the entire contact zone, which shows that the abrasive process is the same in all directions.

The surface structure consists of local indentations with a mean diameter of a few micrometres. Such a uniform structure can be explained by the interaction of loose, rolling particles that are randomly indented into the crystal surface until small silicon pieces are chipped away. Since SiC particles are facetted and contain sharp edges and tips, they can exert very high local pressures on the surface. This 'rolling grain' model forms the physical basis of the wire sawing process. Similar surface structures also form after lapping semiconductor surfaces with loose abrasive particles.

The individual process of the interaction of a single particle with sharp edges and the surface of a brittle material can be studied by micro-indentation experiments. This is shown in Figure 6.14b for a silicon surface. The damage structure of several overlapping micro-indentations with a Vickers diamond indenter resembles the structure of an as-cut wafer. Numerous micro-indentation experiments on monocrystalline silicon have been carried out in the past to investigate the damage structure quantitatively, e.g. [45–49]. The main results are summarised schematically in Figure 6.15 for a 'sharp' Vickers indenter with pyramid geometry. Loading by sharp indenters first leads to the generation of a remnant plastic impression in the surface known as the elastic–plastic zone. Recent Raman investigations of this region have shown that under high pressures the silicon lattice transforms into other crystal structures. Several phase changes have been observed directly under the indenter, in particular a metallic high-pressure phase [50, 51]. Under loading at 11.8 GPa an endothermic transformation to metallic silicon (Si II) occurs ($\Delta G = 38$ kJ/mol), which partly transforms back to another high-pressure phase (Si III at 9 GPa, $\Delta G = -8.3$ kJ/mol). In the metallic state the silicon can plastically deform and the material can be removed by processes known for ductile metals. This is, however, a slow but moderate process.

With increasing pressure the material begins to break and cracks are generated parallel to the load axis emanating from the plastic zone. Median cracks are generated beneath the plastic zone, where the tensile stresses are maximal, in the form of full or truncated circles. At a critical size they become unstable and extend towards the surface. In addition, shallow radial cracks may be generated at the edges of the plastic zone. Both radial and median cracks may coalesce to form halfpenny-shaped cracks that are visible at the surface (as shown in Figure 6.16). Upon unloading, residual stresses from the elastic–plastic zone can lead to lateral cracks parallel to the surface. When these lateral cracks reach the surface, material is chipped away. This is the main process for material removal during sawing. Chipping requires a certain minimum load to occur (chipping threshold). Above the limit when material is removed only the median and radial cracks remain. They are finally part of the saw damage.

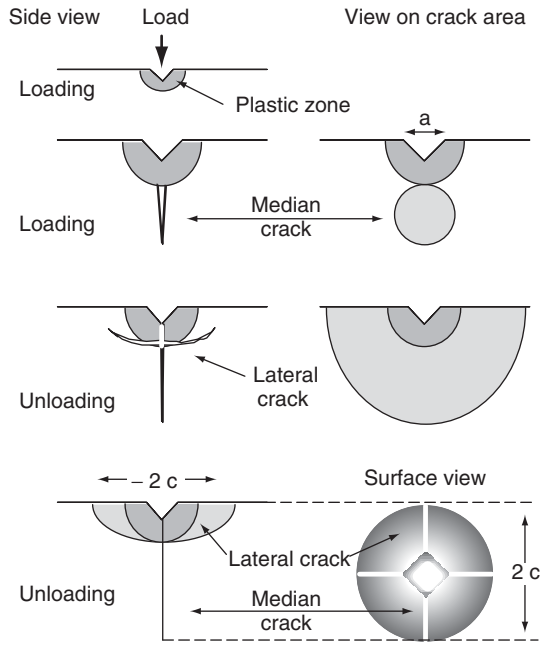


Figure 6.15 Schematic diagram of the development of the crack system below a sharp indenter upon loading and unloading. Dark gray areas indicate the plastic zone below the indenter. The dotted areas are the crack planes of the halfpenny-shaped median crack system. They are viewed from end-on (left side) or perpendicular to the plane (right side). In case radial cracks also occur, they may coalesce with the median crack and form a similar crack pattern

A quantitative model based on the rolling-grain interaction described above has been developed. Results have been compared with experimental investigations of the sawing process on commercial multi-wire saws, allowing for the extraction of useful conclusions. Details can be found in reference [52].

6.4.3 Wafer Quality and Saw Damage

Several factors are currently considered to determine the quality of the wafers: fracture behaviour, crack density, thickness variations, surface roughness and cleanness. After sawing, the surface of the wafer is damaged from the fracture processes and contaminated with organic and inorganic remnants from the slurry. Therefore, the wafers have to be cleaned and the saw damage removed by etching before a solar cell can be fabricated. In addition, the thickness and surface roughness of the wafer may vary, which may be detrimental for some of the further processing steps. All factors are related to the sawing process. Figure 6.17 shows an example of the topology of an as-cut surface. It consists of thickness variations on different length scales. On a scale of millimetres, one can observe grooves parallel to the direction of the wire. They occur particularly under higher loads and can be caused by a deficit of slurry, mechanical vibrations or inhomogeneities of the material. Mostly a large number of parallel wafers are then affected. Grooves on wafers cannot be removed by etching and thus reduce their quality.

On a length scale of about 100 µm, the surface may have a wavy topology that is not detrimental unless sharp steps occur. On the micrometre length scale the surface shows a certain

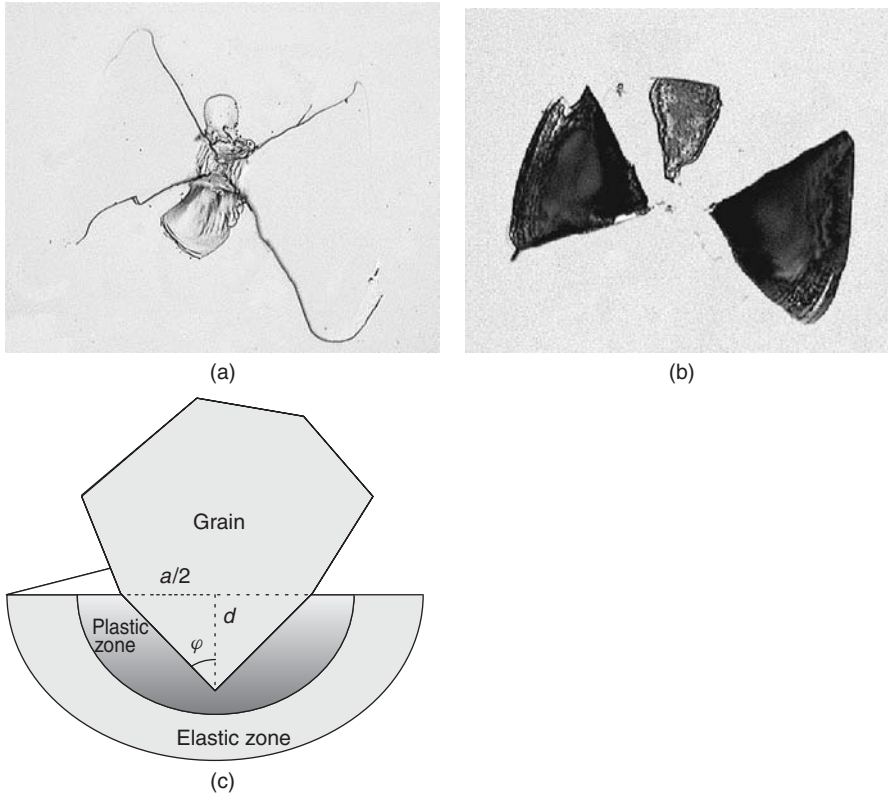


Figure 6.16 (a) Optical micrograph of median and (b) lateral cracks (below the surface) at a Vickers indentation. (c) Schematic representation of the impression of a sharp grain into a surface

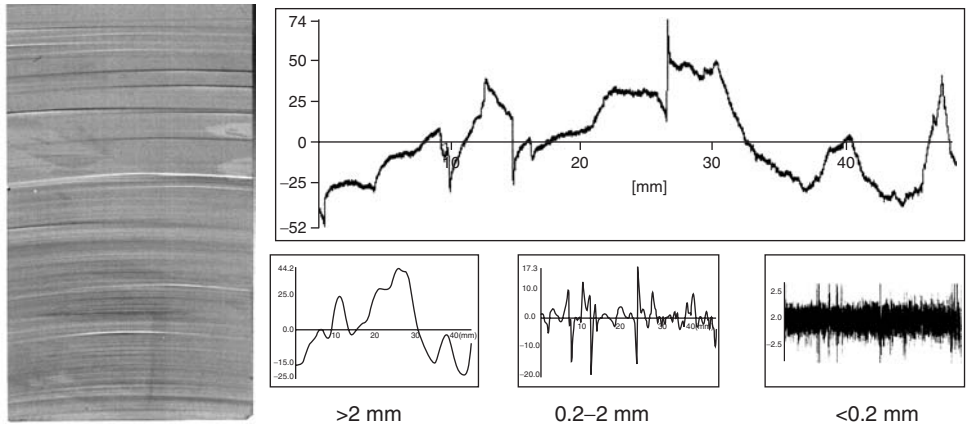


Figure 6.17 Surface structure of a wafer with grooves resulting from uneven cutting. It also shows the bowing of the wire under load during sawing. The surface profile measured by a laser scanner profiler is depicted on the right. Different wavelengths filtered from the profile are shown below

roughness, which is directly related to the microscopic sawing process as described before. The extent of saw damage, which has to be etched away before solar cell processing, lies typically in the range of 5–10 μm.

Saw damage also occurs in the abrasive grains and the wire itself. Although the fracture strength of the SiC particles is higher than that of silicon, the grains eventually lose their sharpness owing to breakage that reduces their sawing performance. To reduce the abrasion of the grains, sawing should be done in a stress range where the load on the individual grains lies above the fracture strength of silicon but below that of SiC.

Typically, for photovoltaic applications the wires have diameters between 120 and 140 μm and a length of about 500 km. They are made of stainless steel and coated with a brass layer. It is important that the thickness is very uniform over the entire length, because sudden changes in the diameter can lead to fracture of the wire or damage to the wafer surface. The abrasion of the steel wires is also due to the interactions with the grains. Excessive wear can lead to breakage, which is undesirable during sawing because it is very time-consuming to build up the wire web inside the machine, and undesirable saw marks might appear on the surface as a result of the wire break in the gap in between wafers. Most of the wire saw manufacturers developed *in situ* detection systems to control the sawing process and thus prevent the wire breakage.

6.4.4 Cost and Size Considerations

The investigations of the microscopic processes of wire sawing have laid the basis for the selection of the best range of parameters and for further modifications. It allows one to increase the sawing performance, to reduce the consumption of slurry, SiC powder, wire material and etchant, and hence directly the costs of slicing. Furthermore, the quality of the wafers concerning roughness, flatness and saw damage of the surfaces can be improved. This is important in view of the development of thinner wafers for solar cells, which will reduce the consumption of expensive silicon. The current sawing technique in production allows the sawing of wafers with thickness down to 180 μm in industrial production and below 100 μm at research level [53]. The goal is to further reduce the thickness below 100 μm. New technologies such as direct film transfer [54] or laser chemical processing [55] demonstrate that sawing of thinner wafers is possible, and through these techniques the mechanical properties of the wafers are improved and the wafers become flexible [54–56]. These technologies and some others are still at R&D level, but show a promising potential. In the next section the most promising ones are described.

6.4.5 New Sawing Technologies

The main issues related to the multi-wire sawing technology are the kerf loss, the expensive slurry consumption, the wafer breakage rates and the challenge to move to wafers below 180 μm thick. Production of wafers as thin as 80 μm by multi-wire sawing has been reported [53], although the mechanical damage created by this technique makes the manipulation of these wafers very challenging and the next steps of the PV value chain would need to adapt their lines to minimize breakage rate.

New slicing techniques are being developed that do not show this behaviour. One example is the direct film transfer technology [54], where a cleavage plane is developed inside the silicon brick at a certain thickness using high-energy hydrogen beam irradiation. Once the cleavage plane is developed, the brick is moved to a cleaving subsystem that uses a two-step initiation–propagation sequence to cleave the wafer from each brick. The surface quality and the mechanical properties of the wafers are excellent since no abrasion system has been used and no kerf loss exists. The energy

consumption at industrial level and the throughput of this technology are the main challenges and it should be noticed that this technology is only applicable to produce monocrystalline silicon wafers.

A different technology also under development to produce thin wafers with no kerf loss is the stress-induced lift-off method [56], where a metal layer is screen printed on top of a silicon material substrate, which is then annealed at high temperature in a belt furnace. Upon cooling down, both the metal layer and the silicon substrate suffer a thermal contraction . Due to their different thermal expansion coefficients, a high-stress field in the substrate is generated, and then the peeling off of a silicon layer attached to the metal layer ocurrs. This metal layer is afterwards removed by metallic chemical etching. This technology is only applicable to produce monocrystalline silicon wafers and high-throughput equipment at high yields is still to be demonstrated.

Laser chemical process on a water media [55] has demonstrated to produce thin wafers with minimal kerf loss, its major problems being the surface quality and the electricity consumption involved.

Finally, promising developments are also taking place on the multi-wire wafering technology, the most important being the introduction of diamond wire slurry-free systems where an iron-based wire is covered by sharp particles of diamond that perform the cut, acting as the abrasive being bonded to the wire by a resin [57]. The main advantages of the diamond wire technology are: the higher throughput (of the order of 2.5 times more than current slurry-based systems), the use of water as coolant and no need of expensive SiC or PEG as well as the improvement of the wafer surface; moreover most of the currently used slicing equipment could be adapted to use the diamond wire without major changes. Also, there is a potential for an easy recycling of the kerf loss created by this technology, thus recovering the silicon related to it. The main challenges are the diamond wire cost, and the difficulty of manufacturing wafers below 100 µm thick.

6.5 SILICON RIBBON AND FOIL PRODUCTION

Research and development on crystal growth technologies for production of crystalline silicon ribbon have been under way for four decades. Interest in methods of crystalline silicon wafer production was initiated during the oil crises of the mid-1970s. Out of this period arose the first large-scale efforts in R&D to develop low-cost methods of producing substrates for solar cell manufacture. A seminal programme was conducted in the US, which was led between 1975 and 1985 by the Jet Propulsion Laboratory (JPL) Flat Plate Array Project [58]. It was the activity in this project in this time period, combined with larger investments from the private sector both in the US and internationally, that developed the seeds of the technology of crystalline silicon ribbon and foil production methods being commercialised today.

R&D on crystalline silicon materials in the 1990s culminated in the expansion of wafer manufacturing at an unprecedented pace. While established methods of production based on Cz growth, directional solidification and ingot casting have flourished, a new generation of ribbon technologies has moved past the R&D stage into large-scale manufacturing and is in competition with these conventional approaches. Ribbon technologies, some of which had already entered R&D in the early 1970s, and reached maturity at the beginning of the 2000s with the start-up of manufacturing on a megawatt (MW) scale, include edge-defined film-fed growth (EFG), string ribbon (STR) and Silicon Film™ (SF). Further technologies like dendritic web (WEB) production and ribbon growth on substrate (RGS) were then moving to pilot demonstration phases. Much has happened in the last few years. A summary of the changes in the status of leading ribbon/foil technologies over the past decades is given in Table 6.2.

Table 6.2 History of R&D and manufacturing status of leading ribbon/foil technologies in recent decades

Wafer process/ year started	1990 status level		2001	2009	Schematic
WEB/1967	R&D	<0.1 MW	Pilot – 0–1 MW	No new data available	Figure 6.19
EFG/1971 (Ribbon) 1988 (Octagon)	Pilot	1.5 MW	Production – 20 MW	All production facilities closed down	Figure 6.20
ESP (STR)/ 1980	R&D		Production <5 MW	170 MW in US; further 400 MW announced in Asia	Figure 6.22
SF/1983			Production >5 MW	No data available	
RGS/1983	R&D		Pilot <1 MW	380 MW announced in the Netherlands	Figure 6.23

Development has not been continuous for most of the methods listed above. The R&D has been interrupted and then restarted in several cases when the technological status changed to generate new opportunities for cost-effective production. EFG development has the longest continuous history. After the initial technology development on EFG started at Tyco Laboratories in 1971, it was subsequently augmented with funding from Mobil Oil, starting in 1974. From 1971 to the present, pilot lines using five different variations of the EFG process have been evaluated, starting with single ribbons in 1971 to the octagonal crystal tube now having been commercialised.

Ownership transferred to ASE Americas in 1994, at which time the transition to manufacturing was initiated. Today, after some years of operation of up to 200 MW, the last technology owner, Schott Solar AG in Germany, has closed down both production sites for EFG in Germany and USA [59]. After periods of decreased activity, WEB, STR and RGS were all strengthened with R&D in the beginning of the millennium, subsequent to being revitalised by new owners. WEB development was initiated with funding from Westinghouse in the 1970s, then carried out by EBARA Solar, which stopped activities by 2004. STR technology underwent an R&D phase in the early 1980s under the name of edge-stabilised ribbon (ESR) and edge-supported pulling (ESP) at the National Renewable Energy Laboratory and at Arthur D. Little, respectively, before being taken up in 1994 by Evergreen Solar, which now has a close to 170 MW capacity factory in the USA, and has recently announced the opening of a 400 MW factory in Asia. RGS development was initiated at Bayer, was then continued by ECN in the Netherlands, and Sunergy and Deutsche Solar joined in to industrialise the project. These two companies founded the joint venture Solwafer, based in the Netherlands. As of 2009, engineering of the first factory was completed, with a 380 MW capacity in six lines, expected to reach full production in 2010, and then expand to 455 MW in 2012 [60].

Ribbon and foil technologies must still meet the challenges of the photovoltaic marketplace and overcome a number of existing technical barriers if they are to continue to expand manufacturing and to position themselves to remain competitive. Challenges to be met are productivity increases on a per furnace basis to drive down labour and overhead (capital) costs, improved mechanical and

electronic quality of ribbon wafers together with the development of low-cost solar cell designs that will raise efficiencies to 18–20%, and reduction of wafer thickness while maintaining high yields in order to reduce demand on silicon feedstock.

Achievement of these goals in the last decade has been incomplete, although some of them are partially achieved. Lab efficiencies of 18.9% in EFG and 17.8% in STR have been obtained [61]. Machine productivity has also been increased [60]. Still, even if absolute production has increased in the last decade as in the rest of the technologies, the market share for ribbon has steadily decreased, as can be seen in Figure 6.18 [1]. In a global situation of feedstock shortage where ribbon has had a good chance of growing, this decreasing share indicates how difficult the challenges are.

Technology description, the status of each of the growth approaches and the barriers for each of the technologies to overcome in order to remain competitive are the topics of the following sections.

6.5.1 Process Description

The ribbon technologies that have been proposed over the past four decades and that have survived until the R&D and commercial manufacturing phases (Table 6.2) may be grouped into two basic approaches: 'vertical' and 'horizontal' growth (pulling) methods. The latter category is further subdivided into methods in which either a substrate is used to assist in the formation of the crystal or foil or those that do not use any substrate material. The vertical or horizontal methods refer not so much to the geometrical aspects related to the ribbon-pulling direction as to the disposition of the temperature gradients, which act at the interface and influence growth characteristics. The EFG, WEB and STR methods are examples of the vertical method category, while both the RGS foil and the SF methods grow crystals in a horizontal pulling configuration with the aid of a substrate. The term 'foil' is used interchangeably with wafer, but here we use it to refer more specifically

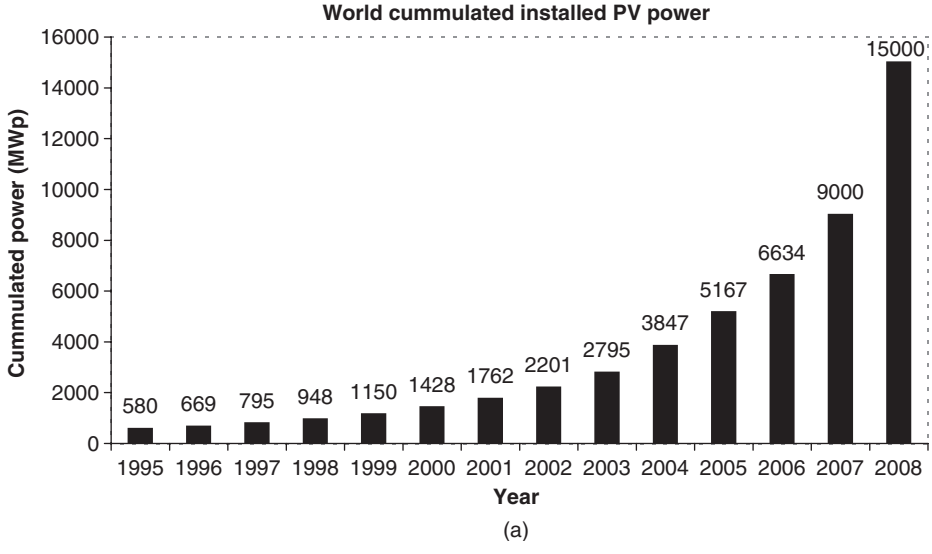
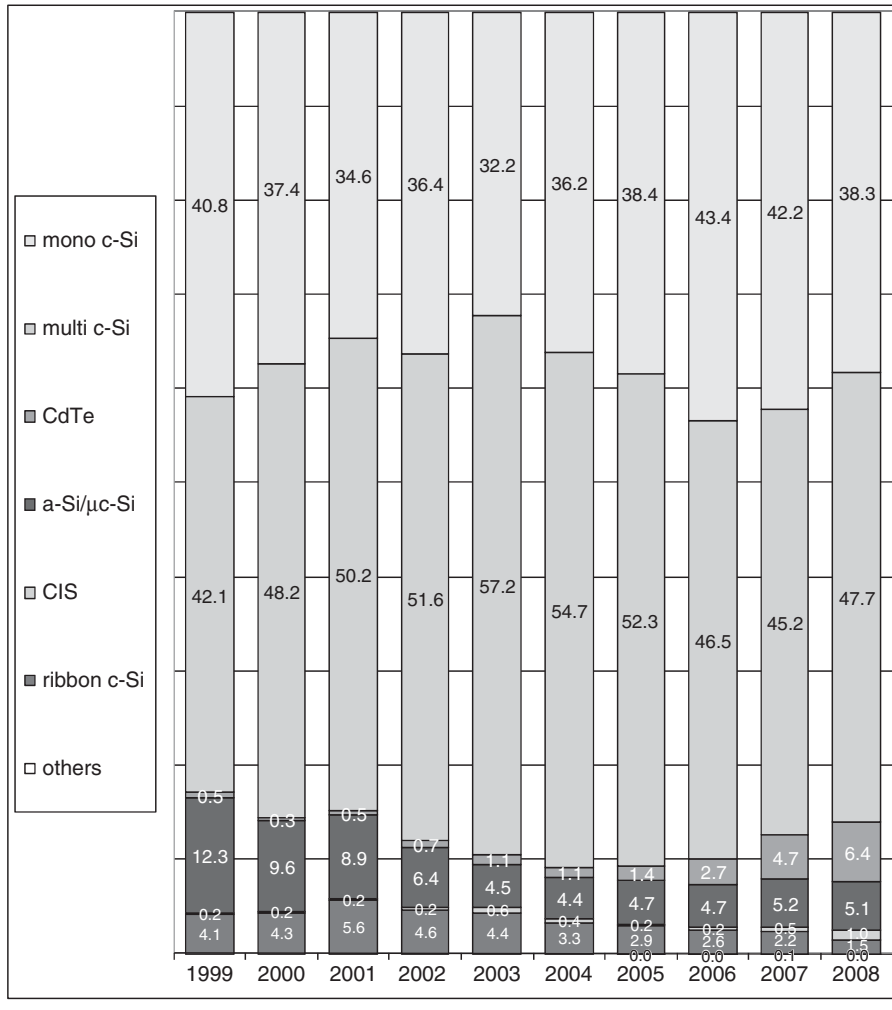


Figure 6.18 (a) PV market growth and (b) technology share in recent years [1]



(b)

Figure 6.18 *(continued)*

to the RGS wafer to distinguish a unique aspect: the wafer is crystallised upon contact with a substrate, and is then detached and the substrate material recycled.

Ribbon/foil growth techniques have historically been evaluated in a number of variants and modifications of the techniques listed in Table 6.2. Successes and failures in many of these variants often spawned new processes or led to evolution and modifications in old variants. A bibliography and descriptions of the techniques and a detailed historical perspective of the many variant ribbon technologies that have been pursued can be found in references [62–64].

Fundamental differences exist in the heat transfer and the interface temperature gradients during growth for these two general categories of ribbon and foil production methods. These lead to very different process limits in several important areas: the capacity, or throughput potential in a single furnace configuration, crystallite or grain nucleation characteristics and the pulling speed.

The speed is constrained as a consequence of the thermoelastic stress acting on the crystal during growth. The pull speed and stress affect the defect density and electronic quality. For example, for the vertical techniques – WEB, EFG and STR – the crystal growth direction and dominant heat transfer of latent heat from the interface are both parallel to the pulling axis of the ribbon and essentially perpendicular to the growth interface. The latent heat conducted along the ribbon is radiated to the environment. The pulling speed and the interface growth velocity are the same. For RGS and SF, crystals nucleate on the substrate and grow nearly perpendicular to the substrate pulling direction, while the growth interface tends to be angled towards the pulling axis of the substrate. Thermal conduction of latent heat from the growth interface is augmented in the direction perpendicular to the pulling axis, that is, through the thickness of the ribbon, because of conduction into the substrate. This augmented heat removal allows very high ribbon production rates, whereby low interface growth rates, v_I, are realised with high pull rates, v_P, that is,

$$v_P = v_I/\cos(\theta)$$

where θ is the angle between the normal to the interface and the pull direction and is close to 90°. The low interface growth rate, in turn, reduces the need for the high interface temperature gradients required to maintain growth stability in vertical ribbon growth. The gradients in the vertical methods are the cause of high thermoelastic stresses and set practical productivity limits when low defect densities and flat ribbon are required. Details on the process limits affecting the horizontal growth techniques may be found in other publications [65, 66]. We next give a description of each of the techniques listed in Table 6.2.

WEB is grown directly from melted silicon in a crucible with no shaping device (Figure 6.19) [67]. A dendritic seed or button is lowered into a supercooled melt. The seed spreads laterally to form a button. When the seed is withdrawn, two secondary dendrites propagate from

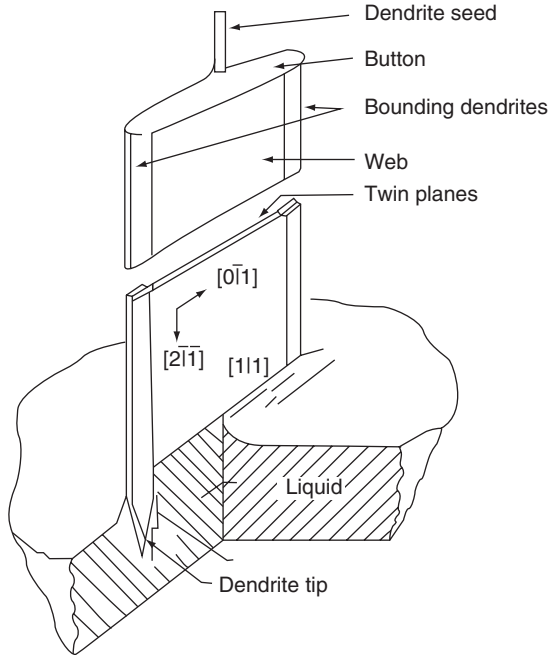


Figure 6.19 Schematic of dendritic web (WEB) growth process for ribbon

the ends of the button into the melt, forming a frame to support the freezing ribbon. The dendrites grow into the melt that has been supercooled by several degrees. Very accurate melt temperature control is required in order to maintain the supercooled interface condition and prevent pull-out, whereby the growth terminates by voiding of the meniscus. The width of the ribbon is controlled by the position of the two dendrites that support the liquid film. The growth velocity is determined by the rate of removal of the latent heat into the ribbon and of the heat conducted through the melt through the meniscus. Typical growth rates are 1–3 cm/min.

In vertical ribbon growth, the meniscus contains the suspended melt volume that connects the bulk melt to the growth interface and crystal. Its shape and the heat conduction taking place within it critically affect impurity segregation and the crystallisation conditions and shape of the crystal. The meniscus height h away from the influence of the dendrites, is fixed by the surface tension, liquid-silicon contact angle and liquid density. This can be calculated from the solution of the Young–Laplace equation requiring a contact angle of $\Theta = 11^\circ$ at the liquid–solid interface that gives

$$h = a[1 - \sin(\Theta)] \quad \text{where} \quad a = (2\sigma/\rho g)^{1/2}$$

where $\sigma = 720$ dyne/cm is the surface tension, $\rho = 2.53$ g/cm^3 is the density of liquid silicon and g is the gravitational acceleration. Because thermal radiation dominates the heat flow from the ribbon, the geometry of the heat shields and the susceptor lid controls the isotherms in the melt and the ribbon.

Accurate temperature control, within a few tenths of a degree, is necessary to ensure a uniform ribbon width and thickness. The temperature of the melt surface must be constant over the width of the growing web to prevent the dendrites from growing in or out. The dominant impurity in the WEB in production today is oxygen since a quartz crucible is used. Typical WEB thicknesses range from 100 to 150 μm and widths up to 8 cm have been grown. Growth lengths between seedings in pilot production extend to many tens of metres.

EFG. In this technique, the geometry of the ribbon is controlled by a slotted graphite die through which silicon is fed via capillary action (Figure 6.20) [68]. A seed crystal is lowered until it contacts the liquid in the capillary. The liquid spreads out over the top of the die to the edges where it is pinned by surface tension. The seed is withdrawn, pulling the liquid up while more liquid flows upward through the capillary. As the ribbon is withdrawn, the liquid freezes on the solid crystal. The die and the crucible are integral, that is, made of the same piece of graphite. The thickness of the sheet material is fixed by the width of the die top, distance between the die tip and melt level, meniscus shape, heat loss from the sheet and the pull rate. The shape of the liquid–gas interface, or meniscus, which connects the die to the solidifying ribbon, is described by the Young–Laplace or the capillary equation. As with WEB, the growth rate is controlled by how fast heat can be conducted away from the interface and lost by radiation or convection from the solid crystal. Growth is self-stabilising because the meniscus height increases with an increase in pull rate. The curvature of the meniscus causes the thickness of the crystal to decrease. This increases the rate of heat removal per unit area of the interface, thus increasing the growth rate until it is again equal to the pull rate.

The dominant impurity in EFG ribbon is carbon, which is in supersaturation. Temperature control of a few degrees along the interface is sufficient to prevent ribbon pull-out or freezing of a growing ribbon to the die top. Over time, the die becomes eroded, affecting ribbon properties and leading to a nonuniform ribbon thickness and growth difficulties.

Ribbons with thicknesses from 400 μm to as little as 100 μm have been grown. Rather than a single flat ribbon, hollow EFG polygons are grown to enhance the rate of throughput. The favoured geometries for commercial development today are octagons with 10- or 12.5-cm-wide

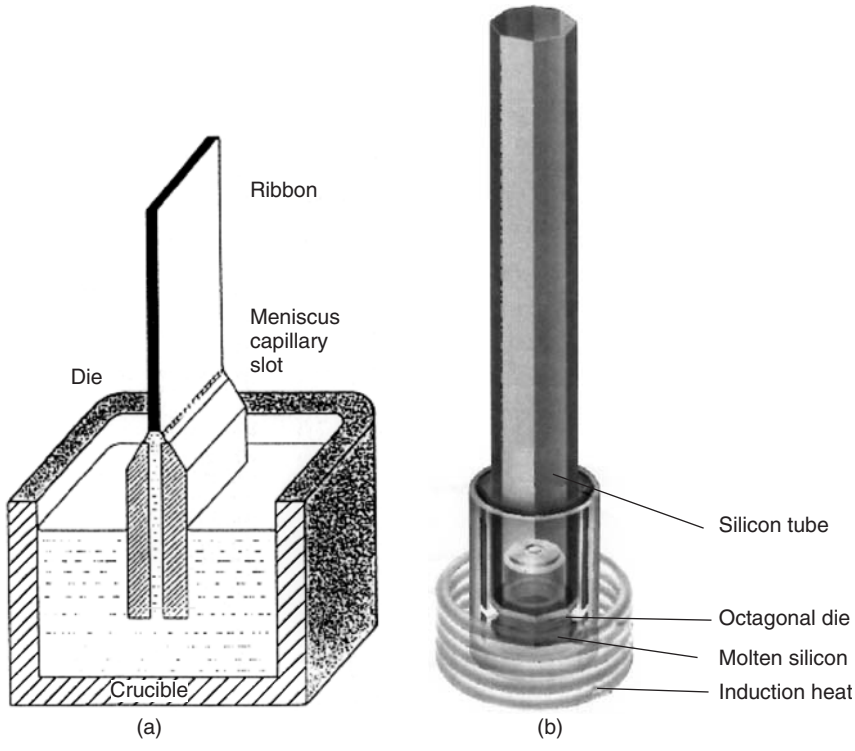


Figure 6.20 Schematics of edge-defined film-fed growth (EFG) growth process: (a) ribbon die and crucible configuration; (b) octagon configuration

faces, equivalent to growth of up to a 100-cm-wide ribbon from a single furnace. Various closed geometries, including nonagons with 5-cm faces, and large-diameter cylinders have been grown. Growth velocities for the EFG octagon are 1.7 cm/min. The relationships between the EFG process parameters and the silicon ribbon characteristics, including thermal stress and the influence of impurities and defects on the quality of the material, have been extensively examined and are reviewed in reference [69].

An extension of the EFG process to growth of 50-cm-diameter cylinders has been demonstrated [70]. An example of such a cylinder 1.2 m in length is shown in Figure 6.21. The cylindrical geometry offers some relief from the large thermoelastic stresses generated in plane ribbon. This allows consideration of higher-productivity furnaces from a combination of larger perimeters and potentially higher growth speeds. Growth of EFG cylinders with average wall thickness down to 100 μm has been demonstrated, and solar cells have been made on this material [71].

STR. In this technique, ribbon growth takes place directly from a pool of molten silicon without a die (Figure 6.22) in a process mirroring the WEB geometrically. Rather than dendrites, as with WEB, the position of the ribbon edges in STR is maintained by two strings fed through holes in the bottom of the crucible. The strings are drawn upward out of the melt to support the meniscus and the ribbon, and their pull rate determines the growth speed of the ribbon. The thickness of the ribbon is controlled by surface tension, heat loss from the sheet and pull rate. An important difference of the STR process from WEB growth is that the constraints of maintaining propagating

Figure 6.21 Experimental 50-cm-diameter EFG cylinder exiting from furnace

dendrites and a supercooled melt are eliminated, and this relaxes the high degree of temperature control required in the WEB furnace. The high meniscus, 7 mm, allows simple control of the growth process and maximises its stability to mechanical and thermal fluctuations. Depending upon the wetting qualities of the strings and their diameter, the meniscus height at the strings near the edges differs from that at the centre, and it is usually much lower at the edges [72, 73].

For comparable thicknesses, the growth velocities of STR are similar to EFG and WEB. Careful adjustment of the growth parameters can allow very thin ribbon, down to 5 μm, to be grown [74]. Generally passive after-heaters are used, but some work on an active after-heater has been carried out to allow low-stress, 100-μm-thick ribbon to be grown. This material was sufficiently flat to be made into solar cells. Because of the concave downward meniscus curvature at the string, any grains nucleated at the strings can propagate into the ribbon.

RGS. In this growth technique, the silicon melt reservoir and die are placed in close proximity to the top surface of a substrate, on which the ribbon/foil grows. The substrate may be graphite or ceramic (Figure 6.23) [75]. The principle is to have a large wedge-shaped crystallisation front.

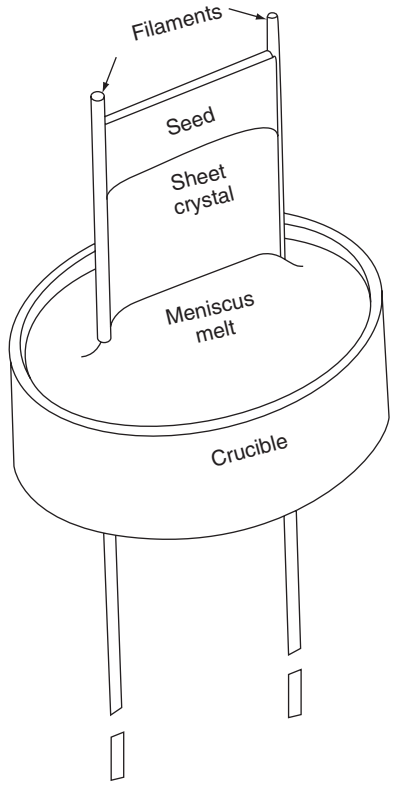


Figure 6.22 Schematic of string ribbon (STR) growth system

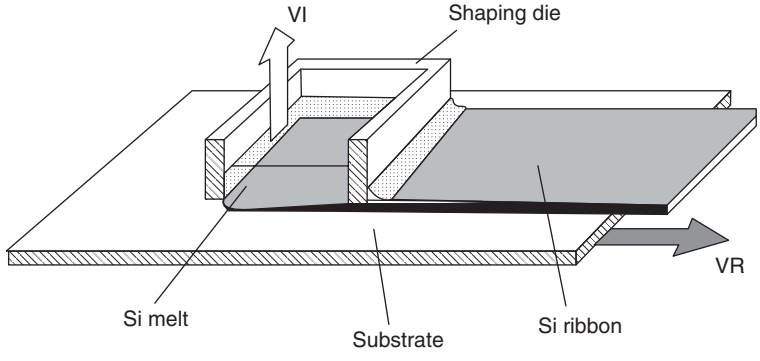


Figure 6.23 Schematic of ribbon growth on substrate (RGS) configuration

The die contains the melt and acts to fix the width of the foil. The thickness of the foil is controlled by the heat-removal capacity of the substrate, pull rate and surface tension. The direction of crystallisation and growth are nearly perpendicular. The area of the growth interface now can be very large compared with the foil thickness. The latent heat is extracted by conduction into the substrate. The thermal gradients near the interface are small, thus reducing thermally induced stress in the

wafer. Growth rates from 4 to 9 m/min have been demonstrated. One example was an 8.6-cm-wide foil, 300 μm thick, grown at 6.5 m/min [76].

An important goal in the R&D phase of RGS has been to make a substrate that can be reused. After cooling, the silicon foil may be separated from the substrate by stresses arising from differences in thermal expansion between substrate and silicon. Thicknesses between 100 and 500 μm have been grown. By working with the lower thermal gradients in the foil thickness direction, but still large enough for rapid growth, fluctuations in the pulling speed and gradient only affect the foil thickness slightly [75].

SF. The details of the SF process are proprietary. The silicon crystal is grown in a thin layer directly upon either an insulating or a conducting substrate, with a barrier layer that promotes nucleation [77]. In the case of an insulating substrate, the barrier layer must also act as a conductor to collect the current generated in the cell. In the case of a conducting substrate, the substrate can also act as an electrical conductor if *vias*, or holes, are provided to connect the thin silicon crystal layer and the substrate. The SF thin film and barrier layer do not separate from the substrate on cooling as in RGS, but become the active part of the solar cell. The grown multicrystalline silicon layer is made very thin ($\ll$100 μm), thus reducing the amount of silicon required. Currently, layers of 20 μm thickness are under development. A variety of substrate materials have been used, including steel, ceramics and graphite cloth [77, 78]. It is necessary that the barrier layer prevent the transport of impurities from the substrate into the silicon. The barrier layer allows wetting and nucleation during growth. It should also act to electrically passivate the back surface and have a high optical reflectivity.

An insulating barrier layer has been reported that promotes growth of large columnar grains (greater than 1 mm) through the thickness of the grown SF silicon film [77]. As-grown films on coated ceramic substrates exhibit very low diffusion lengths of less than 10 μm. The new barrier layer and the substrate result in longer diffusion lengths, 20–40 μm, and the silicon has an improved response to phosphorus gettering.

6.5.2 Productivity Comparisons

Ribbon crystal growth technology for production of silicon wafers has been historically faced with the evaluation of the trade-off between bulk electronic quality and throughput (productivity per furnace). The choice of crystal growth conditions is made on the basis of wafer cost parameters and the premium imposed by the marketplace on solar cell efficiency. Material quality that can translate into high solar cell efficiencies has always been a primary market driver guiding ribbon growth process development. Ribbon wafers have inherently lower wafer production costs than those obtained from directionally solidified and cast ingots, or Cz boules, because ribbon growth avoids the large material losses due to sawing, which are as high as 50% of the starting feedstock. However, by the recent reduction in polysilicon prices, the feedstock saving advantage becomes less important. If process rates are compared with commercial cast and sawn wafers, as seen in Table 6.3, and assuming growth in directional solidification systems (DSS) is the limiting step for this process (wire sawing parameters are also shown for comparison), only RGS has a higher throughput, therefore also a process rate advantage. Not being the limiting step in the commercial casting and sawing technology, also wire sawing at 1550 cm^2/min is much faster than all the ribbon technologies, except RGS. Considering the higher bulk electronic quality, and higher solar cell and module efficiencies for cast and sawn wafers, ribbon technologies are nowadays probably more than ever faced with severe competition. But their asset is the better use of the silicon as it is not subject to kerf losses.

A summary of performance metrics for ribbon technologies currently under development is given in Table 6.3. A critical driving parameter in technology development of ribbon methods for

Table 6.3 Single-furnace performance metrics for ribbon technologies under development and in commercialisation. Comparison with usual wire sawing and DSS growth is given

Method/parameter	Pull speed [cm/min]	Width [cm]	Throughput [cm^2/min]	Furnaces per 100 MW[a]
WEB	1–2	5–8	5–16	2000
EFG Octagon	1.65	8 × 15.6	165	100
STR	1–2	8	5–16	900
SF	[b]	15–30	[b]	[b]
RGS	600–1000	15.6	7500–12 500	2
DSS Growth (25 wafers/cm)[c]	0.0125	15.6	1901	16
Wire saw (4 × 500 wires)	100[d]	15.6	1550	10

[a]Furnace data are taken from References [79–81] and [60], where throughput is normalized for comparison purposes to an overall yield of 90% and cell efficiency of 15% for all processes.
[b]Pulling rate parameters for the SF process are not available.
[c]For comparison, *equivalent* 'pull speed' and 'throughput' data for usual DSS growth; the pull rate is small, but the area is very large. To convert volume growth rate into wafer surface the number of wafers per cm in the ingot (25) must be given.
[d]This surface is cut by a wire saw cutting 2000 wafers at the same time (in four bricks). The wire is actually a single one wound 500 times around four bricks.

large-scale manufacture and commercialisation is the productivity per furnace. Productivity governs the capital cost of installed capacity and direct labour costs, which constitute significant barriers to ongoing commercialisation on a large scale for all ribbon technologies.

6.5.3 Manufacturing Technology

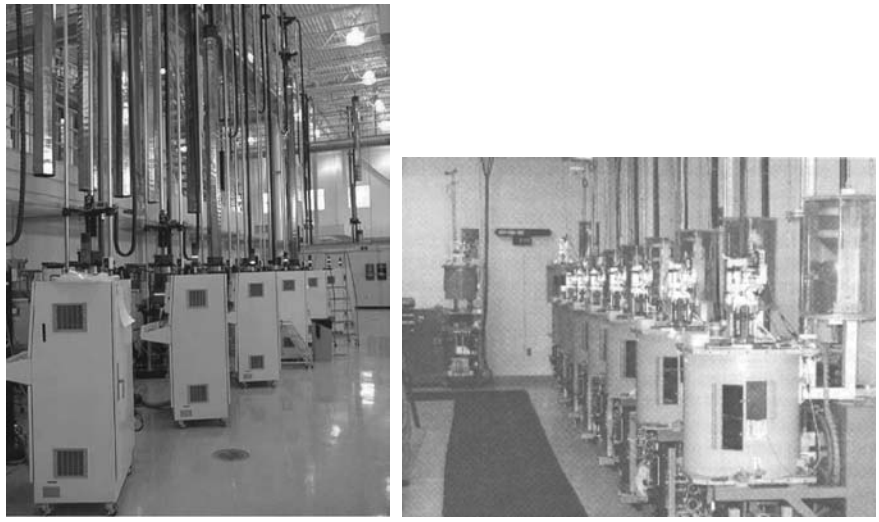
Table 6.3 shows that the development of ribbon technologies is proceeding along two distinct paths. WEB and STR rely on low furnace cost to remain competitive in wafer costs. Development of SF and RGS technologies is focused on achieving superior throughput per furnace.

Scaling of ribbon factories significantly beyond the manufacturing levels practised now, for example, 1 GW, poses different challenges for technology development in these two cases. The low-throughput ribbon furnace (WEB, STR) requires a simple and low-cost furnace design, a high level of automation and low infrastructure costs. In contrast, high reliability and uptime of furnaces and the growth process are most critical for the high-throughput technologies SF and RGS. EFG technology development is moving in a direction that is trending towards the middle of these two extremes.

Examples of commercial installations and equipment now in manufacturing for EFG and STR technologies are shown in Figure 6.24. Figure 6.24a pictures a group of furnaces in the EFG octagon manufacturing line, with a high bay area to accommodate the 5.4 m octagon growth lengths. EFG wafers of 15.6 × 15.6 cm are standard products and are cut from the octagon tubes using high-speed lasers (not shown). More detail on this technology is given in references [59, 82].

Single-ribbon furnaces for the growth of 8-cm-wide ribbon of the STR manufacturing line are shown in Figure 6.24b. Ribbon sections up to a meter long are scribed from the ribbon while it is growing, and then further cut into wafers prior to processing.

WEB and SF ribbon technologies are in various stages of R&D and pilot demonstrations leading to commercialisation. Little public information has been released in recent years. At the beginning of the 2000s, SF was perhaps the closest to successfully scaling up the technology, as a wider (20 cm) and higher throughput furnace was reaching the final demonstration phase. WEB



(a) (b)

Figure 6.24 Manufacturing crystal growth equipment in commercialisation for: (a) EFG; (b) STR technology

was basing its expansion to a 1- to 2-MW pilot operation on a single ribbon furnace for producing 5-cm-wide wafers. No further data have been disclosed in the last years.

RGS moved towards the use of 15.6-cm-wide ribbon in which high sustained throughput technology and consistent high-quality material is now to be proven industrially with the new 380 MW factory, which should be operative by the end of 2010.

6.5.4 Ribbon Material Properties and Solar Cells

Except for WEB, all the growth processes produce multicrystalline ribbon. WEB ribbon grows with (111) crystallographic faces (see Figure 6.21). It typically does not have any grain boundaries, but has a single multiple-twin boundary located about midway through the ribbon thickness. Each (111) surface is made up of a single grain, and the dislocation density is the lowest of any ribbon, $10^3 - 10^4/cm^2$.

For the other two vertical ribbon growth cases, EFG and STR, extraneous crystals are generated most often at the sides of the ribbon (i.e. octagon tube corners) and propagate along the growth axis of the ribbon. These crystals form elongated grains, often many centimetres in length along the growth axis, and which extend through the ribbon thickness. In EFG, the grains are interspersed with numerous twin boundary arrays. The grains at the ribbon edge in STR are generally smaller than in the centre. Because the meniscus near each string is concave downward, the grains nucleated at the string can propagate into the ribbon centre. In both EFG and STR, the grain dimensions typically are large compared with the ribbon thickness and the as-grown diffusion length, and the charge collection and solar cell efficiency are minimally influenced by grain boundary recombination.

For SF and RGS, the substrates provide the dominant nucleation sites for crystals. The grains are usually columnar, and extend through the ribbon thickness with dimensions that can be made large compared with the diffusion length with proper adjustments of the pulling speed and interface inclination.

Table 6.4 Comparison of silicon ribbon material characteristics. In all cases the columnar grains extend through the thickness of the ribbon. An equilibrium segregation coefficient of $k_0 \sim 10^{-5}$ is typical of the most detrimental impurities for ribbon bulk lifetime

Material	Crystallinity	Dislocation density[1/cm^2]	Effective segregation	Thickness [μm]
EFG	Columnar grains in growth direction	10^5–10^6	$k_0 < k_{\text{eff}} < 10^{-3}$	250–350
WEB	Single (111) face central twin planes	10^4–10^5	$k_0 < k_{\text{eff}} < 10^{-3}$	75–150
STR	Columnar grains in growth direction	5×10^5	$k_0 < k_{\text{eff}} < 10^{-3}$	100–300
SF	Columnar grains through thickness	10^4 to 10^5	$k_{\text{eff}} < 1$	50–100
RGS	Columnar grains through thickness	10^5–10^7	$k_{\text{eff}} < 1$	100–500

Fast movement of the solid–liquid interphase reduces the ability of the freezing process to segregate impurities to the melt. This is represented in Table 6.4 together with the crystalline aspect and the dislocation density of the different technologies.

Stresses produced by thermal gradients during growth generate most of the intragranular dislocations in ribbon material. In the best-quality EFG material, it has been shown that the loss in background current is highly correlated with the dislocation density and not grain boundaries [83]. It is not clear if the intrinsic qualities of the dislocations act as recombination centres or if the associated impurity cloud is responsible. Dislocations decorated with SiO_x precipitates have been reported to limit the lifetime in WEB [84] and RGS material [85]. Photoluminescence studies suggest similar causes for dislocation recombination activity in EFG material [86].

A critical parameter for solar cell efficiency is the ribbon thickness. As shown by Bowler and Wolf [87], an optimum thickness for peak efficiency occurs. This thickness is dependent on the fabrication technique and material properties, including front and back surface recombination velocities, minority-carrier lifetime and base resistivity among other parameters. For a 'typical' n^+pp^+ structure with a 200-μm diffusion length, L_d, a back surface field and a single layer antireflection coating, the optimum thickness region consists of a broad peak near 80 μm (Figure 6.25) as calculated using PC-1D [88]. For a 100-μm L_d the optimum thickness peaks at about 50 μm and for a 400-μm L_d it is near 120 μm. A front surface recombination velocity of 10^5 cm/s is assumed.

The thickness of WEB and SF is closer to the optimum than the other three ribbon growth techniques. Light trapping will move the optimum to a thinner base, much research has been done on optical structures in the last decade, see for instance reference [89]. For growth on a substrate, it is possible to texture the substrate to trap light. This has been shown on SF wafers [78].

Solar cell fabrication processes used for conventional Cz and cast material wafers are commonly applied to silicon ribbon. If the ribbon is doped p-type with boron, the n-type emitter typically is formed by phosphorus diffusion, either from $POCl_3$, PH_3 or a spin- or spray-on source. Front contacts are usually screen printed or evaporated with diffused or alloyed aluminium as back contact to produce a back surface field. A double- or single-layer antireflection coating may be used. For SF grown on an insulating substrate, contact with the back surface requires etching holes to allow contact with the conducing barrier layer. Table 6.5 summarises material characteristics and solar cell performance potential of the various forms of ribbon material. Boron dopes the ribbon p-type. Antimony can be used to produce n-type material as reported for WEB.

Ribbon cell processing, nevertheless, needs to recognise the unique growth constraints of all of the techniques. As noted above, a common material characteristic for ribbon materials historically has been the compromised as-grown bulk electronic quality, dictated by the development of

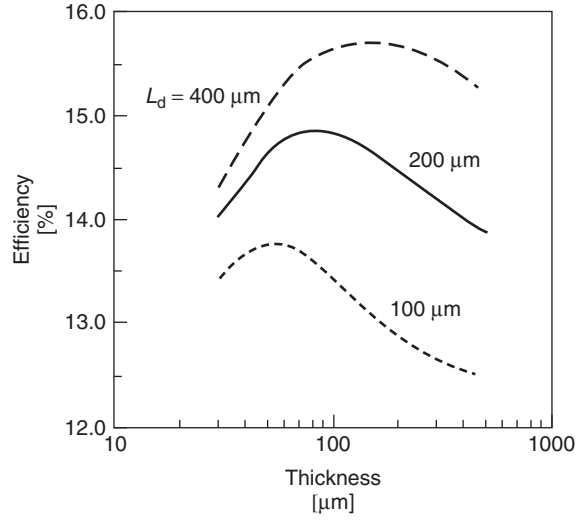


Figure 6.25 Solar cell efficiency versus thickness (see the text for a description of the solar cell parameters)

Table 6.5 Some 'best' solar cell efficiency levels for various ribbon technologies

Material (reference)	Resistivity [Ω cm]	Carbon [1/cm^3]	Oxygen [1/cm^3]	Efficiency [%]
EFG [61, 90]	2–4, *p*-type	10^{18}	$<5 \times 10^{16}$	18.9
WEB	5–30, *n*-type	Not detected	10^{18}	17.3
STR [61]	1–3, *p*-type	4×10^{17}	$<5 \times 10^{16}$	17.8
SF [91]	1–3, *p*-type	5×10^{17}	5×10^{17}	16.6
RGS [92]	2, *p*-type	10^{18}	2×10^{18}	14.4

desired high-throughput growth configurations. This strategy has its base in a commercial demand for very-low-cost wafers that must compete at relatively low volumes with already established dominant products based on single-crystal Cz wafers or directionally solidified and cast ingots. As-grown ribbon diffusion lengths most often are less than 100 μm. To obtain the maximum cell efficiency on this ribbon, with higher dislocation densities and contaminating impurities than in competitor wafers, a solar cell processing strategy was devised early in the history of ribbon technology development that incorporates special bulk lifetime upgrading steps. For example, bulk lifetime upgrading via aluminium alloying and hydrogen is particularly effective for EFG material [70, 93, 94]. Another approach is to use plasma-enhanced chemical vapour deposition (PECVD) of silicon nitride to generate hydrogen for passivating the silicon bulk [95]. This technique has become common in recent years, also in conventional solar cell processing, with Cz or cast material.

6.5.5 Ribbon/Foil Technology: Future Directions

Ribbon/foil wafer production is poised to move on to face a new round of challenges in the construction of large (500–1000 MW) manufacturing facilities for crystalline silicon ribbon wafers.

The RGS foil technology, with the greatest potential of all ribbon methods for cost reduction on the basis of a high throughput per furnace, is entering the industrial phase, with a first 380 MW in six lines, which should be operative at the end of 2010. This process has overcome some challenges in process and equipment development before it could enter high-volume manufacturing of wafers. Nowadays, RGS claims to have a consistent material quality, sufficient for improving cell efficiency to greater than 13%, process control capable of reproducibly producing a low-stress, regular structure, with a shaped 15.6 cm-wide wafer suitable for high-yield cell processing at between 150 and 300 μm thickness and a reliable production furnace with melt replenishment to enable continuous production.

The future focus of WEB ribbon development is still on process automation and capital cost reduction for furnaces and infrastructure based on a concept of a low-throughput (per furnace) process. There are no figures of its evolution in recent years, therefore prediction for future capacities is uncertain. STR has expanded its manufacturing with a 300-μm-thick wafer. Although wafer bulk quality is demonstrated on an R&D level to be capable of achieving almost 18% cell efficiencies for STR, quality and cell efficiency levels on a multi-megawatt scale are yet to be established with the new 170 MW factory started in 2008. This approach will attempt to demonstrate the cost-effectiveness of operating on a multi-MW level in the next few years. R&D directions, which would appear to have the most potential for the reduction of wafer material costs for these techniques, are growth of wider ribbons and more ribbons per furnace.

EFG and SF ribbon technology successfully completed their initial scale-up of wafer production to the multi-megawatt level, although EFG factories recently closed down and SF technology has kept silent in the recent year. Process control and equipment reliability improvements, which can drive throughput and yield higher, become increasingly more dominant in determining the manufacturing cost. As throughput per furnace increases, capital cost impact on variable manufacturing costs from the growth furnace decreases. There is pressure on all ribbon technologies to concentrate on reducing the capital cost of the wafer production equipment if the transition to large-scale wafer manufacturing of 500–1000 MW annual capacity factories is to be sustained.

The benefits of the savings in silicon feedstock and potential gains in cell efficiency with a reduction of the ribbon/foil thickness are well understood for all these technologies. However, the pressure to carry out R&D in this direction for the case of wafers made from ribbons is not as acute as for conventional crystalline silicon wafer manufacturing methods because of the large benefit in feedstock savings already realised for ribbons on account of their favourable geometry. The R&D for the next generation of vertical ribbon technology beyond about five years will target the demonstration of production methods for very thin wafers. A strong motivator driving thickness reduction will be the pressure to increase the cell efficiency, which is seen to be capped in the 17–19% range (see Table 6.5) for current cell designs and wafer bulk quality. Low-cost cell designs, which can break this barrier and achieve desired targets of 18–20% for ribbon, are most easily found for thinner wafers, but this also requires improvements in bulk electronic quality to be achieved concurrently.

The major problem in this development for all vertical growth techniques will be to find methods to reduce the effects of thermal stress. At present, the only means by which this can be done is to reduce the pull speed. The cylinder geometry has the potential to offer some relief to the EFG process at the expense of having to work with thin curved wafers in cell and module processing. Although thermal stress is not a problem with substrate-assisted growth techniques, there probably will be a trade-off between good bulk quality with large grains and throughput.

6.6 NUMERICAL SIMULATIONS OF CRYSTAL GROWTH TECHNIQUES

Commercial finite element simulation tools for structural analysis in computer-aided engineering started to develop at the beginning of the 1970s. Today, simulation tools are an essential part in various industry productions; e.g. crash test simulation for automobile development or airflow simulations in the aerospace industry. As an advanced application the descriptions of whole production processes are the goal of the strategies for simulations. If these strategies are successful, computer modelling opens the opportunity to shorten development time for production facilities, to reduce the costs for the engineering and to speed up process optimisation. In this chapter we will report on those simulation tools, various thermal models and examples of numerical simulations of silicon crystallisation processes.

6.6.1 Simulation Tools

Numerical simulation tools can be distinguished in universal and special-purpose programmes. Examples of commercial universal-purpose programmes are ABAQUS [96], ANSYS [97] and MARC [98] with a wide range of applications in structural analysis, thermal and fluid-flow simulations or electromagnetic field simulations. The number of special-purpose programmes cannot be estimated seriously. Many universities and companies are working with specially developed software tools to obtain solutions of their specific problems. Recently, the large commercial programmes both compromise and enable the user to add their specific subroutines to a programme run.

The main structure of most of the simulation tools is similar: a pre-processing is designed to define the initial and boundary conditions of a simulation run and includes the generation of the simulation domain (finite element mesh) as well as a set of physical data that describes the material properties. The main processing is normally not interactive and contains the solver of the mathematical formulations. The post-processing visualises the simulation results.

The demand for simulation tools depends on the complexity of the physical problem or on the technical process that the user wants to simulate. In general, the description of all physical relationships is reachable only in relatively simple and well-known problems. The full description of an industrial production facility by numerical simulations is not possible today, and neither is it reachable in the near future because too many details are too complex to be described by the numerical models. Therefore, the development of useful simplifications is one of the important keys to a successful simulation. This requires an integrated teamwork between the user of a simulation tool and the operators at the production facility and other process specialists.

Another important requirement is the validation of simulation results by experimental data. At least two experiments are necessary to validate simulation results concerning the process behaviour of a production facility. This means that the simulation model should be validated by measurements during a standard process and in a worst-case scenario to ensure the correctness of the results in an enlarged area of validity. Normally, these experiments are expensive and difficult to realise during a running production, but otherwise, running a nonoptimised production would be quite more expensive. Anyhow, the validation of simulation results is necessary to ensure the success of the simulation method.

6.6.2 Thermal Modelling of Silicon Crystallisation Techniques

The wafer material for crystalline silicon solar cells can be divided into those from ribbon and bulk crystals. For most of wafer production processes, numerical simulations are in use to describe

the thermal conditions during the crystallisation. In the case of ribbon crystals, the EFG [99, 100] and the STR process [101], and recently RGS, have reached a market production. The Cz crystal-pulling technique is the standard process for microelectronic single-crystal wafers and covers an essential part of the PV market share [102, 103]. The characteristics of ingot crystallisation can be explained by the shape of their liquid–solid interface. Anyhow, today's ingot crystallisation is tending more and more towards a mostly planar solidification. For the use of numerical simulations of the cold wall process, see [104, 105]; for the heat exchange method (HEM), see [106] and for the solidification by planar interface (SOPLIN) processes, see [107, 108].

To simulate the temperature history during crystallisation, various thermal effects must be taken into account. In Figure 6.26 the scheme of thermal conditions for the ribbon growth and ingot crystallisation is presented. The biggest difference between the two is the strong variation of the cooled surface-to-volume relation (SV) during crystal growth. This relationship can be used to qualify the cooling behaviour of the different crystals in an equivalent surrounding. For ribbon growth SV is given as 2/ribbon thickness and for ingots as 1/ingot height. The high number for ribbon growth (e.g. SV = 66/cm) means that the surface affects the crystallisation, while the low number for an ingot geometry (e.g. SV = 0.033/cm) shows that volume effects are more important for crystallisation. By this, the SV parameter characterises the requirements for the modelling of different crystallisation techniques. In the case of bulk crystallisation, the latent heat at the liquid–solid interface must be led away by a heat sink at the bottom of the ingot. By this, the crystallisation is propagated by a conductive heat flow through the solid ingot volume, and the temperature gradients inside the volume have to be simulated with great care. In the case of ribbon growth, heat flow by convection and radiation at the silicon surface is the dominant heat transport mechanism to lead away latent heat and propagate the solidification. Therefore, simulation results are very sensitive to heat transition coefficients and the emission behaviour at the ribbon surface.

Furthermore, both techniques can be distinguished into quasi-steady state and moving boundary processes. Assuming a constant pulling speed, the ribbon growth is characterised by a steady-state temperature field, and the liquid–solid phase boundary can be modelled by a fixed-temperature boundary condition, at the silicon melting temperature of 1410 °C. In the case of the ingot casting, the phase boundary moves through the crystal and the release of latent heat can be modelled by an enthalpy formulation. By this, the release of the latent heat of finite elements can be taken into account directly after their total solidification, or the fraction of latent heat must be considered for partly solidified elements [109, 110]. The importance of an accurate modelling

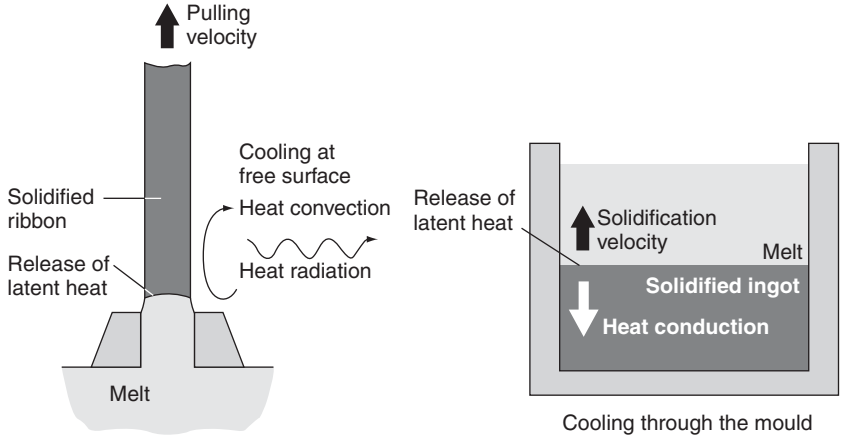


Figure 6.26 Scheme of thermal effects during ribbon growth and ingot crystallisation

of the release of latent heat may become more clear by estimating the crystallised volume rate in typical ingot processes to be around 9000 cm^3 per hour, which means a latent heat source of more than 8 kW at the location of the phase boundary. Pulling one 10-cm-wide ribbon, the crystallised volume rate is about 30 cm^3 per hour with a latent heat release of 0.03 kW.

The Czochralski pulling technique can be classified as ribbon or ingot crystallisation. The temperature profile can be assumed to be stationary and the SV parameter, given as 1/crystal diameter, lies, for example, in the range of 0.066/cm.

In general, the heat flow in silicon during crystallisation can be described by the heat transport equation [111–113]:

$$\rho c_p \frac{\partial T}{\partial t} = \lambda \nabla^2 T + L \frac{\partial f_c}{\partial t}$$

with the silicon data:

Density of solid silicon	$\rho_{(1410\,^\circ C)}$	=	2.30 [g/cm^3]
Density of liquid silicon	$\rho_{(1411\,^\circ C)}$	=	2.53 [g/cm^3]
Heat capacity	$c_{p(20\,^\circ C)}$	=	0.83 [J/g K]
	$c_{p(1410\,^\circ C)}$	=	1.03 [J/g K]
Heat conductivity	$\lambda_{(20\,^\circ C)}$	=	1.68 [W/cm K]
	$\lambda_{(1410\,^\circ C)}$	=	0.31 [W/cm K]
Latent heat of phase change	L	=	3300 [J/cm^3]

The time t and the temperature T are variable and result from the simulation. For the moving boundary case the solid fraction f_c becomes important. This parameter depends on the finite element temperature and varies between zero for a completely liquid finite element and one for a solid element.

Additionally to these material properties and the heat flow mechanisms in the silicon material, the description of the internal furnace construction must be taken into account to perform simulations of crystallisation facilities. This includes the geometrical description and material properties of the internal set-up as well as the radiative heat exchange with heaters and cooling facilities.

6.6.3 Simulation of Bulk Silicon Crystallisation

As an example of the temperature simulation of silicon ingot crystallisation, the SOPLIN casting technique is selected. To simulate this process, a finite element mesh of about 230 000 elements was built to describe the furnace geometry. This mesh includes the silicon ingot, the mould, all insulation materials and active heating and cooling facilities, as shown in Figure 6.27. Because of confidentiality agreements with the industry, the heating and cooling systems are not shown in detail in this figure. All heat conductance and capacity effects as well as the nonstationary release of latent heat are taken into account. All material contact regions between silicon, mould or insulation materials are modelled by heat flow–resistance parameters. To describe the heat flux by radiation inside the furnace, a view-factor model is included in the software. All material data are treated in their temperature dependency and all the internal control systems of the furnace are added to the simulation software.

To start one simulation run, only the cooling water temperature and the time-dependent process control information are necessary as input data, as they are entered in the crystallisation

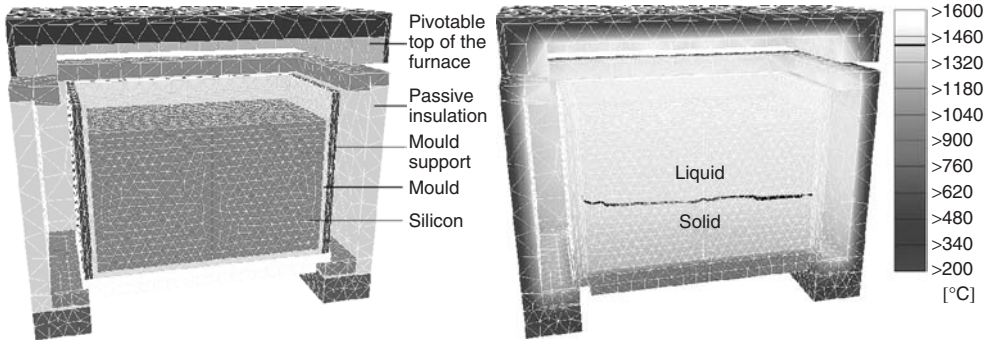


Figure 6.27 Finite element geometry of an ingot casting furnace and simulated temperature distribution during a reference process. The liquid–solid interface is marked by the black line

furnace. Output from one calculation is the three-dimensional temperature history in the furnace, beginning after pouring the melt and ending with a homogeneous temperature of about 300 °C inside the ingot. This calculation needs less than 6 h on a common one-processor workstation.

In Figure 6.27, an example of the temperature distribution during a reference process is shown in the middle, cut through the furnace. The liquid–solid interface is marked by the melting temperature isotherm. The solidification front is mostly flat, and a slight asymmetry is caused by the specific construction of the heating system. These simulation results are verified in an experimental crystallisation furnace, in good agreement with the measurement in the ingot volume during crystallisation.

In general, the shape of the solidification front is controlled by the lateral heat flux, while the vertical heating and cooling conditions control the solidification velocity. To investigate these general reflections for the described furnace, variations of the process control were simulated. In Figure 6.28, two variations are presented. By a 30% raise of heating power at the side walls of the ingot, the shape of the solidification front becomes more convex. Otherwise, a reduction of heating power by about 20% turns the solidification front to a more concave shape. Additionally, to this more or less predictable effect, simulation results show an increase in the solidification time for the convex crystallisation of 44%, and a 30% reduced processing time for the concave solidification. Both effects are due to the total varying power input in the furnace. These simulation results enable

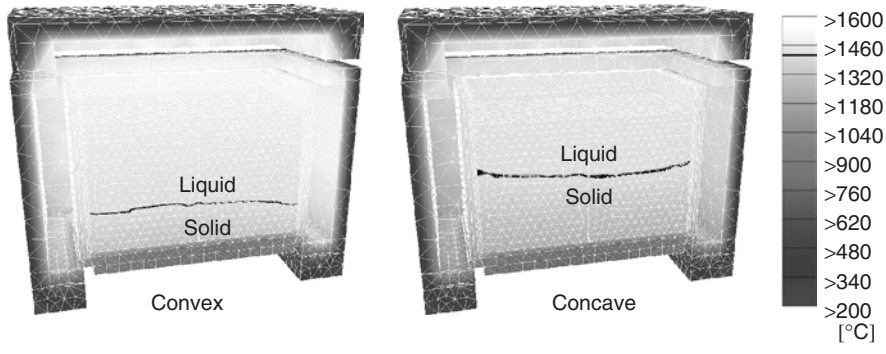


Figure 6.28 Examples of convex and concave liquid–solid interfaces due to the variation of side wall heating power. Both pictures are taken at the same process time

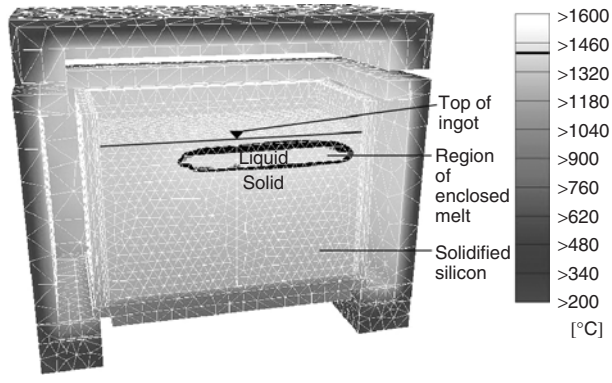


Figure 6.29 Simulation result representing a marked decrease of heating power at the top region of the ingot. By this, the solidification time is reduced to half, but solidification ends with an inclusion of silicon melt

the furnace operator to find the balance between material quality, which is known to be high in the case of planar solidification, and process economy.

In Figure 6.29, simulation results are shown, representing a noticeable reduction of heating power at the ingot top. By this, the solidification velocity can be speeded up and the time for solidification is reduced to half. However, in the simulated case study, the solidification ends with an encapsulation of the melt by solidified silicon. Because of the 10% higher density of liquid silicon with respect to the solid phase, this process scenario causes a burst out of the melt from the ingot volume. This can lead to a crack in the mould and to damage of the furnace. These case-study simulations can find worst-case process conditions that must be avoided in production. For more simulation results of the SOPLIN process, see [114–116].

6.6.4 Simulation of Silicon Ribbon Growth

As a second example of silicon crystallisation processes, the RGS is taken. The basic idea of this process is the decoupling of the pulling direction of substrates on which silicon solidifies and the solidification direction of the silicon itself, as shown in Figure 6.30. By this, a very high production rate of one wafer per second is realisable. For this process numerical simulations were performed describing the crystallisation of the silicon ribbon in further detail. This simulation is realised by a phase-field approach [117]. In Figure 6.31 two nucleation states of the silicon on the substrate are compared by their temperature field in two-dimensional simulations. In the first case a supercooled region in front of the tip of the growing crystal leads to a more dendritic growth mode of the

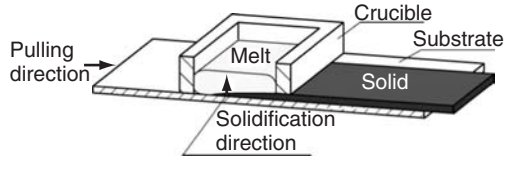


Figure 6.30 Scheme of the RGS process. The latent heat is removed through the substrate. By this, solidification is propagated vertical to the substrate and is decoupled from the substrate pulling direction

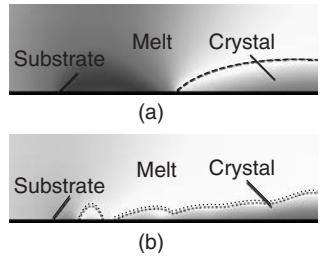


Figure 6.31 Simulated temperature profile on the RGS substrate. Light grey indicates high temperature. The liquid–solid interface is marked with the dotted line. Unstable crystal growth into a supercooled melt at the tip of the ribbon; nucleation of new grains limits the supercooling and leads to a stable columnar grain growth

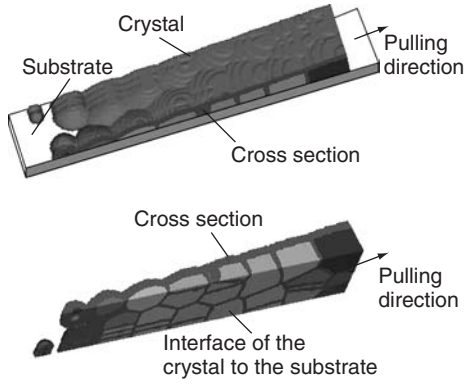


Figure 6.32 Simulation results of the growth of silicon grains on the RGS substrate during solidification. Different grains are marked by different shades of grey. The liquid–solid interface is marked by a darker grey scale and envelops the grains. For better visibility, the melt is not shown in this figure

crystal, because the liquid–solid interface is morphologically unstable. In the second crystallisation mode, the supercooling decreases owing to a nucleation of new grains on the substrate surface, which leads to a more columnar growth of the silicon sheet. Both crystallisation modes depend on the surface of the substrate and on the heating power controlled temperatures of the substrate and the melt. Numerical simulations allow investigating this crystal growth tendency as well as the temperature field at the contact region of the ribbon on the substrate. In Figure 6.32 an example of the simulation of single silicon grain growth on the substrate is given [118]. In this way, numerical simulations can study the grain growth under various temperature conditions and grain-selection mechanisms of silicon crystals.

6.7 CONCLUSIONS

At the present time, the PV industry relies on solar cells made on crystalline silicon wafers, which provide around 90% of the total PV power installed. It is expected that monocrystalline and multicrystalline solar cells will continue dominating the industry for the next 10–20 years.

With rejects from the microelectronic industry as feedstock, and today with high-purity polysilicon, the PV industry grows monocrystalline ingots by the Czochralski technique. Owing to more relaxed specifications than in microelectronics, the throughput of PV Cz pullers can be increased and still produce high-quality silicon, allowing the achievement of 16–18% efficient solar cells.

Multicrystalline Si can be manufactured at a lower cost than monocrystalline Si, but produces less efficient cells, mainly owing to the presence of dislocations and other crystal defects. With the introduction of new technologies, the gap between multicrystalline and monocrystalline solar cells is reducing, being now at around 1–1.5% absolute efficiency.

Si ingots are sliced into thin wafers with multi-wire saws, with throughputs of about 10 000 wafers per day and per machine. Sawing is responsible for high material losses, and amounts to a substantial part of the wafer production cost. Understanding the microscopic processes of wire sawing allows optimisation of the technique and improvement of sawing performance.

Silicon ribbon wafer/foil technologies have matured in the past decade. Several ribbon technologies have already demonstrated robust and reproducible processes, which have been scaled up to megawatt levels. However, they have not reached the expected status of serious contenders for competing on a scale equal to that of conventional wafers. EFG and STR technologies have built up hundred-megawatt factories in the last years, although EFG has recently shut down all factories. RGS intends to reach this scale soon, while SF and WEB have not reached this goal.

During the last decade, computer power has been increasing strongly and there is great progress in the modelling of physical phenomena and in the development of simulation tools. By this, computer simulation has become a powerful tool in science and industrial applications. The simulation results of industrial crystallisation processes that are shown and the detailed study of the crystallisation mode by numerical simulations are some examples of the possibilities today. These numerical simulations offer a wide range of possibilities to increase the knowledge of the basic physics of crystallisation and technical processes. One insistent demand on computer simulations is to close the gap between science and engineering to get a closer picture of reality.

REFERENCES

1. Hirshman WP, *Photon International*, 170–206, March (2009).
2. Knobloch J *et al*., *Proc. 13th Euro. Conf. Photovoltaic Solar Energy Conversion*, 9 (1995).
3. Li J *et al*., *Appl. Phys. Lett.* **60**, 2240 (1992).
4. Zundel M, Csaszar W, Endrös A, *Appl. Phys. Lett.* **67**, 3945 (1995).
5. Sinke W, Cañizo C, del Coso G, *Proc. 23rd EU-PVSEC*, 700–3705 (2008).
6. Güttler G, Queisser H, *Energy Conversion* **10**, 51 (1970).
7. Hahn G *et al*., *Solid State Phenomena* **156–158**, 343–349 (2010).
8. Rogol M, *Photon's 8th Solar Silicon Conference*, Stuttgart (2009).
9. Rinio M *et al*., *Proc. 23rd EU-PVSEC*, 10014–1017 (2008).
10. Hofstetter J *et al*., *Materials Science and Engineering B* **150–160**, 299–304 (2008).
11. Laades A *et al*., *Solid State Phenomena* **156–158**, 381–386 (2010).
12. Klingshirn H, Lang R, *U.S. Patent 5324488* (1994).
13. Fickett B, Mihalis G, *Journal of Crystal Growth*, **225**, 580–585 (2001).
14. Spangler MV, Seburn CD, *U.S. Patent 7141114* (2006).
15. Hu T *et al*., *Microelectronics Engineering*, **56** (1–2), 89–92 (2001).
16. www.longi-silicon.com.
17. Koch W *et al*., *Solid State Phenomena* **401**, 57–58 (1997).
18. Häßler C *et al*., *Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, 1886 (1998).
19. Schmidt J *et al*., *Proc. 22nd EU-PVSEC*, 998–1001 (2007).

20. Nage., H. *et al*., *Proc. 22nd EU-PVSEC*, 1547–1551 (2007).
21. Cuevas A *et al*., *Applied Physics Letters*, **81**, 4952 (2002).
22. Tucci M *et al*., *Proc. 23rd EU-PVSEC*, 1847–1850 (2008).
23. Uzum A *et al*., *Proc. 23rd EU-PVSEC*, 1625–1628 (2008).
24. MacDonald D *et al*., *Proc. 23rd EU-PVSEC*, pp 1475–1477 (2008).
25. Chen J *et al*., *Solid State Phenomena* **156–158**, 19–26 (2010).
26. Koch W, Krumbe W, Schwirtlich I, *Proc. 11th EU Photovoltaic Specialist Conf.*, 518 (1992).
27. Bertoni M *et al*., *Solid State Phenomena* **156–158**, 11–18 (2010).
28. Fuller C, Logan R, *J. Appl. Phys.* **28**, 1427 (1957).
29. Kaiser W, Frisch H, Reiss H, *Phys. Rev.* **112**, 1546 (1958).
30. Wagner P, Hage J, *Appl. Phys. A: Solids Surf.* **49**, 123 (1989).
31. Cazcarra V, Zunino P, *J. Appl. Phys.* **51**, 4206 (1980).
32. Pensl G *et al*., *J. Appl. Phys. A.* **48**, 49 (1989).
33. Häßler C *et al*., *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, 720 (1997).
34. Du G *et al*., *Semicond. Sci. Technol.*, **23**, 055011 (2008).
35. Hudelson S *et al*., *Proc. 23rd EU-PVSEC*, 963–964 (2008).
36. Hofstetter J *et al*., *Solid State Phenomena* **156–158**, 387–393 (2010).
37. EPIA Report, *Global Market Outlook for Photovoltaics until 2013*, see www.epia.org (2009).
38. Chonan S, Jiang Z, Yuki Y, *J. Vib. Acoustics* **115**, 529 (1993).
39. Wells R, *Solid State Technol.* **30**, 63 (1987).
40. Sahoo R *et al*., in Subramania K (ed.), *ASME – IMECE Manufacturing Science and Engineering*, 131, ASME press, New York (1996).
41. Li J, Kao I, Prasad V, *ASME – IMECE Manufacturing Science and Engineering*, 439, ASME press, New York (1997).
42. Yang F, Kao J, *J. Electron. Packaging* **121**, 191 (1999).
43. Bhagavat M, Kao I, Prasad V, *ASME J. Tribology* **122**, 394–404 (2000).
44. Kao I, Wie S, Chiang P, *Proc. of NSF Design & Manufacturing Grantees Conf.*, 239 (1997).
45. Chen C, Leipold M, *J. Am. Ceram. Soc.* **59**, 469 (1980).
46. Lawn B, Marshall D, *J. Am. Ceram. Soc.* **62**, 347 (1979).
47. Anstis G, Chantikul P, Lawn B, Marshall D, *J. Am. Ceram. Soc.* **64**, 533 (1981).
48. Evans A, Charles E, *J. Am. Ceram. Soc.* **59**, 371 (1976).
49. Lawn B, Evans A, *J. Mater. Sci.* **12**, 2195 (1977).
50. Gogots Y, Baek C, Kirscht F, *Semicond. Sci. Technol.* **14**, 936 (1999).
51. Weppelmann E, Field J, Swain M, *J. Mater. Res.* **8**, 246 (1993).
52. Borst C, Möller H, *German BMBF VEDRAS Report*, 23 (1999).
53. Beesley JG, Schönholzer U, *Proc. 22nd EU-PVSEC*, 956–962 (2007).
54. Henley F *et al*., *Proc. 23rd EU-PVSEC*, 1090–1093 (2008).
55. Hopman S *et al*., *Proc. 23rd EU-PVSEC*, 1131–1135 (2008).
56. Dross F, *et al*., *Proc. 33rd IEEE-PVSC*, 1–5, (2008).
57. Bye, JI, Jensen SA, Aalen F, Rohr C, Nielsen Ø, Gäumann B, Hodsden J, Lindemann K, *Proc. 24th EU-PVSEC*, 1269–1272 (2009).
58. Electricity from Solar Cells, *Flat Plate Array Project*, 10 Years of Progress, JPL Publication 400–279 10/85 (October 1985).
59. www.schott.com/photovoltaics.
60. www.solwafer.eu/nl.
61. Rohatgi A *et al*., *Applied Physics Letters*, **84**, 145 (2004).
62. For comprehensive reviews see bibliographies in *J. Cryst. Growth*, **50** (1980); *J. Cryst. Growth*, **82** (1987); *J. Cryst. Growth*, **104** (1990).
63. Ciszek T, *J. Cryst. Growth* **66**, 655 (1984).
64. Bell R, Kalejs J, *J. Mat. Res.* **13**, 2732 (1998).
65. Chalmers B, *J. Cryst. Growth* **70**, 3 (1984).

66. Thomas P, Brown R, *J. Cryst. Growth* **82**, 1 (1987).
67. Hopkins R *et al*., *J. Cryst. Growth* **82**, 142 (1987).
68. Ravi K, *J. Cryst. Growth* **39**, 1 (1977).
69. Kalejs J, in Khattak C, Ravi K (eds), *Silicon Processing for Photovoltaics II*, 185–254, North Holland, Amsterdam (1987).
70. Garcia D *et al*., *J. Cryst. Growth* **225**, 566 (2001).
71. Mackintosh B *et al*., *Proc. 28th IEEE Photovoltaic Specialist Conf.*, 46 (2000).
72. Sachs E, *Proceedings. of the Flat-Plate Solar Array Project Research Forum on the High-Speed Growth and Characterisation of Crystals for Solar Cells*, JPL, 279 (1984).
73. Sachs E, Ely D, Serdy J, *J. Cryst. Growth* **82**, 117 (1987).
74. Wallace R *et al*., *Sixth Workshop on the Role of Impurities and Defects in Silicon Device Processing – Extended Abstracts and Papers*, NREL/SP-413-21550, 203 (1996).
75. Lange H, Schwirtlich I, *J. Cryst. Growth* **104**, 108 (1990).
76. Schönecker A *et al*., *12th Workshop on Crystalline Silicon Solar Cells, Materials and Processes* (2002).
77. Cotter J *et al*., *Proc. 13th Euro. Conf. Photovoltaic Solar Energy Conversion*, 1732 (1995).
78. Barnett A *et al*., *Proc. 18th IEEE Photovoltaic Specialist Conf.*, 1094 (1985).
79. Ruby D, Ciszek T, Sopori B, *NCPV Program Review Meeting*, April, 16–19, Denver, CO, unpublished.
80. www.evergreensolar.com.
81. *Evergreen annual report 2008*, in www.evergreensolar.com (2009).
82. Schmidt W, Woesten B, Kalejs J, *Proc. 16th EPVSEC*, 1082–1086 (2000).
83. Sawyer W, Bell R, Schoenecker A, *Solid State Phenomena* **37–38**, 3 (1994).
84. Meier D, Hopkins R, Campbell R, *J. Propulsion Power* **4**, 586 (1988).
85. Koch W *et al*., *Proc. 2nd World Photovoltaic Conf.*, 1254 (1998).
86. Koshka Y *et al*., *Appl. Phys. Lett.* **74**, 1555 (1999).
87. Bowler D, Wolf M, *IEEE Trans. Components, Hybrids Manufacturing Technology* **3**, 464 (1980).
88. Basore P, Clugston D, PC1D Version 4.2 for Windows, Copyright University of New South Wales.
89. Van Nieuwenhuysen K *et al*., *Journal of Crystal Growth*, **287**(2), 438–441 (2006).
90. Bathey B *et al*., *Proc. 28th IEEE Photovoltaic Specialist Conf.*, 194 (2000).
91. Sims P *et al*., *Annual Report on DOE/NREL Subcontract No. DE-AC36-98-GO20337*, Publication No. NREL/SR-520-28547 (2000).
92. Seren R *et al*., *Proc. 22nd EU-PVSEC*, 854–858 (2007).
93. Sana P, Rohatgi A, Kalejs J, Bell R, *Appl. Phys. Lett.* **64**, 97 (1994).
94. Bailey J, Kalejs J, Keaveny C, *Proc. 24th IEEE Photovoltaic Specialist Conf.*, 1356 (1994).
95. The earliest work is reported in the *3rd Photovoltaic Science and Engineering Conference* by Morita H *et al*., *Jpn. J. Appl. Phys.* **21** 47 (1982); *U.S. patent No. 4,640,001* (1987) and *Australian patent No. 609424* (1991); a comprehensive review of recent developments is given by Szlufcik J *et al*., *Proc. E-MRS Spring Meeting*, E-VI.1 (2001).
96. www.abaqus.com.
97. www.ansys.com.
98. www.marc.com.
99. Lambropoulos J *et al*., *J. Cryst. Growth* **65**, 324–330 (1983).
100. Kalejs J, Schmidt W, *Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, 1822–1825 (1998).
101. Wallace R, Janoch R, Hanoka J, *Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, 1818–1821 (1998).
102. Chang C, Brown R, *J. Cryst. Growth* **63**, 343–352 (1983).

103. Zulehner W, Huber D, *Crystals*, Vol. 8, 3–143, Springer Verlag, Berlin, Heidelberg, New York (1982).
104. Helmreich D, The Wacker Ingot Casting Process, in Khattak C, Ravi K, (eds), *Silicon Processing for Photovoltaics*, Vol. II, 97–115, North-Holland, Amsterdam (1987).
105. Schätzle P *et al*., *Proc. 11th Euro. Conf. Photovoltaic Solar Energy Conversion*, 465–468 (1992).
106. Khattak C, Schmid F, *Proc. 25th IEEE Photovoltaic Solar Energy Conversion*, 597–600 (1996).
107. Koch W *et al*., *Proc. 12th Euro. Conf. Photovoltaic Solar Energy Conversion*, 797–798 (1994).
108. Häßler C *et al*., *Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, 1886–1889 (1998).
109. Zabaras N, Ruan Y, Richmond O, *Computer Methods in Applied Mechanics and Engineering*, Vol. 8, 333–364, Elsevier Science Publisher B.V., Amsterdam (1990).
110. Diemer M, Franke D, Modelling of Thermal Stress Near the Liquid–Solid Phase Boundary, in Thomas B, Beckermann C, (eds), *Modelling of Casting, Welding and Advanced Solidification*, Vol. 8, 907–914, TMS (1998).
111. Zienkiewicz O, Taylor R, *The Finite Element Method*, McGraw-Hill, London (1989).
112. Bathe K, *Finite Element Procedures in Engineering Analysis*, München Verlag, Wien (1984).
113. Kurz W, Fischer D, *Fundamentals of Solidification*, Trans Tech Publications, Zürich (1998).
114. Franke D *et al*., *Proc. 25th IEEE Photovoltaic Specialist Conf.*, p 545–548 (1996).
115. Häßler C *et al*., *Solid State Phenomena* **67–68**, 447–452 (1999).
116. Franke D, Apel M, Häßler C, Koch W, *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, 1317–1320 (2000).
117. Steinbach I *et al*., *Physica D* **94**, 135–147 (1996).
118. Steinbach I, Höfs H, *Proc. 26th IEEE Photovoltaic Solar Energy Conversion*, 91–93 (1997).

7

Crystalline Silicon Solar Cells and Modules

Ignacio Tobías[1], Carlos del Cañizo[1] and Jesús Alonso[2]

[1]*Instituto de Energía Solar, Universidad Politécnica de Madrid, Spain*
[2]*Isofotón S. A., Malaga, Spain*

7.1 INTRODUCTION

Crystalline silicon solar cells and modules have dominated PV technology from the beginning. They constitute more than 85% of the PV market today, and although their decline in favor of other technologies has been announced a number of times, they presumably will retain their leading role for a long time, at least for the next decade.

One of the reasons for crystalline silicon to be dominant in PV comes from the fact that microelectronics has developed silicon technology to a great extent. Not only has the PV community benefited from the accumulated knowledge, but also silicon feedstock and second-hand equipment have been acquired at reasonable prices. On the other hand, microelectronics has taken advantage of some innovations and developments proposed in photovoltaics.

For several decades, the terrestrial photovoltaic market has been dominated by p-type Czochralski silicon substrates. Continuous improvements in performance, yields and reliability have allowed an important cost reduction and the subsequent expansion of the PV market. Due to the lower cost of mc-Si wafers, multicrystalline silicon cells emerged in the 1980s as an alternative to single-crystal ones. However, their lower quality precluded the achievement of similar efficiencies to those of Cz, so that the figure of merit \$/W has been quite similar for both technologies for a long time.

Deeper understanding of the physics and optics of the mc-Si material led to improved device design, which allowed a wider spread of the technology. A combination of improved material quality and material processing has allowed higher efficiencies at a still low cost, increasing multicrystalline share of the PV market. Recent evolution of the market can be seen in Table 7.1 [1].

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

Table 7.1 Market share of monocrystalline and multicrystalline solar cells

Year	Cz-Si solar cells		MC-Si solar cells	
	Output (MW)	Market share (%)	Output (MW)	Market share (%)
1996	48.7	55	28.4	32
2000	92.0	32	146.7	51
2007	1805	42	1934	45

This chapter offers an overview on silicon solar cell and module technology. First, Si properties justifying its use as photovoltaic material are presented. Then, design of Si solar cells is reviewed, highlighting the benefits and limits of different approaches. Manufacturing processes are described, paying special attention to technologies that are currently implemented at the industrial level, mostly based on screen-printing metallization technology. Considerations on ways of improving solar cell technology are also specified, including other approaches which are already in industrial production. Next, crystalline Si modules are reviewed, attending to electrical performance, fabrication sequence and reliability concerns.

7.2 CRYSTALLINE SILICON AS A PHOTOVOLTAIC MATERIAL

7.2.1 Bulk Properties

Crystalline silicon has a fundamental indirect bandgap $E_G = 1.17$ eV and a direct gap above 3 eV [2] at ambient temperature. These characteristics determine the variation of optical properties of Si with wavelength, including the low absorption coefficient for carrier generation for near-bandgap photons [3]. At short (UV) wavelengths in the solar spectrum, the generation of two electron–hole pairs by one photon seems possible, though quantitatively this is a small effect [4]; at the other extreme of the spectrum parasitic free carrier absorption competes with band-to-band generation [5]. The intrinsic concentration is another important parameter related to the band structure; it links carrier disequilibrium with voltage [6].

At high carrier densities, doping- or excitation-induced, the band structure is altered, leading to an increase in the effective intrinsic concentration: this is one of the so-called heavy doping effects that degrade the PV quality of highly doped regions [7].

Recombination in Si is usually dominated by recombination at defects, described with Shockley–Read–Hall (SRH) lifetimes. The associated lifetime τ (which can also be described in terms of diffusion length L) increases for good quality materials. Auger recombination, on the contrary, is a fundamental process that becomes important at high carrier concentration [8]. The Auger coefficients are reported to be higher at moderate carrier densities due to excitonic effects [9]. Band-to-band direct recombination is also a fundamental process, but quantitatively negligible (it is instructive, however, to notice that record-efficiency solar cells have so extraordinarily low SRH recombination levels that they perform as 1%-efficient light-emitting diodes, i.e. radiative recombination is significant [10]).

At low and moderate doping, electrons present mobilities about three times higher than holes, both limited by phonon scattering at room temperature. Impurity scattering dominates for higher doping densities [11]. Carrier–carrier scattering affects transport properties in highly injected material [12].

7.2.2 Surfaces

7.2.2.1 Contacts

Contacts are structures built on a semiconductor surface that allow charge carriers to flow between the semiconductor and the external circuit. In solar cells, contacts are required to extract the photogenerated carriers from the absorbing semiconductor substrate. They should be selective, i.e. should allow one type of carriers to flow from Si to metal without energy loss while blocking the transport of carriers of the opposite type.

Direct Si–metal contacts in general do not behave this way. As an exception, good hole contacts to highly doped p-Si substrates with aluminum are possible. But the most used approach is to create heavily doped regions under the metal, p-type for hole extraction and n-type for electron extraction. Majority carriers in this region can flow through the contact with low voltage loss. The transport of minority carriers is described by a surface recombination velocity (SRV), S. Although the SRV is high, limited only by thermal diffusion so that $S \cong 10^6$ cm·s^{-1} [13], the concentration of minority carriers, for a given pn product, is suppressed by the high doping and the flow is reduced. As it will be seen later on, the contact for the minority carriers is usually placed at the front (illuminated) face of the substrate, and the corresponding heavily doped layer is usually called emitter. The doped region under the majority carrier contact at the back is called a back surface field (BSF).

Recombination at these heavily doped regions is described by the saturation current density J_0 which includes volume and true contact recombination. The thickness of these layers w should be much higher than the minority-carrier diffusion length L so that few excess carriers reach the contact, and the doping level must be very high to decrease contact resistance and the minority-carrier concentration, although heavy doping effects may limit the doping level advisable for such regions. The recombination activity of BSF layers is often described in terms of an effective SRV instead of the saturation current density.

Typical 10^{-13}–10^{-12} A·cm^{-2} J_0 values are achieved [14, 15]. Diffused phosphorus is used for n-contacts. Aluminum alloying has the advantages over boron for p-contacts that very thick p^+ layers can be formed in a short time at moderate temperatures and that gettering action is achieved [16]. As a shortcoming, the p^+ layer is nonhomogeneous and can even be locally absent; the obtained J_0 is larger than expected for a uniform layer. In comparison with aluminum, boron offers higher doping levels because of a larger solubility [17] and transparency to light so that it can be used also at illuminated surfaces.

Heterojunctions to wide-bandgap materials, such as the a-Si:H/ITO transparent contact in Sanyo HIT cells [18], can produce selective contacts if proper alignment of the bands is achieved [19]. Other structures tested are metal–insulator–semiconductor (MIS) contacts [20] and polysilicon contacts [21].

7.2.2.2 Non-contacted surfaces

Due to the severe alteration of the bonding of Si atoms, a large number of bandgap states exist at a bare Si surface which, acting as SRH recombination centers, make the SRV to be very large, up to 10^5 cm s^{-1}, though highly dependent on substrate resistivity and surface condition [22]. In order to reduce surface recombination two main approaches are followed [23].

In the first one, the density of electron surface states in the gap is decreased. This is accomplished by depositing or growing a layer of an appropriate material that partially restores the bonding environment of surface Si atoms. This material must be an insulator.

Thermal SiO_x is grown in an oxygen-rich atmosphere at the expense of substrate Si atoms at high temperatures around 1000 °C. Plasma-enhanced chemical vapor deposition (PECVD) is a low-temperature process widely used in the industry for producing SiN_x layers in the 300–450 °C range [24] that can also be used for amorphous silicon and silicon oxide deposition. The quality of both techniques is very sensitive to subsequent treatments, with hydrogen playing a major role in obtaining low SRV values below 100 cm s^{-1}.

Other passivating systems with a potential for solar cells are PECVD silicon oxide [25, 26], silicon carbide [27] and aluminum oxide [28, 29]. Low-temperature-deposition materials are of greatest interest for multicrystalline substrates because of their sensitivity to thermal processing.

As a general rule, S increases with the doping of the substrate. It depends on injection level and doping type too, because the interfaces contain positive charges that affect the number of carriers at the surface and because the capture probability for electrons and holes is different. n-type or intrinsic surfaces are usually better than p-type ones [30–32]. Stability under UV exposure is another fundamental concern.

In the second approach, the excess minority carrier density at the interface is decreased relative to the bulk. This effect has been already commented with respect to contacts and results in a reduced effective SRV at the bulk edge of the corresponding space-charge region. It can be produced by doping, by the electrostatics associated to a MOS structure [33] or by charges in the surface layer; this mechanism is thought to be very important in SiN_x passivating layers that contain a large amount of positive charge.

The surface layer can be accumulated or inverted, or, correspondingly, doped the same or opposite to the substrate. Its recombination activity is better described by a constant saturation current density J_0, whose minimization follows the same rules as mentioned for the contacts if S at the surface is high. On the contrary, if S is low compared with D/L for minority carriers (with D the diffusion constant and L the diffusion length in the substrate) in the surface layer, it is better for it to be thin or "transparent" to minority carrier flow ($w < L$). The optimum doping level is a compromise between reduction of the excess carriers, and heavy doping effects and the increase of the SRV with doping. Moderate doping levels are favored in this case. Under passivated surfaces, J_0 values around 10^{-14} A·cm^{-2} are achievable, with phosphorus-doped substrates, giving better results than boron doped ones [14, 15].

In conclusion, recombination at a non-contacted surface can be made much smaller than at a metallized one, and this has a deep influence in the evolution of Si solar cell design. Low-temperature passivation layers and heterojunction contacts have become very important in the last years.

7.3 CRYSTALLINE SILICON SOLAR CELLS

7.3.1 Cell Structure

Several studies have been carried out to find the limiting efficiency and the optimum structure of a Si solar cell [34–37]. All avoidable losses are assumed to be suppressed:

1. No reflection losses and maximum absorption as achieved by ideal light trapping techniques.
2. Minimum recombination: SRH and surface recombination are assumed avoidable and only Auger and radiative recombination remain.
3. The contacts are ideal: neither shading nor series resistance losses.

4. No transport losses in the substrate: the carrier profiles in the substrate are flat so that recombination is the minimum possible for a given voltage.

The optimum cell should use intrinsic material, to minimize Auger recombination and free carrier absorption, and should be around 80 μm thick, result of the trade-off between absorption and recombination. It could attain nearly 29% efficiency at one sun AM 1.5 Global, 25 °C [35].

This ideal case does not tell where to put the contacts. To realize condition 4 above, they should be located at the illuminated or front face, closest to photogeneration (Figure 7.1a). Due to metal shading losses this threatens condition 3. This solution is being considered for concentration [38]. Putting both contacts at the back works the other way round (Figure 7.1b). Back-contacted cells exhibit record efficiencies under concentration and very high values near 23% have been demonstrated at one sun [39]. As will be shown below in Section 7.6.3, highest efficiency at the industrial level is also achieved by cells following this approach.

In most cells each contact is placed on a different face, which is technologically simpler (Figure 7.1c). Minority carriers in the substrate are usually collected at the front since their extraction is more problematic due to their low density. The diffusion length describes the maximum distance from which they can be collected. Majority carriers can drift to the back contact with low loss. Several designs extract minority carriers both at the front and at back (Figure 7.1d) [40], thus increasing the volume of profitable photogeneration.

Bifacial cells are designed to collect light incident at both faces, which allows a boost in output power if there is a significant albedo component. They can be implemented with any structure in Figure 7.1 with the condition that both surfaces allow light through [41].

Both record-efficiency, laboratory solar cells [42] with nearly 25% efficiency (Figure 7.2a) and mass-produced industrial devices (Figure 7.2b), typically 15.5–16% (multi-Si) or 16.5–17% (mono-Si) efficient, display the contacting structure in Figure 7.1c. They will be described in the paragraphs that follow with the aim of illustrating the amplitude and the reasons for the performance gaps between the ideal and the best Si cell, and between this and industrial cells.

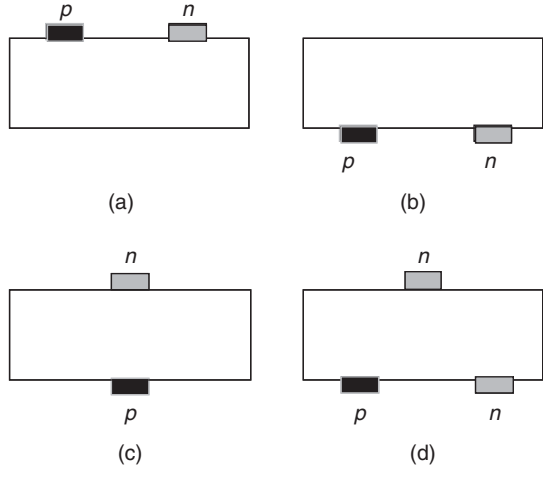


Figure 7.1 Contacting structures: (a) both contacts at the front and (b) at the back; (c) both faces contacted; (d) one carrier extracted at both faces. The structures with interchanged *n* and *p* types are also possible

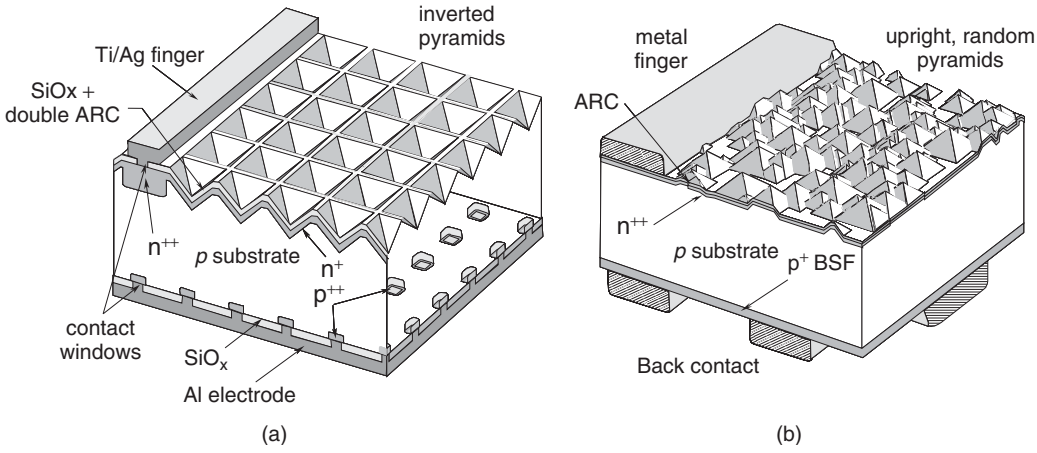


Figure 7.2 (a) Passivated emitter and rear locally diffused (PERL) cell; (b) industrial cell with screen-printed contacts (not to scale)

7.3.2 Substrate

7.3.2.1 Materials and processing

Highest efficiencies are achieved with monocrystalline float zone (FZ-Si) material, which in addition to extreme crystalline perfection shows the lowest contamination levels of both metallic and light (O, C, N) impurities. This translates into longest post-processing SRH lifetimes in the millisecond range, but still shorter than the Auger limit. Magnetic Czochralski (MCz) material contains much less oxygen than conventional Czochralski allowing very high efficiencies to be obtained too [42].

Industrial cells use Czochralski (Cz-Si) wafers because of their availability. Cz wafers are also perfect crystals, but they contain a high concentration of oxygen which affects lifetime in several ways [43]. In recent years, a majority of commercial devices are made on multicrystalline (mc-Si) substrates grown in blocks or ribbons with procedures specially developed for photovoltaics. In addition to crystal defects such as grain boundaries and dislocations, the potential content of metals is higher because of lower segregation to the melt during the faster solidification process. As a result, the lifetime of mc-Si is lower.

But lifetime is important at the end of solar cell fabrication during which it can undergo strong variations. This issue is handled in a different way in a laboratory and a factory environment. In the former, measures are taken to maintain long initial lifetimes by avoiding contamination during high-temperature steps: furnace cleaning, ultra-pure chemicals, etc. In a rough industrial environment and with defect-containing (Cz- and mc-Si) materials the problem is more complex: in addition to contamination from the surroundings, impurities and defects in the substrate move, interact and transform at high temperature. The solution is to integrate gettering steps [44] in the fabrication flow that reduce the impact of contamination, and to tailor the thermal treatments to the peculiarities of the material. Final substrate lifetimes of industrial cells are in the $10-20\,\mu s$ range.

Gettering techniques eliminate or reduce contaminant impurities in a wafer, so neutralizing the effect of lifetime reduction. In a gettering process a sink region is formed, able to accommodate the lifetime-killing impurities in such a way that they are not harmful for the device being manufactured, or at least where they can be easily removed.

In solar cell fabrication we take advantage of the fact that phosphorus and aluminum diffusions, appropriate candidates for emitter and BSF layers, respectively, produce gettering in certain conditions [45]. Other techniques have been explored [46, 47], but their integration in a solar cell process is not so straightforward.

The P gettering effect has been proved for a wide variety of P diffusion techniques (spin-on, $POCl_3$, PH_3, ...), provided diffusion is done in supersaturation conditions (i.e. over its solid solubility in silicon). Unfortunately, this leaves a "dead layer" of electrically inactive phosphorus near the surface, which reduces UV response of the cells in the case it is not etched away [48]. Another phenomenon related to this supersaturated P is the injection of silicon self-interstitials to the bulk of the material, which is responsible of an enhancement of the gettering effect [49].

When Al is deposited on Silicon (by different techniques such as sputtering, vacuum evaporation or screen-printing) and annealed over the eutectic temperature (577 °C) a liquid Al–Si layer is formed, where impurities tend to segregate due to their enhanced solubility [50]. They will remain in this gettering layer when cooling, so that bulk lifetime will improve after the process.

Gettering conditions (temperature, process duration, ...) for multicrystalline substrates differ from those of single crystal, due to the interaction among metal impurities, crystalline defects and other impurities present in mc-materials (mainly O and C).

For mc-Si, it has been realized that gettering efficiency is strongly material dependent [51, 52]. This is explained by the fact that different techniques to grow mc-Si ingots produce wafers with different number and distribution of defects. Differences are even found in regions of the same ingot [53].

Additionally, a single mc wafer may exhibit nonuniform properties, both areal and with depth, so that response to a gettering process can be inhomogeneous, affecting the final electrical performance of the solar cell [54, 55].

Another approach to improve material quality is "bulk passivation", a treatment with hydrogen where atomic H interacts with impurities and defects in the bulk of Si, neutralizing their recombination properties to a certain extent [56]. Hydrogenation is typically induced during SiN_x deposition by PECVD.

A third possibility, mainly associated to the greater extent of defects and impurities in multicrystalline substrates, which is called *defect engineering*, is based on the idea "if you cannot get rid of the impurities, minimize their impact", by inducing their precipitation thanks to tailored thermal processes or cooling steps [57].

7.3.2.2 Doping level and type

Laboratory record-efficiency and low-cost industrial cells use boron-doped substrates. Rather than fundamental advantages [58], there are practical (properties of P diffusion, easiness of Al alloying) and historical reasons for the preponderance of *p* silicon [59]. Evidence is being collected in favor of *n*-type silicon: the role of boron in the degradation of Cz-Si cells under illumination [60], the higher activity of structural defects and of iron and other important contaminants of mc-Si in p-type material [61, 62]. The subject is however controversial and no definite conclusion for all material and solar cell types seems to be valid [63]. Industrial HIT and back-contact high-efficiency devices are fabricated on *n*-type silicon.

The optimum substrate doping depends on the cell structure and dominant recombination mechanism. Though intrinsic substrates present the advantage of highest Auger limiting lifetimes, higher doping is favored when SRH recombination is present, since recombination is proportional to

the excess density which decreases, for a given voltage, as doping increases [34]. This is balanced with a reduction of lifetime itself.

High doping also helps to minimize the series-resistance losses associated with the transport of carriers to the back face in thick cells with the majority-carrier contact at the back.

Doping levels in the 10^{16} cm^{-3} range are found in the substrate of industrial cells [61]. Very high efficiencies have been obtained with both low (1 Ω cm for PERL cells) and high substrate resistivities, as in the point-contact cells [39].

7.3.2.3 Thickness

From the point of view of electrical performance, the choice of the optimum substrate thickness also depends on the structure and the quality of the materials and involves several considerations.

In cells with diffusion lengths longer than thickness the most important issue is surface recombination: if S at the back is higher than D/L for the minority carriers in the substrate (around 250 cm s^{-1} for the best cells), thinning the cell increases recombination at a given voltage, and vice versa. Thinner cells always absorb less light as well, which is attenuated by light trapping techniques. PERL cells were reported to improve when going from 280 to 400 μm thickness because of a (relatively) high rear surface recombination and non-ideal light trapping [64].

The losses associated with the transport of carriers extracted at the non-illuminated face decrease with thinning; in conventional structure cells this leads to decreased series resistance. In back-contacted cells, both types of carriers benefit from thinning and the trade-off with absorption leads to lower w values, around 150–200 μm.

In industrial cells, the main issues related to thickness are cost and fabricability. Thinner cells save expensive feedstock material and advanced wafering techniques and procedures to process very thin, large-area substrates without breaking are being developed. Light trapping and surface recombination are becoming increasingly important. Today, the typical thickness is around 200 μm, which can be comparable to the minority-carrier diffusion length so that a BSF rear passivation is implemented.

7.3.3 The Front Surface

7.3.3.1 Metallization techniques

Metal grids are used at the front face to collect the distributedly photogenerated carriers. The compromise between transparency and series resistance requires metallization technologies able to produce very narrow, but thick and highly conductive metal lines and with a low contact resistance to Si.

Laboratory cells use photolithography and evaporation to form 10- to 15-μm-wide metal fingers. Ti/Pd/Ag structures combine low contact resistance to n-Si and high bulk conductivity. These processes are not well suited to mass production that relies on thick film technologies. In the majority of the industrial cells, Ag pastes are screen-printed, resulting in over 100-μm-wide lines with higher bulk and contact resistance. A finer metallization technology can be implemented by Ni plating on grooves delineated by laser, giving finger dimensions of around 40 μm deep and 20 μm wide [65]. This so called laser-grooved buried-grid approach had some relevance in the past because it was used in the highest-efficiency cells in the market, but its presence is fading with respect to screen-printed metallization due to the simplicity of the latter, even though coarser metallization

techniques imply higher shading and resistance losses, and restrict the efficiency enhancement that could be achieved by internal cell design. Alternative developments include pad printing, aerosol jet printing, metal powder sintering with laser or laser ablation followed by electroplating [66].

7.3.3.2 Homogeneous emitters

Under the metal lines, the substrate must be heavily doped to make the contact selective. Usually the doped region, the emitter, extends all across the front surface, acting as a "transparent electrode" by offering minority carriers in the substrate a low resistance path to metal lines.

When the exposed surface is not passivated (Figure 7.3a), the emitter should be as thin as possible, because the high SRV causes the light absorbed in this region to be poorly collected, and also highly doped to decrease recombination. On the other hand, a sufficiently low sheet resistance has to be achieved. The solution is to make very thin and highly doped emitters.

If the surface is passivated (Figure 7.3b), the collection efficiency of the emitter can be high by lowering the doping level, thus avoiding heavy doping and other detrimental effects. This must be balanced with contact resistance. Etching off the passivating layer before metallization is usually needed (not in screen-printed cells). To maintain low sheet resistance and diminish recombination at the metallized fraction, the emitter is deep (around 1 μm). Note that the collection of carriers near the surface implies that the emitter is thin in terms of the minority-carrier diffusion length ($w < L$) and so very sensitive to surface recombination. Recombination is further reduced by making the contact window narrower than the finger width, as illustrated in Figure 7.3 [67].

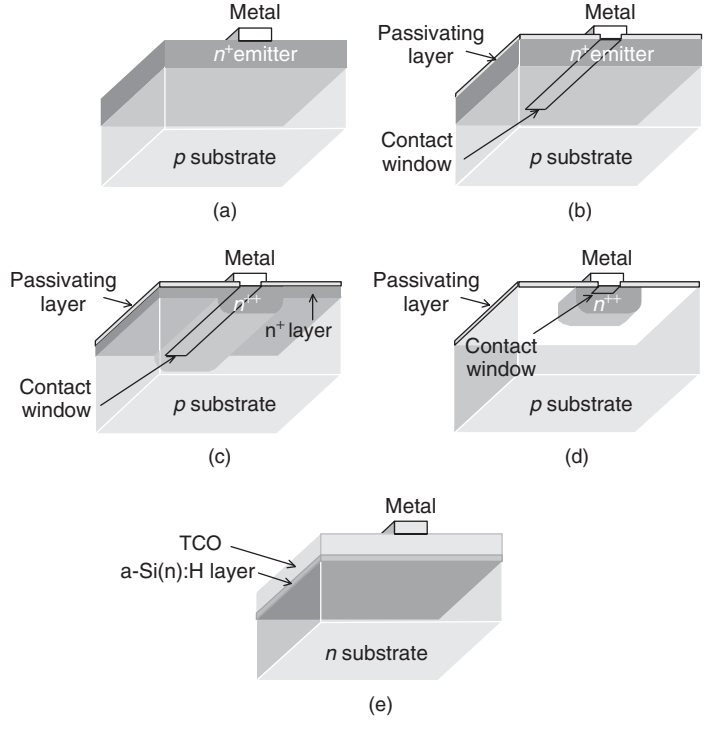


Figure 7.3 Different emitter structures: (a) homogenous emitter without and (b) with surface passivation; (c) selective emitter and (d) localized emitter; (e) heterojunction emitter in HIT solar cells

Control of both the surface concentration and the depth of the emitter is achieved by depositing in a thermal step the desired amount of phosphorus or boron (predeposition) and then diffusing it into the substrate (drive-in) during subsequent furnace steps. The J_0 of the emitter is the average, weighted by the contacted area, of the J_0 of contacted and non-contacted portions.

7.3.3.3 Selective and point emitters

A further improvement involves making separate diffusions for the different regions since the requirements are so different (Figure 7.3c) [68]: a heavily doped and thick region under the contacts, a thin and lightly doped region under the passivating layer. These structures, known as "selective emitters", come at the expense of more complicate processing, usually involving photolithographic delineation and alignment of the diffusions.

If a very low SRV is possible, it would be best to have no emitter at all since doping always degrades bulk lifetime (Figure 7.3d). Examples are the back point contact solar cell and the point emitter design with bifacial contact [69], originally designed for concentration but capable of very high one-sun efficiency as well.

With localized contacts, surface recombination decreases, with the penalty of an increase in transport losses in the substrate: deeper gradients for minority carriers, or increased series resistance for majority carriers, because of current crowding near the contacts. The trading is more favorably solved as the contact size shrinks [70]. Light and/or localized diffusions have also the drawback of decreased gettering action.

7.3.3.4 Heterojunction solar cells

The structure of heterojunction solar cells is described in Section 7.6.2. It features (Figure 7.3e) a continuous transparent conducting electrode in contact with the substrate with an intervening amorphous silicon layer, reducing interface recombination to very low levels. An additional screen-printed grid is needed to provide sufficient lateral conductance.

7.3.3.5 Industrial cells

Screen-printing drastically affects the design of the emitter: it must be very highly doped to decrease the high contact resistance and not very shallow so that it is not perforated during paste firing which would short-circuit the junction. Besides, since the wide metal lines must be placed well apart in order to keep shading losses moderate, the emitter lateral conductance must be high, which also requires deep and highly doped regions. These characteristics are good at decreasing recombination at the contacts, but far from optimum at the exposed surface.

Industrial phosphorus emitters typically feature surface concentrations over 10^{20} cm^{-3} and 0.4 μm depth, resulting in a sheet resistance around 60 Ω. As already mentioned, the very highly doped region exhibits almost no photovoltaic activity due to the presence of precipitates ("dead layer"). As a result, the collection of short-wavelength light is very poor and J_0 large, irrespective of a hypothetical surface passivation which is therefore not implemented. As an advantage, heavy phosphorus diffusions produce very effective gettering. Some ways of incorporating selective emitters to screen-printed cells are being considered, SiN_x appearing very well suited for surface passivation. This, however, must be accompanied by a decrease in the finger width so that higher sheet resistances are tolerated.

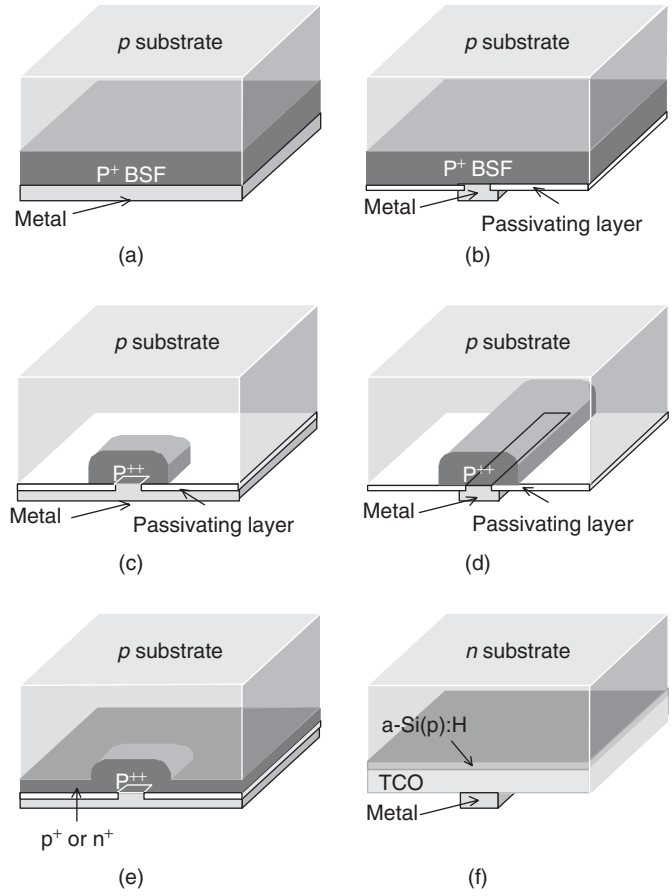


Figure 7.4 Rear contact structures: (a) continuous BSF; (b) bifacial cell; (c) local BSF; (d) local BSF, bifacial cell; (e) selective emitter or floating junction passivation; (f) HIT cell

7.3.4 The Back Surface

A p^+ layer is useful in decreasing contact recombination for cells with $w < L$, as explained. A BSF is a first step (Figure 7.4a). Today, most industrial solar cells feature an screen-printed Al BSF layer followed by an Ag or Ag-Al electrode.

Localized contacts, as shown in Figure 7.4b, further reduce recombination. This structure is presented by some bifacial cells [71]. If the surface passivation is good, the BSF is restricted to point contacts, some micrometers in size, as in PERL and similar cells (Figure 7.4c) [72]. The back face of bifacial cells passivated with SiN_x is shown in Figure 7.4d.

A shallow and light diffusion helps to decrease surface recombination (Figure 7.4e). The diffusion can be the same type of the substrate, or the opposite one: so-called PERT [42] and PERF [73] cells demonstrate these concepts. The latter structure benefits from the lower J_0 values of n^+ layers, and it is essential that no electron flow is injected from the n region to the p contact: the junction must be in open-circuit (a "floating" junction). A similar effect takes place in passivated boron-doped surfaces due to positive dielectric charges and boron surface depletion that lead to the

formation of inversion layers. These can produce shunting between electron and hole quasi Fermi levels, degrading the passivation.

Industrially feasible localized back contact designs are being addressed by several groups. Laser firing of aluminum through dielectric passivation layers is an interesting technique producing point contacts with local BSF [74].

7.3.5 Size Effects

Substrate edges are highly recombining surfaces that adversely affect cell performance, especially for small-size, large-diffusion-length devices. For laboratory cells, efficiency is defined on the basis of a design area. The emitter is limited to it by planar masking or mesa etching. The true edge is thus placed far away from the cell limit, and then recombination is reduced. For real applications, on the contrary, only the substrate area counts, and edge optimization is more complex. Advanced passivation schemes such as edge diffusions are being considered [75, 76]. In large industrial cells this recombination is much less important.

Large cell sizes are preferred by the industry, 12.5×12.5 cm or 15.6×15.6 cm being standard. Apart from fabricability concerns, a bigger cell means that more current must be collected to the terminals, making Joule losses grow: the longitudinal resistance of the metal lines increases quadratically with their length. This problem, more severe for coarser metallization techniques, along with a loss of homogeneity, makes efficiency decrease with increasing size. To alleviate series resistance at the price of increased shading, terminals are soldered to metal busbars inside the cell active area, thus decreasing the distance that current must be collected from along the fingers.

7.3.6 Cell Optics

Flat plate solar cells in operation are illuminated from a large portion of the sky, not only because of the isotropic components of radiation, but because of the Sun's apparent motion over the day and the year if no tracking system is used. So, regarding angular distribution, these cells must accept light from the whole hemisphere. The spectral distribution also varies with time, weather conditions, etc. For calibration purposes a standard spectral distribution AM 1.5 Global is adopted as a representative condition, generally specified at $0.1\,\mathrm{W\ cm^{-2}}$.

A solar cell should absorb all useful light. For non-encapsulated cells, the first optical loss is the shading by the metal grid at the illuminated face, if any. This loss is of the order of 7% for industrial cells while for laboratory cells using fine metallization it is much lower. Though several techniques have been proposed to decrease the effective shading, such as shaped fingers, prismatic covers, or cavities [77], their efficacy depends upon the direction of light and so they are not suited to isotropic illumination.

7.3.6.1 Antireflection coatings

The next loss comes from the reflectance at the Si interface, more than 30% for bare Si in air due to its high refraction index. A layer of non-absorbing material with a lower refractive index (n_{ARC}) on top of the Si substrate decreases reflectance: this is a step towards the zero-reflection case of a smoothly varying refractive index [78]. If the layer is thick in terms of the coherence length of the illumination, around 1 μm for sunlight, there are no interference effects inside it. The encapsulation (glass plus lamination) belongs to this category.

The term antireflection coating (ARC) is used to refer to an optically thin dielectric layer designed to suppress reflection by interference effects. Reflection is a minimum when the layer thickness is (an odd multiple n_{ARC} of) $\lambda_0/4$, with λ_0 the free space wavelength, since in this case reflected components interfere destructively. At other wavelengths reflection increases, but is always below the value with no ARC or, at most, equal [79]. The ARC is usually designed to present the minimum at around 600 nm, where the flux of photons is a maximum in the solar spectrum. For reflection to become zero at the minimum, the coating index should be the geometric average of those of air and silicon, i.e. 2.4 at 600 nm for non-encapsulated cells. Today, the industry uses SiN_x deposited by PECVD or by sputtering for this purpose.

By using double-layer coatings with $\lambda/4$ design, with growing indices from air to silicon, the minimum in reflection is broader in wavelength. Evaporated ZnS and MgF_2 are used by laboratory high-efficiency cells. The low index of passivating SiO_x in contact with Si degrades the performance of the ARC. The SiO_x layer is then made as thin as possible, compatible with efficient passivation [80].

7.3.6.2 Texturing

Alkaline solutions etch a Si crystal anisotropically, exposing {111} planes on which the etching rate is lowest. On [1 0 0]–oriented wafers, randomly distributed, square pyramids are formed whose size is adjusted to a few micrometers by controlling etching time and temperature. In a textured face, a ray can be reflected towards a neighboring pyramid (Figure 7.5a) and hence absorption is enhanced. Though calculation of reflection requires ray tracing, a rough estimate for near-normal incidence can be derived by assuming each ray strikes the Si surface twice so that reflection is the square of the untextured case. As multicrystalline substrates lack a single crystal orientation, alkaline etches are not efficient, and several alternatives are proposed to achieve similar effects of reflection reduction, as will be described in Section 7.4.1.

Texturing is incorporated in both industrial and laboratory Si solar cells, and, in combination with AR coating, reduces reflection losses to a few percent. In the latter case, in order to better control the pyramid geometry and to allow delineation of fine features on the surface, photolithographic techniques are used to define inverted or upright pyramids at the desired positions. It has to be noted that in this case the reflectivity is similar to that of a random texture [81].

Light entering the substrate at a textured surface is tilted with respect to the cell normal. This means that photogeneration takes place closer to the collection junction, which is very beneficial for

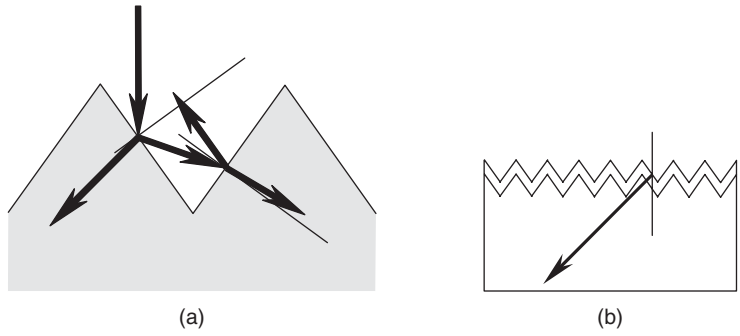


Figure 7.5 Effects of surface texturing: (a) decreased reflection; (b) increased photogeneration in the base

low-diffusion-length cells by enhancing the collection efficiency for medium to long wavelengths (Figure 7.5b). The effect is equivalent to an increase of the absorption coefficient. As a drawback, textured surfaces present higher SRVs.

7.3.6.3 Light trapping

Long-wavelength photons are weakly absorbed in silicon, and, unless internal reflectances are high, they will escape the substrate without contributing to photogeneration. The aim of light trapping or light confinement techniques is to achieve high internal reflectances.

Practical back mirrors can be implemented that are fully compatible with the cell electrical design [82], such as those schematized in Figure 7.4. A metal can make a good reflector, but Al, especially after heat treatment, gives low reflectance. The Si–oxide–metal structure in Figure 7.4c can present a high reflectance by capitalizing on interference effects [83].

At the front the metal mirror is not applicable because the ray paths must be kept open for the entering light. High front reflectance can still be achieved because of total internal reflection [77]. Rays striking the surface at angles larger than the critical air-Si one are totally reflected. Texturing one or both surfaces with macroscopic or microscopic features serves this purpose by tilting the rays. Even for the geometric texturing case with well-defined surface orientations, after a few internal reflections the direction of rays inside the wafer is randomly distributed: this is the lambertian case, a useful analytic approximation to light trapping. Bifacial structures in Figure 7.4 can, for the same reason, be very efficient at confining the light [84].

Light trapping increases the effective thickness of the wafer for absorption. In the geometrical optics regime, it has been shown that, for one-side isotropic illumination, the maximum enhancement factor (though perhaps not realizable) is $4\,(n_{Si}/n_{air})^2$, i.e. each ray traverses 50 times the cell thickness before escaping [85]. The corresponding enhancement in photogeneration will be lower because of the competition of the absorption by free carriers at long wavelength.

Light trapping is essential for thin cells. Even in the thick PERL design, it can offer around 1 mA cm^{-2} enhancement in short-circuit current with respect to the case where the internal reflectances were zero.

7.3.7 Performance Comparison

For illustration purposes, Table 7.2 collects relevant parameter of the Auger-limited ideal Si solar cell [35], the best one-sun PERL cell [42] and a typical screen-printed, industrial cell on mc-Si. The different concept behind each set of data must be accounted for when comparing the figures. For instance, the ideal cell is assumed to be isotropically illuminated, while measurements are made for near-normal incidence.

The most striking difference between the best and the ideal solar cell is the difference in design: thick and low injection *versus* thin and high injection. The PERL cell is surely the best design for the currently achievable levels of surface recombination, that limits open-circuit voltage and shifts the optimum thickness to high values. The low resistivity follows then from transport considerations for the chosen structure. The very high fill factor of the ideal cell is characteristic of high injection, Auger-limited operation.

Reduction of surface recombination in the best laboratory cells relies on surface passivation and the restriction of very heavily doped regions to a minimum. In the end, this is possible because of the possibility of defining and aligning very small features on the surfaces.

Table 7.2 Cell performance (25 °C, AM 1.5 Global 0.1 W cm^{-2})

Cell type	Ideal (calculated)	PERL (measured)*	Industrial (typical)
Size (cm^2)		4	225
Thickness (μm)	80	450	250
Substrate resistivity (Ω cm)	Intrinsic	0.5	1
Short circuit current density, J_{SC} (A cm^{-2})	0.0425	0.0422	0.034
Open circuit voltage, V_{OC} (V)	0.765	0.702	0.600
Fill factor, *FF*	0.890	0.828	0.740
Efficiency, η (%)	28.8	24.7	15.0

*The efficiency of this cell has been revised to 25% (±0.5%) after changes in the reference spectra (Green MA, Emery K, Hishikawa, Warta W, *Prog. Photovolt: Res. Appl*., **17**, 85–94 (2009)

The very heavily doped emitter, along with lower substrate lifetimes, is responsible for the reduced short-circuit current and open-circuit voltage in the industrial cell. The fill factor is affected by the large device area in conjunction with the limitations of the metallization technique, which reduce further the current because of shading.

Continuous improvement in material quality and cost-driven thinning of the substrates will increase the need for industrial cells to implement surface passivation schemes. This will require the refinement of the metallization technique; another issue is that substrate lifetimes in an industrial environment depend on the gettering action of very heavy diffusions, which are not compatible with optimum surface performance. The PERL approach – high-temperature processes and delineation of fine features – is the most successful path to high efficiency, but it is not the only one that can inspire the forthcoming developments in industrial cells. It is worth noting, in this respect, that back-contacted solar cells and HIT solar cells have entered the exclusive club of more than 20% efficiency with different approaches that will be explained in Section 7.6.

7.4 MANUFACTURING PROCESS

7.4.1 Process Flow

Figure 7.6 shows the main steps of a simple process for solar cell fabrication based on screen-printing. With more or less minor modifications, this process is currently in use by most of PV manufacturers [86]. The main virtues of this 35-year-old PV technology are easy automation, reliability, good usage of materials and high yield. The drawback, as explained in preceding sections, is the efficiency penalty derived from the coarse and aggressive metallization technique.

Each step is briefly described in the following with illustrative purposes: values of temperature, time, etc. will be only indicative.

7.4.1.1 Starting material

The industry uses so-called solar grade Cz-Si wafers, round in origin, but very often trimmed to a pseudo-square shape, or multicrystalline square wafers. Wafer dimensions are 12.5–15.6 cm side

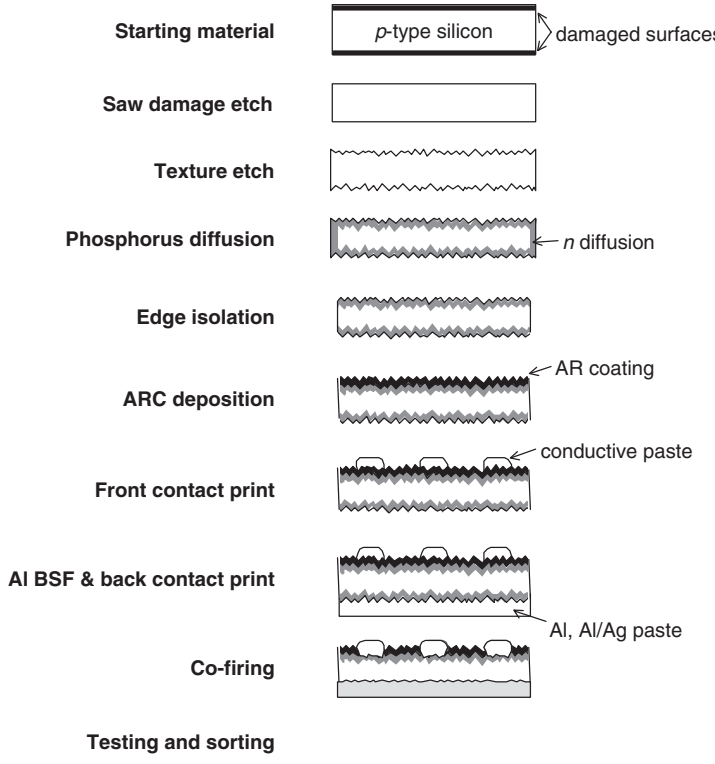


Figure 7.6 A typical processing sequence with schematic illustrations of the resulting structures

and 180–210 μm thickness, with a steady tendency to reduce in thickness. Doping is *p*-type (boron) to a resistivity in the range of 0.5–6 Ω cm.

7.4.1.2 Saw damage removal

The sawing operation leaves the surfaces of "as-cut" wafers with a high degree of damage. This presents two problems: the surface region is of very bad quality and the defects can lead to wafer fracture during processing [87–89]. For this reason about 10 μm are etched off from each face in alkaline or acidic solutions. The wafers, in Teflon cassettes, are immersed in tanks containing the solution under temperature and composition control. Alkaline etches are preferred to acidic solutions from considerations on waste disposal, but the latter are advantageous for multicrystalline material because they can provide isotropic texturing, as will be explained next. Plasma etching can also be used for saw damage removal.

7.4.1.3 Texturization

KOH etching, leading to microscopic pyramids, is commonly employed for monocrystalline material. Their size must be optimized, since very small pyramids lead to high reflection while very large ones can hinder the formation of the contacts. Figure 7.7 shows a SEM picture of a textured surface.

To ensure complete texturing coverage and adequate pyramid size, the concentration, the temperature and the agitation of the solution and the duration of the bath must be controlled (in

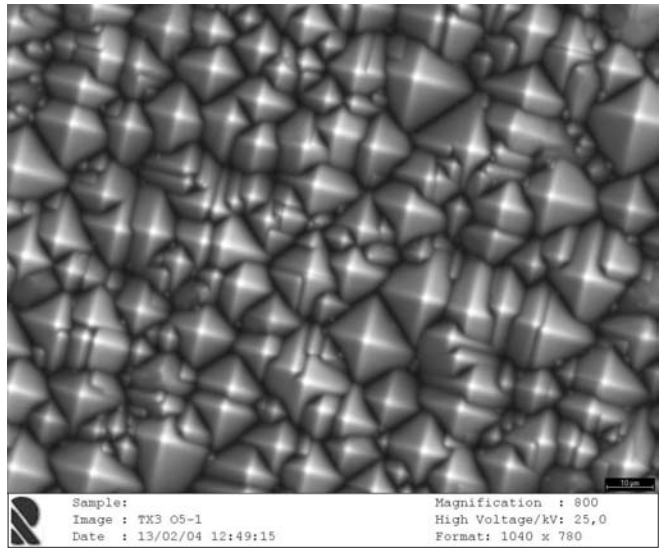


Figure 7.7 Mono-Si surface after alkaline texturing

fact KOH at a higher concentration and at a higher temperature is commonly used as an isotropic etch for saw damage removal). Alcohol is added to improve homogeneity through an enhancement of the wettability of the silicon surface. Typical parameters are 5% KOH concentration, 80 °C, 15 minutes [90].

Anisotropic texture with alkaline solutions is also applied to multicrystalline wafers, but with much poorer results. Reflectance of the textured wafers is relatively high because, for randomly oriented grains, etch rate is not that of (100) crystals. Another drawback is the existence of steps between grains, that may cause interruption in the screen-printed metal contacts.

That is why alternative processes are implemented. Their evaluation should take into account not only gains in reflectivity, but also surface damage and compatibility with metallization. The potential of some of these methods have been proved and developed for industrial applications, while others need still more research. The surface features produced by some of them are comparable in size or smaller than the wavelength, and geometrical optics is no longer applicable. They act as diffraction gratings, as scattering media or, in the limit of very small feature size, as graded index layers.

7.4.1.3.1 Acidic texturing

Several chemical techniques have been proposed. Some of them result in an inverted pyramid structure, but need photolithographic patterning, which is a serious drawback for compatibility with industry [91]. The result of nearly 20% for a mc-Si solar cell relies on an oxide-forced acidic texturing scheme [92]. A simpler approach is based on isotropic etching with an acidic solution containing nitric acid, hydrofluoric acid and some additives. The resulting etch pits of diameter 1–10 μm are uniformly distributed, giving a homogeneous reflectance over the surface of the wafer and the absence of steps between grains (Figure 7.8). An absorbent surface porous silicon layer needs to be removed, typically in an alkaline solution. Increases of short circuit current of about 5–7% are reported for solar cells processed on isotropic textured wafers compared with cells processed on anisotropic textured wafers [93, 94]. Some technical difficulties, such as depletion of the solution

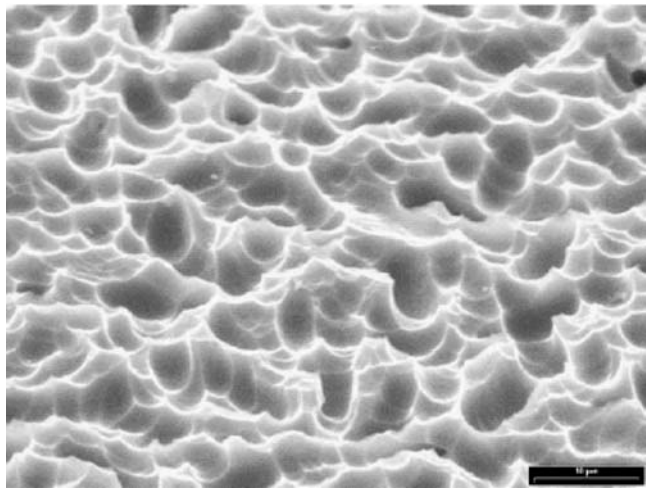


Figure 7.8 Mc-Si surface after acid etching. Reprinted from *Solar Energy Materials and Solar Cells*, **74**, 155–164, (2002) Szlufcik J, Duerinckx F, Horzel J, Van Kerschaver E, Dekkers H, De Wolf S, Choulat P, Allebe C and Nijs J, High-efficiency low-cost integral screen-printing multicrystalline silicon solar cells, with permission from Elsevier Science

and exothermic effects, can be encountered, and automatic wet benches, with temperature control of the etching solution and automatic replenishment of chemicals, have been designed to overcome these problems. In-line industrial equipment can be found in the market, in some cases performing both the saw damage removal and the texturing, simultaneously or sequentially.

Reduction in reflection by forming porous silicon is also under development [95]. A detailed analysis, taking into account the sum of reflectance and absorption within the porous Si layer shows an optimum of about 5–6% total optical loss. Besides its potential, the compatibility of the porous Si formation with screen-printed contacts needs to be addressed.

7.4.1.3.2 Plasma texturing

Reactive ion etching (RIE) texturization of silicon in chlorine or fluorine plasma is a dry isotropic etching process that creates a surface with a high density of steep etching pits, with typical dimensions less than 1 μm [96–99]. Increases in the range of 10% in short-circuit current compared with anisotropic texture have been reported with maskless techniques [100, 101]. RIE can also be performed in conjunction with a masking layer to produce more regular features [102]. The main obstacles in industrial implementation are the high global warming potential of the chemicals involved and the excessively low process throughput.

7.4.1.3.3 Mechanical texturing

V-grooves about 50 μm deep can be formed in Si wafers by mechanical abrasion using a conventional dicing saw and beveled blades, followed by alkaline etching to decrease surface damage. With this technique, average reflectivities in the range 6–8% have been obtained [103–105], as well as efficiency gains of 5% (relative) after encapsulation [106]. Contact fingers should be screen-printed parallel to the grooves, on plateaux left untextured to ensure easy printing, so that some kind of alignment is needed. Other contacting alternatives can be implemented, such as roller printing [107] or buried contact [108]. Automated systems are being developed to check industrial feasibility.

Table 7.3 Comparison of weighted AM1.5 reflectivities for mc-Si wafers with several surface treatments [201]

Reflectivity(%)	Alkaline textured	Acidic textured	Maskless RIE
Bare	34.4	27.6	11.0
With SiN AR coating	9.0	8.0	3.9
SiN and encapsulated	12.9	9.2	7.6

Another approach is scribing grooves by laser [109]. Upright pyramids of height 7 μm can be created by two orthogonal sets of parallel grooves, followed by a chemical etch to remove the silicon residues. Combined with a single-layer antireflection coating, laser texturing can reduce weighted reflectivity to 4%, half of that given by anisotropic etching and the same AR coating. Adjustments have been made to obtain smoother and smaller grooves, in order to adapt the technique to a screen-printed process.

7.4.1.3.4 AR coating and encapsulation

It has to be taken into account that cell reflection properties differ from those of texturing because it is usually complemented with AR coating (see below), and cell encapsulation, so that relative difference between several texturing method normally reduce, as can be seen in Table 7.3.

7.4.1.4 Phosphorus diffusion

Phosphorus is universally used as the *n*-type dopant for silicon in solar cells. Since solid-state diffusion demands high temperature, it is very important that the surfaces are free of contamination before processing. To this end, after the texturing the wafers are subjected to an acid etch to neutralize alkaline remains and eliminate adsorbed metallic impurities.

The industry uses a number of procedures to perform the phosphorus diffusion. The following classification is based on the type of furnace in which the high-temperature step takes place.

7.4.1.4.1 Quartz furnaces

The cells to be diffused, loaded in quartz boats, are placed in a quartz tube with resistance heating and held at the processing temperature (Figure 7.9). The cells enter and exit the furnace through one end, while gases are fed through the opposite one. Phosphorus itself can be supplied in this way, typically by bubbling nitrogen through liquid $POCl_3$ before injection in the furnace. Solid dopant sources are also compatible with quartz furnace processing. Tubes are typically open and operated at atmospheric pressure, although low-pressure equipment is available that may improve uniformity and throughput [110]. 20–30 minutes at temperatures in the range 830–860 °C can be considered representative. As suggested in Figure 7.6, both surfaces and the edges of the wafer will be diffused.

7.4.1.4.2 Belt furnaces

In this case a phosphorus source is applied to one or both wafer faces and, after drying, the wafer is placed in a conveyor belt passing through the furnace (Figure 7.9). Dopant sources can be screen-printed [111], spun-on [112], sprayed-on [113–115], deposited by CVD [116] or by vaporization [117].

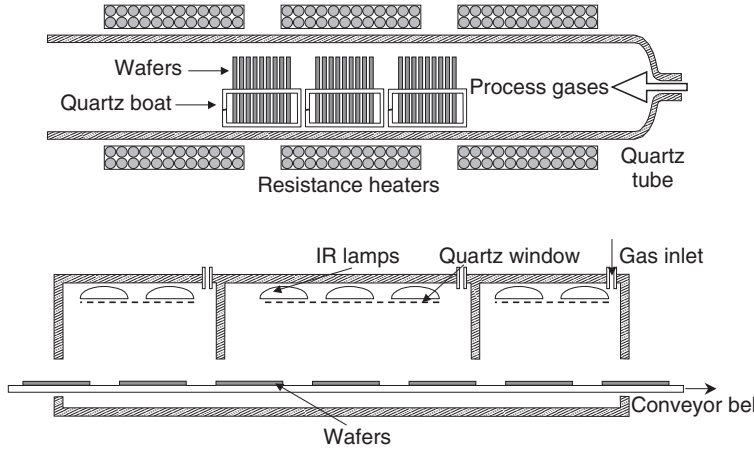


Figure 7.9 A quartz furnace (top) and a belt furnace (bottom) for the diffusion of phosphorus

Infrared heating is usually employed, offering the possibility of faster heating/cooling rates of the wafer. The temperature inside the furnace can be adjusted in several zones and, though it is open, gases can be supplied. The temperature cycle undergone by the wafer will mimic the temperature profile along the furnace with the timescale set by the advancing speed of the belt. The process cycles in terms of time and temperature are similar to those described for quartz furnace diffusion. In principle, one face of the wafer is diffused, but as the parasitic junction formed at the edges is also present in this technique due to diffusion from the gas phase and should be removed (see below), sometimes both faces are diffused to take advantage of an enhanced gettering effect.

The main benefit of the quartz furnace is cleanliness, since no metallic elements are hot and no air flows into the tube. Though this is a batch step, high throughput can be achieved since many wafers can be simultaneously diffused in each tube, commercial furnaces consisting of stacks of four tubes. In a belt furnace, the ambient air can get into the furnace and the hot conveyor belt is a source of metallic impurities. The assets of belt furnaces are found in automation and in-line production, throughput and the ability to implement temperature profiles. New designs try to reconcile the advantages of both types of equipment [118].

After diffusion, an amorphous glass of phosphosilicates remains at the surface that is usually etched off in dilute HF because it can hinder subsequent processing steps.

7.4.1.5 Junction isolation

The *n*-type region diffused at the back in the quartz furnaces need not be removed, as it is compensated by the subsequent aluminum deposition step. But the *n*-type region at the wafer edges present in both quartz and belt furnaces would interconnect the front and back contacts: the junction would be shunted by this path translating in a very low shunt resistance. To remove this region, several procedures can be used.

Laser grooving of the wafer edges is the most widely used option [119]. 20-μm-deep and 60-μm-wide V-grooves are sufficient for an effective isolation. In this case, the step is performed after metallization to avoid losing the isolation due to the thermal treatment.

Another option is single-side etching by acidic solutions, thanks to the availability of equipment where the wafers are floating in the chemical bath, supported by the surface tension of the

liquid, so that only the rear side and edges are wetted [120]. An advantage of the chemical approach is that it can be combined with the removal of the phosphorus silicate glass in in-line equipment.

7.4.1.6 ARC deposition

Traditionally, titanium dioxide (TiO_2) was used for creating the antireflection coating due to its near-optimum refractiven index for encapsulated cells. A popular technique was atmospheric pressure chemical vapor deposition (APCVD) from titanium organometallic compounds and water [121]. This process is easily automated in a conveyor-belt reactor. Other possibilities included to spin-on or screen-print appropriate pastes.

But nowadays, hydrogenated silicon nitride films is the preferred option, as it combines its antireflection properties with others of bulk and surface passivation. Films can be deposited by several techniques, but the most commonly used process is chemical vapor deposition (CVD), involving the reaction of silane gas and amonnia. Plasma-enhanced chemical vapor deposition (PECVD) is preferred to other CVD technologies (atmospheric-pressure CVD or low-pressure CVD) because it is a low-temperature process ($T < 500\,^{\circ}C$), and that means reducing complexity and preventing lifetime degradation.

PECVD techniques induce hydrogenation, whose benefits for silicon are well known [122, 123]. Amorphous silicon nitride films are produced by PECVD with up to 40 atom % of hydrogen (i.e. although these films are usually referred to as SiN_x they are really a-SiN_x:H). A subsequent thermal step is needed to activate hydrogenation, and in an industrial process metal firing step fulfills this objective [124].

Additionally, surface passivation due to SiN_x deposition by PECVD has also been reported [125]. Achievable surface recombination velocity on a phosphorus-doped emitter is similar to that of a high-quality oxide passivated one, and a value as low as 4 cm/s has been obtained on a polished 1.5 Ω cm FZ *p*-type silicon wafer [24].

These three different properties (AR coating, bulk passivation and surface passivation) cannot be varied independently, an optimization of processing parameters (temperature, plasma excitation power and frequency, gas flow rate) is necessary, and a compromise should be reached [126, 127]. Furthermore, there are different PECVD techniques giving different results.

In "direct" PECVD, schematized in Figure 7.10a, the processing gasses are excited by means of an electromagnetic field, and the wafers are located within the plasma. Bulk is effectively passivated, but surface damage is produced due to direct exposition of wafers to plasma, precluding the achievement of good surface passivation. Furthermore, surface passivation degrades with exposition to UV light.

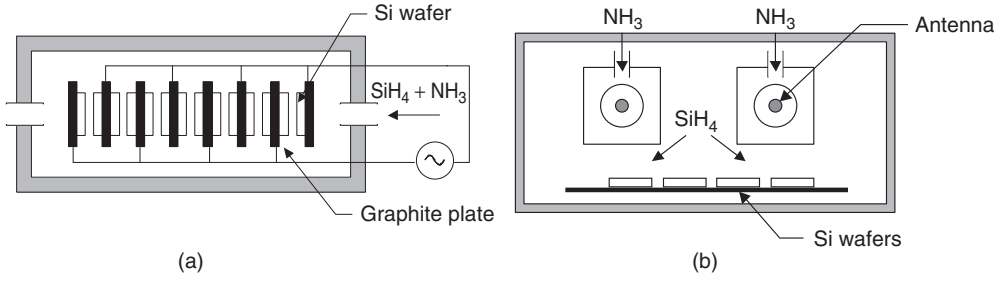


Figure 7.10 Industrial PECVD reactors. (a) Direct-plasma reactor; (b) remote-plasma system

There is a high-frequency direct PECVD (13.56 MHz) and a low-frequency one (in the range 10–500 kHz), the former being better in terms of surface passivation and UV stability. On the other hand, it is more difficult to obtain uniform layers.

A different approach is the "remote" PECVD, where wafers are located outside the region where the plasma is formed. Surface damage is avoided in this way, so that better surface passivation is achieved. On the other hand, bulk passivation is reduced. Figure 7.10b shows a sketch of an industrial remote PECVD. It implements a continuous feed of wafers, an advantage that should be compared with the batch-type direct PECVD.

Another technology able to achieve similar surface and bulk passivation properties as those of PECVD is sputtering [128], with the advantage of avoiding the use of the pyrophoric silane gas. For this process, wafers are moved horizontally through the in-line system, where silicon targets are alternately sputtered in argon and nitrogen to deposit the silicon nitride film onto the silicon wafer. Nitrogen and ammonia may be added to vary the refractive index and the hydrogen content of the film.

7.4.1.7 Front contact print and dry

The requirements for the front metallization are low contact resistance to silicon, low bulk resistivity, low line width with high aspect ratio, good mechanical adhesion, solderability and compatibility with the encapsulating materials. Resistivity, price and availability considerations make silver the ideal choice for the contact metal. Copper offers similar advantages, but it does not qualify for screen-printing because subsequent heat treatments are needed, during which its high diffusivity will produce contamination of the silicon wafer.

Screen-printing compares very unfavorably with vacuum evaporation in the first three requirements. It has been already commented how this affects the cell design and leads to a significant efficiency gap between laboratory and commercial cells, but throughput and cost compensate for this.

Screen-printing will be described in more depth in the following section. It is used to stick a paste containing silver powder to the front face of the wafer in the comb-like (fingers plus busbars) pattern. Automatic screen-printers are available capable of in-line, continuous operation with high throughput. These machines accept wafers from packs, cassettes or a belt line, place them with sufficient accuracy under the screen and deliver the printed wafers to the belt line. The paste is a viscous liquid due to the solvents it contains; these are evaporated in an in-line furnace at 200–250 °C. The dried paste is apt for subsequent processing.

7.4.1.8 Aluminum layer print and dry

A highly doped *p*-type region at the back can be easily formed by screen-printing an aluminum paste, forming the back-surface field (BSF) layer [129]. The low eutectic temperature of the Al–Si system (577 °C) means that some silicon will dissolve in the Al and recrystallize epitaxially upon cooling after the firing step, forming the *p*-type BSF layer. The characteristics of this layer (thickness, uniformity, reflectivity) depend on the amount of paste, which is in the mg/cm^2 range.

7.4.1.9 Back contact print and dry

As the soldering onto the Al contacts is not possible, a silver and aluminum paste is used to print busbars which will be soldered to the tabs to form the arrays of cells in the module, as will be explained briefly.

7.4.1.10 Co-firing of metal contacts

A high-temperature step is still needed: organic components of the paste must be burnt-off, the metallic grains must sinter together to form a good conductor, and they must form an intimate electric contact to the underlying silicon. As Figure 7.6 shows, the front paste is deposited on an insulating layer (the AR coating) and the Al BSF plus the back contact on the parasitic *n*-type rear layer (or directly onto the base if this parasitic layer has been eliminated during the isolation).

Upon firing, the active component of the front paste must penetrate the ARC coating to contact the *n*-emitter without shorting it: too mild a heat treatment will render high contact resistance, but too high a firing temperature will motivate the silver to reach through the emitter and contact the base. In extreme situations this renders the cell useless by short-circuiting it. In more benign cases, small shunts appear as a low shunt resistance or dark current components with a high ideality factor that reduce the fill factor and the open-circuit voltage.

The back paste, in its turn, must completely perforate the parasitic back emitter to reach the base during the firing, when present.

In order to comply with these stringent requirements the composition of the pastes and the thermal profile of this critical step must be very carefully adjusted.

Another aspect that should be considered is that thermal behavior of the Al paste is different from that of silicon, causing bowing for wafers in the range of 200 μm thickness. This effect can be minimized by quenching the wafers after firing, which is performed with coolers placed at the end of the IR furnace.

7.4.1.11 Testing and sorting

The illuminated $I-V$ curve of finished cells is measured under an artificial light source with a spectral content as similar as possible to sunlight (a flash lamp is the preferred option). Cell temperature is also measured, to extrapolate results to 25 °C. Defective devices are then rejected, and the rest are classified according to their output.

The manufacturer establishes a number of classes attending, typically, to the cell current at a fixed voltage near the maximum power point. Modules will subsequently be built with cells of the same class, thus guaranteeing minimal mismatch losses.

If, for instance, cell currents within a class must be equal within 5%, the accuracy and stability of the system must be better than that. Automatic testing systems are available that meet the very demanding requirements of high-throughput processing.

7.4.2 Screen-printing Technology

Screen-printing is a thick film technology, a terminology that opposes it to the usual microelectronic procedures of evaporation of thin films. It consists in translating a layer of a material in a desired pattern to the surface of the wafer. Though it can be employed for virtually any step in solar cell manufacturing, contact formation constitutes the most demanding, frequent and conspicuous application of screen-printing. Screens and pastes are the essential elements of the technology [87].

7.4.2.1 Screens

Screens are tight fabrics of synthetic or stainless steel wires stretched in an aluminum frame, as sketched in Figure 7.11. The screen is covered with a photosensitive emulsion, which is treated

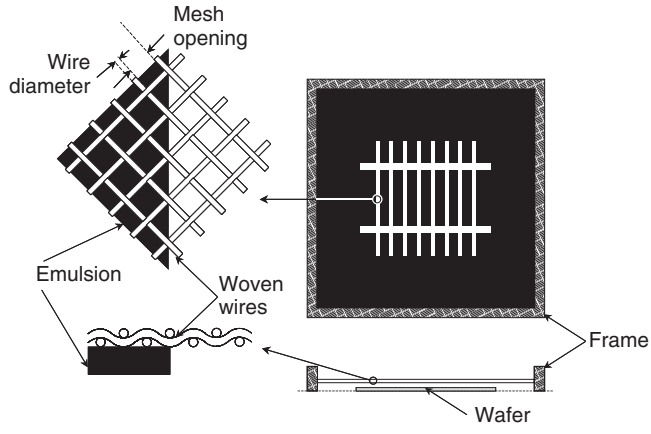


Figure 7.11 A screen for transferring the top contact pattern to a solar cell

by photographic techniques in such a way that it is removed from the regions where printing is desired.

For printing fine and thick layers, as needed for the front contact of a solar cell, the wires must be very thin and closely spaced [130]. On the other hand, the opening of the reticule must be several times larger than the largest particle contained in the paste to be printed. Screens for the front contact typically feature 325 wires per inch, wire diameter around 30 µm, mesh opening around 50 µm, corresponding to near 40% open surface, i.e. not intercepted by wires, and a total thickness (woven wires plus emulsion) around 90 µm. For the rear contact, figures are somewhat different: 200 wires per inch, wire diameter 40 µm, mesh opening 90 µm, corresponding to 50% open surface and total thickness of 110 µm.

7.4.2.2 Pastes

The pastes are the vehicles that carry the active material to the wafer surface. Their composition is formulated to optimize the behavior during printing. A paste for the metallic contacts of the solar cell is composed of:

Organic solvents that provide the paste with the fluidity and rheology required for printing.
Organic binders that hold together the active powder before its thermal activation.
Conducting material, which is a powder of silver (or aluminum, or both) composed of crystallites of a size of tenths of micrometers (this amounts to 60–80 % in weight of the paste).
Glass frit, 5–10% in weight, a powder of different oxides (lead, bismuth, silicon...) with a low melting point and high reactivity at the process temperature, that enables movement of the silver grains and etches the silicon surface to allow intimate contact.

The paste composition is extremely important for the success of the metallization and is critically linked to the heat treatment.

7.4.2.3 Printing

Figure 7.12 illustrates the process of printing a paste through the patterned emulsion on a screen. The screen and the wafer are not in contact, but a distance apart, called the snap-off. After dispensing

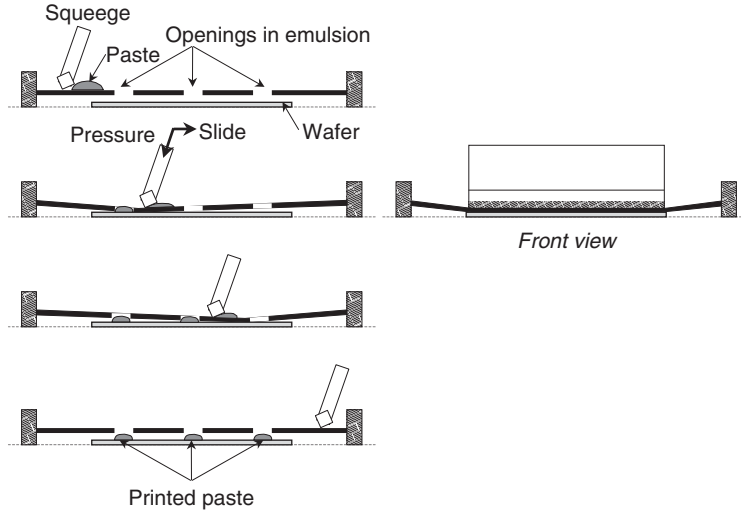


Figure 7.12 Illustration of a printing sequence

the paste, pressure is applied to the squeegee, which can be made of metal or rubber: this puts the screen in contact with the wafer. The squeegee is then moved from one side of the screen to the opposite one, dragging and pressing the paste in front of it. When an opening is reached, the paste fills it and sticks to the wafer, remaining there after the squeegee has passed and the screen has elastically retired.

The amount of printed paste depends on the thickness of the screen material and the emulsion and the open area of the fabric. It also depends on the printed line width.

The viscous properties are of the utmost relevance: when printing, the paste must be fluid enough to fill without voids all the volume allowed by the fabric and the emulsion, but after being printed it must not spread over the surface.

Critical parameters of this process are the pressure applied on the screen, the snap-off distance and the velocity of the squeegee.

7.4.2.4 Drying

Solvents are evaporated at 200–250 °C right after printing so that the wafer can be manipulated without the printed pattern being damaged.

7.4.2.5 Firing

Firing of the pastes is usually done in an IR belt furnace. After a preheating step, the organic compounds that bind the powder together are burnt in air at 500–600 °C, then temperature is raised to 750–800 °C to form the BSF, and a peak over 900 °C is reached to form the frontal contact. The complete process lasts a few minutes. Crystal orientation and paste composition must be considered too. In a last step the wafer is cooled down.

The phenomena taking place during firing are very complex and not completely understood. The oxides forming the glass frit melt, enabling silver grains to sinter and form a continuous

conductor so that the layer can present low sheet resistance. Neither the silver melting point nor the silicon–silver eutectic temperature is reached, sintering consisting in the intimate contacting of solid silver crystallites. At the same time, the reactive molten glass etches some silicon, and silver grains are allowed to form intimate contact with the substrate. The amount of etched silicon is of the order of 100 nm. When a layer of SiN is present the glass frit is able to etch through it. In fact, the quality of the contact improves because of a better homogeneity.

The picture of the contact after cooling down shows two zones [131]. In the inner one, crystallites of silver are plugged into silicon forming crystalline interfaces and presumably very good electrical contact in a sort of "point contact". These grains are embedded in a compact amorphous glass. The outer zone is more porous and contains silver grains and glass frit: this porosity explains why the resistivity of silver paste is much larger than that of pure silver.

Besides, the contact resistance of printed contacts is much higher than that of an evaporated contact to *n*-Si of the same doping. It seems that, although enough silver grains make good contact to silicon, not all of them are connected to the grains in the outer layer, but many remain isolated by the glass.

When the paste, as in the case of the back metallization, contains aluminum, the Al–Si eutectic formed and recrystallized ensures a good contact. With dielectric layers the contact appears to be localized as well, and some beneficial role is attributed to the metal atoms in the frit [132].

7.4.2.6 Limitations and trends in screen-printing of contacts

As explained in previous sections, the high contact resistance and the etching action of the glass frit require the front emitters to be highly doped and not very thin if screen-printing is used. Only improved paste formulation and processing can overcome this limitation.

Narrow, but thick fingers with good sheet conductance are also needed. Well-defined lines must be much wider than the pitch of the woven fabric; 60 µm lines seem achievable, with 100 µm ones being standard (Figure 7.13). Incrementing the amount of transferred paste implies increasing the thickness of the emulsion or the pitch-to-diameter ratio of the wires, which are both limited. Besides, screens become deformed with usage and a continuous deterioration of printed patterns is observed.

In screen-printing the wafer is subjected to considerable pressure. This can pose a problem with very thin or irregular wafers, such as those obtained by sheet growth of silicon, which can break down.

Metal stencils [133] can outperform screens: they produce finer lines with better aspect ratio, endure more printing operations without degradation and need less cleaning and maintenance. Roller printing [134] or pad printing [135] have also been proposed as high-throughput alternatives giving narrower lines.

Other metallization concepts are also investigated to overcome the limitations of screen-printing technology, an overview can be found in reference [136].

7.4.3 Throughput and Yield

Because of the rapidly growing demand, photovoltaic factories are quickly expanding their production volumes so that there is a strong driving force to increase the throughput of processes and equipment. Automation is being extensively applied to the fabrication of solar cells and in-line, continuous processing tends to displace batch steps: in the process outlined above, only chemical

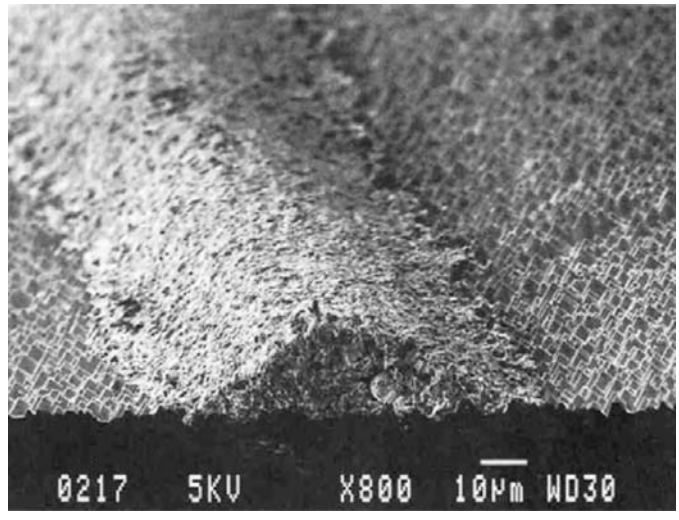


Figure 7.13 Screen-printed contact finger; texture pyramids are also visible. Reprinted from *Solar Energy Materials and Solar Cells*, **41/42**, 101–107, (1997), Nijs J, Demesmaeker E, Szlufcik J, Poortmans J, Frisson L, De Clerq K, Ghannam M, Mertens R, Van Overstraeten R, Recent improvements in the screenprinting technology and comparison with the buried contact technology by 2D-simulation, with permission from Elsevier Science

etches and tube diffusion are batch steps. Automation and large-scale production lead to reduced costs [137, 138].

Most processes described above have been borrowed from the electronics industry: diffusion, plasma etching, etc. are standard in microelectronics, while screen-printing was extensively used by thick film technology for hybrid circuits. The character of the industries being different, the requirements for equipment differ, and it is to be expected that substantial improvements for photovoltaics will take place now that the business volume makes it attractive for equipment manufacturers to get involved.

A modern fabrication line is capable of processing around 1500–3000 wafers per hour, i.e. an operation in a cell takes 1–2 s. Of course, the slowest operation along the flow line will limit the overall throughput. In order to get an estimate of how this translates into yearly production, let us consider $15.6 \times 15.6\,\mathrm{cm}^2$ cells with 15.8% efficiency ($3.85\,\mathrm{W_p}$ power per cell). If the line operates without interruption and all wafers are successfully processed, during one year it will produce:

$$3.85\mathrm{W_p/cell} \times 2000\ \mathrm{cells/hour} \times 24\mathrm{hours/day} \times 365\mathrm{days/year} \cong 67.5\mathrm{MW_p/year}$$

This number has to be decreased by: (i) the downtime of the equipment due to maintenance, repair, etc.; and (ii) the yield, i.e. the percentage of defective or broken wafers. Allowing for both would give a throughput in the range of $50–55\,\mathrm{MW_p}$/year per production line with available, commercial equipment.

Yield is a most important parameter for cell production: it can be defined as the ratio of successful finished cells to starting wafers. Since PV technology is material-intensive, yield has a strong influence on cost. Breakage and poor electrical performance are the causes of low yield, which is, generally speaking, benefited by automation. In this respect, in-line quality control acquires a great relevance to quickly detect and amend problems affecting yield.

For a given time per operation per cell, the throughput increases if the power of the cell increases. This is attained by increasing the cell area and the efficiency, which also helps decreasing costs. When going to larger areas, series resistance and the uniformity of the obtained layers (emitter, AR coating), that may compromise the electrical performance, become important issues. Besides, larger cells are more difficult to handle without breaking and the yield may be affected.

There is a lot of room for efficiency improvement for industrial solar cells, and many of the processes to realize it are proven in the laboratory. Their potential to reduce production costs is high if they are successfully transferred to the industrial environment [139].

7.5 VARIATIONS TO THE BASIC PROCESS

This section introduces some variations to the basic process described above that aim at improving the efficiency, the throughput or the cost. While some modifications are already in production, others are still being developed at the laboratory.

7.5.1 Thin Wafers

Wafering and sheet-growth techniques improve and produce thinner substrates, with wafer thickness around 150 μm or below being envisaged for the near term [140, 141]. When processing these thin cells several relevant issues appear.

The probability of fracture during handling increases, especially in conjunction with a larger size. Adequate handling tools must be designed. Some steps appear to be critical: for instance, in chemical baths convection can exert significant torque on the wafers. This issue is fostering the study of the mechanical properties of silicon [142–144] and even the development of new crystallization procedures.

The behavior during heat treatments is modified due to a decreased thermal mass. On the other hand, wafers can more easily become bowed [145]. Processes need to be specifically optimized for thin cells [146–148].

Thin cells largely depend on surface passivation and optical confinement. If attained to reasonable degrees, efficiency improvement comes as a bonus for thin cells, but otherwise the performance is degraded. New optimal structures must be developed.

7.5.2 Back Surface Passivation

The enhancement of material quality and the decrease of wafer thickness will make it necessary to improve the passivation of the back surface. Several approaches are feasible to be incorporated to the basic screen-printing process:

- *Boron back surface field*. The use of boron instead of aluminum as BSF has the potential to increase cell performance, and at the same time avoids the bowing of thin cells [149]. The diffusion from a liquid source in a quartz tube is problematic because of the high temperatures needed, so that diffusion from solid sources is preferred for easier integration in the basic process. It would be very attractive to diffuse both phosphorus and boron during the same thermal step, but simultaneous optimization of the dopant profiles is not so straightforward [150].
- *Dielectric passivation*. A number of techniques for rear passivation based on deposition of dielectric layers (silicon oxide, nitride and carbide, or aluminum oxide) have proven their

excellent performance of surface recombination velocity, as already explained in Section 7.2.2, and efforts are devoted to implement them on industrial solar cells. To do so, the compatibility of the deposition conditions with the solar cell process should be considered, as well as their ability to reflect light back to the silicon wafer, and how demanding they are in terms of rear surface morphology [136]. For instance, high- temperature steps can degrade the passivating properties of the layer, or shunting can be produced by interaction between the layer and the metallization [151].

- *Amorphous Si passivation*. Amorphous hydrogenated silicon is used as a passivating layer in the HIT solar cell, that will be described in Section 7.6.2, and it has been also applied on more conventional cell structures, reaching high efficiencies [152, 153].

7.5.3 Improvements to the Front Emitter

Quantitative improvements in both recombination currents and spectral response will be derived from front surface passivation only when better paste formulations and finer line prints allow more resistive emitters – thinner and/or less doped – to be used [154]. In that case, tube oxidation and silicon nitride deposition appear as good candidates for industrial use.

Selective emitters are being developed for screen-printed solar cells. A number of techniques have been proposed that are compatible with the screen-printing process:

- Two separate diffusions for metallized and unmetallized regions, the heavy diffusion being restricted to the regions to be metallized by a screen printed or deposited mask [155, 156].
- A homogeneous thick emitter is diffused by conventional means; a screen-printed or inkjet mask is applied to protect the regions to be metallized from the plasma etch that follows. In this way, in the unprotected regions the emitter is thinner and less doped [157, 158].
- Self-aligned selective emitter by diffusion from a patterned solid dopant source: under the dopant source a highly doped emitter is formed, while a much lighter diffusion from the gas phase takes place at the uncovered regions [159].
- First a high sheet resistance emitter is formed to which self-doping pastes containing phosphorus as well as silver are applied. Firing is performed above the silver–silicon eutectic temperature, thus leading to the formation of an alloyed layer heavily doped with phosphorus [160].

Most of these techniques require some kind of pattern aligning to print the front contact fingers on the heavily-doped regions. Automatic screen-printers feature enough accuracy to perform this task. Another aspect that should be considered is that some of these alternatives lose the gettering effect associated with supersaturated P diffusion.

Obviously, the unmetallized part of these emitters is sensitive to surface passivation, which must thus be implemented by oxidation or nitride deposition.

7.5.4 Rapid Thermal Processes

In conventional furnaces of the closed-tube or conveyor-belt types not only the wafers are heated to the process temperature, but the equipment itself: chambers, substrates or boats, etc. This brings about: (i) long heating/cooling times, due to the large thermal masses involved; (ii) a high potential for contamination, since a lot of parts, some of them metallic, are held at a high temperature; and (iii) high energy consumption.

On the other hand, the microelectronics industry has developed in recent years the so-called rapid thermal processes (RTP), whereby only the wafers, and not their environment, are

heated to high temperature. Selective heating is accomplished by intense UV illumination of the semiconductor. The interest of RTP for solar cell fabrication comes from very short thermal cycles, down to a few minutes including heating and cooling, so that throughput can be boosted. Besides, the absence of hot parts in the equipment diminish potential contamination and energy consumption.

At the laboratory scale rapid thermal diffusion of very thin emitters (which is beneficial from the electrical point of view) has been demonstrated. Rapid thermal firing of screen-printed contacts and aluminum alloying have also been successfully implemented, as well as other promising techniques such as rapid thermal nitridation and oxidation of the surfaces for passivation purposes [161, 162]. It can be said that every thermal step in the solar cell process can be made by RTP.

A possible drawback of the technique is a degradation of substrate lifetime compared with conventional processing for some materials, due to the formation and quenching in of defects because of the very fast heating and cooling cycles, and possibly to precluded gettering action [163].

A hurdle to industrial deployment of RTP is the lack of suitable equipment, since microelectronics uses one-wafer reactors while photovoltaics would need large-capacity batch reactors or, better, in-line continuous equipment. It seems possible to furnish conveyor-belt furnaces with UV lamps to obtain these industrial RTP reactors, but some problems such as temperature uniformity must be solved [164].

7.6 OTHER INDUSTRIAL APPROACHES

Other commercially available technologies will be described in this section. They all look for a decrease in the $/W figure of merit, following different approaches:

- using ribbons, presented in Chapter 6, as substrates;
- implementing techniques that do not need high-temperature processes: the HIT cell, based on a-Si/x-Si heterojunction emitter;
- performing both *n* and *p* contacts at the back of the solar cells;
- processing a series of Si strips that are partially attached to a 1-mm-thick Si wafer: the "sliver cell".

7.6.1 Silicon Ribbons

Ribbon technologies offer a cost advantage over crystalline silicon, thanks to the elimination of the slicing process (see Chapter 6). They cover for the moment around 3% of the PV market, edge-defined film-fed growth (EFG) being the most mature of them, while string ribbon (STR), dendritic web (WEB, recently revisited as ribbon on a sacrificial template, RST) and other approaches are also into industrial production at different levels of development.

To account for the high density of defects (dislocations, grain boundaries, impurities, etc.), a specific solar cell process is needed for ribbon substrates, optimizing the P diffusion and Al screen-printing steps to enhance the gettering effect, and the PECVD silicon nitride deposition to induce hydrogenation for bulk defect passivation. Other aspects related to the uneven surface of the sheets that should be addressed are the difficulty of texturing with both alkaline and acidic etches, or the problems of using screen-printing metallization, that in some cases is replaced by pad printing and direct writing (extrusion) of silver pastes and inks [165].

For EFG and STR, efficiencies in the range 15–16% have been reached in large-area substrates with industrial techniques, which in some cases include screen-printed contacts fired

with RTP [166, 167], while more sophisticated processing allows the 18% efficiency level to be achieved [168, 169].

Regarding dendritic web, an n^+np^+ structure (phosphorus front diffusion and rear Al alloyed emitter) has been industrially implemented on a high-resistivity antimony-doped substrate. Due to the low substrate thickness (100 µm), it can benefit from the location of the *pn* junction at the back, performing an effective front surface field, enabling a high diffusion length and being immune to light-induced degradation. Using only production-worthy, high-throughput processes, dendritic web cells have been fabricated with efficiencies up to 14.2% [170].

Other ribbon approaches are in a development stage, with efficiencies in the 12–13% range, but with the potential of low silicon utilization (see for example [171] for the RGS technology).

7.6.2 Heterojunction with Intrinsic Thin Layer

The HIT (heterojunction with intrinsic thin layer) structure makes use of the cheaper amorphous silicon (a-Si) technology, depositing a-Si layers on crystalline silicon by PECVD [18]. It provides an excellent surface passivation with very-low-temperature processes (below 200 °C), avoiding lifetime degradation of the bulk material. On the other hand, passivation is very dependent on surface morphology and cleanliness, and it can be lost if the temperature is raised in the subsequent process steps.

Figure 7.14 shows the structure of the HIT cell. A textured *n*-type Cz-Si substrate is used. The emitter and BSF are made of *p*-type and *n*-type a-Si layers, respectively. Very thin intrinsic a-Si layers are inserted between a-Si and the crystalline substrate, to improve the characteristics of the a-Si/c-Si interface. Thickness of these amorphous layers is of the order of 10–20 nm. On both doped layers, transparent conductive oxide (TCO) layers are formed by sputtering, and metal fingers are screen-printed. Back metallization is also comb-like to reduce thermal and mechanical stresses, making the cell symmetrical and able to perform as a bifacial cell.

Cell efficiencies over 21% have been demonstrated [172]. HIT modules have been industrially produced since 1997, and special modules exist for roof-tile and bifacial applications.

7.6.3 All-rear-contact Technologies

Several cell structures have been proposed that bring both contacts to the back face, either adopting the scheme of Figure 7.1b (interdigitated back contact, IBC cells), or that of Figure 7.1d

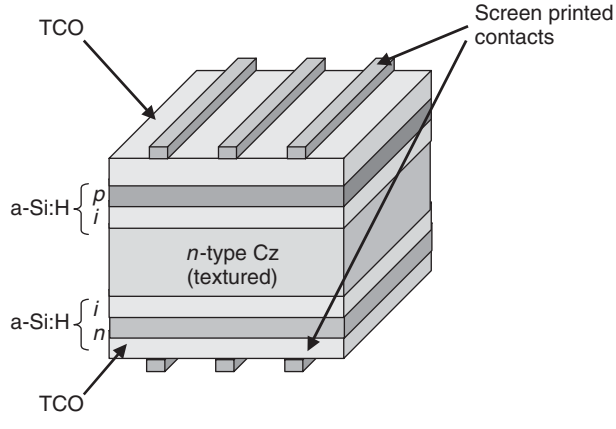


Figure 7.14 Structure of the HIT cell

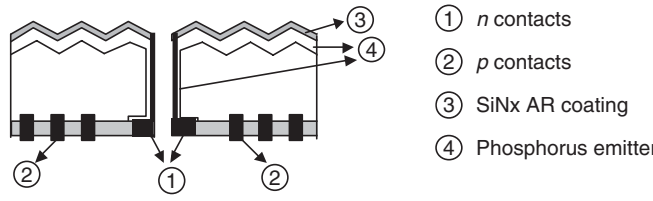


Figure 7.15 Sketch of the metallized wrap through concept

(metallized wrap through, MWT, or emitter wrap through, EWT cells) [40]. Although the approach and the performance are different in both cases, similarities can be found in module assembly (see Section 7.7).

As already mentioned in Section 7.3.1, record efficiencies for concentration silicon solar cells have been obtained with an IBC concept [39], a technology that is now industrially exploited by several companies around the world. The more sophisticated concentration solar cell process has been adapted to an industrial process for one sun operation, using low-cost screen-printing technology to pattern the rear boron and phosphorus diffusions, silicon dioxide for surface passivation and Ni plating for metallization [173]. The more complex process is compensated by the achievement of remarkable efficiencies, which have already surpassed the 22% level in the industrial line [174].

The IBC structure needs high-quality substrates to permit carrier collection at the rear junction, so, in order to implement an all-rear-contacted structure on low-quality substrates, such as multicrystalline silicon, phosphorus layers are diffused at both faces, and afterwards internally connected by drilling holes in the silicon, either metallized to connect the front fingers to the rear busbars (in the MWT approach [175]), or phosphorus diffused, locallizing the complete negative electrode grid on the rear (EWT [176]). An example is sketched in Figure 7.15. Efficiencies in the range 15–16% have been reached in both cases with industrial-type processing [177–179].

7.6.4 The Sliver Cell

In the sliver cell process, thick Si wafers (in the range of 1 mm) are micromachined to create narrow grooves that separate very thin (~50 μm) Si strips ("slivers"). Each wafer can contain several thousand slivers with an effective combined surface area of 20–50 times the wafer surface area. While still supported by the wafer at their edges, the solar cell processing is performed, and the increase in the active surface area per unit volume of silicon makes affordable to implement many high-efficiency features, achieving solar cell efficiencies over 19% on high-quality float zone substrates [180]. Due to the characteristics of the slivers, bifacial, transparent and flexible modules can be fabricated with them.

7.7 CRYSTALLINE SILICON PHOTOVOLTAIC MODULES

The power of a single solar cell being small, several of them must be electrically associated to make a practical generator. The module is the building unit for generators that can be purchased in the market, i.e. it is the real PV product. Performance and lifetime of PV systems depend on the protection that module construction offers to the active photovoltaic devices.

The basic module fabrication procedure in use by most manufacturers was developed three decades ago and is briefly described below. Modules for special applications (building integration, marine operation, etc) require slight modifications of the process and the materials.

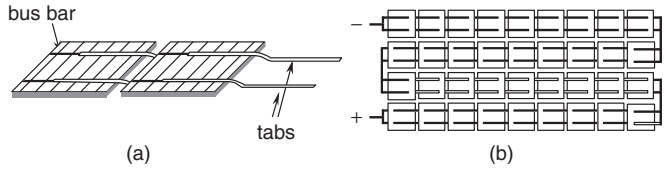


Figure 7.16 Cell interconnection with tabs: (a) two cells in series; (b) layout of 36 series-connected cells

7.7.1 Cell Matrix

In a module the cells are usually arranged in series. For that, tinned copper ribbons (tabs) are used to connect the front of one cell to the rear side of the adjacent one (Figure 7.16a). In such a way strings of typically 9–12 series-connected cells are formed.

It has to be noted that tabs must overlap a long distance along the busbar length since the conductance of the printed busbars is too low. Two (or three for larger cells) tabs per cell are employed, thus providing redundancy which allows current to flow in case electrical continuity is broken due to some failure [181]. Besides, the effective length of grid fingers is one-fourth (or one-sixth) of the cell side, and series resistance is alleviated. Tabs provide a non-rigid link between cells, allowing thermal expansions to be accommodated.

In the past, stringing was done in a two-step process: tabs were first soldered to the front side, followed by the soldering to the rear side of another cell. However, it has to be taken into account that during the last few years wafer thickness has been decreased to below 200 μm. In this case a large bowing of the solar cells during the soldering due to the difference in thermal expansion coefficient between silicon and copper is produced (Figure 7.17a). This bowing would cause, above all, an increased breakage during the subsequent rear soldering and during lamination step. In order to overcome this problem the attachment of the tabs (front and rear) on the same cell is done simultaneously. Nevertheless, thermal stress induced during the heating and cooling would lead to microcracks (Figure 7.17b), which degrades module performance after some time in the field.

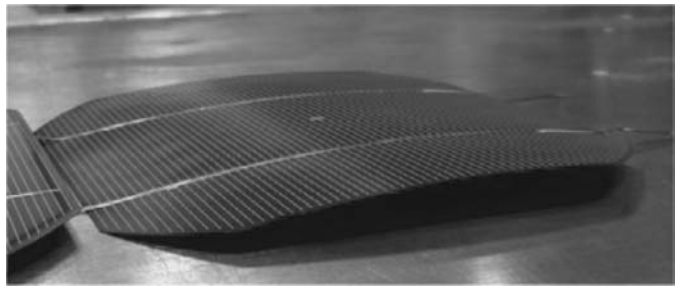
Conductive epoxies or low-temperature solder alloys [182] can replace conventional ones and illumination or induction is widely used instead of former iron heating, in order to reduce the stress during the stringing process.

Several strings are interconnected using auxiliary ribbons or a printed circuit board (PCB) to form the cell matrix. Typically it consists of a single series string (Figure 7.16b) although several strings internally paralleled is also a possibility.

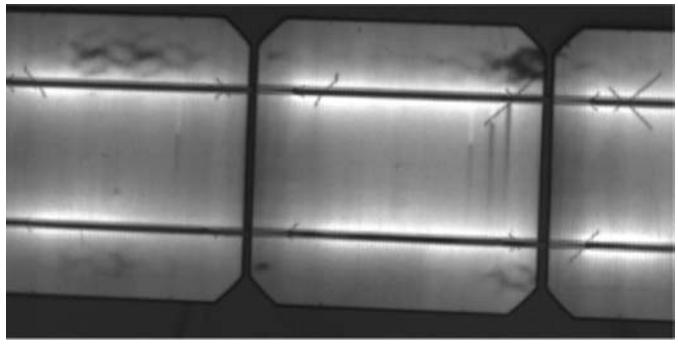
Traditionally, a typical module configuration used 36 series-connected cells, which, under operating conditions, would produce around 15 V at maximum power, appropriate for 12 V battery charging [183]. As grid-connected applications and, to a less extent building-integrated systems, grow, modules with different electrical configurations enter the market, so that the standard now is 72 cells of $125 \times 125\,\text{cm}^2$ and 60 cells of $156 \times 156\,\text{cm}^2$.

7.7.2 The Layers of the Module

The array of cells must be properly encapsulated for reliable outdoor operation for more than 25 years, attending to rigidity to withstand mechanical loads, protection from weather agents and humidity, protection from impacts, electrical isolation for people safety, etc.



(a)



(b)

Figure 7.17 (a) Bowing on a solar cell after tabbing and (b) thermography of microcracks induced by stress during stringing

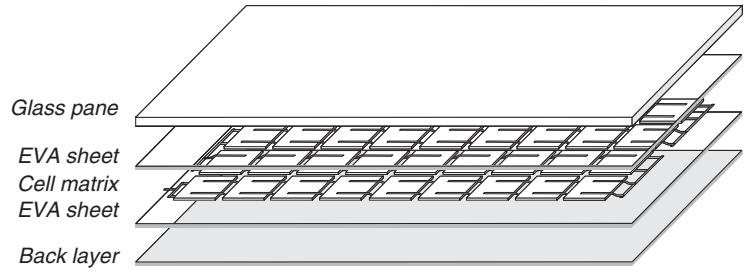


Figure 7.18 Stack of materials to be laminated

The different layers the module is made of are then stacked. A common structure is sketched in Figure 7.18.

A 3- to 4-mm-thick soda-lime glass is used as a superstrate that provides mechanical rigidity and protection to the module while allowing light through. It must have low iron content or otherwise the light transmission will be low. Tempered glass must be employed to increase the resistance to impacts and for safety reasons if a module breaks.

The cell matrix is sandwiched between two layers of the encapsulant or pottant material. Except for special application such as double-glazed modules for building integration the encapsulant is the copolymer EVA (ethylene vinyl acetate), a plastic composed of long molecules with

a backbone of carbon atoms with single covalent bonding. EVA is a thermoplastic, i.e. its shape changes under heating are reversible. It is sold in rolls of extruded film around 0.5 mm thick. Along with the polymer, the film contains: (i) curing agents whose role will be described later and (ii) UV stabilizers.

The back layer, at the non-illuminated module side, is usually a composite plastic sheet acting as a barrier for humidity and corroding species that also should provide electrical isolation for safety reasons. Typically it is formed by a laminate of three layers. The outer one is Tedlar® (commercial name of a flouropolymer by DuPont) that is an excellent barrier for external agents, however it does not provide high-voltage electrical isolation; for that a second layer made of polyester is used. For the inner part another layer of Tedlar® or a layer of EVA glued to polyester is commonly used. Some novel backsheets are based only in polyester.

7.7.3 Lamination

This step is carried out in a laminator, a table that can be heated and furnished with a cover that tightly closes the edges. The cover has an internal chamber and a diaphragm that separates this from the chamber containing the module. Both chambers can be independently evacuated: this configuration allows the module to be kept in a vacuum while mechanical pressure is exerted on it.

During the lamination process, both chambers are evacuated while the temperature is raised above the EVA melting point (around 120 °C). Vacuum is important to extract air (to prevent voids from forming), moisture and other gases. The EVA flows and soaks the cells. After a few minutes, with the module chamber still in vacuum, the upper chamber is filled with air so that the diaphragm presses the laminate. When temperature reaches 150 °C, the curing stage begins: the curing agents induce cross-linking of the EVA chains, i.e. chemical bonds are formed transversally among the long molecules that before curing are only weakly linked to one another. The plastic then acquires elastomeric, rubber-like properties, and indeed the curing step is a close analogy to the vulcanization of rubber. The total process takes around 12–15 min for standard EVA [184, 185].

Lamination used to be a bottleneck in the module fabrication process. To improve throughput several solutions have been followed by the industry: (i) commercial ultra-fast-curing EVA formulations that have allowed the reductions of curing time to less than 10 min nowadays; (ii) alternative materials based on silicones, polyurethanes, ionomers or polyolefins with process time in the range of 2–4 min are under evaluation; and (iii) a large lamination area – up to 10 m^2 – or in a multilevel layout enabling simultaneous process of several modules or very large ones.

7.7.4 Post-lamination Steps

These include: (i) trimming the edges of the laminate to remove spread-out encapsulant; (ii) sealing them with silicone rubber to prevent different thermal expansion of the aluminium frame and the glass; (iii) sticking the plastic junction box at the back of the laminate and performing the connections; and (iv) when required, installing the anodized aluminum frame (Figure 7.19). The frame must be electrically insulated from the active cell circuit so that high voltage differences can be sustained between electrical terminals and the frame without current flow.

Besides, among other final tests, the $I-V$ curve of all modules under standard conditions is measured in a solar simulator to check they fulfill specifications. Flash simulators with electronic equipment able to record a complete $I-V$ curve in a fraction of a second are commonly utilized for saving energy, increasing the throughput and avoiding differences in temperature during the

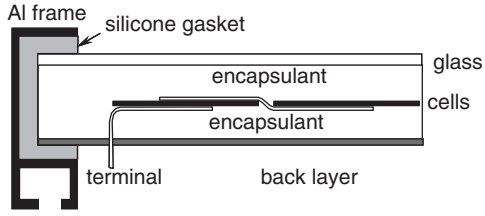


Figure 7.19 Cross- section of a standard module

measurement. They must have a spectral content matched to the AM 1.5 standard, or else, they must be calibrated with a reference cell of the same technology in order to correct the spectral mismatch.

7.7.5 Automation and Integration

Up to mid-1990s, with factory capacity of a few tenths of MW, all the operations in the module manufacturing were performed manually. Even more, most of the equipment used, such as laminators or sun simulators, was in most cases developed in-house by the manufacturers.

In a second step, with factories capacity in the range of several tens of MW, sophisticated equipment that performs most of the operations automatically was used. Both throughput and yield benefit from automation since the connected cells are very fragile and difficult to handle, leading to a dramatic cost reduction. Equipment manufacturers from other sectors with less growing potential (for example, automotive industry) developed specific equipment for the PV industry. Nevertheless, most of the processes (stringing, lamination, framing, etc.) remain in island configuration, with intermediate buffers to facilitate the management of the production routine.

Current manufacturing plants have a capacity in the range of hundreds of MW, with expectations of reaching the GW level in very near future. At this scale, full automation and integration of the whole module manufacturing chain is essential. In such a way, operation processing is improved by minimizing the work in progress and handling, thus reducing costs. On the other hand, standardization of both materials and models is required in order to get all the benefits for this integrated concept. Finally, production management should be supported by automatic process control.

7.7.6 Special Modules

7.7.6.1 BIPV products

Building integration of PV modules (BIPV) is thought to become an important application of photovoltaics in the near future. Modules perform two tasks: construction materials as well as power generators. Modules can be incorporated to a building in a number of ways and special products are being developed so that the typical framed module will be no longer the only PV product. Very large modules with special fixing for roof or façade integration, roof tiles with cells, semi-transparent modules allowing light through, etc. Visual appearance is enhanced by module shape, encapsulation and cell color [186]. Besides, these products must comply with building norms such as fire resistance.

In general modules for these applications use two glass panes, and another polymeric material as encapsulant (polyvinyl butyral, PVB) is an alternative to EVA. It was used in the early days

of module fabrication and it is very common in the glass industry. It is processed in a similar way to EVA and can present some advantages over it [187], but it requires low temperature storing.

7.7.6.2 Bifacial modules

Several cell structures have been presented that can operate with bifacial illumination. By encapsulation with a transparent backsheet (usually plastic one although glass is also possible), bifacial modules offering increased power output per unit cell area can be produced without technology changes. In spite of their potential, their presence in the market is very small for the moment.

7.7.6.3 Modules with back contact cells

Back contact cells are interconnected without tabs by soldering them to a layer with the connection paths printed, similar to PCB practice in electronic circuits. These experimental designs offer simplified module fabrication and enhanced visual appeal.

7.8 ELECTRICAL AND OPTICAL PERFORMANCE OF MODULES

7.8.1 Electrical and Thermal Characteristics

The voltage of the module is, in principle, the number of series-connected cells times the voltage of the single cell, and the module current the number of paralleled cells times the single-cell current. Whichever the combination is, the module power theoretically equals the power of a single cell times their number; however due to ohmic losses (mainly in the tabs), optical losses and the mismatch effects a reduction in power of around 3–4% is observed. Mass-produced modules offered in the catalogs of manufacturers show power ratings that typically range from 150 to 300 W_p, delivered at current levels 5–10 A and at voltages 20–40 V. Lower and higher values in all the parameters are possible for special applications.

The manufacturer usually provides values of representative points (short-circuit, open-circuit and maximum power) of the module $I-V$ curve measured at standard cell conditions (STC), i.e. 1 kW m^{-2} irradiance ($= 0.1 W \cdot cm^{-2}$), AM 1.5 spectral distribution and 25 °C cell temperature. The maximum power of the module under STC is called the peak power and given in watts peak (W_p). While efficiency has the greatest importance for a solar cell, for a module it is less relevant since part of the area is not occupied by the expensive solar cells.

The conditions in real operation are not the standard ones; instead, they vary strongly and influence the electrical performance of the cell, causing an efficiency loss with respect to the STC nominal value. This loss can be divided into four main categories [188]:

1. *Angular distribution of light*. Due to the movement of the Sun and the diffuse components of the radiation, light does not fall perpendicular to the module, as is the case when measurements are done and the nominal efficiency is determined.
2. *Spectral content of light*. For the same power content, different spectra produce different cell photocurrents according to the spectral response. And the solar spectrum varies with Sun's position, weather and pollution, etc. and never exactly matches the AM 1.5 standard.
3. *Irradiance level*. For a constant cell temperature, the efficiency of the module decreases with diminished irradiance levels. For irradiances near one sun, this is primarily due to the logarithmic

dependence of open-circuit voltage on photocurrent; at very low illumination the efficiency loss is faster and less predictable. Low shunt resistance in the device implies larger decreases in the efficiency at low irradiance levels.

4. *Cell temperature*. The ambient temperature changes and, because of the thermal insulation provided by the encapsulation, light causes cells in the module to heat above the ambient value; higher temperature means reduced performance. This is usually the most important performance loss.

On the other hand, prediction of the module response under different conditions is required to correctly assess the yearly production of a PV system in the field. The physical mechanisms of influence of temperature and irradiance on cell performance are well known, so that in principle prediction of module output could be rooted on physical models. This is however unpractical and a different approach is followed by PV system engineers.

Instead, very simple methods are used for translating the $I-V$ performance to different operating conditions and standardized procedures have been developed for PV modules of industrial technologies [189]. These methods are applicable within a limited range of temperature and irradiance conditions that are not very far from those met when testing the module, and require a small number of easily measurable parameters. The module datasheets from the manufacturers use to include some of these allowing simplest estimates to be made:

1. The steady-state power balance determines cell temperature: input is the absorbed luminous power, which is partially converted into useful electrical output and the rest dissipated to the surroundings. Convection is the main mechanism for heat dissipation in terrestrial, flat-plate applications, and radiation is the second non-negligible mechanism of heat dissipation. A common simplifying assumption is made that the cell-ambient temperature drop increases linearly with irradiance. The coefficient depends on module installation, wind speed, ambient humidity, etc., although a single value is used to characterize a module type. This information is contained in the nominal operating cell temperature (NOCT), which is defined as the cell temperature when the ambient temperature is 20 °C, irradiance is 0.8 kW m^{-2} and wind speed is 1 m s^{-1}. NOCT values around 45 °C are typical. For different irradiance values G this will be obtained:

$$T_{cell} = T_{ambient} + G \times \frac{NOCT - 20^{\circ}\mathrm{C}}{0.8\mathrm{kW}\cdot\mathrm{m}^{-2}}$$

2. The module short-circuit current is assumed strictly proportional to irradiance. It increases slightly with cell temperature (this stems from a decrease in bandgap and an improvement of minority carrier lifetimes). The coefficient α gives the relative current increment per °C. By combining both assumptions, the short-circuit current for arbitrary irradiance and cell temperature is calculated as:

$$I_{SC}(T_{cell}, G) = I_{SC}(\mathrm{STC}) \times \frac{G}{1\mathrm{kW}\cdot\mathrm{m}^{-2}} \times (1 + \alpha(T_{cell} - 25^{\circ}\mathrm{C}))$$

For crystalline Si α is around or 0.025% per °C (derived from an increase of 9 μA/cm^2·°C for a single cell).

3. The open-circuit voltage strongly depends on temperature (the main influence is that of the intrinsic concentration), decreasing linearly with it. Knowledge of the coefficient, called β, allows the open-circuit voltage to be predicted:

$$V_{OC}(T_{cell}, G) = V_{OC}(\mathrm{STC}) - \beta(T_{cell} - 25^{\circ}\mathrm{C})$$

The irradiance dependence is implicit in T_{cell}. For crystalline Si, β is a little above 2 mV/°C per series-connected cell, that is, around 0.4% per °C.

4. A lot of factors affect the variation of the maximum power (or, equivalently, the efficiency) with irradiance and temperature. The parameter γ is defined as the relative decrease in module efficiency per °C of cell temperature increase:

$$\eta(T_{cell}, G) = \eta(\text{STC}) \times (1 - \gamma(T_{cell} - 25^{\circ}\text{C}))$$

Usual γ values are near 0.5% per °C.

7.8.2 Fabrication Spread and Mismatch Losses

So-called mismatch losses arise when cells with different $I-V$ characteristics are interconnected because of the fewer degrees of freedom left to bias the devices, so that the array output is less than the sum of the powers the individual cells could deliver. The differences come from the unavoidable fabrication spread or from nonuniform irradiance or working temperature within the array.

To minimize mismatch losses, finished cells are measured and sorted in the factory. For series connection the important parameter is the current at the maximum power point (mpp). It is common practice to measure the current before encapsulation at a fixed voltage close to the mpp, and to classify the cells accordingly, though other classification criteria are possible [190]. Inside each class all devices present similar currents within the specified tolerance that ensures that, when connected in series to form the module, the mismatch loss will be below the desired limit [191]. Depending on the class being processed, the power rating of the resulting module will vary and this explains why manufacturers offer different module families although they are built in exactly the same way.

7.8.3 Local Shading and Hot Spot Formation

Due to local shading or failure, one or several solar cells can present a much smaller short-circuit current than the rest of devices in the series string. If the defective cells are forced to pass a current higher than their generation capabilities, they become reverse-biased, even entering the breakdown regime, and sinking power instead of sourcing it.

Figure 7.20 illustrates this behavior for an 18-cell string with one cell shaded so that its short-circuit current is half that of the remaining devices. String short-circuit is marked with an horizontal line, showing that in this condition the shaded cell is strongly reverse-biased, and dissipating the power produced by the unshaded cells. This effect of course severely degrades the efficiency of the module, but more important is the fact that it can get damaged.

Avalanche breakdown is characterized by a nonuniform distribution of current across the junction, breakdown occurring preferentially at localized regions, possibly correlated to damage during processing. Intense local heating can produce very high temperatures (a hot spot). If around 150 °C are reached, the lamination material becomes degraded and the module deteriorates irreversibly [192, 193]. Due to the localized nature of the process, solar cells show large scattering in their reverse characteristics so that the module behavior under partial shading is not accurately predictable.

In order to devise the means of preventing hot spot failure, the worst case is considered. This occurs when the N-cell series string is short-circuited and a shaded solar cell is reverse-biased with the voltage of the remaining $N-1$ good devices, as shown in Figure 7.20. The minimum N that will lead to hot spot formation (i.e. the maximum N for safe operation) depends on rather uncontrollable factors, as explained. For Si solar cells of standard technology it is around 15–20. Devices fabricated on upgraded metallurgical (UMG) silicon show a breakdown voltage of 10–12 V.

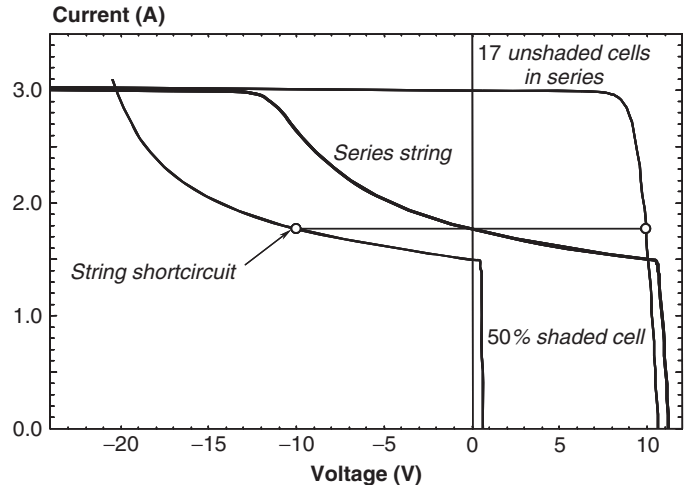


Figure 7.20 Computer simulation of the $I-V$ curves of a 50% shaded cell, showing the typical "soft" reverse breakdown, and of 17 identical cells, unshaded, in series. When series-connected with the shaded cell, they curve labeled "series string" is obtained

Since larger series strings are generally used, the approach followed is to put a diode (bypass diode) in parallel, but in opposite polarity, with a group cells. The number of cells in the group is chosen so that hot spots cannot be formed. When one or several cells are shaded, they are reverse-biased only to the point where the diode across the group starts forward conduction. The diode carries away the necessary current to keep the group near short-circuit.

Figure 7.21 illustrates the operation of the bypass diodes. When the current forced through the shaded sub-string is such that the reverse bias equals the diode threshold voltage, the bypass diode sinks all necessary current to keep the string at this biasing point, hence preventing the power dissipated in the shaded cell from increasing. It is also apparent that the bypass diode leads to a significant increase of output power, allowing the module to keep delivering the power generated by the unaffected groups.

It is clear then that the smaller the number of cells per bypass diode, the lower the efficiency loss for a shading condition, but this means a higher cost and more complex fabrication. It has been proposed to integrate a bypass diode in each cell so that these effects will be minimized at the expense of a more complicated cell processing [194].

The practice is to take electrical terminals outside the encapsulation, not only for the extremes of the series string, but for intermediate points as well so that bypass diodes are connected in the junction box each 18–20 cells (Figure 7.22), or 12 cells for UMG-Si. Endurance to shading is a standard test for module qualification.

The influence of local shading on the module output depends on the details of the $I-V$ curve of the cells as well. Under certain circumstances of partial shading, it is beneficial that the cells show some shunt resistance. However, tight control of leakage currents by processing is not easy.

7.8.4 Optical Properties

The encapsulation affects the optical properties of the cells in several ways. The optical properties of the cells must be optimized, attending to cost and performance after encapsulation.

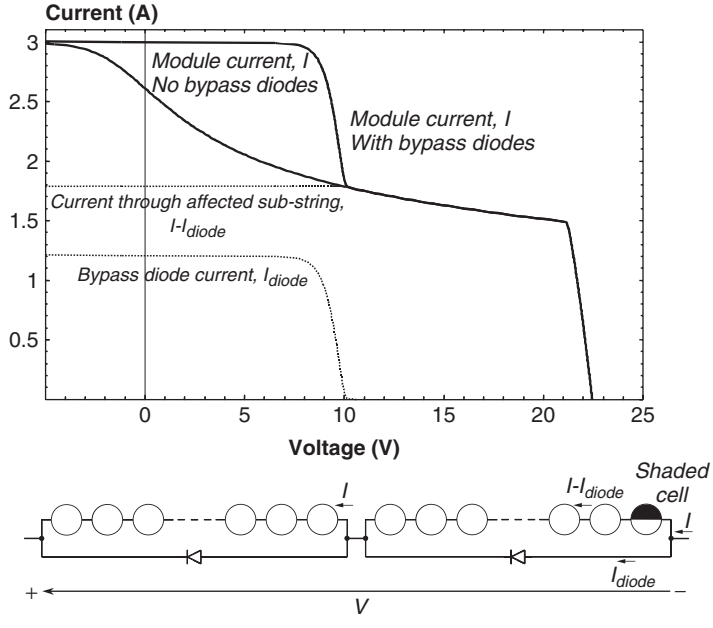


Figure 7.21 Computer simulation of the $I-V$ curves of a 36-cell series string without and with two bypass diodes, connected as shown in the bottom of the figure, when one cell is 50% shaded. The currents through the shaded sub-string and its bypass diode are also shown

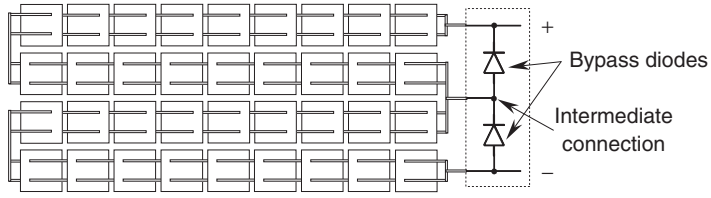


Figure 7.22 Two bypass diodes in a 36-cell module. The connections are done in the junction box

Some effects of encapsulation are [195]:

- The refractive index of glass and EVA is similar, around 1.5, between those of air and Si. Encapsulation acts then as a thick AR.
- The design of the ARC coating must account for the fact that the cell is illuminated from a medium with this index. The optimum ARC refractive index is slightly larger than in air.
- Glass and EVA absorb some light in the short-wavelength range.
- Typically, 4% reflection occurs at the air–glass interface (Figure 7.23(1)). ARC coatings and texturing can be applied to decrease this loss and to increase energy output [196].
- The light reflected by the metal fingers and the cell surface, if the reflected rays are tilted with respect to the normal to the glass surface, can be partly recovered by total internal reflection at the glass–air interface (Figure 7.23(2)). This effect could be enhanced by texturing the cell surface with tilted pyramids, instead of the upright pyramids obtained by alkaline etching of (1 0 0) surfaces [197].

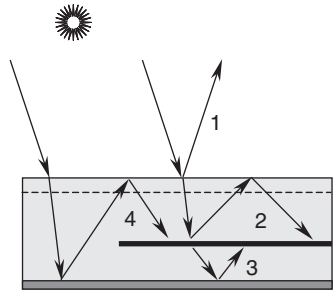


Figure 7.23 Optical effects of encapsulation: (1) glass reflection; (2) trapping of cell reflectance; (3) trapping of cell transmittance; (4) collection of peripheral light

- Although the trapping capabilities of the cell, due to the lower difference in refractive index, appear to worsen with encapsulation, the escaped rays are trapped in the glass so that the absorption enhancement in the ideal case is not affected.
- The white backsheet, since it reflects diffusively, allows some of the light incident between the cells to be collected (Figure 7.23(4)).

7.9 FIELD PERFORMANCE OF MODULES

7.9.1 Lifetime

Long lifetime is claimed as one of the main virtues of PV, and some manufacturers currently offer warranty for more than 20 years, with 30-year lifetime being the objective for short-term development. This should mean that for this period of time the module will keep working, i.e. producing electrical power with an efficiency of 80% compared with the starting efficiency, and without deterioration that compromises the safety or the visual appearance. Two factors determine lifetime: reliability, that refers to premature failure of the product, and durability, that refers to slow degradation that eventually decreases production to unacceptable levels. Cost-effectiveness, energy pay-back balance and public acceptance of photovoltaic energy strongly rely on the reliability and the long lifetime of modules.

PV systems worldwide have been working for more than 20 years, and this is allowing us to gather information concerning degradation mechanisms. Modules in the field are subjected to static and dynamic mechanical loads, thermal cycling, radiation exposure, ambient humidity, hail impact, dirt accumulation, partial shading, etc. Common failure modes [181, 198] are related to the action of weather agents in combination with deficiencies in fabrication.

Location-dependent steady degradation of module output is also observed, with short-circuit current and fill factor being the most affected parameters. In many cases this has been proved to correlate with degradation of EVA encapsulation [185]. EVA, like most polymers, is known to undergo photothermal degradation; UV radiation breaks molecular chains. Diffusion of chemical species through it is also relatively easy, so that moisture and corroding agents can enter while absorbers and stabilizers can out-diffuse.

Yellowing or browning of EVA reduces its optical transmission, affecting module current. For this reason EVA incorporates UV absorbers in its novel formulations. Degradation also decreases the strength of the encapsulant, leading to loss of adhesion to the cells and even detachment of

the layer (delamination). This is promoted by the shear stress that accompanies different expansion coefficients upon diurnal thermal cycles. Delamination brings about optical and thermal degradation. Besides, the degraded encapsulant is more easily penetrated by moisture and chemicals. Among these, sodium from the glass and phosphorus from the cell emitter are known to precipitate at the cell surface, corroding solder joints and increasing series resistance [198]. Encapsulant formulations are being continuously improved to address these problems.

7.9.2 Qualification

Several organisms, such as the International Electrotechnical Commission (IEC), the Institute of Electrical and Electronic Engineers (IEEE), have designed tests aimed at guaranteeing the quality of PV products [199]. Test procedures have been defined that, if successfully passed by a product, should guarantee the reliability of the PV module.

Manufacturers voluntarily submit their products to qualification tests in an accredited laboratory. These include verification of the module performance claimed in the datasheets as well as reliability tests. The certifications obtained are intended as a quality assurance for the customer.

Qualification tests consist in verifying the module integrity by visual inspection, measurement of the electrical performance at STC and of the electrical isolation before and after treatments that simulate, in an accelerated manner, real operation conditions. For instance, the IEC Standard 61215 [200] specifies:

- Ultraviolet exposure using xenon lamps.
- Thermal cycling (-40 °C to 50 °C, 50 cycles) in climate chamber.
- Humidity freeze cycling (thermal cycling with 85% relative humidity).
- Damp heat (1000 hours at 85 °C and 90% relative humidity).
- Twist test for testing resistance to torques.
- Pressure is applied to the module to test resistance to static mechanical loads.
- Hail impact test, where the module is struck by 25-mm-diameter ice balls at a speed of 23 m/s.
- Outdoor exposure.
- Hot spot tests, where the module is selectively shaded.

Different test combinations are applied to a sample of a few modules. The modules will qualify if no major failures are found, the visual inspection reveals no damage, the electrical power is within 90% of specifications, and isolation is maintained.

7.10 CONCLUSIONS

This chapter has reviewed the current state of crystalline silicon solar cells and modules. The main lines defining the structure of the described situation can be summarized as:

- *Changing scale*. The current boom in the markets enables and fosters technological and processing improvements.
- *Laboratory–industry gap*. There is a mature technology at the laboratory that has led to impressive performance levels, on one hand, and a reliable, fast, 30-year-old industrial process producing modest efficiency, on the other one: closing this gap is the key to a lower $\$/W_p$ figure of merit.
- *Novel silicon materials*. Market growth and the threat of silicon shortage stimulates new materials and very thin substrates that demand new technological solutions.

- *Technology diversification*. These two challenges are to be faced by solar cell production technology in the next years. Intensive preindustrial research is being conducted and solutions are being developed along several different lines.
- *Quality*. Product reliability and durability and environmental and aesthetic quality are as important as cost for the growth of PV industry, and this also influences technology.
- *Long-term scenario*. Alternatives to crystalline silicon technology are being thoroughly pursued, and presumably some of them will succeed in reducing photovoltaic costs to competitive ones. Nevertheless, for these so-called leapfrogs to take place, a mature PV market should consolidate, for which silicon technology is essential at least for the next decade.

REFERENCES

1. Schmela M, *Photon International* **3**, 140 (2008).
2. Turton R, *Band structure of Si: Overview* in Hull R (ed.), *Properties of Crystalline Silicon*, INSPEC, Stevenage, UK (1999).
3. Green M, Keevers M, *Prog in Photovoltaics* **3**, 189–192 (1995).
4. Kolodinski S, Werner J, Wittchen T, Queisser H, *Appl. Phys. Lett.* **63**, 2405–2407 (1993).
5. Clugston D, Basore P, *Prog in Photovoltaics* **5**, 229–236 (1997).
6. Sproul A, Green M, *J. Appl. Phys.* **70**, 846–854 (1991).
7. Altermatt P *et al*., *Proc. 16th EC PVSEC* 102–105 (2000).
8. Dziewior J, Schmid W, *Appl. Phys. Lett.* **31**, 346–351 (1977).
9. Altermatt P, Schmidt J, Kerr M, Heiser G, Aberle A, *Proc. 16th EC PVSEC,* pp 243–246 (2000).
10. Green M *et al*, *Nature* **412**, 805–808 (2001).
11. Thurber W, Mattis R, Liu Y, Filliben J, *J. Electrochem. Soc.* **127**, 1807–1812 (1980), and *J. Electrochem. Soc.* **127**, 2291–2294 (1980).
12. Kane D, Swanson R, *Proc. 20th IEEE PVSC*, pp 512–517 (1988).
13. Sze S, *Physics of Semiconductor Devices*, Chap.5, John Wiley & Sons, Inc., NY (1981).
14. King R, Sinton R, Swanson R, *IEEE Trans. Electron Devices* **37**, 1399–1409 (1990).
15. King R, Swanson R, *IEEE Trans. Electron Devices* **38**, 365–371 (1991).
16. Narashima S, Rohatgi A, Weeber A, *IEEE Trans. Electron Devices* **46**, 1363–1370 (1999).
17. Honsberg C, Slade A, McIntosh K, Vogl B, Cotter J, Wenham S, *Proc. 16th EC PVSEC*, pp 1655–1658 (2000).
18. Taguchi M *et al*, *Prog in Photovoltaics* **8** 503–514 (2000).
19. Wang TH *et al*, *Proc. 14th Workshop on Crystalline Silicon Solar Cells and Modules*. NREL/CP-520-36669 (2004).
20. Grauvogl M and Hezel R, *Prog in Photovoltaics* **6**, 15–24 (1998).
21. Gan J, Swanson R, *Proc. 21st IEEE PVSC*, pp 245–250 (1990).
22. Mäckel H and Cuevas A, *Proc. 3rd World Conf on Photovoltaic Energy Conversion*, pp 71–74 (2003).
23. Aberle A, *Prog in Photovoltaics* **8**, 473–488 (2000).
24. Aberle A, Hezel R, *Prog in Photovoltaics* **5**, 29–50 (1997).
25. Choulat P, Agostinelli G, Ma Y, Duerinckx F, Beaucarne G, *Proc. 22nd European PVSEC*, pp 1011–1014 (2007).
26. Yamamoto H *et al*., *Proc. 22nd EC PVSEC*, pp 1224–1226 (2007).
27. Martín I, Vetter M, Orpella J, Cuevas A, Alcubilla R, *Applied Physics Letters* **79**(14), 2199–2201 (2001).
28. Agostinelli G, Delabie A, Vitanov P, Alexieva Z, Dekkers HZW, De Wolf S, Beaucarne G, *Solar Energy Materials & Solar Cells* **90**, 3438–3443 (2006).
29. Schmidt J. *et al*., *Prog. in Photovoltaics* **16**, 461–466 (2008).

30. Cuevas A, Basore P, Giroult-Matlakowski G, Dubois C, *Proc. 13th EC PVSEC*, pp 337–342 (1995); Cuevas A, Stuckings M, Lay J, Petravic M, *Proc. 14th EC PVSEC*, pp 2416–2419 (1997).
31. Eades W, Swanson R, *J. Appl. Phys.* **58**, 4267–4276 (1985).
32. Kerr, MJ, Cuevas A, *Proc. IEEE PVSC*, pp 103–107 (2002).
33. Aberle A, Glunz S, Warta W, *Sol. Energy Mat. Sol. Cells* **29**, 175–182 (1993).
34. Luque A, The requirements of high efficiency solar cells, in Luque A and Araújo G (eds), *Physical limitations to photovoltaic energy conversion*, pp 1–42, Adam Hilger, Bristol (1990).
35. Green M, *Silicon solar cells. Advanced principles and practice*, Chap. 7, Centre for Photovoltaic Devices and Systems, University of New South Wales, Sydney (1995).
36. Tiedje T, Yablonovitch E, Cody G, Brooks B *IEEE Trans. Electron Devices* **31**, 711–716 (1984).
37. Kerr MJ, Campbell P, Cuevas A, *Proc. IEEE PVSC*, pp 439–441 (2002).
38. Luque A, Tobías I, Gidon P, Pirot M, del Cañizo C, Antón I, Jausseaud C, *Progress in Photovoltaics* **12**, 503–516 (2004); Luque A, Gidon P, Pirot M, Antón I, Caballero LJ, Tobías I, del Cañizo C, Jausseaud C, *Progress in Photovoltaics* **12**, 517–528 (2004).
39. Verlinden P, Sinton R, Wickham K, Crane R, Swanson R, *Proc. 14th EC PVSEC*, pp 96–100 (1997).
40. Van Kerschaver E, Beaucarne G, *Prog in Photovoltaics* **14**, 107–123 (2006).
41. Luque A, Ruiz J, Cuevas A, Agost M, *Proc. 1st EC PVSEC*, pp 269–277 (1977).
42. Zhao J, Wang A, Green M, *Prog in Photovoltaics* **7** 471–474 (1999).
43. Saitoh T, Hashigami H, Rein S, Glunz S, *Prog in Photovoltaics* **8** 535–547 (2000).
44. Myers S, Seibt M, Schröter W, *J. Appl. Phys.* **88** 3795–3819 (2000).
45. McHugo S, Hieslmair H, Weber E, *Appl. Phys. A* **64**, 127–137 (1997).
46. Ohe N, Tsutsui K, Warabisako T, Saitoh T, *Sol. Energy Mat. Sol. Cells* **48**, 145–150 (1997).
47. Martinuzzi S *et al*., *Mat. Science and Eng. B* **71**, 229–232 (2000).
48. Wenham S, Green M, *Prog. in Photovol* **4**, 3–33 (1996).
49. Schröter W, Kühnpafel R, *Appl. Phys. Lett.* **56**, 2207–2209 (1990).
50. Joshi S, Gösele U, Tan T, *J. Appl. Phys.* **77**, 3858–3863 (1995).
51. Périchaud I, Floret F, Martinuzzi S, *Proc. 23rd IEEE PVSC*, pp 243–247 (1993).
52. Narasimha S, Rohatgi A, *IEEE Trans. on Electron Devices* **45**, 1776–1782 (1998).
53. Macdonald D, Cuevas A and Ferraza F, *Solid State Electronics* **43**, 575–581 (1999).
54. Gee J and Sopori B, *Proc. 26th IEEE PVSC*, pp 155–158 (1997).
55. del Cañizo C, Tobías I, Lago R, Luque A, *J. of the Electrochem. Soc.* **149**, 522–525 (2002).
56. Sopori B *et al*, *Sol. Energy Mat. Sol. Cells* **41/42** 159–169 (1996).
57. Buonassisi T, Istratov A, Marcus M, Lai B, Cai Z, Heald S, Weber E, *Nat. Mater* **4**, 676–679 (2005).
58. Waver P, Schmidt A, Wagemann H, *Proc. 14th EC PVSEC*, pp 2450–2453 (1997).
59. Green M, *Prog in Photovoltaics* **8**, 443–450 (2000).
60. Schmidt J, Bothe K, Hezel R, *Proc 29th IEEE PVSC*, pp 178–181 (2002).
61. Geerligs LJ, Macdonald D, *Prog in Photovoltaics* **12**, 309–316 (2004).
62. Cotter JE *et al*, *IEEE Trans on Electron Devices*, **53** 1893–1901 (2006).
63. Schmidt J *et al*., *Proc. 22nd EC PVSEC*, pp 998–1001 (2007).
64. Zhao J, Wang A, Green M, *Prog in Photovoltaics* **2**, 227–230 (1994).
65. Wenham S, *Prog in Photovoltaics* **1**, 3–10 (1993).
66. Glunz SW *et al*., *Proc. 21st' European PVSEC*, pp 746–749 (2006).
67. Cuevas A, Russell D, *Prog in Photovoltaics* **8**, 603–616 (2000).
68. Green M, *Silicon solar cells. Advanced principles and practice*, Chap. 10, Centre for Photovoltaic Devices and Systems, University of New South Wales, Sydney (1995).
69. Cuevas A, Sinton R, Swanson R, *Proc. 21st IEEE PVSC*, pp 327–332 (1990).

70. Luque A, *Solar cells and optics for photovoltaic concentration*, Chap.6, Adam Hilger, Bristol (1989).
71. Moehlecke A, Zanesco I, Luque A, *Proc. 1st World CPEC*, pp 1663–1666 (1994).
72. Glunz S, Knobloch J, Biro D, Wettling W, *Proc. 14th EC PVSEC*, pp 392–395 (1997).
73. Wenham S *et al*, *Proc. 1st World CPEC*, pp 1278–1282 (1994).
74. Schneiderlöchner E, Preu R, Lüdemann R, Glunz SW, *Progress in Photovoltaics* **10**, 29–34 (2002).
75. Mulligan W *et al*, *Proc. 28th IEEE PVSC*, pp 158–163 (2000).
76. Zhao J, Wang A, Altermatt P, Zhang G, *Prog in Photovoltaics* **8** 201–210 (2000).
77. Luque A, *Coupling light to solar cells* in Prince M Ed, *Advances in Solar Energy Vol 8*, American Solar Energy Society, Boulder, pp 151–230 (1993).
78. Kuo ML *et al*, *Optics Letters*, **33**, 2527–2529 (2008).
79. Born M, Wolf E, *Principles of Optics*, 7th edn, Chap. 1, Cambridge Universtity Press, Cambridge, UK (1999).
80. Zhao J and Green M, *IEEE Trans. on Electron Devices* **38**, 1925–1934 (1991).
81. Rodríguez J, Tobías I, Luque A, *Sol. Energy Mat. Sol. Cells* **45**, 241–253 (1997).
82. Kray D, Hermle M, Glunz SW, *Prog in Photovoltaics* **16**, 1–15 (2008).
83. Green M, *Silicon solar cells. Advanced principles and practice*, Chap. 6, Centre for Photovoltaic Devices and Systems, University of New South Wales, Sydney (1995).
84. Moehlecke A, *Conceptos avanzados de tecnología para células solares con emisores p^+ dopados con boro*, Chap. 5, *Ph D Thesis*, Universidad Politécnica de Madrid (1996).
85. Miñano J, *Optical confinement in Photovoltaics* in Luque A and Araújo G Eds, *Physical limitations to photovoltaic energy conversion*, Adam Hilger, Bristol. pp 50–83 (1990).
86. Neuhaus D-H, Münzer A, *Advances in Optoelectronics* vol 2007, article ID 24521, doi 10.1155/2007/24521.
87. Van Overstraeten R, Mertens R, *Physics, Technology and Use of Photovoltaics*, Chap. 4, Adam Hilger Ltd., Bristol (1986).
88. Funke C, Wolf S, Stoyan D, *J. Sol. Energy Eng.*, **131**(1), 11012–11017 (2009).
89. Barredo J, Fraile A, Jimeno JC, Alarcón E., *Proc. 20th European PVSEC*, pp 298–301 (2005).
90. Hylton J, Kinderman R, Burgers A, Sinke W, Bressers P, *Prog in Photovoltaics* **4**, 435–438 (1996).
91. Shirasawa K *et al*., *Proc 21st IEEE PVSC*, pp 668–73 (1990).
92. Zhao J, Wang A, Campbell P, Green M, *IEEE Trans. Electron Devices* **46**, 1978–1983 (1999).
93. De Wolf S *et al*., *Proc. 16th EC PVSEC*, pp 1521–1523 (2000).
94. Hauser A, Melnyk I, Fath P, Narayanan S, Roberts S, Bruton TM, *Proc 3rd WCPEC*, pp 1447–1450 (2003).
95. Bilyalov R, Stalmans L, Schirone L, Lévy-Clement C, *IEEE Trans. Electron Devices* **46**, 2035–2040 (1999).
96. Ruby D *et al*., *Proc. of the 2nd WCPEC*, pp 39–42 (1998).
97. Lüdemann R, Damiani B, Rohatgi A, Willeke G, *Proc 17th EC PVSEC*, pp 1327–1330 (2001).
98. Dekkers HFW, Agostinelli G, Deherthoghe D, Beaucarne G, Traxlmayer U, Walter, G, *Proc. 19th European PVSEC*, pp 412–415 (2004).
99. Mrwa A, Ebest G, Erler K and Rindelhardt U, *Proc 21st European PVSEC*, pp 815–817 (2006).
100. Inomata Y, Fukui K, Shirasawa K, *Sol. Energy Mat. and Solar Cells* **48**, 237–242 (1997).
101. Ruby DS, Zaidi SH, Narayanan S, *Proc IEEE 28th PVSC*, pp 75–78 (2000).
102. Winderbaum S, Reinhold O, Yun F, *Sol. Energy Mat. and Solar Cells* **46**, 239–248 (1997).
103. Narayanan S, Wohlgemuth J, Creager J, Roncin S, Perry M, Proc 12th European PVSEC, pp 740–742 (1994).

104. Szlufcik J, Fath P, Nijs J, Mertens R, Willeke G, Bucher E, Proc 12th European PVSEC, pp 769–772 (1994).
105. Fath P *et al*., *Proc 13th European PVSEC*, pp 29–32 (1995).
106. Spiegel M *et al*., *Sol. Energy Mat. Solar Cells*, **74**, 175–182 (2002).
107. Huster F, Gerhards C, Spiegel M, Fath P, Bucher E, *Proc 28th IEEE PVSC*, pp 1004–1007 (2000).
108. Joos W *et al*., *Proc 16th EC PVSEC*, pp 1169–1172 (2000).
109. Pirozzi L, Garozzo M, Salza E, Ginocchietti G, Margadonna D, *Proc. 12th EC PVSEC*, pp 1025–1028 (1994).
110. Goaer G, Le Quang N, Bourcheix C, Pellegrin Y, Loretz JC, Martinuzzi S, Perichaud I, Warchol F, *Proc 19th European PVSEC*, pp 1025–1028 (2004).
111. Salami J, Pham T, Khadilkar Ch, McViker K, Shaikh A, *Technical Digest Int. PVSEC-14*, pp 263–264 (2004).
112. Tool CJJ, Coletti G, Granek FJ, Hoornstra J, Koppes M, Kossen EJ, Rieffe HC, Romijn IG, Weeber AW, *Proc 20th European PVSEC*, pp 578–583 (2005).
113. Bentzen A, Schubert G, Christensen JS, Svensson BG, Holt A, *Progress in Photovoltaics: Research and Appl*, **15**, 281–289 (2007).
114. Voyer C, Biro D, Wagner K, Benick J, Preu R, Koriath J, Heintze M, Wanka HN, *Proc 20th European PVSEC*, pp 1415–1418 (2005).
115. Kim DS, Hilali MM, Rohatgi A, Nakano K, Hariharan A, Matthei K, *Journal of Electrochemical Society* **153**, A1391–A1396 (2006).
116. Benick J, Rentsch J, Schetter C, Voyer C, Biro D, Preu R, *Proc 21st European PVSEC*, pp 1012–1015 (2006)
117. http://www.schmid-group.com/en/business-fields/photovoltaic/cell/phosphorus-doper.html
118. Horzel J *et al*., *Proc. 17th EC PVSEC*, pp 1367–1370 (2001).
119. Emanuel G, Schneiderlöchner, Stollhof J, Gentischer J, Preu R and Lüdemann R, *Proc 17th EC PVSEC*, pp 1578–1581 (2001).
120. Melnyk I, Wefringhaus E, Delahaye F, Vilsmaier G, Mahler W, Fath P, *Proc 19th EC PVSEC*, pp 416–418 (2004).
121. Richards B, Cotter J, Honsberg C and Wenham S, *Proc. 28th IEEE PVSC*, pp 375–378 (2000).
122. Johnson J, Hanoka J, Gregory J, *Proc. 18th PVSC*, pp 1112–1115 (1985).
123. Sopori B, Deng X, Narayanan S, Roncin S, *Proc 11th EC PVSEC*, pp 246–249 (1992).
124. Szulfcik J *et al*., *Proc. 12th EC PVSC*, pp 1018–1021 (1994).
125. Leguijt C *et al*., *Sol. Energy Mat. Sol. Cells* **40**, 297–345 (1996).
126. Soppe W *et al*., *Proc. 29th IEEE PVSC*, pp 158–161 (2002).
127. Ruby D, Wilbanks W, Fieddermann C, *Proc. IEEE 1st WPEC*, pp 1335–1338 (1994).
128. Ruske M, Liu J, Wieder S, Preu R, Wolke W, *Proc. 20th European PVSEC*, pp 1470–1473 (2005).
129. Meemongkolkiat V, Nakayashiki K, Kim DS, Kopecek R and Rohatgi A, *Journal of Electrochemical Society* **153**(1), G53–G58 (2006).
130. Nijs J *et al*, *Proc. 1st World CPEC*, pp 1242–1249 (1994).
131. Ballif C, Huljić F, Hessler-Wyser A, Willeke G, *Proc. 29th IEEE PVSC*, pp 360–363 (2002).
132. Lenkeit B *et al*., *Proc. 16th EC PVSEC*, pp 1332–1335 (2000).
133. Hoornstra J, de Moor H, Weeber A and Wyers P, *Proc. 16th EC PVSEC*, pp 1416–1419 (2000).
134. Huster F, Fath P, Bucher E, *Proc. 17th EU-PVSEC*, pp 1743–1746 (2001).
135. Huljić DM, Thormann S, Preu R, Lüdemann R, Willeke G, *Proc. 29th IEEE PVSC*, pp 126–129 (2002).
136. Glunz SW, *Advances in Optoelectronics* vol 2007, article ID 97370, doi 10.1155/2007/97370.
137. Bruton T *et al*, *Proc. 14th EC PVSC*, pp 11–19 (1997).

138. Podewils C, Dreaming of dinosaurs, *Photon International* **3**, 114–123 (2007).
139. del Cañizo C, del Coso G, Sinke W, *Prog. in Photovoltaics* **17**(3), 199–209 (2009).
140. Tool C *et al*., *Prog in Photovoltaics* **10**, 279–291 (2002).
141. Upadhyaya A, Sheoran M, Ristow A, Rohatgi A, Narayanan S, Roncin S, *Proc. 4th WCPVEC*, pp 1052–1055 (2006).
142. Mueller A, Cherradi N, Nasch PM, *Proc. 3rd World CPEC*, pp 1475–1478 (2003).
143. Funke C, Sciurova O, Kaminski S, Fütterer W, Möller HJ, *Proc. 21st EC PVSEC*, pp 171–175 (2006).
144. Cereceda E, Gutiérrez JR, Jimeno JC, Barredo J, Fraile A, Alarcón E, Ostapenko S, Martínez A, Vázquez MA, *Proc. 22nd EC PVSEC*, pp 1168–1170 (2007).
145. Huster F, *Proc. 20th EC PVSEC*, pp 560–563 (2005).
146. Finck von Finckenstein B *et al*., *Proc. 28th IEEE PVSC*, pp 198–200 (2000).
147. Kim S, Shaikh A, Sridharan S, Khadilkar C, Pham T, *Proc. 19th European PVSEC*, pp 1289–1291 (2004).
148. Jimeno JC *et al*., *Proc. 22nd European PVSEC*, pp 875–878 (2007).
149. Munzer KA, Holdermann KT, Schlosser RE, Sterk S, *IEEE Trans. Electron Devices* **46**, pp 2055–2061 (1999).
150. Recart F, Freire I, Pérez L, Lago-Aurrekoetxea R, Jimeno JC, Bueno G, *Sol. Energy Mat. Sol. Cells* **91**, 897–902 (2007).
151. Dauwe S, Mittelstädt L, Metz A, Hezel R, *Prog. in Photovoltaics* **10**(4), 271–278 (2002).
152. Hoffman M, Glunz SW, Preu R, Willeke G, *Proc. 21st European PVSEC*, pp 609–613 (2006).
153. Rostan PJ, Rau U, Nguyen VX, Kirchartz T, Schubert MB, Werner JH, *Sol. Energy Mat. Sol. Cells*, **90**(9), 1345–1352 (2006).
154. Moschner J *et al*., *Proc. 2nd WCPEC*, pp 1426–1429 (1998).
155. Einhaus R *et al*., *Proc. 14th EC PVSEC* (1997).
156. Raabe B *et al*., *Proc. 22nd EC PVSEC*, pp 1024–1029 (2007).
157. Ruby DS, Yang P, Ro M, Narayanan S, *Proc. 26th IEEE PVSC*, pp 39–42 (1997).
158. Zerga A, Slaoui A, Muller JC, Bazer-Bachi B, Ballutaud D, Lê Quang N, Goaer G, *Proc. 21st European PVSEC*, pp 865–869 (2006).
159. Horzel J, Szlufcik J, Nijs J, *Proc. 16th EC PVSEC*, pp 1112–1115 (2000).
160. Hilali M, Jeong JW, Rohatgi A, Meier DL, Carroll AF, *Proc. of the 29th IEEE PVSC*, pp 356–359 (2002).
161. Sivoththaman S *et al*., *Proc. 14th EC PVSEC*, pp 400–403 (1997).
162. Doshi P, Mejia J, Tate K, Kamra S, Rohatgi A, Narayanan S, Singh R, *Proc. 25th IEEE PVSC*, pp 421–424 (1996).
163. Doshi P, Rohatgi A, Ro M, Chen M, Ruby D, Meier D, *Sol. Energy Mat. Solar Cells* **41/42**, 31–39 (1996).
164. Biro D *et al*., *Sol. Energy Mat. Sol. Cells* **74**, 35–41 (2002).
165. Schmidt W, Woesten B, Kalejs J, *Prog. in Photovoltaics* **10**, 129–140 (2002).
166. Horzel J, Grupp G, Preu R, Schmidt W, *Proc. 20th EC PVSEC*, pp 895–898 (2005).
167. Hahn G, Gabor AM, *Proc 3rd WCPEC*, pp 1289–1892 (2003).
168. J. Junge, *et al*., *Proc. 33rd IEEE PVSC*, pp 1–5 (2008).
169. Rohatgi A *et al*., *Proc. 14th International PVSEC*, pp 635–638 (2004).
170. Meier D, Davis H, Garcia R, Salami J, Rohatgi A, Ebong A, Doshi P, *Sol. Energy Mat. Sol. Cells* **65**, pp 621–627 (2000).
171. Seren S, Kaes M, Gutjahr A, Burgers AR, and Schönecker A, *Proc. 22nd European PVSEC*, Milan, pp 855–859 (2007).
172. Tanaka H, Okamoto S, Tsuge S, Kiyama S, *Proc 3rd WCPEC*, pp 955–958 (2003).
173. Mulligan W *et al*., *Proc. 19th PVSEC*, pp 387–390 (2004).
174. De Ceuster D *et al*., *Proc. 22nd EC PVSEC*, pp 816–819 (2007).
175. Van Kerschaver E, DeWolf S, Szlufcik J, *Proc 16* th *EPVSC*, Glasgow, pp 1517–1520 (2000).

176. A. Kress, R. Kühn, P. Fath, G. P. Willeke, and E. Bucher, *IEEE Trans. Electron. Devices* **46**, 2000–2004 (1999).
177. Hacke P, Gee JM, Hilali M, Dominguez J, Dundas H, Jain A, Lopez G, *Proc 21st European PVSEC*, pp 761–764 (2006).
178. Knauss H, Haverkamp H, Jooss W, Steckemetz S, Nussbaumer H, *Proc. 21st EC PVSEC*, pp 1192–1195 (2006).
179. Romijn I, Lamers M, Stassen A, Mewe A, Koppes M, Kossen E, Weeber A, *Proc. 22nd European PVSEC*, pp 1043–1049 (2007).
180. Blakers AW, Deenapanray PNK, Everett V, Franklin E, Jellett W, Weber KJ, *Proc 20th EC PVSEC*, pp 682–685 (2005).
181. Wenham S, Green M and Watt M, *Applied Photovoltaics*, Chap 5, Centre for Photovoltaic Devices and Systems, University of New Soth Wales, Sydney (1995).
182. B. Lalaguna, P. Sánchez-Friera, I.J. Bennett D. Sánchez L. J. Caballero, J. Alonso *Proc. 2nd EC PVSEC*, pp 2712–2715 (2007).
183. Van Overstraeten R and Mertens R, *Physics, technology and use of photovoltaics*, Chap. 8, Adam Hilger, Bristol (1986).
184. Galica J and Sherman N, *Proc. 28th IEEE PVSC*, pp 30–35 (2000).
185. Czanderna A and Pern J, *Sol. Energy Mat. Sol. Cells* **43**, 101–181 (1996).
186. Ishikawa, N *et al*., *Proc. 2nd World CPEC*, pp 2501–2506 (1998).
187. Schmidhuber H and Krannich K, *Proc. 17th EC PVSEC*, pp 663–663 (2001).
188. Parreta A, Sarno A, Schloppo R and Zingarelli M, *Proc. of the 14th EC PVSC*, pp 242–246 (1997).
189. Herrmann W, Becker H and Wiesner W, *Proc 14th EC PVSEC*, pp 224–228 (1997).
190. Appelbaum J and Segalov T, *Prog. in Photovoltaics* **7**, 113–128 (1999).
191. Zilles R and Lorenzo E, *International Journal of Solar Energy* **13**, 121–133 (1993).
192. Herrmann W, Wiesner W and Vaaßen W *Proc. 26th IEEE PVSC*, pp 1129–1132 (1997).
193. Herrmann W, Adrian M and Wiesner W, *Proc. 2nd World CPEC*, pp 2357–2359 (1998).
194. Roche D, Outhred H and Kaye R, *Prog. in Photovoltaics* **3**, 115–127 (1995).
195. Hauselaer P, van den Bossche J, Frisson L and Poortmans J, *Proc 17th EC PVSEC*, pp 642–645 (2001).
196. Sánchez-Friera P, Montiel D, Gil JF, Montañez JA, Alonso J, *Proc 4th World CPEC*, pp 2156–2159 (2006).
197. Green M, *Surface texturing and patterning in solar cells*, in Prince M Ed, *Advances in Solar Energy Vol 8*, American Solar Energy Society, Boulder, pp 231–269 (1993).
198. Quintana M, King D, McMahon T and Osterwald C, Commonly observed degradation in field-aged photovoltaic modules, *Proc. 29th IEEE PVSC*, pp 1436–1439 (2002).
199. Wilshaw A, Bates J and Oldach R, *Proc. 16th EC PVSEC*, pp 795–797 (2000).
200. IEC Standard 61215, Crystalline Silicon Terrestrial PV Modules – Design Qualification and Type Approval (1993).
201. McDonald D, Cuevas A, Kerr M, Samundsett C, Ruby D, Winderbaum S, Leo A, *Sol. Energy* **76**(1–3), 277–283 (2004).

8

High-efficiency III–V Multijunction Solar Cells

D. J. Friedman, J. M. Olson and Sarah Kurtz

National Renewable Energy Laboratory, Golden, CO 80401, USA

8.1 INTRODUCTION

The large-scale use of photovoltaics is becoming a reality. The total worldwide solar cell production in 2009 was >10 gigawatts (GW), mostly in the form of flat-plate Si solar cells. This is more than ten times the worldwide production in 2000, representing remarkable progress. Silicon modules have reached efficiencies of 20%, and the cost has been reduced to under $5/watt. However, in the context of world energy consumption, 3 GW is a small number. The phenomenal growth of the photovoltaics (PV) industry in the last seven years has been limited by the availability of purified silicon. The capital investment to ramp up production of purified silicon is daunting, especially since the use of Si by the PV industry has surpassed that by the integrated-circuit industry. One solution to this problem is to use "concentrator technology." Concentrators are discussed in detail in Chapter 10, but the principle is simple: lenses or mirrors focus (concentrate) the sunlight from a larger area onto a smaller solar cell. Concentration ratios of 500 or more are often used in actual systems. At these concentration ratios, high efficiency is more important than the cost of the cell. High cell efficiency is also of value for space applications (Chapter 9), providing reductions in size and weight of the PV module.

For both concentrator and space applications, there is thus the need for the highest-efficiency cells possible. In the quest for high efficiencies, the fundamental limitation on the efficiency of a conventional single-junction cell is a significant obstacle. For such a cell, photons with energy greater than the cell's bandgap have their excess energy lost as heat, while photons with energy below the bandgap are not absorbed and all their energy is lost. The multijunction approach provides an effective way of overcoming this obstacle, by dividing up the spectrum into several spectral regions and converting each region with a cell (junction) whose bandgap is tuned to that region. The concept is illustrated in Figure 8.1, which shows the solar spectrum divided up into two regions for conversion by a two-junction cell (specifically, a GaInP/GaAs cell, the basis of the industry-standard GaInP/GaAs/Ge three-junction design, and which will be discussed extensively in this chapter).

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

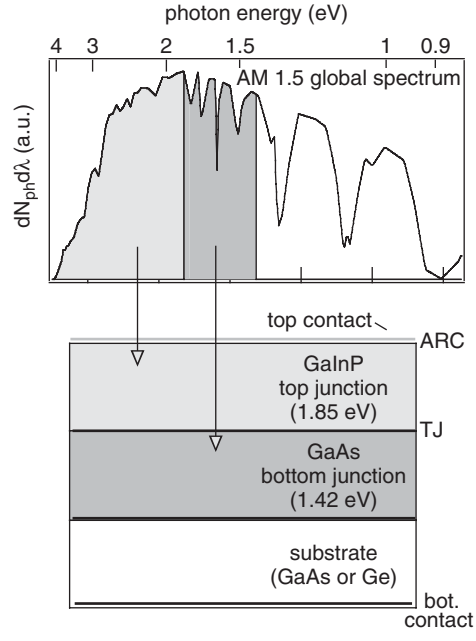


Figure 8.1 Schematic of GaInP/GaAs two-junction solar cell, showing the spectral regions converted by each junction. When the structure is grown on a Ge substrate, there is an option for introducing a third junction in the low bandgap (0.7 eV) Ge substrate, thus boosting the voltage and efficiency of the overall device. ARC = Antireflection Coating; TJ = tunnel junction. Dimensions are not to scale

While the multijunction concept is straightforward, the path to an actual high-efficiency commercially viable multijunction cell is quite complex. In 1984, researchers at the National Renewable Energy Laboratory (NREL) conceived and began work on the GaInP/GaAs two-junction solar cell, which would become the first commercially viable multijunction cell [1]. A schematic of the cell is shown in Figure 8.2. The cell consists of a $Ga_xIn_{1-x}P$ top cell (with a bandgap of 1.8 to 1.9 eV) grown monolithically on a lattice-matched interconnecting tunnel junction and a GaAs bottom cell. As shown in Figure 8.2, for $x \approx 0.5$, $Ga_xIn_{1-x}P$ has the same lattice constant as GaAs with a bandgap energy between 1.8 and 1.9 eV. Prior to this, several groups were working on tandem device designs that theoretically should achieve efficiencies approaching 36–40%. These included mechanical stacks of a high-bandgap top cell on a Si bottom cell and monolithic combinations of AlGaAs, GaAs, and GaInAs or GaAsP on Si. However, the mechanical stacks were viewed as too costly and cumbersome. Minimizing the defects generated by the lattice mismatch between top and bottom cells in some of the monolithic structures (i.e. GaAs on GaInAs or GaAsP on Si) was a challenging problem. The AlGaAs/GaAs tandem cell is lattice-matched with a theoretical efficiency of 36% [2]. However, the sensitivity of AlGaAs to trace levels of oxygen present in all growth systems and source materials made it difficult to produce this cell with high yield in a production environment and, thus, limited its use. The novel GaInP/GaAs cell design proposed by NREL trades manufacturability (i.e. lattice-matched top and bottom cells and oxygen-tolerant device materials) for a slightly lower theoretical efficiency of 34%.

By most standards, progress was rapid (see Figure 8.3). Despite initial problems with the growth of GaInP by metalorganic chemical vapor deposition (MOCVD) and complications associated with an anomalous red shift of the bandgap energy, by 1988 reasonably good GaInP top

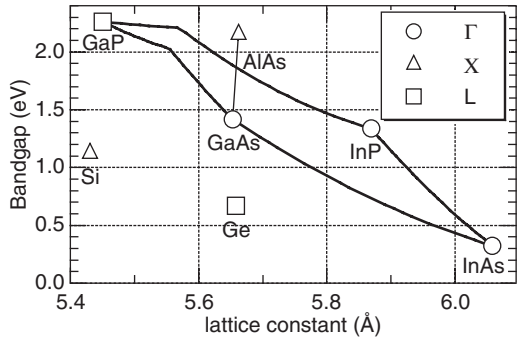


Figure 8.2 Estimated bandgap as a function of lattice constant for Si, Ge, and selected III–V binaries and their ternary alloys. The legend indicates the location of the bandgaps in the Brillouin zone; bandgaps located at the Γ point are direct

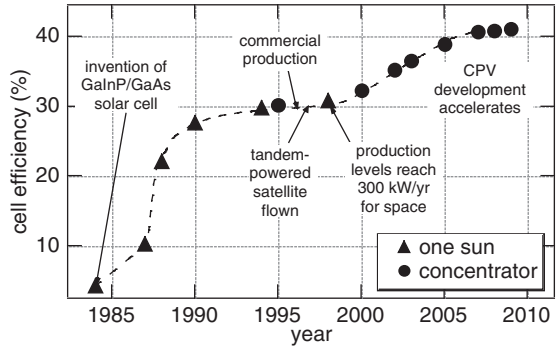


Figure 8.3 Champion multijunction cell efficiencies, showing the historic development of these cells. The most recent champions are summarized in Table 8.1

cells could be fabricated [3–5]. In 1990, efficiencies greater than 27% one-sun air-mass 1.5 global (AM1.5G) were achieved by changing the top cell thickness to achieve current matching [6, 7]. This tuning of the top cell thickness can also be used to achieve current matching under different solar spectra for different applications, as NREL demonstrated over the next three years, setting records at AM1.5G with an efficiency $\eta = 29.5\%$ [8], at 160-suns AM1.5D with $\eta = 30.2\%$ [9], and at one-sun AM0 with $\eta = 25.7\%$ [10] (see Chapter 17 for discussion of AM0 and AM1.5 spectra.). In 1994, it was discovered that the GaInP/GaAs tandem cells had very good radiation tolerance for operating in space. Kurtz and coworkers published results for a GaInP/GaAs cell with $\eta = 19.6\%$ (AM0) after irradiation with 1 MeV electrons at a fluence of $10^{15}\,\mathrm{cm}^{-2}$ [10], a standard radiation dose used to compare various solar cells. This efficiency was higher than the beginning-of-life efficiency for an unirradiated Si solar cell. These attributes soon attracted the attention of the commercial sector. Production of GaInP/GaAs solar cells (on Ge substrates) began around 1996, and the first GaInP/GaAs-powered satellite was launched in 1997.

Today, multijunction cells using GaInP are standard for both space and terrestrial concentrator applications, with more than a dozen companies capable of growing these cells. The cell

Table 8.1 Record solar-cell efficiencies. Unless otherwise specified, the cells were fabricated from single-crystal materials and the measurements were two-terminal [11, 12]

Cell	Efficiency (%)	Area (cm^2)	Intensity (suns)	Spectrum	Description
GaAs	25.9 ± 0.8	1.0	1	Global	Radboud U. Nijmegen
GaAs (thin film)	24.5 ± 0.5	1.0	1	Global	Radboud U. Nijmegen
GaAs(poly)	18.2 ± 0.5	4.0	1	Global	RTI Ge substrate
InP	21.9 ± 0.7	4.0	1	Global	Spire, epitaxial
$Ga_{0.5}In_{0.5}P$/GaAs	30.3	4.0	1	Global	Japan Energy
$Ga_{0.5}In_{0.5}P$/GaAs/Ge	32.0 ± 1.5	4.0	1	Global	Spectrolab
$Ga_{0.5}In_{0.5}P$/GaAs/ $Ga_{0.73}In_{0.27}As$	33.8 ± 1.5	0.25	1	Global	NREL, inverted metamorphic [13]
Si	24.7 ± 0.5	4.0	1	Global	UNSW, PERL
GaAs	27.8 ± 1.0	0.20	216*	Direct	Varian, ENTECH cover
GaInAsP	27.5 ± 1.4	0.08	171*	Direct	NREL, ENTECH cover
InP	24.3 ± 1.2	0.08	99*	Direct	NREL, ENTECH cover
$Ga_{0.5}In_{0.5}P$/ $Ga_{0.99}In_{0.01}As$/Ge	40.1 ± 2.4	0.25	135*	Low-AOD	Spectrolab, lattice-matched
$Ga_{0.44}In_{0.56}P$/ $Ga_{0.92}In_{0.08}As$/Ge	40.7 ± 2.4	0.27	240*	Low-AOD	Spectrolab, metamorphic
$Ga_{0.35}In_{0.65}P$/ $Ga_{0.83}In_{0.17}As$/Ge	41.1	0.051	454*	Low-AOD	Fraunhofer, metamorphic
$Ga_{0.5}In_{0.5}P$/ $Ga_{0.96}In_{0.04}As$/ $Ga_{0.63}In_{0.37}As$	40.8 ± 2.4	0.1	326*	Low-AOD	NREL, inverted metamorphic [14]
GaAs/GaSb	32.6 ± 1.7	0.053	100*	Direct	Boeing, 4-terminal mechanical stack
InP/GaInAs	31.8 ± 1.6	0.063	50*	Direct	NREL, 3-terminal, monolithic
$Ga_{0.5}In_{0.5}P$/GaAs	32.6	0.01	1000*	Low-AOD	UPM, monolithic
Si	27.6 ± 1.0	1.0	92*	Low-AOD	Amonix, back contact

*One sun is defined as 1000 W/m^2.

structures continue to evolve and other material combinations may become important, as described in Section 8.9. Current record solar-cell efficiencies are given in Table 8.1.

This chapter discusses the principles and operation of multijunction solar cells fabricated from III–V semiconductor compounds and alloys, with a particular emphasis on multijunction cells containing GaInP and Ga(In)As. III–V semiconductors have several characteristics that make them especially suitable for solar cells. A wide selection of these materials is available with direct bandgaps, and therefore, high absorption coefficients, in the ~0.7 to 2 eV range of interest for solar cells; GaAs, with a bandgap of 1.42 eV, and $Ga_{0.5}In_{0.5}P$, with a bandgap of ~1.85 eV, are especially

important examples. Both *n*- and *p*-type doping of these materials are generally straightforward, and complex structures made from these materials can be grown with extremely high crystalline and optoelectronic quality by high-volume growth techniques. As a result, III–V cells have achieved the highest single-junction efficiencies. Although these *single-junction* efficiencies are only slightly higher than the impressive efficiencies achieved by the best silicon cells, the ease of fabricating complex III–V structures (including layers with different bandgaps) makes possible the creation of III–V *multijunction* cells with efficiencies in excess of 40%, significantly exceeding the efficiencies of all single-junction devices. As shown in Figure 8.3, the champion efficiencies of these cells have been increasing at a rate of almost 1%/yr. *Production* versions of the cells have stayed within a few percent of the champion efficiencies. With further development, probably including the addition of a fourth junction, it is likely that the efficiencies will continue to increase, ultimately surpassing 45 or even 50%.

8.2 APPLICATIONS

8.2.1 Space Solar Cells

The higher efficiencies and radiation resistance of III–V cells have made them attractive as replacements for silicon cells on many satellites and space vehicles. Over the years, III–V multijunction cells have largely replaced silicon cells on new satellite launches. The GaInP/GaAs/Ge cells are integrated into modules very much like single-junction solar cells, and have the added advantage of operating at high voltage and low current, as well as having excellent radiation resistance. They also have a smaller temperature coefficient than silicon cells, which implies better performance under the operating conditions encountered in space applications.

Space applications of GaInP/GaAs/Ge and other III-V solar cells are discussed in detail in Chapter 9.

8.2.2 Terrestrial Electricity Generation

The PV industry currently services a wide range of terrestrial applications, from power for small consumer products to larger grid-connected systems. III–V solar cells are currently too expensive for most one-sun applications. While satellites represent an example of an application for which the extra cost is acceptable, for bulk electricity generation, a concentration of 400 suns or greater may be needed to achieve an acceptable cost. Multiple companies have now implemented multijunction III–V cells into terrestrial concentrator systems. Concentrator cells and systems are discussed in detail in Chapter 10.

Use of GaInP/Ga(In)As/Ge cells in high-concentration systems (500× and above) has the potential of generating electricity at 7 cents/kWh [15]. The price of multijunction solar cells has been reported to be as low as $24 M for 105 MW, or $0.23/W (http://cleantech.com/news/1713/emcore-gets-24m-purchase-order). The current space-cell production capacity of ∼1 MW/yr translates into a 1000× concentrator cell production capacity of ∼1 GW/yr. The high efficiency, the projected low cost, and the ease with which production of concentrator cells could be initiated with existing production equipment make terrestrial applications attractive for these cells. As the champion efficiencies of these cells have increased, the interest in using these in concentrator systems for terrestrial electricity generation has grown impressively. The specific issues associated with development of concentrator solar cells are discussed below.

8.3 PHYSICS OF III–V MULTIJUNCTION AND SINGLE-JUNCTION SOLAR CELLS

8.3.1 Wavelength Dependence of Photon Conversion Efficiency

As a prelude to the detailed examination of the design and performance of multijunction cells, it is useful to review briefly the fundamental factors that limit the efficiency of single-junction cells. Consider an ideal single-junction cell with characteristic bandgap E_g. A photon incident on this cell with photon energy $h\nu > E_g$ will be absorbed and converted to electrical energy, but the excess energy $h\nu - E_g$ will be lost as heat. The greater the excess, the lower the fraction of that photon's energy will be converted to electrical energy. On the other hand, a photon of energy $h\nu < E_g$ will not be absorbed and converted to electrical energy. Thus, the efficiency of photon conversion is a maximum efficiency at $h\nu = E_g$. Note that this maximum efficiency is less than 100%; the maximum work per absorbed photon is calculated by Henry [16].

Since the solar spectrum is broad, containing photons with energies ranging from near 0 to 4 eV, single-junction solar cell efficiencies under the unconcentrated spectrum of a black body at the temperature of the sun are limited to 31%, called the Shockley–Queisser limit (see Chapter 4). The solution to this problem is (in principle) simple: rather than trying to convert all the photon energies with one cell with one bandgap, divide up the spectrum into several spectral regions and convert each with a cell whose bandgap is tuned for that region. For instance, suppose the spectrum is divided up into three regions $\infty - h\nu_1$, $h\nu_1 - h\nu_2$, and $h\nu_2 - h\nu_3$, where $h\nu_1 > h\nu_2 > h\nu_3$. The light from these spectral regions would be converted by cells with bandgaps $E_{g1} = h\nu_1$, $E_{g2} = h\nu_2$, and $E_{g3} = h\nu_3$, respectively. The more spectral regions allowed, the higher the potential overall efficiency.

8.3.2 Theoretical Limits to Multijunction Efficiencies

Henry has calculated the limiting terrestrial one-sun efficiencies for conversion with 1, 2, 3, and 36 bandgaps; the respective efficiencies are 37, 50, 56, and 72% [16]. The improvement in efficiency from one to two bandgaps is considerable, but the returns diminish as more bandgaps are added. This is fortunate since the practicality of a device with five or more junctions is questionable. Note that the promise of the multijunction efficiency improvements will not be realized unless the bandgaps of the multiple junctions are correctly chosen; this choice will be discussed below in detail. Theoretical efficiency limits for multijunction devices based on thermodynamic fundamentals are presented in Chapter 4.

8.3.3 Spectrum Splitting

The multijunction approach requires that an incident photon with a given energy be directed onto the correct subcell. Perhaps the conceptually simplest approach would be to use an optically dispersive element such as a prism to spatially distribute photons with different energies to different locations, where the appropriate cells would be placed to collect these photons. This approach is illustrated in Figure 8.4a. Although conceptually simple, in practice the mechanical and optical complexities of this scheme make it undesirable for most circumstances. A preferable approach is to arrange the cells in a stacked configuration, as illustrated in Figure 8.4b, arranged so that the sunlight strikes the highest bandgap first, and then strikes the progressively lower-bandgap junctions. This arrangement

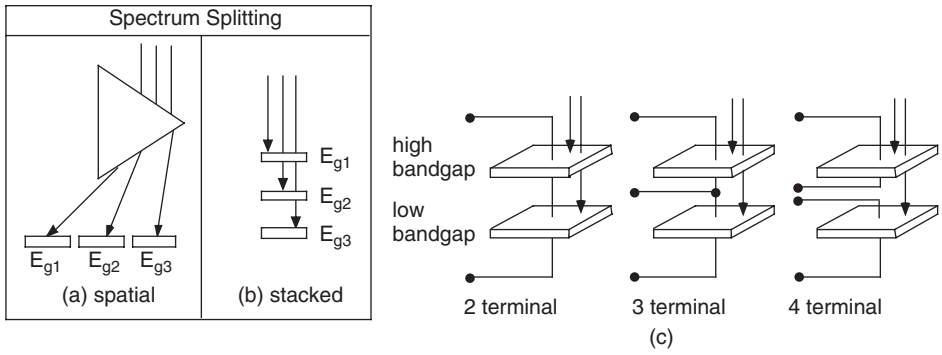


Figure 8.4 Schematic comparison of (a) spatial- and (b) stacked-configuration approaches to distributing light to subcells of different bandgaps. (c) Illustration of two-, three-, and four-terminal connection to a two-junction cell. The figure shows the subcells as mechanically separate, but the two- and three-terminal devices can be monolithic

makes use of the fact that junctions act as low-pass photon energy filters, transmitting only the sub-bandgap light. Thus, in Figure 8.4b, photons with $h\nu > E_{g1}$ are absorbed by that subcell, photons with $E_{g2} < h\nu < E_{g1}$ are absorbed by the E_{g2} subcell, and so on. In other words, the junctions themselves act as optical elements to distribute the spectrum to the appropriate junctions for multijunction photoconversion. The bandgaps *must* decrease from top to bottom of the stack. The stacked arrangement avoids the necessity for a separate optical element such as a prism to distribute the spectrum. Also, even if the junctions are physically separate from each other, they can be mechanically brought together into a relatively compact package, called a mechanical stack. The stacked configuration requires, of course, that all the junctions in the stack except the bottom one be transparent to light below their bandgaps, which, in practice, can set challenging constraints on the substrates and the back-contact metallizations of these junctions through which sub-bandgap light must pass. An elegant approach to this problem, which has several other advantages as well, is to fabricate all the junctions, each one atop the last, monolithically on a single substrate. This monolithic stack approach is the emphasis of this chapter.

8.4 CELL CONFIGURATION

8.4.1 Four-terminal

There are several ways to connect power leads to the junctions comprising a multijunction stack. These configurations, which provide for varying degrees of electrical isolation of the subcells, are illustrated in Figure 8.4c for a two-subcell stack. In the four-terminal configuration, each subcell has its own two terminals and is electrically isolated from the other subcells. This configuration has the advantage that it sets no constraints on the polarities (p/n vs n/p) of the subcells, or on their currents or voltages. However, the terminals and the electrical isolation between subcells in the four-terminal configuration would be inconvenient to accomplish monolithically, requiring a complicated cell structure and processing. Generally, a four-terminal device is implemented as a mechanical stack, whose complexities of fabrication and assembly make it a significantly less desirable structure than the monolithic device. In addition, unique power controls are needed to keep each cell at its maximum power point and then to sum the power “off-chip”.

8.4.2 Three-terminal

In contrast, in the three-terminal configuration the subcells are not electrically isolated; the bottom of each cell is electrically connected to the top of the cell beneath it. The fabrication of a monolithic three-terminal device is relatively straightforward, although more complex than the fabrication of a two-terminal device. The semiconductor structure must be designed to provide a layer for contact to the intermediate terminal, and to accommodate the processing steps necessary to put the intermediate terminal in place. With this intermediate terminal, the different subcells in the stack do not need to have the same photocurrents. Furthermore, in this three-terminal configuration, the different subcells in the stack may have different polarities, e.g. p/n for the top cell and n/p for the bottom cell.

Module-level interconnection of four- and three-terminal devices is discussed in detail by Gee [17].

8.4.3 Two-terminal Series-connected (Current-matched)

The two-terminal series-connected configuration provides the fewest possibilities for interconnection of the devices. This configuration requires that the subcells be of the same polarity and that the photocurrents of the subcells be closely matched, since in this series connection the subcell with the least photocurrent limits the current generated by the entire device. This current-matching constraint, about which more will be said shortly, puts relatively tight constraints on the selection of bandgaps for the various junctions in this structure. Against these disadvantages, however, are critical advantages. The existence of high-quality monolithic tunnel-junction subcell interconnects means that these stacks can be made as monolithic two-junction structures with metallization at the top and bottom of the stack only. This, in turn, means that such devices can be integrated into modules with the same simplicity afforded by single-junction devices. Figure 8.5 shows a schematic cross-section of a monolithic two-terminal series-connected three-junction solar cell, along with typical materials parameters for the realization of this device structure as a GaInP/GaAs/Ge cell. The two-terminal, series-connected configuration will be the focus of the following sections, where we analyze in detail the dependence of the cell performance on the cell design parameters.

8.5 COMPUTATION OF SERIES-CONNECTED DEVICE PERFORMANCE

8.5.1 Overview

This section discusses the quantitative modeling of the performance of series-connected, two-terminal, multijunction devices, as well as the quantitative design of these devices. Emphasis is placed on selecting bandgap pairs, and predicting the efficiency of the resulting structures. This modeling also lays the groundwork for the analysis of the dependence of the device performance on spectrum, concentration, and temperature. Following the treatments of references [7, 18], we make simplifying assumptions, which include: (1) transparent zero-resistance tunnel-junction interconnects; (2) no reflection losses; (3) no series-resistance losses; (4) junctions that collect every absorbed photon; (5) junctions whose current–voltage ($J-V$) curves are described by the ideal ($n = 1$) diode equation; and (6) no reabsorption of emitted photons (often called photon recycling). Later, we relax assumption (2) to analyze the effect of antireflection coatings. The first five assumptions neglect practical losses that are minimized as the design is perfected. The sixth

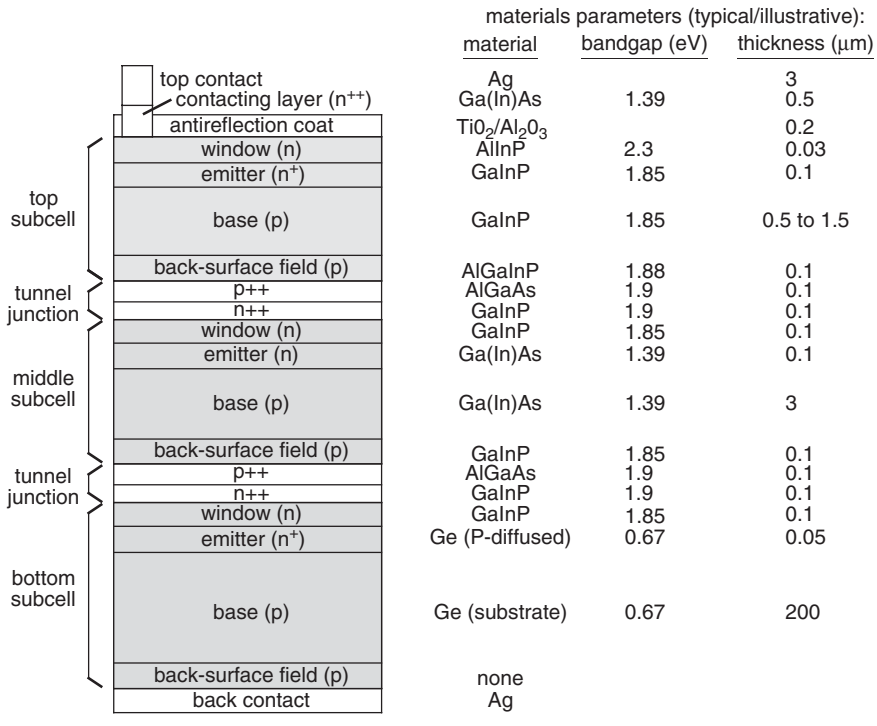


Figure 8.5 Schematic cross-section of a monolithic two-terminal series-connected three-junction solar cell. An n-on-p configuration is illustrated. Doping indicated by n^{++}, n^{+}, n (or p^{++}, p^{+}, p) corresponds to electron (or hole) concentrations of the order of $10^{19}-10^{20}$, 10^{18}, 10^{17} respectively. Typical materials, bandgaps, and layer thicknesses for the realization of this device structure as a GaInP/GaAs/Ge cell are indicated. A back-surface field is not needed for the Ge junction because the base thickness is significantly greater than the minority-carrier diffusion length. Note that not all layers in an actual device (e.g. tunnel-junction cladding layers) are included in the illustration. The figure is not to scale

assumption neglects an effect that can enhance performance. Historically, modeling based on these assumptions has proven very useful in developing high-efficiency multijunction solar cells. As the perfection of the cells improves, inclusion of photon recycling in the model (relaxing assumption (6)) may lead to new insights. It should be noted that high-quality III–V cells have achieved ~80% of these predicted efficiencies.

8.5.2 Top and Bottom Subcell QE and J_{SC}

The short-circuit current density (J_{sc}) of each subcell is determined by the quantum efficiency of the subcell, QE(λ), and by the spectrum of light incident on that cell $\Phi_{inc}(\lambda)$ in the usual way:

$$J_{SC} = e \int_0^\infty \mathrm{QE}(\lambda)\Phi_{inc}(\lambda)d\lambda. \tag{8.1}$$

The QE for an ideal cell of finite base thickness x_b, emitter thickness x_e, and depletion width W (for a total thickness $x = x_e + W + x_b$,) is given by the standard equations

$$QE = QE_{emitter} + QE_{depl} + \exp[-\alpha(x_e + W)]\, QE_{base}, \tag{8.2}$$

where

$$QE_{emitter} = f_\alpha(L_e)\left(\frac{\ell_e + \alpha L_e - \exp(-\alpha x_e)[\ell_e \cosh(x_e/L_e) + \sinh(x_e/L_e)]}{\ell_e \sinh(x_e/L_e) + \cosh(x_e/L_e)} - \alpha L_e \exp(-\alpha x_e)\right), \tag{8.3}$$

$$QE_{depl} = \exp(-\alpha x_e)\,[1 - \exp(-\alpha W)] \tag{8.4}$$

$$QE_{base} = f_\alpha(L_b)\left(\alpha L_b - \frac{\ell_b \cosh(x_b/L_b) + \sinh(x_b/L_b) + (\alpha L_b - \ell_b)\exp(-\alpha x_b)}{\ell_b \sinh(x_b/L_b) + \cosh(x_b/L_b)}\right), \tag{8.5}$$

$$\ell_b = S_b L_b/D_b,\ \ell_e = S_e L_e/D_e,\ D_b = kT\mu_b/e,\ D_e = kT\mu_e/e, \tag{8.6}$$

$$f_\alpha(L) = \frac{\alpha L}{(\alpha L)^2 - 1}. \tag{8.7}$$

The photon wavelength dependence is not explicit in these equations, but enters through the wavelength dependence of the absorption coefficient $\alpha(\lambda)$. The quantities $\mu_{b(e)}$, $L_{b(e)}$, and $S_{b(e)}$ are, respectively, the mobility, diffusion length, and surface recombination velocity for the minority carriers in the base (emitter); T is the absolute temperature. Later in this chapter, we will illustrate the use of these equations in the analysis of real-world III–V cells. However, in this section, we shall make the simplifying assumption that each absorbed photon is converted to photocurrent, a remarkably good first approximation for high-quality III–V junctions. In this case, the QE depends very simply on the total thickness of the device, $x = x_e + W + x_b$, as

$$QE(\lambda) = 1 - \exp[-\alpha(\lambda)x], \tag{8.8}$$

because a fraction $\exp[-\alpha(\lambda)x]$ of the incident light is transmitted through the cell instead of being absorbed. (Although this last equation is self-evident, it can also be deduced from Equations (8.2–8.5) by setting $S = 0$, $L \gg x$, and $L \gg 1/\alpha$.)

We now consider the light absorbed and converted by each of the junctions in a multijunction cell. Consider a cell with n junctions numbered from top to bottom as 1/2/. . ./n and with corresponding bandgaps $E_{g1}/E_{g2}/\ldots/E_{gn}$, as illustrated in Figure 8.4b. For sub-bandgap photons, $\alpha(\lambda) = 0$, and thus $\exp[-\alpha(\lambda)x] = 1$. The light Φ_{inc} incident on the top subcell is simply the solar spectrum, Φ_S. In contrast, the light hitting the mth subcell is filtered by the junctions above it, so that the mth subcell sees an incident spectrum $\Phi_m(\lambda) = \Phi_S(\lambda)\exp\left[-\sum_{i=1}^{m-1}\alpha_i(\lambda)x_i\right]$, where x_i and $\alpha_i(\lambda)$ are the thickness and absorption coefficient, respectively, of the ith subcell. The short-circuit current density of the mth subcell is then

$$J_{SC,m} = e\int_0^{\lambda_m}(1 - \exp[-\alpha_m(\lambda)x_m])\Phi_m(\lambda)\,d\lambda, \tag{8.9}$$

where $\lambda_m = hc/E_{gm}$ is the wavelength corresponding to the bandgap of the mth subcell. In the case that the mth subcell is optically thick, i.e. is thick enough to absorb essentially all above-bandgap light incident on it (i.e. $\alpha_m x_m \gg 1$), the exponential term goes to zero for all photon energies above the bandgap E_{gm}.

In order to understand how J_{SC} depends on the bandgaps of the various junctions, it is helpful to consider the simple case of a two-junction cell. Because the bottom junction is filtered by the top junction, the bottom-junction current density $J_{SC,2}$ depends on both E_{g1} and E_{g2}, whereas $J_{SC,1}$ depends only on E_{g1}. Equation (8.9) shows this dependence with special clarity in the case of optically thick junctions; in this case,

$$J_{SC,1} = e\int_0^{\lambda_1} \Phi_S(\lambda)\,\mathrm{d}\lambda, \quad J_{SC,2} = e\int_{\lambda_1}^{\lambda_2} \Phi_S(\lambda)\,\mathrm{d}\lambda. \tag{8.10}$$

8.5.3 Multijunction $J-V$ Curves

For any set of m series-connected subcells (or, indeed, any sort of two-terminal element or device) whose individual current–voltage ($J-V$) curves are described by $V_i(J)$ for the ith device, the $J-V$ curve for the series-connected set is simply

$$V(J) = \sum_{i=1}^{m} V_i(J); \tag{8.11}$$

i.e., the voltage at a given current is equal to the sum of the subcell voltages at that current. Each individual subcell will have its own maximum-power point $\{V_{mp,i}, J_{mp,i}\}$ which maximizes $J \times V_i(J)$. However, in the series-connected multijunction connection of these subcells, the currents through each of the subcells are constrained to have the same value, and therefore *each subcell will be able to operate at its maximum-power point only if $J_{mp,i}$ is the same for all the subcells*, i.e. $J_{mp,1} = J_{mp,2} = \ldots = J_{mp,m}$. If this is the case, then the maximum power output of the combined multijunction device is the sum of the maximum power outputs $V_{mp,i}J_{mp,i}$ of the subcells. On the other hand, if the subcells do not all have the same value for $J_{mp,i}$, then in their series-connected multijunction combination, some of the subcells must necessarily operate away from their maximum power points.

The consequences of this last point are especially important when, as is the case for high-quality III–V junctions, the subcells do not leak or quickly break down in reverse bias. The adding of series $J-V$ curves in this case is illustrated graphically in Figure 8.6, which shows $J-V$ curves for a GaInP top subcell, a GaAs bottom subcell, and the two-junction series-connected combination of these two subcells. In this example, the bottom subcell has a higher J_{SC} than the top subcell; the top subcell is slightly shunted, to make the illustration of its behavior at the tandem J_{SC} easier to see. For any given value of current, the tandem voltage satisfies $V_{tandem} = V_{top} + V_{bottom}$, as can be verified by inspection of the figure. The region of current near the tandem-cell $J_{SC} = -14\,\mathrm{mA/cm^2}$, shown in expanded scale in the bottom panel of the figure, is of special interest. At $J = -13.5\,\mathrm{mA/cm^2}$, both subcells are in forward bias, with voltages only slightly less than their respective open-circuit voltages (V_{OC}). As the magnitude of the current density is further increased to -14 mA/cm^2 and beyond, the bottom subcell remains in forward bias near its V_{OC}. At the same time, in contrast, the top subcell voltage becomes rapidly more negative, so that at $J = -14\,\mathrm{mA/cm^2}$, it has reached a negative bias of about -1 V, equal in magnitude, but opposite in sign to the top subcell's forward bias of $+1$ V. At this point, marked in the figure with $\times$ symbols, the tandem cell is at zero bias; i.e. the tandem is at its J_{SC}. This behavior illustrates the general principle that for subcells without significant leakage or reverse-bias breakdown, *the tandem J_{SC} is constrained to be, to a very good approximation, the lesser of the J_{SC} values of the subcells*. (Note that this current-limiting characteristic makes series-connected multijunction cells of the type considered here much *worse* than single-junction cells for conversion of narrow-band spectra such as the light from a laser!)

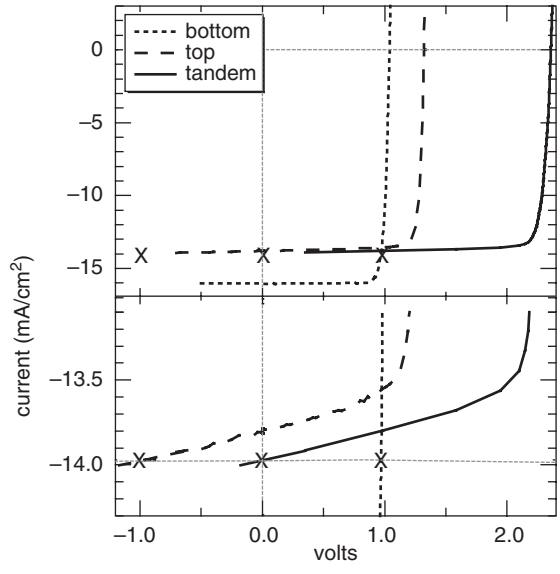


Figure 8.6 Illustration of the addition of $J-V$ curves for two series-connected subcells. The lower panel is an expanded view of the current range in the vicinity of the current-limiting top subcell J_{SC}, showing how the tandem J_{SC} is limited to the lesser of the subcell currents. The $J-V$ of the top subcell in this example is slightly leaky, which makes the addition of the subcell $J-V$ curves near J_{SC} easier to see

To model multijunction devices quantitatively, we need expressions for the subcell $J-V$ curves, $V_i(J)$. To proceed, we use the classical ideal-photodiode $J-V$ equations (neglecting the depletion region) [19],

$$J = J_0\,[\exp(\mathrm{eV}/k_\mathrm{B}T) - 1] - J_\mathrm{SC}, \tag{8.12}$$

where e is the electron charge. An important special case of this is

$$V_\mathrm{OC} \approx (kT/e)\ln(J_\mathrm{SC}/J_0), \tag{8.13}$$

because, in practice, $J_\mathrm{SC}/J_0 \gg 1$. The dark current density J_0 is given by

$$J_0 = J_{0,\mathrm{base}} + J_{0,\mathrm{emitter}}, \tag{8.14}$$

where

$$J_{0,base} = e\left(\frac{D_b}{L_b}\right)\left(\frac{n_i^2}{N_b}\right)\left(\frac{(S_bL_b/D_b) + \tanh(x_b/L_b)}{(S_bL_b/D_b)\tanh(x_b/L_b) + 1}\right), \tag{8.15}$$

and a similar equation describes $J_{0,\mathrm{emitter}}$. The intrinsic carrier concentration n_i is given by

$$n_i^2 = 4M_\mathrm{c}M_\mathrm{v}(2\pi kT/h^2)^3\,(m_e^*m_h^*)^{3/2}\exp(-E_\mathrm{g}/kT), \tag{8.16}$$

where m_e* and m_h* are the electron and hole effective masses, and M_c and M_v are the number of equivalent minima in the conduction and valence bands respectively. $N_{\mathrm{b(e)}}$ is the base (emitter) ionized-impurity density.

These equations fully define the multijunction $J-V$ curve, as described by Equation (8.11). The maximum power point $\{J_{mp}, V_{mp}\}$ can be calculated numerically as the point on the $V(J)$ curve that maximizes $J \times V(J)$. The various solar cell performance parameters of interest can be extracted from the $J-V$ curve in the usual way; e.g. $V_{OC} = V(0)$, $FF = J_{mp}V_{mp}/(V_{OC}J_{SC})$.

8.5.4 Current Matching and Top-cell Thinning

The relative magnitudes of the top- and bottom-subcell short-circuit current densities J_{SCt} and J_{SCb} depend on the bandgaps of the subcells, as Equation (8.10) shows explicitly for the case of optically thick subcells. For this case, Figure 8.7a shows J_{SCt} and J_{SCb} as a function of E_{gt} for $E_{gb} = 1.42\,\text{eV}$ for the AM1.5 global spectrum. The figure shows that as E_{gt} decreases, J_{SCt} increases and J_{SCb} decreases, becoming less than J_{SCt} for $E_{gt} < 1.95\,\text{eV}$. The J_{SC} for the series-connected combination of these two cells will be the lesser of J_{SCt} and J_{SCb}. This quantity is a maximum at the current-matched bandgap $E_{gt} = 1.95\,\text{eV}$, and falls off rapidly as E_{gt} is decreased below 1.95 eV.

Because the absorption coefficient $\alpha(h\nu)$ for solar-cell materials is not infinite, a cell of finite thickness (i.e. not optically thick) will not absorb all the incident above-bandgap light. Some light will be transmitted, especially at photon energies near the bandgap where α is small; the thinner the cell, the greater the transmission. Therefore, for a two-junction cell, thinning the top subcell will reapportion the light between the two subcells, increasing the bottom-subcell current at the expense of the top-subcell current. If, before thinning, $J_{SCb} < J_{SCt}$, then the top subcell can be thinned to make $J_{SCb} = J_{SCt}$. Because the series multijunction cell current J_{SC} is limited to the lesser of J_{SCb} and J_{SCt}, J_{SC} and hence the cell efficiency will be maximized when the top subcell is thinned to achieve this current matching. (This criterion does not rigorously maximize the efficiency, since a cell's maximum power point is not at short-circuit, but it is a very good approximation for high-quality, high-bandgap cells.) Figure 8.7b illustrates this top-cell thinning for the tandem device and spectrum of Figure 8.7a, for the case that $E_{gt} = 1.85\,\text{eV}$ and $E_{gb} = 1.42\,\text{eV}$. The tandem current $J_{SC} \approx \min(J_{SCt}, J_{SCb})$ is maximized at a top-subcell thickness of 0.7 μm, for which the subcells are current-matched. At this thickness, $J_{SC} = J_{SCt} = 15.8\,\text{mA/cm}^2$, or about 85% of what J_{SCt} would be for an infinite-thickness top subcell. The ability of such a thin cell to absorb such a high fraction of the incident light is due to the large absorption coefficient for this direct-gap material.

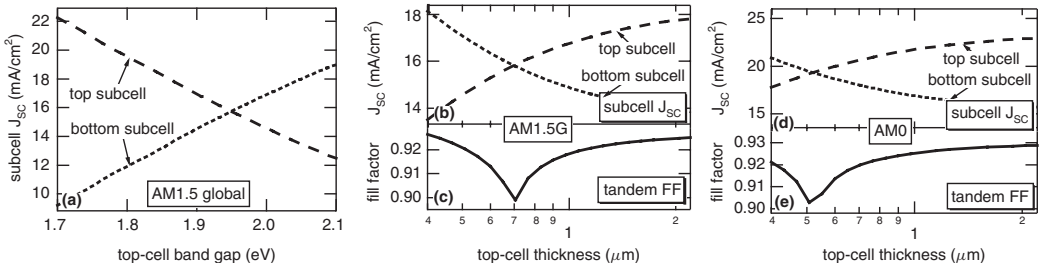


Figure 8.7 (a) J_{SCt} and J_{SCb} for an infinitely thick top subcell as a function of top-subcell bandgap E_{gt} for a bottom-subcell bandgap $E_{gb} = 1.42$ eV. (b) J_{SCt} and J_{SCb} as a function of top-subcell thickness for the AM1.5 global spectrum, with $E_{gb} = 1.42$ eV and $E_{gt} = 1.85$ eV. The subcells are current-matched at a base thickness of 0.7 μm. (c) The corresponding tandem-cell fill factor. (d) J_{SCt} and J_{SCb} as a function of top-subcell thickness as in (b), but for the AM0 spectrum. The current-matching thickness is significantly less than for the global spectrum. (e) Corresponding fill factor for AM0

8.5.5 Current-matching Effect on Fill Factor and V_{OC}

The fill factor (FF) of the tandem cell depends on the top- and bottom-subcell photocurrents. Figure 8.7c, e show the fill factor as a function of top-cell thickness, and thus effectively as a function of J_{SCt}/J_{SCb}, for the device of Figure 8.7b. The fill factor is a minimum at the current-matched condition, an effect that holds in general for reasonably ideal (non-leaky) subcells. This effect slightly undermines the efficiency gains that accrue from the increase in J_{SC} at the current-matched condition; however, the decrease in fill factor at current-matching is roughly half the increase in J_{SC}. This dependence of fill factor on the ratio of the subcell currents is important, because it implies that correctly measuring the fill factor of an actual device requires correctly light-biased subcells. This subject is discussed further in Chapter 17.

As Equations (8.13–8.15) show, V_{OC} also depends on cell thickness. Figure 8.8 shows how finite base thickness x_b and base surface-recombination velocity S_b affect the V_{OC} of a GaInP cell (neglecting photon recycling). These curves were calculated using Equations (8.13–8.15), assuming $D_b/L_b = 2.8 \times 10^4$ cm/s, a typical value for a GaInP cell. The figure shows that for a cell with a well-passivated base, i.e. S_b small enough that $S_b \ll D_b/L_b$, thinning the cell results in a meaningful increase in V_{OC}. On the other hand, for a cell whose base is so poorly passivated that $S_b > D_b/L_b$, thinning the cell lowers V_{OC}. For the GaInP/GaAs tandem structure, with the thin top subcell required for current-matching, the passivation of the base of the top subcell is thus an important consideration for the overall device efficiency. The passivation of GaInP surfaces will be discussed later in this chapter.

8.5.6 Efficiency vs Bandgap

8.5.6.1 Two-junction cells

To obtain concrete, numerical values for the cell performance from the concepts described above, we need to choose numbers for the relevant materials properties to determine J_0 for each junction. Reference [7] provides a reasonable model of a two-junction n/p cell, in which the bottom junction has the properties of GaAs, except that the bandgap is allowed to vary. The absorption coefficient

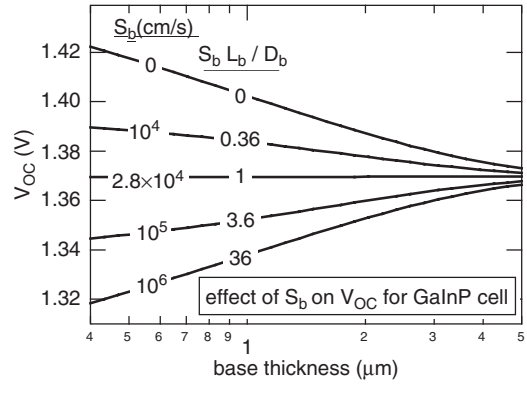


Figure 8.8 Effect of base thickness x_b and surface-recombination velocity S_b on V_{OC} for a GaInP top cell with $J_{SC} = 14$ mA/cm^2. The base is characterized by a *bulk* recombination velocity $D_b/L_b = 2.8 \times 10^4$ cm/s. Note that when the bulk and surface recombination velocities are equal, V_{OC} is independent of base thickness. Photon recycling is neglected

is shifted rigidly with bandgap so that it goes to zero as the photon energy decreases below the bandgap energy. Likewise, for the top subcell, the model uses the materials properties of GaInP, again allowing the bandgap to vary. The diffusion lengths at 300 K for the GaAs cell are $L_b = 17\,\mu m$ and $L_e = 0.8\,\mu m$; for the GaInP cell, $L_b = 3.7\,\mu m$ and $L_e = 0.6\,\mu m$. For simplicity, and to give results representing the maximum possible performance, all surface recombination is taken as zero. The emitters for both subcells have thickness $x_e = 0.1\,\mu m$ and ionized dopant concentration $N_e = 2 \times 10^{18}/cm^3$, and the bases for both subcells have $N_b = 10^{17}/cm^3$. These values are comparable to those used in actual GaInP/GaAs multijunction cells, providing an optimal combination of high quantum efficiency, low dark current, and low series resistance. Using this model, Figure 8.9a plots contours of cell efficiency for a two-junction series-connected cell with infinitely thick subcells, calculated for the one-sun standard AM1.5 global spectrum. Similar contours are shown for a variety of spectra and concentrations by Nell and Barnett [20] and by Wanlass *et al*. [18]. At the optimal bandgap combination of $\{E_{gt} = 1.75\,eV,\ E_{gb} = 1.13\,eV\}$ an efficiency of almost 38% is predicted, well in excess of the 29% efficiency that the model would predict for the best single-junction device. Even well away from the optimal bandgap combination, at a bandgap combination of $\{E_{gt} = 1.95\,eV,\ E_{gb} = 1.42\,eV\}$, the efficiency is still much higher than the best single-junction efficiency. But as E_{gt} decreases from 1.95 eV to the GaInP bandgap of 1.85 eV (with E_{gb} held at the GaAs bandgap of 1.42 eV) the efficiency falls very rapidly, from 35 to 30%. This dropoff is due to the dependence of the top- and bottom-subcell photocurrents on the top-subcell bandgap shown in Figure 8.7a. As noted above, lowering the top-subcell bandgap while holding the bottom-subcell bandgap constant increases the top-subcell J_{SC} at the expense of the bottom-subcell J_{SC}.

We now consider the effect of thinning the top cell to the thickness required for current matching. Figure 8.9b shows contours of cell efficiency versus top- and bottom-subcell bandgap, for optimal top-subcell thickness. The figure shows that the cell efficiency at $\{E_{gt} = 1.85\,eV$ $E_{gb} = 1.42\,eV\}$ is about 35%, a gain of about 5% absolute over the case without top-cell thinning. The figure also shows contours of the optimal top-subcell thickness. The optimal thickness decreases with increasing E_{gb} or decreasing E_{gt}, as required to maintain current matching. The thick dashed line is the contour of infinite top-subcell thickness; above this contour, the tandem cell current is

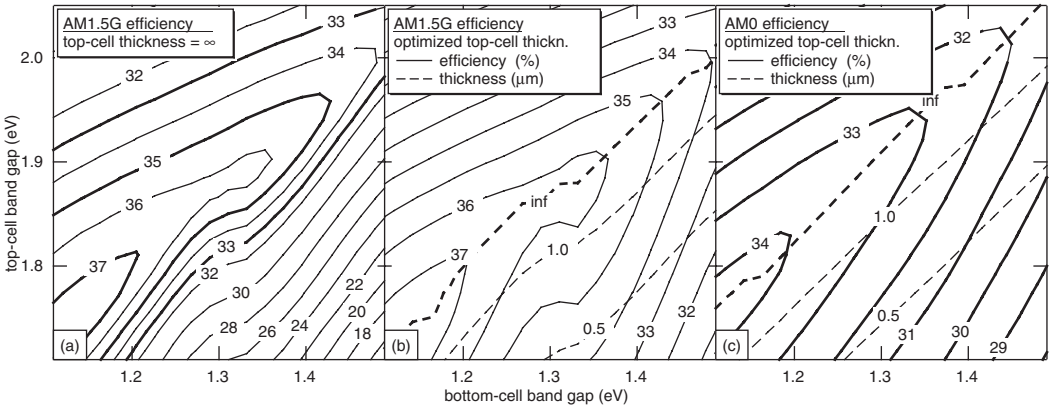


Figure 8.9 Contour plots of efficiency vs subcell bandgaps for a series-connected two-terminal two-junction tandem cell, adapted from reference [7] Panel (a) is calculated for the AM1.5 global spectrum, with an infinitely thick top subcell. Panel (b) is calculated for the same spectrum, but for the top-subcell thickness that optimizes the tandem-cell efficiency at each combination of top- and bottom-cell band gap; the dashed contours show this optimal thickness. In the region of the graph above the infinite-thickness contour, the tandem current is limited by the top cell. Panel (c) shows efficiencies and optimal top-cell thicknesses calculated as in (b), but for the AM0 spectrum

always limited by the top subcell, whereas below this contour, thinning the top subcell improves the tandem-cell efficiency. Comparing the optimized efficiencies with the infinite-thickness efficiencies of Figure 8.9a, we see that the top-subcell thinning greatly reduces the sensitivity of the tandem efficiency on subcell bandgap, in effect widening the range of bandgaps that can be selected.

8.5.6.2 Three-junction cells

The concepts and approach used above for two-junction cells can be applied to cells with any number of junctions. Figure 8.10 shows the calculated efficiency of a three-junction cell as a function of the three subcell bandgaps. The bottom subcell is taken to be optically thick, and the two subcells above it are optically thinned to optimize the efficiency. The figure is calculated at 500-suns concentration using the low-aerosol-optical-depth (AOD) AM1.5 direct spectrum. Under these conditions, a maximum efficiency of almost 53% is obtained at a bandgap combination of {1.86, 1.34, 0.93 eV}; there is also a local maximum at {1.75, 1.18, 0.70 eV} with comparable efficiency. Several bandgap combinations of particular interest are indicated. The bandgap combination {1.86, 1.39, 0.67 eV} indicated by the cylinder-shaped marker corresponds to the lattice-matched GaInP/Ga(In)As/Ge three-junction structure which is commercially available from several vendors; practical considerations for this device are discussed later in this chapter. Two other bandgap combinations are indicated as well, corresponding to device structures under development; these structures are discussed below as well.

8.5.7 Spectral Effects

The amount of light distributed to each subcell, and thus, the photocurrents generated by each subcell, is determined by the spectrum of the incident light. (See Chapter 17 for more complete

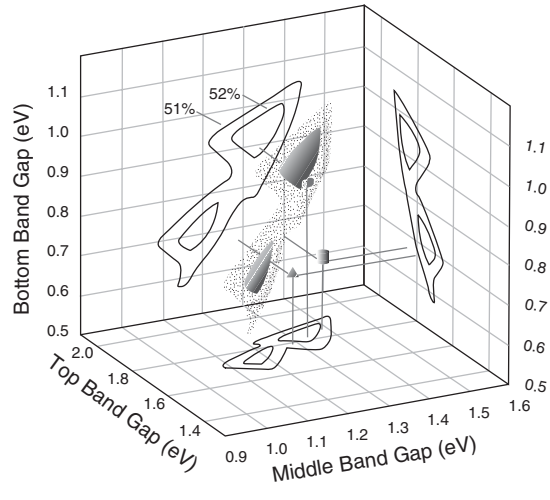


Figure 8.10 Efficiency vs subcell bandgaps for a series-connected two-terminal thickness-optimized three-junction cell under the low-AOD direct spectrum at 500-suns concentration and 300K temperature. 52 and 51% isoefficiency surfaces are shown, as indicated by their projections onto the two-dimensional contours. Bandgap combinations of actual champion devices discussed in the text are also shown with markers of various shapes: {1.86, 1.39, 0.67} eV cylinder; {1.80, 1.29, 0.67} eV, cone; and {1.83, 1.34, 0.89} eV, sphere. Adapted from reference [14]

discussion of spectra and absorption.) Therefore, the optimal bandgaps and the optimal top-subcell thinning depend on the incident spectrum. Figure 8.9c shows the efficiency versus top- and bottom-subcell bandgap for the standard AM0 spectrum for the same two-junction device as was modeled for the global spectrum in Figure 8.9b. For a given bottom-subcell bandgap, the optimal top-subcell bandgap E_{gt} is higher for the AM0 spectrum than for the global spectrum. This difference arises because the AM0 spectrum has more blue light than the global spectrum, resulting in a greater J_{SCt}/J_{SCb}; increasing E_{gt} compensates for this by directing more light to the top cell. Likewise, for a given E_{gt} and E_{gb}, the optimal top-subcell thickness is lower for the AM0 spectrum than for the global spectrum. Figure 8.7d shows J_{SCt} and J_{SCb} for the {1.85, 1.42} eV bandgap pairing as a function of top-subcell thickness as in Figure 8.7a, but calculated for the AM0 spectrum instead of the global spectrum. Because the AM0 spectrum is blue-rich compared with the AM1.5 global spectrum, the optimal top-subcell thickness, about 0.5 μm, is correspondingly less.

8.5.7.1 Spectral fluctuations

While the analysis above shows how to choose a top-cell thickness for a given spectrum, no one spectrum precisely represents the actual spectrum seen by a terrestrial solar cell. Variations in the spectrum due to the changing position of the sun in the sky, and changing atmospheric conditions, are quite significant. The detailed implications for tandem-cell design are complicated [21, 22]. In general, it is found that series-connected tandem cells are sensitive to fluctuations in air mass in particular. Fortunately, efficiency at high air mass is relatively unimportant because the net power output is small under these conditions. Overall, a well-designed multijunction cell is expected to outperform a single-junction cell by a comfortable margin, even when spectral fluctuations are taken into account. (It is interesting to note that spectral fluctuations are less of a concern for voltage-matched devices than for current-matched devices, because a change in a subcell current is reflected only logarithmically in the corresponding change in voltage.)

8.5.7.2 Chromatic aberration

A related issue of concern for series-connected multijunction cells is the chromatic aberration of the spectrum that can be caused by the concentrating optics in concentrator modules, especially the Fresnel lenses used in some concentrator configurations [23]. Chromatic aberration will result in a position-dependent variation in the spectrum across the cell. For low-concentration applications, the detrimental effect of such spatial variations may be mitigated by making the emitters of the bottom subcells highly conductive [24]. This aids current matching within a given cell by allowing more lateral conduction to equalize photocurrent variation between adjacent areas on the cell.

8.5.8 AR Coating Effects

8.5.8.1 Introduction

Our discussion has thus far assumed for simplicity that there is no reflection of the incident light from the front surface of the cell. However, without an antireflective (AR) coating, III–V cells typically have large reflectances of the order of 30% in the spectral region of importance for converting the solar spectrum. AR coatings can reduce this reflectance to ∼1%, but only over a limited spectral range; this limitation has important implications for tandem-cell current matching. An examination of AR coating issues is therefore especially important for multijunction cells, and is provided in this section. In subsequent sections, we revert to the assumption of no reflection (i.e. an ideal AR coat).

8.5.8.2 Calculation

In analyzing the effect of AR-coating parameters on multijunction performance, it is useful to have a realistic numerical model for the reflectance. Here, we use the relatively simple model of Lockhart and King [25]. This model calculates the normal-incidence reflection of a three-layer coating. Each layer is assumed to be lossless so that the index of refraction n_j of the jth layer, along with its thickness d_j, fully characterizes the optical properties of the layer. The layers are numbered $j = 1$ to 4, with $j = 4$ being the top layer and $j = 1$ being the substrate. For example, for a two-layer MgF_2/ZnS AR coat on a GaInP cell with an AlInP window layer, layers 4/3/2/1 are respectively MgF_2/ZnS/AlInP/GaInP. The reflection R as a function of wavelength λ is given by

$$R = |(X - 1)/(X + 1)|^2, \tag{8.17a}$$

where

$$X = [n_2(n_3n_4 - n_2n_4t_2t_3 - n_2n_3t_2t_4 - n_3^2t_3t_4) + i\, n_1(n_3n_4t_2 + n_2n_4t_3 + n_2n_3t_4 - n_3^2t_2t_3t_4)]/ \\ [n_1n_4(n_2n_3 - n_3^2t_2t_3 - n_3n_4t_2t_4 - n_2n_4t_3t_4) + i\, n_2n_4(n_2n_3t_2 + n_3^2t_3 + n_3n_4t_4 - n_2n_4t_2t_3t_4)], \tag{8.17b}$$

and

$$t_j = \tan(2\pi\, n_j d_j/\lambda). \tag{8.17c}$$

Although this approach incorrectly assumes no absorption from the layers of the top subcell, and completely neglects all the deeper layers in the cell stack, in practice it gives results that agree reasonably well with more rigorous approaches [26], and has the virtue of simplicity. The AR-coating calculations, which illustrate the following discussion, were done using Equations (8.17). Note, however, that more complex problems such as the calculation of back-surface AR coatings needed for mechanical stacks require the more rigorous formalism.

8.5.8.3 Current-matching

Figure 8.11a shows the modeled reflectance for a GaInP cell with a MgF_2/ZnS AR coating, for several different combinations of the layer thicknesses [27], using optical constants from the literature [28, 29]. The dependence of the reflectance on the layer thicknesses can, roughly, be broken down into two parts: the ratio of the two thicknesses, and the total thickness of each layer. Proper choice of the ratio, as for the layer thicknesses in Figure 8.11a, yields a reflectance with a flat, low, notch-shaped minimum. With this ratio held constant, the total thickness of the coating determines the position of the minimum, with increasing thickness shifting the notch position to lower photon energy. The width of the notch is less than the solar spectral range, so no matter what the position of the notch, the photocurrents of the subcells will be less than what they would be for the ideal case of zero reflectance. Shifting the notch to higher photon energy will send more light to the top subcell at the expense of the bottom subcell, and vice versa; the AR coating thus affects the current matching in the cell. Figure 8.11b shows the photocurrent of a GaInP(1.85 eV)/GaAs(1.42 eV) tandem cell as a function of ZnS/MgF_2 AR-coating thicknesses, for the AM1.5 global spectrum. As the top-subcell thickness increases, the optimal AR-coating thickness increases to compensate by directing more of the light to the bottom subcell.

8.5.9 Concentration

Terrestrial application of high-efficiency multijunction solar cells is generally in concentrator systems, given the high solar cell costs. These types of cells are well suited for concentrator operation,

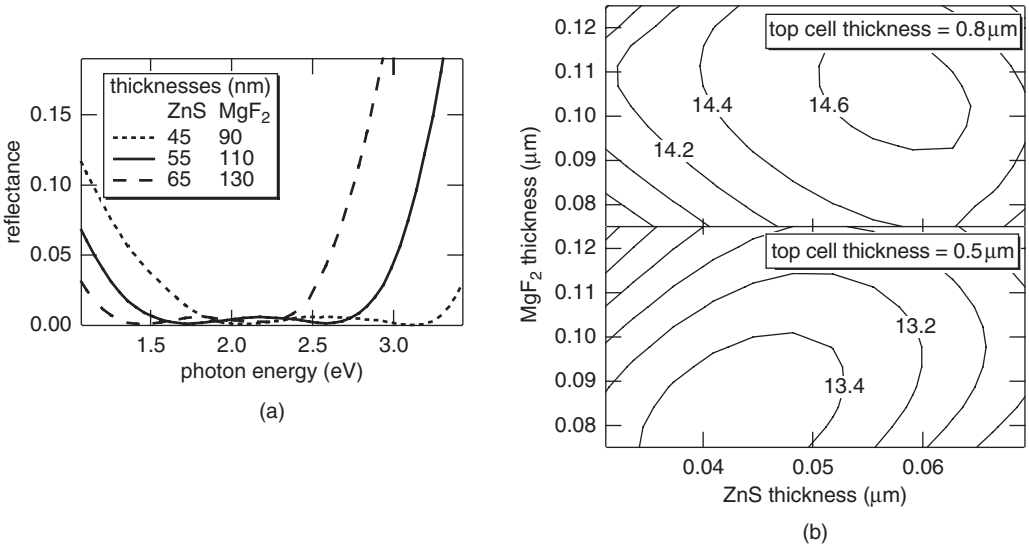


Figure 8.11 (a) Calculated reflectance for a GaInP cell with a MgF_2/ZnS AR coating, for several different combinations of the layer thicknesses. (b) Calculated photocurrent (in mA/cm^2) of a GaInP/GaAs two-terminal tandem cell under the AM1.5 global spectrum as a function of ZnS/MgF_2 AR-coating thicknesses, for two different top-subcell thicknesses. The top subcell has a 25-nm-thick AlInP window layer in both cases

not only because of their high one-sun efficiencies, but also because these high efficiencies can be maintained up to concentration levels exceeding 1000 suns. This section discusses the adaptations of the one-sun devices needed for suitable concentrator operation, and describes the resulting concentrator performance to be expected from these devices. A detailed discussion of general issues in concentrator PV is given in Chapter 10; see also reference [30].

8.5.9.1 Spectrum

Cells in a terrestrial concentrator module will be exposed to a spectrum containing significantly less high-energy light than the AM0 spectrum to which multijunction devices are exposed in space applications. This difference calls for the thickness of the terrestrial-concentrator top subcell to be greater than that of a space cell, to satisfy current-matching requirements. Current-matching top-subcell thicknesses for GaInP/GaAs tandem cells are on the order of 0.5 μm for the AM0 spectrum, as noted above, and 0.9 μm for the standard concentrator spectrum [31]. In practice, the situation is not so simple, because the spectrum that a terrestrial cell sees will vary as a function of time. The cell design also depends on cell temperature, which will depend on the details of the module (and which, of course, will also vary with time). Further discussion of these issues for multijunction concentrators is given in reference [32].

8.5.9.2 Concentration dependence of efficiency

Equation (8.13) shows that, for each decade of increase in J_{SC} due to a corresponding increase in the incident light flux, V_{OC} will increase by $(kT/e)\ \ln(10) = 60\,\text{mV}$ for an ideal $n = 1$ junction at 300 K. For a series-connected multijunction device, each junction will contribute this amount

to the net increase in V_{OC} with concentration. This increase in V_{OC} gives a significant boost to cell efficiency with concentration, a boost that is relatively greater for low-bandgap junctions. For instance, a two-junction GaInP/GaAs cell with a one-sun V_{OC} of 2.4 V will go to $V_{OC} = 2.76$ V at 1000 suns for an increase of 15%, whereas a three-junction GaInP/GaAs/Ge cell with a one-sun V_{OC} of 2.6 V will go to $V_{OC} = 3.14$ V at 1000 suns for an increase of 21%. For a cell with negligible series resistance, the fill factor will also increase with concentration, although not in as numerically simple a fashion as does V_{OC}. The increase with concentration is proportionally much less than for V_{OC}; the fill factor typically increases by around 1–2% as the concentration is raised from 1 to 1000 suns, for the ideal case of no series resistance.

It is interesting to note that while series-connected multijunction devices maintain their current matching with increasing light intensity (assuming that the spectrum does not change), the increase in junction voltage with concentration means that voltage-matched devices are voltage-matched only for a fixed concentration ratio.

8.5.9.3 Series resistance and metallization

In practice, of course, series resistance is unavoidable. The resulting J^2R power loss scales as the square of the current, and thus eventually becomes a dominant factor in the cell efficiency with increasing current. This series resistance will manifest itself as a loss in the fill factor (and, at very high currents or for very high resistance, in J_{SC} and V_{OC} as well [33]). Series-connected multijunction cells, which distribute the spectrum into several subcells and are thus inherently lower-current devices than single-junction cells, are therefore at a great advantage in minimizing J^2R losses at high concentrations. For example, the GaInP/GaAs tandem operates at half the current of a single-junction GaAs cell, and thus suffers only one-quarter the J^2R loss for a given resistance and concentration.

Even with this low-current advantage of multijunction cells, in adapting a cell from one-sun to concentrator operation, series resistance must be addressed. One of the most vital adaptations of a cell design for concentrator operation is the front-contact metallization [34]. The series resistance of the cell depends on the density of front-contact grid fingers [35]; a grid design optimized for 1000 suns will have a much higher density of grid fingers than a one-sun grid. Grid-finger spacings of 200 μm or less are not unusual at 1000 suns. Naturally, decreasing the grid-finger spacing increases the device shadowing and so decreases the current; thus, concentrator grid design involves careful trade-offs of the shadowing versus the series resistance. Fortunately, with the sophisticated photolithography/evaporation/lift-off metallization processing used for high-efficiency devices, grid-finger widths of the order of 3 μm, with height/width aspect ratios of two or more, can be achieved, though widths of 5 μm with aspect ratios of 0.6–1 are more typical. Such finger dimensions allow a very dense packing of high-conductivity grids, while maintaining a reasonably low shadow loss.

An additional approach to decreasing the series resistance is to raise the emitter conductivity in the top subcell by raising the emitter doping and/or thickness (for uniformly illuminated monolithic two-terminal devices, there is no lateral conduction in the emitters of the other subcells, so only the top-cell emitter conductivity is important). Because doing so may decrease the quantum efficiency of the top subcell, increasing the conductivity is a trade-off that must be made carefully. Achieving a sufficiently high emitter conductivity is easier for n/p than for p/n devices, because of the higher *majority*-carrier mobility in n-type material than in otherwise comparable p-type material.

At least as important to the concentrator operation of monolithic two-terminal multijunction devices is the necessity for tunnel-junction interconnects with low series resistance, high peak tunneling currents, and low absorption losses. These considerations are discussed in more detail later in this chapter.

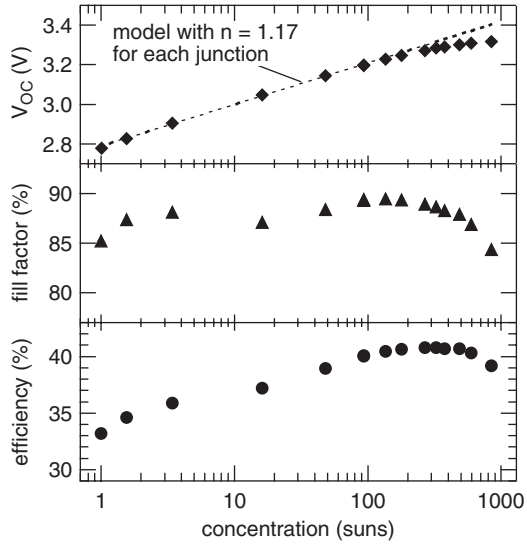


Figure 8.12 Efficiency, V_{OC}, and fill factor of state-of-the-art GaInP/$Ga_{0.96}In_{0.04}As$/$Ga_{0.63}In_{0.37}As$ three-junction cell as a function of concentration. J_{SC}, not shown, is assumed to increase proportionally to concentration

8.5.9.4 Measured performance of multijunction concentrator cell

The considerations of device performance versus concentration that we have just described are well illustrated by the plot of efficiency versus concentration for a state-of-the-art GaInP/$Ga_{0.96}In_{0.04}As$/$Ga_{0.63}In_{0.37}As$ {1.83, 1.34, 0.89 eV} three-junction concentrator device [14] (Figure 8.12; this cell structure is discussed in more detail in Section 8.9.2). For concentrations up to about 400 suns, V_{OC} increases with concentration at a rate corresponding to a near-ideal $n = 1.17$ ideality factor for each junction, as illustrated by the dashed line in the top panel of the figure. The fill factor increases with concentration roughly as would be expected for an ideal junction and then starts to roll over at higher concentrations due to series resistance, resulting in an efficiency that peaks at 40.8% efficiency at 326 suns.

8.5.9.5 Linearity

In measuring the concentration dependence of device performance, it is usually assumed that J_{SC} is linear with concentration, so that J_{SC} can be used as the measure of the concentration level. The assumption of linearity is generally considered to be quite good for III–V devices [36]. Although a detailed discussion of linearity is outside the scope of this chapter, it is worth mentioning because different degrees of nonlinearity for different subcells in the device could lead to a cross-over between top-cell-limited and bottom-cell-limited performance as a function of concentration.

8.5.10 Temperature Dependence

Typical cell operating temperatures are ~40–80 °C for concentrator cells, and ~55–85 °C for space cells in earth orbit. To predict device performance at realistic operating temperatures, and to be

able to interpret measured device characteristics at these temperatures, it is useful to analyze the temperature coefficients of these devices using the basic cell equations (8.1–8.16) [37–39]. In the following we neglect the temperature dependence of materials parameters such as diffusion lengths and minority-carrier lifetimes, so that the explicit temperature dependence of J_0 becomes

$$J_0(T) \approx \text{const} \times T^3 \exp(-E_g/kT). \tag{8.18}$$

We shall see that the current-matching constraint for series-connected multijunction cells leads to effects in the temperature coefficients that are not seen for single-junction devices.

8.5.10.1 V_{OC}

For a single (sub)cell, combining Equations (8.13) and (8.18), differentiating with respect to temperature and omitting a numerically negligible term, we get the convenient analytical expression [37, 40]

$$\frac{dV_{OC}}{dT} \approx \frac{1}{T}\left[V_{OC} - \frac{3kT}{e} + \frac{T}{e}\frac{dE_g}{dT} - \frac{E_g}{e}\right]. \tag{8.19}$$

Taking the GaInP/GaAs two-junction cell as an example, the application of Equation (8.19) yields $dV_{OC}/dT \approx -2.2$ mV/°C for the GaInP subcell and -2.0 mV/°C for the GaAs subcell. In general, because the series-connected multijunction V_{OC} is simply the sum of the subcell V_{OC} values, the temperature coefficient dV_{OC}/dT of the multijunction V_{OC} is likewise the sum of the dV_{OC}/dT values for the subcells. The GaInP/GaAs cell therefore has $dV_{OC}/dT \approx -4.2$ mV/°C [37]. In Table 8.2 we compare temperature coefficients for several types of cells. The GaInP/GaAs/Ge three-junction cell is seen to have a more negative $1/V_{OC}dV_{OC}/dT$ value than the GaInP/GaAs two-junction cell, due to the contribution of the low-bandgap Ge junction. However, going to high

Table 8.2 Modeled V_{OC} temperature coefficients dV_{OC}/dT at 300 K for multijunction cells and their single-junction component subcells, assuming junction ideality factors $n = 1$. Also shown are the relative temperature coefficients $1/V_{OC}\ dV_{OC}/dT$, in units of percentage change in V_{OC} per degree. One-sun illumination is assumed except as noted. Values are guidelines for comparison with actual cells, but precise agreement should not be expected, especially for junctions whose ideality factor deviates significantly from $n = 1$. For comparison, data are also shown for a passivated-emitter rear locally diffused (PERL) Si cell, which has a higher V_{OC} and hence a less-negative temperature coefficient than those of standard Si cells [42]

Cell	V_{OC} (mV)	dV_{OC}/dT (mV/K)	$1/V_{OC}\ dV_{OC}/dT$ (%/K)
Ge	200	−1.8	−0.90
GaAs	1050	−2.0	−0.19
GaInP	1350	−2.2	−0.16
GaInP/GaAs	2400	−4.2	−0.17
GaInP/GaAs/Ge	2600	−6.0	−0.23
GaInP/GaAs/Ge (500 suns)	3080	−4.5	−0.15
PERL Si	711	−1.7	−0.24

concentrations significantly reduces the magnitude of $1/V_{OC}\ dV_{OC}/dT$, due to the increase in V_{OC} with concentration. Furthermore, the magnitude of dV_{OC}/dT itself decreases under concentration, as can be seen from Equation (8.19): the only concentration-dependent term is V_{OC}/T, which increases under concentration, thus making dV_{OC}/dT less negative. Under many circumstances, Equation (8.19) gives remarkably good agreement with experiment, with small deviations attributed to the neglect of the temperature dependence of the materials parameters [39]. However, there are circumstances for which measurements of dV_{OC}/dT for a GaInP/GaAs/Ge three-junction cell showed concentration dependence not accounted for by this simple model [41].

8.5.10.2 J_{SC}

Although the V_{OC} temperature coefficients for the subcells of a series-connected multijunction cell are independent and additive, the multijunction J_{SC} temperature coefficient is more complex. Again taking the GaInP/GaAs tandem as an example, recall that the GaAs subcell J_{SC} depends not only on the GaAs bandgap, but also on the GaInP bandgap, because the GaInP subcell filters the light to the GaAs subcell. When the tandem-cell temperature is raised, the bottom-subcell bandgap decreases, tending to increase its J_{SC}; at the same time, however, the top-subcell bandgap also decreases, which decreases the amount of light going to the bottom cell and thus minimizes the increase in the bottom-subcell J_{SC} with temperature.

The tandem J_{SC} is limited by the smallest subcell J_{SC}. In general, these subcell J_{SC} values will not have identical temperature coefficients. For a tandem cell that is nearly current matched, there will therefore be a crossover temperature below which the tandem J_{SC} is limited by one subcell and above which the tandem J_{SC} is limited by the other subcell. Figure 8.13 illustrates this cross-over for a modeled GaInP/GaAs tandem cell that is slightly top-subcell limited at 300 K. The top-subcell J_{SC} increases faster with temperature than does the bottom-subcell J_{SC}, leading to a cross-over from top- to bottom-subcell limited as the temperature is raised above 350 K. The tandem dJ_{SC}/dT likewise crosses over from dJ_{SCt}/dT to dJ_{SCb}/dT.

8.5.10.3 Fill factor

Because the tandem-cell fill factor is determined more by the current-limiting subcell than by the other subcell(s), the current-matching crossover has similar implications for dFF/dT as for dJ_{SC}/dT. For the cell of Figure 8.13, the cross-over from top- to bottom-subcell limited causes dFF/dT to change, as shown in panel (c).

8.5.10.4 Efficiency

Because the efficiency is proportional to $V_{OC} \times J_{SC} \times FF$, and dJ_{SC}/dT and dFF/dT change in opposite directions as the temperature goes through the current-matched temperature, $dEff/dT$ is a relatively smooth function of temperature. Figure 8.13 shows that for a GaInP/GaAs two-junction cell at one sun and room temperature, $dEff/dT \sim -0.05\%$/K (absolute efficiency percentage points) which corresponds to a relative efficiency loss $(1/Eff)(dEff/dT) \sim -0.15\%$/K. By comparison, Si cells typically have a larger relative efficiency loss $(1/Eff)(dEff/dT) \sim -0.5\%$/K, due to the low Si bandgap compared with that of GaAs or GaInP. Under concentration, the temperature coefficient is even less, as quantified in Table 8.2. Also, concentrator modules are often assigned a power rating according to ambient temperature, instead of module temperature (as is usually used for Si modules), implying that the difference between rated and observed performance caused by elevated temperatures is surprisingly small for concentrator modules.

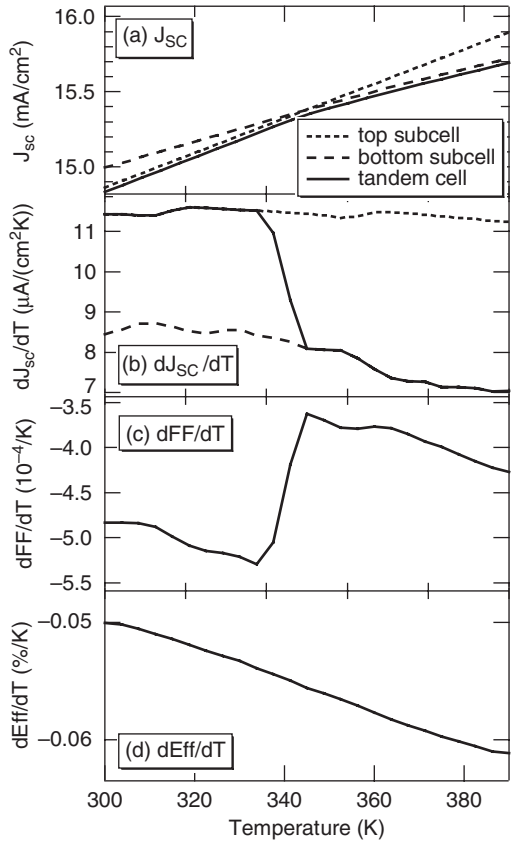


Figure 8.13 (a) Subcell and corresponding tandem-cell J_{SC}s values calculated as a function of temperature for a GaInP/GaAs tandem cell that is slightly top-subcell current-limited at 300 K. (b) The corresponding temperature derivatives dJ_{SC}/dT. As the cell temperature is raised above ~340 K, the cell crosses over from top limited to bottom limited, and dJ_{SC}/dT changes correspondingly. (c) Tandem-cell fill-factor temperature derivative dFF/dT. (d) Efficiency temperature coefficient $dEff/dT$

8.6 MATERIALS ISSUES RELATED TO GaInP/GaAs/Ge SOLAR CELLS

8.6.1 Overview

In the previous section, we discussed the basic elements of a monolithic, multijunction solar cell, such as that shown in Figure 8.5, including subcell bandgaps and thicknesses, metallization, and AR coating effects. We also assumed unity collection efficiency for all photogenerated carriers, implying that component semiconductor materials, interfaces, and junctions are virtually perfect. In practice, however, numerous intrinsic and extrinsic factors tend to limit the quality and performance of multijunction solar cells. In this section, we review various aspects of real materials and devices, including issues associated with their growth.

8.6.2 MOCVD

The GaInP/GaAs/Ge dual- and triple-junction solar cells are produced commercially in relatively large volume by several companies in the United States, Germany and Taiwan. These cells are fabricated in large metal–organic chemical vapor deposition (MOCVD) reactors made by Veeco Corporation in the United States and Aixtron in Germany. Although these devices can be grown by other techniques such as molecular beam epitaxy (MBE) [43], the predominant growth technique is MOCVD and is, therefore, the focus of this section. A description of the basic MOCVD reactors used at NREL can be found elsewhere [44]. Briefly, most of the results presented here are from layers and devices grown by MOCVD using trimethylgallium (TMG), trimethylindium (TMI), arsine, and phosphine in a Pd-purified H_2 carrier gas. The dopants sources included H_2Se, Si_2H_6, diethylzinc (DEZ) and CCl_4.

8.6.3 GaInP Solar Cells

8.6.3.1 Lattice matching

One of the major advantages of the monolithic GaInP/GaAs/Ge solar cell is that it is composed of semiconductors that are all closely lattice-matched. The fabrication of such a monolithic structure is achieved by the generic process of heteroepitaxy, the specific process being MOCVD. Close lattice matching makes the job of heteroepitaxy much easier, especially for chemically similar materials such as AlGaAs or GaInP on GaAs. The heteroepitaxy of lattice-mismatched materials is generally more difficult. The lattice mismatch is accommodated by nucleation and propagation of dislocations in concentrations that depend on the amount of mismatch and thickness of the individual layers. These dislocations are often centers for nonradiative recombination, in effect limiting the minority-carrier lifetime or diffusion length, and ultimately the efficiency of the device.

The lattice constant of the semiconductor alloy $Ga_xIn_{1-x}P$ is linearly related to the composition x by:

$$a_{Ga_xIn_{1-x}P} = xa_{GaP} + (1-x)a_{InP}, \tag{8.20}$$

where $a_{GaP} = 0.54512\,\text{nm}$ and $a_{InP} = 0.58686\,\text{nm}$ are the lattice constants of GaP and InP, respectively (see Table 8.3). An epitaxial layer of $Ga_xIn_{1-x}P$ on GaAs with $a_{GaAs} = 0.565318\,\text{nm}$ will be lattice matched to GaAs at 25 °C for $x = 0.516 = x_{\text{LM}}$. The quality of a thin, epitaxial $Ga_xIn_{1-x}P$ layer is relatively good for small variations of x around x_{LM}. This case is shown in Figure 8.14, where a broad-spectrum photocurrent from an electrolyte/$Ga_xIn_{1-x}P$ junction (see Section 8.7.1) is plotted as a function of $\Delta\theta$. The quantity $\Delta\theta$ is measured by double-crystal X-ray rocking-curve diffraction and is a measure of x [45]. If the thickness of the $Ga_xIn_{1-x}P$ layer is less than the x-dependent critical thickness, then the epilayer is said to be coherent with the substrate and

$$\Delta\theta = \tan\theta_B \left(\frac{xa_{GaP} + (1-x)a_{InP} - a_{GaAs}}{a_{GaAs}} \right) \left(\frac{1 + (\nu_{GaP}x + \nu_{InP}(1-x))}{1 - (\nu_{GaP}x + \nu_{InP}(1-x))} \right), \tag{8.21}$$

where θ_B is the Bragg angle and $\nu_{GaP}x + \nu_{InP}(1-x)$ is the Poisson ratio for $Ga_xIn_{1-x}P$ obtained from Poisson ratios for GaP and InP (see Table 8.3). (The Poisson ratio is defined as the negative of the ratio of the lateral to longitudinal strains under uniaxial stress.) If the epilayer is fully relaxed, the last multiplicative term of Equation (8.21) is replaced by unity. A plot of $\Delta\theta$ versus x for $Ga_xIn_{1-x}P$ on GaAs for these two extremes is shown in Figure 8.15a. The critical layer thickness is the epilayer thickness for which the elastic energy created by strain exceeds the energy associated

Table 8.3 Lattice constants and Poisson ratios for selected III–V binary compounds. Values from reference [47]

Material	Lattice constant [nm]	Poisson ratio, ν
AlP	0.546354	
GaP	0.54512	0.307
InP	0.58686	0.360
$Ga_xIn_{1-x}P$ [x = 0.516]	0.56532	0.333
GaAs	0.565318	0.311
InAs	0.60583	0.352
Ge	0.5657906	0.273

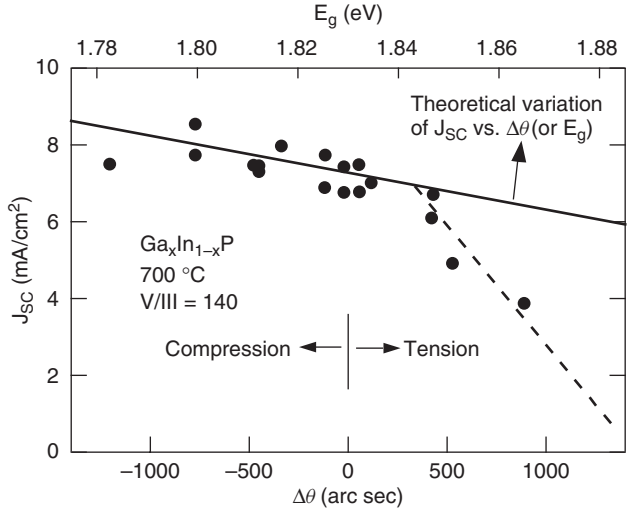


Figure 8.14 Saturated photoelectrochemical-current density of $Ga_xIn_{1-x}P$, a measure of the optoelectronic properties of the epilayer, as a function of the lattice mismatch as measured by X-ray rocking-curve peak separation in units of arc seconds, or equivalently as a function of resulting E_g [45]. The growth temperature is 700 °C and V/III ratio (PH_3/(TMGa + TMIn)) is 140. The dashed line is included to guide the eye. The intersection of the dashed line with the solid line is the critical $\Delta\theta$ above which the epilayer strain is relieved by dislocations, which in turn degrade the optoelectronic quality of the epilayer

with the creation of dislocations that relieve the elastic strain. Below the critical layer thickness, the lowest energy state of the system is an epilayer with a lattice constant, in the plane of the interface between the epilayer and the substrate, equal to the lattice constant of the substrate. The layer is then said to be coherent with the substrate. Above the critical epilayer thickness, the lowest energy state is one composed of some epilayer strain and some strain-relieving dislocations. The problem was first solved by Matthews and Blakeslee [46]. Note that for a given epilayer thickness, there also exists a critical lattice mismatch (or critical $\Delta\theta$) that defines the boundary between coherent (dislocation-free) and incoherent growth. The relationship between lattice mismatch and layer thickness is shown in Figure 8.15b.

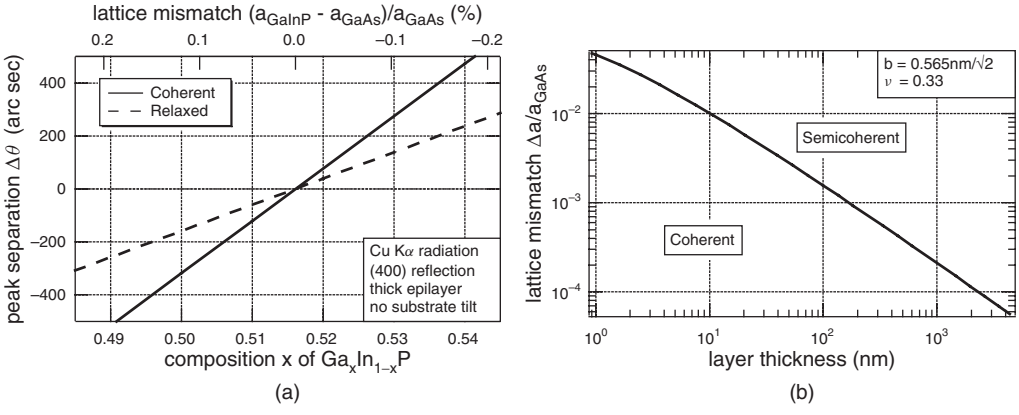


Figure 8.15 (a) A plot of $\Delta\theta$ versus x for $Ga_xIn_{1-x}P$ on GaAs; (b) plot of equilibrium boundary between coherent and semi-coherent epitaxy as a function of the absolute value of the epilayer lattice mismatch and thickness for a Poisson ratio ν and a Burgers vector b, characteristic of $Ga_xIn_{1-x}P$

Referring to Figure 8.14, for $\Delta\theta = 0$, the critical layer thickness is infinite and J_{sc} is a measure of the intrinsic minority-carrier transport quality of the epilayer in the absence of misfit dislocations. The solid line with negative slope is the theoretical variation of J_{sc} with $\Delta\theta$. For $\Delta\theta < 0$, the epilayer is In-rich ($x < x_{LM}$), and its bandgap is lower than that of lattice-matched $Ga_xIn_{1-x}P$. Hence, J_{sc} increases with decreasing $\Delta\theta$. For $\Delta\theta > 0$, the epilayer is Ga-rich ($x > x_{LM}$). At first, J_{sc} decreases with increasing $\Delta\theta$ in line with the In-rich portion of the curve, but then falls off rapidly with increasing $\Delta\theta$. The (critical) $\Delta\theta$ at which this occurs is a function of not only the layer thickness, but also the growth temperature and growth rate. As can be seen in Figure 8.14, the critical $\Delta\theta$ also depends on the sign of $\Delta\theta$. Indium-rich material is under compression. It is generally more difficult (and requires a larger compressive strain) to generate misfit and threading dislocations in compressively strained material compared with material under tension; hence, the difference in strain-relaxation behavior. Intuitively, the critical layer thickness for compressively strained material will generally approach that of tensively strained material as the growth temperature is increased and/or the growth rate is decreased. This behavior is relatively common. Note that this is contrary to the theoretical calculation shown in Figure 8.15b, which considers only the equilibrium state of the epilayer; the sign of the strain is, therefore, immaterial.

The thickness of the $Ga_xIn_{1-x}P$ top cell for most conditions will be of the order of 1 μm or less to match the photocurrent to that of the GaAs junction. From Figure 8.15b this would imply that the critical lattice mismatch should be less than 2×10^{-4} or $|\Delta\theta| \leq 50$ arcsec. There are several factors that tend to increase or decrease this tolerance limit:

- Material lattice-matched at room temperature is lattice-mismatched at growth temperature. This is due to a difference in the thermal expansion coefficients between $Ga_xIn_{1-x}P$ and GaAs (see Table 8.4). For kinetic reasons, it is probably more important that the layers be lattice-matched at growth temperature. A layer that is lattice-matched at a growth temperature of 625 °C will exhibit a lattice mismatch of $\Delta\theta \sim -200$ arcsec at room temperature [48], or alternatively, a layer that is lattice-matched at room temperature, would exhibit a $\Delta\theta = 200$ arcsec at 625 °C. Because it is easier to introduce misfit dislocations at high temperatures, it is probably better to grow the layer lattice-matched at the growth temperature. Hence, a ±50 arcsec tolerance at growth temperature would yield a room-temperature tolerance of $-250 < \Delta\theta < -150$ arcsec.

Table 8.4 Important Properties of Ge, GaAs, and GaInP at 298 K

	Ge	GaAs	$Ga_xIn_{1-x}P$	$Al_xIn_{1-x}P$
Atoms/cm^3	4.42×10^{22}	4.44×10^{22}		
Lattice constant [Å]	5.657906 [47]	5.65318 [47]	$= a_{GaAs}$ for $x = 0.516$	$= a_{GaAs}$ for $x = 0.532$
Energy gap [eV]	Indirect 0.662 Direct 0.803 [47]	1.424 [47]	Disordered 1.91 [51]	Indirect 2.34 Direct 2.53 [47]
Density of states				
Conduction band N_C [cm^{-3}]	1.04×10^{19}	4.7×10^{17}		
Valence band N_V [cm^{-3}]	6.0×10^{18}	7.0×10^{18}		
Intrinsic carrier concentration [cm^{-3}]	2.33×10^{13}	2.1×10^{6}		
Linear coefficient of thermal expansion [K^{-1}]	7.0×10^{-6} [47]	6.0×10^{-6} [47] 6.63×10^{-6} [48]	5.3×10^{-6} [48]	

- As mentioned above, material grown under compression is usually more stable to relaxation than material under tension, allowing one to err more toward negative values of $\Delta\theta$.
- Because of dynamical scattering effects, the measured $\Delta\theta$ for a thin ($\leq 0.1\,\mu m$) epilayer will be less than that of a thicker layer with the same composition and lattice mismatch [49].
- The value of $\Delta\theta$ for epilayers grown on nonsingular or vicinal (100) substrates (i.e. misoriented from exact (100)) is not unique, but depends on the orientation of substrate with respect to the X-ray beam. The effective $\Delta\theta$ is the average of two measurements of $\Delta\theta$. The first measurement is made in the conventional manner; the second measurement is with the sample rotated by 180° [50]. For vicinal substrates close to (100), this effect is small, usually ~10% at a misorientation of 6°; however, for {511} substrates, the effect is closer to 50%.

8.6.3.2 Optical properties of GaInP

8.6.3.2.1 Ordering in GaInP

Prior to 1986, it was generally assumed that the bandgap of a III–V ternary alloy semiconductor such as $Ga_xIn_{1-x}P$ was a unique function of the composition, and most publications showed $Ga_xIn_{1-x}P$, lattice-matched to GaAs, as having a bandgap of 1.9 eV. However, in 1986 Gomyo *et al.* [52] reported that the bandgap of $Ga_xIn_{1-x}P$ grown by MOCVD was usually less than 1.9 eV and depended on the growth conditions. In a subsequent paper [53], they showed that the bandgap shift was correlated with ordering of Ga and In on the group III sublattice. The ordered structure is CuPt-like, with alternating {111} planes of $Ga_{0.5+\eta/2}In_{0.5-\eta/2}P$ and $Ga_{0.5-\eta/2}In_{0.5+\eta/2}P$, where η is the long-range order parameter. Perfectly ordered GaInP ($\eta = 1$) would be composed of alternating {111} planes of GaP and InP. The first theoretical treatments of ordering in $Ga_xIn_{1-x}P$ were put forward by Kondow and coworkers [54] using tight binding theory and by Kurimoto and Hamada [55] using "first principles" linearized augmented plane wave (LAPW) theory.

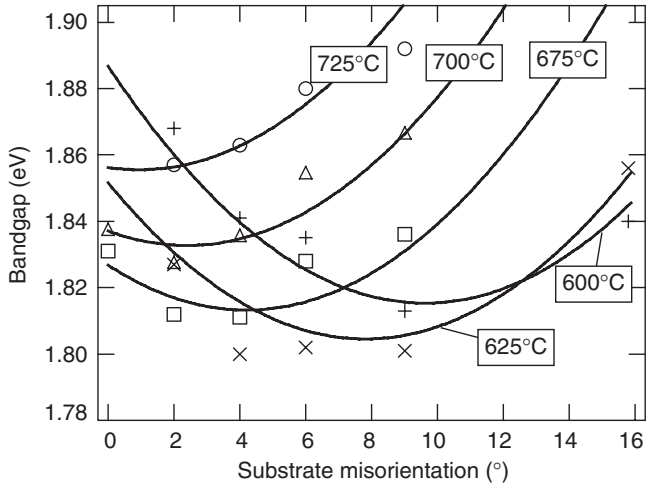


Figure 8.16 Bandgap of $Ga_xIn_{1-x}P$ vs growth temperature and substrate misorientation from (100) toward (111)B

The functional relationship between the bandgap shift defined above, ΔE_g, and the order parameter for GaInP was first published by Capaz and Koiller [56]:

$$\Delta E_g = -130\eta^2 + 30\eta^4 \text{ [in meV]}. \tag{8.22}$$

A more recent result [57] suggests that

$$\Delta E_g = -484.5\eta^2 + 435.4\eta^4 - 174.4\eta^6 \text{ [in meV]}. \tag{8.23}$$

The effects of various growth parameters on the ordering and the bandgap of $Ga_xIn_{1-x}P$ have been studied extensively. The bandgap of $Ga_xIn_{1-x}P$ is a function not only of the growth temperature T_g, but also the growth rate R_g, phosphine partial pressure P_{PH3}, substrate misorientation from (100), and doping level. Some of these effects are illustrated in Figure 8.16. Although the behavior is very complicated, there are a few characteristics that stand out. For example, for substrates that are closely oriented to within a few degrees of (100), the bandgap of GaInP, using typical values for $T_g \sim 675\,^{\circ}C$, $R_g \sim 0.1\,\mu m/min$, and $P_{PH3}/III \sim 100$, is closer to 1.8 eV than 1.9 eV. One can obtain bandgaps closer to 1.9 eV using extreme values for T_g, R_g or P_{PH3}/III, but the material quality typically suffers in some other way, e.g. minority-carrier diffusion length, composition, or morphology. The most straightforward way to higher bandgaps is by using substrates that are strongly misoriented from (100) toward {111}. The {111} surface in the zinc-blende system is the group-III-terminated surface and is often referred to as (111)A. Substrates misoriented towards $\{11\bar{1}\}$, or (111)B, generally enhance the degree of ordering. When growing on Ge there is no distinction between A and B misoriented substrates, and it is typically very difficult to control the A/B character of the III–V (GaAs or GaInP) epilayer. Hence, the easiest way to obtain high-bandgap GaInP on Ge is to use misorientation angles larger than about 15°, coupled with high R_g, moderate T_g, and low P_{PH3}/III, or to expose the growing GaInP surface to trace levels of Sb during growth. The Sb tends to "poison" the ordering process, leading to disordered high-bandgap GaInP [58]. The effect of Sb on the properties of GaInP solar cells was studied by Olson and coworkers [59].

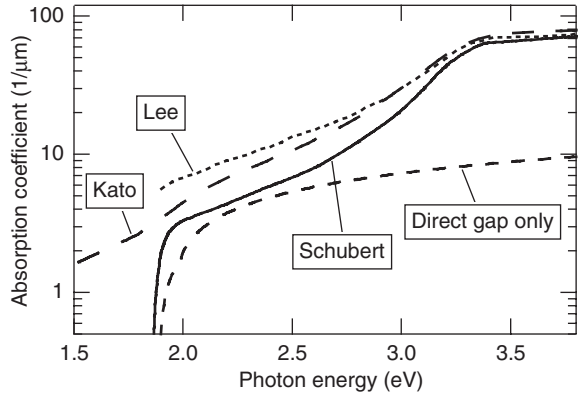


Figure 8.17 Comparison of absorption coefficients of $Ga_xIn_{1-x}P$ published in the literature. The data of Lee *et al*. [68] and Schubert *et al*. [69] are ellipsometric data. The model of Kato *et al*. [70] is a reasonably good fit at high energies, but is a poor fit for sub-bandgap photon energies. The curve marked "Direct gap only" is a plot of Equation 8.24

There are other material properties that are affected by ordering, including optical anisotropy [60–63], transport anisotropy [64, 65], and surface morphology [66, 67].

8.6.3.2.2 Absorption coefficient

For modeling and characterization of GaInP epilayers and solar cells, accurate models for the optical properties of GaInP are required. The optical constants of GaInP have been measured using spectroscopic ellipsometry and modeled by several groups and using transmission measurements by Kurtz *et al*. [10]. These results are summarized in Figure 8.17. For the most part, there is no single model that adequately describes the broad-band optical properties of lattice-matched GaInP with some arbitrary degree of ordering. The model of Kato [70] appears adequate for short wavelengths, but fails for near-gap optical properties. The models of Schubert and Kurtz only account for near-bandgap transitions. The model of Kurtz *et al*. for the absorption coefficient of GaInP, α_{GaInP}, is given by:

$$\alpha_{GaInP} = 5.5\sqrt{E - E_g} + 1.5\sqrt{E - (E_g + \Delta_{so})}\ [\mu\mathrm{m}^{-1}], \tag{8.24}$$

where the photon energy E, bandgap energy E_g, and spin-orbit energy Δ_{so} are in units of eV. The value of E_g, of course, varies with the degree of order, η, and the value of Δ_{so} is typically set to 0.1 eV, independent of η. This model reasonably accounts for the absorption associated with the two near-band transitions at E_g and $E_g + \Delta_{so}$ and is useful for deducing the minority-carrier diffusion length from near-gap photoresponse measurements [71]. It is not a good model for absorption at higher photon energies.

8.6.3.3 Doping characteristics

8.6.3.3.1 n*-type dopants*

Selenium. The element Se is a commonly used *n*-type dopant in III–V materials and is usually obtained from the decomposition of H_2Se. The doping behavior of H_2Se has been studied by a

number of investigators [72–77]. Under most growth conditions, the electron concentration increases with H_2Se flux or partial pressure and then saturates at about $2 \times 10^{19}\,cm^{-3}$, depending on T_g. The incorporation of Se also depends on the partial pressure of PH_3 (for GaInP) or AsH_3 (for GaAs). These data are best fit with an equation of the form

$$n^{-1} = (1 + \alpha P_V)(\beta P_{Se})^{-1} + k^{-1}, \tag{8.25}$$

where n is the electron concentration and P_V and P_{Se} are group V and Se partial pressures, respectively. The coefficients α and β depend on T_g and carrier flow rate (residence time) and therefore must be determined for each reactor system. This doping behavior can be derived from a modified Langmuir adsorption model that accounts for the competitive adsorption of Se and the group V species on a fixed number of sites k, which will depend, for example, on the density of steps and kinks, i.e. the substrate misorientation. This model is a much better fit than the *ad hoc* form $n \propto P_{Se}^x P_V^y$ most often published in the literature.

For electron concentrations greater than about $2 \times 10^{18}\,cm^{-3}$, the bandgap energy of GaInP increases, the ordering decreases, and the morphology of the growth surface becomes very smooth [75]. At sufficiently high fluxes of Se, the surface again begins to roughen. At the same time, the electron concentration begins to decrease [72] and Se precipitates are observed by transmission electron microscopy [1]. Selenium has been linked to DX-like centers in $(Al_xGa_{1-x})_{0.5}In_{0.5}P$ with Al concentrations greater than about $x = 0.4$ [78].

Silicon is another widely used dopant in III–V materials and devices and the most popular source is Si_2H_6. The first report of using Si_2H_6 to dope GaInP was by Hotta and coworkers [79]. They found that for $T_g < 640\,^{\circ}C$, the electron concentration n decreased with decreasing T_g due presumably to a decrease in the Si_2H_6 pyrolysis rate. For $T_g > 640\,^{\circ}C$, n saturates at about $n = 5 \times 10^{18}\,cm^{-3}$ with Si_2H_6, presumably due to formation of non-ionized complexes, such as $(Si_{III}^{+}\text{-}Si_V^{-})$ or $(Si_{III}^{+}\text{-}V_{III}^{-})$ where the subscript/superscript refers to the III–V lattice site/charge state and V_{III}^{-} is a negatively charged vacancy on a group III site. The results for Si-doped GaAs were quantitatively similar. Scheffer and coworkers [80] saw no evidence of saturation for electron concentrations up to $8 \times 10^{18}\,cm^{-3}$ using Si_2H_6, whereas Minagawa and coworkers [81] found that the electron concentration saturated at about $1 \times 10^{19}\,cm^{-3}$, essentially independent of substrate orientation and growth temperature.

It has been shown that Si delta doping (where the doping is confined to a single atomic layer or series of layers) in GaInP increases the maximum electron concentration and increases the electron mobility relative to that of uniformly doped layers [82, 83]. The conclusion from these studies is that Si delta doping yields fewer Si shallow acceptor defects. Silicon apparently does not introduce any deep states in GaInP, but does so in $(Al_xGa_{1-x})_{0.5}In_{0.5}P$ for $x > 0.3$. As with Se, Si concentrations above some critical level tend to disorder $Ga_xIn_{1-x}P$, causing the bandgap to increase. However, the details can be quite varied. Gomyo and coworkers [74] reported that a lower concentration of Si was required to disorder GaInP than that for Se. However, Minagawa and coworkers [81] found that Si concentrations closer to $1 \times 10^{19}\,cm^{-3}$, depending on T_g, were required to inhibit the ordering in $Ga_xIn_{1-x}P$.

8.6.3.3.2 p-*type dopants*

Zinc. The most common *p*-type dopant in GaInP is Zn. The typical sources are dimethylzinc (DMZ) and diethylzinc (DEZ). The Zn-doping characteristics have been studied by a number of investigators [42, 72, 73, 85]. The incorporation efficiency is typically sublinear with input flow, and increases with lower growth temperature and higher growth rate R_g. A model that accounts for some of these effects has been proposed by Kurtz *et al.* [85].

High Zn concentrations cause several problems in GaInP. Carrier concentrations in the neighborhood of $1 \times 10^{18}\,\mathrm{cm}^{-3}$ destroy the ordering in $Ga_xIn_{1-x}P$ and increase the bandgap [86]. And high Zn concentrations (or more accurately, high DEZ flows) cause a loss of In from both GaInP [45] and AlGaInP [87]. The problem is probably associated with some parasitic gas-phase reaction involving DEZ, TMIn, and PH_3. The effect can be large enough to measurably change the growth rate and the Ga/In ratio in the material. The Ga/In ratio can be so far from lattice-matched conditions as to affect the morphology. High DEZ flows also inhibit the incorporation of Ga, but to a lesser extent.

Diffusion of zinc during epilayer growth can cause degradation of the performance of solar cells [88]. The zinc dopant in the substrate, back-surface field, or tunnel-junction layers can serve as a reservoir for zinc diffusion into the base region of an n-on-p cell. The diffusion is largely driven by point defects that are injected during the growth of the n-type layers. The diffusion can be reduced by reducing the doping levels of any of the layers (including the n-type layers), by adding diffusion barriers, and/or by using Se instead of Si doping of the n-type layers [88]. A side effect of the Zn diffusion is a disordering of any ordered structures [89].

The effect of changing the cap or overlying layer and cooling atmosphere on the hole concentration in Zn-doped $(Ga_{1-x}Al_x)_{0.5}In_{0.5}P$, $x = 0.7$, has been studied by Minagawa and coworkers [90]. Cooling in an H_2 atmosphere containing AsH_3 or PH_3 reduces the hole concentration. Hydrogen radicals from the decomposition of group V hydrides easily diffuse into the epilayers and passivate the Zn acceptors. Cap layers of n- or p-GaAs help impede the indiffusion of H, and underlying layers can enhance the indiffusion of H, a special problem with p-on-n cells [88].

Magnesium doping (with cyclopentadienyl magnesium) in AlGaInP has been studied by a number of investigators [86, 91–94]. It is useful for achieving higher hole concentrations in AlGaInP and AlInP layers. However, the Mg incorporation efficiency decreases with decreasing temperature. Therefore, higher growth temperatures are favored, which is an advantage for AlGaInP. However, since dopant diffusion rates increase rapidly with temperature, this is a disadvantage for maintaining dopant profiles in tunnel junctions and IIIV/Ge heterointerfaces. For GaInP, it has no obvious advantages over Zn, and suffers from rather severe memory effects [92]. Also, relatively good-quality Zn-doped AlInP can be obtained by paying careful attention to source material and system purity [95].

Carbon. The use of C (from CCl_4 or CBr_4) has not been studied extensively. The halide tends to etch GaInP [96], with InP etching faster than GaP. Also, C-doped GaInP exhibits poor minority-carrier transport properties [96, 97].

8.6.3.4 Window layers and back-surface fields

8.6.3.4.1 AlInP window layers

The function of an emitter window layer is to passivate the surface states associated with the emitter surface. These states are minority-carrier traps. Their effect is characterized by a quantity called the surface- (or interface-) recombination velocity S. The value of S can range from 10^7 cm/s for an unpassivated GaInP emitter to less than 10^3 cm/s for a high-quality AlInP/GaInP interface. As with any solar cell, a high surface-recombination velocity will reduce the photoresponse of the GaInP solar cell, most strongly in the blue portion of the spectral response. To be an effective window layer for an n-on-p cell, the material should have a:

- Lattice constant close to that of GaInP;
- E_g much larger than that of the emitter;

- Large valence band offset with respect to the emitter to provide a potential barrier for minority holes;
- Relatively high electron concentration (of the order of $n \geq 10^{18}\,\mathrm{cm}^{-3}$); and
- Material quality sufficient to produce low interface-recombination velocity.

The semiconductor AlInP has most of the required characteristics. $Al_xIn_{1-x}P$ is lattice-matched to $Ga_{0.516}In_{0.484}P$ for $x = 0.532$. The indirect band edge of AlInP at this composition is 2.34 eV, 0.4–0.5 eV larger than that of GaInP. The AlInP/GaInP band alignment appears to be type 1 with $\Delta E_c \sim 0.75\Delta E_g$ and $\Delta E_v \sim 0.25\Delta E_g$ [98]. This implies that AlInP should provide reasonable confinement of the holes in the GaInP emitter of an n-on-p device. AlInP is easily doped n-type with either Si or Se. The internal quantum efficiency of a GaInP cell with a good $Al_xIn_{1-x}P$ window layer is greater than about 50% at a photon energy of 3.5 eV. However, there is a strong affinity between the Al in $Al_xIn_{1-x}P$ and oxygen, and oxygen is a deep donor in $Al_xIn_{1-x}P$. Hence, if the reactor chamber or the source materials are contaminated with water vapor or other oxygenated compounds, the quality of the $Al_xIn_{1-x}P$, including its conductivity, will suffer. Poor-quality $Al_xIn_{1-x}P$ will degrade the blue response of a GaInP cell and degrade the fill factor (via contact resistance) [8].

8.6.3.4.2 Back-surface barrier

The function of the top-cell back-surface barrier, often referred to as a back-surface field or BSF, is to passivate the interface between the top-cell base and the tunnel-junction interconnect (TJIC). Also, in some cases, it may help to reduce outdiffusion of dopants from the TJIC [99]. The high recombination velocity at this interface will affect both the photoresponse (in particular, the red photoresponse) and the V_{OC}. The V_{OC} effect is shown in Figure 8.8. Note that the magnitude of the effect can be quite large, and is also affected by the base minority-carrier diffusion length and thickness. The requirements of a good BSF are similar to those for a front-surface window layer. For an n-on-p GaInP cell, a BSF should have a:

- Lattice constant close to that of GaInP;
- E_g larger than E_g of GaInP;
- Large conduction-band offset with respect to GaInP;
- Relatively high hole concentration (of the order of $p = 1 \times 10^{18}\,\mathrm{cm}^{-3}$);
- Relatively good minority-carrier transport properties; and
- High transparency to photons destined for the underlying GaAs cell.

Initial results [100] implied that disordered or high-E_g GaInP was a better BSF for a low-E_g GaInP top cell compared with an AlGaInP BSF. This was probably due to oxygen contamination in the AlGaInP layer. As a deep donor, oxygen is a bigger problem in p-type AlGaInP. Other researchers found that strained, Ga-rich $Ga_xIn_{1-x}P$ was superior to either disordered lattice-matched GaInP or AlGaAs [101]. Recently, however, the best commercial tandem solar cells use AlGaInP [102] or AlInP [103] BSFs. The growth of high-quality Zn-doped AlInGaP and techniques for assessing the quality thereof have been published by several groups [93, 95, 104].

8.6.3.5 Characteristics of state-of-the-art GaInP cells

Because the bandgap of GaInP can change so dramatically with growth conditions, and because the optimal design for a GaInP subcell in a multijunction cell is not the same as the optimal design for a single-junction GaInP cell, it is not meaningful to talk of "state-of-the-art" GaInP cells using only efficiency as the measure of quality. As the bandgap of a single-junction GaInP cell increases,

V_{OC} should increase, but J_{sc} and efficiency should decrease. (Note however, that the efficiency of an ideal current-matched GaInP/GaAs tandem cell would increase slightly with increasing GaInP bandgap up to ~2 eV, as shown in Figure 8.9 above.) Also, in any optimized multijunction solar cell the thickness of the GaInP cell will likely be optically thin, i.e. will absorb fewer photons and hence *may* be lower in efficiency than an optically thick cell. Thus, thickness and bandgap are two important parameters that must be considered when comparing single-junction efficiencies. In general, relative measures of quality, e.g. V_{OC} at a given E_g, are more useful.

8.6.4 GaAs Cells

8.6.4.1 Quality of GaAs on Ge(100) substrates

Despite the close lattice matching between GaAs and Ge substrates, the quality of GaAs grown on Ge can be quite variable. (See Section 8.6.5.3 for a discussion of the heteroepitaxy of GaAs on Ge.) The primary criterion for good-quality heteroepitaxial GaAs is, of course, the efficiency of the overlying GaAs and GaInP solar cells. Generally, a good indicator of quality is a specular episurface with little or no haze, often caused by antiphase domains (APDs), and few extended defect features such as pits, hillocks, or slip lines. For a specular, epitaxial GaAs layer, one should observe a faint "cross-hatch" pattern. This "cross-hatch" is a replica or shadow of the misfit dislocation array located in the GaAs/Ge interface plane. Sometimes, the absence of this "cross-hatch" pattern is an indicator that the misfit is being relaxed by threading dislocations. The density of these threading dislocations can become high enough to affect the minority-carrier transport properties of the GaAs and GaInP solar cells, and should, therefore, be avoided.

The morphology of $Ga_xIn_{1-x}P$ grown on GaAs is an even more sensitive indicator of the quality of the original GaAs surface. Morphologically faint defects in or on the GaAs will be "decorated" by the growth of $Ga_xIn_{1-x}P$. This is probably caused by differences in the attachment of Ga and In to the different surface orientations offered by the defect.

GaAs grown on Ge can be "lattice-matched" to the Ge substrate by the addition of about 1% indium. This eliminates the "cross-hatch" in good heteroepitaxy, but does not appear to make the task of heteroepitaxy any easier. Under the best conditions, a $Ga_{0.99}In_{0.01}As$ solar cell will be slightly better than a GaAs cell on Ge [105].

8.6.4.2 Optical properties

The optical parameters of GaAs are tabulated in the work of Aspnes and coworkers [106] and a model for the optical dielectric function of GaAs (and $Al_xGa_{1-x}As$) has been proposed [107].

8.6.4.3 Window layers and back-surface fields

$Ga_xIn_{1-x}P$ and $Al_xIn_{1-x}P$ should both make excellent window layers and back-surface-field layers for GaAs solar cells [108, 109]. Both have type-I band alignment with GaAs, with reasonably adequate conduction- and valence-band offsets [98, 110]. Ideally, $Al_xIn_{1-x}P$ would make a better window layer than $Ga_xIn_{1-x}P$ because of its larger bandgap energy. However, because of its sensitivity to oxygen contamination, $Al_xIn_{1-x}P$ will probably never produce as good an interface with GaAs as does $Ga_xIn_{1-x}P$. (This is the main problem with the AlGaAs/GaAs interface used widely for single-junction GaAs solar cells [19].) The undoped $Ga_xIn_{1-x}P$/GaAs interface has one of the lowest interface recombination velocities ($S < 1.5$ cm/s) of any heterostructure ever

measured, including the SiO_2/Si interface [108]. In addition, it is difficult to dope $Al_xIn_{1-x}P$ p-type at a level of $p > 1 \times 10^{18}\,cm^{-3}$. For these reasons, $Ga_xIn_{1-x}P$ is usually the preferred window layer and BSF layer for GaAs solar cells in GaInP/GaAs tandem-cell structures.

8.6.5 Ge Cells

8.6.5.1 Optical properties of Ge

The optical and electronic properties of Ge are well documented [47]. Germanium has a lattice constant close to that of GaAs and has a diamond structure. It is also mechanically stronger than GaAs and, hence, has long been viewed as an excellent substitute for GaAs substrates. With a bandgap of 0.67 eV, it is current-matched to a thin GaAs top cell [7] and is also a good bottom-cell candidate in a four-junction stack. However, in both of these cases, it has several properties that put it at a disadvantage:

- Due to the low bandgap, the V_{OC} of a Ge junction is limited by its indirect bandgap to a theoretical V_{OC} of 0.3 V [112] and is relatively more sensitive to temperature.
- Ge is relatively expensive, and cannot be viewed as a one-sun solar cell material (with the exception of its use in space).
- Ge is an *n*-type dopant in GaAs and GaInP. In GaInP, Ge also exhibits amphoteric behavior with a compensation ratio $N_a/N_d = 0.4$ [113] and has been associated with a deep acceptor state [114].
- Ga, As, In, and P are all shallow dopants in Ge. Hence, the control of the junction formation process becomes complicated when it is combined with the III–V heteroepitaxy process (see Section 8.6.5.3).

8.6.5.2 Junction formation in Ge

Diffusion of a group V or a group III dopant into a Ge substrate is the most common junction formation process for Ge subcells. Indeed, due to the proximity of III–V epilayers and the high temperatures involved in heteroepitaxial process, diffusion of both Group III and Group V atoms into the Ge substrate is unavoidable. The challenge is to control the process so as to obtain a Ge subcell with good photovoltaic properties and simultaneously form a defect-free heteroepitaxial layer of GaAs with the appropriate conductivity type and level. A detailed description of the optimum process is beyond the scope of this chapter, but a few critical factors to consider are listed below:

- Diffusion coefficients are thermally activated. So, in general, dopants and junctions are less mobile and more stable at lower growth temperatures.
- As noted by Tobin *et al*. [115], the diffusion coefficient of As in Ge at 700 °C is higher than that of Ga, but the solid solubility of Ga is larger than that of As.
- For three-junction GaInP/GaAs/Ge devices with a reasonably good-quality Ge subcell, the only Ge-device parameter that is of consequence is the V_{OC}, because the J_{sc} of the Ge subcell is potentially much greater than that of the GaInP (or GaAs) subcell so it is not the current-limiting junction.
- One of the highest V_{OC}s of a Ge solar cell reported to date is 0.239 V [112]. This V_{OC} is a sensitive function of process conditions and is most sensitive to the quality of the III–V/Ge interface and its fabrication.

- AsH_3 etches Ge. The etch rate increases with temperature and AsH_3 partial pressure. Heavily etched, singular and vicinal Ge(100) surfaces are microscopically rough [116]. Hence, prolonged AsH_3 exposures should be avoided.
- The etch rate for PH_3 is much lower, and there appears to be much less roughening of the surface from PH_3 exposure [116]. The diffusion coefficient of P at 600 °C is about two orders of magnitude lower than that of As [47]. Hence, PH_3 may be a better group V, *n*-type dopant source than AsH_3.

8.6.5.3 III–V heteroepitaxy

Although there are a number of "recipes" for the growth of GaAs on Ge(100) with specular morphologies or low antiphase-domain (APD) or low stacking-fault densities, many present contradictory results. For example, Pelosi *et al*. [117] found that the GaAs surface morphology is best for very low V/III ratio (of the order of 1), using a moderate growth rate ($R_g \sim 3.5\,\mu m/hr$) and low growth temperature ($T_g = 600\,°C$). On the other hand, Li *et al*. [118] found the lowest APD density occurs for high V/III, low R_g, and high ($\sim$700 °C) T_g. Chen *et al*. [119] showed that "good" morphology could only be obtained for growth temperatures in a range of 600–630 °C.

The cause of this striking difference is not known with certainty. It may be due to differences in reactor design or purity. It may be related to the quality of the Ge substrates. Other research [116] would suggest that it is related to the prenucleation conditions or the structure of the (100) Ge surface immediately prior to the GaAs nucleation step.

Much has been published about the structure of the (100) Ge surface, but most of it is in regard to surfaces prepared in ultra-high-vacuum (UHV) or molecular beam epitaxy (MBE) environments. It has been shown, however, that under most conditions, AsH_3-treated surfaces in an MOCVD reactor are quite different, as explained below [116, 120, 121]. Arsenic on a (100) Ge surface terrace forms rows of dimers, similar to As on a (100) GaAs surface [120, 121]. This reduces the (1 × 1) symmetry of the unreconstructed Ge surface to a surface that now has (2 × 1) or (1 × 2) surface symmetry. For the (2 × 1) reconstruction, the dimer bonds are parallel and rows of dimers run perpendicular to the step edges. Adjacent terraces on a (100) As-terminated Ge surface can be composed of orthogonal reconstructions; adjacent terraces on a (100) GaAs are always of the same type. An As-terminated Ge surface prepared in a UHV or MBE environment usually exhibits a single-domain, (1 × 2) symmetry. An MOCVD-prepared surface will initially be (1 × 2), but tends toward (2 × 1) with a transition time that ranges from 1 minute to tens of minutes and depends on temperature, AsH_3 partial pressure, and substrate temperature. Intermediate states, of course, are composed of a mixture of (1 × 2) and (2 × 1) domains, a condition that is conducive to the formation of APDs in a GaAs heterolayer. Also, as mentioned above, AsH_3 etches Ge. This etching causes significant step bunching or faceting and microscopically rough surfaces.

8.6.6 Tunnel-junction Interconnects

The purpose of the tunnel-junction interconnect (TJIC) between the subcells of a multijunction cell is to provide a low-resistance connection between the *p*-type BSF of a subcell and the *n*-type window layer of the subcell beneath it. Without the TJIC, this *pn* junction has a polarity or forward turn-on voltage that is in opposition to that of the top or bottom cells and, when illuminated, would produce a photovoltage that could roughly negate the photovoltage generated by the top cell. A tunnel junction is simply a $p^{++}n^{++}$ junction where p^{++} and n^{++} represent degenerately doped material. The space-charge region for a $p^{++}n^{++}$ junction should be very narrow, $\sim$10 nm. For small forward biases and any reverse bias, the normal thermal current characteristic of a $p^{++}n^{++}$

junction is "shorted" by tunneling through the narrow space-charge region, and the tunnel junction behaves much like a resistor. In forward bias, for current densities greater than some critical value, called the peak tunneling current J_p, this resistor-like behavior disappears and the behavior of the tunnel-junction switches to that dominated by the usual thermionic emission and the voltage drop across the tunnel junction increases to that of a typical *pn* junction. The functional form of J_p is dominated by an exponential term of the form:

$$J_p \propto \exp\left(-\frac{E_g^{3/2}}{\sqrt{N^*}}\right), \tag{8.26}$$

where E_g is the bandgap and $N^* = N_A N_D/(N_A + N_D)$ is the effective doping concentration [122]. The value of J_p must be larger than the photocurrent of the tandem cell. For a concentrator cell operating at 1000 suns, $J_{sc} \sim 14\,\mathrm{A\,cm^{-2}}$.

The best tunnel junctions for very-high-efficiency solar cells are relatively defect free. Lifetime-limiting, midgap defects usually only add to the excess current. There is no evidence in the literature that point or extended defects add to J_p or increase the conductivity in the tunneling portion of the $I-V$ curve. High excess currents can mask a low J_p, but usually the junction conductivity is also unacceptably low. On the other hand, it is possible that high concentrations of point or extended defects can compensate donors or acceptors in the junction, leading to increased depletion width and lower tunneling currents. In addition, defects can reduce the thermal stability of the tunnel junction and the quality of overlying layers. Therefore, in general, it is usually best to grow the TJICs free of point or extended defects.

The first high-efficiency GaInP/GaAs dual-junction solar cells were fabricated using an optically thin GaAs TJIC. The best tunnel junctions were doped with C and Se. Hence, they were reasonably stable under the thermal conditions required to grow the top cell, and were capable of operating at more than 1000 suns, i.e. $J_p > 14$ A/cm^2. They were also less than 30 nm thick and obscured less than 3% of the light destined for the lower cell. With optically thick, unannealed devices, peak tunneling currents were greater than 300 A/cm^2 with excess current densities close to zero [8].

8.6.6.1 AlGaAs/GaInP TJIC

Despite the higher bandgap and its concomitant penalty, the p^{++}-AlGaAs/n^{++}-GaInP heterojunction tunnel diode proposed by Jung and coworkers is the preferred TJIC for one-sun operation and may be suitable for concentration [123]. It takes advantage of the innate propensity for AlGaAs to incorporate C and for GaInP to incorporate Se, so that high values of N^* are easily achieved. Hence, peak tunneling currents as high as 80 A/cm^2 were reported. The devices are also thermally stable; J_p is reduced to about 70 A/cm^2 for 30-min anneal at 650 °C and to about 30 A/cm^2 for a 30-min anneal at 750 °C. This TJIC is more optically transparent than a thin GaAs TJIC and therefore should yield a higher tandem cell photocurrent.

8.6.7 Chemical Etchants

The processing of epitaxial products into finished devices is beyond the scope of this chapter. Most of the processes used by industry are proprietary, and there are numerous laboratory processes, such as evaporation of metals and optical coatings, that are suitable for research. One very useful area that is common to both industrial and laboratory processes is the use of selective and nonselective etchants for the various materials used in GaInP/GaAs-based multijunction solar cells. A list of

these etchants is given below (etch rates are at room temperature). Note that solutions containing H_2O_2 typically exhibit an etch rate that depends on the age of the solution [124].

- Mixtures of ammonia, hydrogen peroxide, and water etch GaAs, but do not etch GaInP and AlInP. A common formulation is 2 parts NH_4OH, 1 part 30% H_2O_2, and 10 parts H_2O (2:1:10). Also, a solution of H_3PO_4, H_2O_2, and H_2O combined in a ratio of 3:4:1 etches GaAs and not GaInP.
- Concentrated HCl rapidly etches GaInP, but GaInP exposed to dilute HCl and HCl vapor may not etch at all. HCl does not etch GaAs.
- Dilute HCl:H_2O etches AlInP [125].
- Au metallization is impervious to both 2:1:10 and concentrated HCl.
- A 1:20 solution of HCl and CH_3COOH (acetic acid) etches GaInP at a rate of 70 nm/min and GaAs at a rate of <5 nm/min [124].
- $5H_2SO_4$: $1H_2O_2$:$1H_2O$ at room temperature etches GaInP at a rate of about 25 nm/min. It etches GaAs much more rapidly (>1 μm/min).
- Mixtures of HCl:H_3PO_4:H_2O etch GaInP [126]. For high HCl compositions, the etch rate is ~1 μm/min.

8.6.8 Materials Availability

A question of interest for all solar cell technologies is the availability of the component materials required for very-large-scale, long-term production of the cell. Predicting the long-term availability of such natural resources as gallium, indium and germanium is very difficult. This issue has been studied periodically over the years, including a recent thorough work by Andersson [127]. It appears that the material whose availability constrains production for the GaInP/GaAs/Ge structure may prove to be germanium. If so, one approach would be to forfeit the relatively small additional contribution of the Ge third junction by using the two-junction GaInP/GaAs structure grown on GaAs. Also, reuse of the substrate by lift-off of the active junctions may prove practical. In any case, high-concentration operation of these cells makes the best use of their constituent materials.

8.7 EPILAYER CHARACTERIZATION AND OTHER DIAGNOSTIC TECHNIQUES

The standard procedures for measuring light and dark $I-V$ curves and QE curves are described in Chapter 17. Here, we describe additional techniques for characterizing materials and devices.

8.7.1 Characterization of Epilayers

A modified electrochemical capacitance–voltage (ECV) profiler, available from Nanometrics, Inc. (http://www.nanometrics.com/products/ecvpro.html) can measure the carrier concentration, bandgap, and minority-carrier diffusion length of an epilayer [71, 128]. The sample is mounted in a special holder that allows the formation of front and back contacts in a fraction of the time that is required for a typical solid-state device. An ohmic contact is made to the back of the wafer by passing a surge of current (something like a spot weld). A junction is formed between the epilayer and an aqueous electrolyte (e.g., 0.1 M HCl). The capacitance–voltage ($C-V$) characteristics of this junction provide a measure of the carrier concentration as a function of depth. In a finished device, the carrier concentrations of the individual layers can be checked by profiling (etching) through the structure in the electrochemical cell, or by using selective etches to uncover the layer(s) of interest. CV measurements on a processed single-junction device (rather than on an

aqueous–semiconductor junction) tend to give lower dissipation factors because of lower series resistance, but only give information about the lightly doped side of the junction. Thus, it can be difficult to ascertain that the emitter is underdoped by a $C-V$ profile on the solid-state device. Although an ECV profiler is designed to measure the doping level of each layer as a multilayered stack is etched, etch profiles of a multijunction solar cell require considerable skill, and some luck. Nonuniform etching can distort the results, especially if the material has some defects. Also some layers may be completely depleted. Often, an ECV profiler gives the best information when the etching is partly done by the profiler and partly by applying selective etches *ex situ* (see Section 8.6.7). A (dilute) mixture of ammonia and hydrogen peroxide etches GaAs, but stops at GaInP and AlInP, whereas concentrated HCl etches GaInP and AlInP, but not GaAs. GaInP does not always etch in concentrated HCl, especially if the surface is wet, the HCl is not full strength, and/or if the GaInP surface has previously been in contact with a dilute HCl solution.

A window is provided in the ECV cell for shining light on the aqueous–semiconductor junction. The internal photocurrent ($QE_{Internal}$) from the junction at long wavelengths can be fit to the form

$$QE_{Internal} = \alpha(h\nu)L/[1 + \alpha(h\nu)L], \tag{8.27}$$

where $h\nu$ is the photon energy, and it is assumed that L is much longer than the depletion width, but less than the layer thickness.

The fit value for L will reflect the minority-carrier diffusion length when L is greater than the depletion width, but less than the thickness of the layer. For small $\alpha(h\nu)L$, the QE is proportional to $\alpha(h\nu)$, allowing an easy fit to determine E_g from $\alpha(h\nu) = A(h\nu - E_g)^{0.5}$.

Photoluminescence (PL) intensity is commonly used to test the quality of a material, but the intensity depends strongly on carrier concentration and surface recombination. Time-resolved PL measurements on double heterostructures (passivated layers) of different thicknesses can quantify both the minority-carrier lifetime and the interface-recombination velocity [129, 130].

When working with alloys such as GaInP, rocking-mode X-ray diffraction is very helpful toward confirming that the desired lattice constant (alloy composition) was realized (see Section 8.6.3.1).

8.7.2 Transmission-line Measurements

Once a device is made, in addition to the $I-V$ and QE measurements, characterization of the device contacts using a transmission-line measurement is useful toward diagnosing problems. Transmission lines [131] allow determination of the specific contact resistance (ρ_c) and the sheet resistance (R_s). The resistance, R, between any two pads as a function of x, the distance between the pads, is estimated by:

$$R = 2\rho_c/(w^2) + xR_s/w, \tag{8.28}$$

where w is the width of the transmission line and the dimension of the square pads. ρ_c and R_s are calculated from the intercept and slope of the line [131]. If the sheet resistance is large compared with the contact resistance, then the current across the semiconductor/pad interface is not uniform, and the estimated ρ_c can be refined to a more accurate value of the contact resistance, ρ_c', using the equation:

$$\rho_c{}' = \rho_c^2/(w^2R_s)\tanh^2\{w^2R_s/[\rho_c\tanh(w^2R_s/\rho_c)]\}. \tag{8.29}$$

8.7.3 I–V Measurements of Multijunction Cells

The chapter on measurements (Chapter 17) describes how to measure the light $I-V$ curve for a standard reference spectrum and for systematically varied spectra. Complete diagnosis of a multijunction cell requires characterization (an $I-V$ curve) of each of the active junctions to quantify their photocurrents and shunting. It would also be useful to establish the photovoltages and series resistances of each junction, but these are difficult for a two-terminal, series-connected cell. We describe here some alternative approaches for characterizing the individual junctions.

In some situations, it is possible to apply a contact between the series-connected junctions. The three-terminal configuration allows measurement of each junction of a two-junction cell. For a three-junction cell, a three-terminal measurement allows independent measurement of the top or bottom cells, but measurement of the middle cell by itself may require four connections. A primary advantage of the three-terminal approach is that it allows measurement of the photovoltage of each junction. The measured photocurrents should be adjusted for the change in junction area that occurred when the third connection was applied. Also, the increased perimeter area of the top cell sometimes affects the dark current.

Underlying junctions may be investigated by first chemically removing the upper junctions. In this case, the underlying junction will respond to a wider spectral range.

It is also useful to be able to characterize two-terminal multijunction cells. The shape of the $I-V$ curve of a series-connected multijunction cell is dominated by the characteristics of the junction that generates the smallest photocurrent (this concept can be understood by reviewing Figure 8.6 above). The $I-V$ curve for each junction can be measured by adjusting the spectrum so that junction has the smallest photocurrent. Using these measurements, the estimated individual-junction $I-V$ curves can be derived mathematically, as described in reference [132]. An example is given below. From the individual $I-V$ curves, one can calculate the $I-V$ curve for the multijunction cell under an arbitrary spectrum. For some samples, the J_{sc} values may be measured by reverse biasing the cell beyond the breakdown points of the junctions [133]. Finally, Kirchartz *et al.* [134] have recently described a method to obtain the $I-V$ curves of the individual subcells by combining electroluminescence and quantum efficiency measurements.

8.7.4 Evaluation of Morphological Defects

Careful examination of the devices with a microscope can identify many problems, especially when the device is forward biased so that it emits light, or when an OBIC (optical-beam-induced current) image is available. GaInP junctions emit red light that is usually visible to the naked eye. GaAs emission may be observed with an infrared (IR) imaging device. If the emission shows dark or bright spots, these can usually be correlated with a morphological defect, giving an explanation for the problem. Also, metal (e.g. a contact pad) that extends to the very edge of the pad may touch a layer that is nearby and short the device. This failure mode can sometimes be detected by microscopic examination.

8.7.5 Device Diagnosis

In general, a low J_{sc} may be evaluated from the energy dependence of the photocurrent loss. It is useful to measure and model the internal QE. The external QE measurements are described in Chapter 17. The internal QE is modeled (neglecting photon recycling) according to

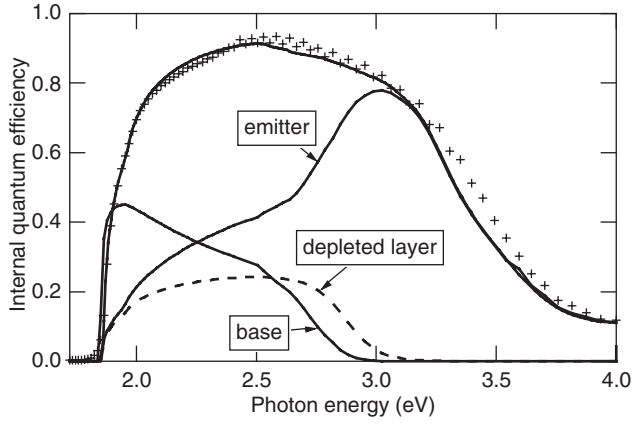


Figure 8.18 Measured (crosses) and modeled (lines) quantum efficiency of a GaInP solar cell. The contributions from the different layers of the solar cell are labeled and demonstrate how the emitter dominates the blue response, whereas the base dominates the red response. The relatively large contribution from the emitter is a result of the strong absorption of direct-gap materials

Equations (8.2–8.8) (Figure 8.18), and is determined experimentally from

$$\mathrm{QE_{Internal}} = \mathrm{QE_{external}}/(1 - \mathrm{Reflectivity}). \tag{8.30}$$

Accurate knowledge of the absorption coefficient is essential to successfully model the QE. The absorption coefficients of GaAs and GaInP were discussed above. High-quality samples may exhibit non-negligible photon recycling, implying that the omission of photon recycling in the model may need to be revisited. The application of Equations (8.2–8.8) is most useful when there are significant losses within the cell. Figure 8.19a compares the QE of a typical GaInP-cell (solid line) with what would be expected if there were no loss in the AlInP window (top curve) or

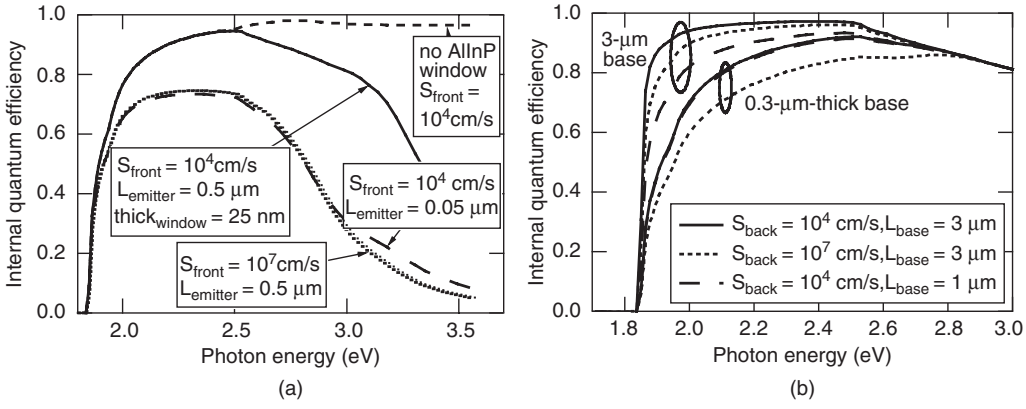


Figure 8.19 Modeled QE of GaInP cell. (a) The solid line, relative to the "no AlInP window" line, shows the effect of absorption of 25 nm of AlInP. The two lower curves show the degradation from an increased front-surface recombination velocity S_{front} or decreased emitter diffusion length L_{emitter}. (b) Comparison of a thin (0.3 μm base) and a thick (3 μm base) GaInP cell

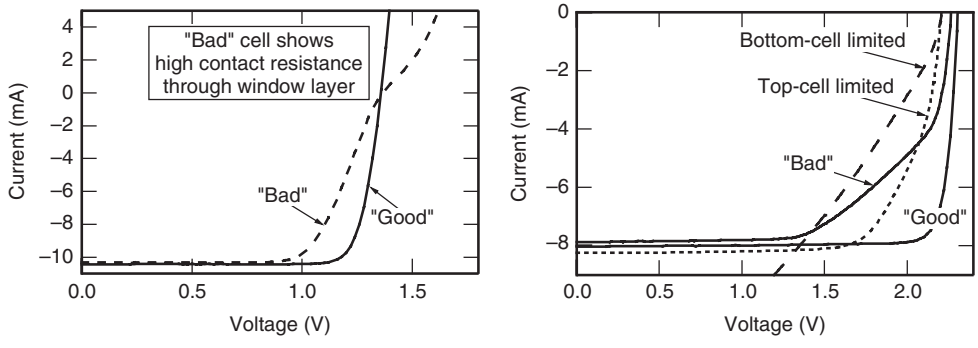


Figure 8.20 (a) Comparison of $I-V$ curves for a "good" GaInP cell and a GaInP cell with an extra junction caused by the AlInP window. Transmission-line measurements showed that the non-ohmic contact resistance was high through the window layer. (b) $I-V$ curves for good and bad GaInP/GaAs tandem cells. The "bad" curve shows a shunt which appears across the corner of the tandem $I-V$ curve; measurements of the same cell under bottom-cell-limited and top-cell-limited conditions (dashed curves) show that the shunt is in the bottom junction

poor collection in the emitter (bottom curves). Although window-absorption losses can easily be distinguished from emitter losses, the similarity of the two lower curves demonstrates the difficulty of distinguishing poor front-surface recombination from poor emitter-material quality. However, it is somewhat easier to differentiate poor base-material quality from poor rear passivation using a series of devices with variable base thickness. A cell with a thick base layer is more sensitive to the diffusion length in the base, whereas a cell with a thin base layer is relatively more sensitive to the quality of the back-surface field (see Figure 8.19b).

There are numerous reasons why the V_{OC} or FF may be degraded. Figure 8.20 illustrates two examples.

Extra junctions are more likely to be problems when working with Ge because III–V elements dope Ge and Ge dopes the III–V materials. The back of the Ge wafer must be etched before processing to avoid an extra junction at the back [135]. Accidental junctions in Ge are often highly shunted, with nearly ohmic $I-V$ characteristics. In this case, these are easiest to observe at high concentration because the V_{OC} of the Ge junction increases faster with photocurrent than that of the intended junction. Spurious junctions in the Ge may add, subtract, or both (in the case of back-to-back junctions) to the V_{OC}.

When a two-junction cell shows evidence of shunting, it is useful to determine which of the two junctions is shunted. This can be determined by measuring the light $I-V$ curve under two different spectra, a red-rich spectrum which reduces the photocurrent of the top junction, and a blue-rich spectrum which reduces the photocurrent of the bottom junction [132]. The example in Figure 8.20b shows a case for which the bottom cell is shunted. This sort of problem is often related to defects originating from particulates or poor wafer quality. Particulate exposure before or during growth is often a bigger problem for GaInP/GaAs cells than for single-junction cells.

8.8 RELIABILITY AND DEGRADATION

Successful deployment of multijunction cells in a concentrator system will require development of stable products. Reliability of concentrator systems is treated in Chapter 10. Using the relative

operating temperatures, current densities and other operating conditions of concentrator cells and light-emitting diodes (LEDs), it has been predicted that operational lifetimes longer than 10^5 hours (~34 years of operation) may be achievable for concentrator cells operating at about 1000 suns [136]. Because the multijunction cells are grown epitaxially at temperatures of the order 600 °C and above, the crystal stability at room temperature is excellent as long as the cell is not subject to illumination or electrical biasing.

However, under cell operating conditions, crystalline structure defects such as dislocations, stacking faults, and pinholes may compromise the stability of the cell. Dislocations act as localized recombination centers and their densities can increase with injected current. Dislocations densities higher than 10^4–10^6 cm^{-2} may cause reliability problems in cells of arbitrary size. Stacking faults and pinholes tend to act as localized current shunts. The shunt resistance also tends to decrease with time as current flows through it, especially for high shunt currents. Fortunately, densities of stacking faults and other current shunting defects in state-of-the-art lattice-matched cells are typically fairly low, of the order of 1 cm^{-2}. The probability of a cell containing one of these defects is proportional to the cell area, so the defects only become a problem for large-area cells. The large cell area may additionally contribute to the shorting of these defects by providing a large photocurrent, e.g. of the order of 5–10 A for a 1-cm^2 cell at 500 suns. With the aid of the front grids, which (by design) greatly enhance lateral current transport, the photocurrent may be effectively directed to the defect, especially when the cell is biased near the maximum power point or near open-circuit, so that there is a voltage across the defect. The defect area may be $(100\,\mu m)^2$ or less, implying a current density through the defect on the order of 10^5 A/cm^2 or greater, enough to destroy the cell.

8.9 FUTURE-GENERATION SOLAR CELLS

The GaInP/Ga(In)As/Ge cell is close to maturity, with champion concentrator cells achieving 40.1% efficiency at 236 suns [137].[1] The modeled efficiency for this structure from Figure 8.10 is 48% at 500 suns under the concentrator AM1.5 direct spectrum; historically, III–V based multijunction cells have achieved 80–90% of their theoretical efficiencies [111]. A very promising route to higher efficiencies is the development of structures with bandgap combinations more optimal than the {1.86, 1.39, 0.67 eV} bandgaps of the lattice-matched GaInP/Ga(In)As/Ge structure. For three-junction cells, Figure 8.10 shows the {1.86, 1.34, 0.93 eV} bandgap combination to be ideal with a corresponding efficiency of 53%, and {1.75, 1.18, 0.70 eV} to be a local maximum with 52.5% efficiency. For four-junction cells, the same type of calculation gives an optimal bandgap combination of {1.93, 1.44, 1.04, 0.70 eV} with a corresponding 57% efficiency, and {2.03, 1.56, 1.21, 0.92 eV} to be a local maximum with 56% efficiency. (The above numbers and the discussion below are for cells designed for terrestrial concentrators, operating at 500 suns at 300K.) The challenge is to actually make cells that not only have the desired bandgaps, but also realize the promised performance.

8.9.1 Lattice-mismatched GaInP/GaInAs/Ge Cell

For the AM0 spectrum, which is blue-rich compared with the terrestrial spectrum and thus sends more light to the GaInP top cell, an improvement in the efficiency of the GaInP/GaAs/Ge cell is predicted when the GaInP bandgap is increased. However, addition of aluminum to the GaInP cell has been shown to increase the bandgap, but not the efficiency, of the top cell because the J_{sc} was reduced by more than 10% while the V_{OC} increased only slightly, if at all. This effect is presumably due to the adverse effect of Al (and the associated oxygen contamination) on minority-carrier properties [138].

[1] As this chapter goes to press, 41.6% has been achieved.

One very rewarding approach to achieving a high-performing cell approaching the {1.75, 1.18, 0.70 eV} bandgap combination mentioned above has been to lower the bandgaps of the top and middle cells by increasing the indium content. An efficiency of 40.7% at 240 suns has been demonstrated [139] for a $Ga_{0.44}In_{0.56}P/Ga_{0.92}In_{0.08}As$/Ge cell with bandgaps of {1.80, 1.29, 0.67 eV}; this bandgap combination is marked in Figure 8.10 as a cone symbol. With these bandgaps, the top and middle cells have a 0.5% lattice constant misfit with the Ge substrate. In a different device design with an even higher misfit of 1.2%, an efficiency of 41.1% at 454 suns has been demonstrated [140] for a $Ga_{0.35}In_{0.65}P/Ga_{0.83}In_{0.17}As$/Ge cell with bandgaps of {1.67, 1.17, 0.67 eV}. The high efficiencies demonstrated for these structures are notable achievements, requiring mitigation of the harmful effects of dislocations due to the lattice mismatch.[2] The effects of lattice mismatch on the manufacturability and reliability of solar cells remain to be fully evaluated.

8.9.2 Inverted Lattice-mismatched GaInP/GaInAs/GaInAs (1.83, 1.34, 0.89 eV) Cell

The recent introduction of the inverted, lattice-mismatched GaInP/$Ga_{0.96}In_{0.04}As/Ga_{0.63}In_{0.37}As$ device structure, sometimes referred to as an inverted metamorphic multijunction (IMM) cell, has shown significant promise for achieving next-generation efficiencies [14, 141]. The basic concept is to obtain high-performance junctions with bandgaps near the optimal {1.86, 1.34, 0.93 eV} bandgap combination (see Section 8.5.6.2) by growing the junctions in increasing order of lattice mismatch to the substrate, thus minimizing the propagation of strain-induced defects through the device structure. Thus the lattice-matched GaInP junction (1.86 eV) is grown first, leaving an essentially strain- and defect-free surface for the growth of the slightly mismatched $Ga_{0.96}In_{0.04}As$ middle junction (1.34 eV); the highly mismatched $Ga_{0.63}In_{0.37}As$ junction (0.93 eV) is grown last, so that its strain-induced defects have little effect on the other junctions. Growing the junctions in this order, which is the inverse of the growth direction for conventional multijunction cells, requires removing the substrate in order to let the light enter the highest-bandgap subcell first, as is required in the standard stacked-subcell multijunction configuration. (Prior to the substrate removal, a "handle mount" foreign substrate is bonded to the other side of the cell for mechanical support. The handle is likely to be much less expensive than the original substrate, and can be selected to have properties such as flexibility that might be desirable for a given application.) Step-graded buffer layers are used between the mismatched junctions to relieve strain and confine dislocations away from the active regions of the junctions. The compositions of the buffer layers are chosen to be transparent to light which is intended for the underlying junctions. Figure 8.21a shows the as-grown device structure, while Figure 8.21b shows the finished device after mounting on the handle and removing the substrate.

The success of this approach depends on the tolerance of GaInAs to threading dislocations and on advances in our understanding of the growth of mismatched structures [142] which in turn has been aided by advances in high-resolution X-ray diffraction, transmission electron microscopy and *in situ* stress measurement techniques [143]. Extremely high efficiencies have been demonstrated with this approach: under concentration, this device demonstrated 40.8% at 326 suns [14], as shown in Figure 8.12. With continued development, efficiencies approaching 45% should be achievable. Furthermore, the extra processing steps required to remove the substrate may more than pay for themselves by enabling recycling or reuse of the substrate.

8.9.3 Other Lattice-matched Approaches

With the aim of avoiding the challenges of growing high-quality lattice-mismatched junctions, considerable effort has been invested toward the development of materials lattice-matched

[2] As this chapter goes to press, an efficiency of 42.3% has been reported for a 3-junction cell with a GaInAs junction grown on the back of the wafer.

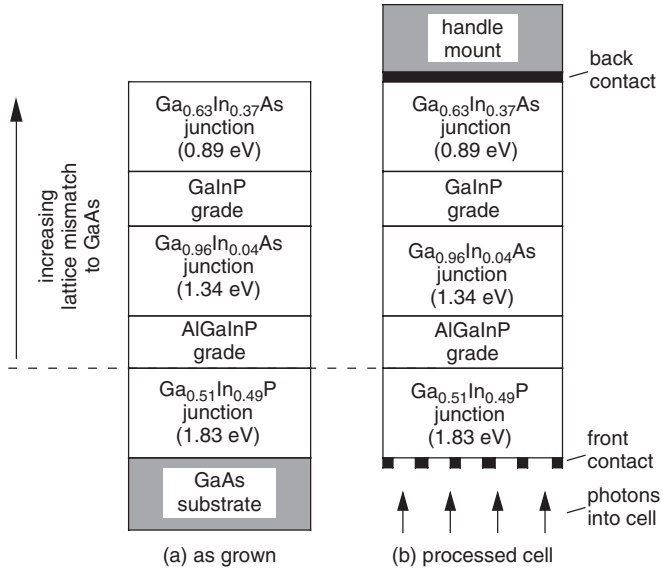


Figure 8.21 (a) Inverted metamorphic three-junction cell as grown. The tunnel-junction interconnects are not shown. The drawing is not to scale. (b) Finished device structure after back-contact deposition, handle mounting, substrate removal, and front-contact deposition

to GaAs and with a bandgap between that of GaAs and Ge. Unfortunately, this has proven difficult.

- $Ga_{1-x}In_xAs_{1-y}N_y$ can be grown lattice-matched ($x = 3y$) to GaAs with a bandgap of about 1 eV [144], but the minority-carrier diffusion length is too small to make a sufficiently high-quality junction [145–147].
- $ZnGeAs_2$ is somewhat difficult to grow (especially at low pressures) and can cause cross-contamination (e.g., Zn contamination of subsequent growth) [148].
- $Ga_{0.5}Tl_{0.5}P$ was reported to be lattice-matched to GaAs with a bandgap of about 0.9 eV [149], but a number of laboratories have been unable to duplicate the original report [150, 151].
- The bandgap of BGaInAs lattice-matched to GaAs has been pushed down to 1.35 eV, but so far, not to 1.0 eV [152]; this material also exhibits inferior material quality [153].

8.9.4 Mechanical Stacks

A high-efficiency result may also be obtained by a mechanical stack, which relaxes the need for lattice matching. The most probable candidates for this are GaInP/GaAs stacked over either GaInAsP (1 eV)/GaInAs (0.75 eV) or GaSb [154, 155]. The difficulties with implementing these stacks are associated with making the upper cell very transparent to the sub-bandgap light (use of a transparent GaAs substrate, nonconventional approach for the back contact, and a good AR coating on the back, as well as the front, of the upper cell) and with finding a way to mount both cells with simultaneous heat-sinking and electrical isolation, a much greater problem at 500–1000× than at 10–50×. An advantage of this approach is the decoupling of the photocurrents of the two pieces (assuming that four-terminal measurements are made), allowing for greater flexibility in the choice of materials and higher efficiency when the spectrum is changed.

For two-terminal operation, the mechanical stack may most easily be accomplished by bonding the two semiconductor materials directly [156]. Because wafer bonding is now routinely used for integrating many devices, techniques are available, and the wafer bond avoids the need to use a transparent substrate, avoids reflection losses, and removes the difficulty of heat sinking and electrically isolating the stacked cells. If a method for reusing the substrate can be made economical, wafer bonding also has the potential to reduce the substrate cost.

There are many more approaches to making a multijunction cell than can be discussed in this chapter. All approaches are variations on the structures shown in Figure 8.4. Wafer bonding of III–V multijunction cells to silicon would provide a lighter substrate (an advantage for space cells). A method for making a GaAs–Si bond with ohmic character between a GaAs cell and a silicon wafer has been reported [157]. Wafer bonding has not yet been developed for high-yield manufacturing of solar cells, but large-area wafer bonding is a possibility, given that 8-inch wafer-bonded silicon-on-insulator substrates are commercially available. The cost of these wafers is currently high (comparable to the cost of 4-inch Ge wafers), but may be reduced in the future.

8.9.5 Growth on Other Substrates

A silicon–III–V stack may also be made by growing III–V epitaxial layers directly on silicon. Growth of GaAs on Si has always been problematic because of the large lattice and thermal-expansion mismatch between the two materials. However, growth of a lattice-matched III–V alloy on silicon might be more similar to the growth of GaAs on Ge. High-quality GaAsNP lattice-matched to silicon substrates has been reported [158]. The use of Si as the 1-eV material in a multijunction higher-efficiency stack is compromised by the poor red QE of most Si cells, but the lower cost and weight of the silicon substrate might make it attractive even if higher efficiencies are not achieved.

High efficiencies have been achieved with two-junction (InP/GaInAs) structures on InP [18]. A three- or four-junction structure based on InP could, potentially, achieve higher efficiencies. This approach is limited by the availability of high-bandgap materials that are lattice-matched to InP and by the weight and current cost of the InP substrates.

8.9.6 Spectrum Splitting

Over the years, a number of groups have proposed to separate the light, directing each portion onto a solar cell optimized for that wavelength range as illustrated in Figure 8.4a. If four or five single-junction solar cells are used, the theoretical efficiency is quite high. However, the balance-of-system issues imply that this approach only makes sense for space missions for which high efficiency is essential, and the economics allow for the added cost of multiple substrates without requiring a high concentration ratio. Spectrum splitting can also be used with multijunction or mechanically stacked cells [159].

8.10 SUMMARY

Multijunction solar cells offer the potential for cell efficiencies higher than are possible for conventional single-junction cells. This potential has been realized by lattice-matched GaInP/GaInAs/Ge three-junction cells, which have demonstrated laboratory efficiencies of 40%, and which are commercially available as the cell of choice for space and high-concentration terrestrial applications. The implementation of lattice-mismatched alloy compositions offers a

viable path to next-generation cells with even higher efficiencies. Both the GaInP/GaInAs/Ge and the inverted GaInP/GaInAs/GaInAs lattice-mismatched approaches have already demonstrated efficiencies higher than the best lattice-matched efficiencies. With further development, including implementation of a suitable fourth junction, efficiencies approaching 50% appear possible.

REFERENCES

1. Olson JM, Gessert T, Al-Jassim MM, *18th IEEE Photovoltaic Specialists Conference* 552 (1985).
2. Fan JCC, Tsaur BY, Palm BJ, *16th IEEE Photovoltaic Specialists Conference* 692 (1982).
3. Olson JM, Kurtz SR, Kibbler AE, *18th IEEE Photovoltaic Specialists Conference* 777 (1988).
4. Kurtz SR, Olson JM, Kibbler A, *Solar Cells* **24**, 307 (1988).
5. Kurtz SR, Olson JM, Kibbler A, *Appl. Phys. Lett*. **57**, 1922 (1990).
6. Olson JM, Kurtz SR, Kibbler AE, Faine P, *Appl. Phys. Lett*. **56**, 623 (1990).
7. Kurtz SR, Faine P, Olson JM, *J. Appl. Phys*. **68**, 1890 (1990).
8. Bertness KA, Kurtz SR, Friedman DJ, Kibbler AE, Kramer C, Olson JM, *Appl. Phys. Lett*. **65**, 989 (1994).
9. Friedman DJ, Kurtz SR, Bertness KA, Kibbler AE, Kramer C, Olson JM, King DL, Hansen BR, Snyder JK, *Prog. Photovolt*. **3**, 47 (1995).
10. Kurtz SR, Bertness KA, Friedman DJ, Kibbler AE, Kramer C, Olson JM, *1st World Conference on PV Energy Conversion* 2108 (1994).
11. Green MA, Emery K, King DL, Igari S, Warta W, *Prog. Photovolt*. **9**, 49 (2001).
12. Green M, Emery K, Hishikawa Y, Warta W, *Prog. Photovolt*. **16**, 435 (2008).
13. Geisz JF, Kurtz SR, Wanlass MW, Ward JS, Duda A, Friedman DJ, Olson JM, McMahon WE, Moriarty T, Kiehl J, *Appl. Phys. Lett*. **91**, 023502 (2007).
14. Geisz JF, Friedman DJ, Ward JS, Duda A, Olavarria WJ, Moriarty TE, Kiehl JT, Romero MJ, Norman AG, Jones KM, *Appl. Phys. Lett*. **93**, 123505 (2008).
15. Swanson RM, *Prog. Photovolt. Res. Appl*. **8**, 93 (2000).
16. Henry CH, *J. Appl. Phys*. **51**, 4494 (1980).
17. Gee JM, *Solar Cells* **24**, 147 (1988).
18. Wanlass MW, Coutts TJ, Ward JS, Emery KA, Gessert TA, Osterwald CR, *22nd IEEE Photovoltaic Specialists Conference* 38 (1991).
19. Hovel HJ, *Solar Cells*. Willardson RK, Beer AC, (eds), *Semiconductors and Semimetals*, Academic Press, New York, (1975), vol. 11.
20. Nell ME, Barnett AM, IEEE Trans. *Electron Devices* **ED-34** 257 (1987).
21. Faine P, Kurtz SR, Riordan C, Olson JM, *Solar Cells* **31**, 259 (1991).
22. McMahon WE, Emery KE, Friedman DJ, Ottoson L, Young MS, Ward JS, Kramer CM, Kurtz S, Duda A, *Proceedings of the 31st IEEE Photovoltaic Specialists Conference* 715 (2005).
23. Nishioka K, Takamoto T, Agui T, Kaneiwa M, Uraoka Y, Fuyuki T, *Japanese Journal of Applied Physics Part 1 – Regular Papers & Short Notes* **43**, 882 (2004).
24. Kurtz SR, O'Neill MJ, *25th IEEE Photovoltaic Specialists Conference* 361 (1996).
25. Lockhart LB, King P, *J. Opt. Soc. Am*. **37**, 689 (1947).
26. Bader G, Ashrit PV, Girouard FE, Truong VV, *Applied Optics* **34**, 1684 (1995).
27. Friedman DJ, Kurtz SR, Bertness KA, Kibbler AE, Kramer C, Emery K, Field H, Olson JM, *12th NREL Photovoltaic Program Review* 521 (1993).
28. Palik ED, Addamiano A, "Zinc Sulfide", *Handbook of Optical Constants of Solids*, Palik ED, (ed.), Academic Press, San Diego, 1998, vol. I, p. 597.
29. Cotter TF, Thomas ME, Tropf WJ, Magnesium Fluoride (MgF_2), *Handbook of Optical Constants of Solids*, Palik ED, (ed.), Academic Press, San Diego, 1998, vol. II, p. 899.
30. Andreev VM, Grilikhes VA, Rumyantsev VD, *Photovoltaic conversion of concentrated sunlight*, John Wiley & Sons, Ltd, Chichester, Sussex, (1997).

31. Myers D, Emery K, Gueymard C, *ASME Journal of Solar Energy Engineering* **126**, 567 (2004).
32. Kurtz SR, Olson JM, Faine P, *Solar Cells* **30**, 501 (1991).
33. Gray JL, Schwartz RJ, Nasby RD, *IEEE International Electron Devices Conference* 510 (1982).
34. Rey-Stolle I, Algora C, *IEEE Trans. Electron Electron Devices* **49**, 1709 (2002).
35. Gessert TA, Coutts TJ, *J Vac Sci Technol A* **10**, 2013 (1992).
36. Emery K, Meusel M, Beckert R, Dimroth F, Bett A, Warta W, *28th IEEE Photovoltaic Specialists Conference* 1126 (2000).
37. Friedman DJ, *25th IEEE Photovoltaic Specialists Conference* 89 (1996).
38. Nishioka K, Takamoto T, Agui T, Kaneiwa M, Uraoka Y, Fuyuki T, *Solar Energy Materials and Solar Cells* **90**, 57 (2006).
39. Kinsey GS, Hebert P, Barbour KE, Krut DD, Cotal HL, Sherif RA, *Prog. Photovolt*. **16**, 503 (2008).
40. Fan JC, *Solar Cells* **17**, 309 (1986).
41. H. Cotal, Sherif R, *4th IEEE World Conference on Photovoltaic Energy Conversion* 845 (2006).
42. Zhao J, Wang A, Robinson SJ, Green MA, *Prog. Photovolt*. **2**, 221 (1994).
43. Lammasniemi J, Kazantsev AB, Jaakkola R, Aho R, Mäkelä T, Pessa M, Ovtchinnikov A, Asonen H, Robben A, Bogus K, *Second World Conference and Exhibition on Photovoltaic Energy Conversion* 1177 (1998).
44. Bertness KA, Friedman DJ, Kibbler AE, Kramer C, Kurtz SR, Olson JM, *12th NREL Photovoltaic Program Review Meeting* 100 (1993).
45. Olson JM, Kibbler A, Kurtz SR, *19th IEEE Photovoltaic Specialists Conference* 285 (1987).
46. Matthews JW, Blakeslee AE, *J. Cryst. Growth* **27**, 118 (1974).
47. O. Madelung, (ed.), *Semiconductors: Group IV Elements and III-V Compounds*; Springer-Verlag: Berlin, (1991).
48. Kudman I., Paff RJ, *J. Appl. Phys*. **43**, 3760 (1972).
49. Wie CR, *J. Appl. Phys*. **66**, 985 (1989).
50. Tanner BK, Miles SJ, Peterson G, Sacks RN, *Mater. Lett*. **7**, 239 (1988).
51. Delong MC, Mowbray DJ, Hogg RA, Skolnick MS, Williams JE, Meehan K, Kurtz SR, Olson JM, Schneider RP, Wu MC, Hopkinson M, *Appl. Phys. Lett*. **66**, 3185 (1995).
52. Gomyo A, Kobayashi K, Kawata S, Hino I, Suzuki T, Yuasa T, *J. Cryst. Growth* **77**, 367 (1986).
53. Gomyo A, Suzuki T, Kobayashi K, Kawata S, Hino I, Yuasa T, *Appl. Phys. Lett*. **50**, 673 (1987).
54. Kondow M, Kakibayashi H, Minagawa S, Inoue Y, Nishino T, Hamakawa Y, *J. Cryst. Growth* **93**, 412 (1988).
55. Kurimoto T., Hamada N, *Phys. Rev. B* **40**, 3889 (1989).
56. Capaz RB, Koiller B, *Phys Rev B – Condensed Matter* **47**, 4044 (1993).
57. Zhang Y, Mascarenhas A, Wang LW, *Phys. Rev. B* **63**, 201312 (2000).
58. Shurtleff JK, Lee RT, Fetzer CM, Stringfellow GB, *Appl. Phys. Lett*. **75**, 1914 (1999).
59. Olson JM, McMahon WE, Kurtz S, *IEEE 4th World Conference on Photovoltaic Energy Conversion*; IEEE: Waikoloa, HI, 787 (2006).
60. Mascarenhas A, Olson JM, *Physical Review B* **41**, 9947 (1990).
61. Mascarenhas A, Kurtz S, Kibbler A, Olson JM, *Physical Review Letters* **63**, 2108 (1989).
62. Luo JS, Olson JM, Bertness KA, Raikh ME, Tsiper EV, *J Vac Sci Technol B* **12**, 2552 (1994).
63. Luo JS, Olson JM, Kurtz SR, Arent DJ, Bertness KA, Raikh ME, Tsiper EV, *Phys Rev B – Condensed Matter* **51**, 7603 (1995).
64. Friedman DJ, Kurtz SR, Kibbler AE, Bertness KA, Kramer C, Matson R, Arent DJ, Olson JM, *Evolution of Surface and Thin Film Microstructure* 493 (1993).

65. Chernyak L, Osinsky A, Temkin H, Mintairov A, Malkina IG, Zvonkov BN, Safanov YN, *Appl. Phys. Lett*. **70**, 2425 (1997).
66. Friedman DJ, Zhu JG, Kibbler AE, Olson JM, Moreland J, *Appl. Phys. Lett*. **63**, 1774 (1993).
67. Friedman DJ, Horner GS, Kurtz SR, Bertness KA, Olson JM, Moreland J, *Appl. Phys. Lett*. **65**, 878 (1994).
68. Lee H, Klein MV, Olson JM, Hsieh KC, *Phys Rev B – Condensed Matter* **53**, 4015 (1996).
69. Schubert M, Gottschalch V, Herzinger CM, Yao H, Snyder PG, Woollam JA, *J. Appl. Phys*. **77**, 3416 (1995).
70. Kato H, Adachi S, Nakanishi H, Ohtsuka K, *Jpn. J. Appl. Phys. Pt 1* **33**, 186 (1994).
71. Kurtz SR, Olson JM, *19th IEEE Photovoltaic Specialists Conference* 823 (1987).
72. Iwamoto T, Mori K, Mizuta M, Kukimoto H, *J. Cryst. Growth* **68**, 27 (1984).
73. Ikeda M, Kaneko K, *J. Appl. Phys*. **66**, 5285 (1989).
74. Gomyo A, Hotta H, Hino I, Kawata S, Kobayashi K, Suzuki T, *Jpn. J. Appl. Phys*. **28**, L1330 (1989).
75. Kurtz SR, Olson JM, Goral JP, Kibbler A, Beck E, *J. Electron. Mater*. **19**, 825 (1990).
76. Goral JP, Kurtz SR, Olson JM, Kibbler A, *J. Electron. Mater*. **19**, 95 (1990).
77. Kurtz SR, Olson JM, Friedman DJ, Kibbler AE, Asher S, *J. Electron. Mater*. **23**, 431 (1994).
78. Watanabe MO, Ohba Y, *J. Appl. Phys*. **60**, 1032 (1986).
79. Hotta H, Hino I, Suzuki T, *J. Cryst. Growth* **93**, 618 (1988).
80. Scheffer F, Buchali F, Lindner A, Liu Q, Wiersch A, Prost W, *J Cryst Growth* **124**, 475 (1992).
81. Minagawa S, Ishitani Y, Tanaka T, Kawanaka S, *J Cryst Growth* **152**, 251 (1995).
82. Wang CJ, Wu JW, Chan SH, Chang CY, Sze SM, Feng MS, *Japanese Journal of Applied Physics Part 2 – Letters* **34**, L1107 (1995).
83. Malacky L, Kudela R, Morvic M, Cerniansky M, Peiner E, Wehmann HH, *Appl. Phys. Lett*. **69**, 1731 (1996).
84. Suzuki M, Ishikawa M, Itaya K, Nishikawa Y, Hatakoshi G, Kokubun Y, Nishizawa J, Oyama Y, *J Cryst Growth* **115**, 498 (1991).
85. Kurtz SR, Olson JM, Kibbler AE, Asher S, *Proc. of the InP and Related Materials Conf.* (1992).
86. Suzuki T, Gomyo A, Hino I, Kobayashi K, Kawata S, Iijima S, *Jpn. J. Appl. Phys*. **27**, L1549 (1988).
87. Nishikawa Y, Ishikawa M, Tsuburai Y, Kokubun Y, *J. Cryst. Growth* **100**, 63 (1990).
88. Kurtz SR, Olson JM, Bertness KA, Sinha K, McMahon B, Asher S, *25th IEEE Photovoltaic Specialists Conference* 37 (1996).
89. Dabkowski FP, Gavrilovic P, Meehan K, Stutius W, Williams JE, Shahid MA, Mahajan S, *Appl. Phys. Lett*. **52**, 2142 (1988).
90. Minagawa S, Kondow M, Yanagisawa H, Tanaka T, *J Cryst Growth* **118**, 425 (1992).
91. Hino I, Gomyo A, Kawata S, Kobayashi K, Suzuki T, *Inst. Phys. Conf. Ser.* **79**, 151 (1985).
92. Kondo M, Anayama C, Sekiguchi H, Tanahashi T, *J Cryst Growth* **141**, 1 (1994).
93. Bauhuis GJ, Hageman PR, Larsen PK, *J Cryst Growth* **191**, 313 (1998).
94. Stockman SA, Huang JW, Osentowski TD, Chui HC, Peanasky MJ, Maranowski SA, Grillot PN, Moll AJ, Chen CH, Kuo CP, Liang BW, *J. Electron. Mater*. **28**, 916 (1999).
95. Bertness KA, Kurtz SR, Asher SE, Reedy RC, *J Cryst Growth* **196**, 13 (1999).
96. Kibbler AE, Kurtz SR, Olson JM, *J. Cryst. Growth* **109**, 258 (1991).
97. Friedman DJ, Kibbler AE, Reedy R, *Appl. Phys. Lett*. **71**, 1095 (1997).
98. Ishitani Y, Minagawa S, Kita T, Nishino T, Yaguchi H, Shiraki Y, *J. Appl. Phys*. **80**, 4592 (1996).
99. Suguira H, Amano C, Yamamoto A, Yamaguchi M, *Jpn. J. Appl. Phys. Pt 1* **27**, 269 (1988).
100. Friedman DJ, Kurtz SR, Kibbler AE, Olson JM, *22nd IEEE Photovoltaic Specialists Conference* 358 (1991).

101. Rafat NH, Bedair SM, Sharps PR, Hills JS, Hancock JA, Timmons ML, *1st World Conference on Photovoltaic Energy Conversion* 1906 (1994).
102. Karam NH, King RR, Haddad M, Ermer JH, Yoon H, Cotal HL, Sudharsanan R, Eldredge JW, Edmondson K, Joslin DE, Krut DD, Takahashi M, Nishikawa W, Gillanders M, Granata J, Hebert P, Cavicchi BT, Lillington DR, *Solar Energ Mater Solar Cells* **66**, 453 (2001).
103. Chiang PK, Chu CL, Yeh YCM, Iles P, Chen G, Wei J, Tsung P, Olbinski J, Krogen J, Halbe S, Khemthong S, Ho F, *28th IEEE Photovoltaic Specialists Conference* 1002 (2000).
104. Kadoiwa K, Kato M, Motoda T, Ishida T, Fujii N, Hayafuji N, Tsugami M, Sonoda T, Takamiya S, Mitsui S, *J Cryst Growth* **145**, 147 (1994).
105. Takamoto T, Agui E, Ikeda E, Kurita H, *28th IEEE Photovoltaic Specialists Conference* 976 (2000).
106. Aspnes DE, Studna AA, *Phys. Rev. B* **27**, 985 (1983).
107. Kim CC, Garland JW, Raccah PM, *Phys Rev B – Condensed Matter* **47**, 1876 (1993).
108. Olson JM, Ahrenkiel RK, Dunlavy DJ, Keyes B, Kibbler AE, *Appl. Phys. Lett*. **55**, 1208 (1989).
109. Kurtz SR, Olson JM, Kibbler A, *21st IEEE Photovoltaic Specialists Conference* 138 (1990).
110. Oshea JJ, Reaves CM, Denbaars SP, Chin MA, Narayanamurti V, *Appl. Phys. Lett*. **69**, 3022 (1996).
111. Kurtz SR, Myers D, Olson JM, *26th IEEE Photovoltaic Specialists Conference* 875 (1997).
112. Friedman DJ, Olson JM, *Prog. Photovolt*. **9**, 179 (2001).
113. Lee JB, Kim I, Kwon HK, Choe BD, *Appl. Phys. Lett*. **62**, 1620 (1993).
114. Yoon IT, Han SY, Park HL, Kim TW, *J Phys Chem Solids* **62**, 607 (2001).
115. Tobin SP, Vernon SM, Bajgar C, Haven VE, Geoffroy LM, Sanfacon MM, Lillington DR, Hart RE, Matson RJ, *20th IEEE PVSC* 405 (1988).
116. Olson JM, McMahon WE, *2nd World Conf. on Photovoltaic Energy Conversion* (1998).
117. Pelosi C, Attolini G, Bocchi C, Franzosi P, Frigeri C, Berti M, Drigo AV, Romanato F, *J. Electron. Mater*. **24**, 1723 (1995).
118. Li Y, Salviati G, Bongers MMG, Lazzarini L, Nasi L, Giling LJ, *J. Cryst. Growth* **163**, 195 (1996).
119. Chen JC, Ristow ML, Cubbage JI, Werthen JG, *J. Electron. Mater*. **21**, 347 (1992).
120. McMahon WE, Olson JM, *Phys Rev B – Condensed Matter* **60**, 2480 (1999).
121. McMahon WE, Olson JM, *Phys Rev B – Condensed Matter* **60**, 15999 (1999).
122. Sze SM, *Physics of Semiconductor Devices*, John Wiley & Sons, Inc., New York, (1969).
123. Jung D, Parker CA, Ramdani J, Bedair SM, *J. Appl. Phys*. **74**, 2090 (1993).
124. Flemish JR, Jones KA, *J Electrochem Soc* **140**, 844 (1993).
125. Lothian JR, Kuo JM, Hobson WS, Lane E, Ren F, Pearton SJ, *J Vac Sci Technol B* **10**, 1061 (1992).
126. Lothian JR, Kuo JM, Ren F, Pearton SJ, *J. Electron. Mater*. **21**, 441 (1992).
127. Andersson BA, *Prog. Photovolt*. **8**, 61 (2000).
128. Blood P, *Semicon. Sci. Technol*. **1**, 7 (1986).
129. Ahrenkiel RK, *Solid State Electron* **35**, 239 (1992).
130. Ahrenkiel RK, Minority-Carrier Lifetime in III–V Semiconductors, *Minority Carriers in III–V Semiconductors: Physics and Applications*, Ahrenkiel RK, Lundstrom MS (eds), Academic Press Inc, San Diego, 1993, vol. 39, pp 39.
131. Berger HH, *J. Electrochem Society* **119**, 507 (1972).
132. Kurtz SR, Emery K, Olson JM, *1st World Conference on Photovoltaic Energy Conversion* 1733 (1994).
133. King DL, Hansen BR, Moore JM, Aiken DJ, *28th IEEE Photovoltaic Specialists Conference* 1197 (2000).
134. Kirchartz T, Rau U, Hermle M, Bett AW, Helbig A, Werne J.H, *Appl. Phys. Lett*. **92**, 123502 (2008).

135. Friedman DJ, Olson JM, Ward S, Moriarty T, Emery K, Kurtz S, Duda A, King RR, Cotal HL, Lillington DR, Ermer JH, Karam NH, *28th IEEE Photovoltaic Specialists Conference* 965 (2000).
136. Vazquez M, Algora C, Rey-Stolle I, Gonzalez JR, *Prog. Photovolt*. **15**, 477 (2007).
137. King RR, Law DC, Edmondson KM, Fetzer CM, Kinsey GS, Krut DD, Ermer JH, Sherif RA, Karam NH, *Proceedings of the 4th International Conference on Solar Concentrators (ICSC-4)* 5 (2007).
138. King RR, Karam NH, Ermer JH, Haddad M, Colter P, Isshiki T, Yoon H, Cotal HL, Joslin DE, Krut DD, Sudharsanan R, Edmondson K, Cavicchi BT, Lillington DR, *28th IEEE Photovoltaic Specialists Conference* 998 (2000).
139. King RR, Law DC, Edmondson KM, Fetzer CM, Kinsey GS, Yoon H, Sherif RA, Karam NH, *Appl. Phys. Lett*. **90**, 183516 (2007).
140. Guter W, Schöne J, Philipps SP, Steiner M, Siefer G, Wekkeli A, Welser E, Oliva E, Bett AW, Dimroth F, *Appl. Phys. Lett*. **94**, 223504 (2009).
141. Wanlass MW, Geisz JF, Kurtz S, Wehrer RJ, Wernsman B, Ahrenkiel SP, Ahrenkiel RK, Albin DS, Carapella JJ, Duda A, Moriarty T, *Proceedings of the 31st IEEE Photovoltaic Specialists Conference* 530 (2005).
142. Ahrenkiel SP, Wanlass MW, Carapella JJ, Gedvilas LM, Keyes BM, Ahrenkiel RK, Moutinho HR, *J. Electron. Mater*. **33**, 185 (2004).
143. Geisz JF, Levandor AX, Norman AG, Jones KM, Romero MJ, *J. Crystal Growth* **310**, 2339 (2008).
144. Kondow M, Uomi K, Niwa A, Kitatani T, Watahiki S, Yazawa Y, *Jpn. J. Appl. Phys*. **35**, 1273 (1996).
145. Geisz JF, Friedman DJ, Olson JM, Kurtz SR, Keyes BM, *J Cryst Growth* **195**, 401 (1998).
146. Kurtz SR, Allerman AA, Jones ED, Gee JM, Banas JJ, Hammons BE, *Appl. Phys. Lett*. **74**, 729 (1999).
147. Friedman DJ, Geisz JF, Kurtz SR, Olson JM, *J. Cryst. Growth* **195**, 409 (1998).
148. Timmons M, private communication.
149. Asahi H, Fushida M, Yamamoto K, Iwata K, Koh H, Asami K, Gonda S, Oe K, *J. Cryst. Growth* **175**, 1195 (1997).
150. Friedman DJ, Kurtz SR, Kibbler AE, *NREL/SNL PV Program Review Meeting* 401 (1998).
151. Antonell MJ, Abernathy CR, Sher A, Berding M, Van Schilfgaarde M, *InP and Related Materials* 444 (1997).
152. Geisz JF, Friedman DJ, Olson JM, Kurtz SR, Reedy RC, Swartzlander AB, Keyes BM, Norman AG, *Appl. Phys. Lett*. **76**, 1443 (2000).
153. Geisz JF, Friedman DJ, Kurtz S, *28th IEEE Photovoltaic Specialists Conference* 990 (2000).
154. Moto A, Tanaka S, Tanabe T, Takagishi S, *Solar Energ Mater Solar Cells* **66**, 585 (2001).
155. Fraas L, Daniels B, Huang HX, Avery J, Chu C, Iles P, Piszczor M, *28th IEEE Photovoltaic Specialists Conference* 1150 (2000).
156. Tanabe K, Morral AFi, Atwater HA, Aiken DJ, Wanlass MW, *Appl. Phys. Lett*. **89**, 102106 (2006).
157. Arokiaraj J, Okui H, Taguchi H, Soga T, Jimbo T, Umeno M, *Solar Energ Mater Solar Cells* **66**, 607 (2001).
158. Fujimoto Y, Yonezu H, Utsumi A, Momose K, Furukawa Y, *Appl. Phys. Lett*. **79**, 1306 (2001).
159. Barnett A, Honsberg C, Kirkpatrick D, Kurtz S, Moore D, Salzman D, Schwartz R, Gray J, Bowden S, Goossen K, Haney M, Aiken D, Wanlass M, Emery K, *4th World Conf. on Photovoltaic Energy Conversion* 2560 (2006).

9

Space Solar Cells and Arrays

Sheila Bailey[1] and Ryne Raffaelle[2]

[1]*NASA Glenn Research Center, Cleveland, OH, USA,* [2]*National Center for Photovoltaics, National Renewable Energy Lab, Golden, CO, USA*

9.1 THE HISTORY OF SPACE SOLAR CELLS

9.1.1 Vanguard 1 to Deep Space 1

In the mid 1950s, the development of single-crystal photovoltaic (PV) solar cells based on Si, as well as GaAs, had reached solar conversion efficiencies as high as 6% [1, 2]. By 1958, small-area silicon solar cells had reached an efficiency of 14% under terrestrial sunlight. These accomplishments opened the door to the possibility of utilizing solar power on board a spacecraft. On March 17, 1958 the world's first solar-powered satellite was launched, Vanguard 1 [3]. It carried two separate radio transmitters to transmit scientific and engineering data concerning, among other things, performance and lifetime of the 48 p/n silicon solar cells on its exterior. The battery powered transmitter operated for only 20 days, but the solar-cell-powered transmitter operated until 1964, at which time it is believed that the transmitter circuitry failed. Setting a record for satellite longevity, Vanguard 1 proved the merit of space solar cell power. The solar cells used on Vanguard 1 were fabricated by Hoffman Electronics for the US Army Signal Research and Development Laboratory at Fort Monmouth. In 1961, many of the staff from the silicon cell program at Fort Monmouth transferred to the National Aeronautics and Space Administration (NASA) Lewis Research Center (now Glenn Research Center) in Cleveland, Ohio. From that time to the present, the Photovoltaic Branch at Glenn has served as the research and development base for NASA's solar power needs. Impressed by the light weight and the reliability of photovoltaics, almost all communication satellites, military satellites, and scientific space probes have been solar-powered. It should be noted that the history presented here focuses on the United States space program. NASA was created in 1958; the Institute of Space and Astronautical Sciences (ISAS) and the National Space Development Agency (NASDA) in Japan were created in 1965 and 1969, respectively; the European Space Agency (ESA) was created in 1975 by the merger of the European Organization for the Development and Construction of Space Vehicle Launchers (ELDO) and the European Space Research Organization (ESRO),

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

which had begun in the early sixties. There are notable achievements in photovoltaics from these multiple agencies.

As the first PV devices were being created, there were corresponding theoretical predictions emerging that cited ~20% as the potential efficiency of Si and 26% for an optimum bandgap material (~1.5 eV) under terrestrial illumination [4]. In addition, it was not long before the concept of a tandem cell was proposed to enhance the overall efficiency. An optimized three-cell stack was soon to follow with a theoretical optimum efficiency of 37% [5]. Early solar cell research was focused on understanding and mitigating the factors that limited cell efficiency (e.g., minority carrier lifetime, surface-recombination velocity, series resistance, reflection of incident light, and non-ideal diode behavior).

The first satellites needed only a few watts to several hundred watts. They required power sources to be reliable and ideally to have a high specific power (W/kg), since early launch costs were ~$10 000/kg or more. The cost of the power system for these satellites was not of paramount importance since it was a small fraction of the satellite and the launch cost. The size of the array, and therefore the power, was limited for many early satellites owing to the body-mounted array design. Thus, there were multiple reasons to focus on higher-efficiency solar cells. Explorer I launched in 1958 discovered the Van Allen radiation belts, adding a new concern for space solar cells (i.e. electron and proton irradiation damage). The launch of Telstar in 1962 also ushered in a new era for space photovoltaics (i.e. terrestrial communications) [6]. Telstar's beginning of life (BOL) power was 14 W, but high radiation caused by the Starfish high-altitude nuclear weapon test reduced the power output [7]. This test caused a number of spacecraft to cease transmission. The lessons learned from Explorer I and Telstar prompted a surge of activity in radiation protection of space solar cells and prompted the use of *n*-on-*p* silicon semiconductor type (rather than *p*-on-*n*) for superior radiation resistance. Radiation damage studies at the Naval Research Laboratories in the 1960s provided much in the way of guidance to spacecraft designers in accounting for cell degradation [8].

As communication satellites evolved throughout the 1960s, so did their power requirements and thus the size and mass of the solar arrays. There were some early attempts to address the issue of mass by developing thin film cells such as CdS on CuS_2 heterojunction devices [9]. Unfortunately, their use was prohibited by severe degradation over time. CdTe cells were developed reaching efficiencies of ~7% [10]. However, the higher efficiency and stability of the silicon solar cells assured their preeminence in satellite power for the next three decades. Research on thin film cells for space applications, because of their higher specific power and projected lower costs, is still an area of intense research today.

In 1973, the largest solar array ever deployed up to that time was placed in low-Earth orbit (LEO) of Skylab 1 [11]. Skylab was powered by the Orbital Workshop array and the Apollo Telescope Mount array. The Orbital Workshop array had two deployable wings, each with 73 920 (2×4 cm) *n*-on-*p* Si cells that provided over 6 kW of power. Unfortunately, one of these wings was lost during launch. The Apollo Telescope Mount array had four wings with 123 120 (2×4 cm) cells and 41 040 (2×6 cm) cells providing over 10 kW of power. The 1970s also saw the first use of shallow-junction silicon cells for increased blue response and current output, the use of the back-surface field, the low–high junction theory for increased silicon cell voltage output, and the development of wraparound contacts for high-efficiency silicon (HES) cells to enable automated array assembly and to reduce costs.

In the 1980s, the gap between theoretical efficiencies and experimental efficiencies for silicon, gallium arsenide, and indium phosphide became almost nonexistent (Figure 9.1) [12]. New thin film cells of amorphous silicon and $CuInGaSe_2$ brought the possibility of higher thin film efficiencies and flexible, lightweight substrates that excited the space community. However, silicon

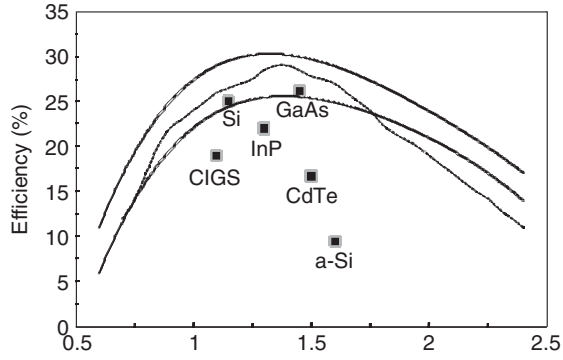


Figure 9.1 Comparison of measured cell efficiencies to the theoretical limits as a function of bandgap [12]

still provided the majority of the power for space and eventually the solar arrays for the International Space Station (ISS) (Figure 9.2).

The ISS has the largest PV power system ever present in space. It has 262 400 (8×8 cm) silicon solar cells with an average efficiency of 14.2% on eight Lockheed Martin US solar arrays (each $\sim 34 \times 12$ m), each producing $\sim$32 kW [13]. The array deployment was initiated in 1998 and completed in March 2009. The system generates about 110 kW of average power, which after battery charging, life support, and distribution, supplies $\sim$46 kW of continuous power for research experiments. The Russians also supply an additional 20 kW of solar power to ISS. The ISS is in low-Earth orbit or LEO, thus the array power output will degrade over time with the degradation primarily due to low-energy electrons trapped in the aforementioned Van Allen radiation belts.

Space solar cell research in the 1990s focused on the III–V and multijunction (MJ) solar cells that had higher efficiencies and were more tolerant of the radiation environment. Satellites continued to grow in both size and power requirements, and structures were designed to deploy large solar arrays. The mass and fuel penalty for attitude control of these large arrays continued to drive the space photovoltaics community to develop more efficient cells. Costs for satellite power systems remained at about a $1000/W.

The Deep Space 1 spacecraft, launched in October 1998, was the first spacecraft to rely upon SCARLET concentrator arrays to provide power for its ion propulsion engines [14]. Concentrator arrays use either refractive or reflective optics to direct concentrated sunlight onto a smaller active area of solar cells. Deep Space 1 had two such arrays and each was capable of producing 2.5 kW at 100 V (DC). The SCARLET arrays were developed by AEC-ABLE Engineering, Inc., under a program sponsored by the Ballistic Missile Defense Organization (BMDO). These arrays performed flawlessly under this inaugural demonstration. Unfortunately, a more recent example of the use of concentrator arrays did not fare as well. The concentrator solar arrays used on the Boeing 702 communication satellites exhibited faster than expected degradation. The problem was isolated to the concentrator reflector surfaces, which degraded after becoming coated due to something outgassing from the array.

The state-of-the-art (SOA) space solar cells available today are triple-junction III–V semiconductor cells. However, high-efficiency Si cells are still utilized in a number of space applications. Table 9.1 summarizes the SOA in space solar cells [15].

The production and deployment of terrestrial photovoltaics has been increasing at an almost astonishing $\sim$30% compounded annual growth rate (cagr) over the past decade. This has not been

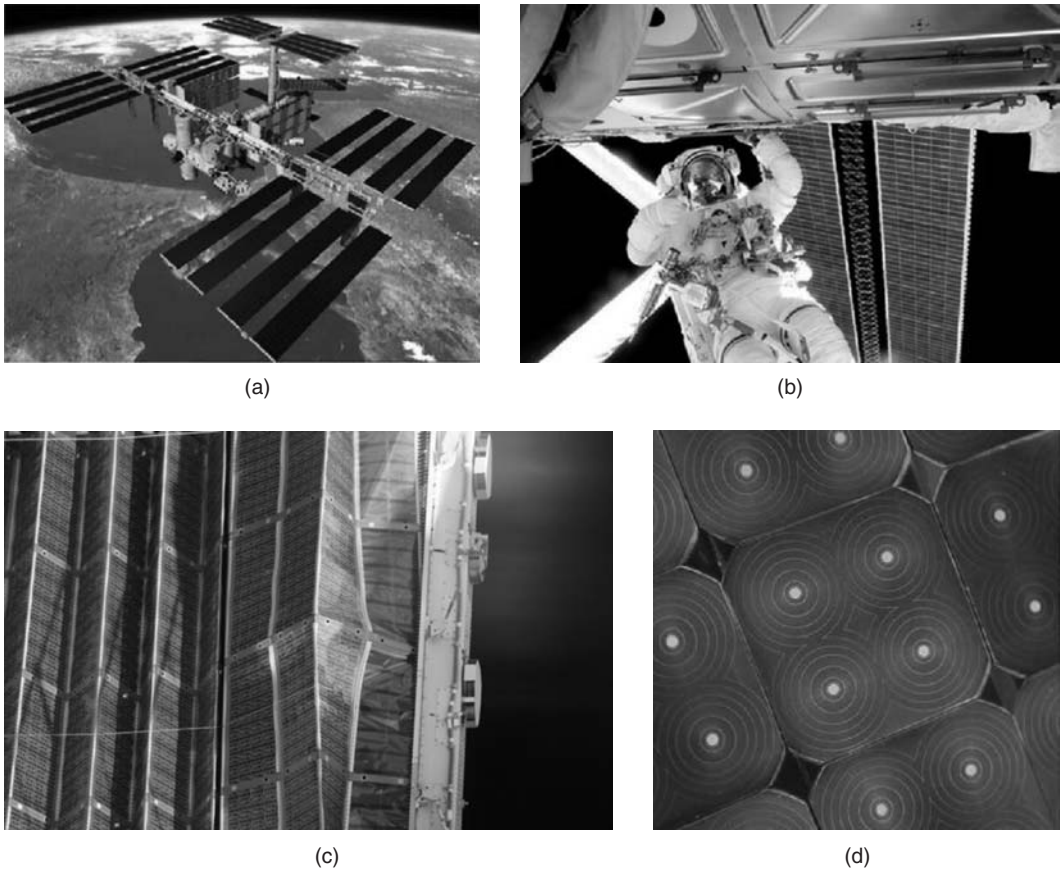


Figure 9.2 International Space Station space solar arrays: (a) upon completion; (b) up-close a single array next to an astronaut performing maintenance; (c) array deployment; and (d) close-up of the Si solar cells used in the arrays. (Pictures courtesy of NASA)

true within the space power world. The demand for space solar cells has been essentially flat, at ~27 geosynchronous, GEO, satellites per year. Today the vast majority of this demand is for conventional lattice matched triple junction III–V solar cells on Ge. Ironically, there are currently some very attractive terrestrial markets for these types of cells due to the emergence of large two-dimensional concentrator systems. The historic focus in the space world of increasingly higher-efficiency cells has merit in the terrestrial market. Solar cells that can essentially meet the dual-use requirements for both space and terrestrial markets could have some advantages. However with the size of terrestrial markets dwarfing that of space, this does raise some concern over whether or not the manufacturers will find it still attractive to meet the needs of space power.

There are several potential things that could change the current status quo and open the playing field to players emerging to serve new niche markets. The argument about spacecraft flying low efficient thin film cells is well documented in numerous trade studies. In 1994 when GaAs/Ge cells sold for ~$644/W compared with $432/W for Si cells, it was not cost-effective to consider thin film cells unless the cells were at least 13% air mass zero (AM0), efficient and there were low-cost space-qualified array designs as well. We now have thin film modules that have reached

Table 9.1 Summary of existing commercial space solar cell performance [15] obtained at AM0

Parameter	Silicon	High-efficiency silicon	Single-junction GaAs	Dual-junction III–V	Triple-junction III–V
Status	Obsolete	SOA	Obsolete	Obsolete	SOA
STC Efficiency (%)	12.7–14.8	16.6	19	22	29.9
STC operating voltage (V)	0.50	0.53	0.90	2.06	2.29
Cell weight (mg/cm^2)	13–50	13–50	80–100	80–100	80–100
Temperature coefficient at 28 °C	−0.55%/C	−0.35%/C	−0.21%/C	−0.25%/C	−0.19%/C
Cell thickness (μm)	50–200	76	140–175	140–175	140–175
Radiation tolerance	0.66–0.77		0.75	0.80	0.84
Absorptance	0.75		0.89	0.91	0.92

that efficiency. A recent mass and cost comparison of lightweight and rigid array structures showed that there could certainly be an advantage in cost and mass savings, particularly in high radiation orbits for thin film cells. The Boeing high-power solar array (HPSA) concept was used in this study and compared with a Boeing 702 rigid array structure ($750/W was assumed at array level for a standard 28% multijunction (MJ) rigid array with an array performance at 83 W/kg and $400/W was assumed for the baseline HPSA array up to 30 kW with 13% CIGS cells at 180 W/kg.

The US, through NASA and the Air Force, has dramatically reduced funding for the development of lightweight array structures using what would be considered thin films (a-Si, CIGS, CdTe) and there are currently no definitive plans for any deployment activities of such arrays. However with the new developments surrounding thin cells of high-efficiency multijunction III–V, this may provide resurgence in planning for very high specific power arrays. The US Air Force is holding a workshop shop devoted to looking at the integration issues of the new thin III–V cells for space arrays at the 2010 Space Power Workshop.

9.2 THE CHALLENGE FOR SPACE SOLAR CELLS

In 2002, a team from NASA, Department of Energy (DOE), and the Air Force Research Laboratory (AFRL) engineers reviewed the power technology needs of mid- and long-term proposed space science missions and assessed the adequacy of SOA solar cell and array technologies [17]. It concluded that low-cost roll-to-roll-produced thin film cells were yet to present a viable space power option. However, since that time progress has continued to be made. United Solar Ovonic (USO) has produced a-Si alloy-based TJ space solar cells on ~25-μm-thick polymer substrate using roll-to-roll and PECVD with an aperature area efficiency of 9.84% and initial specific power as high as 1200 W/kg [18].

The best space solar arrays are currently triple-junction III–V cells with an AM0 efficiency around 30%, and conventional arrays have reached panel level specific powers of over 100 W/kg. These arrays meet the needs of many near-Earth missions, but fail to meet some critical NASA Office of Space Science (OSS) mission needs in three ways. These are: (1) missions that utilize solar

electric propulsion (SEP) and require much higher specific power (150–200 W/kg); (2) missions that involve harsh environments (low solar intensity/low temperature (LILT), high solar intensities (HIHT), high radiation exposure, and Mars environments); and (3) Sun–Earth connection missions that require electrostatically clean arrays that do not allow the array voltage to contact and thereby distort the plasma environment of the array. The entire surface of an electrostatically clean array is maintained at approximately the same potential as the spacecraft structure. Examples of such arrays are those designed for the Solar Probe Plus (SPP) spacecraft. There are a variety of approaches to achieve this such as the use of conducting apertures and indium tin oxide coatings [19]. The majority of this work has been supported by the US AFRL, DOE, and other governmental agencies. Table 9.2 compares space solar power (SSP) drivers and current SOA technology.

Work is in progress at several US National Labs, universities, and solar cell companies to develop space solar cells with efficiencies over 30%. Approaches include the development of four-junction cells for conventional triple-junction solar cells, new materials such as InGaAsN, multiple quantum well and quantum dot devices, mechanically stacked devices and/or beam splitting to separate junctions (rainbow approach), and metamorphic (lattice-mismatched) devices. However, perhaps the most exciting new opportunity that has been presented to the space power community is the inverted metamorphic solar cell or IMM [20].

The IMM involves essentially inverting the growth or growing a conventional triple-junction essentially upside-down. An AlInGaP junction is grown on Ge, followed by an InGaAs junction, and then a series of grading layers is used to form a bottom InGaAs junctions. The cell is finalized by the removal of the Ge substrate. This approach enables a superior response to the solar

Table 9.2 Comparison of technology requirements with state-of-the-art space solar cells [17]

Technology	Driving missions	Mission application	State of the art
High-power arrays for solar electric propulsion (SEP)	Comet nucleus sample return, outer planet missions, Venus surface sample return, Mars sample return	>150 W/kg specific power Operate to 5 AU	50–100 W/kg Unknown LILT effect
Electrostatically clean arrays	Sun Earth connection missions	<120% of the cost of a conventional array	~300% of the cost of a conventional array
Mars arrays	Mars smart lander, Mars sample return, scout missions	26% efficiency > 180 sols @ 90% of full power	24% 90 sols @ 80% of full power
High-temperature solar arrays	Solar probe, sentinels	≥350 °C operation (higher temperatures reduce risk and enhance missions)	130 °C steady state; 260 °C for short periods
High-efficiency cells	All missions	30+%	27%
Low intensity low temperature (LILT), resistant arrays	Outer planet missions, SEP missions	No insidious reduction of power under LILT conditions	Uncertain behavior of MJ cells under LILT conditions
High-radiation missions	Europa and Jupiter missions	Radiation resistance with minimal weight and risk penalty	Thick cover glass

spectrum compared with conventional multijunction architecture. Soon after the first IMM's were demonstrated, both Spectrolab and Emcore began developing IMM cells for space application and are now producing CICs at over 32% 1 sun AMO efficiency [21–22].

The improved flexibility in device design and efficiency are obviously attractive, however one of the most intriguing aspects of the IMM approach is that the resulting cell is extremely thin, lightweight and flexible. These cells are typically about one-fifteenth the thickness of the conventional multijunction solar cell. These cells may enable a new class of extremely lightweight, high-efficiency, and flexible solar arrays for space applications.

The US Air Force Research Laboratory, Space Vehicles Directorate, has been leading the US efforts at the development of space solar arrays that can capitalize on the tremendous potential space power system mass savings afforded through the use of the IMM solar cell. Boeing has developed a flexible solar panel which is suitable for the implementation of 33% IMM which they call the Integrated Blanket/Interconnect System (IBIS) [23].

There are also a number of flat-folding "flexible" arrays that were produced by Lockheed Martin that have incorporated conventional high-efficiency 3J solar cells with specific power over 100 W/kg, such as the Mars Phoenix Lander solar array and the EOS-AM (Terra) solar array, which are obvious candidates for IMM integration [24].

JAXA has also developed a "Space Solar Sheet" which currently incorporates thin film 2J cells (InGaP/GaAs) with a 1 sun AM0 efficiency of 25% into a flexible laminate of either a transparent resin polymer sheet for LEO applications or thin coverglass for GEO applications. Sheets have been produced with a specific power of 500 W/kg and designs to achieve 700 W/kg are in production. Future versions of the space solar sheet will presumably take advantage of the IMM efficiency will undoubtedly push this specific power to even higher levels [25].

9.2.1 The Space Environment

All solar cells that are developed for use in space must take into consideration the unique aspects of the space environment. The spectral illumination that is available in space is not filtered by our atmosphere and thus is different from what is experienced on Earth. Space solar cells are designed and tested under an air mass zero (AM0) spectrum (see Figure 18.1). A more complete discussion of air mass (AM) can be found elsewhere in this handbook.

In the terrestrial PV world, cost is still the driver in PV development, and this has generated interest in several thin film material systems (i.e. amorphous silicon, $CuInGaSe_2$, CdTe). The smaller material costs and higher production potential for thin film arrays may well drive PV modules below current costs. The current National Photovoltaics Goal is the development of a 20% thin film cell. The problem is more complicated for space applications since these cells must also be developed on a lightweight flexible substrate that can withstand the rigors of the space environment. This suggests that a minimum of 15% AM0 efficiency will be needed to be competitive with current satellite power systems. The current benchmark for the space PV world are commercially available multijunction III–V cells of GaInP/GaAs/Ge described later in this chapter. Table 9.3 lists the current status of cell efficiencies measured under standard conditions (AM1.5 global) as well as extraterrestrial (AM0) spectra.

The major types of radiation damage in solar cells that are of interest to designers are associated with ionization and atomic displacement due to high-energy electrons and protons (although, low-energy protons can cause problems in the unshielded gap areas on the front of solar cells and on the unshielded back). The solar wind is the source of both electrons and protons and the Van Allen belts and thus the radiation environment does vary with solar activity. Solar flares can cause

Table 9.3 Measured global AM1.5 and estimated* AM0 efficiencies for small-area cells

Cells	Efficiency(%) Global AM1.5	Efficiency(%) AM0	Area (cm^2)	Manufacturer
c-Si	25.0	22*	4.00	UNSW
Poly-Si	20.4	18*	1.00	FhG-ISE
a-Si	9.5	8*	1.07	U. Neuchatel
GaAs	26.1	23*	1.00	Radboud U. Nijmegen
InP	22.1	19*	4.02	Spire
GaInP/GaAs/Ge	32.0	29*	3.99	Spectrolab
CIGS	19.4	17*	0.99	NREL
CdTe	16.7	15*	1.03	NREL
a-Si/uc-Si	11.7	10*	14.23	Kaneka

*Estimated AM0 efficiencies based on cells measured under standard conditions. The calculated efficiency used the ASTM E490-2000 reference spectrum and assumes that the fill factor does not change for the increased photocurrent. Quantum efficiencies corresponding to the table entries were used in the calculations.

intense fluxes of highly charged particles. Less important is radiation from cosmic rays originating outside our solar system.

Ionization effects can reduce the transmittance of the solar cell cover glasses through the development of color centers. Ionized electrons caused by the radiation become trapped by impurity atoms in the oxide to form stable defect complexes. Ionizing radiation is also a large detriment to the other materials associated with space solar arrays. It causes trapped charges to be created in silicon dioxide passivating layers that can lead to increased leakage currents. Ionizing radiation, including ultraviolet photons, is particularity bad for organic materials such as polymers used in array development as it can produce ions, free electrons, and free radicals that can dramatically change the optical, electrical, and mechanical properties of these materials.

The loss of energy of the high-energy protons and electrons due to interactions with electrons in a material accounts for a large fraction of the dissipated energy. In fact, these collisions are used to determine the penetration range for the electrons and protons in the 0.1–10 MeV range. However, it is the atomic displacements created by irradiation that are the major cause of degradation in space solar cells.

The displacement of an atom from a lattice site requires energy similar to that necessary to sublime an atom or to create a vacancy. The energy of sublimation for Si is 4.9 eV and for vacancy formation is 2.3 eV. The displacement of an atom requires the formation of a vacancy, an interstitial, and usually the creation of some phonons. Therefore, to create a displacement will require energy several times larger than that needed to create a vacancy.

The main importance of displacement defects due to irradiation is their effect on minority-carrier lifetime. The lifetime in the bulk *p*-type material of a Si solar cell is the major radiation-sensitive parameter. This was the basis for the switch from *p*-on-*n* to *n*-on-*p* Si solar cells in the 1960s [31]. The minority-carrier lifetime or diffusion length in an irradiated solar cell may be a function of excess or nonequilibrium minority carriers. This behavior is referred to as injection level dependence. This is usually associated with damage due to high-energy protons.

The primary radiation defects in Si are highly mobile. The radiation damage effects in Si are primarily due to the interaction of primary defects with themselves and with impurities in the material. Radiation damage in these cells can be mitigated to a certain extent by removing some of the damage before it becomes consolidated. Radiation-resistant Si cells use intrinsic gettering

to remove a part of the radiation damage while it is still mobile. These cells contain a relatively pure region near the surface or "denuded zone" with a gettering zone rich in oxygen deeper in the wafer away from the junction. Although this approach decreases the beginning-of-life (BOL) output, it increases the end-of-life (EOL) output. The cells are much more radiation-resistant, which can dramatically extend the mission lifetime.

Annealing of irradiated solar cells can be used to remove some of the damage, although not at temperatures that would be considered practical for space applications. Temperatures of nearly 400 °C are required for significant improvement in Si cells. However, there is some amount of ambient annealing of radiation damage that can occur. In space the damage and annealing processes are occurring simultaneously and are thus hard to quantify. However, in the lab, ambient annealing improvement of as much as 20% in the short-circuit current has been observed after 22 months.

The main method for mitigating radiation damage in space solar cells is to prevent damage by employing a coverglass. The coverglass not only stops the low-energy protons, but also slows down the high-energy particles. It can also serve to stop micro-meteors, act as an antireflection coating, provide resistance to charging, and even provide added thermal control to the spacecraft. In the 1970s, manufacturers began adding a nominal 5% cerium oxide to the coverglasses. This was shown to significantly improve the resistance of the glass to darkening from radiation or ultraviolet light [32]. The protection offered by the cerium also will improve the lifetime of the adhesives that are used to bond the cover glasses. The SOA coverglass is a drawn cerium-doped borosilicate glass. Research today is focused on improving the transmission of the glasses over a wider spectral range to accommodate the development of new MJ devices.

Current methods for calculating damage to solar cells are well documented in the *GaAs Solar Cell Radiation Handbook* (JPL 96-9). Recently the displacement damage dose (Dd) method has been developed to model radiation degradation. This method is currently being implemented in the SAVANT radiation degradation modeling computer program [33].

The bombardment of cells by charged particles can also lead to dangerously high voltages being established across solar arrays. These large voltages can lead to catastrophic electrostatic discharging events. This is especially true in the case of large arrays and poses significant problems for future utilization of large-area arrays on polymeric substrates. Much work has been done on the grounding and shielding of arrays to mitigate the effects of array charging [34]. Progress in addressing these effects has been made through the development of plasma contactors that ground the arrays to the space plasma.

Solar cell performance is also diminished over time due to neutral particle or micrometeor bombardment. These events account for approximately a 1% decrease in EOL space solar cells performance [35]. These events may also have a correlation to the initiation of discharging events discussed above. Space solar cells and arrays must also be equipped to contend with the plasma environment (radiation and charging). Removed from much of the shielding effects of the Earth's magnetic field, solar cells in space are continually bombarded with high-energy electrons and protons. The radiation damage caused by these particles will degrade solar cell performance and can dramatically limit spacecraft life. This is especially true for mid-Earth orbits (MEO, defined as $\sim$2000–12 000 km) in which cells must pass through the Van Allen radiation belts and thus get a much higher dose of radiation than would be experienced in low-Earth orbits (LEO, defined as $<$1000 km) or geosynchronous Earth orbits (GEO, defined as 35 780 km). LEO orbits vary in their radiation dose depending on their orientation, with, for example, polar orbits yielding a higher radiation dose than equatorial orbits. Figure 9.3 shows a comparison of equivalent fluence on a silicon solar cell in a variety of orbits. Figure 9.4 shows the dramatic decline in EOL power of cells in MEO orbit [36]. The degradation of cells in space due to radiation damage can be mitigated through the use of coverglasses at the expense of added mass to the spacecraft. Figure 9.5 shows the decline in power density as a function of time over 10 years in an 1853-km, 103° sun-synchronous orbit.

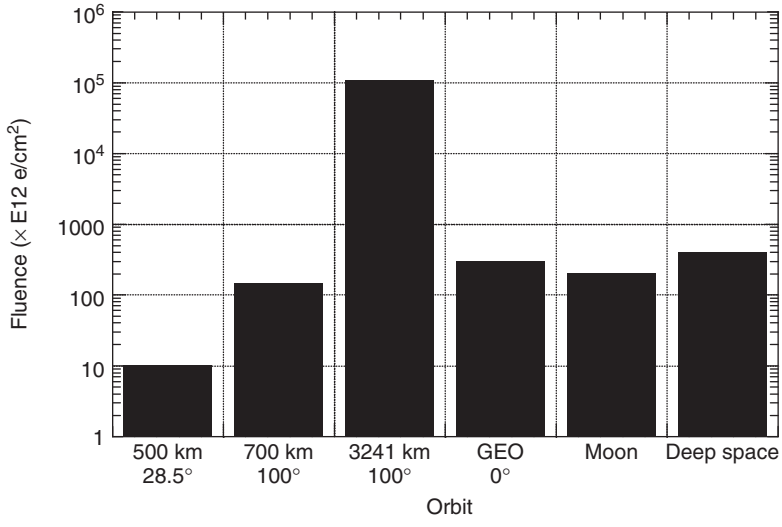


Figure 9.3 Equivalent 1-MeV electron fluence for a silicon solar cell in a variety of orbits (altitude in km) and inclinations (°)

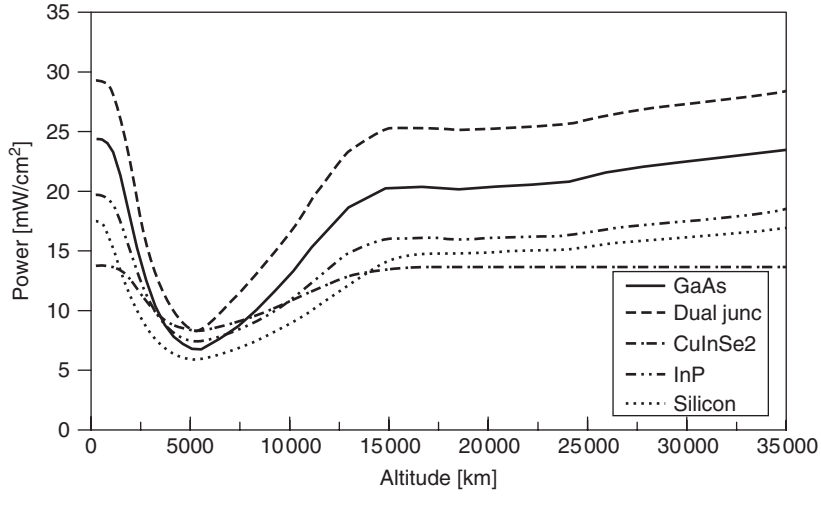


Figure 9.4 Solar cell power density as a function of altitude after 10 years in a 60° orbit with a coverglass thickness of 300 μm (GaAs: BOL 24.4 (mW/cm^2) [37]; Dual Junc – GaInP/GaAs/Ge [38]: BOL 29.3 (mW/cm^2); CuInSe2: BOL 14.4 (mW/cm^2) [39]; InP: BOL 19.7 (mW/cm^2) [40]; Si: BOL 17.5 (mW/cm^2) [41]). (Graph courtesy of Tom Morton, Ohio Aerospace Institute)

9.2.2 Thermal Environment

There is a considerable range of temperatures and intensities that may be encountered for the space use of photovoltaics. The temperature of a solar cell in space is largely determined by the intensity

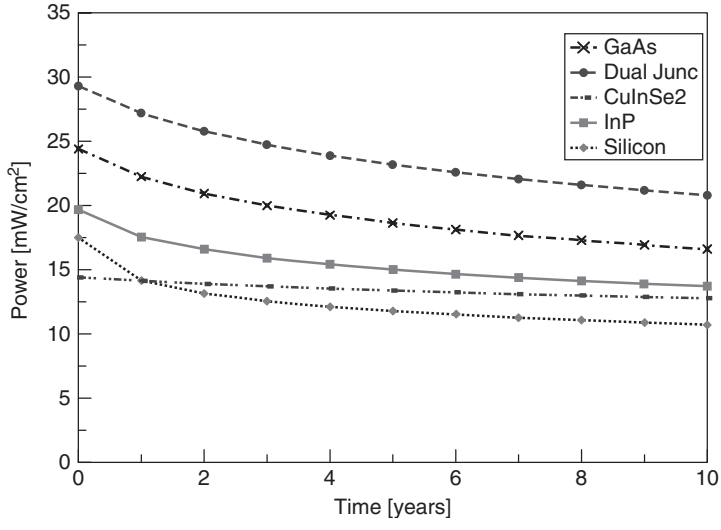


Figure 9.5 Solar cell power density as a function of time in a 1853-km altitude, 103° sun-synchronous orbit for the same cells as in Figure 9.6. (Graph courtesy of Tom Morton, Ohio Aerospace Institute)

and duration of its illumination [42]. In a typical LEO, such as the orbit of the ISS, the operating temperature of the silicon solar cells is 55 °C when sun tracking and −80 °C when in the longest eclipse. The average illuminated temperature at the orbit of Jupiter is −125 °C, whereas at the average orbit of Mercury the temperature is 140 °C. Similarly, the average intensity at the orbit of Jupiter is only 3% of the solar intensity at the Earth's radius, whereas the average intensity at Mercury is nearly double that at Earth. The solar intensity in the orbit around the Earth will vary seasonally because of the ellipticity of the Earth's orbit around the sun. Orbital characteristics are also a major source of thermal variation. Most spacecraft experience some amount of eclipse that will vary as the orbit precesses. This can result in very large and rapid temperature changes, as shown above for the space station cell. The temperature of a solar cell in space is also affected by the fraction of incident solar radiation returned from a planet or albedo. The average albedo from the Earth is 0.34, but can range anywhere from 0.03 (over forests) to 0.8 (over clouds) [7]. Typically, the range of temperatures experienced by solar cells in orbit about the Earth is from about 20 °C to as high as 85 °C.

An increase in solar cell temperature will cause a slight increase in the short-circuit current, but a significant decrease in the open-circuit voltage (Figure 9.6). Therefore, the overall effect is a reduction in power of a solar cell with an increase in temperature. It is typically less than 0.1%/ °C,

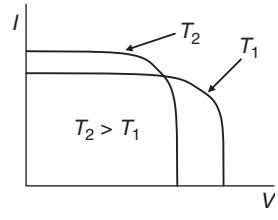


Figure 9.6 Effect of increasing temperature on solar cell $I-V$ photoresponse

Table 9.4 Theoretical normalized efficiency temperature coefficients [43]

Cell type	Temperature (°C)	η(28 °C)	$1/\eta$ dη/dT ($\times 10^{-3}$ °C^{-1})
Si (calc.)	27	0.247	−3.27
Ge (calc.)	27	0.106	−9.53
GaAs (calc.)	27	0.277	−2.4

Table 9.5 Measured temperature coefficients for various types of solar cells used in space [36]

Cell type	Temperature (°C)	η(28 °C)	$1/\eta$ dη/dT ($\times 10^{-3}$ °C^{-1})
Si	28–60	0.148	−4.60
Ge	20–80	0.090	−10.1
GaAs/Ge	20–120	0.174	−1.60
2-j GaAs/Ge	35–100	0.194	−2.85
InP	0–150	0.195	−1.59
a-Si	0–40	0.066	−1.11(nonlinear)
$CuInSe_2$	−40 to 80	0.087	−6.52

but can vary dramatically depending on cell type. This change coupled with rapid changes in temperature associated with eclipses can result in power surges that may be problematic.

The degradation of solar cell performance as a function of temperature is expressed in terms of temperature coefficients. There are several different temperature coefficients used to describe the thermal behavior of solar cells. These coefficients are generally expressed as the difference in a device parameter (i.e. I_{SC}, V_{OC}, I_{mp}, V_{mp}, or η) measured at a desired temperature and at a reference temperature (traditionally 28 °C, although the new International Space Organization (ISO) standard is 25 °C) divided by the difference in the two temperatures. Solar cell response to temperature is fairly linear for most cells over the range of −100 to 100 °C. Unfortunately, for cells of amorphous silicon or low-bandgap cells such as InGaAs, the response is only linear with temperature for small temperature differences. Another frequently used definition for the temperature coefficient is the normalized temperature coefficient. In the case of the efficiency it is defined as

$$\beta = \frac{1}{\eta}\frac{\mathrm{d}\eta}{\mathrm{d}T} \tag{9.1}$$

or the fractional change in efficiency with temperature. Theoretical values for the normalized efficiency temperature coefficients for Si, Ge, and GaAs are given in Table 9.4. Representative temperature coefficients for the various types of cells used in space are given in Table 9.5. In general, the temperature coefficient decreases in magnitude with the increasing bandgap, but is always negative, except in the case of a-Si, which can have a positive coefficient.

9.2.3 Solar Cell Calibration and Measurement

Calibration of solar cells for space is extremely important for satellite power system design. Accurate prediction of solar cell performance is critical to solar array sizing, often required to be within 1%. Calibration standards as a function of bandgap are required to perform simulated AM0 efficiency measurements on earth. The calibration standards are produced by evaluating various cell types

in space via the Shuttle or in the near future aboard the ISS. A less-costly means of developing standards is through the use of high-altitude aircraft or balloon flights. NASA Glenn Research Center solar cell calibration airplane facility has been in operation since 1963 with 531 flights to date [44]. The calibration includes real data to AM0.2 and uses the Langley plot method plus an ozone correction factor to extrapolate to AM0. Comparison of the AM0 calibration data indicates that there is good correlation with balloon- and Shuttle-flown solar cells within 1%.

Solar intensity is a function of the thickness of the atmosphere that the sunlight must pass through (AM). Plotting the logarithm of solar cell short-circuit current, proportional to solar intensity, as a function of AM permits extrapolation to an unmeasured AM and AM0 (Langley Plot Method). Early ground-based measurements were based on the change in atmosphere that the sun would pass through as it moves across the sky (i.e. more atmosphere at dawn and dusk and a minimum at solar noon). This is the same basic method that is used with an airplane, changing altitude to vary the AM.

Data analysis from early flights between 1963 and 1967 showed that the AM0 extrapolation was slightly lower than what was expected from radiometer data. This was found to be due to ozone absorption of sunlight in the upper atmosphere. A change in the data linearity was also noticed when the plane flew below the tropopause. This was later correlated with Mie scattering from particulate matter in the atmosphere and with absorption by moisture. These effects are primarily manifested in the higher-energy region of the solar spectrum. Calibration flights currently are performed above the troposphere and a correction for ozone absorption is used. Today calibration runs are performed with a Lear 25A jet housed at the NASA Glenn Research Center (Figure 9.7). It has flown 324 flights since 1984. The plane can fly up to ~15 km and gets above AM0.2 at 45 °N latitude. The data is now gathered using a continuous descent rather than remaining level over a range of altitudes.

The current Lear test setup has a 5:1 collimating tube in place of one of the original aircraft windows. This tube illuminates a 10.4-cm-diameter temperature-controlled plate. The tube angle can be adjusted to the sun angle. During descent $I-V$ curves for up to six cells, a pressure transducer, a

Figure 9.7 The NASA Glenn Research solar cell calibration aircraft. (Photo courtesy of NASA)

thermopile, and a temperature sensor are all measured. The cells are held at a constant temperature of ± 1 °C. A fiber optic connected to the test-plate is also connected to a spectroradiometer that can measure the solar spectrum from 250 to 2500 nm with 6-nm resolution, and a second spectrometer is used to measure the spectrum from 200 to 800 nm with 1-nm resolution. Both of these spectrometers are used to check for any spectral anomalies and to provide information on the ozone absorption.

There are currently several commercially available steady-state and pulse solar simulators that can simulate the sun's light in a variety of conditions (i.e. AM1.5, AM0). The NASA Glenn Research Center uses a Spectrolab X-25 Mark II xenon arc-lamp solar simulator. Steady-state solar simulators are generally used in laboratory or production environments for precision testing of PV devices. Solar simulators are also used on terrestrial, aerospace, and satellite products as a long-term simulated sunlight exposure system to test optical coatings, thermal control coatings, paints, and so on. Pulse simulators make it possible to test large solar cell assemblies and solar array blankets.

NASA Glenn uses a Spectrolab Spectrosum Large Area Pulsed Solar Simulator. It has a xenon arc lamp that is flashed to produce approximately 1-sun illumination with a nearly AM0 spectrum. The flash lasts approximately 2 ms, during which time the voltage across the cell is ramped and the resulting current is measured. At the same time the short-circuit current from a standard solar cell of a similar type is used to adjust the measured test sample current for the slightly changing illumination during the flash.

The standards for space solar cell calibration fall under the auspices of the ISO technical committee 20: Aircraft and Space Vehicle, sub committee 14: Space Systems and Operation. The working draft ISO 15 387 addresses the requirements for reference solar cells, the extraterrestrial solar spectral irradiance, and the testing conditions. Round-robin testing procedures, which rotate the cell measurements from agency to agency, involving NASA, CAST, and ESA for space solar cell calibration are currently under way. AIAA standards, AIAA S-111-2005, "Qualification and Quality Requirements for Space-Qualified Solar Cells," and AIAA S-112-2005, "Qualification and Quality Requirements for Space-Qualified Solar Panels," are used in the US. These standards are currently undergoing revision.

9.3 SILICON SOLAR CELLS

Silicon solar cells are the most mature of all space solar cell technologies and have been used on practically every near-Earth spacecraft since the beginning of the US space program. In the early 1960s, silicon solar cells were $\sim$11% efficient, relatively inexpensive, and well suited for the low-power (hundreds of watts) and short mission duration (3–5 years). The conversion efficiency of current "standard-technology" silicon ranges from around 12 to 15% under standard AM0 test conditions [26]. The lower efficiency cells are generally more resistant to radiation.

Cell efficiencies for any application should be adjusted for the array packing factor, radiation damage, ultraviolet degradation, assembly losses, and for corrections due to variations in intensity and temperature from standard conditions. At operating temperature, a silicon solar cell will degrade about 25% over 10 years in GEO orbit owing to charged particle irradiation damage [45]. The performance of these cells degrades significantly (often exceeding a 50% loss) in very-high-radiation environments such as experienced near Jupiter or in MEO. The relatively large temperature coefficient of silicon cells also results in large reductions in efficiency at high temperatures [36].

There have been many enhancements to silicon cells over the years to improve their efficiency and make them more suitable to space utilization. Textured front surfaces for better light absorption, extremely thin cells with back-surface reflectors for internal light trapping, and

passivated cell surfaces to reduce losses due to recombination effects are just a few examples, and are discussed in more detail in Chapters 7 and 8. Currently, high-efficiency Si (HES) cells approaching 17% AM0 efficiency in production lots are available from several producers. The advantage of the HES cell compared with a III–V cell lies in its relatively lower cost and lower material density. However, silicon solar cells are less tolerant of the radiation environment of space.

9.4 III–V SOLAR CELLS

The efficiency of space solar cells achieved dramatic improvements as the focus shifted from Si toward GaAs and III–V semiconductor systems. The 1.43 eV direct bandgap is nearly ideal for solar conversion (see Figure 9.1). By 1980 several types of III–V cells had been tested in space, with a 16% GaAs solar cell being developed by 1984 and a 18.5% efficient GaAs/Ge solar cell developed by 1989. It was also found that the GaAs cells had significantly better radiation resistance than Si cells. GaAs cells with efficiencies in excess of 80% of their theoretical maximum were routinely available commercially by 1998. Single-junction GaAs on Ge cells are currently commercially available with an AM0 efficiency of 19% and V_{OC} of 0.9 V. III–V solar cells for space applications are currently grown on Ge wafers because of the lower cost and higher mechanical strength over GaAs wafers.

Investigations on further efficiency improvement toward the end of the 20th century turned toward the development of multiple-junction cells and concentrator cells. Much of the development of "multijunction" GaAs-based photovoltaics was supported by a cooperative program funded by the Air Force Manufacturing Technology (ManTech) program and Space Vehicles Directorate, the Space Missile Center, and NASA [46]. This work resulted in the development of a "dual-junction" cell that incorporates a high-bandgap GaInP cell grown on a GaAs low-bandgap cell. The 1.85 eV GaInP converts short-wavelength photons and the GaAs converts the lower-energy photons. Commercially available dual-junction GaInP/GaAs cells have an AM0 efficiency of 22% with a V_{OC} of 2.06 V.

The III–V cells have now long held the record for efficiency from a monolithic cell structure. There are currently two main US manufacturers of lattice-matched III–V space cells and terrestrial concentrator cells, Spectrolab and Emcore. Both Spectrolab and Emcore have record cells of monolithic triple junction cells (GaInP/GaAs/Ge) slightly above 30% AM0. These cells, the ZTJ for Emcore and the XTJ for Spectrolab, are currently undergoing AIAA S-111 testing and are part of a ManTech program for the US government. Spire Semiconductor and Microlink in the US, Azur in the UK, and Sharp in Japan are also now producing MJ III–V with similar levels of performance to that of Emcore and Spectrolab.

Spectrolab, founded in 1956, is the supplier of more than half of the world's spacecraft solar cells, and has deployed over 25 000 triple-junction GaInP/GaAs/Ge solar cells in space. The Spectrolab Ultra-Triple-Junction solar cells have a current minimum average efficiency of 28.3%. Their newest design or XTJ solar cells have a minimum average efficiency of 29.9% (see Figure 9.8).

The new multijunction III–V cells have allowed a reduction in solar array size and mass over the previously used Si cells while maintaining comparable power levels. The alternative way to view the efficiency increase offered by the III–V cells is that they have increased available payload power over the value available from a comparably sized Si array. Scientists expect the majority of the commercial and military spacecraft launched in the next five years to continue to use conventional lattice-matched multijunction III–V technology.

At the Space Power Workshop in April 2009, both Spectrolab and Emcore showed progress on an inverted metamorphic (IMM) triple junction cell. Spectrolab measured 31.5%, with 20 of

Figure 9.8 Commercially available 30% GaInP/GaAs/Ge triple-junction cell (photograph courtesy of Spectrolab, Inc)

25 cells exceeding 30%. Both companies are addressing the not insubstantial issues of cell contact configuration, interconnect attachment, scaling-up in area, handling concerns, and bonding. The low-mass (0.012 g/cm^2), high-efficiency (>32.4% AM0) flexible cell is an excellent candidate for applications where high specific power (1 W/g), high areal power (400 W/m^2) and cell flexibility are required. These could include unmanned aerial vehicles as well as low-mass flexible satellite solar panels.

There are a number of on-going programs that are attempting to produce a fourth junction in a conventional lattice-matched III–V MJ approach to push the efficiencies into the low- to mid-30% range. However, these efforts are currently being somewhat overshadowed by the aforementioned excitement regarding IMM cell designs. The US Air Force initiated a program beginning in 2009 to develop a 37% AM0 IMM cell by a target date of 2011.

Theoretical estimates for a 4J IMM with optimum sub-gaps (2.09/1.58/1.21/0.93 eV) yields an efficiency of 41.3% [47]. Measurements on an actual 4J IMM using (1.92/1.42/1.02/0.70 eV) structure produced at Emcore have experimentally yielded 33.9% under a 1 sun AM0 simulated spectrum as measured at the NASA Glenn Research Center [48].

The IMM cells still require the use of a single-crystal substrate template that is later removed. If one wanted to preserve that substrate for further use, thus reducing cost, there are several ways that might be accomplished. MicroLink has developed proprietary technology based on epitaxial lift-off (ELO) to effectively peel a 4-inch wafer with epitaxial cells [49].

Microlink fabricated DJ ELO solar cells with efficiency of 28% at AM1.5 illumination of 1 sun. They have also succeeded in peeling an InP wafer as well. That opens the possibility for new cell configurations. There are other techniques for peeling cells using a “damaged” layer in the substrate. Years ago we had a “cleft” process to remove cells. It should be noted that you can put these IMM cells back on a substrate of your choice. A recent study looked at a TJ cell grown on thin (100 mm) Ge and a CIGS cell on thin metal substrates. Cells were tested against a set of performance indicators for an array and the thin TJ cells proved remarkably robust. It was noted

also that to fully benefit from the specific power (W/kg) of these arrays the TJ cells needed to have an alternative to the relative thick coverglass required by some orbits.

The conventional lattice-matched triple junction III–V approach (GaInP/GaInAs/Ge) has reached maturity. Even with the improvements offered by moving to four-junction and five-junction cells, we are already reaching a law of diminishing returns [50].

This is especially true as we balance the improvements in efficiency with additional junctions with the requirements that these cells maintain mission specific radiation hardness. The IMM technology is a breakthrough with regard to performance and is the most obvious path forward to higher cell efficiencies in space. However, this is qualified with the fact that some high-radiation environments will still require thick coverglass. However, not all space-based missions are in high-radiation environments. In this case one can benefit from the light weight and potential flexibility of an IMM cell. A good example of such an application is the power system of high-altitude long-endurance aircraft. They require high conversion efficiency coupled with power/weight ratio. For applications such as these the thin film IMM cell could be manufactured on a flexible lightweight substrate that can be integrated with the structural components of the aircraft. Thermal management and advanced arrays will be required to take full advantage of this potentially breakthrough technology.

Unfortunately, multijunction III–V cells are still expensive to produce. The added processing and array development involving IMMs may actually increase cost. The development of large-area arrays using these cells can become cost-prohibitive. One option to reduce the overall cost is to use the cells in solar concentrators, where a lens or mirror is used to decrease the required cell area. The previously mentioned SCARLET concentrator arrays used on Deep Space 1 used GaInP/GaAs high-efficiency dual-junction cells. A more advanced SCARLET-II array will build on the SCARLET technology with lower cost, easier fabrication, and simplified assembly and testing [14].

In addition to adding additional junctions or using metamorphic approaches to continue to enhance III–V solar cell performance, there are a number of other approaches being pursued. Efforts to enhance efficiency through new cell structures such as the quantum dot cells, intermediate band cells, up-and-down photon conversion, hot carriers in multi-quantum well cells, superlattice, and even mechanically stacked or lithographically integrated approaches are underway, albeit at low technology readiness levels at this point. Lastly there are efforts to produce a virtual single-crystal multijunction cell on a polycrystalline foil substrate or yet another way of producing high-efficiency yet thin lightweight and flexible devices.

9.4.1 Thin Film Solar Cells

One of the first thin film cells, Cu_2S/CdS, was developed for space applications. Reliability issues eliminated work on this particular cell type for both space and terrestrial considerations, even though AM1.5 efficiencies in excess of 10% were achieved. Thin film cells require substantially less material and thus lower mass, and promise the advantage of large-area, low-cost manufacturing. Until recently, the focus in space cells has been on efficiency rather than cost. In a several billion dollar spacecraft the solar cell cost is relatively small at even a thousand dollars per watt, which is approximately the current array cost. This has primarily been true for spacecraft with power needs from a few hundred watts to tens of kilowatts. However, deployment of a large Earth-orbiting space power system or some of the proposed SEP missions will require major advances in the PV array weight, stability in the space environment, efficiency, and ultimately the cost of production and

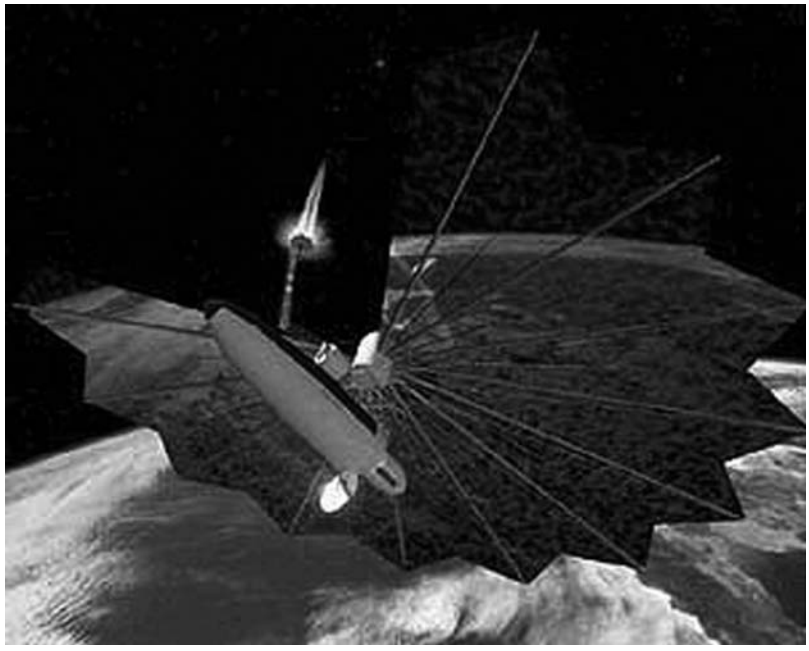


Figure 9.9 Proposed Mars solar electric propulsion vehicle. (Picture courtesy of NASA)

deployment of such arrays. The development of viable thin film arrays has become a necessity for a host of space missions [51]. Mission examples include ultralong duration balloons (e.g. Olympus), deep space SEP "tug" array, Mars surface power outpost, and Mars SEP Array (Figure 9.9). Solar electric propulsion missions are those in which the solar array is used to provide power for an electric propulsion system such as a Hall thruster or an ion thruster. These missions require large, lightweight, low-cost power systems. Studies have shown that the specific power or power per mass that will be required (i.e. 1 kW/kg) cannot be achieved with single-crystal technology [40]. The specific power required is almost 40 times what is presently available in commercial arrays. While high-efficiency ultra-lightweight arrays are not likely to become commercially available anytime soon, advances in thin film photovoltaics may still impact other space technologies (i.e. thin film integrated power supplies) and thus support a broad range of future missions.

Lighter power generation will allow more mass to be allocated to the balance-of-spacecraft (i.e. more payload). In addition, less expensive power generation will allow missions with smaller budgets and/or the allocation of funds to the balance-of-spacecraft. This is an essential attribute in enabling such missions as the Mars Outpost SEP Tug. An example of the benefits of thin film PV arrays for the now-canceled ST4/Champollion indicates a $50 million launch cost saving and 30% mass margin increase when thin film solar array power generation was combined with advanced electric propulsion. A parametric assessment showed similar advantages for other solar system missions (e.g. main belt asteroid tour, Mars SEP vehicle, Jupiter orbiter, Venus orbiter, lunar surface power system) [40].

Much of the original development of thin film PV arrays was performed with the terrestrial marketplace in mind. This has been a tremendous benefit to researchers hoping to develop such arrays for space. Features such as cell efficiency, material stability and compatibility, and low-cost and scalable manufacturing techniques are important to both environments. However, many key array aspects necessary for space utilization are not important for terrestrial use and thus have

not experienced a similar progress. Features such as radiation tolerance, air mass zero (AM0) performance, use of lightweight flexible substrates, stowed volume and lightweight space deployment mechanisms must be developed before a viable space array can become a reality. Unfortunately, the costs associated with developing these features, along with the subsequent space qualification studies mitigate the savings of using a thin film array for space, and thus have inhibited their application.

Previous efforts by NASA and the US Air Force have investigated the issues associated with the development of thin film arrays for space. Copper indium gallium diselenide (CIGS), cadmium telluride (CdTe), and amorphous silicon (a-Si) thin film materials appear to have a good chance of meeting several proposed space power requirements [52]. Reasonably efficient (~10% AM1.5) large-area flexible blankets using a-Si triple junction technology have already been manufactured by USO as previously mentioned.

Table 9.6 summarizes the thin film solar cell technologies that are currently available. Further details on a-Si, CIGS, and CdTe solar cells can be found elsewhere in this handbook. Several device structures offer specific power exceeding 1000 W/kg, but these values only include the device and substrate, not the entire module and array. This table shows the importance of lighter or thinner substrates in achieving higher specific power. Note that multijunction cells (such as the a-Si triple cells) whose thicknesses and bandgaps have been optimized for the terrestrial AM1.5 spectra should be reoptimized for AM0 since the distribution of photons between top, middle and bottom cells will be different. This changes the current matching. Reoptimization is not required for single-junction thin film devices. References [39] and [40] showed no substrate dependence for a-Si devices; that is, the same efficiency resulted between deposition on thin (10−25 μm) or thick (125 μm) stainless steel or between stainless steel or Kapton. This is good news for obtaining lighter thin film modules with higher specific power. In contrast, efficiency of Cu(InGa)S devices decreased from 10.4 to 4.1% as the stainless steel substrate decreased in thickness from 128 to 20 μm [53]. An 11% efficiency for CdTe on 10 μm polyimide was reported, but is unpublished [54].

Development of other wide-bandgap thin film materials to be used in conjunction with CIGS to produce a dual-junction device is underway. As has already been demonstrated in III–V cells for space use, a substantial increase over a single-junction device efficiency is possible with a dual-junction device. NASA and the National Renewable Energy Lab (NREL) have both initiated dual-junction CIS-based thin film device programs. The use of Ga to widen the bandgap of CIGS and thus improve the efficiency is well established [58]. The substitution of S for Se also appears to be attractive as a top cell material. In particular, AM0 cell efficiencies 8.8% have been measured for $CuIn_{0.7}\ Ga_{0.3}S_2$ (E_g1.55 eV) thin film devices on flexible stainless steel substrates [60].

The majority of thin film devices developed for terrestrial applications have been on heavy substrates such as glass. However, progress is being made in reducing substrate mass through the use of thin metal foils and lightweight flexible polyimide or plastic substrates [63]. The use

Table 9.6 Thin film solar cell mass specific power [17]

Cell Type	Thickness (cell + substrate) (μm)	Nominal AM1.5 efficiency (%)	Nominal cell specific power (W/kg)
a-Si/polyimide	50	5.0	700
a-Si/stainless	127	7.0	70
CdTe/glass	5000	7.0	6
CIS/glass	3000	8.0	11
a-Si/glass	3000	5.0	7
CIS/polyimide	50	7.0	985

of such plastic substrates as Upilex or Kapton puts a slight restriction on the processing temperatures. This of course can be obviated by the use of metal foil if one is willing to accept the mass penalty.

In addition to cost and weight savings for spacecraft, thin film solar cells have potential for improved radiation resistance relative to single-crystal cells, possibly extending mission lifetimes. For example, after a dose of 10^{16} 1 MeV *electrons*/cm^2, the maximum power generated by a GaAs cell can decrease to less than half of its BOL value. By contrast, after a dose of 10^{13}/cm^2 10 MeV *protons* (which would degrade GaAs cell power performance to less than 50% of BOL), the power generated by CIS cells has been shown to retain more than 85% of its BOL value [48]. In addition a recent study funded by the Italian Space Agency used CIGS cells that were 12% almost stable over a wide range of radiation fluences [65]. The cells used in the aforementioned study were acquired from an Italian University and also from the production line of Global Solar Energy.

In addition to thin film cell development, there is the problem of making thin film arrays for space. The flexibility of a thin film cell on a polymeric substrate is a great advantage when it comes to stowability and deployment; however, it must be rigidly supported after deployment. Work is still required to be done to determine how best to deploy and maintain thin film arrays, regardless of the cell technology. The same issues concerning stability in the space environment that were addressed for arrays with crystalline cells must now be addressed for thin film arrays. As outlined in US Government Military Standard 1540C, the array must pass a number of qualification tests, including those for integrity and performance after exposure to elevated temperatures, radiation (particularly electrons and protons), thermal cycling, vibration and mechanical stress, and atomic oxygen.

The US Air Force has led a large, multidisciplinary team under the auspices of the Air Force Dual-Use Science and Technology program to develop a functional thin film array for space [66]. That program has the specific goals of demonstrating:

1. a stabilized 10–15% AM0 efficient thin film submodule;
2. submodule and module electrical architectures including bypass and blocking diode technology;
3. submodule and module mechanical interface architectures, module strength, and structural support requirements;
4. array support structure design for an array with a wide range of power levels;
5. space environment and thermal control protection/qualification standards for thin film arrays.

This effort will culminate in the design of a 1-kW LEO and 20-kW GEO thin film solar array.

9.5 SPACE SOLAR ARRAYS

Solar array designs have undergone a steady evolution since the Vanguard 1 satellite. Early satellites used silicon solar cells on honeycomb panels that were body-mounted to the spacecraft. Early space solar arrays only produced a few hundred watts of power. However, satellites today require low-mass solar arrays that produce several kilowatts of power. Several new solar array structures have been developed over the past 40 years to improve the array specific power and reduce the stowed volume during launch.

The most important characteristics of solar arrays required for space applications are

- high specific power (W/kg)
- low stowed volume (W/m^3)

- low cost ($/W)
- high reliability.

In addition, several proposed space missions have put other constraints on the solar arrays. Several proposed Earth-orbiting missions designed to study the sun require "electrostatically clean" arrays. Inner planetary missions and missions to study the sun within a few solar radii require solar arrays capable of withstanding temperatures above 450 °C and functioning at high solar intensities (HIHT). Outer planetary missions require solar arrays that can function at low solar intensities and low temperatures (LILT). In addition to the near-sun missions, missions to Jupiter and its moons also require solar arrays that can withstand high-radiation levels.

The EOL power generated by an array is impacted in a variety of ways. Radiation damage will result in an 8% loss of BOL power/m^2 after 5×10^{14} 1 MeV electrons (i.e. typical EOL fluence in geosynchronous orbit). The temperature correction due to the operation of the solar cell at 75 °C rather than at the 25 °C test conditions will reduce the power/m^2 by ∼9%. Degradation due to UV exposure is around 1.7%. Loss in power over time due to micrometeors and ordinary surface contamination are each around 1% [34].

The solar arrays presently in use can be classified into six categories:

- body-mounted arrays
- rigid panel planar arrays
- flexible panel array
- flexible roll-out arrays
- concentrator arrays
- high-temperature/intensity arrays
- electrostatically clean arrays.

A summary of the important typical characteristics of these arrays is given in Table 9.7.

9.5.1 Body-mounted Arrays

Body-mounted arrays are preferred for small satellites that only need a few hundred watts. Early spherical satellites and spin-stabilized cylindrical satellites used body-mounted arrays of silicon solar cells on the honeycomb panels. This type of array is simple and has proven to be extremely reliable. One of the limitations of this type of array is that it puts a constraint on the direction the spacecraft must point. This type of array is still used on smaller spacecraft and spin-stabilized spacecraft. The Mars Pathfinder Sojourner rover, the first rover used on Mars used body-mounted solar arrays with high-efficiency MJ III–V solar cells (Figure 9.10).

Table 9.7 Space solar array mass specific power, cost, and inverse areal specific power [17]

Technology	Specific power W/kg (BOL) @ cell efficiency	Cost $K/W	Area per power (m^2/kW)
High-efficiency silicon (HES) rigid panel	58.5 @ 19%	0.5–1.5	4.45
HES flexible array	114 @ 19%	1.0–2.0	5.12
Triple-junction (TJ) GaAs Rigid	70 @ 26.8%	0.5–1.5	3.12
TJ GaAs Ultraflex	115 @ 26.8%	1.0–2.0	3.62
CIGS thin film	275 @ 11%	0.1–0.3	7.37
Amorphous-Si MJ/thin film	353 @ 14%	0.05–0.3	5.73

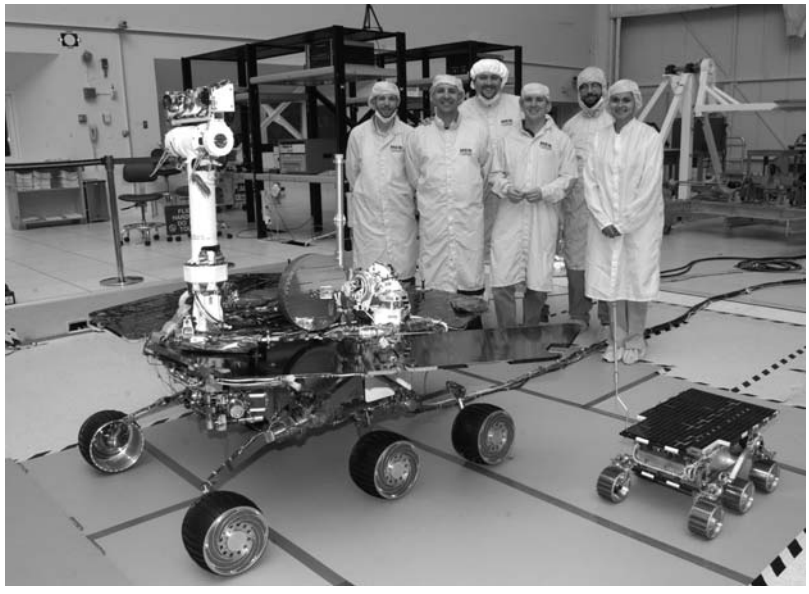


Figure 9.10 Body-mounted array on the Mars rovers. (Picture courtesy of NASA JPL)

Subsequently two other rovers have been used on Mars which also utilized body-mounted arrays which generate about 140 W for up to four hours per Martian day (sol). Spirit successfully landed on Mars on January 4, 2004, three weeks before its twin, Opportunity, landed on the other side of the planet. Its name was chosen through a NASA sponsored student essay contest. Spirit completed its planned 90-sol mission and aided by "cleaning events" or wind storms that blew the solar arrays clean. This resulted in higher power from its solar panels than predicted, it went on to function effectively over twenty times longer than NASA planners expected, logging about 10 km of driving (the mission was originally planned for about 1 km). The scientific results from the first phase of the mission were published in a special issue of the journal *Science* [67]. On May 1, 2009 (5 years, 3 months, 27 Earth days after landing; 21.6 times the planned mission duration), Spirit became stuck in soft soil.

9.5.2 Rigid Panel Planar Arrays

Rigid panel arrays have been used on many spacecraft, requiring several hundred watts to many tens of kilowatts of power. They consist of rigid honeycomb core panels that are hinged such that they can be folded against the side of the spacecraft during launch (Figure 9.11). Each panel is rigid and quite strong, but can add considerably to the overall weight of the array. There has been much development recently on panels of materials other than aluminum (i.e. graphite/epoxy sheets and ribbons). Hybrid panels with aluminum honeycombs and epoxy/glass face sheets have also been developed. The folded arrays are deployed by means of pyrotechnic, paraffin, or knife blade actuators and damper-controlled springs.

The BOL power density of the rigid panel array is extremely dependent on the type of solar cell used. BOL power densities range from 35 to 65 W/kg for silicon cells and from 45 to 75 W/kg for GaAs/Ge cells. The panel assembly of a rigid array accounts for 75–80% of the total mass, with the stowed and deployment structure making up the balance [17]. The Tropical Rainfall

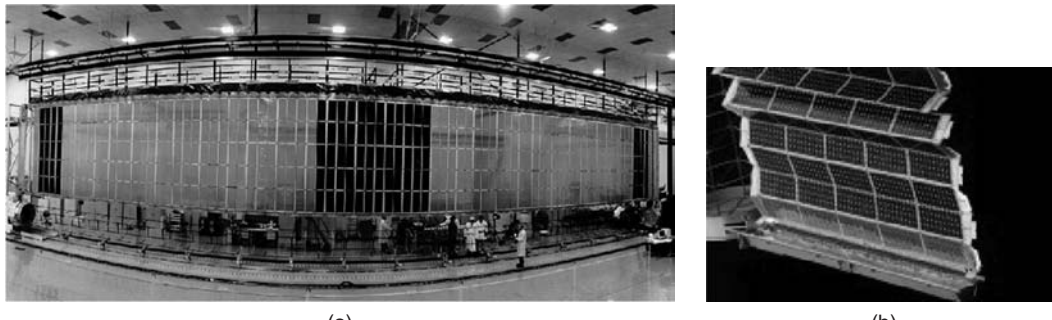
Figure 9.11 Rigid panel GaAs solar array (Picture courtesy of NASA). (The PV array consists of the two rectangular panels of five modules each)

Measuring Mission (TRMM) and Rossi X-Ray Timing Explorer (XTE) both employ rigid panel arrays. Available power supplied by typical rigid panel arrays range from very small to in excess of 100 kW.

9.5.3 Flexible Fold-out Arrays

Flexible fold-out arrays are attractive for missions that require several kilowatts of power because of their high specific power, high packaging efficiency (low stowed volume), and simple deployment system. These arrays are generally designed in two basic configurations:

1. Flexible flat panel array with linear deployment as shown in Figure 9.12.
2. Flexible round panel array with circular deployment as shown in Figure 9.13.



(a) (b)

Figure 9.12 (a) ISS array and (b) during its linear deployment. (Figure courtesy of NASA)

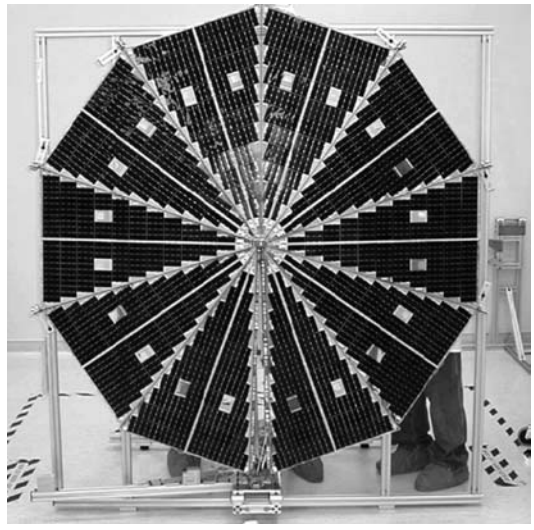


Figure 9.13 Flexible round panel array with circular deployment. (Picture courtesy of AEC-ABLE)

These arrays have flexible or semi-flexible panels that are stowed for launch with accordion folds between each panel. On reaching an appropriate orbit, these are unfurled by means of an Astromast™, an Ablemast™, or some other similar device. The specific power of these types of arrays varies from 40 to 100 W/kg, depending on the cell type, power, mission-reliability requirements, spacecraft orientation and maneuverability capabilities, and safety requirements. Initially, they were marketed as a significant improvement in power produced per unit mass. However, even though flexible arrays have an excellent figure of merit in this regard, the best rigid honeycomb panels have thus far matched their specific power performance. Very large flexible blanket solar arrays present complex structural and spacecraft design issues. This type of array is used on the MILSTAR series of spacecraft, on the TERRA spacecraft, and on the ISS (see Figure 9.2).

TRW, Inc., developed a flexible flat panel/rectangular array in the late 1980s, known as the Advanced Photovoltaic Solar Array (APSA), under a contract with NASA/JPL [68]. A similar array development was performed with DOD support. These arrays were based on the same fundamental concept of using polyimide panels stretched between lightweight hinges that could be deployed by an extendible mast. Silicon cells with an average AM0 efficiency of 14% were used for the arrays. The structure (mast, release motor, containment box) accounted for ∼51% of the array mass with the panel assembly (polyimide substrate, solar cells, cover glass, and interconnect tabs, hinges, wiring harness) making up the balance.

The original APSA design was for 130 W/kg at 5.3 kW BOL in GEO. However, the specific power of this type of array does not scale linearly. The low-power arrays of this design had a specific power on the order of 40–60 W/kg. Projected BOL and EOL specific power of the APSA array for both Si and GaAs cells available in the early 1990s are given in Figure 9.14 [14].

The Terra satellite uses an APSA-type array. The specific power of this array is only ∼40 W/kg. This was due to the necessity of reinforcing the stowage box because the box could not be stiffened by the spacecraft structure as APSA had assumed. Also, a stronger and heavier substrate than that APSA had assumed possible was used. The ISS array also had a BOL specific energy of 40 W/kg owing to additional maneuverability and safety and reliability requirements put on the arrays [13].

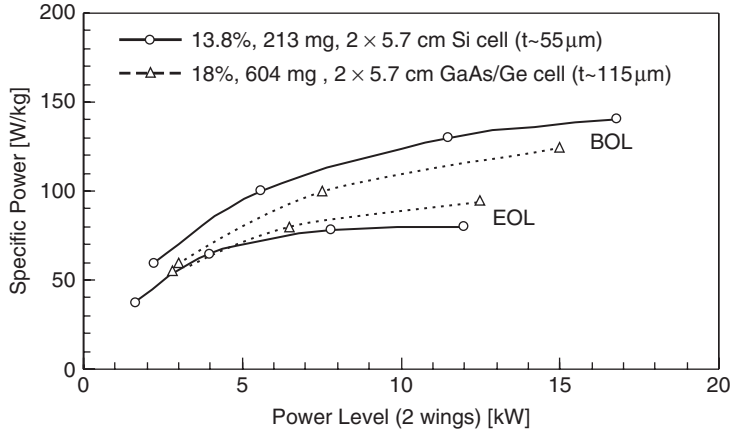


Figure 9.14 Projected specific power of the APSA array with various cell technologies [14]

9.5.4 Thin Film or Flexible Roll-out Arrays

The flexible roll-out array is similar to the accordion-folded array mentioned earlier, except for the fact that the semi-flexible or flexible substrate is rolled onto a cylinder for launch. The Hubble space telescope used such a roll-out array (Figure 9.15). It contained a polyimide blanket in a rolled-up stowed configuration. The array was deployed by a tubular, extendible boom (Bi-STEM) deployment system. The flexible role-out array design was developed for the US Air Force.

After eight years in orbit, the solar arrays on Hubble were replaced on orbit due to degradation [69]. During the repair mission, delamination of the solar array busbars was observed and it was also noticed that two of the hinge pins had started to creep out. One of the arrays was returned to Earth to be studied, while the other array was jettisoned into space. The returned solar array was shipped to ESA for further study. These roll-out arrays were replaced with rigid panels that were thought to be more reliable.

AFRL has begun a $6 M, three-year program with two prime contractors (Boeing and Lockheed Martin) to investigate and design complete arrays uniquely tailored to thin film solar cells. The SquareRigger^{TM} solar array being developed by AEC-ABLE is a flexible blanket system composed of modular "bays." This array is attempting to combine an ultra-high-power capability (>30 kW) with a high stowed packaging efficiency. The SquareRigger^{TM} solar array system is projected to achieve a specific power between 180 and 260 W/kg BOL, depending on the type of cells used. A SquareRigger^{TM} system using thin film cells is projected to offer an order-of-magnitude reduction in cost over conventional rigid panel systems.

Ultraflex Solar Arrays were recently used to provide power to the Mars Phoenix Lander. Launched in August 2007, the Phoenix Mars Mission is the first in NASA's Scout Program. Phoenix is designed to study the history of water and habitability potential in the Martian arctic's ice-rich soil. This was the first flight for this unique solar array technology developed by ATK's Goleta, California facility. Each Ultraflex array unfolds like an oriental fan into a circular shape 2.1 m in diameter, and will generate 770 watts of power from sunlight at the distance of Earth from the sun. Since Mars is approximately 1.5 times farther from the sun, the solar arrays will produce less than half the power possible on Earth. These same arrays have been projected as the solar arrays for use on Orion, the new proposed crew exploration vehicle to replace the Shuttle. The AFRL/RV High Power Solar Array (HPSA) and DARPA High Power Generation System (HPGS) solar array

Figure 9.15 Roll-out arrays used on the Hubble space telescope. (Photo courtesy of NASA)

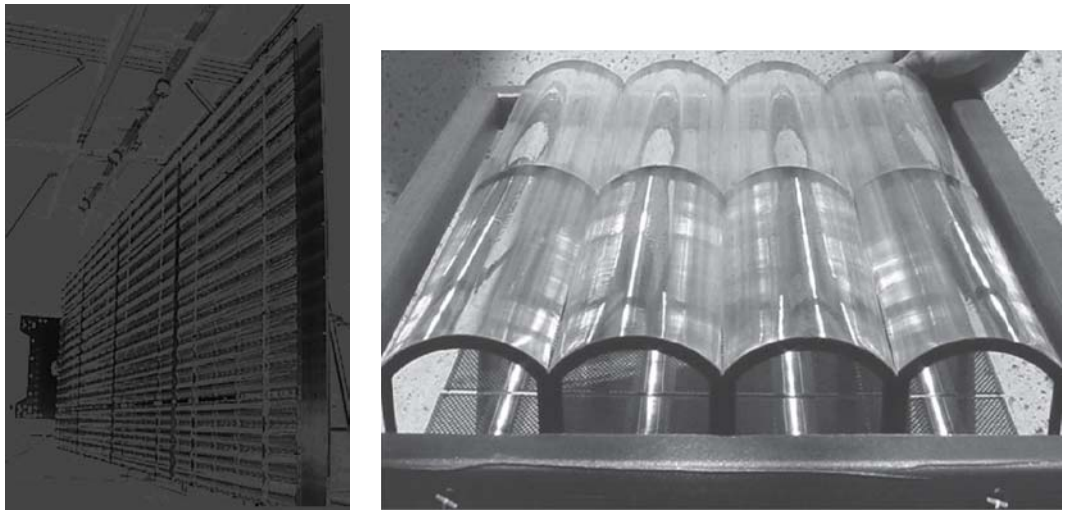
projects demonstrate promise to deliver up to 500 kW on-orbit in existing launch vehicle fairings using thin, flexible arrays of 12- to 14-μm-thick solar cells.

9.5.5 Concentrating Arrays

Photovoltaic concentrating arrays have been proposed for outer planetary missions, solar electric propulsion missions, and missions that operate in high-radiation environments. These arrays are attractive for these missions because they have the potential to provide a high specific power, higher radiation tolerance, and improved performance in LILT environments. The technical issues in using concentrating arrays are precision pointing, thermal dissipation, nonuniform illumination, optical contamination, environmental interactions, and complexity of deployment. They can also decrease overall spacecraft reliability because a loss of pointing may cause significant power loss to the spacecraft.

Reflective systems can have concentration ratios from 1.6× to over 1000×, with a practical limit of around 100×. Refractive designs are generally limited to the range of about 5× to 100×, with a practical limit of around 20×. Solar energy may be focused on a plane, line, or point, depending on the geometry of the concentrator design. These concentrators may be small and numerous if used in a distributed focus design or they may be a single large concentrator as in a centralized focus design.

AstroEdge^{TM} array on the NRO STEX spacecraft, launched in October 1998, was the first spacecraft to use a concentrator as its main power source. This system used a reflective trough



(a) (b)

Figure 9.16 (a) The SCARLET array used on Deep Space 1; (b) a next-generation flexible SCARLET prototype (Pictures courtesy of NASA)

design with a nominal 1.5× concentration. The arrays were successfully deployed and cell currents were slightly higher than predicted. Thermal problems did occur on some of the panels owing to the higher concentrator operating temperature.

The Deep Space 1 spacecraft launched in October 1998 used SCARLET concentrator arrays to provide power to its ion propulsion engines (Figure 9.16) [14]. Its two arrays were capable of producing 2.5 kW at 100 V DC. The Scarlet array was developed by AEC-ABLE under a program sponsored by BMDO.

The SCARLET array has a refractive linear distributed focus with a 7.5× concentration ratio. The array has 720 lenses to focus sunlight onto 3600 solar cells. Deep Space 1 has two SCARLET solar array wing assemblies. Each assembly is made up of a composite yoke stand-off structure, four composite honeycomb panel assemblies, and four lens frame assemblies. High-efficiency triple-junction $GaInP_2$/GaAs/Ge cells were used in this array.

The first commercial concentrator array developed for space was the Boeing 702. It was used on the Galaxy XI spacecraft and was deployed on January 12, 2000. It has a reflective planar centralized focus concentrator design in which the sun's rays are reflected onto a single rectangular plane of solar cells. It used thin film reflectors and had a 1.7× concentration. It was designed for power levels of 7–17 kW over a 16+ year design life. The array deployed as expected and its initial power output was within the expected range. However, its concentrator surfaces degraded very quickly while in orbit. The specific power of this array was ~60 W/kg using 24% efficient MJ solar cells. The similar Boeing 601 bus, which uses an ordinary planar solar array, is limited to about 15 kW of power owing to the array stowed volume limitations.

9.5.6 High-temperature/Intensity Arrays

Missions to Mercury and other missions with close encounters to the sun (i.e. solar probe) have generated the need for cells and arrays that are capable of operating in high-light-intensity,

high-radiation, and high-temperature environments. Two missions that had to contend with such an environment have already flown. Helios A, which reached 0.31 astronomical units (AU) – the average Earth to sun distance – was launched on December 10, 1974 and Helios B, which reached 0.29 AU, was launched on January 15, 1976. These spacecraft used ordinary silicon cells that were modified for high-intensity use and had second surface mirrors to cool the array. The remainder of their technology was very similar to what is used on standard arrays. In addition to these missions, the upcoming MESSENGER Discovery mission is planned for travel to 0.31 AU. Its solar array design is already under development.

The current solar array technology can meet the needs of MESSENGER or other spacecraft that approach the sun to about 0.3 AU, but with reduced performance and increased risk compared to other applications. Further progress is required in both cell and array development for closer encounters to the sun. The common feature to the high-temperature and high-intensity solar arrays that have operated thus far is the replacement of a significant fraction of the solar cells by optical solar reflectors (OSRs). These are mirrors that help control the array temperature near the sun at the cost of reduced power at larger distances.

The MESSENGER design also off-points the array as the spacecraft nears the sun to keep the array below 130 °C. The array is designed to tolerate pointing at the sun for a maximum of 1 h (probably much longer). However, it will be unable to function under this extreme condition (i.e. 260 °C).

The US Air Force and BMDO also developed some high-temperature arrays in the late 1980s. The survivable concentrating photovoltaic array (SCOPA) and survivable power system (SUPER) were designed to be capable of surviving laser attack. These were concentrator arrays that directed the incident laser light away from the solar cells. Although the laser light would not impinge directly, the array's temperature would increase dramatically and thus the arrays needed to withstand several hundred degrees Celsius.

The high-temperature survivability of SCOPA and SUPER was achieved through changes to the contact metallization and through the use of diffusion barriers in the GaAs cells used. Both Tecstar and Spectrolab developed the cells in conjunction with this effort. Other smaller companies such as Astropower, Kopin, and Spire have also worked on developing high-temperature cells. GaAs cells reaching an AM0 efficiency of 18% were produced that degraded less than 10% under one-sun after annealing in vacuum for 15 min at 550 °C. Concentrator cells were produced that survived repeated 7-min excursions to 600 °C. These same cells exhibited only 10% loss with exposure to 700 °C. NASA is also currently funding an effort to develop wide-bandgap solar cells for high-temperature/high-intensity environments. Cells using materials such as SiC, GaN, and AlGaInP are being developed [69]. These cells may also benefit from high-emissivity selective coatings that will limit the unusable IR entering the solar cells and reduce their steady-state temperature.

9.5.7 Electrostatically Clean Arrays

There is an entire class of proposed missions designed to study the Sun–Earth Connection (SEC). These spacecraft typically measure the fields and particles associated with the solar wind. This requires that arrays be developed that do not distort the local environment or be electrostatically "clean." These arrays must have their voltage separated from the space plasma and the array must be maintained at the same potential as the spacecraft. This is usually accomplished by coating the cell coverglass and arrays between the cells with a conductor. Since the coating for the coverglass must be transparent, a transparent conducting oxide (TCO) such as indium tin oxide is used. The coatings between the cells must not short them out, so an insulating coating must first be applied to

all of the interconnects before the conductive coating or "V" clips. All of this must be done within a thickness of ~0.08 mm and within a width of about 0.8 mm.

Fabricating an electrostatically clean array presently costs three to six times as much as a typical array. This is due in large part to the hand labor involved in developing such arrays. These arrays are also less reliable due to the lack of robustness of the conductive coatings used to maintain the equipotential. In addition, these arrays are also generally body-mounted, which cuts down on the available power to the spacecraft (i.e. pointing issues, etc.). The power is also limited due to the thicker coverglass that is employed owing to the high-radiation environment associated with SEC missions. Unfortunately, there is not a wide knowledge base on how to develop electrostatically clean arrays. This was demonstrated in the cost of developing the fast auroral snapshot (FAST) solar array. The electrostatically clean body-mounted solar panels for FAST cost in excess of $7400 per test condition watt.

The use of monolithic diodes on the latest generation of MJ solar cells could prove to be a tremendous advantage in developing electrostatically clean arrays. The presence of antennas, booms and outcroppings from a body-mounted array, requires that solar cells have bypass diodes to reduce the shadowing losses and potential damage to the arrays. The new built-in diodes will obviate the need and the expense involved in adding the diodes to the array circuitry. The NASA Goddard Space Flight Center (NASA-GSFC) recently funded Composited Optics Incorporated (COI) to study electrostatically clean arrays through the Solar Terrestrial Probe (STP) Program's Magnetospheric Multiscale (MMS) and Geospace Electrodynamic Connection (GEC) projects. COI will be supplying the electrostatically clean solar panels for the Communication/Navigation Outage Forecast System (CNOFS).

9.5.8 Mars Solar Arrays

Mars orbiters have used PV arrays that are quite similar to those used in Earth orbit with good results. However, Mars surface missions, in which the solar spectrum is depleted at short wavelengths, causes the efficiency of the cells to be lower than that if the cells were operated above the atmosphere of Mars. The cell efficiency is reduced by about 8% (relative %). In addition, the effect of dust accumulating on arrays was observed on the Mars Pathfinder mission by monitoring the J_{SC} of cells exposed to the environment whose short-circuit current could be monitored on a routine basis. One cell indicated an increase in obscuration of about 0.3%/sol for the first 20 sols (note that a "sol" is a Martian day of 24.6 h). The other cell indicated that over a longer period of ~80 sols, the obscuration flattened out and seemed to be approaching an asymptote of around 20% obscuration [69]. Cells that are "tuned" to the Martian solar spectrum and methods for mitigating dust obscuration will be necessary to produce efficient arrays for Mars surface power.

9.5.9 Power Management and Distribution (PMAD)

There are a number of different devices involved in efficiently connecting a space solar array to its intended loads. A system for managing and distributing the power consists of regulators, converters, charge controllers, blocking diodes, and wiring harness [70]. This system must condition the power to maintain the appropriate current and voltage levels to the power subsystems under varying illumination, temperature, and with cell degradation over the mission lifetime. The electrical bus for this system must also be able to isolate individual panel faults in such a way that the entire spacecraft will not lose total power in the case of a panel failure. The entire power management and distribution (PMAD) typically will account for 20–30% of the entire power system mass in the case of a conventional array. This can be reduced if an unregulated system is used.

Very often solar power generation is combined with a battery storage element that can be used in eclipse. In order to provide the appropriate charging conditions for the batteries and to avoid overcharging and heating, peak-power tracking (PPT) or direct energy transfer (DET) are used. PPT controls the arrays, so they only produce the power levels required by means of a DC–DC converter in series with the array. Peak-power tracking is typical on missions that need less power at EOL. A PPT system uses about 5% of the power produced by the array. Systems using DET operate using the fixed voltage of the array and shunt the excess power through shunt resistors. The fixed voltage of the array is chosen to be close to the EOL maximum power point voltage. These systems generally have a higher EOL efficiency and therefore are used on longer missions.

On an unregulated bus with a battery storage component, the loads will experience whatever voltage is currently on the batteries. This can lead to a large swing in voltage (i.e. 20%) to the load, depending on the battery chemistry and the depth of discharge. In a quasi-regulated system that employs a simple battery charger, the loads will be maintained at a voltage that is higher than the voltage on the batteries during charging. However, the loads will track the decrease in battery voltage as they are discharged (i.e. during eclipse). A fully regulated system that uses a regulator will maintain a constant load voltage, independent of the charging or the discharging cycle. A fully regulated system requires more elements and thus increases the PMAD complexity and mass. There will also be a decrease in system efficiency due to the power loss from overall bus resistance. However, it does provide more reliability and increases battery life. The resistive power losses can be minimized by using a higher bus voltage. The maximum voltage limits on an array is set by the voltage that the exposed parts of the power system can hold off without discharging through the space plasma (i.e. $\sim$50 V for LEO).

9.6 FUTURE CELL AND ARRAY POSSIBILITIES

9.6.1 Low-intensity Low-temperature (LILT) Cells

The term LILT is used to refer to solar arrays operating under conditions encountered at distances greater than 1 AU from the sun. Typical Earth-orbiting solar arrays have steady-state illuminated temperatures of approximately 40–70 °C. The efficiency of most cells increases down to about −50 °C. This temperature would correspond to around 3 AU. Currently available solar cells have uncertain performance under LILT conditions. NASA Glenn Research Center has initiated a program to evaluate solar cells under LILT conditions and to look for ways of enhancing their performance.

9.6.2 Quantum Dot Solar Cells

A recent approach to increasing the efficiency of thin film PV solar cells involves the incorporation of quantum dots [71]. Semiconductor quantum dots are currently a subject of great interest, mainly due to their size-dependent electronic structures, in particular the increased bandgap and therefore tunable optoelectronic properties. To date these nanostructures have been primarily limited to sensors, lasers, LEDs, and other optoelectronic devices. However, the unique properties of the size-dependent increase in oscillator strength, due to the strong confinement exhibited in quantum dots and the blue shift in the band gap energy of quantum dots, are properties that can be exploited for developing PV devices that offer advantages over conventional photovoltaics. Theoretical studies predict a potential efficiency of 63.2%, for a single size quantum dot, which is approximately a factor of two better than any SOA device available today. For the most general case, a system

with an infinite number of sizes of quantum dots has the same theoretical efficiency as an infinite number of bandgaps or 86.5%. Some recent work has indicated that quantum dot solar cells may have some favorable radiation resistance which could also be useful for space utilization [72]. See Chapter 4 for a more complete discussion of quantum dots and theoretical efficiencies.

9.6.3 Integrated Power Systems

NASA has also been working to develop lightweight, integrated space power systems on small-area flexible substrates [73]. These systems generally consist of a high-efficiency thin film solar cell, a high-energy density solid-state lithium-ion battery, and the associated control electronics in a single monolithic package. These devices can be directly integrated into microelectronic or micro-electromechanical systems (MEMS) devices and are ideal for distributed power systems on satellites or even for the main power supply on a nanosatellite. These systems have the ability to produce constant power output throughout a varying or intermittent illumination schedule, as would be experienced by a rotating satellite or "spinner" and by satellites in a LEO by combining both generation and storage.

An integrated thin film power system has the potential to provide a low-mass and low-cost alternative to the current SOA power systems for small spacecraft. Integrated thin film power supplies simplify spacecraft bus design and reduce losses incurred through energy transfer to and from conversion and storage devices. It is hoped that this simplification will also result in improved reliability. The NASA Glenn Research Center has recently developed a microelectronic power supply for a space flight experiment in conjunction with the Project Starshine atmospheric research satellite (http://www.azinet.com/starshine/).

This device integrates a seven-junction small-area GaAs monolithically integrated photovoltaic module (MIM) with an all-polymer $LiNi_{0.8}Co_{0.2}O_2$ lithium-ion thin film battery. The array output is matched to provide the necessary 4.2 V charging voltage, which minimizes the associated control electronic components. The use of the matched MIM and thin film Li-ion battery storage maximizes the specific power and minimizes the necessary area and thickness of this microelectronic device. This power supply was designed to surface-mount to the Starshine 3 satellite, which was ejected into a LEO with a fixed rotational velocity of 5° per second. The supply is designed to provide continuous power even with the intermittent illumination due to the satellite rotation and LEO [74].

9.6.4 High Specific Power Arrays

To achieve an array specific power of 1 kW/kg, a much higher cell specific power will be necessary. Similarly, the blanket specific power (i.e. interconnects, diodes, and wiring harnesses) must be over 1 kW/kg as well. The APSA assessment determined that the mass of the deployment mechanism and structure is essentially equal to the blanket mass for a lightweight system [31]. Therefore, a blanket specific power of approximately 2000 W/kg would be necessary to achieve a 1-kW/kg array. NASA is currently sponsoring an effort by AEC-ABLE Engineering to develop lightweight thin film array deployment systems.

Gains in array specific power may be made by an increase in the operating voltage. Higher array operating voltages can be used to reduce the conductor mass. The APSA was designed for 28 V operation at several kilowatts output, with the wiring harness comprising ~10% of the total array mass, yielding a specific mass of ~0.7 kg/kW. If this array was designed for 300-V operation, it could easily allow a reduction of the harness specific mass by at least 50%. This alone would increase the APSA specific power by 5% or more without any other modification.

The extremely high specific power arrays that need to be developed for SEP and SSP applications will require lightweight solar arrays that are capable of high-voltage operation in the space plasma environment. SEP missions alone will require 1000–1500 V to directly power electric propulsion spacecraft (i.e. no voltage step-up is required to operate the thrusters). NASA has proposed a thin film stand-alone array specific power that is 15 times the SOA III–V arrays, an area power density that is 1.5 times that of the SOA III–V arrays, and specific costs that are 15 times lower than the SOA III–V arrays [75].

9.6.5 High-radiation Environment Solar Arrays

There are several approaches to mitigating the effects of a high radiation on a solar array. The simplest is to employ thick coverglass (assuming that a commercial source could be developed). Thick coverglasses protect the cell from the highly damaging low-energy protons, but will cause a significant decrease in the specific power of the array. However, this can be reduced if one adopts a concentrator design, assuming of course that the additional elements associated with the concentrator can withstand the high-radiation environment as well. A different approach is to try to develop cells that are more radiation-resistant. Several of the materials that are being investigated for high-temperature/high-intensity missions have also shown good radiation resistance. However, these will not be suited to the high-radiation missions involving LILT. Many of the high-radiation NASA missions being considered occur at distances much greater than 1 AU. Thin films may offer a possibility since they have demonstrated some advantages with regard to radiation tolerance as previously mentioned, provided the problem of their low efficiencies is solved.

9.7 POWER SYSTEM FIGURES OF MERIT

There are many figures of merit that must be considered in developing an SSP system (i.e. specific mass, specific power, cost per watt, temperature coefficients, and anticipated radiation degradation of the solar cells used).

The radiation hardness and the temperature coefficients for the III–V multijunction cells are significantly better than Si cells, as previously discussed. This leads to significantly higher EOL power level for a multijunction cell compared with a Si cell. This is shown in Table 9.8, where the BOL cell efficiencies at room temperature and the typical EOL cell efficiencies for LEO and GEO operating temperatures and radiation environments are presented.

The difference in radiation degradation can have a huge impact on power system design. For example, if the area for a typical rigid panel is approximately $8\,m^2$ and the area of a typical solar cell is $24\,cm^2$, using a panel packing factor of 0.90 will allow the panel to have 3000 cells. Under GEO conditions, this panel populated with high-efficiency Si cells will produce 1.2 kW of EOL power. The EOL power could almost be doubled to 2.2 kW if it were populated with SOA triple-junction cells. Alternatively, the solar arrays populated with high-efficiency Si cells would need to be 77% larger than arrays using triple-junction cells in order to deliver the equivalent amount of EOL power in GEO and 92% larger in LEO.

The large difference in size between solar arrays populated with Si and MJ cells is very significant in terms of stowage, deployment, and spacecraft attitudinal control. This is especially true for very high-powered GEO communication satellites in which Si solar array area can exceed $100\,m^2$. The comparable array with triple-junction cells, although by no means small, would have an area of $\sim 59\,m^2$. The array size will impact the spacecraft's weight, volume (array stowage), and system requirements on spacecraft attitude control systems (additional chemical fuel). Three

Table 9.8 A comparison of relative radiation degradation of 75-μm multijunction cells and high-efficiency Si under GEO and LEO operation [68]

Solar cell technology	BOL efficiency @ 28 °C (%)	EOL efficiency on orbit (%)
GEO conditions (60 °C)−1-MeV, 5×10^{14} e/cm^2		
HE Si	14.1	12.5
2J III–V	20.9	20.0
3J III–V	23.9	22.6
LEO conditions (80 °C)−1-MeV, 1×10^{15} e/cm^2		
HE Si	13.4	10.6
2J III–V	19.7	18.1
3J III–V	22.6	20.3

Table 9.9 EOL area power density (W/m^2), specific weight (W/kg), and normalized cost ($/W) for high-efficiency Si, dual-junction (2J), and triple-junction (3J) bare solar cells [68]

Solar cell technology	W/m^2	W/Kg	Normalized cell cost ($/W)
GEO conditions (60 °C)−1-MeV, 5×10^{14} e/cm^2			
75 μm HE Si	169	676	1.00
2J III–V	271	319	1.38
3J III–V	306	360	1.22
LEO conditions (80 °C)−1-MeV, 1×10^{15} e/cm^2			
75 μm HE Si	143	574	1.00
2J III–V	245	288	1.29
3J III–V	275	323	1.15

important figures of merit used in power system optimization are EOL area power density (W/m^2), specific weight (W/kg), and cost ($/W). Representative values for the various SOA cell technologies are listed in Table 9.9.

The EOL power per unit area for a MJ cell is significantly better than a Si cell. However, the EOL specific weight for Si is almost a factor of two greater than a MJ cell. This results in a slightly smaller EOL cost per watt for a high-efficiency Si. This demonstrates the dramatic reduction in cost of MJ cells over the past few years.

If one considers the mass of the necessary array components (i.e. panel substrate, face sheet, adhesive, hinges, insulators, wiring, etc.) along with equivalent power per area for the different cell types, and also the cost involved in having the cells interconnected and covered (CIC) and laid on

Table 9.10 EOL specific weight (W/kg) at the CIC and panel levels for three-mil high-efficiency Si, dual-junction (2J), and triple-junction (3J) cells [68]

Solar cell technology	CIC specific power (W/kg)	Panel specific power (W/kg)	Normalized panel cost ($/W)
GEO conditions (60 °C)−1-MeV, 5×10^{14} e/cm^2			
75 μm H.E. Si	261	75	1.00
2J III–V	219	95	0.9
3J III–V	248	108	0.8
LEO conditions (80 °C)−1-MeV, 1×10^{15} e/cm^2			
75 μm H.E. Si	221	63	1.00
2J III–V	199	86	0.84
3J III–V	223	97	0.75

rigid panels, then the cost for developing an array using MJ cells is slightly less than that for HES cells. The EOL specific weight values at the CIC (with 100-μm ceria-doped microsheet coverglass) and the panel levels for these cells and the normalized cost per watt for the panels is shown in Table 9.10. A similar comparison with somewhat less expensive 100-μm high-efficiency Si cells (at the panel level) shows a slightly smaller cost advantage for the multijunction cells. The 100-μm Si cells are less radiation-hard and have lower specific power (W/kg) than 75-μm Si cells, but they cost about 35% less.

Currently conventional space Si cells are less expensive than MJ cells at the panel level. However, their EOL power is much lower than either the high-efficiency Si cells or the MJ cells. The increased mass and area that their usage entails would have to be considered against the cost savings and other mission considerations in any comparative study.

9.8 SUMMARY

Engineers have worked on ways to improve space solar cells and arrays in terms of all the important figures of merit since the early days of our space program. Numerous mission studies have shown that even extremely high array costs can be worth the investment when they result in lower array mass. In general, mass saving in the power system can often be used by payload. If the revenue generated by this payload (i.e. more transmitters on a communications satellite) is greater than the cost of higher-efficiency solar cells, the choice is rather an easy one to make. However, more instrument capabilities will often require more support from the spacecraft (e.g. command and data handling, structure, attitude control, etc.) as well as more power. These additions can negate any apparent advantage to the overall spacecraft.

REFERENCES

1. Chapin D, Fuller C, Pearson G, *J. Appl. Phys*. **25**, 676–681 (1954).
2. Jenny D, Loeferski J, Rappaport P, *Phys. Rev*. **101**, 1208–1212 (1956).
3. Easton R, Votaw M, *Rev. Sci. Instrum*. **30**, 70–75 (1959).
4. Loeferski J, *J. Appl. Phys*. **27**, 777–785 (1956).

5. Jackson E, *Trans. of the Conf. on the Use of Solar Energy*, Vol. 5, pp 122–128 (Tucson, AZ, 1955).
6. *Bell Syst. Tech. J*. **42** (1963).
7. Solar Cell Array Design Handbook, *JPL* **SP43-38**, Vol. 1, pp 1.1–2 (1976).
8. Statler R, Curtin D, *Proc. International Conf. on the Sun in the Service of Mankind*, pp 361–367 (1973).
9. Reynolds D, Leies G, Antes L, Marburger R, *Phys. Rev*. **96**, 533 (1954).
10. Lebrun J, *Proc. 8th IEEE Photovoltaic Specialist Conf.*, pp 33–37 (1970).
11. North N, Baker D, *Proc. 9th IEEE Photovoltaic Specialist Conf.*, pp 263–270 (1972).
12. Bailey S, Raffaelle R, Emery K, *Proc. 17th Space Research and Technology Conf*. (2001).
13. Hague L, *et al*., *Proc. 31st Intersociety Energy Conversion Engineering Conf*., pp 154–159 (1996).
14. Stella P, *et al*., *Proc. 34th Intersociety Energy Conversion Engineering Conf*. (1999).
15. Bailey S, Landis G, Raffaelle R, *Proc. 6th European Space Power Conf*. (2002).
16. Ralph E L, High Efficiency Solar Cell Arrays System Trade-Offs, *First WCPEC*, pp 1998–2001 (1994)
17. Bailey S, *et al*., *Solar Cell and Array Technology for Future Space Science Missions*, Internal NASA report to Code S (2002).
18. Banerjee A, Xu X, Beernink K, Liu F, Lord K, DeMaggio G, Yan B, Su T, Pietka G, Worrel C, Ehlert S, Beglau S, Yang J, Guha S, *Proc. 35th IEEE Photovoltaics Specialists Conf.* (2010).
19. Stern T, Krumweide D, Gaddy E, Katz I, Proceedings of the *28th IEEE Photovoltaic Specialists Conference*, Anchorage, AK (September 2000).
20. Wanlass M W, Ahrenkiel S P, Albin D S, Carapella J J, Duda A, Geisz J F, Kurtz S, Moriarty T, Wehrer R J, and Wernsmann B, *Proc. 31st IEEE Photovoltaics Specialists Conference*, p. 230 (2005).
21. Yoon H, Hadda M, King R R, Law D C, Fetzer C M, Sherif R A, Edmondson K M, Kurtz S, Kinsey G S, Cotal H L, Krut D D, Ermer J H, and Karam N H, *Proc. 20th European Photovoltaic Solar Energy Conference*, p. 118 (2005).
22. Boisvert J C, Law D C, King R R, Bhusari D M, Liu X-Q, Zakaria A, Mesropian S, Larrabee D C, Woo R L, Boca A, Edmondson K M, Krut D D, Peterson D M, Rouhani K, Benedikt B J, Karam N H, *Proc. 35th IEEE Photovoltaics Specialists Conf.* (2010).
23. Breen M L, Street A R, Cokin D S, Stribling R, Mason A V, Sutton S C, *Proc. 35th IEEE Photovoltaics Specialists Conf.* (2010).
24. Kurlan R, Schurig H, Rosenfeld M, and Herriage M, *28th IEEE PVSC*, pp 1061–1066 (2000).
25. Imaizumi M, Takahashi M, Takamoto T, *Proc. 35th IEEE Photovoltaics Specialists Conf.* (2010).
26. Bücher K, Kunzelmann S, *Proc. 2nd World Conference and Exhibition on Photovoltaic Solar Energy Conversion*, 2329–2333 (1998).
27. Green M, *et al*., *Prog. Photovolt*. **6**, 35–42 (1998).
28. Green M, *et al*., *Prog. Photovolt*. **9**, 287–293 (2001).
29. King R, *et al*., *Proc. 28th IEEE Photovoltaic Specialist Conf.*, pp 982–985 (2000).
30. Reynard D, Peterson D, *Proc. 9th IEEE Photovoltaic Specialist Conf.*, p. 303 (1972).
31. Crabb R, *Proc. 9th IEEE Photovoltaic Specialist Conf.*, pp 185–190 (1972).
32. Bailey S, *et al*., *Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, pp 3650–3653 (1998).
33. Ferguson D, *Interactions between Spacecraft and their Environments*, AIAA Paper 93-0705, NASA TM 106115 (1993).
34. Landis G, Bailey S, *AIAA Space Sciences Meeting*, AIAA-2002-0718 (Reno, NV, 2002).
35. Flood D, *Proc. NHTC'00, 34th National Heat Transfer Conf*. (2000).
36. Spanjers G, Winter J, Cohen D, Adler A, Guarnieri J, Tolliver M, Ginet G, Dichter B, Summers J, *Proc. 2006 IEEE Aerospace Conference*, pp 1–10 (2006).

37. Anspaugh B, *GaAs Solar Cell Radiation Handbook*, 6–54, NASA JPL Publication 96-9 (1996).
38. Marvin D, Nocerino J, *Degradation Predictions for Multijunction Solar Cells on Earth-Orbiting Spacecraft*, Aerospace Report No. TOR-2000(1210)-2, 9 (2000).
39. Walters R, *et al*., *Technical Digest of the International PVSEC-11*, pp 813–814 (1999).
40. Walters R, *Proc. 15th SPRAT*, pp 30–34 (1997).
41. Tada H, Carter J, Anspaugh B, Downing R, *Solar Cell Radiation Handbook*, 3–82, JPL Publication 82–69 (1982).
42. Fahrenbruch A, Bube R, *Fundamentals of Solar Cells*, Chap. 2, Academic Press, Boston (1983).
43. Landis G, *Proc. 13th SPRAT*, pp 385–399 (1994).
44. Scheiman D, *et al*., *Proc. 17th SPRAT Conference* (Cleveland, OH, Sept. 11–13, 2001).
45. Bailey S, Flood D, *Prog. Photovolt*. **6**, 1–14 (1998).
46. Keener D, *et al*., *Proc. 26th IEEE Photovoltaic Specialist Conf.*, pp 787–281 (1997).
47. Aiken D J, Cornfeld A B, Stan M A, Sharps P R, *Proc. 4th World Conference of Photovoltaic Energy Conversion*, p. 838 (2006).
48. Cornfeld A, Aiken D, Cho B, Ley V, Sharps P, Stan M, Varghese T, *Proc. 35th IEEE Photovoltaics Specialists Conf.* (2010).
49. Tatavarti R, Hillier G, Dzankovic A, Martin G, Tuminello F, Navaratnarajah R, Du G, Vu D P and Pan N, Lightweight, Low Cost GaAs Solar Cells On 4 Epitaxial Liftoff (ELO) Wafers, *33rd IEEE PVSC* (2008).
50. Sharps P R, Cho B, Cornfeld A, Diaz J, Newman F, Stan M, and Varghese T, "Next Generation High-Efficiency Space Solar Cells", *Space Power Workshop* (2009).
51. Hoffman D, *et al*., *Proc. 35th IECEC*, AIAA-2000-2919 (2000).
52. Bailey S, Hepp A, Raffaelle R, *Proc. 36th Intersociety Energy Conversion Engineering Conference*, pp 235–238 (2001).
53. Guha S, *et al*., *Proc. 2nd World Conf. PV Solar Energy Conversion*, pp 3609–3612 (1998).
54. Deng X, Povolny H, Han S, Agarwal P, *Proc. 28th IEEE Photovoltaic Specialist Conf.*, pp 1050–1053 (2000).
55. Arya R, *et al*., *Proc. 1st World Conf. Photovoltaic Solar Energy Conversion*, pp 394–400 (1994).
56. Contrera M, *et al*., *Prog. Photovoltaics* **7**, 311–316 (1999).
57. Hanket G, *et al*., *Proc. 29th IEEE Photovoltaic Specialist Conf.*, pp 567–570 (2002).
58. Dhere N, Ghongadi S, Pandit M, Jahagirdar A, Scheiman D, *Prog. Photovoltaics* **10**, 407–416 (2002).
59. Romeo A, Blatzner D, Zogg H, Tiwari A, *Mat. Res. Soc. Symp. Proc*. Vol. 668, pp H3.3.1–3.6 (2001).
60. Ferekides C, *et al*., *Thin Solid Films* **361–362**, 520–526 (2000).
61. Marshall C, *et al*., *IECEC*, 1999-01-2550 (1999).
62. Messenger S, *et al*., *Proc. 16th European Photovoltaic Energy Conference*, pp 974–977 (2000).
63. Tringe J, Merrill J, Reinhardt K, *Proc. 28th IEEE Photovoltaic Specialist Conf.*, pp 1242–1245 (2000).
64. Stella P, West J, *Proc. 21st IEEE Photovoltaic Specialist Conf.*, pp 1362–1366 (1990).
65. Contini R, Ferrando E, Hazan D, Romani R, Campesato R, Casale M C, Gabetta G, "Comparison Between CIGS And Triple Junction GaAs On Thin Ge Solar Cell Assemblies and Related Development Strategies", *33rd IEEE PVSC* (2008).
66. Gerlach L, Fournier-Sirce A, Fromberg A, Kroehnert S, *Proc. 21st IEEE Photovoltaic Specialist Conf.*, pp 1308–1312 (1990).
67. *Science*, Vol. 305, issue 5685, pp 737–900 (6 August 2004).
68. Scheiman D, Landis G, Weizer V, *AIP Conf. Proc*. **458**, 1–6 (1999).
69. Landis G, *Acta Astronautica* **38**, 1 (1996).
70. Larson W, Pranke L, Eds, *Human Spaceflight Mission Analysis and Design*, McGraw Hill, New York (1999).

71. Luque A, Marti A, *Phys. Rev. Lett*. **78**, 5014 (1997).
72. Leon R, *et al*., *Appl. Phys. Lett*. **76**, 2071 (2000).
73. Hoffman D, Raffaelle R, Landis G, Hepp A, *Proc. 36th IECEC*, IECEC-2001-AT-21 (2001).
74. Raffaelle R, *et al*., *Proc. 36th IECEC*, IECEC2001-AT-66 (2001).
75. Bailey S, *et al*., *Proc. 17th Euro. Photovoltaic Solar Energy Conference*, pp 2137–2143 (2001).

10

Photovoltaic Concentrators

Gabriel Sala and Ignacio Antón

Instituto de Energía Solar, Universidad Politécnica de Madrid, Spain

10.1 WHAT IS THE AIM OF PHOTOVOLTAIC CONCENTRATION AND WHAT DOES IT DO?

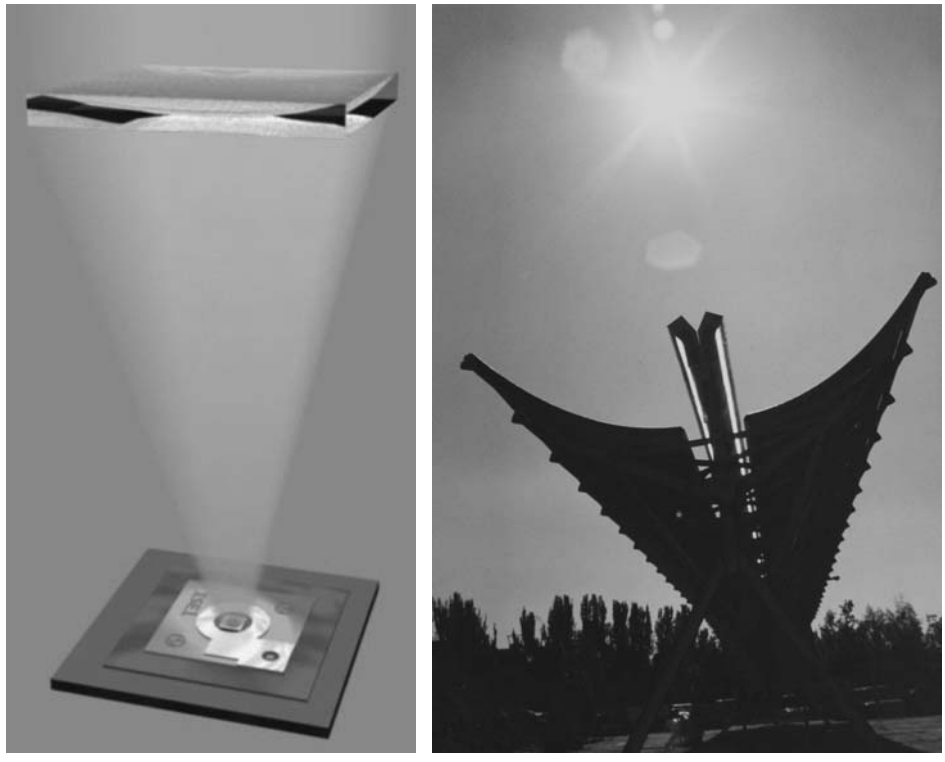
Photovoltaic concentration is an alternative technology to flat panels, consisting of substituting a flat surface of solar cells with optical elements which increases the density of the luminous power coming from the sun (irradiance W/m^2) and projects it onto the special cells whose area is much less than the capturing surface. Figure 10.1.

Although the theory of thermodynamics applied to photovoltaic cells predicts an increase in conversion efficiency [1], which experience confirms, the only reason to justify the development of concentrator photovoltaic (CPV) technology is the plausibility that with it, electrical energy will be produced, both sustainably and acceptably in terms of the environment. Concentration does not reduce the area of the collector except if it increases its efficiency.

Concentration photovoltaic technology came about at more or less the same time as terrestrial photovoltaics (PV). In 1975, PV cells were so expensive that the idea of reducing their area and substituting them with optical elements became an immediate and attractive option. That is why in 1975 a national program was launched in the USA, led by Sandia National Laboratories (DOE) to develop ideas and concentration photovoltaic prototypes [2]. Some of the prototypes demonstrated the feasibility of the concept and the evolution of one of them became the most installed commercial proposal which, in 2009, accumulated up to 10 MW [3] (Figure 10.2).

In spite of the rapid progress of CPV technologies from 1976 to 1985, it soon became clear that the market was not prepared to build the large facilities that CPV would require, given its little modularity (>1 MW) and the need to compete at this scale with the price of conventional energy from power stations. The modularity limit was not strictly a result of the concentration technology, but arose from the use of structures with tracking, which is only profitable in large plants of several MW.

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8



(a) (b)

Figure 10.1 The idea of photovoltaic concentration. (a) A lens with a rotational symmetry concentrates the direct sun light from onto a point focus (Courtesy Fraunhofer-ISE). (b) A parabolic cylinder concentrator forms a linear image of the sun on receiving solar cells. (courtesy IES-UPM, Euclides Prototype 1996)

Combining the appearance of large plants for grid connection on the market and the adoption of over 40% efficient multijunction cells has brought about a revisiting of concentration technology, with renewed development and investment.

With the continued fall in the prices of conventional photovoltaic modules, the basic principle of CPV established by Sandía Labs, that is, a reduction in the amount of cells, has become less attractive, and the principle of reducing the total area of the system through the use of high-efficiency multijunction cells that operate at several hundreds of suns currently prevails [4–6]. However, proposals persist that try to achieve competitiveness through simple concentration systems operating at less than 10× on cells with efficiencies in the range of 20% [7].

10.2 OBJECTIVES, LIMITATIONS AND OPPORTUNITIES

10.2.1 Objectives and Strengths

To help readers carry out their own analysis of the capacity of a technology to achieve certain economic objectives and its sensitivity in respect to the most significant parameters, below we

Figure 10.2 Evolution and commercialisation of the Fresnel lens technology initiated in Sandia Labs (1976): 750 kW_p plant, manufactured and installed by Guascor Fotón in Cáceres, Spain, 2007) (Courtesy Guascor Fotón) See Plate 2 for the colour figure

present a simple equation to give the cost of the energy produced by a PV concentrator, in annual terms for a given climate [8]

$$COST\ kWh(€) = \left[\frac{B.O.S.\left(\frac{€}{m^2}\right) + \frac{Cell\left(\frac{€}{m^2}\right)}{C}}{E_{in}\left(\frac{kWh_{sol}}{m^2}\right) \cdot \eta_{sys}\left(\frac{kWh_{elect}}{kWh_{sol}}\right)} \times (ADR) \right] \tag{10.1}$$

where:

BOS: is the cost per m^2 of all of the elements of the concentrator, including the installation in the field, except for the cost of the solar cells and those elements directly affected by the concentration factor (for example: the cell substrate, the secondary optics, etc.).

Cell: is the cost of the cell per m^2.

C: is the effective concentrator factor.

E_{in}: is the useful accumulated radiation that reaches 1 m^2 of the receptor in a year (kW h/m^2 per year). It is a function of the place. Only the direct radiation is usually useful, except for gains of less than 5× that a certain amount of diffuse radiation could capture.

η_{sys}: is the efficiency of the complete system measured in terms of electrical energy delivered in respect to the luminous energy available on the surface of the array collector. This figure is the combined result of the nominal cell efficiency, of their operating temperature in the field, of the optical efficiency of the modules on the tracker and of the efficiency in the dc/ac conversion.

(ADR): is the annual discount rate: annual distribution of the cost of the investment up to the obsolescence of the system (usually 25 years).

The formula demonstrates that the sunniest places generate cheaper energy, and that the unit price of the most efficient cells, although they might be very large, can be reduced drastically through the use of high concentration. Meanwhile, the conversion efficiency appears as a direct divisor of the cost of the system. It could be said that thanks to the concentration factor the increase in efficiency costs almost nothing.

The reader must realise that some, not yet explicit, price has to be paid to use the concentration, brought about by the inevitable losses produced by the optic system.

As will be justified later, the CPV needs to keep the collectors accurately pointed at the sun in order to work at an optimum level. That is why its mobile structure could be more expensive than that of the static flat modules.

In spite of all of this, the concentrators have an enormous potential deriving from the high efficiency of their cells, because the materials that make up a CPV, which are up to 80% of the system cost, are conventional and widely available. For example, a huge 100 GW yearly production of CPVs will just impact on 2% of the steel market and 10% of glass (assuming the collectors are made of glass).

As well as the availability of materials, there are a large number of manufacturers that have the capacity and the workers adapted to this technology (for example, companies in the automobile or domestic appliance market). The solar cells could be produced by a few high-technology companies, capable of making rapid decisions in an expanding, and low-risk, market scenario [9].

The proof of this scenario can be seen in the pages of the main professional magazines in the sector: up to 40 businesses have been created since 2006 worldwide with the objective of manufacturing CPV systems of the most varied aspects, efficiencies and sizes.

In spite of this, just a few companies are commercializing concentrators at the MW level. In the group using primary reflexive collectors in high concentration with multijunction cells we find Solfocus (USA) and Solar Systems (AUS) (Figure 10.3).

However the majority of commercial concentrators in 2009 are based on PMMA Fresnel lenses, as the refractive primary collector, and cells operating over 30 W/cm^2. For example, Guascor-Fotón, a Spanish company which uses modified Amonix technology and Back Point Contact silicon cells [10] at 27% efficiency, is the leader in the CPV market, having installed 9 MW between 2007 and 2008 (see Figure 10.2). It has now adopted multijunction cells instead of silicon cells. The remaining large manufacturers of modules with lenses as primary collectors are Concentrix Solar (GE), which uses hybrid glass-silicone Fresnel lenses, and Arima Eco, Taiwan (Figure 10.4).

10.2.2 The Analysis of Costs of Photovoltaic Concentrators

We shall now develop the effect of the level of production on the cost, of the components that make up the module and the optimum size of the cells. This analysis will be limited to the most common current module, that is, Fresnel lenses with multijunction cells, with a secondary optical concentrator (SOC) and two axis tracking system. Revolutionary proposals may differ from these results, but we believe it to be illustrative.

The size of the lens and the concentrator factor are very important aspects in modules of this type. With very small cells we increase both the wafer losses and the complexity of the production, however, we save on materials, storage and transportation because the module and housing are thinner and the lenses smaller. We are therefore going to study the costs in accordance with the size of the cell for a fixed gain of 1000× (80 W/cm^2).

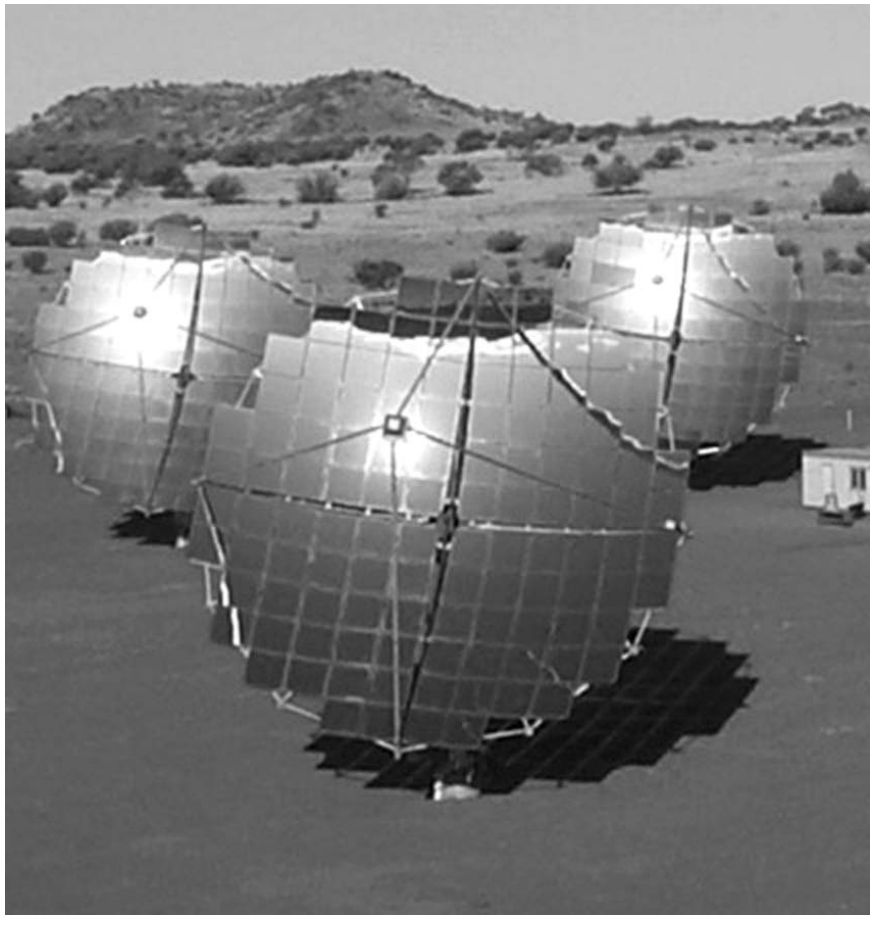


Figure 10.3 Faceted parabolic disk reflector from Solar Systems (AUS) that has a large-area focus to provide 500× on an actively cooled receiver (Courtesy of John Lasich)

The type of receptor for the module with a primary Fresnel lens that we are using for the analysis is shown in Figure 10.5.

We are going to analyse two very prudent scenarios, significant in the starting phase of a factory, relative to the levels of production: 10 MW/year and 30 MW/year.

By exploring cell sizes from 1 to 100 mm^2, for 1000× and f-number = 1, we have obtained the curve detailed in Figure 10.6, which shows a minimum module cost for cells between 10 and 20 mm^2, but increasing slowly for a larger size. This slow increase in cost versus cell size after the minimum is brought about by the worse heat dissipation in larger cells and increased I^2R_s loses.

A specific analysis for cells 1 and 9 mm^2 is presented in Table 10.1, the latter being the most convenient. Costs of 1.0€/W at 30 MW/year manufacturing rate can be obtained with a module efficiency of just 20% at 850 W/m^2, at a normal operating cell temperature (NOTC). Concentrix Solar has already demonstrated 27% efficiency modules and 23% of arrays in the field [11].

The cost of the tracking system and that of the installation in the field have been added, as has the power conditioning and connection to the network.

Figure 10.4 (a) FLATCON-module developed and manufactured at Fraunhofer-ISE. The FLATCON concept is now commercialised by Concentrix Solar GmbH (Courtesy Fraunhofer-ISE). (b) Commercial concentration system from Concentrix using the FLATCON module (Reproduced by permission of Concentrix Solar GmbH)

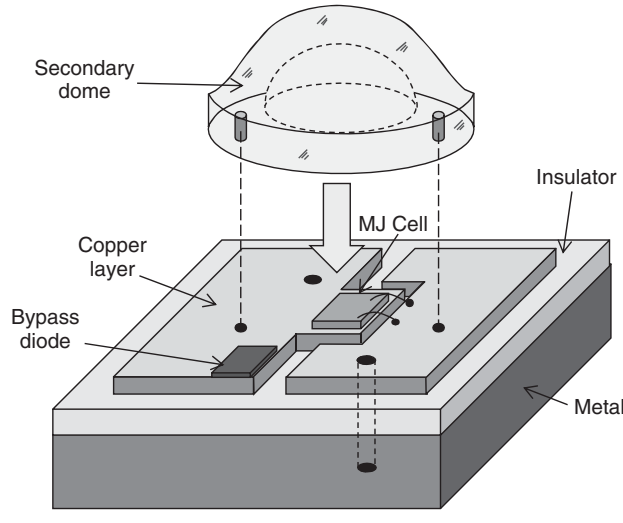


Figure 10.5 Typical receptor for a two-stage optics with a Fresnel lens in the primary. The substrate provides insulation and good thermal transmission. It houses a bypass and a "silo" or "domo type" secondary. (Courtesy UPM-IES)

It is interesting to point out that the cost of the machinery and labour is not significant in relation to the module materials, as it can be reduced as production volume increases, mainly the cells themselves. The latter may also contribute through their increase in efficiency, for which there is still a margin.

We can see from the Table 10.1 that when incorporating the field costs, installation and tracking system, valued at between 0.7 and 0.9 €/W_p, the cost gets close to that of flat panels, which indicates that the tracking plays still a too big a role in the cost and that it should continue

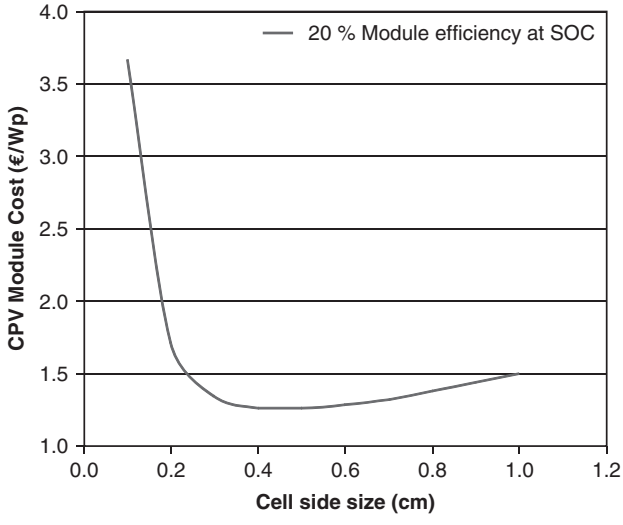


Figure 10.6 Cost of the module with an f:1 Fresnel lens and multijunction cell assuming of 20% field efficiency versus the size of the cell. The increase to the right of the minimum is the result of the increase in temperature and loss due to the Joule effect (I^2R) (Courtesy UPM-IES)

Table 10.1 Costs per W_p of the 20% efficient CPV modules under standard operating conditions (850 W/m^2 and cells at operating temperature)

	Cells 1 × 1 mm		Cells 3 × 3 mm	
	10 MW €/W_p	30 MW €/W_p	10 MW €/W_p	30 MW €/W_p
Machinery	0.13	0.04	0.09	0.03
Labour	0.08	0.05	0.06	0.04
Materials	2.82	2.70	0.976	0.93
Total module cost	3.02	2.80	1.13	1.00
Installation and tracking	0.9	0.7	0.9	0.7
System in the field (DC)	3.39	3.50	2.03	1.70
Grid-connected (AC)	4.27	3.75	2.28	2.05

to be reduced. With a module efficiency of 30%, and 27% in the field, the cost of these CPVs is lower than that of any other photovoltaic technology.

10.3 TYPICAL CONCENTRATORS: AN ATTEMPT AT CLASSIFICATION

10.3.1 Types, Components and Operation of a PV Concentrator

A photovoltaic concentrator always has two main and inescapable elements: A *collector* capable of redirecting the rays of the sun towards a *special solar cell* with a smaller area.

There are systems sometimes called "concentrators" that use flat mirrors to intensify the light onto conventional panels [12]. We prefer to call them "systems with enhancing mirrors" in

order to distinguish them from CPV's in which the cells and receptors are designed while taking into account their demanding operating conditions as well as ensuring their reliability and complying with the specific official regulations.

In order to illustrate the basic concept of PV concentrators we are going to focus first on the system shown in Figure 10.1a. It consists of a rotational convex lens that focuses the disk of sunlight onto the receiving solar cells, when the sun beam is maintained normal to the lens surface with suitable sun tracking driver.

As in all photovoltaic systems, a significant part of the light that reaches the cells will be converted into heat since the efficiency of the conversion is less than 100%. Having increased the power by a factor close to the relationship of the collector/cell area, it might be necessary to include special heat evacuation elements to pass the waste heat to the surrounding air.

The lens type collector allows the heat to be dissipated across the whole rear side of the CPV module. Compared with a flat module the concentrators in operation must dissipate less heat per unit area than the former because it only collects direct radiation and the latter are more effective in generating electricity.

The lenses need to be enclosed, both as a safeguard for the internal side of the lenses and to provide structural support: This enclosure also serves to protect the receptors from the elements, although condensation of water inside the modules with Fresnel lenses is one of the difficulties associated with this technology; Figure 10.7. Some manufacturers have adopted direct actions to avoid it, such a blowing in dry air when there is the daily condensation.

In the previous example we have looked at an optical system, also known as point focus, that projects the circular *image* of the sun onto the cell. It is also possible however to concentrate the light onto a linear receptor by means of cylindrical collectors. The most commonly used concentrator type for this geometry is the trough reflector assembly, which operates at low or medium concentration level; Figure 10.1b.

This concentrator must move so that its plane, normal to the aperture, always contains the solar disk. If the collector is a mirror, the tracking is achieved by turning on just one axis. The focal point is a line, and the receptor usually consists of a group of cells placed in a focal line and connected in series; Figure 10.8.

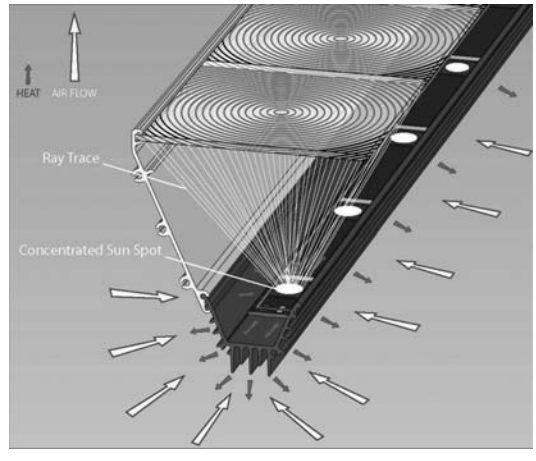


Figure 10.7 Principle and components of a classic one stage Fresnel lens CPV module (Reproduced by permission of Whitfield Solar Ltd)

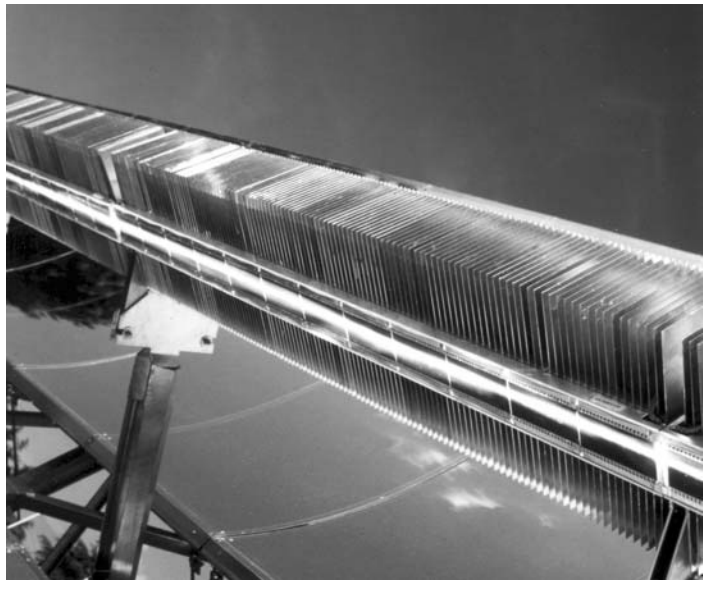


Figure 10.8 Fins in a linear CPV concentrator receiver to dissipate the heat (Courtesy UPM-IES)

Receptors with linear Fresnel lenses have also been made, but the "mechanism" of Snell's law requires the tracking to be made using two axes in order to keep the focal point permanently on the plane of the cells. Linear systems can use active or enhanced passive refrigeration with fins (Figure 10.8); the active refrigeration is usually accompanied by a thermal use at very low temperature.

Although single-axis systems are simpler and use less space for a nominal given power, they also collect less energy as the sun is not always perpendicular to the collector.

New designs with improved optical performance are expected based on non-imaging optics, a new discipline of optics which aims to project the luminous power onto a target without preserving the image on any plane in the system. Non-imaging optics allows ideal concentrators to be designed, which can reach the maximum gain allowed by physics.

Using near ideal concentrators, it is possible to make lineal concentrators up to 5×, called static concentrators, which can "see and concentrate" the sun disk rays all the year as well as a great part of the sky. They are usually cylindrical and oriented East–West. They are not yet commercial although several were developed in early 1980s.

A revision of the know-how on concentrators previous to the present rebirth of CPCs can be found in the classic paper "The Promise of Concentrators" [13] written by the multiple prize winner, Dr Richard Swanson. In this excellent article the reader can find the best bibliography reference list, covering everything relevant to CPV until the year 2000.

10.3.2 Classification of Concentrators

The PV concentration generators have more characteristic parameters than the flat modules, which is why they can be classified under many more possible criteria. For example: geometrical shape, optics, concentration level, cell type, tracking method, heat dissipation, etc.

If we permute all of the options for each criterion we will find around 500 design possibilities. Although the number of realistic proposals, when they are filtered through expected or measured performance and economic factors, is reduced greatly.

So, in summary, the two currently dominant trends are:

(A) Very-high-concentration (>300×) point-focus systems with highly efficient cells (>35%) with a high specific cost which are built with III–V materials.
(B) Low or medium concentration (2–60×) with silicon cells of up to 20–22% efficiency and at low cost.

Once these principal trends are clear, Table 10.2 shows the realistic combinations of components that are found in current feasible CPVs. (The table is better understood after reading the definitions of the concepts "*module*" and "*assembly*" given in Section 10.6).

10.3.3 Concentration Systems with Spectral Change

We believe it interesting to mention a family of concentrators that try to get round the limitations to the concentration linked to the brightness and angular extension of the sun as seen from the Earth (see Section 10.4) through the absorption of solar light by dyes (or pigments) and its later isotropic-fluorescent emission [15, 16].

The process seems, in principle, to be doomed to failure because the temperature of the original source is reduced or degraded (the dyes re-emit at a lower temperature than the sun), but given that the single-junction cells are badly adapted to the natural solar spectrum (they also degrade the source temperature in the absorption), this disadvantage is fairly well compensated. On the other

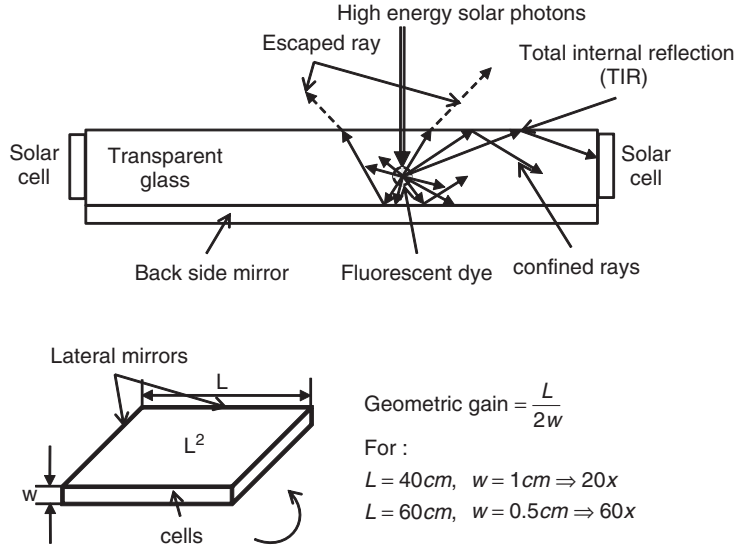


Figure 10.9 Basic principles of operation of a luminescent concentrator. (a) Many rays emitted by a dye molecule are confined in the transparent dielectric plate in which edges are located small solar cells. (b) Concept of theoretical geometric gain in a fluorescent concentrator (Courtesy UPM-IES)

Table 10.2 Table of the C-RATING Project (EU, 5thFP) for the classification of Concentrators [14]

Reference concentrator	Primary optics	Cell assembly	Cell type	Concentration ratio	Cooling	Tracking	Secondary optics
CPV module (point-focus)	Fresnel lens or small bulk lens or small parabolic dish or RXI devices	One single cell or several cells with spectral beam splitting	Single-junction silicon or single-junction III–V or multijunction	$50 < \times g < 500$ for silicon cells >500 for all other cells	Passive	Two-axis	Yes/No
Point-focus assemblies	Big or medium-size parbolic dish or central tower power plant	Parquet of cells	Single-junction silicon or single-junction III–V or multijunction	$150 < \times g < 500$	Active	Two-axis	Yes/No
Linear systems	Linear lens or parabolic trough	Linear array of cells	Silicon or III–V (with 3D Secondary)	$15 < \times g < 60$ (without secondary) $60 < \times g < 300$ (with secondary)	Passive	One-axis for parabolic troughs; two-axis for lenses	Yes/No
Static systems	Non-imaging device	Usually linear array of cells	Silicon	$1.5 < \times g < 10$	Passive	No or manual	No

hand, the light is absorbed and emitted in a space of index $n > 1$ which allows a partial, but reasonable confinement of the light. Figure 10.9 details the basic layout of a luminescent concentrator.

We see in Figure 10.9 that the collimated rays of the sun (incident from any angle) produce a re-emission which is confined at around 70% within the transparent dielectric that contains the pigments. If we place cells just on the sides of the dielectric plate, a gain in geometric concentration equivalent to the $L/2w$ coefficient is produced. This gain could be between 20 and 60 times. Unfortunately the efficiency of the pigments and re-absorption of the fluorescent light reduce these values to less than $6\times$.

The stability of the pigments must also be improved so that this idea, still under development, could turn into practical commercial applications.

10.4 CONCENTRATION OPTICS: THERMODYNAMIC LIMITS

10.4.1 What is Required in Concentrator Optics?

Concentrator optics for PV applications have become the tool necessary to access the very-high-efficiency solar cells for terrestrial applications by means of a significant decrease in the collecting area which should consequently bring about a real reduction in cost. In order to deal with the cost contribution of the expensive high-efficiency cells, the optics must provide a sufficient concentration level, sometimes called "gain", which should be as high as $500-1000\times$ for modern III–V cells.

Another quality of optics is usually called angular acceptance and refers to the degree of tolerance in pointing toward the sun, which is a condition that all concentrators require for casting the incoming direct light onto the receiving cell. The larger the angular acceptance of the concentrator optics the more light will impinge on the receiver throughout the year because the optics will send much light to the cells in spite of tracking errors and structure deformations. As we will see below, the concentration and the angular acceptance are opposite merits, governed by fundamental physical limitations.

In addition to these qualities, the ability of the optics to cast the light as uniformly or homogeneously as possible onto the solar cells is appreciated, mainly for single-junction concentrator cells. The uniformity can be achieved with independence of the aforementioned fundamental limitations, just with special design methods linked to non-imaging optical rules. A deeper analysis of the latter exceeds the objectives of this chapter but is available in [19].

10.4.2 A Typical Reflexive Concentrator

To begin with the concentration, we will consider a rotational parabolic disk mirror assumed to be perfect in shape and reflectivity. The reflectors comply with a simple optical law:

- angle of incidence = angle of reflection;
- the ray is kept within a plane with the normal to the surface, both before and after reflection.

From our use of the word "ray", the reader will have guessed that we are going to use what is known as a "geometric optics" for the analysis of the concentrators.

In Figure 10.10 we represent the curve of the dish as cut for a plane that contains the rotational axis. We obtain a parabola on the plane. The most well known property of this curve

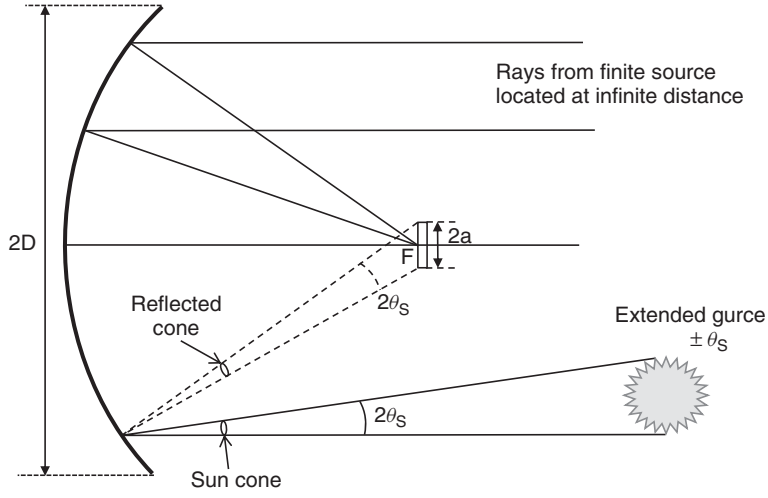


Figure 10.10 Relation between the angular width of the light source and the concentration gain (for any concentrator). The cone of angle $2\theta_s$ requires a receiver of dimension $2a$ for total collection of rays. The gain is $C = D^2/a^2$. For wider sources the value of a increases and the gain C is lower (Courtesy UPM-IES)

(as well as the dish) is that all of the rays parallel to the axis that impinges on the inner mirrored surface of the parabola are reflected and pass through a point, F, known as the focus.

The rays that come from the sun, which is 150 million km. from the Earth, reach each point of the planet as a conical bundle of rays within $\pm 0.27^\circ$ around the axis of this cone. It is not a "*point source*", rather it is an "*extensive source*". We can assume that the brightness is uniform all over the disk.

By orienting the parabolic disk mirror towards the centre of the solar disc we can see that the central ray of the cone heads towards the point F, but the remaining rays are extended around this point F. If we want to collect everything we will have to use an extensive receptor, of diameter a. Therefore, considering that this occurs on all points of the disk we will see that the geometric gains on all of the disk will be, at most

$$C_G = \left(\frac{D}{a}\right)^2 \tag{10.2}$$

which is no longer an infinite number.

It is clear to understand that if the angular width of the cone were greater, $\theta'_s > \theta_s$ (for example, if we were closer to the sun) the size of the receptor necessary would be greater, $a' > a$. Thus we can see why the maximum level of the concentration depends on the size of the light source, and that the concentration for real sources has limits (it cannot be infinite).

These characteristics are not exclusive to the parabolic disk, but general in all concentration systems based on geometric optics, although the relationship between both magnitudes are different for some concentrators and not others. Thus for a parabolic disk the optimum ratio, that is, the one which provides the maximum concentration is

$$C_{\max} = \frac{1}{4\sin^2\theta_s} \tag{10.3}$$

An interesting property is derived from Snell's law of refraction: If the receptor were in a medium of index n, and the source in the air (the Sun for example) our receptor could see the Sun at an angle $\theta_{Dielectric}$, which fulfils $\sin\theta_{Dielectric} = (\sin\theta_{air})/n$. So that this receptor could be smaller, especially in the ratio n for linear systems and n^2 for point-focus.

Therefore, the maximum concentration achievable in a parabolic disk, accepting the complication surrounding the receptor with a dielectric index n, becomes

$$C_{\max} = \frac{n^2}{4\sin^2\theta_{air}} \tag{10.4}$$

10.4.3 Ideal Concentration

If we want to have a good concentrator the first thing we must try to do is make all the available rays coming from a source reach the entrance of the concentrator, and ensure that they are collected by it. This condition could be written as a simple equation for the conservation of the power from the source to the collector.

If we make this calculation for the light coming from a very remote disk source, with an angular extension $\pm\theta_s$, and uniform brightness B_s, that a concentrator with an entry area A_s, can achieve the following value for a total incident power on A_s : $P_s = \pi A_s B_s \sin^2\theta_s$. (Figure 10.11) [17].

The action of the concentrator will consist of taking all of this power to area A_r of the receptor through convergent beams of rays, all shaped like a cone $\pm\theta_r$ and with brightness B_r. By calculating, as before, all of the beams coming from the collector on the surface of the receptor we get the power $P_r = \pi A_r B_r \sin^2\theta_r$ which must equal the entry power.

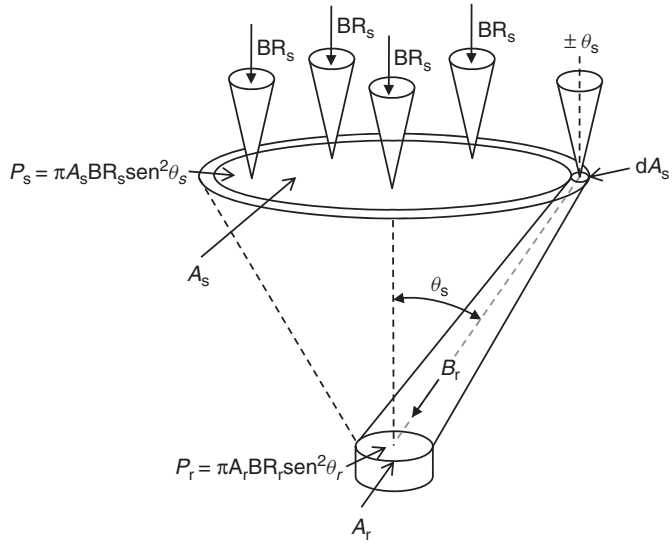


Figure 10.11 The conservation of the power along the optical system (first law). After Poincaré, in optics, this is usually called conservation of the "etendue". The decrease of the light beam section (area) is compensated by an increase of the angular width (by $\sin^2\theta$). BR is the brightness (W m^{-2}sterad^{-1}) of the source which can be preserved or degraded, but never increased according to the second Principle (Courtesy UPM-IES)

We define the concentration as the quotient of areas A_s/A_r in these conditions, $P_s = P_r$, for equaling the aforementioned power

$$C = \frac{A_s}{A_r} = \frac{B_r \sin^2 \theta_r}{B_s \sin^2 \theta_{sr}} \tag{10.5}$$

In this expression it appears that it is in our hands to act solely on the numerator since the properties of the source (in the denominator), are untouchable.

But as the brightness is only a function of the temperature, the second law of thermodynamics prevents us from having brightness greater than that of the source. So the best case will be when $B_r = B_s$.

Now the only thing that we can optimise is $\sin^2 \theta_r$ whose maximum value is 1 for $\theta_r = 90^\circ$. That is, the maximum concentration will occur for a receptor illuminated isotropically (2π steradians) at each point on its surface for the concentrator with a brightness equal to the surface of the source.

The absolute maximum concentration is

$$C_{\max} = \frac{1}{\sin^2 \theta_s} \tag{10.6}$$

which becomes $n^2/\sin^2 \theta_s$ if the receptor is immersed in a material with index n times greater than the source. The concentrator that complies with this equation will be known as the *ideal* concentrator [18].

We can see again the dependency between the angular width of the entrance to the concentrator θ_s and the maximum concentration (losses brought about by absorption or other limitation of the real materials themselves are not included in this expression).

If we were capable of building a concentrator with the characteristics of Equation (10.6) for the sun, $\theta_s = \pm 0.27^\circ$ seen from the Earth, with a system with a point-focus, taking advantage of all of the rays that reach our receptor we would have

$$C_{max} = n^2 \cdot 45032 \tag{10.7}$$

In this receptor we would have the same irradiance and angular distribution of the rays that we have on the surface of the sun. In general we can say that "to concentrate" is like getting closer to the sun centre as many times as the square root of C_{ideal}. In this case we would have the exact brightness $B_r = B_s$ and the light reaching the receptor would be at $\pm 90^\circ$. These are precisely the conditions of the maximum irradiance possible that express the ideal concentration concept.

10.4.4 Constructing an Ideal Concentrator

We can now ask ourselves how an ideal concentrator capable of achieving an ideal gain as that defined in the following equation

$$C_{ideal} = n^2/\sin^2 \theta_s \tag{10.8}$$

can be made, and even if it exists.

Clearly the parabolic disk is only capable of reaching a quarter of this value if the receiver is immersed in index n. In an ideal concentrator the light should arrive to the receptor with a uniform irradiance on each point and from $\pm 90^\circ$.

With this criterion we can see that effectively the parabolic disk cannot be ideal because: (a) it does not send rays from all of the angles $\pm 90^\circ$ and cannot fill the whole receiver from all points of its surface. These limitations distance it from the ideal.

R. Winston was the first person to find a linear concentrator that complies with the ideal equation. It is called a CPC (compound parabolic concentrator) and is represented in Figure 10.12a. This concentrator is capable of taking all of the rays included in beams of $\pm\theta_s$ to the receptor in beams of $\pm\pi/2$.

The set θ_s is known as the angle of acceptance and is maximum in an ideal concentrator for a given gain. The transfer curve of an ideal concentrator is shown in Figure 10.12b.

The main consequence of the Equation (10.8) is deduced by calculating the value of θ_s for different values of the concentration. We would see that for concentrations greater than $10\times$ in an ideal linear concentrator only the rays impinging with less than 5.74° will reach the receptor.

Since the sun covers much greater angles in the sky, we must conclude that in order to achieve a permanent functioning of the concentrators they must move in order to keep the optics pointed at the sun with a precision that demands the value of its angle of acceptance.

10.4.5 Optics of Practical Concentrators

Although Winston's CPC concentrator allows the maximum angular aperture for a given concentration it has significant practical and economic limitations of use. In principle it consumes a lot of material in relation to the capture aperture. That is why practical concentrators have been directed to the use of more compact and simple technologies that offer the perspective of a lower cost, such as Fresnel lenses and aspherical mirrors instead of the ideal CPCs.

10.4.5.1 The Fresnel lens

The Fresnel lens used in CPVs is a refractive collector device with a low f-number (the f-number is the ratio between the focal distance and the diameter of the lens or the mirror). The convex

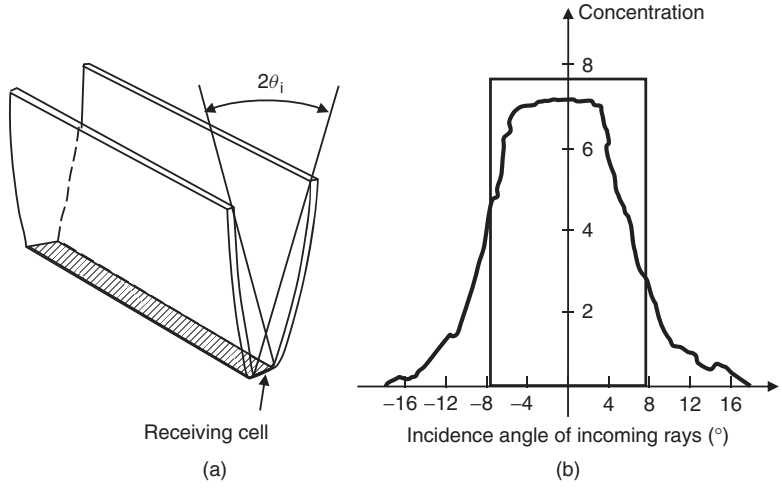


Figure 10.12 (a) The compound parabolic concentrator (CPC): the first discovered ideal concentrator. (b) The ideal and experimental transfer function of a reflective CPC (IES-UPM, Spain)

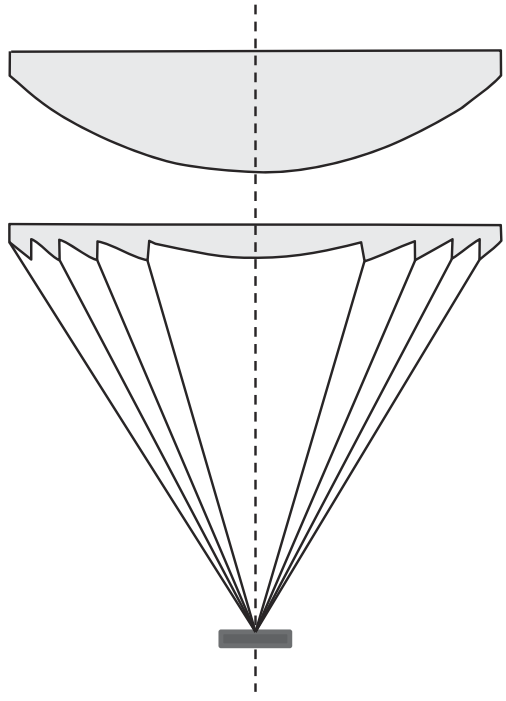


Figure 10.13 The concept of Fresnal lens: the thick conventional lens of low f-number is obtained by projection of the lens curvature. The result is equivalent performance and less material consumption (Courtesy UPM-IES)

conventional lenses with small f-numbers are very thick and heavy. This is why the Fresnel lens concept consists of projecting the slope of the bulky lens, creating discontinuous elements, as shown in Figure 10.13

It is the most used refractive concentrator as a result of its easy manufacture lightweight and convenience of use. The Fresnel lens modifies the beam of rays, increasing its convergence, but without inverting the direction of the rays. Thus the receptor can be positioned below the lens and has a large surface in order to dissipate the excess heat without creating a shadow.

The lens protects the receptor from the elements since the lower face of the lens, where the ridges are located are susceptible to permanent soiling. The upper face is normally flat for the same reason. That is why the modules with refractive elements are housed in enclosed boxes whose outer surface must be effective for the dissipation of heat.

The concentrators with a Fresnel lens are less compact than the reflectors and allow lower concentration levels. We can see why, when we design a Fresnel lens and calculate its maximum gain in the next section.

10.4.5.2 Design and limitation of Fresnel lenses

We are going to "design" a flat circular Fresnel lens with rotational symmetry. The discontinuous nature of discrete teeth of this lens gives us more freedom of design than a conventional lens because the latter is conditioned by the continuity of the derivative of the surface.

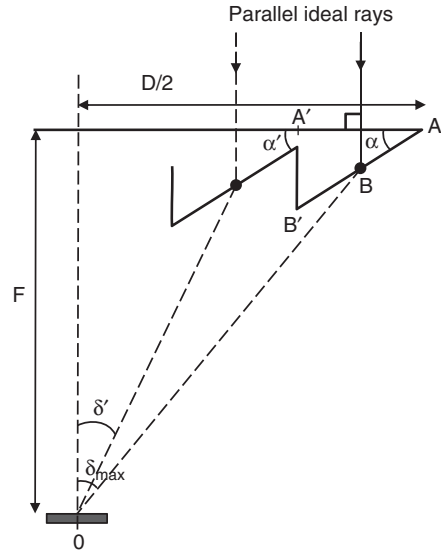


Figure 10.14 The principles for designing a Fresnel lens for PV concentration: the angle δ is related with the tooth angle α (Courtesy UPM-IES)

We are going to calculate first the slope, α, of the most external tooth, the one that will modify most the direction of the incident beam; Figure 10.14.

As the lens is perpendicular to the rays of the sun there is no refraction on the surface AA′ at point B. The outgoing ray must reach the centre of the receptor (it is a possible criteria of the design). Therefore it should form an angle δ_{max} which will comply approximately with

$$\tan \delta_{\max} = \frac{D}{2F} = \frac{1}{2f} \tag{10.9}$$

f being the "f-number". (This expression is more approximate as the size of the AA′ tooth is smaller.)

Using Snell's law, we obtain a relationship between $\alpha_{\max}$ and the f of the lens by means of Equation (10.9)

$$\sin^2 \alpha_{\max} = \frac{\sin^2 \delta_{\max}}{n^2 + 1 - 2n \cos \delta_{\max}} \tag{10.10}$$

In order to calculate the slope of the second tooth we only need to substitute the corresponding $\delta_{\max} \rightarrow \delta$ and $\alpha_{\max} \rightarrow \alpha$.

Although the tooth tilts are designed to go to the centre of the lens, the image will not be perfectly exact for three reasons:

1. The width of the AA′ tooth is not null and therefore many displaced parallel rays will extend over a length

$$CC' = AA'(1 - \tan \alpha \tan \delta) \tag{10.11}$$

on the receptor.

2. The solar source is not made up of parallel rays, but $\pm 0.27^\circ$ beams. This effect will increase the size of the beam at the focus in a diametrical interval

$$\Delta_s \approx \frac{F}{\tan\delta} 2\theta_s = 4\theta_s \frac{F^2}{D} = 4\theta_s f F \tag{10.12}$$

   We can see that in order to reduce this widening it is necessary to make very compact systems, with very small teeth and with a very short focal distance (we can see that it is degraded with F^2). A very compact system needs the slope α of the last ridge to be greater, but for $\alpha \approx 42^\circ$ a total internal reflection comes about and this tooth no longer works through refraction. Even before reaching $\alpha \approx 42^\circ$ significant losses known as Fresnel loses, come about in the A′B′ interface, as a result of the change of index in the air and plastic. That is why it is not possible to make effective refractive Fresnel lenses with $f \leq 0.9$, a disadvantage in comparison with the mirrors, which are more compact.
3. Finally, another significant phenomenon comes about as a result of the dispersion of white light on the ridges (they are small prisms) brought about by the variation of the refraction index of the material from which the lens is made. This well-known consistent chromatic aberration comes about, in our case, as the violet rays converge at an angle δ and the useful infrared rays at $\delta' < \delta$. So we have a third reason to widen the beam of light.

By combining all of these limitations we obtain a gain of just $80\times$ for a PMMA Fresnel lens with 1 mm fingers (if we do not want to lose any ray and want the angle of acceptance of $\pm 1^\circ$) [18].

The first Fresnel lenses for PV were designed to be used below $70\times$ and therefore would achieve their objectives with an acceptable angular acceptance. More recently, the exploitation of very-high-efficiency cells, but with a specific very high cost (€/cm^2) have required much greater levels of concentration, greater than $300\times$.

The restrictive physical effects of the aforementioned lenses have consequently brought about a reduction in the angle of acceptance at low levels that demand an almost "perfect" tracking.

The final consequence is that, for gains larger than 80, Fresnel lenses are usually combined with some type of secondary concentrator whose mission is to increase the optical efficiency and furthermore the angular acceptance.

10.4.6 Two-stage Optical Systems: Secondary Optics

We have already said that usual concentrators differ greatly through having the combined properties of gain and angular acceptance of the ideal ones. There is, however, a strategy that allows the characteristics of the angular acceptance to increase. It consists of placing a small "secondary" concentrator with "ideal" or "quasi-ideal" characteristics in the focus of an extensive collector, like a Fresnel lens or a parabolic mirror). The secondary concentrator can be a dielectric lens optically coupled or just a reflector filled with air.

Although the number of variations in the design is greater (information can be found in the specialised bibliography [19]), we can say, in general, that the secondary optical elements uses "non-imaging" devices which are designed with large acceptance angles at the entrance to see the primary collector as the source of light.

The truncated-cone or truncated-pyramid secondary device made using a simple sheet of folded aluminum has been very successful in achieving more than $80\times$ gain by using Fresnel lenses (Figure 10.15a).

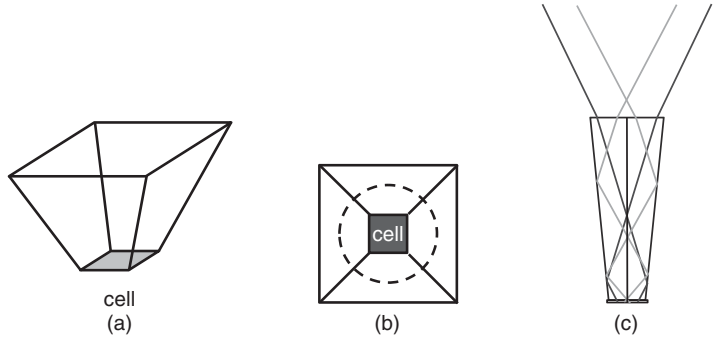


Figure 10.15 The most usual secondary optical devices. They are used to increase the limited gain of real Fresnel lenses (Courtesy UPM-IES)

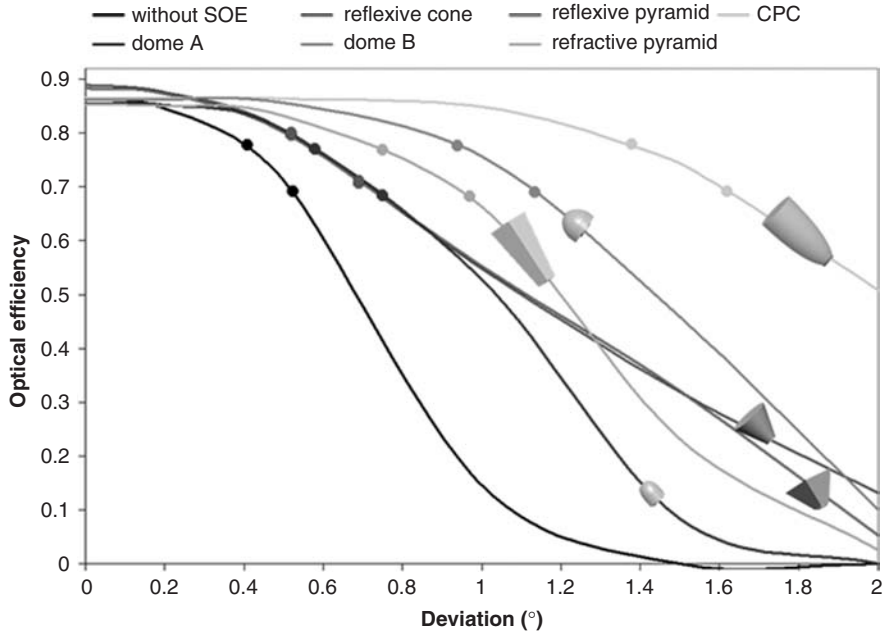


Figure 10.16 The function of transference of several concentrator systems made of the same Fresnel lens, but with different secondary optics. The secondaries made of dielectric material operating by TIR are theoretically better but the reflective ones are the most used up today (Courtesy UPM-IES)

With the cone (for round cells) or the pyramid (for square cells) the part of the spot that exceeds the size of the cell is diverted through reflection (Figure 10.15.b). With this procedure the size of the cell can be reduced but about 15% of the light is absorbed in the aluminum. This secondary collector also contributes something to the increase in the angular acceptance (Figure 10.16) [20].

With bulk dielectric secondaries a greater concentration gain for the same angular acceptance is achieved as a result of the n^2.

The secondary lenses are used likewise to achieve a kaleidoscopic effect to make the irradiance on the cell uniform. This action is very effective with a fairly slender dielectric truncated pyramids operating through total internal reflection (TIR), as shown in Figure 10.15c. The drawback of dielectric secondaries is the losses caused by the appearance of new air–dielectric or air–reflector interfaces.

The potentiality of the ideal concentrators, capable of taking the light to ±90% on the receiver cannot be used because beyond 65° the Fresnel reflection on the cells is excessive. That is why the maximum practical gain of the secondary device is in general

$$G_{\text{sec}} \leq \frac{n^2 \cdot \sin 65^\circ}{\sin^2 \delta_{\max}} \tag{10.13}$$

δ_{max} being the maximum angle coming from the primary lens, which is related to the f-number, in Equation (10.9).

Therefore, the gain from the ideal secondary optics, but with such a practical and efficient design will be

$$C_{\text{sec}} \leq 1,84(1 + 4f^2) \tag{10.14}$$

For a lens with $f = 1.2$ the maximum value of C_{sec} is 12.5×, increasing very rapidly with f. The non-ideal secondary optics, such as pyramids and cones cannot achieve these values.

The graph in Figure 10.16 shows the transfer function (or optical efficiency versus light incidence angle) of several typical secondary devices positioned below a Fresnel lens with $f = 1$. The Fresnel losses in the interfaces and the absorption in the real mirrors are accounted for in the graph. It can be seen that the rotational CPC, the closest to ideal of all of them, contributes a greater angular acceptance [20].

In general, secondary systems whose outgoing area is less than the cell have provided better results than the dome-shaped ones. However the latter are easier to manufacture and assemble and constitute a hopeful solution whose cost projection is about 0.1$/$W_p$. The synergy of this technique with similar massive LED applications will be positive for CPV.

Until now secondary lenses have been used by, at least, the following module manufacturing companies: Isofoton (2 types), Amonix, Daido, Emcore, Solfocus, Sol3G and Opel Solar. But almost all of the modules with a Fresnel lens will end up using them as they need a greater angular acceptance and concentration.

Although a significant effort has been made in designing two-stage optical systems which provide uniform illumination, in addition to gain and acceptance angle, the small currents and the high voltage of multijunction cells have reduced the importance of achieving this characteristic with respect to what happened with single-junction cells [21].

10.5 FACTORS OF MERIT FOR CONCENTRATORS IN RELATION TO THE OPTICS

10.5.1 Optical Efficiency

The light captured by a real concentrator goes through several modifications before reaching the solar cell: some of them are desirable, but others are not and cause losses of power in the

transformation. We define the *optical efficiency* as:

$$\eta_{op} = \frac{\text{Luminous power on the receptor}}{\text{Luminous power at the entrance of the concentrator}} \tag{10.15}$$

If the area of the entrance is A_i and the area of the receptor is A_r, we define the geometric concentration as

$$C_g = \frac{A_i}{A_r} \tag{10.16}$$

Thus, returning to the definition of η_{op} we have

$$\eta_{op} = \frac{< E_r > A_r}{B(n) A_i} = \frac{< E_r x >}{B(n)} \frac{1}{C_g} \tag{10.17}$$

where $B(n)$ is the direct normal irradiance (DNI) on the collector and $<E_r>$ is the average irradiance on the receptor surface (in W/m^2). So $\eta_{op} = 1$ means that we have an ideal optical systems without absorption, dispersion and geometric imperfections.

The optical elements of a concentrator usually modify the spectrum of the light entering through the preferential absorption of certain wavelengths and dispersion of colours: this is why we must explain, that the optical efficiency is in reality a spectral parameter, so that we have a value for η_{op} for each wavelength range $\Delta\lambda$ around λ (Figure 10.17). This is:

$$\eta_{op}(\lambda) = \frac{< E(\lambda, \Delta\lambda) >}{B(\lambda, \Delta\lambda) C_g} \tag{10.18}$$

For a given concentration collector, the optical efficiency decreases in accordance with the size of the receptor: indeed if we want to save a significant cell area it is possible that many rays

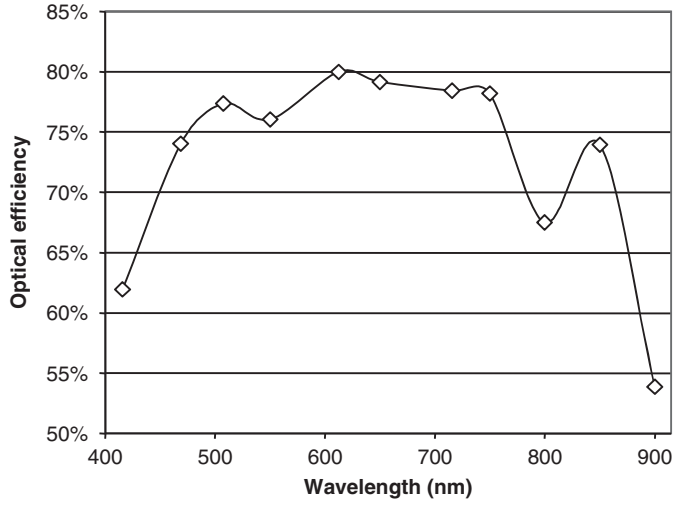


Figure 10.17 Spectral optical efficiency of a system made up of an acrylic Fresnel lens and a rotational dielectric CPC secondary on top of an AsGa cell at 600×. The valley in the curve near 800 nm. is a result of the absorption of the PMMA, which demonstrates the modification of the original solar spectrum. (Courtesy UPM-IES)

do not impinge on it. On the other hand from a certain cell size we do not gather much light, even though we increase its area. The relation of η_{op} in accordance with the size of the cells is called "*encircled power*". It allows us to calculate the economic interest of the size of the cell on the nominal power of the system.

The values of the optical efficiency of the normal systems are not usually more than 85%. It is a fundamental parameter because it influences both the efficiency of the cell and the cost of the system. The η_{op} values falls as the number of interfaces that the light has to cross is increased.[1]

10.5.2 Distribution or Profile of the Light on the Receptor

The concentrator optical elements do not usually produce a homogeneous "spot" of light on the surface of the receptor: the most usual is to have profiles such as that in Figure 10.18. With special designs and free form optics it is possible to reduce the nonuniformity.

As the total luminous power is usually measured by means of the short-circuit current of the solar cell or a thermal receptor, we only know the total power on the sensor area, not its distribution: Thus the measured effective concentration is therefore an "average concentration". However we should know the *Profile* (x, y) function to express the distribution of the irradiance on a plane.

The lack of uniformity gives rise to a reduction in the efficiency in respect to that obtained with homogeneous light for which the cell is usually designed. In many practical cases the real profile can be approximated by means of a "pillbox" profile, which produces similar results and is easier to analyse [22] (Figure 10.19).

If we call U the ratio of the maximum irradiance on the cell versus the average irradiance it can be shown that the loss of power fits well with the substitution of the cell series resistance R_s for a new $R'_s \approx U^* R_s$ in the cell performance models. In Figure 10.19 we can see the classic curve of the variation of the efficiency versus concentration for different irradiance distributions at constant current. The slope of the curves in the decreasing range is actually proportional to the effective R_s.

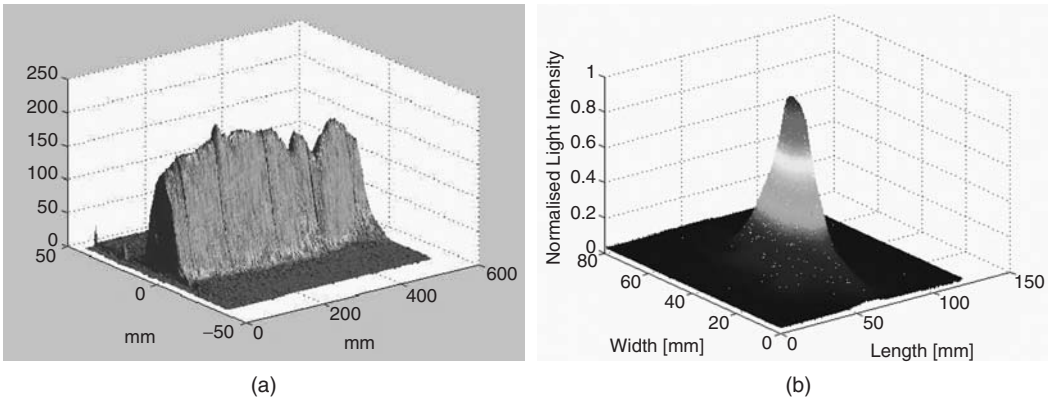


Figure 10.18 Profiles in the focus of two different concentrators. (a) Linear parabolic mirror. (b) rotational parabolic disk (Courtesy UPM-IES)

[1] It must be remembered that, ideally, for each air–glass interface that the light crosses, the power is reduced to 96%, the reflections on the aluminum are reduced by 85% and with silver by 92%.

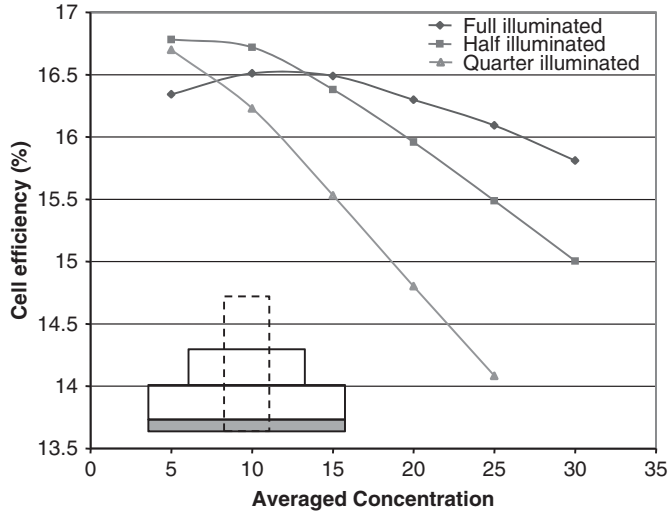


Figure 10.19 Variation of linear focus cells efficiency (experimental) under concentration for uniform and nonuniform illumination. In the experiment the short-circuit current is constant, only the distribution of the light is varied with opaque masks, not the overall power cast on the cell. (Courtesy UPM-IES)

10.5.3 Angular Acceptance and Transfer Function

The angular acceptance or angle of acceptance of a concentration system is possibly the most important quality after the optical and cell conversion efficiencies. It indicates the angle of deviation at which the power output decreases by a given value, with respect to the optimum orientation yielding the maximum power. The values in the efficiency or that of the output power for each orientation value is usually called the *transfer function* (Figure 10.20).

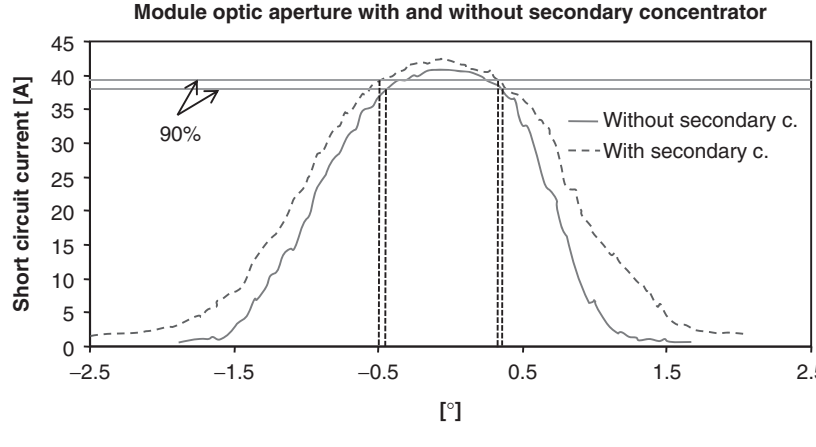


Figure 10.20 Transfer curve (experimental) of a Euclides module (linear) with and without a V-type metallic secondary lens. The angles of acceptance has been marked in. (Courtesy UPM-IES)

The previous formulas of the ideal concentrator (Equations 10.6 and 10.8) and that of the parabolic disk (Equation 10.3) specified the maximum theoretical angle of acceptance for a given concentration under the condition that no ray is lost. In practice and bearing in mind that the distribution of the light is not uniform, the angular acceptance is usually defined as the angle, within a meridian plane, for which the output electrical power falls to 90% in respect to the maximum. It is a relative measurement and therefore easy to take.

Since the sun has $\pm 0.27^\circ$ of angular width it could be sufficient that the concentrators have this angular acceptance in order to see all of the rays coming from the solar disk. However, in engineering practice a certain margin of angular acceptance has to be taken into account in order to be able to compensate the effects resulting from the lack of flatness of the supports, deformation from the loads and tracking errors.

When an individual module or a solar cell and its optics is analysed, the transfer function can be obtained from the register of the short-circuit current for different pointing angles. However in systems with many modules or series-connected cells and with bypass diodes the measurement with I_{sc} is not valid since in this case we roughly register the current of the best illuminated cell for each turning angle.

Figure 10.21 shows a real measurement on a linear concentrator made up of 69 modules (EUCLIDES, single axis, N/S-oriented) each one illuminated with a cylindrical parabolic mirror. The mirrors are not all perfectly aligned for reasons of structure and construction [23].

Curves A and C are those corresponding to the two mirrors most turned to the East and most turned to the West. The register with I_{sc}, as a result of the 69 bypass diodes, gives us a very wide transfer curve, which is that surrounding 69 individual transference curves.

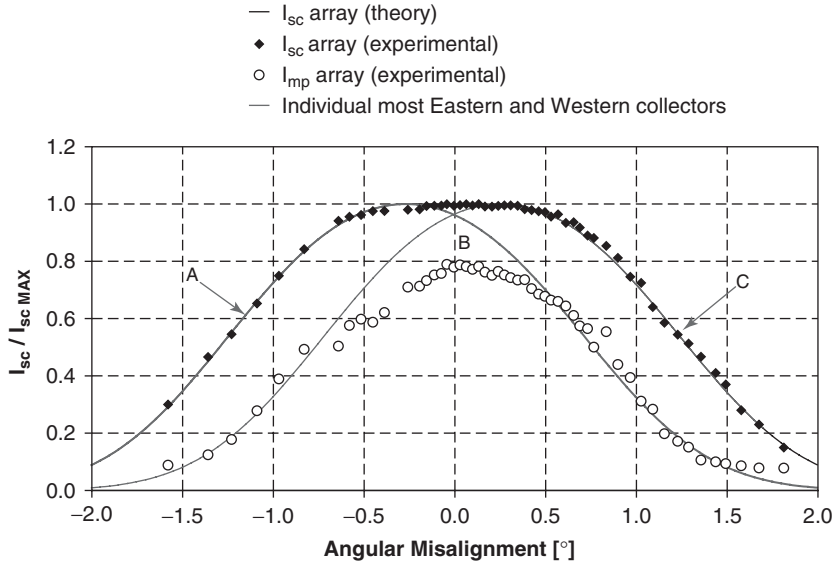


Figure 10.21 Transference curve of a linear CPV concentrator consisting of 69 modules series-connected, each with a bypass diode. The top, wider curve was obtained from I_{sc} recording. The lower curve (small circles) is the I_{mp} recording versus the deviation angle. The second one gives the true system performance. (IES-UPM, Spain, 2009)

The angle of acceptance that we could obtain from this "I_{sc} based" curve is much greater than that of an individual module, and it would therefore be absurd to believe that the worse the single collectors are aligned, the better the functioning of the system. If we have been paying attention to the current at the maximum power point, I_{np}, we would have obtained approximately the curve defined by the intersection of A and C with a maximum in B. The real test is consistent with the reduction of power as defined by the experimental circular dots. The acceptance angle becomes only $\pm 0.2^{\circ}$ at 90%. In the point-focus systems the same thing happens in two angular directions.

In the $I-V$ curves, these misalignment effects or in other words, the reduction in the angle of acceptance are shown as a parallel pseudo-resistance which determines that I_{mpp} is much less than I_{sc}.

Figure 10.21 describes a poor module alignment, far from current commercial products.

10.6 PHOTOVOLTAIC CONCENTRATION MODULES AND ASSEMBLIES

10.6.1 Definitions

The first regulation document for the qualification of concentrators, IEC-62108, distinguish between "modules" and "assemblies". The definitions transcribed in this document are:

> ***Concentrator module***: *A group of receivers, optics, and other related components, such as interconnection and mounting, that accepts unconcentrated sunlight. All of the above components are usually prefabricated as one unit, and the focus point is not field adjustable* [24].

The nominal power and thermal properties are well defined when exiting the factory. The modules are usually subsystems designed to operate using passive refrigeration (Figures 10.4 and 10.7).

> ***Concentrator assembly***: *A group of receivers, optics, and other related components, such as interconnection and mounting, that accepts unconcentrated sunlight. All of the aforementioned components would usually be shipped separately and need some field installation, and the focus point is field adjustable* [24].

The receptor can have secondary optics or not, but it always needs special refrigerating devices (passive fins or active cooling). The module power rating only makes sense if performed in the field, for the whole array. The most usual examples of assemblies are shown in Figures 10.1b, 10.3 and 10.8.

The majority of the modern commercial "modules" have been made with Fresnel lenses as a collector, following the line set out by Sandía Labs. and followed in the USA by Amonix, Opel Solar, Emcore, etc. and in Europe by Guascor-Foton, Concentrix, Sol3g, Isofoton as well as Arima Eco in Taiwan and Daido Steel in Japan.

Only Solfocus is currently commercializing a module based on reflexive collectors (mirrors) which are effectively measured in the factory (Figure 10.22). The modules with reflectors can operate efficiently at shorter focal distances than the lenses, which is why they can be three times more compact than lenses, on average.

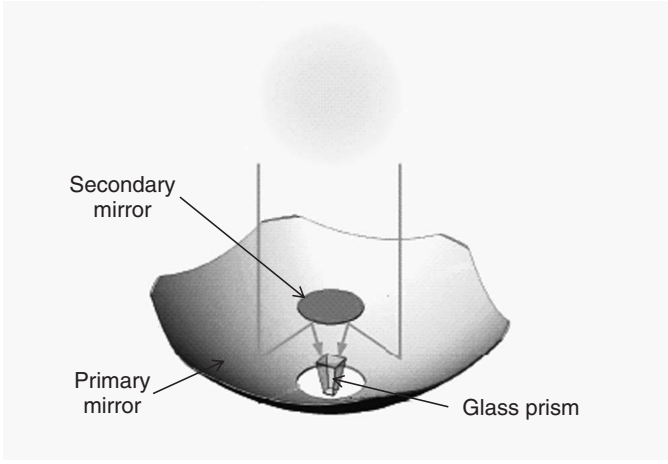


Figure 10.22 The Solfocus concentrator (USA) used initially hexagonal glass reflectors in a Cassegrain structure (two mirrors) and a glass truncated pyramid coupled to multijunction cells: scheme of the principle of the concentrator (Courtesy Solfocus)

10.6.2 Functions and Characteristics of Concentration Modules

The CPV module must mainly provide: optics, housing, electrical insulation, thermal dissipation and cell interconnection. In this section we will consider modules based on Fresnel lenses.

One function of the module is to fix and support the Fresnel lenses in position and correct distance to the cell, without forcing deformation as a result of thermal changes, as well as keeping it waterproof. The series of lenses can be made up of a single piece parquet of several elements made through the addition of individual lenses.

The cells are placed on the focal plane of the lenses and their position in relation to their optical axis is critical. The different expansion of the lenses and substrate can give rise to misalignments between them. That is why seals made from soft materials are advisable in order to eliminate mechanical stresses.

There are interesting solutions in using glass on both the upper and lower faces. Since it is not possible to make quality and high-concentration lenses from glass these lenses are hybrid, that is, they are made up of a fine layer of "plastic polymer" on which the teeth of the lens are engraved or molded, and to which the glass substrate is adhered. Transparent silicone is a suitable material because of its flexible nature and stability. This solution was patented and published [25, 26] by one of the authors in 1979 and is currently being used by Concentrix having been manufactured by Reflexite.

The optical efficiency of the Fresnel lenses reduces with the f-number, in such a way that it is unusual to find modules with Fresnel lenses for $f < 1$. When it is required to make a thin module, it is necessary to use small lenses and for $f = 1$ the cell size becomes approximately linked to the lens area and to the concentrator gain by the relation

$$\text{Cell size}(\text{cm}^2) = \frac{L^2}{C_g} \cong \frac{F^2}{2C_g} \tag{10.19}$$

That is, that the size of the cells will vary with the square of the focal distance, for a given C_g. In practice C_g is fixed by the price of the cells.

The plane of the receptors is placed parallel to the plane of the lenses: This is in charge of dissipating the heat away from the module, to the exterior air. It is usually made of metal and of a thickness sufficient to achieve the spreading the heat from the cell towards the area surrounding the substrate. Sometimes fins are added close to the cell for an easier exchange of heat. The thermal connection of the cells with the backing plate must be good, yet maintain the regulatory electrical insulation.

The electrical insulation in CPV modules had not been evaluated sufficiently at the beginning because the example of the flat panels. These, made of glass, EVA, Tedlar, etc., all of the insulating materials and without humid air on the inside, have been fairly immune from this problem. But in conventional CPV, since the receivers are surrounded by conducting materials, leaks can appear through disruption or conduction resulting from a manufacturing error or induced later through thermal cycling. For any of these reasons, the high-concentration systems, that could pay for high-quality substrates and which reduce the receptive area, use the solution shown in Figure 10.11.

This consists of insulating plate made of sintered materials such as alumina or AlN that combines its excellent electrical resistance with a high thermal conductivity. Although aluminum nitride is better, alumina provides sufficient thermal conductivity at a much lower price: commonly alumina of only 0.3 mm thickness is used metallised with thick copper of similar thickness (Direct Bond Copper plate).

The soldering of the cell to the substrate is usually done with a conducting paste (silver or copper epoxy), in order to avoid thermal–mechanical stresses and even the degradation of the cell when exceeding 220 °C which requires the use of soft solder.

The latter is favourable in respect to the series resistance, but dangerous as a result of thermal shock. The viability of the chip-cell solder with pastes is guaranteed using tests that give rise to specific temperature cycles between the cells and the substrate while the entire unit is kept at a high temperature. There are only a few references in respect to the behavior of these tests that are included in the IEC62108 Regulation.

10.6.3 Electrical Connection of Cells in the Module

The series resistance of good concentrator cells operating at nominal concentration must fulfill the condition $kT \geq R_s I_{sc}(C)$. For silicon cells operating at 30 W/cm^2, this condition demands a specific resistance of only 2.5 mΩ/cm^2. If we reach this excellent level we would only lose 5% in respect to the ideal cell with $R_s = 0$.

The connections and wires between the cells should not introduce extra significant losses. For example, if we want to have only 0.25 mΩ more for a 1 cm^2 silicon cell and the lenses are 18 × 18 cm, we must put copper wires as thick as 4 mm diameter (or a 13 mm^2 section).

Fortunately multijunction cells are easing these needs for thick wires. If we approximate, somewhat roughly, that in a cell of N junctions the current is N times less than a single-junction cell and its voltage is N times greater, we can state that the series resistance requirements to have the same relative losses will be approximately:

$$R_s(N) \approx N^2 R_s(1) \tag{10.20}$$

That is, we can relax the series resistance design of the multijunction cells up to N^2 times the series resistance of a single junction cell, $R_s(1)$, for the same concentration level. In summary,

we could establish that as a minimum the losses resulting from series resistance effects be 10% in CPV systems.

A bypass diode in parallel with each cell is usually mounted on the DCB plate to protect the cell in case of severe current mismatch.

10.6.4 Thermal–Mechanical Effects Related to Cell Fixing

The size of the concentration cells plays an important role in relation to fixing and cooling. As we have already seen in the previous section high-concentration cells have to be small for reasons of series resistance, as well as for thermal reasons.

As the cells are metallised on the rear surface it is more convenient from the thermal and electric point of view that they be soldered to a metal substrate plated with copper or nickel. A model that assumes a perfect soldering between the cell and the substrate indicates that the stress and the curvature of the whole unit as it cools is independent of the size of the cell and only depends on the thicknesses of the two elements. Cells should break according to this model for copper plates 2 mm thick. Practical experience does not agree with this result. But if we adopt a *slip model* in the interface between both materials (taking into account the plasticity of the solder) then the resulting tensions are proportional to the square of the side of the cell, which are well tuned to the lesser problems found in smaller cells in respect to the thermal stress. All together, the thermal–mechanical analysis demonstrates that the solder material would be working in a plastic mode in cell sizes greater than 5 mm. The repetition of plastic deformations should be rejected as it ends up giving rise to breaks. Therefore it would be better to carry out suitable qualification tests that show the real effects experimentally.

The glues that conduct both heat and electricity fall into the group of materials that are going to work at the elastic–plastic limit in accordance with the size of the cells: The use of glues has the advantage of avoiding the cooling from the liquid phase to the solid one typical of solders, and only has to tolerate the thermal operation cycling in the field. As a result the glues that work elastically are the safest.

From this discussion we can conclude that the encapsulation of the large cells, those that work at medium or low concentration (several centimeters long and wide), pose a difficult economic and technical problem: on the one hand the cost of the metallised insulating substrate (DCB) is prohibitive while the use of insulating adhesive tape for both faces has shown to be not very reliable and a poor heat conductor [27].

Table 10.3 shows several materials used or proposed for the fixing and insulation of small and large concentration cells. Likewise, the thermal drop that is brought about between the cell and the heat sink for several degrees of concentration is shown.

Thus we see that for 500× it will be difficult to reduce it to less than 15 °C, even using the best substrates; for large cells, even though the irradiance is less we will have to use lower quality products for reasons of cost.

10.6.4.1 Extraction of heat in concentration solar cells

The thermal conductivity of single-crystal semiconductors is high, approximately a third that of good conductors. That is why the cells, mounted on a substrate that allows the evacuation of heat, do not overheat in spite of almost all of the power being adsorbed in the first few micrometers of the device.

Table 10.3 Thermal properties of the most used insulators and conductors for CPV receivers

Materials	σ W/°C cm	R_{th} @ 0.6 mm. thick(°C cm^2 W^{-1})	Drop @ 100× (°C)	Drop @ 500× (°C)
BeO	2.2	2.73×10^{-2}	0.27	1.35
NAl	1.8	3.33×10^{-2}	0.37	1.85
Al_2O_3	0.21	2.86×10^{-1}	2.9	14.5
Copper	4.01	1.50×10^{-2}	0.15	0.75
Aluminum	2.37	2.53×10^{-2}	0.25	1.25
Glass	1.0	6.00×10^{-2}	0.6	3.0
Silicon	0.9	6.67×10^{-2}	0.68	3.4
IMS substrate		1.67 (thick)	16.7	83.5
"Thermattach" film		4–6 (thick)	40–60	200–300

It is known that the exchange of heat between the surface of a body of area A and the air is established through the Newton heat equation,

$$P = A.h.\Delta T \tag{10.21}$$

where h is a function of the fluid, of the surface of the interface, of the wind, etc. The value of h is calculated by means of the dynamic theory of fluids and the heat exchange equations. For practical effects we can establish the figure of $h = 5\text{–}10 \times 10^{-4}\,\mathrm{W\,cm^{-2}\,^{\circ}C^{-1}}$ for natural convection from a plate to air.

With this value, and assuming that a concentrator module has a rear heat exchange surface equal to the entry surface, we can estimate the increase in the temperature expected in the interface thus:

$$\Delta T = \frac{P_{in}}{hA} = B(n) \cdot \eta_{op} \cdot \frac{A}{hA} \geq \frac{850 \cdot 0,80}{10} = 68\,^{\circ}\mathrm{C} \tag{10.22}$$

As it has been demonstrated experimentally [28] that 25% of the heat escapes through the front and side faces we can estimate that the real drop at the back plate to air interface will be equal or larger than $0.75 \times 68\,^{\circ}\mathrm{C} = 51\,^{\circ}\mathrm{C}$ in calm air.

In the operation of the concentrators the refrigerating contribution of the wind is very important. In effect, with barely 2 m/s on the rear face the value of h is doubled and thus the plate will be about 26 °C over the ambient temperature. The reader will discover, through the simplicity of the calculations, that these are approximate figures, but they give a fast estimate of the back plate temperature.

Compared with the temperature drop between the layer and the air, especially in the absence of wind, we can see (Table 10.3) that the thermal leap between the cell and the layer is small compared with the former; therefore little money should be spent on reducing this fall by just a few degrees, but it is better to concentrate on achieving good insulation and reliability in the fixing or soldering.

The temperature of the cell is established through a simple equation

$$T_{cell} = \left(R_{thspread} + \frac{1}{A_{hs} \cdot h} \right) B(n)\eta_{op} \cdot A_i + T_{amb} \tag{10.23}$$

where h is a function of the speed of the wind and $R_{thspread}$ (°C/W) is a function of the geometry of the heat distributor and the adhesives. A_i and A_{hs} are the areas of light entry and heat interchange with air respectively.

10.6.4.2 Thermal resistance of the heat distributor

The distribution of heat between the cell or the insulating substrate into the air must be carried out by means of the back plate of the modules. If we want the entire rear surface plate to be equally effective we must make this layer practically isothermal. This solution may demand both excessive thickness and cost. To calculate approximately the thickness of the layer that would turn out to be effective, and its relationship with the average temperature of the layer, could lead to the solution of the equation of the heat for the case of a circular layer of r_c in radius and thickness w [29].

First we see that the plate of thickness w is not effective from a radius larger than

$$r_{\max} = 0,66\sqrt{\frac{w \cdot \sigma}{h}} = \frac{0.66}{\alpha} \tag{10.24}$$

The temperature between the cell or receiver bottom and the coldest point of the sheet is:

$$T_{cell} - T_{plate} \cong B(n)\eta_{op}\frac{A_i}{\pi w\sigma}\ln(\alpha R_c) \tag{10.25}$$

where σ is the thermal conductivity of the plate material (W °C^{-1} cm $^{-1}$) and R_C the cell radius.

It turns out that the smaller-size cell is recommendable, but the size of the collecting lens even more so: systems with large lenses will need a greater mass (greater w) in the back plate than those with small lenses for the same heat drop.

If we decide to use only a plate (that is, with no additional finned device) for heat dissipation and that the collector area A_i is equal to the effective thermal area of the plate we obtain, by means of $A_i = \pi {r_{max}}^2$, the temperature drop across the distributor as:

$$T_{cell} - T_{plate} \cong 0.44B(n)\eta_{op}\frac{\ln(\alpha R_C)}{h} \tag{10.26}$$

where the most significant parameters are h and the cell radius. (In this case the h should include the heat radiation from the plate to the surrounding).

10.6.5 Description and Manufacturing Issues of Concentration Modules

The main objectives consist of guaranteeing the watertightness, the electrical insulation and the alignment of all of the cells in the focus of all of the lenses. Ultimately, the cells to the substrate cell unit must be fixed using a uniform thermal contact and a guaranteed durability for many cycles during its lifetime (it can be estimate at some 100 000–200 000 rapid variations of around 15 °C in the lifetime of the system).

In those models with Fresnel lenses, they are usually made up of a parquet with several elements: the receptors are placed on a substrate sheet with high precision using robots. If we keep the base plate fixed, the arm of the robot has to move and carry out many movements as the cells are assembled. So if the cells are small it is necessary to place a greater number of them, and move the robot more quickly. The precision in placing the cells must also be greater as they are

smaller. In general it is necessary to have an optical control for the alignment between the parquet of the lenses and the cells in the substrate. That is why these processes with refractive or reflexive parquets, can have advantages in the manufacturing process and quality in respect to the methods that use isolated individual lenses.

In spite of the assembling complexity, the cost of the module is clearly determined by the cost of the materials and components. Although not very sophisticated, the cost of the aluminum of the rear layer is significant, and allows alternative solutions, like glass, to become cost competitive.

10.6.6 Adoption of Secondary Optics

In Section 10.4 we can see that the use of secondary optics is almost essential with the lenses if we want to have a reasonable angular acceptance and with it tolerable mismatches. We also see that the dielectric secondaries have optical advantages in respect to those which are purely reflexive: however, the integration of these secondary lenses gives rise to some problems. For example, the linking of the dielectrics to the cell must be done using a transparent glue capable of tolerating high densities of radiation: The non-dielectric optical devices are easy to assemble and more reliable over time,

Although systems with optimum dielectric secondary optics have not yet been industrialised, an efficiency of 80% effective at 650 suns has however been demonstrated with an experimental system called RESET from Isofotón that mounted a secondary DCPC-type dielectric and in FLATCON modules, by Fraunhofer-ISE, with quasi-optimum dielectric lenses.

Clear glass is an ideal material for the secondary optics although its price has been a limiting factor up until now. The truncated pyramid made of glass that provides a 2–4 concentration ratio and kaleidoscope effects is the most used dielectric secondary in current commercial modules.

Transparent silicone to make secondaries of any shape (Domo, Silo, CPC, etc.) is profiled as a suitable industrial alternative for the modules with Fresnel lenses.

In summary, it can be said that dielectric secondaries could be suitable for small cells as the volume of the material required varies with the cube of the diameters of the cells.

10.6.7 Modules with Reflexive Elements (Mirrors)

Modules based on concentrator mirrors must, in principle, be better than the lenses because they can be more compact.

The commercial module from Solfocus is an example of this technology (Figure 10.22). The primary mirror and the entire unit are protected by a glass surface. With this, 8% of the light is lost through Fresnel reflection. In order to have cells on a rear dissipation plate a second mirror is placed, in a Cassegrain configuration, so as to divert the rays upwards.

A maximum transmission of 84% is achieved with two reflective surfaces. In order to increase the angular acceptance and make the light on the cell uniform, a truncated pyramid dielectric is available which works through total internal reflection. The pyramid is joined to the cell optically using a transparent elastic adhesive. The system is very compact and has almost 1.2° angular acceptance and its theoretical optical efficiency is 74% before entering the cell.

A similar shape was developed by Miñano *et al*. with a very compact concentrator, but designed using the "successive surfaces method" and which is shown in Figure 10.23. In this

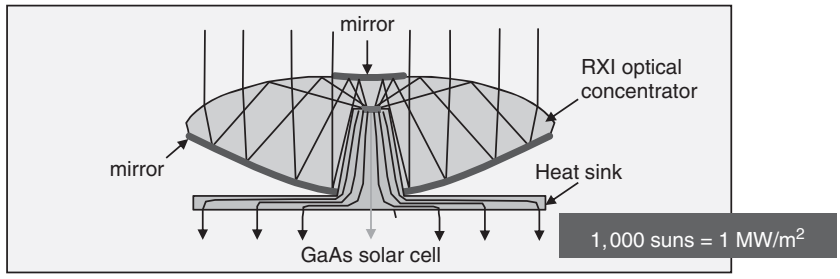
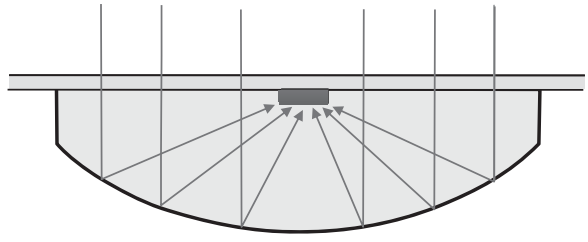


Figure 10.23 Ultra-compact RXI concentrator providing 1.2° acceptance angle for 1000× concentration. (Miñano, IES-UPM, 1994)



Concentration	
500X	±1.47°
1000X	±1.04°
1700X	±0.8°

Figure 10.24 The Fluidrefex CPV concept proposed by Sala, Antón, IES-UPM, 2007. Optical index, cooling and insulation are provided by an inert transparent fluid. There is only one interface with Fresnel loses. (Courtesy of IES-UPM)

design refraction, reflection and TIR combines to reach very high concentration in a very compact optics [30].

Recently, Light Prescriptions Innovations (LPI) designed a reflexive free-form concentrator (that is not rotational, but still point-focus) for Boeing with special emphasis on the uniformity of the illumination on the cells. It uses a dielectric secondary, but pursues refrigeration on one face as the majority of CPVs [31]. One recent proposal from IES-UPM uses a fluid both as a refrigerator and index material. It has only one Fresnel interface reflection. The overall optical efficiency is 82%. It has the advantage of avoiding condensations and does not require a heat distributor since it uses both sides to dissipate the heat (Figure 10.24) [32].

10.6.8 Description and Manufacturing Issues of Concentrators Based on Assemblies

In these concentrators the main technological problems centre on the receptor. In general, large systems have always been made and therefore the construction of a quality optical system has ended up becoming a challenge.

An example of an assembly has been the EUCLIDES technology that installed 450 kW_p in Tenerife in 1998. It was seen that even though individual mirrors with an almost perfect profile were achieved, the assembly and alignment of these collectors gave rise to losses as a result of optical mismatch (Figure 10.21).

The receptors of the assemblies need special refrigerating elements as the surface in contact with the air is small. In the EUCLIDES a heat sink with many fins of less than 1 mm is necessary for the heat exchange into the air. An alternative to this solution in linear systems is active cooling with the circulation of a fluid, such as that used in the CHAPS system which is capable of saving a lot of aluminum with an active heat removal, and to use the heat for preheating domestic hot water, if it is useful in the location [33].

The case of low concentration, that is, for gain levels of up to 10×, the determining requirement of the optics and thermal are fairly relaxed and could give rise, in spite of the low efficiency (≤22%) in the cells used, to efficient and possibly competitive systems. This is the case for Archimedes Systems that locate the receiving cell on the back side of the contiguous metallic mirror which acts as heat dissipater.

Skyline (CA., USA) follows the Euclides steps by relaxing the gain from 32× in the former to 6× in the later (Figure 10.25).

The proposal of the assembly of the Australian company Solar Systems is very different (Figure 10.3). Here they use mult-junction cells joined in a large receiving cell parquet of approximately 55 × 55 cm^2 [34]. The mirror is a parabolic disk made of independent facets that concentrates up to 500×. The receptor is actively refrigerated using a permanently pumped fluid and has an aiming-off system for protection if the pump fails. It does not consume water, and if the climate requires it, one of each eight disks is used to refrigerate the remaining receptors of the plant with a heat pump fed using solar energy.

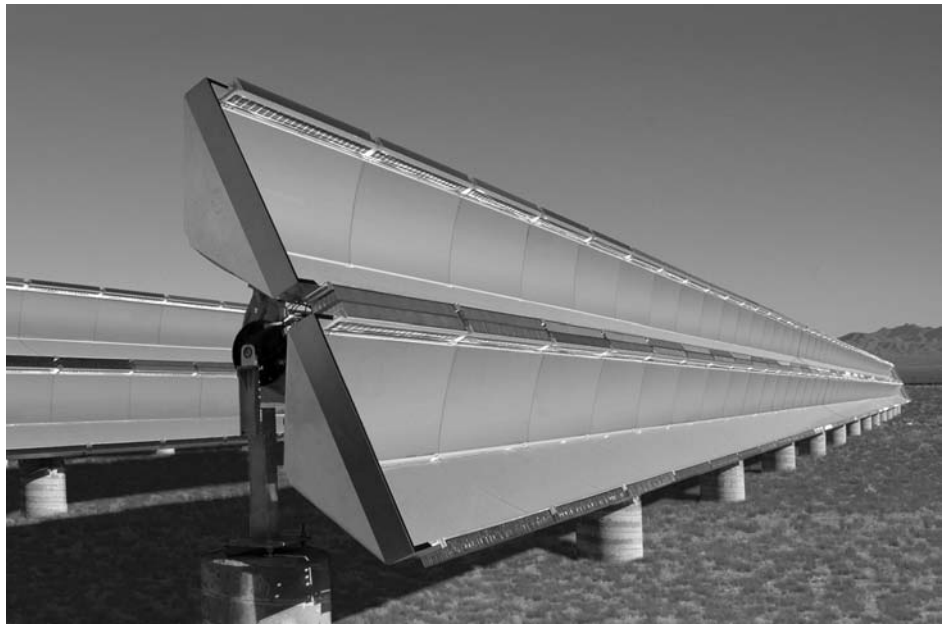


Figure 10.25 The Skyline trough operating at 6×. (Courtesy Skyline Solar, Inc)

CPV assemblies seem to have a significant advantage with respect to modules by means of its hybrid thermal–PV functioning. However, it is greatly apparent because it involves heat at low temperature and it is necessary to have a suitable application *in situ* to take advantage of this thermal contribution which is not very common in high-power multi-MW plants connected to the utility grid.

10.7 TRACKING FOR CONCENTRATOR SYSTEMS

The limited acceptance angle of concentrator systems requires the collectors to point toward the sun disk, and consequently, to vary its position constantly. For concentrators, the accuracy of the pointing combined with the module parallelism must be less than a tenth of a degree, very far from the less demanding $\pm 5^{\circ}$ required to tracking structures for flat plate power plant Figure 10.26 details the real duty of a tracker control system combined with the mechanical fixtures and the module acceptance angle.

The real issue in a CPV tracker is not to point system drivers at the sun all the time, but to allow the sun's rays to reach all the cells in the array equally. This condition substantially changes the design and construction demands of the structure which must be very rigid to restrict further bending of a part of the module acceptance angle. Such requirements state that the tracker accounts for a significant cost of the CPV system and is currently slowing down the cost reduction promised by CPVs.

Both the necessary rigidity and the current cost of steel determine the optimum size range of the CPV trackers: around 30–90 m^2 in the collector area, which at a 24% array efficiency, yields 6–18 kW_p arrays.

10.7.1 Tracking Strategies for CPVs

Point-focus CPV systems require two-axis tracking while linear focus ones need a single axis if the collectors are reflectors and a double axis if they are lenses. The common tracking driving strategies are shown in Figure 10.27.

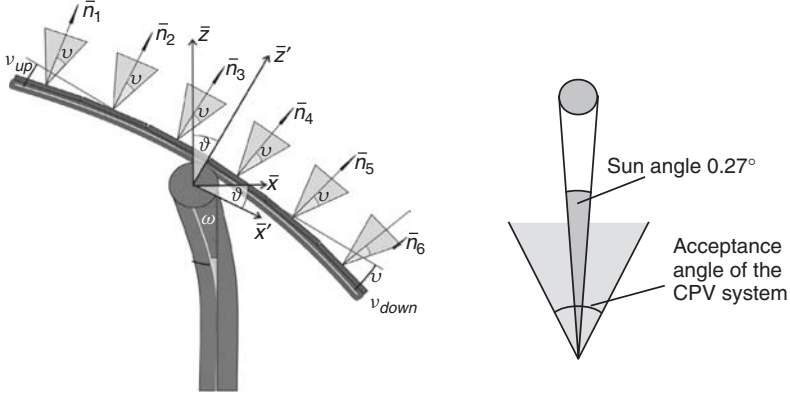


Figure 10.26 The value of the acceptance angle defines the strength and driving accuracy requirements of a tracking system. This picture shows the main sources of error during real tracking: the purpose of tracking is the uniform overall illumination of the cell array, not just to maintain the vector z pointing at the sun. (Courtesy of BSQ Solar)

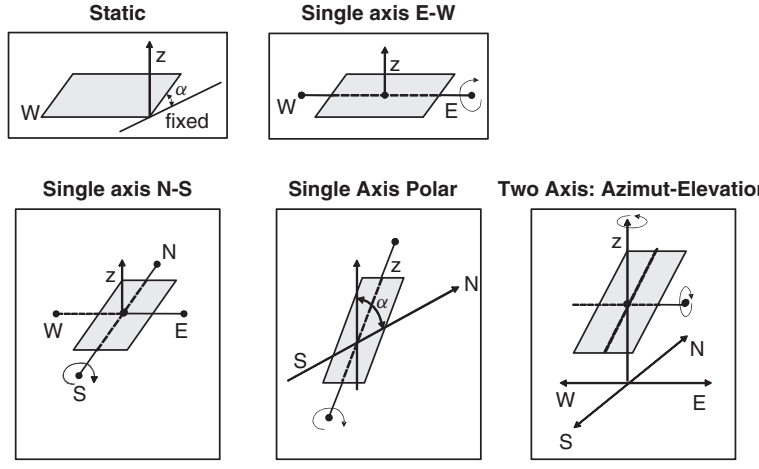


Figure 10.27 The most common strategies to track the sun's path. The fixed position is a reference, but can be identified as the orientation for static concentrators. The one-axis tracking at constant elevation is missing here because it is common in flat plate, but not in CPV (Courtesy UPM-IES)

10.7.1.1 Two-axis tracking systems

Among the two-axis tracking systems we find the:

- Pedestal type with azimuth and elevation driving (Figure 10.2b). It is the most commonly used.
- Turntable type with azimuth and elevation driving. Theoretically the previous type requires more steel than the turntable because the loads are centered on the centre of the pedestal. The turntable has many supporting points, thus the steel bars can be thin and light. But the latter demands a large and well-constructed horizontal foundation while the pedestal only needs a hole surrounded by concrete around the vertical tube. The differences in cost and accuracy have not yet been stated (Figure 10.28).
- End post two-axis tracker. Used originally by Entech for linear lenses systems, and currently for cylindrical mirrors by CHAPS system (ANU).

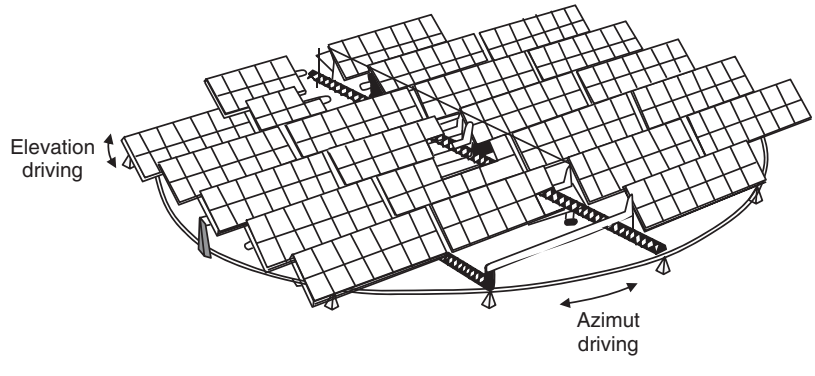


Figure 10.28 The turntable tracker concept. The load is shared between many vertical supports. It reduces the amount of steel necessary, but the basement is more requiring than the pedestal option. The controversy about the optimum solution last from 30 years ago (Courtesy of Sandia Labs)

- Polar axis plus declination driver: This is not used for large systems, but mainly for testing. It has the advantage of simplicity. The main drive is just a "clock", providing constant rpm turning.

10.7.1.2 One-axis tracking systems

These are only suitable for reflective collectors (see Figures 10.1b and 10.25. The only condition for correct operation is to maintain the plane normal to the aperture which includes the axis containing the "sun dish". The N/S horizontal axis tracker collects more energy than E/W axis, The values of the direct radiation collected by different concentrators with different tracking philosophies are presented in Table 10.4 as well as the annual value of the overall irradiance on a surface with tracking in kW h/m^2 in Madrid [35].

10.7.2 Practical Implementation of Tracking Systems

The key elements and the figure of merit of CPV tracking system mechanics are the maximum coupling forces of gears and their limited backlash. The modern fashion for pedestal-type arrays consists of installing a standard screw or hydraulic linear actuator for elevation angle movement. A hydraulic system is preferred in very large arrays (Amonix, Guascor-Foton, 200 m^2, see Figure 11.2) where large forces from gusts of wind are better absorbed and because backlash is avoided.

The azimuth drive is usually carried out by means of motors which operate at a great gear reduction ratio, usually not less than one thousand times. With this degree of gear reduction the errors in the main axis become negligible, but the remaining error is the backlash of the last gear.

With the aforementioned reduction ratios it is evident that the energy consumption of the trackers, which are usually well balanced in respect to the force of gravity, is quite small. In fact the power exerted on the friction of the driving elements is less than 1%, An exception is the hydraulic systems which has a consumption of up to 3–5% of the production.

Table 10.4 Average annual and monthly direct radiation

Month	Two-axis	N–S axis	Polar axis	E–W axis
January	2.31	1.46	2.16	1.93
February	3.23	2.36	3.15	2.56
March	4.24	3.61	4.23	2.95
April	4.73	4.41	4.65	3.03
May	5.95	5.74	5.62	3.74
June	6.98	6.76	6.42	4.37
July	8.61	8.34	8.03	5.46
August	7.89	7.49	7.66	4.83
September	6.38	5.67	6.36	4.20
October	4.51	3.52	4.45	3.39
November	3.26	2.14	3.09	2.68
December	3.04	1.81	2.80	2.60
Direct annual	5.15	4.46	4.90	3.48
Overall Annual	7.08	6.24	6.87	5.61

For its low concentration, Archimedes Systems ZSW succeeded in developing and operating a hydraulic passively powered driver which uses the sun rays both as tracking sensor and power source. [36].

10.7.3 Tracking Control System

This is currently a microcomputer system, which controls the activation of the driving AC or DC motors. This must provide displacement in two opposite directions, in order to restore the morning position of the tracking structure. The movement of each axis is limited both by software and by mechanical sensors at the extremes of the allowed path for redundancy.

10.7.4 Pointing Strategies

There are two basic strategies for steering the tracker perpendicular to the sun's rays: One is based on the direct vision of the sun whilst the other consists of using the astronomical ephemeris of the sun.

10.7.4.1 Characteristics of direct sun vision sensors

The classic sensor follows the concept detailed in Figure 10.29 if the sensor is not pointing at the sun and electrical error signal is generated from the difference of the currents generated by the two light sensors. The signal is used to activate the driver continuously or at time intervals.

This strategy works well on clear days when the sensor is closely aligned to the sun. However additional computing is necessary to guide the sensor to sun vision operation mode and brings the system back to the East after sunset, The accuracy of these systems is quite good, better than 0.05°

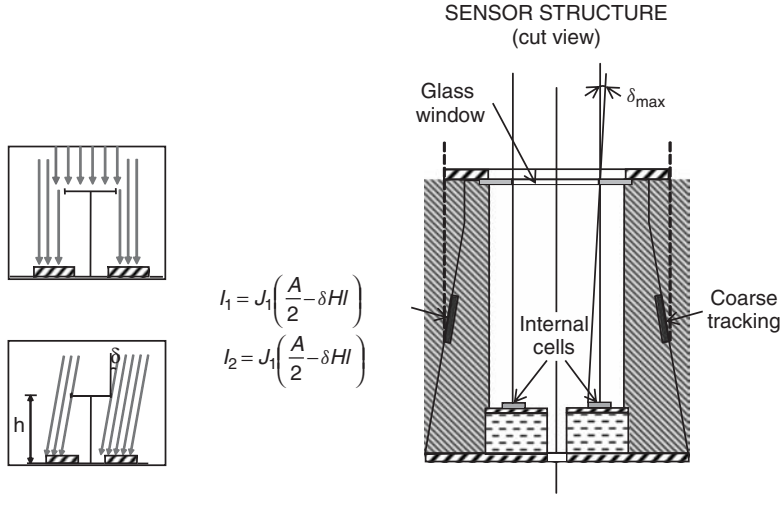


Figure 10.29 (a) Operating principle of a direct sun sensor: the error signal is the difference in the current generated by partially shaded two cells. (b) The real sensors provide coarse tracking from the lateral cells and accurate tracking with the internal ones once the sun is at less than ±10° deviation. (UPM 1980, based on 1977 Sandia Labs. prototype)

Modern sensors consist of *XY* CCD arrays which provide the direct digitalisation of the error signal and avoid the typical offset problems.

10.7.4.2 Astronomic ephemeris method

The complex path of the sun in the sky is well known and exact formulas are available to calculate the instantaneous solar vector at any moment, and at any point on the Earth. GPS systems provide accurate time and site data. The procedure is insensitive to clouds, dust, light intensity, sky, transparency, etc.

The ephemeris is calculated using a software code which also receives the absolute reference orientations of the array (the zero azimuth and the zero elevation angles, for example) as key inputs. The inaccuracies of these data inputs become the absolute pointing deviations.

One company, Inspira, a spin off from the UPM, has reduced this fixed inaccuracy by analyzing the array power output and the array pointing errors with respect to the theoretical position. [37, 38]. The error table created from that analysis allows the initial position errors to be discriminated, and generates the optimum particular ephemeris for a given array. The procedure is programmed for updating the error table and consequently the ephemeris plan. This method has produced tracking within $\pm 0.1\,^{\circ}\mathrm{C}$ for more than 95% of the working time in real arrays.

10.7.5 The Cost of Structure and Tracking Control

The supporting structure of concentrators must be designed well within the elastic strain mode in order to bend just tenths of a degree under gravitational and wind loads. This condition is much stronger than the construction regulations, even for extreme natural conditions.

As a consequence, the cost of the current tracking structures for small power plants or small production levels is usually more than $0.8\,€/\mathrm{W_p}$. Lower figures could be envisaged if design and production are merged in few large specialised companies.

10.8 MEASUREMENTS OF CELLS, MODULES AND PHOTOVOLTAIC SYSTEMS IN CONCENTRATION

The main magnitudes or aspects to be measured in CPV technology are the:

- electrical properties of the cell in nominal concentration;
- thermal resistance between the cell and the air or other heat drain;
- optical properties of a complete concentrator element (e.g. the primary collector, the secondary optics and the cell);
- electrical, optical and thermal properties of a concentration module or assembly;
- electrical, optical and thermal properties of one or several tracking arrays.

10.8.1 Measurement of Concentration Cells

It is important to take the measurements at illumination levels similar to those of nominal operation, because although, for practical purposes, the photovoltaic current is assumed to be proportional to the incident luminous power, it is evident, on the other hand, that the effects of recombination and series resistance are dependent on the irradiance intensity.

10.8.1.1 Measuring cells in concentration (without optics)

High-density illumination is easy to achieve by means of powerful flash sources that are capable of handling more than 3000 J (electrical) per burst. This can provide up to 200 W/cm^2 at 30 cm and provide a good uniformity in cells of 1–2 cm^2.

The measurement using a flash lamp requires one to obtain an $I-V$ curve in the transient regime The duration of the flash is of the order of 1–10 ms. Unfortunately, neither the intensity nor the spectrum are uniform during this time, but vary, as can be seen in Figure 10.30.

High-efficiency silicon concentration cells have a slow response as a result of the long carrier lifetime. This good quality as a solar cell could interfere with the validity of the measurement in the transient regime, in the sense that the sweeping of the $I-V$ curve obtained is not equivalent to the behavior of the cell in the permanent regime. The appearance of these problems derived from the transient regime is easily evidenced by sweeping the $I-V$ curve from $I = I_{sc}$ to $I = 0$ and later the other way round. If there are transient effects the two curves will not coincide.

It is recommended in these cases to take the measurement by means of several flashes (multi-flash mode). It consists of polarizing the cell with a voltage source from $V = 0$ to $V = V_{OC}$ in as many intervals as points they want to register and apply a flash shot for each polarisation. The load is usually a capacitor.

In order to find out the evolution of the flash of light the short-circuit current of a "fast" single junction solar cell is used and calibrated in A/W for a reference spectrum. This cell provides information on the real illumination on the cell and permits each reading to be corrected in accordance with the instantaneous irradiance that really impinges on the cell. Using a control cell of the same technology as the measured one avoids having to make spectral corrections [39].

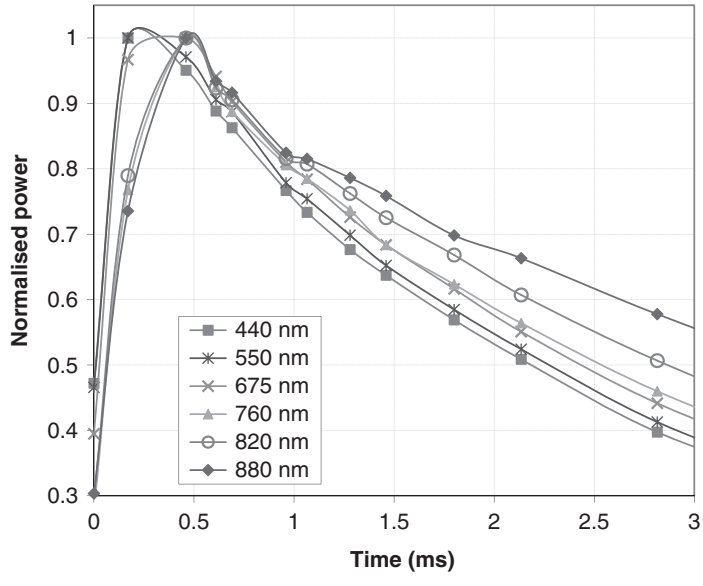


Figure 10.30 Temporary variation (experimental) of a xenon flash burst for different wavelengths: the intensity of the bluish component is extinguished more quickly than the reddish one: the spectral composition varies throughout the discharge (Courtesy UPM-IES)

10.8.1.2 Measurement of bare concentration cells or on a wafer

Probably the main requirement in the measurement of the concentration cells is the taking of low-resistance contacts in order to extract large currents. The flash illumination mode (transitory and of a few milliseconds) eliminates the thermal problem. The very small cells $\leq 10\,\text{mm}^2$ could be measured on the wafer at a great rate with fine probes, but those of the greater area need the use of a multi-contact system.

As this measurement is difficult and dangerous for the cells, it has been opted to encapsulate them first on an insulating substrate and measure the already assembled complete unit at nominal concentration, usually between 30 and $80\,\text{W/cm}^2$. In order to reduce cost in the measurement, several cells should be tested simultaneously in the same single flash. A commercial product has already been developed in the UPM that measures twelve, $1\,\text{cm}^2$ cells at $50\,\text{W/cm}^2$.

10.8.2 Measurement of Concentrator Elements and Modules

We call an elemental concentrator or concentrator unit, one that is made up of an optical element (lens) and a cell including a secondary optics, if it has one. The concentrator module was defined in Section 10.6.1.

In order to reproduce the behavior under real sun in a simple manner in the laboratory, we must have a source that is seen from the elementary CPV with the angular width of the sun.

So as to have $\pm 0.27^\circ$ of angular width we must place, a light source 5 cm in diameter, for example, at 5.3 m from the receptor. This source must provide around $1000\,\text{W/cm}^2$ at this distance and with a good uniformity on the XY plane. The required brightness of the lamp is given approximately by the equation

$$BR = 1000\ (\text{W/m}^2)/\Omega_{sun} \tag{10.27}$$

where Ω_{sun} is the solid angle at which the concentrator sees the whole source.

The value obtained for the brightness BR at the source is about $1.4.10^7$ W m^{-2} sterad $^{-1}$. It is obviously similar to the brightness on the surface of the sun, because we see the source at the same solid angle and we require the nominal solar irradiance at the surface of the Earth.

In spite of the separation of 5 m between the lamp and the receptor, some rays could reach the entry at an angle outside the angular acceptance of the optics. We can see, for example, that the angle δ with which an extreme ray reaches a lens 20 cm in diameter is

$$\delta = \arctan\frac{0.1\text{m}}{5\text{m}} \approx 1.15^\circ \tag{10.28}$$

This value is much greater than the angle of acceptance of the majority of the real elementary concentrators on the market. Consequently, it is also a problem for testing modules which are made up of several elementary units (Figure 10.31). Then the light to cast must consist of identical collimated beams at each point of the module surface. This can be achieved under real sun conditions, but testing the production outdoors is unacceptable for manufacturers. For this reason it is necessary to create a light source to illuminate in the interior just like the sun does.

A parabolic mirror of sufficient quality and low cost, capable of guaranteeing a uniform illumination equivalent to that of the sun on a large area which does not modify the lamp spectrum, is the solution currently adopted by several companies and laboratories [40]. The Fresnel lens approach, as well as being limited in size, produces an inevitable chromatic aberration. The currently

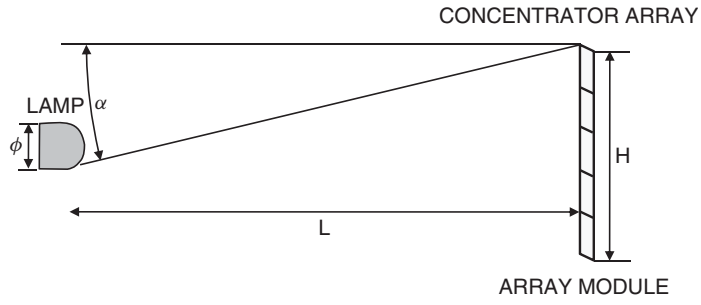


Figure 10.31 Uncollimated indoor source. A single artificial source showing the angular sun width cannot provide the same angle of incidence on all elements of a module: The angle α is much greater than the acceptance angle of the collector, and consequently many cells are not illuminated (Courtesy UPM-IES)

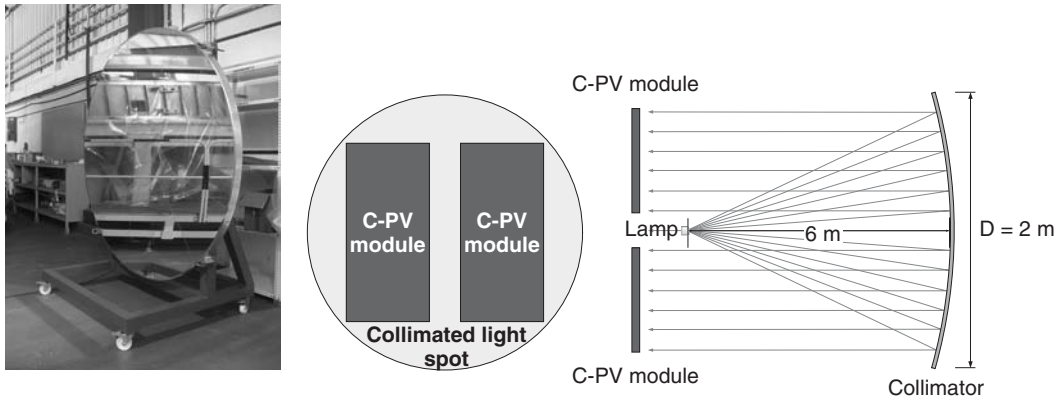


Figure 10.32 Parabolic dish that produces collimated light of $\pm 0.4^{\circ}$. This system developed by UPM at reasonable price allows indoor testing of CPV modules. It is currently a commercial system (Courtesy UPM-IES)

available set-up consists of a parabolic mirror of 6 m focal distance, in such a way that a source 6 cm in diameter located in the focus should provide a beam of $\pm\theta_{sol}$ (i.e.: near $\pm 0.27^{\circ}$) on an area the same as the 2-m-diameter mirror.

However, that reflective disk, with a lower price and quality than an astronomic telescope, produces a certain scattering in such a way that the final beam of light has an amplitude of less than $\pm 0.5^{\circ}$ which is sufficient for the measurements of the majority of the receptors (Figure 10.32). If a lesser angle is necessary, then a source with a smaller diameter will be positioned appropriately.

10.8.3 Absolute and Relative Measurements with Simulators

It is not easy to take absolute measurements of the light coming from a solar simulator and even more so its spectrum. If the light source is not continuous either, more of the flash type, it becomes still more difficult to find out its spectral composition. Modern spectroradiometers based on CCD sensors are presumably "instantaneous", but not very reliable for the very fast variation of the flash.

For this reason the most common and simple method consists of having a calibrated cell of the same technology as the cells being measured. In this way the current of the calibrated cell under the reference spectrum becomes the calibration current for any other spectrum, in particular that of the lamp source. In this case we say that the mismatch factor $M = 1$. It is advisable, however, that the lamp spectrum be continuous (without narrow emission bands) and similar to the solar spectrum.

However the cells of the module or the elementary unit are located under optics whose transmission is dependent on the wavelength. Consequently our reference for testing a CPV module under standard test conditions (STC, direct and normal, 850 W/m^2, AM1.5) must be an STC calibrated elementary unit containing identical optics and cell rather than the module The optical materials and possible dispersion effects can cause a considerable modification of the reference spectral irradiance $E_{ref}(\lambda)$ reaching the cells inside the module. We can represent this variation as $t(\lambda)$ which modifies the cell current according to the integral

$$I_L = A \int E_{ref}(\lambda) t(\lambda) S(\lambda)\, \mathrm{d}\lambda \neq I_{Lref} \tag{10.29}$$

where $S(\lambda)$ is the cell spectral response. I_L is the actual current inside the module which is different to that which a cell alone would yield the cell under the same source of light, not affected by $t(\lambda)$.

The results obtained using this process are quite satisfactory for all of these CPV modules or elementary units which are equipped with single-junction solar cells.

10.8.4 Optical Mismatch in CPV Modules and Systems

It is known that the series connection of cells or modules with a different I_{sc} for the same irradiance causes a reduction in the output power with respect to uniform current generation in the series.

If well-sorted flat modules are mounted roughly parallel on their static or tracking structures, the output is not affected by optical mismatch because the sun irradiance is naturally uniform on all them. Only shadowing caused by trees, bird droppings or other similar sources of shadowing cause "current mismatch".

On the other hand, in concentrators the light that impinges on the cells has been greatly modified by the optics, the pointing accuracy of the tracking, the secondary optics, etc. in addition to shadowing and soiling. This means that we have new sources of current nonuniformity that, we say, could cause "optical mismatch".

The causes and origins of the intensity mismatch in a CPV are so varied that the distribution of the current losses can be expected to correspond to a normal, Gaussian, distribution. The visible effect of this optical mismatch, combined with the effect of the by-pass diodes on the shape of the $I-V$ curve consists of a slope between the I_{sc} value and the maximum power point.

In Figure 10.33 it can be seen that the $I-V$ curve of a parabolic trough concentration system with mirrors and linear receptors equipped with bypass diodes, presents a pseudo-parallel resistance effect as a result of the strong optical mismatch: J.C. Arboiro and I. Antón related the pseudo-parallel resistance to the parameters of the theoretical and experimental Gaussian distribution [41]. However, it has to be said that this is a very extreme case of the defect.

As can be seen in Figure 10.22 the angular acceptance and the maximum power are reduced because of the dispersion of the currents of the module. The source of mismatch is not only the cells inside each module, but the lack of parallelism of the mounted modules. For many systems this parallelism must be better than 0.2°.

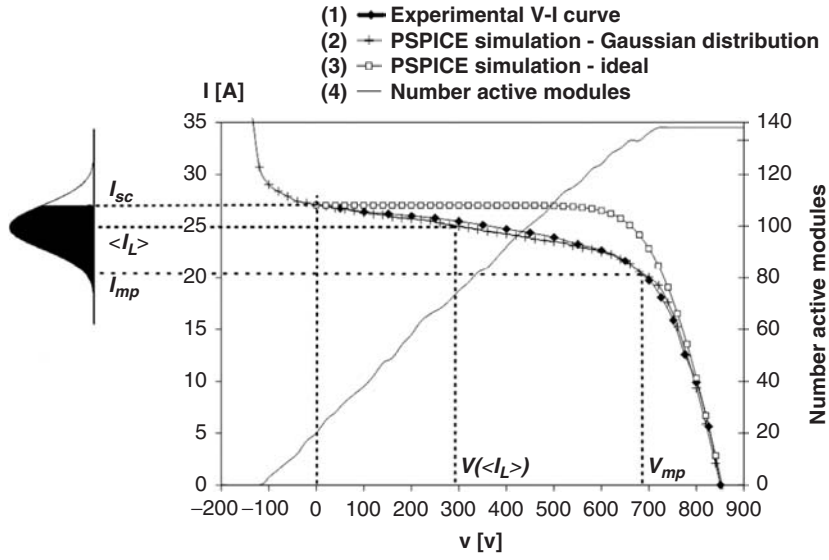


Figure 10.33 When the sources of optical mismatch are due to many causes the current generated by each element, shunted with a bypass diode, follows a Gaussion distribution. The current mismatch appears as a parallel resistance. In short-circuit, several modules are required to bias the non-ideal bypass diodes. In this case 19 modules or 100 V were required. (EUCLIDES Tenerife, Array nº 7, UPM-1999)

The measurement of the acceptance angle can be carried out outside, under real sun, which is a complicated experiment for controlling tenths of a degree on a commercial tracker with the sun "moving" across the sky, and recording the transference curve versus the maximum power. The experiment becomes reasonably easy using an indoor CPV solar simulator where the $I-V$ curve is recorded for any angle of incidence. For large arrays the outdoor test is always required.

10.8.5 Testing CPV Modules and Systems Equipped with Multijunction Solar Cells

The extremely high efficiency of multijunction III–V cells has made them the ideal devices for concentrator systems. These cells, which make better use of the solar spectrum, currently consist of three devices, electrically connected in series which perform simultaneously as solar cells and "bandpass" optical filters. They are more sensitive to spectrum variation than the low-bandgap single-junction solar cells (silicon ones, for example).

Measurement of these cells and the modules made with them requires the simultaneous control of the collimation and the spectrum of the light. The flash lamp must provide the spectrum reference standard or an equivalent to the reference.

Measurement of the $I-V$ curve, as happens with single-junction cells, must be based on a comparison with a reference cell intensity, but now there are, simultaneously, three reference subcells in the game.

The correct measurement at one reference spectrum $E_{ref}(\lambda)$ will be carried out when all the photocurrent cells in the multijunction monolithic stack I_{L1}, I_{L2} and I_{L3} simultaneously equals

the photogenerated current values $I_{Lj}ref$ corresponding to the reference spectrum. That is, when the following relations verify:

$$I_{Lj} = \int_{\lambda j1}^{\lambda j2} E_{ref}(\lambda) \cdot S'_j(\lambda)\mathrm{d}\lambda \equiv I_{Lj}ref \text{ (calibrated value)} \tag{10.30}$$

where the index $j = 1,2$ and 3, are referred to currents and spectral response of each subcell of the multijunction cell.

Of course, to have the exact reference spectrum $E_{ref}(\lambda)$ in a flash shot is practically impossible, as well as being impossible under real sun, but any spectrum, referred to as "local" in the following formulas, which is able to provide the local currents I_{Lj} and I_{Lj} simultaneously that achieve the following equations:

$$\frac{I_{Li}local}{I_{Lj}local} = \frac{I_{Li}ref}{I_{Lj}ref}; \quad \min(I_{Ljref}) = \min(I_{Ljlocal}) \tag{10.31}$$

can be called an "equivalent reference spectrum" and will provide the same $I-V$ curve as multijunction cells illuminated by the standard irradiance.

Although the conditions expressed by Equation (10.31) seem very complex, it is becoming relatively easy in practice because of the third junction. $I_{L3}ref$ and $I_{L3}local$ are usually larger than I_{L1} and I_{L2}. It means that we must only pay attention the top and middle junctions. The condition becomes simpler.

$$\frac{I_{L1}ref}{I_{L2}ref} = \frac{I_{L1}local}{I_{L2}local} \tag{10.32}$$

$$I_{L1}ref = I_{L1}local \text{ if } I_{L1} < I_{L2} \tag{10.33}$$

$$I_{L2}ref = I_{L2}local \text{ if } I_{L1} > I_{L2}$$

As the flash lamp light changes the spectrum throughout its decay, becoming reddish, and the spectrum blue content can be modulated with the power supply, it is possible to find one moment in the flash discharge in which the previous condition (Equations 10.32 and 10.33) are fulfilled (Figure 10.34). This process is carried out in modern solar simulators which allow manufacturers to rate the actual module power indoors, under standard conditions.

10.8.6 Multijunction Cells Inside Module Optics

When the cells are under the module optics, then the calibrated values of the reference sub-cells under the reference spectrum become invalid because the optics modify the source with the spectral transmission function $t(\lambda)$. Then we have

$$I_{Ljref}^{Op} = A_i \int E_{ref}(\lambda)t(\lambda)S_j(\lambda)\mathrm{d}\lambda \neq I_{Ljref} \tag{10.34}$$

for the reference spectrum where the subscript O_p means that the cells are under the optics.

The new ratio assuming the current measurement at equivalent spectrum is

$$\frac{I_{Liref}^{Op}}{I_{jref}^{Op}} = \frac{I_{Lilocal}^{Op}}{I_{Ljlocal}^{Op}}; \quad \min\left(I_{Ljref}^{Op}\right) = \min\left(I_{Ljlocal}^{Op}\right) \tag{10.35}$$

These conditions require that the reference element for that test is a calibrated concentrator element [42].

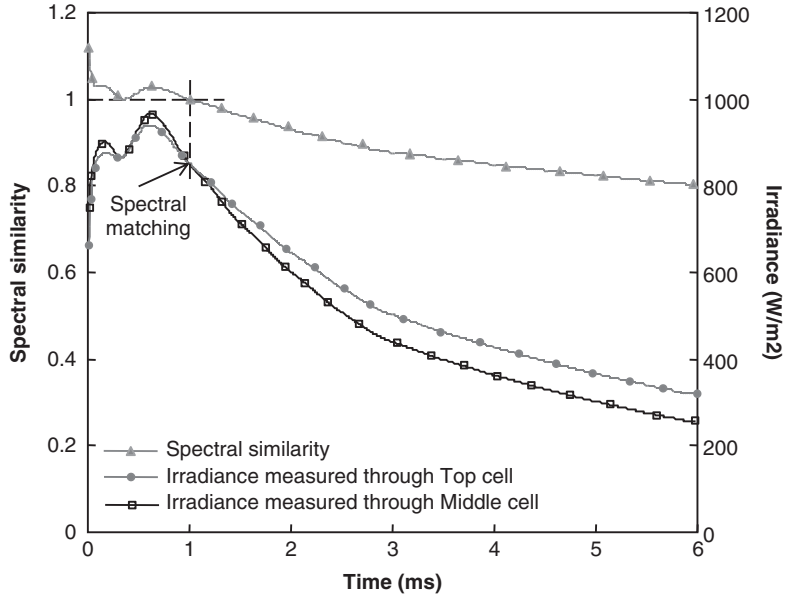


Figure 10.34 The currents I_{L1} and I_{L2} generated simultaneously by each component cell under the same flash shot. The correct measurement of a multijunction cell or system is carried out at the convenient moment, when both currents are equal. In this condition the artificial spectrum is equivalent to the reference spectrum (Courtesy UPM-IES)

10.8.7 The Production of PV Concentrators versus the Effective Available Radiation Accounting for Daylight Spectrum Variations

The yearly productivity of a CPV system should be based mainly on the available normal direct radiation, on the site temperature and on the shadowing between arrays. The operating conditions are quite uniform because the normal incidence of the sunlight and the narrow range of DNI values yield the major part of the incident power.

The reference for this calculation is the power rated under standard test conditions (STC), which defines the irradiance level, the light spectrum and cell temperature for module or system rating.

In spite of the response of the photovoltaic modules being spectrum dependent it has been proven [43] that its influence on the effect in energy production on silicon-based systems is less than 0.5%, negligible in practice. But this effect could be more important in mult-junction cells. The spectral dependence of multijunction cells means that they will perform worse in the periods of sunrise and sunset, because one cell, normally the top one, is limiting the overall current.

Accounting for the amount of the effective radiation used by a multijunction system, $H(\lambda,period)$, with respect to the total solar radiation $H(period)$ on a site is a suitable method to evaluate the energy effects of the spectral mismatch.

In Figure 10.35 we plot the record of the direct irradiance $B_{pirhel}(t)$ throughout one day as measured with a pyrheliometer that collects all radiation from 0.3 to 4 μm and the equivalent

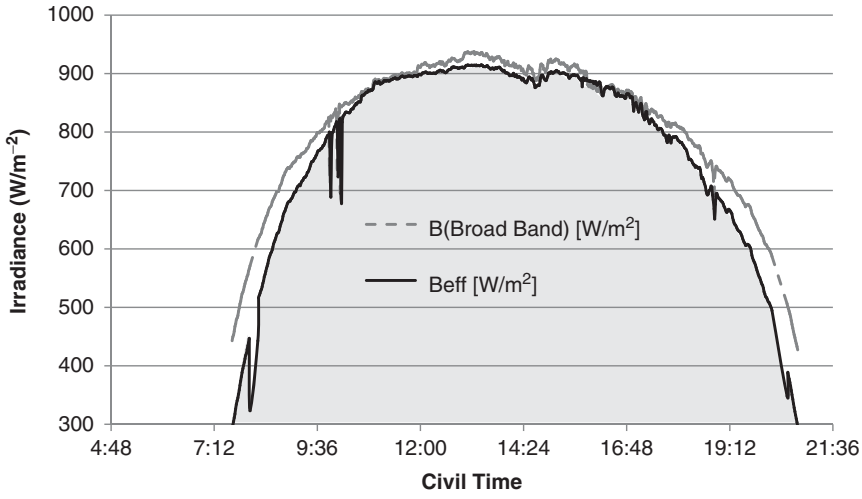


Figure 10.35 The equivalent irradiance is proportional to the actual cell current of the multijunction cells while the pyrheliometer irradiance $B_{\mathrm{BB}}(t)$ is proportional to the multijunction cell current under the reference spectrum. The area under the instantaneous $B_{\mathrm{eq}}(t)$ divided by the area under $B_{\mathrm{BB}}(t)$ provides the factor EMFE, accounting approximately for the losses of energy caused by the spectral mismatch. (Courtesy UPM-IES)

irradiance for a multijunction cell system, calculated as

$$B_{equiv}(t) = \frac{I_{cell}^{Op}(t)}{I_{ref}^{Op}} B_{STC} \tag{10.36}$$

where

$$I_{cell}^{Op}(t) = \min\left(I_{top}^{Op}(t),\, I_{mid}^{Op}(t) \cdot I_{bot}^{Op}(t)\right) \tag{10.37}$$

and I_{ref}^{Op} is the current of the ltijunction cell under STC conditions inside the optics and $I_{top}^{Op}(t)$, $I_{mid}^{Op}(t)$, $I_{bot}^{Op}(t)$ are the instantaneous currents of three calibrated isotype cells measured under the concentrator optics, simultaneously with the pyrheliometer recording. These isotype cells respond to direct irradiance as the top, medium and bottom cells of the multijunction cells located inside the system.

The area under $B_{equiv}(t)$ throughout the day divided by the area under the pyrheliometer curve $B_{pirhel}(t)$ yields a figure called the energy spectral mismatch ratio (ESMR) which represents the effective percentage of all available direct radiation throughout one day.

With these four records over many days, this figure can be defined monthly, yearly, etc. for one site. For example, the useful equation to define $ESMR_{monthly}$ is:

$$H(\lambda, month) = H(month) \cdot ESMR_{monthly} \tag{10.38}$$

where H(month) is the actual monthly direct radiation (kW h/m^2) as recorded by a pyrheliometer. And the spectral dependence effect is synthesised into a new figure called the "equivalent monthly radiation" H (λ, month) which allows the forecast of multijunction cell systems just like single-junction cells, but with a smaller input energy (lower kW h/m^2).

As an example, on a typical clear February day in Madrid (600 m above sea level) we obtained a daily average: $H(\lambda, \text{day Feb-09}) = 0.93\ H(\text{day Feb-09})$ and with the data recorded in Madrid during 2009 we obtained $ESMR_{yearly} = 0.95$.

10.9 SUMMARY

Concentrators have begun to be commercialised. They need to reach maturity by means of a learning curve that will be much faster than that of conventional PV has been. However up to nearly 600 MW of cumulated production is required to reach the cost target, which will be less than 2\$/W_p for the whole system in large plants. Each company will require about 10–30 MW/year to reach the module price level of around 0.8–1.0 \$/W_p. Thus, about ten companies might persist within 3–4 years, the time to reach maturity. However there are many more initiatives open, which can be seen in reference [44].

It is interesting to point out the ISFOC initiative (Puertollano, 2006, Spain), consisting of spending €16 M to develop CPV, gave rise to several calls for offers larger than 200 kW_p and offered the winning companies testing and diagnosis of their performance as well as certifying quality to customers and financing institutions. This is becoming a worldwide reference centre for CPV systems. A similar initiative is being planned in Colorado, USA for PV, CPV and CSP (announced in October 2009), but the amount of resources is not yet known.

The aggressive price lowering campaign which took place by the end of 2009 by the silicon PV sector could damage the development of CPVs as well as other emerging technologies, capturing the market in the period required by the full industrialisation of the new technologies.

The promised high productivity of CPVs has been demonstrated by Concentrix Solar at this station in Puertollano and in Seville, with energy efficiency of over 20% at a yearly level. That is, 20% of the received direct energy has been injected as electricity into de grid. With the next generation modules these figures can be increased up to 23% (remember that the best conventional PV systems are yielding nearly 12%) [45, 46].

REFERENCES

1. This book: Section 3.5.5 and Chapter 4.
2. Burgess EL, Pritchard DA, Performance of a one kilowatt concentrator array utilizing active cooling. *Proc. 13th IEEE Photovoltaic Specialists Conf*. (New York) 1978, pp 1121–1124.
3. Toña J, Guascor Fotón, CPV Systems Concentrated Photovoltaic Summit 08, *CPV Today*, 1–2 April 2008, Madrid.
4. Luther J, Luque A, Bett AW, Dimrot F, Lerchenmüller H, Sala G, Algora C, Concentration photovoltaics for highest efficiencies and cost reduction, *Proc. 20th European Photovoltaic Solar Energy Conf*., pp 1953–1957, Barcelona 2005.
5. Bett AW, Burger B, Dimroth F, Siefer G, Lerchenmuller H, High-Concentration PV using III-V Solar Cells, *PV-IEEE 4th World Conference*, Vol. 1, pp 615–20, Waikoloa, HI, 2006.
6. Algora C., The importance of the very high concentration in 3rd generation solar cells. *Next Generation Photovoltaics for High Efficicncy Through Full Spectrum Utilization*, Institute of Physics Publishing, 2004, Bristol, UK.
7. O'Neill M, Fifth Generation 20X Linear Fresnel Lens/Silicon Cell Concentrator Technology, *International Conference on Solar Concentrators for the Generation of Electricity ICSC-5*, Palm Desert 2008.
8. Sala G, Concentrator Systems, *Practical Handbook of Photovoltaics Fundamentals and Applications*, Chapter IIId, p. 682, Elsevier, 2003.

9. Sala G, Antón I, Photovoltaic concentrator systems facing the problems of commercialization, *Proc. 4th World Conf. on Photovoltaic Energy Conversion*, pp 609–614, Hawaii 2006.
10. Swanson R M, Point contact solar cells: Modeling and Experiment, *Solar Cells*, **17**, 85–118, 1985.
11. Gombert A, Hakenjos A, Heile I, Wullner J, Gerstmaier T, Van Riesen S, FLATCON CPV System, Field Data and New Development, *24th European Photovoltaic Solar Energy Conference*, Hamburg, September 2009.
12. Muñoz JA, Silva D, Payán A, Pereles O, Osuna R, Alonso M, Chenlo F, Two years of operation of the Sevilla PV 1.2 MW Grid Connected Plant, *Proc. 23rd European Photovoltaic Solar Energy Conf.*, pp 3199, Valencia 2008.
13. Swanson, The promise of concentrators, *Progress in Photovoltaics*, **8** (1) 93–111, 2000.
14. Antón I, Pachón D, Sala G, Results and conclusions of C-Rating Project: An European initiative for normalisation and rating of PV concentrator systems, *19th European Photovoltaic Solar Energy Conference*, París 2004, pp 2121–2124.
15. Goetzberger A, Greube W, Solar energy conversion with fluorescents collectors, *Appl. Phys*. **12**, 123–139, 1997.
16. Goldschmidt JC, Peters M, Dimroth F, Bett AW, Hermle M, Glunz SW, Willeke GP, Steidl L., Developing Large and efficient Fluorescents Concentrator Systems, *Proceedings of 24th European Photovoltaic Solar Energy Conference*, Hamburg 2009.
17. Winston R, Welford WT, *Optics and nonimaging concentrators: light and solar energy*, Academic press San Diego, 1978.
18. Luque A, *Solar Cells and Optics for photovoltaic concentration*, Adam Hilger, ISBN 0-85274-106-5, p. 462.
19. Winston R, Miñano JC, Benítez PG, Non Imaging Optics, ISBN 0-12-759751-4 *Elsevier Inc*., 2005.
20. Victoria M, Domínguez C, Antón I, Sala G, Comparative analysis of different secondary optical elements for aspheric primary lenses, *Optics Express*, **17** (8) 6487–6492, 2009.
21. Sala G, Measurement of CPV Components and Systems, CPV Today, Photovoltaic Concentator Summit, 1–2 April 2008, Madrid.
22. Antón I, Solar R, Sala G, Pachón D, IV testing of concentration modules and cells with non-uniform light patterns, *Proc. 17th European Photovoltaic Solar Energy Conf*., pp 611–614, Munich 2001.
23. Antón I, Sala G,. Losses Caused by Dispersion of Optical Parameters And Misalignments in PVConcentrators, Prog. Photovolt: Res. Appl. **13**, 341–352, 2005.
24. Concentrator Photovoltaic (CPV) modules and assemblies- Design qualification and type approval, *IEC 62108 Ed 1.0*, 2007.
25. *Patente de Invención N° 470994*. Concentrador y/o reflector de luz y procedimiento para su Fabricación. Madrid, 21 de Junio de 1978. (Consiste en una lente de fresnel hibrida de vidrio-silicona o vidrio-elastomero transparente). En explotación durante 1981 y 1982.
26. Sala G, Lorenzo E, Hybrid silicone-glass Fresnel lens as concentrator for photovoltaic applications., *2nd European Photovoltaic Solar Energy Conference and exhibition*, Berlin, 1979, pp 1004–1010.
27. Vivar M, Antón I, Pachón D, Sala G *et al*., Third generation of EUCLIDES Systems: First Results and Modelling of Annual Production in Ideoconte Project test sites, *ICSC-4*, El Escorial, March 2007.
28. Martínez M, Antón I, Sala G, Prediction of PV Concentrators Energy Production: Influence of wind in the Cooling mechanism. First Steps, *ICSC-4*, El Escorial, March 2007.
29. Luque A, Solar Cells and Optics for Photovoltaic Concentration, Adam Hilger, ISBN 0-85274-106-5, p. 265.
30. Miñano JC, González JC, Benítez P, (1995). RXI: A high-gain, compact, nonimaging concentrator. *Applied Optics*, Vol. 34, **34**, 7850–7856.

31. Zamora P, Cvetkovic A, Buljana M, Hernández M, Benítez P, Miñano JC, Dross O, Alvarez R, Santamaría A, Advanced PV Concentrators, *Proceedings of the 34th IEEE Photovoltaic Specialists Conference*, 7–12 June 2009, Philadelphia.
32. Victoria M, Domínguez C, Antón I, Sala G, High Concentration Reflexive System With Fluid Dielectric. *Proceedings of the 23rd European Photovoltaic Solar Energy Conference*, Valencia, September 2008.
33. Smeltink JFH, Blakers AW, 40 kW PV Thermal Roof Mounted Concentrator System *Photovoltaic Energy Conversion Conference record of the 2006 IEEE 4th World Conference*, May 2006 ISBN: - 1-4244-0017-1, Vol. 1, pp 636–639.
34. Verlinden PJ, Lewandowski A. Kendall H, Carter S, Cheah K, Varfolomeev I, Watts D, Volk M, Thomas I, Wakeman P, Neumann A, Gizinski P, Modra D, Turner D, Lasich JB, Update on two-year performance of 120 kW_p concentrator PV systems using multi-junction III-V solar cells and parabolic dish reflective optics, *Proc. 33rd IEEE Photovoltaic Specialists Conf*., pp 1–6, 2008.
35. Lorenzo E, *Electricidad Solar Ingeniería de los Sistemas Fotovoltaicos*, p. 218, 1994 PROGENSA, Sevilla, Spain.
36. Klotz FH, Möhring HD, Integrated Parabolic Trough (IPT) for low concentrator PV systems, *Proc. 4th Int. Conf. on Solar Concentrator for the Generation of Electricity or Hydrogen, pp*, El Escorial, 2007.
37. Arboiro JC,. Sala G, A constant self-learning scheme for tracking systems, *Proc. 14th European Photovoltaic Solar Energy Conf*., Barcelona, pp 332–335, 1997.
38. Luque Heredia I, Moreno JM, Magalhaes PH, Cervantes R, Quemere G, Laurent O, Inspira's CPV Sun Tracking, *Concentrator Photovoltaics*, Springer, Berlín, pp 221–251, 2007.
39. Mau S, Krametz T, Influence of solar cell capacitance on the measurement of I-V-curves of PV-modules, *Proc. 20th European Photovoltaic Solar Energy Conf*., Barcelona, Spain, 2005.
40. Domínguez C, Antón I, Sala G, Solar simulator for concentrator photovoltaic systems, *Opt. Express*, **16**, 14894–14901, 2008.
41. Antón I, Sala G, Arboiro JC, Effects of the optical performance on the output power of the EUCLIDES array, *Proc. 16th European Photovoltaic Solar Energy Conf*., Glasgow, 2000.
42. Domínguez C, Askins S, Antón I, Sala G, Indoor characterization of CPV modules using the Helios 3198 Solar Simulator, *Proc. 24th European Photovoltaic Solar Energy Conf*., Hamburg, pp 165–169, 2009.
43. Huld T, Sample T, Dunlop ED, A simple model for estimate the influence of spectrum variations on PV performance, *Proc. 24th European Photovoltaic Solar Energy Conf*., Hamburg, pp 3385–3389, 2009.
44. Wilming W, CPV Systems: Bringing the sun into focus, *Sun and Wind Energy*, **12**, 100, 2009.
45. Gombert A, Hakenjos A, Heile I, Wúllner J, Gerstmaier T, Van Riesen S, DLATCON CPV systems- Field data and new development, *Proc. 24th European Photovoltaic Solar Energy Conf*., Hamburg, pp 156–158, 2009.
46. Martínez M, Sánchez D, Perea J, Rubio F, Banda P, ISFOC demonstration plants: Rating and production data analysis, *Proc. 24th European Photovoltaic Solar Energy Conf*., Hamburg, pp 159–164, 2009.

11

Crystalline Silicon Thin-Film Solar Cells via High-temperature and Intermediate-temperature Approaches

Armin G. Aberle[1] and Per I. Widenborg[2]

[1]*Solar Energy Research Institute of Singapore and Department of Electrical and Computer Engineering, National University of Singapore, Singapore,*
[2]*Formerly with School of Photovoltaic and Renewable Energy Engineering, University of New South Wales, Sydney, Australia. Now with Solar Energy Research Institute of Singapore, National University of Singapore, Singapore*

11.1 INTRODUCTION

11.1.1 Motivation for Thin c-Si Solar Cells

The large majority (>80%) of today's photovoltaic (PV) modules are based on crystalline silicon (c-Si) wafers [1]. The reasons behind c-Si's dominance of the PV market are numerous, the most important ones being a high module efficiency (12–19%), robust and high-yield cell and module fabrication processes, an excellent long-term stability of the modules (> 25 years), and material abundance and nontoxicity. Another important contributor is the fact that wafer-based silicon is the material of choice in the integrated circuit (IC) industry, enabling a sharing of research and infrastructure costs with this mighty partner, and providing an excellent and well-understood technology and equipment base, ranging from the fabrication of silicon wafers to the manufacture of PV cells and modules. There exist very few materials that are as thoroughly researched as crystalline silicon. The global PV production exceeded 3 GW_p/year in 2007, which corresponds to about 30 silicon wafer solar cells every second. Modern silicon wafer PV lines will soon approach production rates

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

of one solar cell per second. Silicon wafer-based PV modules are still getting cheaper (€/W_p) and/or more efficient every year, making them a "moving target" for any competing PV technology.

However, despite the past and present success, there is one problem with silicon wafers – cost. Today's standard Si wafer PV modules (which use multicrystalline wafers and have a total-area efficiency of 12–14%) cost about 2–3 €/W_p to fabricate, and almost half of this cost is due to the starting material, i.e. the unprocessed silicon wafers. The wafers are square (or pseudo-square, i.e. with round corners), have an area in the range of 150–250 cm^2, a thickness in the range of 180–300 µm, and cost about € 2–3 each (i.e., about 1.3 €cents/cm^2). The silicon consumption of these PV modules is about 10 g/W_p. To further reduce the cost of PV electricity, the PV industry is moving towards ever larger wafers. Cell areas of over 300 cm^2 are being tested, and an area of 400 cm^2 might be industry standard within a decade. Simultaneously, industry aims at reducing the wafer thickness as far as possible, as this increases the solar cell area that can be obtained from a given silicon ingot or brick. The minimum wafer thickness used will be determined by the wafer breakage rates in the PV factories and not by the potential cost savings due to further reduced silicon consumption. Larger wafers will require larger minimal thickness. For very large wafers (>400 cm^2), it appears that a minimum thickness of at least 100 µm will be required to keep wafer breakage rates under control. Additionally, the relative fraction of silicon lost due to sawing (kerf loss) increases as the wafer gets thinner. Compared with today's technology, the reduction of the silicon consumption (g/W_p) is thus limited to a factor of about 2 for conventional wafer-based PV technologies (silicon ribbon technologies have a slightly larger reduction potential). In addition to being material intensive, the fabrication of silicon wafers is also energy intensive. Thus, and in light of the rising cost of fossil-fuel-based energy, it will be very difficult to reduce the cost of Si wafer-based PV modules to well below 1 €/W_p using an environmentally friendly manufacturing process. Burning coal for making silicon wafers for PV is not a viable option.

PV module costs of well below 1 €/W_p are, possibly, more easily achievable using thin-film technologies. The main reasons are that the fabrication of such modules requires smaller amounts of energy and materials, thin films can be deposited over very large areas, and thin-film solar cells can be series-connected in a fast and inexpensive way. Importantly, both the efficiency and the long-term stability of thin-film PV modules are improving, while simultaneously their manufacturing costs are decreasing. As a result, PV module costs of well below 1 €/W_p seem possible with thin films, in large-scale production. Thin-film PV modules are particularly attractive for emerging applications such as green-field PV power plants and building-integrated PV.

11.1.2 Classification of c-Si Thin-Film PV Technologies and Materials

A convenient way to distinguish between the different c-Si thin-film PV technologies is the temperature stability of the supporting material, giving low-T (<450 °C), intermediate-T (450–700 °C) and high-T (> 700 °C) approaches. Generally, the crystallographic and electronic quality of c-Si films improves with increasing deposition or growth temperature, which explains the interest in the high-T approaches. It is also important to mention whether the supporting material acts as substrate or superstrate for the solar cells in the finished PV module. In superstrate configuration, the sunlight enters the cells through the supporting material and hence this material must be highly transparent. The superstrate configuration offers cost advantages since the supporting material can also take over the role of the front cover sheet that is required for any PV module. The latter point is one of the reasons why glass superstrates are particularly attractive for thin-film PV. Additional reasons are its excellent UV stability and moisture barrier properties.

The standard low-T supporting material for thin-film PV is soda-lime glass with low iron content, which costs about 10 €/m^2 for thicknesses in the 3–4 mm range. However, as of today, no method has been identified that is capable of producing reasonably efficient crystalline silicon solar cells on soda-lime glass or other low-T supporting materials. The technology closest to this is based on low-T (<200 °C) deposited microcrystalline silicon (a mixed-phase material containing amorphous and crystalline regions), as discussed in Chapter 12 of this book. An important intermediate-T supporting material that has emerged during the last decade for PV applications is a borosilicate float glass (Borofloat33) from Schott AG, Germany [2], which costs about 25 €/m^2 for thicknesses in the 3–5 mm range. Important high-T supporting materials are c-Si wafers, ceramics (such as alumina and mullite), glass ceramics (such as the 9664 product from Corning Inc, USA [3]), graphite, and steel. The cost of the substrate (€/m^2) is a major issue for most high-T approaches. A summary of the high-T and intermediate-T routes for thin-film c-Si solar cells on supporting materials is given in Table 11.1, together with the most advanced technologies in each category.

To simplify the discussion of the results obtained with these two c-Si thin-film PV technology routes, the following material specification is adopted:

- sc-Si (singlecrystalline Si; grain size $g > 10$ cm)
- mc-Si (multicrystalline Si; $g = 1–100$ mm)
- pc-Si (polycrystalline Si; $g = 1–1000$ µm, no amorphous tissue)
- µc-Si:H (hydrogenated microcrystalline or, equivalently, nanocrystalline Si; $g < 1$ µm, material contains amorphous tissue).

Figure 11.1 shows representative examples for some of these silicon materials. The sample in the cross-sectional transmission electron microscope (XTEM) micrograph (a) is a singlecrystalline Si film epitaxially grown by IAD (ion-assisted deposition) at about 600 °C on a (100)-oriented Cz Si wafer [4]. The growth interface is not detectable and the grown film has a low density of structural defects. Another confirmation of the high crystal quality of this Si film is the fact that

Table 11.1 Technology routes for c-Si thin-film solar cells on native and foreign supporting materials. Low-T approaches are omitted as they have not yet produced solar cells with reasonable efficiency

	High-T approaches	Intermediate-T approaches
Substrate	Si wafers, oxidised Si wafers, ceramics, graphite, steel	Borosilicate glass, boro-aluminosilicate glass, metal
Solar cell process	Diffused emitter (~900 °C) or low-T a-Si hetero-emitter	Grown emitter or low-T a-Si hetero-emitter
Post-crystallisation treatment (optional)	Defect anneal not required, hydrogenation (200–700 °C)	Defect anneal (700–1000 °C), hydrogenation (200–700 °C)
Crystallisation	$T_{melt} \approx 1400$ °C (s to min). Step not required for Si wafer substrates.	$T_{SPC} \approx 600$ °C (h), $T_{IAD} \approx 600$ °C (min), $T_{laser} \approx 1400$ °C (µs)
Si deposition (deposition rate, Si film thickness, grain size)	700–1200 °C (up to several µm/min, 10–100 µm, g > 50 µm)	200–700 °C (up to 1 µm/min, 1.5–5 µm, 500 nm < g < 50 µm)
Most advanced PV technology	EpiWE (Fraunhofer, IMEC), Si on graphite (Fraunhofer), Si transfer (several groups)	CSG (CSG Solar), PLASMA (UNSW)

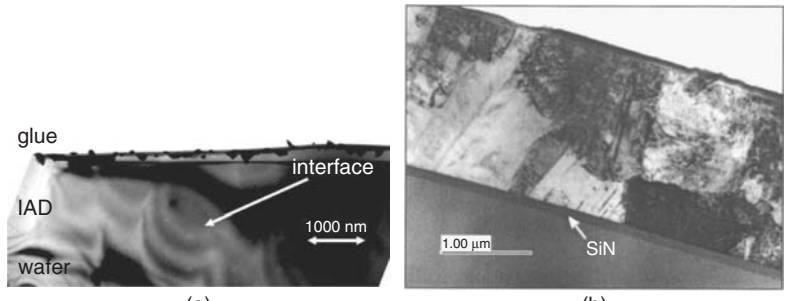


Figure 11.1 (a) XTEM micrograph of a singlecrystalline silicon film grown at 600 °C on a (100)-oriented Cz Si wafer (from reference [4]; reproduced by permission of Elsevier). (b) XTEM micrograph of a polycrystalline silicon film formed by solid phase crystallisation at 600 °C of an amorphous Si precursor diode evaporated at low temperature onto silicon nitride-coated glass (reprinted with permission from reference [5]; Copyright 2005, American Institute of Physics)

interference fringes visible in the Si wafer region run smoothly across the growth interface and into the epitaxially grown film. The reason for these interference fringes are thickness variations in the TEM sample. The sample in the cross-sectional TEM micrograph (b) is a 2-µm-thick polycrystalline silicon film made on silicon nitride (SiN)-coated glass by solid phase crystallisation (SPC) of an evaporated amorphous silicon precursor diode [5]. The average grain size is about 1 µm and the film has a relatively high density of structural defects.

11.1.3 Silicon Deposition Methods

A number of silicon thin-film deposition methods have been developed over the past 40 years. The standard silicon homoepitaxy method is thermal CVD (chemical vapour deposition) at high temperature (~1100 °C) using either silane (SiH_4), dichlorosilane (DCS, SiH_2Cl_2) or trichlorosilane (TCS, $SiHCl_3$) as the silicon source gas [6]. Chlorosilane is generally preferred over silane due to its lower cost and safety requirements. If applied to foreign substrates, thermal CVD produces fine-grained pc-Si films. Thermal CVD can be performed at atmospheric pressure (APCVD) or at low pressure (LPCVD). Liquid phase epitaxy (LPE) is another high-temperature epitaxy method that produces good-quality silicon films. It is usually performed at substrate temperatures in the 700–1000 °C range [7]. The most important intermediate-T homoepitaxy methods are MBE (molecular beam epitaxy) [8] and IAD (ion-assisted deposition) [9]. MBE and IAD are PVD (physical vapour deposition) methods which use a powerful electron beam (> 1 kW) to melt and evaporate high-purity silicon contained within a water-cooled crucible. An alternative to direct homoepitaxial growth is SPE (solid phase epitaxy) where an amorphous Si film is deposited and then crystallised into an epitaxial film using thermal annealing [10]. The standard low-T silicon deposition method is PECVD (plasma-enhanced chemical vapour deposition) using silane [11], giving hydrogenated amorphous silicon (a-Si:H). Standard PECVD systems for a-Si:H deposition use a parallel-plate configuration and a plasma excitation frequency of 13.56 MHz. If the silane is replaced by a hydrogen/silane gas mix with a silane concentration of a few percent, then hydrogenated microcrystalline silicon (µc-Si:H) is obtained instead of a-Si:H [11, 12]. With the exception of APCVD, all these silicon deposition processes are conducted within specific vacuum deposition systems. PVD methods operate at much lower pressures (high or even ultra-high vacuum) than CVD methods and hence require specialised pump systems (turbo pumps and/or cryogenic pumps) capable of achieving a system base pressure of below about 1×10^{-5} Pa. In contrast, the working gas pressure of most CVD methods is of the order of 10 Pa, ensuring that a base pressure of about 10^{-2} Pa is generally

sufficient, and eliminating the need for expensive vacuum pumps. However, to keep contamination levels in the deposited silicon films at tolerable levels, a high flow rate of the ultra-pure silicon source gas is required. As a consequence, the usage of Si atoms ('chemical yield') of CVD systems is generally relatively low (for example, <10% for thermal CVD [6]). For further PV-relevant details of the different silicon deposition methods mentioned above, the interested reader is referred to various PV review articles [7, 11–15] as well as Chapter 12 of this book.

11.1.4 Seeded versus Non-seeded Silicon Film Growth

In the case of foreign or poor-quality native substrates, the 'seed layer concept' may offer advantages. The idea is to form a thin layer (seed layer) of good structural material quality on the substrate and to use the seed layer as a crystal template for the crystalline layer that is subsequently grown on top of it, whereby the crystal properties of the seed layer are transferred to the growing crystalline material via epitaxy. The main advantage of the two-step seed layer method over non-seeded growth methods is that the properties of the seed layer and the subsequently grown crystalline layer can be independently optimised, whereby the focus is on the structural quality and grain size for the seed layer and on the electronic quality and high deposition rate for the much thicker epitaxial layer which must absorb the sunlight [16]. As discussed in Sections 11.3.2 and 11.4.2, various methods exist for the creation of c-Si seed layers for PV applications.

11.2 MODELLING

11.2.1 Impact of Diffusion Length in Absorber Region on Cell Efficiency

The absorber region of c-Si thin-film solar cells is usually doped, because this has the potential to give higher efficiency (via higher open-circuit voltages and fill factors) compared with non-doped ("intrinsic") absorbers. For low-injection conditions, minority charge carrier transport in doped semiconductor layers is dominated by diffusion [17]. Thus, the diffusion length L_{abs} of minority carriers in the absorber has a strong impact on the efficiency of a c-Si thin-film solar cell. To analyse the dependence analytically, let us assume that the thickness of the absorber significantly exceeds its diffusion length ('thick absorber'), ensuring that recombination at the absorber's rear surface negligibly affects the electrical parameters of the solar cell. If the diffusion length L_{abs} becomes much smaller than the absorber thickness, several detrimental effects cause a severe decrease in the PV efficiency. First, the short-circuit current density J_{sc} decreases since more and more light-generated carriers are created too far away from the p-n junction to have a chance of being collected. Second, a decreasing L_{abs} increases the absorber component $J_{0.abs}$ of the diode's dark saturation current density [18]:

$$J_{0.abs} = \frac{qn_i^2}{N_{abs}} \frac{D_{abs}}{L_{abs}}, \tag{11.1}$$

where q is the elementary charge, n_i the semiconductor's intrinsic carrier concentration, N_{abs} the doping concentration in the absorber layer, and D_{abs} the diffusion constant of the minority carriers in the absorber. Neglecting recombination in the emitter and the junction space-charge region of the cell, it follows that the absorber layer limits the solar cell's open-circuit voltage to [17]

$$V_{oc.abs} = \frac{nkT}{q} \ln \frac{J_{sc}}{J_{0.abs}}, \tag{11.2}$$

where n is the ideality factor of the diode, k the Boltzmann constant, and T the absolute temperature. Thus, a decreasing L_{abs} decreases the V_{oc} of the solar cell. Third, as a result of the decreasing open-circuit voltage, the fill factor of the solar cell decreases. An analytical expression describing this dependence, in the absence of parasitic series and shunt resistances and a diode ideality factor of unity, is [19]

$$FF = \frac{[v_{oc} - \ln(v_{oc} + 0.72)]}{(v_{oc} + 1)}, \tag{11.3}$$

where $v_{oc} \equiv V_{oc}/(kT/q)$ is the normalised open-circuit voltage. This simple analytical analysis shows that a decreasing minority-carrier diffusion length in the absorber decreases each of the three electrical parameters (V_{oc}, J_{sc}, FF) that determine the efficiency of the solar cell. A sufficiently high L_{abs} is thus a key prerequisite for a reasonable efficiency of c-Si thin-film solar cells.

As can be seen from Equations (11.1) and (11.2), an increasing doping level in the absorber will also increase the V_{oc} of the device. At high doping levels, however, the diffusion length collapses due to heavy doping effects such as Auger recombination and Shockley–Read–Hall (SRH) recombination [17]. In actual polycrystalline silicon thin-film samples it is impossible to theoretically predict the SRH-related carrier lifetime since it depends on parameters such as the concentration of impurities (for example oxygen, nitrogen, carbon and metals), the type and nature and density of structural defects, the concentration of dopant atoms, the fraction of dopant atoms that are electrically active, the grain size, and the properties of the grain boundaries. Given these complexities, the approach generally taken in the development of c-Si thin-film solar cells is thus to determine *experimentally* which doping level maximises the PV efficiency, by varying the active doping level in the absorber from about $1\times10^{15}\,\mathrm{cm}^{-3}$ to about $5\times10^{17}\,\mathrm{cm}^{-3}$.

An illustrative example for the theoretically expected impact of the diffusion length in the absorber layer on the one-sun efficiency of a p^+nn^+ pc-Si thin-film solar cell is given in Figure 11.2 [20]. The calculations were performed with the one-dimensional device simulator PC1D [21]. The planar cell has a thickness of 2 µm and recombination at both surfaces is set to zero. The field-enhanced recombination model (Hurkx model) in PC1D was switched off. The emitter is heavily p-type doped ($1\times10^{19}\,\mathrm{cm}^{-3}$) and has a thickness of 50 nm, the absorber is n-type doped ($1\times10^{16}\,\mathrm{cm}^{-3}$) and 1900 nm thick, and the BSF (back surface field) layer is heavily n-type doped ($1\times10^{19}\,\mathrm{cm}^{-3}$) and 50 nm thick. The solid lines represent the results obtained for the case of good light trapping properties (internal reflectances of 80% at the front surface and 90% at the rear surface) and a perfect antireflection coating, whereas the dashed lines were obtained for poor light trapping properties (internal reflectances of 10% at the front surface and 50% at the rear surface) and an external reflectance at the front surface of 10% for all wavelengths.

A number of interesting trends can be seen in Figure 11.2, and can be used as design guides when developing pc-Si thin-film solar cells. Regardless of the light trapping properties, J_{sc} increases strongly with increasing diffusion length in the absorber layer until saturating once the diffusion length exceeds about twice the absorber layer thickness. For the present cell, light trapping boosts the J_{sc} by up to 60%, whereby the gain is largest for long diffusion lengths. In contrast to J_{sc}, the open-circuit voltage does not show a saturation behaviour with respect to the diffusion length (although the rate of increase slows once the diffusion length exceeds the absorber thickness). Regardless of the diffusion length, light trapping has a small effect ($<10\%$) on the V_{oc}. Due to the V_{oc} behaviour, the solar cell efficiency keeps improving with increasing diffusion length in the absorber layer. For a diffusion length of 10 µm, the efficiency of the light-trapping cell is about 12.5%.

For the solar cell technologist, Figure 11.2 suggests the following strategy for optimising the efficiency of pc-Si thin-film solar cells: First, ensure that L_{abs} is larger than the absorber layer thickness (this can be realised either by improving the electronic quality of the absorber layer or

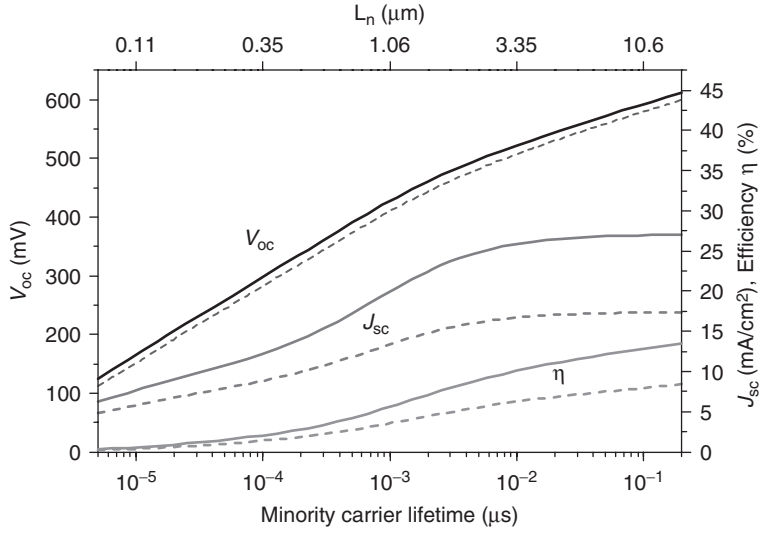


Figure 11.2 PC1D calculated dependence of open-circuit voltage, short-circuit current density and efficiency of a 2-μm-thick p^+nn^+ c-Si thin-film solar cell on the minority-carrier lifetime and diffusion length in the absorber layer. Surface recombination at both surfaces is set to zero. The solid lines assume good light trapping, the dashed lines poor light trapping (from reference [20]; reproduced by permission of Daniel Inns)

by decreasing the absorber layer thickness). Second, introduce an effective light trapping scheme to boost the J_{sc}. Third, focus on further improving the electronic quality of the absorber, emitter and back surface field layers to maximise the V_{oc}.

11.2.2 Impact of Surface Recombination

It is well established that minimising bulk (absorber) recombination in solar cells with $L_{abs} > W_{abs}$ will be of little value if the back surface has a high recombination rate. For an electronically *thick* absorber layer ($L_{abs} < W_{abs}$) of a *p-n* junction diode, assuming spatially uniform doping concentration N_{abs} and minority carrier diffusion length L_{abs}, its dark saturation current density (see Equation 11.1) can be written as

$$J_{0.abs} = \frac{qn_i^2}{N_{abs}} \frac{D_{abs}}{L_{abs}} = \frac{qn_i^2}{N_{abs}} S_\infty, \tag{11.4}$$

where the material parameter $S_\infty \equiv D_{abs}/L_{abs}$ has the dimension of a velocity and can be regarded as the recombination velocity of an infinitely thick, uniformly doped absorber layer [22]. For high-quality single-crystalline silicon, S_∞ depends only weakly on the doping concentration and has values in the 300–600 cm/s range. If the absorber layer is *not* electronically thick (i.e. the thickness of the absorber is not significantly larger than its diffusion length), the expression for $J_{0.abs}$ needs to be modified to consider recombination at the rear surface of the absorber. The result is [23]

$$J_{0.abs} = \frac{qn_i^2}{N_{abs}} S_\infty f_{geo} = \frac{qn_i^2}{N_{abs}} S_\infty \frac{\cosh(W_{abs}/L_{abs}) + (S_\infty/S)\sinh(W_{abs}/L_{abs})}{(S_\infty/S)\cosh(W_{abs}/L_{abs}) + \sinh(W_{abs}/L_{abs})}, \tag{11.5}$$

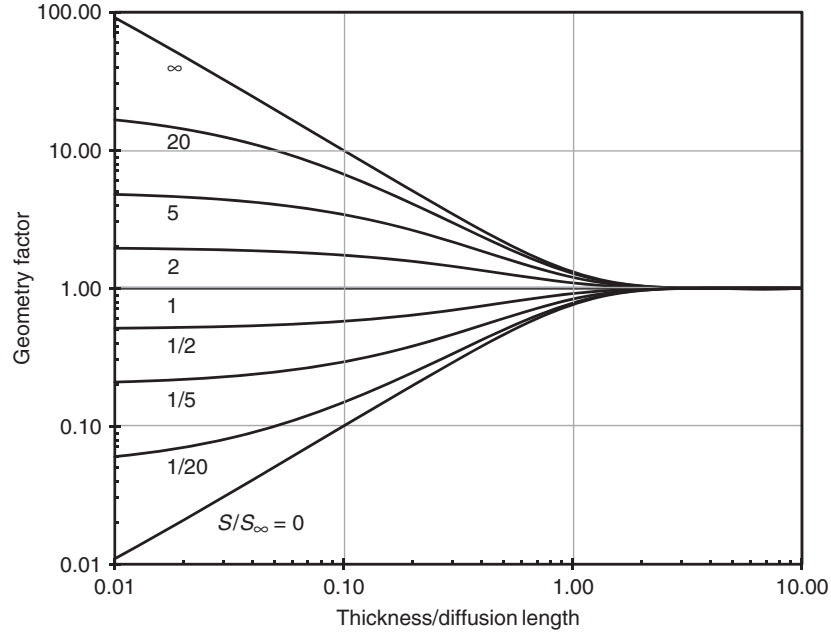


Figure 11.3 Geometry factor f_{geo} of the absorber layer as a function of W_{abs}/L_{abs}, for various values of the parameter S_∞/S

where f_{geo} is the geometry factor of the absorber layer, W_{abs} its thickness, and S the recombination velocity at its rear surface. The geometry factor of the absorber layer is fully determined by the two ratios S_∞/S and W_{abs}/L_{abs}. Figure 11.3 shows the geometry factor as a function of W_{abs}/L_{abs}, for various values of the parameter S_∞/S. Two regions can be distinguished. The first is given by $W_{abs} > 2L_{abs}$. In this case the rear surface has no impact on $J_{0.abs}$ and the geometry factor has a constant value (unity). The second region is given by $W_{abs} < L_{abs}$. In this case, three scenarios are possible: (a) For $S > S_\infty$ the geometry factor is larger than unity. In this case the surface has poorer recombinative properties than the bulk of the absorber. To reduce the negative impact of the surface on $J_{0.abs}$, the absorber layer should be thick. (b) For $S = S_\infty$ the geometry factor has the constant value of unity, regardless of the W_{abs}/L_{abs} ratio. In this case the absorber thickness has no impact on $J_{0.abs}$. (c) For $S < S_\infty$ the geometry factor is smaller than unity. In this case the beneficial impact of the surface on $J_{0.abs}$ can be maximised by thinning of the absorber layer, enabling a significantly higher open-circuit voltage of the solar cell. Of course, thinning of the absorber will lead to a loss in the cell's short-circuit current due to reduced optical absorption. This requires the implementation of an efficient light trapping scheme, as discussed in the next Section.

In parallel to the open-circuit voltage, recombination at the rear surface of the absorber will also reduce the short-circuit current of the solar cell. An illustrative example for the theoretically expected impact of the recombination velocity at the rear surface of the absorber on the short-circuit current density and the efficiency of a n^+p pc-Si thin-film solar cell is given in Figure 11.4, as a function of the cell thickness. Note that the largest thickness shown is 10 µm, which is much smaller than the thickness of a typical Si wafer. The calculations were again performed with PC1D [21], whereby the field-enhanced recombination model (Hurkx model) in PC1D was switched off. The planar cell has a 200-nm-thick emitter with a V_{oc} potential of 690 mV. External reflection at the front surface has a fixed value of 5% for all wavelengths and the cell has relatively good light trapping properties (80% internal reflection at both surfaces). The solid lines represent the results for

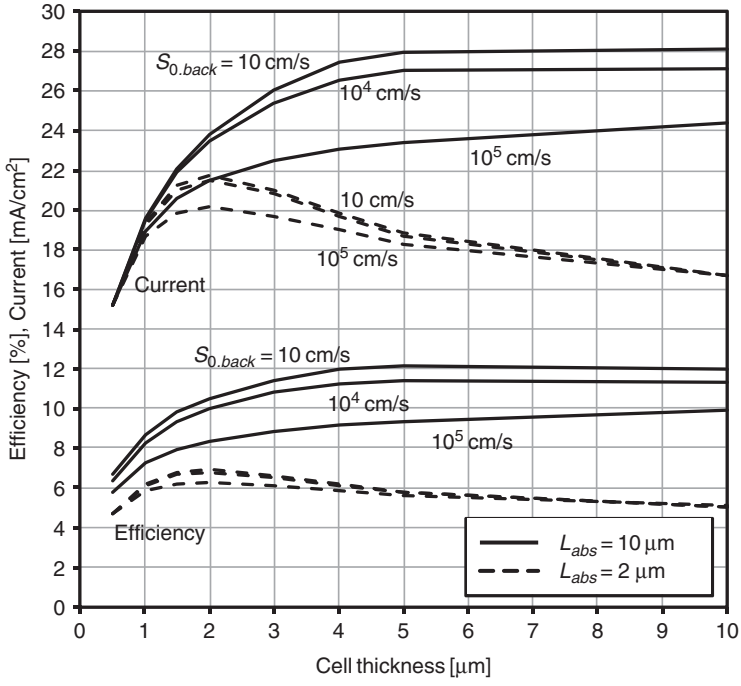


Figure 11.4 PC1D calculations of the one-sun short-circuit current density and efficiency of a planar n^+p c-Si thin-film solar cell as a function of the cell thickness. Parameters are the diffusion length L_{abs} in the absorber (2 and 10 μm) and the recombination velocity parameter $S_{0.back}$ (10, 10^4, and 10^5 cm/s) at the rear surface of the absorber. The cell features an emitter with a V_{oc} potential of 690 mV (junction depth 200 nm, erfc profile, surface concentration 1×10^{20} cm^{-3}, midgap surface states with $S_{n0} = S_{p0} = S_{0.front} = 10^4$ cm/s), relatively good light trapping properties (80% internal reflection at both surfaces) and a wavelength independent front surface reflection loss of 5%

the case of a good diffusion length (10 μm) in the absorber, while the dashed lines were obtained for a modest absorber diffusion length of 2 μm. As can be seen, in the case of the good diffusion length, passivation of the rear surface boosts the J_{sc} by about 4 mA/cm^2 (a 17% gain) and the efficiency by about 2 efficiency points (a 20% gain), with the best efficiencies (up to 12.1%) obtained for cell thicknesses in the 5–8 μm range. There is little to gain from a reduction of the rear surface recombination velocity to values well below 10^4 cm/s. The reason is the thinness of the absorber layer, ensuring that the carrier-collecting *p-n* junction is very close to the light-generated electron–hole pairs in the absorber. As a result, at the one-sun maximum power point, the excess carrier density in the absorber is very low ($\sim10^{12}$ cm^{-3}), ensuring low rear surface recombination rates despite relatively high rear surface recombination velocity (note that the surface recombination rate is the product of the bulk excess carrier density and the rear surface recombination velocity). For the modest diffusion length (2 μm) the impact of the rear surface recombination velocity on current and efficiency is much smaller (<10%) than in the case of the good diffusion length, whereby the best efficiency (6.9%) is obtained for cell thickness of about 2 μm. From Figure 11.4 it follows that, as a rule of thumb, the thickness of a c-Si thin-film solar cell having good light trapping properties should be about 50–80% of the diffusion length in the absorber layer.

For the investigated n^+p cell, the impact of the front surface recombination velocity is significantly smaller than that of the rear surface recombination velocity, even though the highest

carrier generation is at the front. The reason is the electric field associated with the doping profile of the emitter (erfc profile), repelling minority carriers from the front surface of the cell. As a result, the efficiency improves only by about 5% (relative) when the front surface recombination velocity is changed from 10^6 to 10 cm/s. Similar conclusions apply to heterojunction emitter thin-film c-Si solar cells.

Technologically, surface recombination rates at crystalline silicon surfaces can be minimised by various approaches. For a review, the interested reader is referred to the literature [24].

11.2.3 Impact of Light Trapping

Due to the weak absorption of near-infrared light in crystalline silicon, an effective light trapping scheme is essential for c-Si thin-film solar cells. One effective way to obtain light trapping is to texture the supporting material prior to the deposition of the silicon film. Due to the resulting textured silicon film, light is transmitted obliquely into it, significantly enhancing the optical pathlength and thus increasing the optical absorption. The optical absorption is further enhanced by depositing a high-quality reflector (back surface reflector, BSR) onto the rear surface, enabling a second pass through the silicon film. If the texture and the BSR are optimised such that total internal reflection occurs at both the front and the rear surface of the Si film, multiple passes through the silicon film become possible for weakly absorbed light. Apart from the light trapping benefits, the textured silicon film also reduces reflection losses at the front surface of the solar cell (double bounce effect).

Using the method of reference [25] (which is based on Monte Carlo simulation and ray tracing), we calculated the theoretical limit for the absorption of light in an air/glass/SiN/c-Si/air structure, for illumination through the glass (superstrate configuration). The glass/SiN interface, the SiN/c−Si interface and the rear c-Si surface were assumed to have a perfect (i.e. Lambertian) randomising texture, whereas the front glass surface was assumed to be planar. The glass was assumed to be a 3.3-mm-thick Borofloat glass pane from Schott [2] and the SiN (silicon nitride) film to be 70 nm thick, with a fixed refractive index of 2.0 and zero absorption for all wavelengths. Since real SiN films absorb strongly in the UV wavelength range [26], the simulations were limited to wavelengths larger than 400 nm. The optical constants of the c-Si film were assumed to be those of high-quality intrinsic bulk c-Si. The wavelength-dependent absorption results are shown in Figure 11.5, for several c-Si film thicknesses in the 0–2.7 μm range. These are typical thicknesses for c-Si thin-film solar cells on glass superstrates. It can be seen that, for long wavelengths (>1100 nm), absorption occurs predominately in the glass pane [25]. On the other hand, absorption in the glass pane becomes negligible compared with absorption in the Si films for wavelengths below 800 nm. Also shown, for comparison, is the absorption of a corresponding planar structure with a 2.7-μm-thick c-Si film (dashed line). As can be seen, the planar structure has a much lower absorption than all investigated textured structures, including that with the thinnest c-Si film (0.5 μm).

By assuming the standard AM1.5G solar spectrum (1000 W/cm^2) and an internal quantum efficiency (IQE) of 100% for wavelengths from 400 nm up to c-Si's bandgap wavelength of 1130 nm, the absorption limited current of the structure can be calculated. The resulting currents are given in the legend of Figure 11.5. This shows that, theoretically, excellent short-circuit current densities J_{sc} of almost 32 mA/cm^2 are possible for a 2.7-μm-thick c-Si film on textured glass. The curve for zero Si film thickness shows that parasitic absorption in the 3.3-mm-thick Borofloat glass pane corresponds to a J_{sc} of 1.5 mA/cm^2. Given that highly efficient c-Si thin-film solar cells will require a J_{sc} of at least 30 mA/cm^2, Figure 11.5 shows that a c-Si thickness of at least 2.7 μm seems necessary to obtain such currents, assuming optical constants of high-quality intrinsic bulk c-Si.

It should be noted that in these simulations air was assumed as BSR, which is not normally the case for real solar cells. Air, with a refractive index of 1.0, will maximise total internal reflection

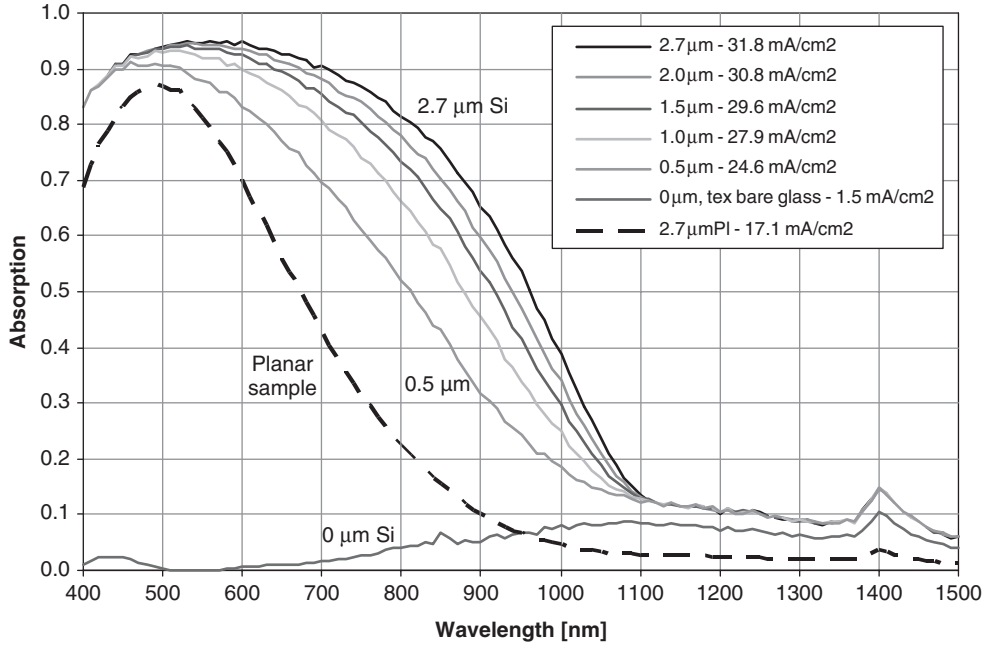


Figure 11.5 Calculated wavelength-dependent absorption of light in an air/glass/SiN/intrinsic c-Si/air structure, for illumination through the 3.3-mm-thick Borofloat glass pane (superstrate configuration). The c-Si film thicknesses are 2.7, 2.0, 1.5, 1.0, 0.5 and 0 µm, whereby the inner glass surface is assumed to be textured in each case. For comparison, the absorption of a corresponding planar structure with a 2.7-µm-thick c-Si film is also shown (dashed line). The current densities shown in the legend are the absorption-limited AM1.5G short-circuit current densities of the corresponding structures, as explained in the text

at the rear. On the other hand, the light that is transmitted through the rear c-Si/air interface can potentially be reflected back into the Si film via a mechanism different from total internal reflection, depending on the BSR properties. Hence, the absorption-limited currents shown in Figure 11.5 are not the true maximum currents available from these structures; nevertheless, they give an excellent indication of the potential of a perfect randomising texture for different c-Si film thicknesses. An additional benefit is that accurate absorption measurements can be performed on air/glass/SiN/c-Si/air samples and compared with the theoretical calculations. Such comparisons will be performed in Section 11.4.2.1.

11.3 CRYSTALLINE SILICON THIN-FILM SOLAR CELLS ON NATIVE AND HIGH-*T* FOREIGN SUPPORTING MATERIALS

11.3.1 Native Supporting Materials

Naturally, the best c-Si thin-films can be grown on native (i.e. c-Si) supporting materials (Si homoepitaxy). A native material eliminates problems associated with a mismatch of the thermal

expansion coefficients of supporting material and c-Si thin-film and allows use of high temperatures (700–1200 °C) for Si film growth. Thin-film growth on native substrates is limited to typical Si wafer sizes and cell interconnection is done wafer by wafer, as in conventional Si wafer PV modules. The standard Si homoepitaxy method is thermal CVD at about 1100 °C, using reactors developed for the microelectronics industry. Solar cell efficiencies of up to 17.6% for 40-μm-thick Si films have been realised with the thermal CVD approach on standard Cz Si wafers [6]. Using liquid phase epitaxy (LPE) at about 900 °C on a thinned (15 μm), lightly doped Fz Si wafer substrate, researchers at the Australian National University (ANU) achieved an efficiency of 18.1% [27] for a 35-μm-thick epitaxial film. These results confirm that the high-*T* approach is capable of high thin-film cell efficiency on high-purity substrates. However, high-purity substrates are incompatible with low-cost thin-film PV cells, and hence a crucial issue with the high-*T* approach is the diffusion of impurities from a lower-purity substrate into the growing silicon film. Application of the above 17.6% CVD solar cell process to lower-purity Si wafer substrates such as block-cast multicrystalline Si wafers and SSP (silicon sheets from powder) wafers reduced the efficiency of the Si thin-film cells to about 13 and 8%, respectively [6]. This shows that in the case of high-*T* silicon thin-film growth, at least a medium-purity silicon substrate seems to be required to obtain reasonable PV efficiency. Alternatively, low-purity Si substrates can be used if coated with a suitable foreign barrier layer material, followed by the preparation of a high-quality seed layer and subsequent epitaxial thickening of the seed layer [28]. This approach has so far produced solar cells with efficiencies of up to 11.3% [28].

The approach of using c-Si wafer substrates for epitaxial thin-film solar cells is now often referred to as EpiWE (epitaxial wafer equivalents) [29]. The approach has recently been thoroughly investigated in a European Union funded project (SWEET), involving several European PV organisations [29]. The SWEET project has shown that, with the exception of the thin-film cell growth step and the surface texturing step, all cell and module fabrication steps are compatible with a standard Si wafer PV line using screen-printed contacts. The EpiWE approach is thus a non-disruptive PV technology that should find its way into the factories, provided it delivers clear cost advantages on a €/W_p basis. The highly doped Si wafer provides the epitaxial template, physical substrate, and ohmic contact for the c-Si thin-film. In the PV industry, surface texturing of Si wafers is done by wet-chemical etching and leads to the removal of a 5- to 10-μm-thick Si layer. This Si loss is unacceptably high for epitaxial Si films because the Si growth rate is only of the order of 1 μm/min and hence significant CVD machine time would be required for the deposition of the 5- to 10-μm-thick sacrificial Si layer. Furthermore, the SWEET project has shown that wet-chemical texture etching of epitaxial Si films gives non-uniform texturisation over large areas. One alternative is to texture the EpiWE substrates prior to epitaxy. Another alternative is plasma texturing of the finished thin-film cells using a fluorine plasma, which has been shown in the SWEET project to provide a uniform texturisation of large-area epitaxial Si films while removing only about 1 μm of silicon. Using large (100 cm^2) heavily doped (~0.02 Ω cm) PV-inactive high-purity block-cast mc-Si wafer substrates and high-*T* silicon growth, screen-printed EpiWE solar cells with good efficiency of about 13.5% have been realised in the SWEET project, using a Si thin-film thickness of about 20 μm. Due to poor light trapping, higher efficiencies would require the deposition of thicker Si films, which is uneconomical with today's methods. The 13.5% efficiency result confirms that the EpiWE technology is compatible with standard PV lines using screen-printed contacts and provides similar cell efficiency. Importantly, the SWEET project also demonstrated that high-quality epitaxial Si films can be grown in an *in-line* CVD reactor. This is an important prerequisite for reducing the cost (€/m^2) of the Si epitaxy step to the low levels required for PV applications.

Compared with the wafers used in standard silicon wafer-based PV technologies, the requirements for EpiWE wafers are greatly relaxed. First, the doping density can be very high and does not have to be precisely controlled. As a matter of fact, any doping that gives a resistivity of less than about 0.5 Ω cm is fine. Second, the average grain size of EpiWE substrates can be as small as 1 mm

without significantly affecting the thin-film cells' efficiency. This grain size will ensure that the grown Si grains (thickness $\ll 100\,\mu m$) are pancake-like, i.e. most cell regions will be located closer to the *p-n* junction than to a grain boundary. Third, the required diffusion length for good collection of light-generated carriers is greatly reduced in thin-film cells, allowing much higher impurity levels (metals, etc.) compared with standard Si wafer cells. As a result, medium-purity multicrystalline silicon wafers made via block-casting or ribbon growth of 'upgraded metallurgical-grade' (UMG) silicon are possibly sufficient for making efficient (>12%) EpiWE solar cells. Assuming low-cost EpiWE substrates and inexpensive silicon epitaxy in an in-line CVD reactor, an economic analysis of PV modules fabricated with the EpiWE approach predicts a significant cost advantage (10–20%) on a €/W_p basis compared to standard Si wafer PV modules [29].

Additional cost savings can possibly be realised by using a low-purity Si wafer substrate in combination with a reflective interlayer [29]. The reflective interlayer acts as a back surface reflector for the thin-film cell, enabling a thinner cell due to light trapping and thus further relaxing the diffusion length requirement. Furthermore, the interlayer acts as a barrier layer between the low-purity substrate and the high-purity epitaxial Si film, enabling a good diffusion length in the thin-film cell. The difficulty with this approach is that the reflective interlayer must not interfere with the silicon epitaxy process, i.e. it must act as a high-quality seed layer for epitaxial thickening. One possible solution is a porous silicon reflector layer [29].

The epitaxial thin-film cells discussed so far are based on low-cost (poly- or multicrystalline) Si wafer substrates, whereby these substrates are destined to end up in the final PV module. An alternative approach is to use a high-quality single-crystal Si wafer substrate, to fabricate a separation layer on/in this substrate, to grow a single-crystalline Si thin-film diode on the separation layer, and then transfer the thin-film diode to a low-cost foreign substrate. The process cycle is then repeated many times, whereby the same Si wafer substrate is reused for epitaxial film growth. To be competitive, the key requirements of these '*silicon transfer processes*' are low material consumption, a simple and robust fabrication process, and a good PV efficiency (>15%).

The so-called ELTRAN (epitaxial layer transfer) process uses a porous Si film as separation layer [30]. The porous silicon layer is formed by anodic electrochemical etching in a solution containing hydrofluoric acid (HF). Depending on the chosen process parameters, 20–90% of the Si within the thickness of the porous Si layer is removed. By means of wafer bonding, the Si film epitaxially grown on the porous silicon layer is attached to a second, oxidised Si wafer. A water jet then splits the starting Si wafer from the epitaxial Si film. The transferred Si film is used in the IC industry for SOI (silicon-on-insulator) applications. The process does not seem to be economic for PV because the Si film is not transferred to a low-cost substrate, however, it has demonstrated that: (i) high-quality Si films can epitaxially be grown on porous Si films; and (ii) that the epitaxial Si film can reliably be separated from the porous Si film and transferred to another substrate.

Another interesting silicon transfer process ('Smart-Cut' [31]) uses hydrogen implantation into a Si wafer substrate for creating a separation layer. The implanted wafer is bonded to another Si wafer or a foreign substrate. During a subsequent high-T anneal the implanted Si layer separates from the parent wafer and is transferred to the new substrate. The process cycle is then repeated, whereby the parent wafer is reused. With respect to PV applications, it appears that the process is suitable for the creation of a single-crystalline silicon seed layer on a high-T stable foreign substrate. However, no solar cell results have as yet been reported for this approach.

In the so-called PSI or Ψ process [32, 33], a (100)-oriented single-crystalline Si wafer is chemically textured (inverted pyramids), followed by the formation of a thin porous Si separation layer. Then a thin-film Si diode is epitaxially grown at high T and transferred to a glass substrate. An efficiency of 12.2% has been realised with this approach [34].

Another layer transfer process was developed by the Sony Corporation in the 1990s. It uses a porous silicon double-layer stack on a singlecrystalline Si wafer substrate [35]. The buried porous layer has a high porosity and serves as the separation layer. The exposed porous layer has a low porosity. During a high-T anneal at above 1000 °C, structural changes occur in the porous double-layer stack. The low-porosity film converts into a single-crystalline film with internal voids (size of the order of 100 nm) and a continuous top surface, whereby this film serves as a seed layer for conventional high-T CVD epitaxy. Following epitaxy, the Si wafer substrate is separated from the thin-film structure and reused. Using a 12-μm-thick epitaxial film that was detached from the substrate, Sony realised 4-cm^2 solar cells with a conversion efficiency of up to 12.5% [35]. Using related approaches, Bergmann *et al*. [36] at the University of Stuttgart achieved 16.6% cell efficiency for a 44.5-μm-thick Si film and Brendel *et al*. [37] at ZAE Bavaria achieved 15.4% efficiency and 32.7 mA/cm^2 short-circuit current density for a 25.5-μm-thick thin-film cell.

In the so-called VEST (via-hole etching for the separation of thin-films) layer transfer process developed by Mitsubishi Electric Corporation [38], a thin fine-grained pc-Si film is deposited by thermal CVD onto an oxidised singlecrystalline Si wafer. Following the deposition of a capping layer, the sample is given a ZMR (zone melt recrystallisation) treatment to convert the Si film into a large-grained, mainly (100)-oriented high-quality pc-Si seed layer. In a ZMR furnace, the sample moves with constant speed through a narrow linear hot zone, created, for example, by concentrated light from a halogen lamp located inside an elliptical mirror [28]. A back-surface field (BSF) layer and a thick absorber layer are then deposited by thermal CVD. An anisotropic etch through a masking layer follows which forms a regular array of 100 μm^2 sized via holes through the Si thin-film structure to the underlying SiO_2. Hole spacing is 1.5 mm. The Si film is then separated from the underlying Si wafer by immersion into HF, which etches the SiO_2 film. The Si thin-film structure is sufficiently thick (~80 μm) to be self-supporting. The front surface is then wet-chemically textured with random upright pyramids, followed by a heavy phosphorus diffusion that generates an n^{++} emitter along the entire front surface and down the via holes. The emitter is then etched back to improve the cell's blue response, followed by the deposition of a SiN antireflection coating by LPCVD. Then, interdigitated contact grids are formed on the back surface by screen-printing. Finally, a hydrogen implantation step is performed to improve the electronic quality of the diode. Using a cell area of 96 cm^2, an efficiency of 16.0% has been realised ($V_{oc} = 589$ mV, $J_{sc} = 35.6$ mA/cm^2, $FF = 76.3\%$).

A summary of the efficiency and voltage results of the most important c-Si thin-film solar cell technologies on native supporting materials developed in recent years is given in Table 11.2. With the exception of ANU's LPE-grown film, all these Si films were grown by high-temperature CVD. None of these technologies has advanced to the pilot line stage as yet. Prototypes of production-type high-T CVD reactors for photovoltaic Si epitaxy purposes are presently being developed [39]. It is expected that such machines will have a Si deposition rate of several μm/min, a throughput of about 150 000 m^2/yr for 10 μm Si film thickness, and costs of about 10 €/m^2 for the silicon epitaxy step.

11.3.2 High-T Foreign Supporting Materials

A list of the most promising c-Si thin-film solar cells on high-T *foreign* supporting materials demonstrated in recent years is given in Table 11.3. The substrates used for these cells include graphite, ceramics, and metals. Despite significant advances, no technology has advanced as yet to the pilot line stage. AstroPower's Silicon-Film™ technology [40] evolved into a thick, self-supporting mc-Si ribbon method and hence is not discussed further here. The most promising high-T supporting materials for thin-film growth appear to be graphite [41] and ceramics such as silicon carbide [42], alumina [43], and glass ceramics [44, 45]. However, major breakthroughs are

Table 11.2 Most important c-Si thin-film solar cells on native supporting materials developed in recent years. Ranking is based on the reported cell efficiencies. Also shown is the best reported V_{oc} of the corresponding technology (note that the best V_{oc} and *Eff* were not necessarily obtained on the same solar cell)

Organisation	*Eff* [%]	V_{oc} [mV]	Reference	Status	Details
Australian National University	18.1	666	[27]	lab	LPE on lightly doped, thinned Fz Si wafer substrate, 35 µm Si film, substrate thinned to 15 µm
Fraunhofer ISE	17.6	661	[6]	lab	EpiWE, Cz Si wafer substrate, 37 µm Si film
University of Stuttgart	16.6	645	[36]	lab	Silicon transfer, 45 µm Si film
Mitsubishi Electric	16.0	589	[38]	lab	Silicon transfer (VEST), 77 µm pc-Si, area 96 cm^2, seed layer formed by ZMR
ZAE Bavaria	15.4	623	[37]	lab	Silicon transfer, 25.5 µm Si film
Fraunhofer ISE, IMEC	15.2	649	[29]	lab	EpiWE, Cz Si wafer substrate, ~20 µm Si film
ZAE Bavaria	12.2	600	[34]	lab	Si transfer (PSI), 15.5 µm Si film
Fraunhofer ISE	11.3	578	[28]	lab	35 µm Si film on a SSP Si substrate coated with an insulating barrier layer. Seed layer formed by ZMR

Table 11.3 Most important c-Si thin-film solar cells on high-temperature foreign supporting materials developed in recent years. Ranking is based on the reported cell efficiencies. Also shown is the best reported V_{oc} of the corresponding technology (note that the best V_{oc} and *Eff* were not necessarily obtained on the same solar cell)

Organisation	Eff [%]	V_{oc} [mV]	Reference	Status	Details
Fraunhofer ISE	11.0	570	[41]	lab	SiC-coated graphite substrate, Si ~55 µm (includes ~40 µm BSF layer), high-T Si deposition, seed layer treated by ZMR
Fraunhofer ISE	9.3	567	[42]	lab	SiO/SiN/SiO-coated SiSiC substrate, Si film ~60 µm (includes ~40 µm BSF), high-T Si deposition, seed layer ZMR treated
IMEC	8.0	536	[43]	lab	AIC-seeded ceramic (alumina), Si thickness ~2.5 µm (p^+p), high-T Si deposition, a-Si hetero-emitter, substrate configuration
IMEC	5.4	539	[44, 45]	lab	AIC-seeded glass-ceramic substrate (Corning 9664), Si film ~2 µm (p^+p), high-T Si deposition, a-Si hetero-emitter

required before these PV technologies can be commercialised. As with the high-T native substrates, the most important high-T Si deposition method is thermal CVD.

The common issues of high-T-resistant foreign supporting materials are cost (€/m^2), availability, and impurity content. Furthermore, many foreign materials (such as graphite) have a porous structure, which makes conventional wet-chemical processing difficult or even impossible. One

way to overcome the problems associated with impurities and porosity is to encapsulate the foreign supporting material with a smoothening barrier layer. Examples for barrier layer materials are silicon oxide, silicon nitride, and silicon carbide.

In 1997, a collaboration between ASE GmbH (now Schott Solar) and Fraunhofer ISE resulted in 11.0% efficient pc-Si thinfilm solar cells on silicon carbide (SiC) coated graphite substrates [41]. Upon encapsulation of the entire graphite substrate with SiC (whereby the SiC film on the top surface was conducting and that on the bottom surface was insulating), a fine-grained 40-μm-thick p^+-doped pc-Si film was deposited by thermal CVD. This p^+ film was then subjected to a ZMR step, producing a large-grained p^+ layer that served both as the seed layer for the epitaxial growth of the absorber layer and as the BSF layer of the finished solar cell. The active layer (thickness in the 15–30 μm range) was deposited by thermal CVD. The n^+ emitter was formed by conventional phosphorus diffusion. The solar cell process used dry processing (reactive ion etching, RIE) instead of wet-chemical processing. Cell area was 1 cm^2. Contact to the bottom surface of the graphite substrate was realised by drilling of an array of holes and subsequent evaporation of an Al film. The 11% cell had a V_{oc} of 570 mV, a J_{sc} of 25.6 mA/cm^2 and a FF of 75.7%. Albeit a respectable efficiency result, significant improvements are required to make the technology economical compared with the EpiWE technology discussed in Section 11.3. The same conclusion applies to the thin-film cells made at Fraunhofer ISE using SiC substrates [42].

Seed layers formed by AIC (aluminium-induced crystallisation) of amorphous silicon are also an interesting route for epitaxial solar cells on high-T-resistant substrates. Although these were developed for intermediate-T substrates such as borosilicate glass (see Section 11.4.2), researchers at IMEC in Belgium have shown that the AIC process also works well on high-temperature substrates such as alumina [43] and glass ceramics [44, 45]. Prior to the AIC process, the alumina substrates are coated with a spin-on flowable oxide (Fox-25 from Dow Corning) to reduce the surface roughness. Using thermal CVD for epitaxial solar cell growth on AIC-seeded alumina substrates, efficiencies of up to 8.0% and voltages of up to 536 mV have been realised [43]. The total Si film thickness is in the range 2–4 μm. The n^+ emitter is a heterojunction emitter formed at low temperature by depositing doped amorphous silicon, similar to Sanyo's HIT Si wafer PV technology [46]. The V_{oc} of the IMEC thin-film cells drops significantly if the heterojunction emitter is replaced by a phosphorus-diffused homojunction emitter, a behaviour attributed to phosphorus smearing along grain boundaries during the high-T diffusion process, increasing the junction area and leading to enhanced $n = 2$ junction recombination losses [43]. The IMEC team has also applied the process to high-temperature-resistant glass ceramics (Corning 9664), giving 5.4% cell efficiency [44]. However, glass ceramics are too expensive for PV applications and hence such substrates are merely useful for process development and/or academic purposes. On the other hand, ceramics such as alumina seem inexpensive enough to be of interest for PV. To become industrially relevant, cell efficiencies of at least 12–13% will be required for the AIC-seeded alumina approach. This efficiency level should be possible because the present 8% cells still have a very high density ($\sim 10^9$ cm^{-2}) of electrically active intragrain defects [47], possibly due to an insufficient quality of the used AIC seed layers.

11.4 CRYSTALLINE SILICON THIN-FILM SOLAR CELLS ON INTERMEDIATE-T FOREIGN SUPPORTING MATERIALS

A summary of the most efficient c-Si thin-film solar cells developed in recent years on *intermediate-temperature* (450–700 °C) foreign supporting materials is given in Table 11.4. The silicon films of all these cells are polycrystalline and the substrate is borosilicate glass, with the exception of the Sanyo cells which were made on metal substrates. In Section 11.4.1, the Sanyo cells will be described in some detail. Section 11.4.2 then describes the solar cells fabricated on borosilicate

Table 11.4 Most efficient c-Si thin-film solar cells on intermediate-temperature foreign supporting materials developed in recent years. Ranking is based on the reported cell efficiencies. Also shown is the best reported V_{oc} of the corresponding technology (note that the best V_{oc} and *Eff* were not necessarily obtained on the same solar cell)

Organisation	*Eff* [%]	V_{oc} [mV]	Reference	Status	Details
CSG Solar	10.4	492/cell	[58]	factory	CSG, pc-Si on borosilicate glass superstrate, SPC of PECVD a-Si, Si film 2.2 µm, grown emitter, module area 94 cm^2
UNSW	9.3	528	[79]	lab	PLASMA, pc-Si on borosilicate glass superstrate, SPC of PECVD a-Si, Si film 2–5 µm, grown emitter, cell 4.4 cm^2
Sanyo Electric	9.2	553	[54]	abandoned	Pc-Si on metal, Si film ~5 µm, SPC (10 h, 600 °C) of PECVD n^+n^- a-Si structure, a-Si made at 550–650 °C, hetero-emitter
UNSW	5.8	517	[79]	lab	SOPHE, pc-Si on SPC-seeded borosilicate glass superstrate, SPE (17 h, 550 °C) of PECVD or evaporated a-Si structure
UNSW	5.2	517	[92]	lab	EVA, pc-Si on borosilicate glass superstrate, SPC of evaporated a-Si, Si film 1.5–3 µm, grown emitter
IPHT	4.8	510	[99]	lab	Pc-Si on borosilicate glass superstrate, seed layer and absorber layer both laser crystallised (ISC-CVD cells). Seed layer and absorber layer by hot PECVD (~600 °C)
UNSW	4.8	480	[79]	lab	ALICE, pc-Si on AIC-seeded borosilicate glass superstrate, SPE (17 h, 550 °C) of PECVD or evaporated a-Si structure

glass. The 10% efficient STAR solar cell on glass [48] developed in the 1990s by Yamamoto *et al.* at Kaneka is not included here because it is a microcrystalline silicon cell. This is a mixed-phase material that contains both crystalline and amorphous silicon regions. Microcrystalline silicon cells are discussed in Chapter 12 of this book.

11.4.1 Solar Cells on Metal

Solid phase crystallisation (SPC) of amorphous silicon films has been researched for more than 30 years, see for example references [49] and [50]. The SPC method is interesting for PV as it is robust, has a high yield, is inexpensive and scalable. Polycrystalline silicon films made by the SPC method have been used in integrated active-matrix liquid crystal displays [51] and hence the SPC process development has benefited enormously from the resources of the semiconductor industry.

Solar cells using SPC of amorphous silicon were pioneered by the Japanese company Sanyo Electric in the late 1980s and early 1990s, using steel and quartz as substrates and PECVD for a-Si:H deposition [52–55]. The Sanyo team developed a 'partial doping method' where a two-layer stack of amorphous silicon material (thickness in the range 1–20 µm) was crystallised during several hours at about 600 °C by means of SPC. The team showed that excellent-quality polycrystalline silicon films can be obtained by SPC if the two-layer amorphous stack consists of a thin n^+

doped (phosphorus) layer and a much thicker undoped (or lightly doped) layer. The reason for the excellent material quality produced by this method is the directional crystallisation of the entire stack, starting in the n^+ doped layer (nucleation layer) and then progressing through the remainder of the structure (crystallisation layer). Matsuyama *et al*. [52–55] showed that a heavily phosphorus-doped a-Si layer crystallises much more rapidly than a lightly doped (or undoped) a-Si layer of the same thickness, and hence acts as an excellent nucleation layer in a n^+n (or n^+i) structure. To realise a diode structure from the solid-phase crystallised n^+n stack, Matsuyama *et al*. deposited an intrinsic a-Si:H/p^+ a-Si:H bi-layer onto the *n*-type pc-Si layer at low temperature, giving a heterojunction diode consisting of both crystalline and amorphous silicon material.

Using the substrate configuration, the Sanyo team fabricated solar cells according to the following layer sequence: Metal substrate/n^+ pc-Si/*n* pc-Si/*i* a-Si:H/p^+ a-Si:H/ITO. Measurements under standard test conditions (one sun) revealed that these low-temperature fabricated heterojunction silicon thin-film solar cells on metal exhibit good PV efficiencies of up to 9.2% [54]. The absorber layer thickness was about 5 µm, the cell area 1 cm^2, the short-circuit current density 25 mA/cm^2, and the minority carrier diffusion length in the absorber region about 10 µm. The Sanyo SPC cells were made on textured substrates [53, 54]. Although the enhanced optical absorption due to the textured substrate was noted and attributed to light scattering [55], it appears that the substrate texture was primarily introduced as a method of improving the SPC silicon material quality, with light trapping as a welcome side effect. It should be noted that no post-crystallisation treatments such as rapid thermal annealing or hydrogenation (see Section 11.4.2.4) were applied to the Sanyo SPC cells [52–55]. The 9% efficiency obtained in 1996 by the Sanyo team represents a significant achievement in the history of photovoltaics. Surprisingly, since 1996 Sanyo has not published any new results on SPC pc-Si thin-film solar cells. It thus seems that the SPC pc-Si thin-film approach has been abandoned at Sanyo around 1996 and replaced by the so-called HIT solar cell which uses a similar heterojunction structure with an a-Si:H emitter, but with a single-crystalline *n*-type silicon wafer as absorber material [46].

11.4.2 Solar Cells on Glass

Glass is an excellent foreign supporting material for thin-film PV. It is inexpensive, mechanically stable, readily available in large quantities and in large substrate sizes, long-term stable in outdoor conditions, and features excellent moisture barrier properties. It is also highly transparent for most of the solar photons, enabling the use of the glass pane as front cover sheet of the PV module (superstrate configuration). Last but not least, glass can be textured in such a way that significant light trapping is realised in c-Si thin-film solar cells formed on the textured glass panes.

In this section, we first discuss the glass texturing methods that have been developed in recent years for thin-film PV applications. Then follow absorption measurements on pc-Si thin-film diodes fabricated on textured glass. We then present two pc-Si thin-film cell metallisation methods that have been shown to be capable of producing one-sun fill factors of over 70% for pc-Si thin-film solar cells on glass. The section concludes with a survey of the pc-Si on glass thinfilm solar cell fabrication methods that have produced solar cells with reasonable efficiencies.

11.4.2.1 Texturing of glass

Glass is an isotropic material – to create a textured glass surface with chemical etching is thus not easy. Nevertheless, researchers at Pacific Solar (now CSG Solar) have developed a method that uses hydrofluoric acid to which another chemical compound (for example $BaSO_4$) was added to partially mask the glass surface during etching, giving a textured surface [56]. Pacific Solar has not

published any solar cell results based on this texturing method and has not revealed further details regarding the texture.

Another texturing method developed by Pacific Solar uses a liquid surface coating (sol–gel) containing SiO_2 spheres ('glass beads') to fabricate a textured glass surface [57]. The glass bead method is currently used in CSG Solar's PV factory in Germany. For mini-modules, CSG Solar has reported J_{sc} values of up to 25.6 mA/cm^2 using this texturing method, for a pc-Si film thickness of 1.6 μm [58].

A more obvious method for texturing glass surfaces is sandblasting [59]. CSG Solar recently realised a pc-Si on glass mini-module with excellent efficiency (10.4%) and current (29.5 mA/cm^2) for a 2.2-μm-thick silicon film formed on a sandblasted and subsequently hydrofluoric acid (HF) etched glass pane [58]. The purpose of the HF etch is to smoothen the extremely damaged and rough glass surface after sandblasting.

Another approach to texture glass surfaces is embossing. The idea behind embossing is to press a tool that has a certain texture into heated glass and thus create a textured glass surface [60]. Pressing at temperatures below the softening point of the glass is necessary for creating small-sized features. However, as yet, the smallest feature size obtained with the embossing method is of the order of 10 μm [61], which is not optimal for the desired c-Si film thicknesses of 1–3 μm.

Another glass texturing method uses plasma etching or reactive ion etching (RIE) with or without natural lithography [62–65]. So far, the light trapping results obtained with these methods (see, for example, reference [64]) cannot compete with those obtained using the glass bead method or the sandblasting/wet etching method described above.

The present authors have developed another glass texturing method [66]. It is referred to as 'aluminium-induced texture' (AIT). A thin sacrificial Al film is deposited (by evaporation or sputtering) onto planar glass, followed by annealing at intermediate temperatures (~600 °C) in an inert atmosphere. The anneal initiates a redox reaction whereby aluminium is oxidised to Al_2O_3 and SiO_2 from the glass is reduced to silicon, as follows:

$$4Al + 3SiO_2 \rightarrow 3Si + 2Al_2O_3 \tag{11.6}$$

A surface texture is imparted to the glass according to the nucleation conditions provided during this anneal. Subsequent wet-chemical etching removes the reaction products from the glass surface and reveals the glass texture. A schematic representation of the AIT process is shown in Fig. 11.6. Figure 11.7 shows an example of the type of surface morphology and cross-section obtained for pc-Si films formed on AIT glass.

Excellent absorption enhancements have recently been achieved with PECVD SPC pc-Si thinfilm diodes formed on AIT glass sheets, for a range of pc-Si thicknesses [67]. As an example, Figure 11.8 shows absorption measurements (in superstrate configuration) on an air/AIT glass/SiN/n^+pp^+ pc-Si diode/air structure, for a pc-Si thickness of 2.7 μm and 1.15 μm. For each pc-Si thickness, an optimised AIT process was used. The absorption was determined via $A = 1 - R - T$, where R and T are the hemispherical reflectance and hemispherical transmission measured with an integrating sphere attached to an optical dual-beam spectrometer (Cary 5G from Varian). As can be seen, at 800 nm, the AIT glass provides an optical absorption of over 80% for a pc-Si film thickness of 2.7 μm. For a pc-Si film thickness of 1.15 μm, the corresponding value is still well above 60%. For comparison, a corresponding planar sample (Si thickness 2.7 μm) is included in Figure 11.8. As can be seen, its optical absorption at long wavelengths (>600 nm) is much poorer than that of all textured samples and is clearly insufficient for realising efficient pc-Si solar cells.

Also shown in Figure 11.8 are the theoretically expected absorption limits, assuming Lambertian scattering at the textured surfaces. These theoretical curves were calculated using the method

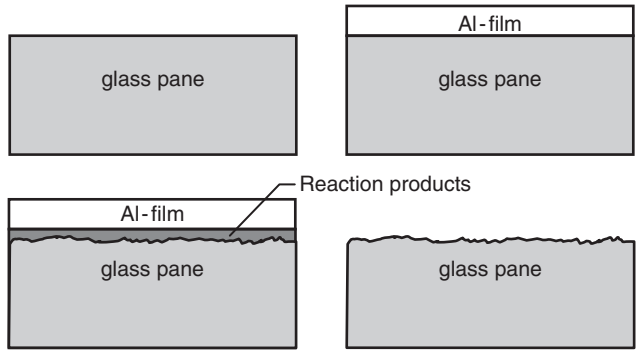
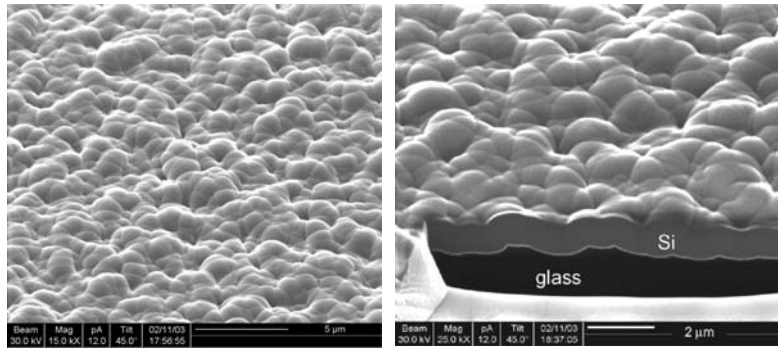


Figure 11.6 Schematic representation of the AIT glass texturing process



(a) (b)

Figure 11.7 Focus ion beam microscope images of (a) the surface morphology (from reference [70]; reproduced by permission of Hindawi) and (b) the cross-section of a pc-Si film formed on SiN-coated AIT glass (from reference [15]; reproduced by permission of Elsevier). The pc-Si film was formed by SPC of PECVD-deposited a-Si. Note that the images have different scales: image (a) shows a 22-µm-wide region, image (b) a 13-µm-wide region

described in Section 11.2.3. As can be seen, for both pc-Si thicknesses (2.7 and 1.15 µm), the samples made on AIT glass have essentially reached the Lambertian absorption limit. The optical model, which assumes the optical constants of high-quality intrinsic c-Si and experimentally determined optical constants of 3.3-mm-thick Borofloat glass panes, significantly underestimates the parasitic absorption in textured samples at long wavelengths (>1100 nm). As shown elsewhere, this difference is largely due to free carrier absorption in the thin heavily doped regions of the silicon diode. Free carrier absorption is insignificant in planar samples for all investigated wavelengths and in textured samples for short wavelengths (<800 nm). As already mentioned in Section 11.2.3, the remaining absorption at long wavelengths (>1100 nm) is largely due to absorption in the textured glass pane. Additional possible reasons include parasitic absorption at grain boundaries and/or intragrain defects, and measurement errors.

11.4.2.2 Back-surface reflector

The combination of a low near-infrared absorption coefficient of c-Si and a desired silicon thickness of well below 10 µm results in a large fraction of the incident photons not being absorbed in their

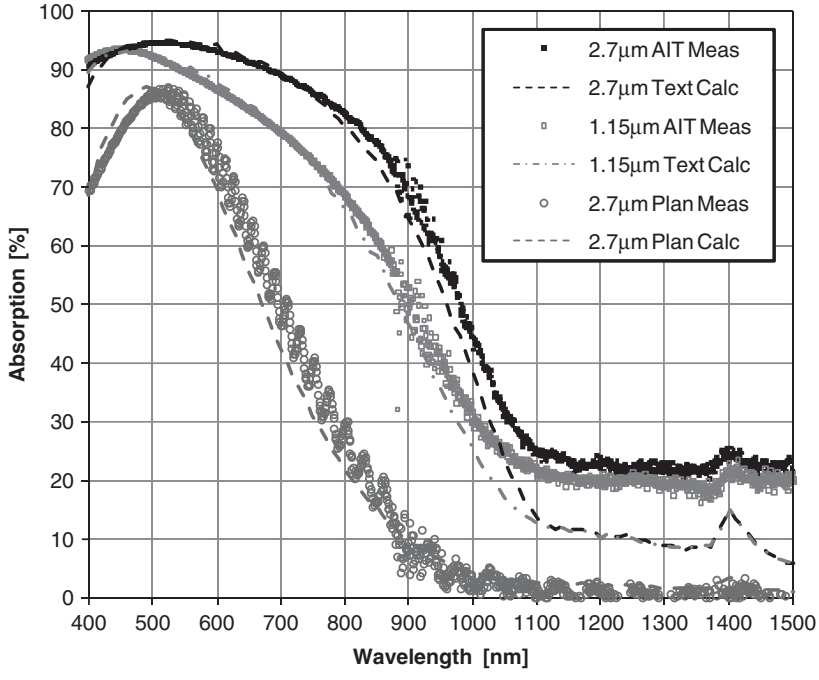


Figure 11.8 Measured absorption (in superstrate configuration) of air/AIT glass/SiN/n^+pp^+ pc-Si diode/air structures, for a pc-Si thickness of 2.7 µm and 1.15 µm. For comparison, a corresponding planar sample (Si thickness 2.7 µm) is included. Also shown (dashed lines) are the theoretically expected absorption limits using the method described in Section 11.2.3. The theoretical curves assume the optical properties of intrinsic c-Si, i.e. free carrier absorption is not included

first pass. These photons reach the rear surface where they interact with the back-surface reflector (BSR). The decay of the photon intensity as a function of the distance from the illuminated c-Si surface is shown in Figure 11.9, for several wavelengths. For example, about 90% of 800-nm photons entering a 1.1-µm-thick c-Si film will reach the rear surface and interact with the BSR. Moreover, for samples with light trapping, this interaction will be repeated several or even many times. Hence, the optical properties of the BSR are of paramount importance for maximising light trapping in pc-Si thin-film solar cells.

One can distinguish between two different strategies when trying to optimise the optical properties of the BSR: The first relies on the mechanism of total internal reflection at the silicon/BSR interface. If the light beam does not arrive perpendicularly at the rear surface (for example due to a textured superstrate), total internal reflection can be used to prevent the photons from entering the BSR and thus enabling multiple passes. The two critical parameters here are the refractive index of the BSR and the angle with which the light beam strikes the rear c-Si surface. The lower the refractive index of the BSR, the more normally can the light beam strike the rear surface and still be totally internally reflected. Snell's law quantifies the critical angle below which the light beam can leave the semiconductor and above which it will be totally internally reflected. The situation is schematically shown in Figure 11.10. For the case of a semiconductor/air interface, the critical angle θ_c is given by

$$\theta_c = \sin^{-1}(n_{air}/n_s), \tag{11.7}$$

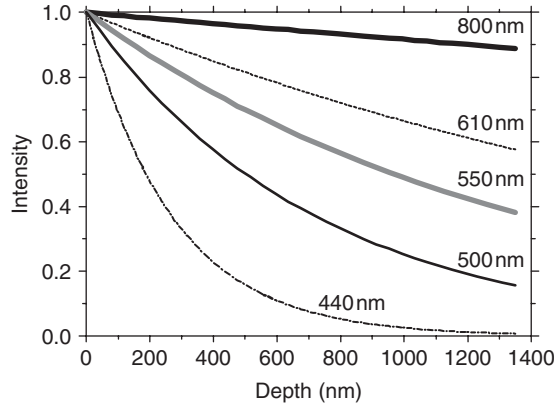


Figure 11.9 Decay of the photon intensity in crystalline silicon, for five free-space wavelengths (from reference [20]; reproduced by permission of Daniel Inns)

where n_{air} is the refractive index of air and n_s the refractive index of the semiconductor, for the wavelength under consideration. For the system c-Si/air, the critical angle for 1000-nm radiation ($n_s = 3.57$) is 16.3°. In three dimensions, the critical angle defines a cone. Rays inside this cone (i.e., $\theta_1 < \theta_c$) can escape from the dense material (escape cone), whereas rays outside the escape cone (i.e., $\theta_1 \geq \theta_c$) are totally internally reflected. Since semiconductors have rather large refractive indices, the escape cone is rather narrow and hence it is difficult for radiation to escape from semiconductors into air. For a BSR other than air, the refractive index of air is simply replaced with that of the BSR. Table 11.5 shows the refractive indices [68, 69] and the associated critical angles θ_c for several important BSR materials. As already mentioned in Section 11.2.3, air as a BSR maximises the chances for total internal reflection. As shown in the literature [70–72], Al is a poor BSR on c-Si, and particularly so on textured Si surfaces. However, if a SiO_2 layer is sandwiched between the c-Si and the aluminium, parasitic optical absorption in the aluminium layer

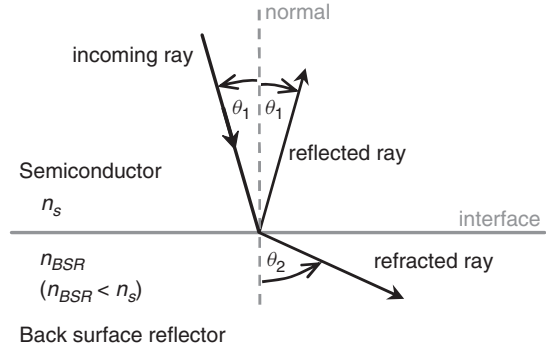


Figure 11.10 At the interface between two different materials, radiation is totally or partly reflected. The direction of the refracted ray is given by Snell's law. The graph shows partial reflection of a ray at the interface between a semiconductor and an optically less dense material (back-surface reflector). For a certain value of θ_1 (the critical angle θ_c), the angle θ_2 reaches 90°. For θ_1 values exceeding the critical angle, refraction is no longer possible and the entire intensity of the incoming ray is reflected at the interface (total internal reflection)

Table 11.5 Refractive indices n_{BSR} at about 1.5 eV photon energy of important back-surface reflector materials [68, 69] and the corresponding critical angle θ_c for a silicon-BSR system as obtained from Equation (11.7)

Material	n_{BSR}	θ_c
Air	1.00	16.9°
SiO_2	1.46	25.1°
ITO	1.65	28.6°
SiN (plasma-deposited)	2.00	35.5°
ZnO	2.02	35.9°
Crystalline silicon	3.44	not applicable

is significantly reduced for both planar and textured c-Si films [73], a behaviour that is at least partly due to total internal reflection at the Si/SiO_2 interface.

The second strategy, which can be used separately or in combination with the first one, is to allow the photons to enter the BSR and design the BSR in such a way that most photons are returned to the c-Si film. For a planar sample without any light trapping features, most non-absorbed photons hit the rear surface about normally (i.e. $\theta_1 \approx 0°$) and hence this second strategy is the only viable light trapping strategy. An excellent example for this approach is to apply a layer of white paint to the rear surface of a planar pc-Si film formed on a glass superstrate. In such devices, a large fraction of the incoming photons enters the BSR. As a result of random optical Mie scattering of photons by the pigments held in suspension by the optical binder in white paint, a large fraction of the photons entering the BSR are diffusely scattered back into the c-Si film, as discussed in reference [71]. The operating principle of such a 'pigmented diffuse reflector' (PDR) as BSR for thin-film silicon solar cells on glass is shown in Figure 11.11. A PDR consist of pigments (usually TiO_2) held in suspension by a medium (usually an organic binder). Due to the pigments in the paint, many of the photons entering the BSR are diffusely scattered back into the silicon. Hence, white paint not only returns a large fraction of the photons back to the c-Si film, but also acts as an active light trapping component.

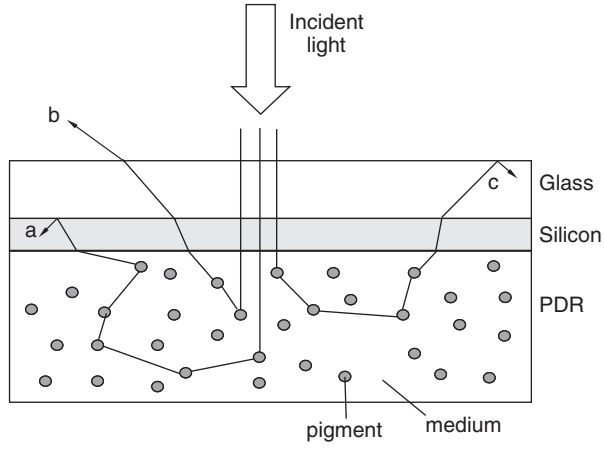


Figure 11.11 The operating principle of a pigmented diffuse reflector (PDR) on the back surface of a thin-film silicon solar cell on glass (from reference [71]; reproduced by permission of Elsevier)

Since 80–95% of white paint consists of an organic binder with a low refractive index in the range of 1.4 [74] to 1.6, application of a layer of white paint to the textured rear surface of a pc-Si solar cell on a glass superstrate: (i) enables a large fraction of the photons arriving at the c-Si/paint interface to be total internally reflected; and (ii) returns a large fraction of the photons entering the paint back to the pc-Si film due to scattering by the pigments. Hence, white paint on a textured pc-Si surface utilises both light trapping strategies mentioned above, resulting in very high short-circuit current densities in experimental devices. CSG Solar uses a white resin layer as a PDR [75] on textured pc-Si surfaces. The resin is believed to contain similar ingredients as white paint (i.e. an organic binder with TiO_2 pigments) and has enabled J_{sc} values of up to 29.5 mA/cm^2 for 2.2 μm thick textured pc-Si solar cells on glass [58].

It should be noted that the PDR concept requires a fairly good lateral conductance of the pc-Si rear surface layer because the PDR is an electrical insulator and hence needs to be opened locally (point or line contacts) to enable electrical contact between the rear metal electrode and the pc-Si film. As a result, a PDR cannot be efficiently applied to a-Si:H and μc-Si:H solar cells. Due to their high sheet resistance ($>10^4$ Ω/sq), these solar cell technologies need to rely on a transparent conductive oxide (TCO)/metal structure for obtaining a good rear conductance and a good rear reflectance. As discussed in the literature [72, 76, 77], the TCO layer must be optimised carefully since otherwise there may be significant parasitic (free carrier) absorption at wavelengths above 600 nm [77].

11.4.2.3 Metallisation

To enable extraction of power from a solar cell, contacts need to be created to the negative and positive terminals of the device and conductive paths (usually made of metal) need to deliver the current and voltage out from the device. Hence, all solar cells have a metallisation process that fabricates such contacts and conductive pathways. Due to the large size of thin-film PV modules on glass, it is important to divide the large (∼1 m^2) initial thin-film solar cell into smaller unit cells and then interconnect them in series to keep ohmic losses at a tolerable level.

There are two reasons why pc-Si thin-film PV on glass can't use the established method of monolithic series interconnection of the individual solar cells that is used in the a-Si:H thin-film PV industry for modules on glass/TCO substrates [78]: First, TCOs are not sufficiently temperature stable to withstand the high temperatures (>600 °C) used during some of the fabrication steps of pc-Si thin-film solar cells, eliminating the possibility to fabricate the pc-Si solar cell on a TCO layer (front TCO). Secondly, doped pc-Si layers have a much higher electrical conductance (i.e. a much lower sheet resistance) than doped a-Si:H layers and hence the thin-film cells would be severely shunted when a TCO film is deposited over their exposed side walls to connect the rear surface of one cell with the front surface of the neighbouring cell.

One method for metallising pc-Si thin-film solar cells on glass has been developed in recent years at UNSW [73, 79]. It involves two photolithography steps. The method is schematically shown in Figure 11.12. A thin (∼100 nm) SiO_2 layer is deposited onto the entire rear surface of the solar cell, followed by the formation of a matrix of round openings (diameter ∼30 μm, spacing ∼80 μm, surface coverage ∼5%) in the SiO_2 layer. Various methods can be used to create these openings, for example via a conventional photolithography-based sequence involving wet-chemical etching (whereby no alignment is required for the photomask) or by controlled deposition of small droplets of hydrofluoric (HF) acid. The SiO_2 layer is deposited via RF sputtering at room temperature. Next, an approximately 600-nm-thick Al layer is blanket deposited over the structure, for example using dc magnetron sputtering at room temperature. The purpose of this SiO_2/Al stack is to provide both the rear electrode of the solar cell and a high-quality back-surface reflector (BSR). As outlined previously, the SiO_2 layer boosts light trapping in the solar cell due to total

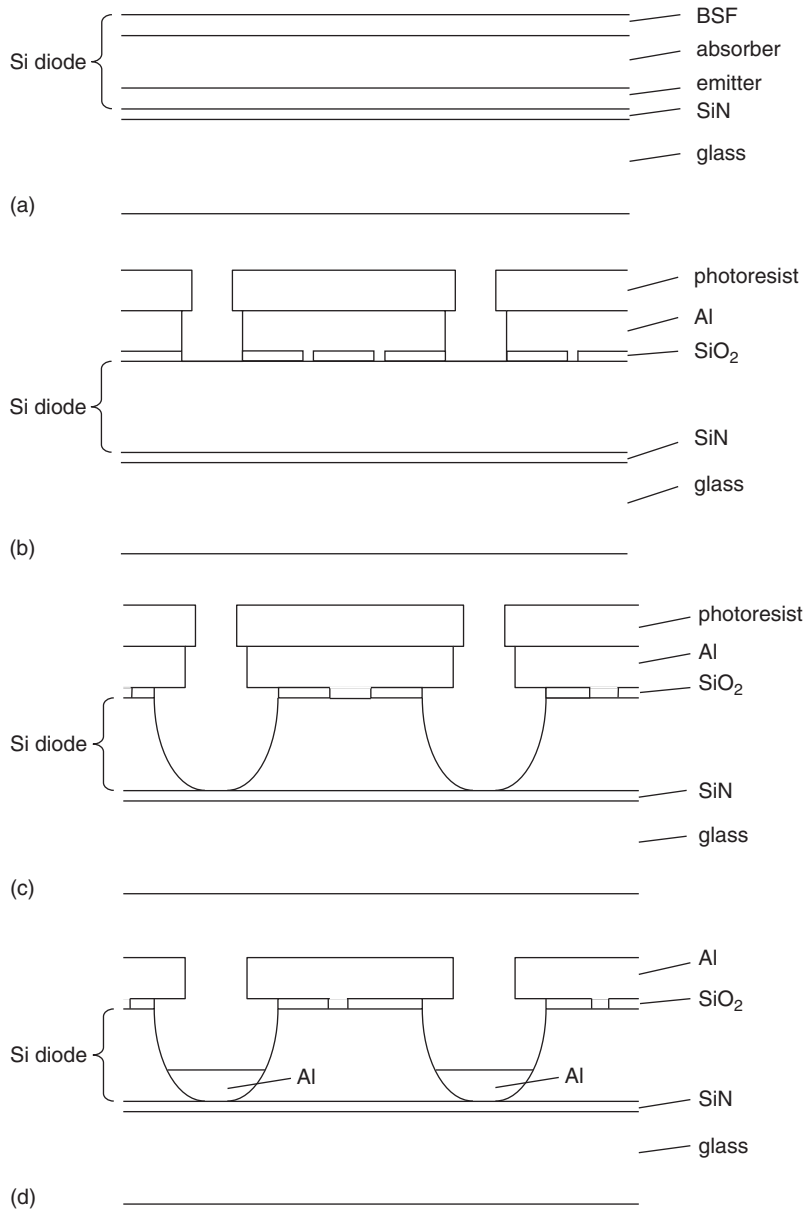


Figure 11.12 Schematic representation of the metallisation method developed at UNSW for pc-Si thin-film solar cells on glass. (a) initial structure; (b) before the plasma etching step for the emitter electrode (the openings in the rear contact stack were formed by photolithography and wet-chemical etching); (c) after the plasma etching step for the emitter electrode; (d) final structure

internal reflection at the c-Si/SiO_2 interface. Next, a photoresist film is blanket deposited onto the Al film and patterned using a photomask with a conventional comb-like structure. This structure defines the location of the emitter (i.e. the glass-side) electrode. This patterning process requires no alignment of the photomask on the sample. A wet-chemical etching step then removes the aluminium and SiO_2 below the openings in the photoresist film, thus locally exposing the pc-Si diode (see Figure 11.12b). Next, U-shaped grooves are etched into the Si film using a dry etching process (plasma etching) in a conventional 13.56-MHz parallel-plate plasma etcher, with SF_6 as etching gas. The resulting structure is shown in Figure 11.12c. After a brief HF dip, a 600-nm-thick Al film is then deposited via e-beam evaporation. The photoresist and the overlying Al film are then lifted off by ultrasonic treatment in an acetone solution, giving the final structure shown in Figure 11.12d.

At UNSW, the interdigitated metallisation process described above is routinely used for the metallisation of four individual solar cells on 5×5 cm glass sheets. The metallisation process has proved to be robust and to routinely generate cells with fill factors of over 70%. The best *FF* realised as yet is 75.9% and has been obtained on a cell with an area of 4.4 cm^2. This is believed to be the highest fill factor ever obtained for a pc-Si thin-film solar cell on glass and is a clear proof of the potential of the method. The metallisation process has two photolithography steps, but neither requires alignment of the photomask. Since the UNSW pc-Si solar cells have a short high-temperature ($>900\,^{\circ}C$) defect anneal step which produces some glass deformation, precise alignment of photolithography masks would be cumbersome, slow and expensive. We also note that a low-cost LED array is used as UV light source for exposure of the photoresist, a method that can easily be scaled to large areas. Large-scale photoresist deposition via slit coating using micro-nozzles which spray the photoresist onto very large ($>1\,m^2$) glass substrates is now a standard process in the LCD flat panel industry [80]. However, drawbacks of this metallisation method are that there is no PDR at the rear solar cell surface, that two separate aluminium depositions are needed, and that an extra processing step is required to series-connect the individual solar cells. Work is underway at UNSW to simplify the process, to incorporate a PDR, and to series-connect neighbouring solar cells.

Another method of forming a series-connected thin-film PV module based on polycrystalline silicon has been disclosed by Basore [75]. The technology is referred to as CSG (crystalline silicon on glass) and is the only pc-Si on glass technique that has entered industrial production. Device fabrication starts by using a pulsed laser to slice the Si layer into a series of adjacent, ~6-mm-wide strip cells. The module is then coated with a thin layer of novolac resin loaded with white pigments to make it more reflective and thus improve light trapping in the cell. Next the openings for the *n*-type emitter contacts (craters) are formed. This involves etching of openings into the resin layer (using an inkjet printhead), followed by chemical etching of the Si. Then the openings for the *p*-type rear contacts (dimples) are formed using the same inkjet process. A blanket deposition of sputtered aluminium provides electrical contact to the n^+ and p^+ Si layers. The aluminium film is then sliced into a large number of individual pads using laser pulses. Each metal pad series connects one line of *p*-type contacts in one cell with a line of *n*-type contacts in the next cell. The final structure is shown in Figure 11.13. It is noted that this metallisation and interconnection scheme again does not involve a TCO layer.

Advantages compared with the UNSW metallisation scheme discussed above are that only one metal deposition step is involved, that a PDR is incorporated, and that the solar cells are automatically series-connected. One challenge with the Basore technique [75] is the large number (millions/m^2) of craters and dimples that need to be created. Another is that all craters and dimples need to be accurately positioned across the entire module, imposing significant challenges with respect to the alignment of the glass sheet and the patterning tools (such as inkjet or laser).

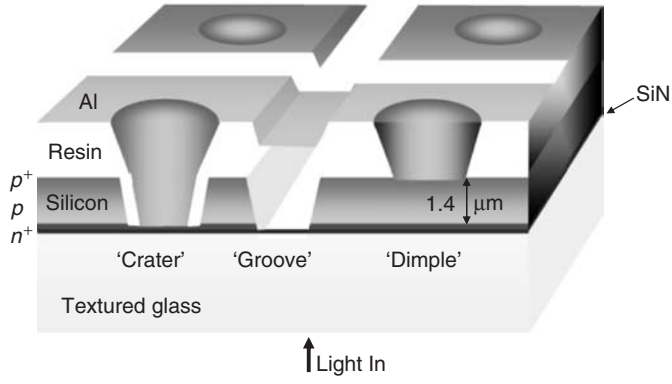


Figure 11.13 Schematic of CSG technology (from reference [75]; reproduced by permission of CSG Solar)

11.4.2.4 Solid-phase crystallised pc-Si cells on borosilicate glass

PECVD cells. Pacific Solar was formed in 1995 (i.e., shortly before Sanyo Corp stopped their SPC research) as a spin-off from the University of New South Wales (UNSW) in Sydney. In 2002 the company reported [81] that it had developed a SPC pc-Si solar cell technology on textured SiN-coated borosilicate glass superstrates [82]. The function of the SiN was two-fold, first as an antireflection layer, but also as a diffusion barrier preventing impurities from the glass sheets reaching the pc-Si diode [83]. Unlike the Sanyo SPC pc-Si cells (which use an a-SiH/pc-Si heterojunction), the Pacific Solar (now CSG Solar) diode (see Figure 11.13) is a homojunction device with the intended structure of glass/SiN/n^+ pc-Si/p^- pc-Si/p^+ pc-Si, whereby the solar cell is designed for the superstrate configuration. Also in contrast to Sanyo, CSG Solar uses two crucial post-SPC process steps: defect anneal and hydrogenation. These will be discussed in more detail below. CSG Solar's advanced light trapping structure, back-surface reflector and metallisation process have been discussed above. The best PV efficiency realised so far is 10.4%, achieved in 2007 for a mini-module with 94 cm^2 aperture area [58]. The J_{sc} is 29.5 mA/cm^2, the pc-Si film has a thickness of 2.2 µm, and the glass sheet was textured by sandblasting and subsequent HF etching, as discussed in Section 11.4.2.1.

The flat panel display industry has induced a rapid development of very-large-scale PECVD tools that are capable of 24/7 operation. CSG Solar has tapped into these developments and is using a modified large-scale PECVD machine designed originally for flat panel displays. The a-Si:H (and the SiN) is deposited in a batch-type multi-chamber PECVD tool (KAI-1200 from Oerlikon) with a silicon deposition rate of approximately 30–40 nm/min [84]. Each KAI-1200 machine processes 20 glass sheets (1.10 m × 1.25 m) in parallel [85]. The CSG Solar factory in Thalheim, Germany has a rated capacity of about 20 MW_p/year [86].

Since 2004, a related SPC pc-Si solar cell (PLASMA) has been developed at UNSW [70, 73], using different methods for glass texturisation (AIT), a-Si:H diode deposition, defect anneal, and metallisation. PLASMA cells with efficiencies of up to 9.3% have been obtained as yet, using a 2.3-µm-thick pc-Si absorber region and a cell area of 4.4 cm^2. Further efficiency improvements are expected by incorporation of recent improvements with the AIT glass texturing method, see Section 11.4.2.1. UNSW uses a conventional parallel-plate PECVD machine (a cluster tool from MVSystems, USA) capable of processing 15 cm × 15 cm glass substrates. The plasma excitation frequency is 13.56 MHz and the resulting a-Si:H deposition rate is similar to that of the KAI-1200.

One advantage of the SPC method is that the Si deposition rate is not limited by the electronic properties of the as-deposited material (as is the case for a-Si:H solar cells) – instead, these are established during subsequent thermal processing steps [85]. As pointed out in reference [8]5, CSG Solar's KAI-1200 PECVD tool costs nearly as much as all other process tools combined. Hence, one possible way of significantly decreasing the cost of CSG modules is to strongly increase the a-Si:H deposition rate. In the literature it is well known that an increasing plasma excitation frequency leads to higher a-Si:H deposition rates. For example, reference [87] reports the in-line deposition of device-grade a-Si:H with a static deposition rate of up to 120 nm/min, using a 81-MHz linear plasma source. A project is under way at UNSW that investigates the benefits of VHF PECVD for SPC pc-Si solar cells on glass. Significant cost reductions of the PECVD SPC technology, due to increased PECVD a-Si:H deposition rates, seem possible in the near future.

Two additional pc-Si thin-film solar cells (ALICE and SOPHE) recently developed at UNSW involve the fabrication of a thin ($\sim$100 nm) pc-Si seed layer on the SiN-coated glass pane [16, 79]. This seed layer is either formed by AIC (aluminium-induced crystallisation) of an intrinsic a-Si layer at about 500 °C (ALICE cells) or by SPC of a heavily doped a-Si layer at about 600 °C (SOPHE cells). The heavily doped a-Si layer (n^+ or p^+) is formed by either e-beam evaporation or PECVD, whereas the intrinsic a-Si layer is formed by dc magnetron sputtering or PECVD. Details on the AIC and SPC seed layers can be found elsewhere [88, 89]. Upon completion of the seed layer, the next step is the deposition (by PECVD or evaporation) of a much thicker pp^+ or nn^+ a-Si structure onto the heavily doped seed layer, followed by crystallisation of the a-Si structure by SPE, post-deposition treatment of the samples (RTA, hydrogenation), and metallisation. The best efficiencies realised as yet with these seed layer solar cells are 5.8% (SOPHE) and 4.8% (ALICE), obtained using PECVD for the absorber layer deposition [79]. These efficiencies are much lower than those obtained with standard (i.e. non-seeded) solar cells, suggesting that the seed layer concept does not look promising. However, the seeded solar cells made as yet are still in an early stage of development and hence it is premature to judge their industrial relevance.

Evaporated cells. E-beam evaporation was the original method of depositing a-Si films as a precursor for SPC [90]. The rationale for using evaporated amorphous silicon is due to very high deposition rates of up to 1000 nm/min and the avoidance of toxic and expensive gases such as silane [91]. UNSW started developing evaporated SPC pc-Si cells (EVA cells) in 2002 [79]. Planar EVA cells now have efficiencies of up to 5.2% [92], i.e. about the same as for planar PECVD-deposited SPC pc-Si solar cells. Further efficiency improvements will require the use of textured glass sheets. As mentioned in references [25] and [93], this could cause problems due to the directional nature of the evaporation process (line-of-sight method), leading to non-conformal coating of textured surfaces. Research towards improved light trapping in evaporated SPC pc-Si solar cells is on-going at UNSW.

Hot-wire CVD cells. The hot-wire CVD method is also capable of high silicon deposition rates ($\gg$100 nm/min). In contrast to evaporation this method should not have problems with coating of textured surfaces, due to the high gas pressure. However, no efficiency results for SPC hot-wire CVD solar cells have as yet been published. As shown by researchers at NREL [94], a-Si:H deposited by hot-wire CVD tends to have very high nucleation rates, which leads to smaller SPC pc-Si grains than for PECVD-deposited material.

Defect removal and passivation. Post-SPC processing steps are well known in the literature to improve the structural and electronic properties of SPC pc-Si films. For example, in 1994 Dyer *et al*. [95] showed that exposure of a PECVD-deposited SPC pc-Si film on a low-strain-point glass substrate to a hydrogen plasma leads to a significant (about 10 times) increase of the film's photoconductivity. They also found a substantially reduced density of recombination centres in SPC pc-Si films after hydrogen plasma treatment. Also in 1994, Morita *et al*. [96] from Sharp Corporation suggested that the excimer-laser-based defect annealing step on SPC pc-Si thin-film

transistors (TFTs) on glass should be replaced by a rapid thermal anneal (RTA) step using halogen lamps. In 1998 Girginoudi *et al*. [97] showed that a RTA step at 850 °C for 45 s significantly decreases the density of intra-grain defects in SPC pc-Si films. In 2004 Green *et al*. [82] reported that CSG Solar uses a RTA defect anneal step, followed by a hydrogenation step to improve the efficiency of their SPC pc-Si solar cells on glass. Keevers *et al*. showed in 2005 [98] that a high-temperature remote-plasma hydrogenation process more than quadruples the efficiency of CSG modules. They also mentioned that the RTA step used by CSG Solar heats the silicon film to above 900 °C. Note that this is well above the softening point of the glass, but it is of very short duration (of the order of 1 minute). Also in 2005 Terry *et al*. [5] showed that a RTA at 900 °C provides point defect annealing and dopant activation for evaporated SPC pc-Si solar cells on glass and that the open-circuit voltage of the samples increased by more than a factor of three using a 900 °C RTA process and a subsequent rf (13.56 MHz) parallel-plate hydrogen plasma treatment at around 480 °C glass temperature. From the results published in the literature, it follows clearly that: (i) a high-T RTA provides point defect annealing and dopant activation in SPC pc-Si solar cells on glass; and (ii) a subsequent 'hot' (500–650 °C) hydrogenation process passivates a large fraction of the remaining electronically active defects.

11.4.2.5 Laser-crystallised PECVD cells on borosilicate glass

An interesting alternative to SPC for forming pc-Si thin-film solar cells on glass is the layered laser crystallisation (LLC) process [99, 100]. The LLC process consists of two steps: (i) a seed layer is generated by laser crystallisation; (ii) epitaxial thickening is achieved by continually depositing a-Si and periodically scanning the surface with a UV laser to induce crystallisation in a layered fashion. Transmission electron microscopy (TEM) analysis of LLC material reveals an excellent structural quality of the pc-Si films [100]. Pc-Si thin-film solar cells on planar glass using the LLC process, whereby the absorber layer was deposited by PECVD, have achieved efficiencies of up to 4.8% [99]. This is a respectable result, considering that the cells do not feature a light trapping scheme.

11.5 CONCLUSIONS

In this chapter, the theoretical fundamentals and the technological status of the most efficient crystalline silicon thin-film solar cells developed during the past 15 years have been reviewed. The strong interest in these solar cells results from the belief that they will enable long-term stable (>25 years) and efficient (>10%) PV modules that can be manufactured at much lower costs (€/W_p) than standard silicon wafer-based PV modules and without any material supply and toxicity issues. From the theoretical analysis of the solar cells it follows that 10% PV module efficiency will require very good light trapping properties, a diffusion length in the absorber layer of at least several micrometres, and a cell thickness of at least 1.5 µm. Progress during the past 15 years has been excellent and the first c-Si thin-film PV technology has entered commercial production (CSG Solar in Germany), delivering large (1.4 m^2) modules with stable efficiencies in the 6–8% range. The rated capacity of the CSG factory is 20 MW_p/year and the estimated fabrication costs of CSG modules are reported to be around 1.5 €/W_p [75]. The CSG technology uses a very thin (>2 µm) pc-Si film on a borosilicate glass superstrate. Progress has also been very good with other c-Si thin-film PV technologies, and several of these (including PLASMA, EVA, EpiWE, silicon on graphite or alumina) are now ready for testing at the pilot line scale. Thus, while c-Si thin-film PV technologies have proved difficult to commercialise in the past, it appears that the rate of progress is accelerating and that there is now a high probability that inexpensive 10% efficient pc-Si thin-film modules will be commercially available within 10 years. This assessment applies to

both silicon thin-film solar cells fabricated via high-temperature approaches (for example EpiWE) and silicon thin-film solar cells fabricated via intermediate-temperature approaches (for example CSG or PLASMA).

ACKNOWLEDGEMENTS

The work at the University of New South Wales (UNSW) described in this chapter has been supported by the Australian Research Council (via its Centres of Excellence scheme), the State Government of New South Wales, and UNSW. The Solar Energy Research Institute of Singapore (SERIS) is supported by the Singapore Government and the National University of Singapore.

REFERENCES

1. Hirshman WP, Hering G, Schmela M, Market survey on cell and module production 2006, *Photon International*, March 2007, 136 (2007).
2. http://www.schott.com/hometech/english/products.
3. http://www.corning.com.
4. Straub A, Harder NP, Huang Y, Aberle AG, High-quality homoepitaxial silicon growth in a non-ultra-high vacuum environment by ion-assisted deposition, *Journal of Crystal Growth*, **268**, 41–51 (2004).
5. Terry ML, Straub A, Inns D, Song D, Aberle AG, Large open-circuit voltage improvement by rapid thermal annealing of evaporated solid-phase-crystallized thin-film silicon solar cells on glass, *Applied Physics Letters*, **86**, 172108–172110 (2005).
6. Faller FR. Hurrle A, High-temperature CVD for crystalline-silicon thin-film solar cells, *IEEE Transactions on Electron Devices*, **46**, 2048–2054 (1999).
7. McCann MJ, Catchpole KR, Weber KJ, Blakers AW, A review of thin-film crystalline silicon for solar cell applications. Part 1: Native substrates, *Solar Energy Materials and Solar Cells*, **68**, 135–171 (2001).
8. Foxon CT, Three decades of molecular beam epitaxy, *J. Crystal Growth*, **251**, 1–8 (2003).
9. Itoh T, Nakamura T, Epitaxial growth of silicon assisted by ion implantation, *Radiation Effects*, **9**, 1 (1971).
10. Zotov AV, Korobtsov VV, Present status of solid phase epitaxy of vacuum-deposited silicon, *Journal of Crystal Growth*, **98**, 519 (1989).
11. Shah AV, Schade H, Vanecek M, Meier J, Vallat-Sauvain E, Wyrsch N, Kroll U, Droz C, Bailat J, Thin-film silicon solar cell technology, *Progress in Photovoltaics*, **12**, 113–142 (2004).
12. Keppner H, Meier J, Torres P, Fischer D, Shah A, Microcrystalline silicon and micromorph tandem solar cells, *Applied Physics A*, **69**, 169–177 (1999).
13. Bergmann RB, Crystalline Si thin-film solar cells: a review, *Applied Physics A*, **69**, 187–194 (1999).
14. Catchpole KR, McCann MJ, Weber KJ, Blakers AW, A review of thin-film crystalline silicon for solar cell applications. Part 2: Foreign substrates, *Solar Energy Materials and Solar Cells*, **68**, 173–215 (2001).
15. Aberle AG, Fabrication and characterisation of crystalline silicon thin-film materials for solar cells, *Thin Solid Films*, **511–512**, 26–34 (2006).
16. Aberle AG, Widenborg PI, Straub A, Harder NP, Polycrystalline silicon on glass thin-film solar cell research at UNSW using the seed layer concept, *Proc. 3rd World Conference on Photovoltaic Energy Conversion*, Osaka, 2003, pp. 1194–1197.
17. Sze SM, Physics of Semiconductor Devices, 2nd edn (John Wiley & Sons, Inc., New York, 1981).

18. Shockley W, The theory of p-n junctions in semiconductors and p-n junction transistors, *Bell System Technical Journal*, **28**, 435 (1949).
19. Green MA, Accuracy of analytical expressions for solar cell fill factors, *Solar Cells*, **8**, 3–16 (1983).
20. Inns D, *PhD Thesis*, School of Photovoltaic and Renewable Energy Engineering, University of New South Wales, Sydney, Australia, 2007.
21. Basore PA, Numerical modeling of textured silicon solar cells using PC1D, *IEEE Transactions on Electron Devices*, **27**, 337 (1990).
22. Aberle AG, Untersuchungen zur Oberflächenpassivierung von hocheffizienten Silicium-Solarzellen, *PhD Thesis*, Faculty of Physics, University of Freiburg, Germany (1991).
23. Fahrenbruch AL, Bube RH, *Fundamentals of Solar Cells* (Academic Press, New York, 1983).
24. Aberle AG, Surface passivation of crystalline silicon solar cells: a review, *Progress in Photovoltaics*, **8**, 473–487 (2000).
25. Campbell P, Widenborg PI, Sproul A, Aberle AG, Surface textures for large-grained polysilicon thin-film solar cells on glass using the AIT method, *Proc. 15th International Photovoltaic Science and Engineering Conference*, Shanghai, 2005, 859–860.
26. Nagel H, Aberle AG, R. Hezel R, Optimised antireflection coatings for planar silicon solar cells using remote PECVD silicon nitride and porous silicon dioxide, *Progress in Photovoltaics*, **7**, 245–260 (1999).
27. Blakers AW, Weber KJ, Stuckings MF, Armand S., Matlakowski G, Stocks MJ, Cuevas A, 18% efficient thin silicon solar cell by liquid phase epitaxy, *Proc. 13th European Photovoltaic Solar Energy Conference*, Nice, 1995, 33–36.
28. Kieliba T, Bau S, Osswald D, Eyer A, Coarse-grained Si films for crystalline Si thin-film solar cells prepared by zone-melting recrystallization, *Proc. 17th European Photovoltaic Solar Energy Conference*, Munich, 2001, 1604–1607 (WIP, Munich, 2001).
29. Reber S, Duerinckx F, Alvarez M, Garrard B, Schulze FW, EU project SWEET on epitaxial wafers equivalents: Results and future topics of interest, *Proc. 21st European Photovoltaic Solar Energy Conference*, Dresden, 2006, 570–576 (WIP, Munich, 2006).
30. Yonehara T, Sakaguchi K, Sato N, Epitaxial layer transfer by bond and etch back of porous Si, *Applied Physics Letters*, **64**, 2108 (1994).
31. Bruel M, Aspar B, Auberton-Herve AJ, Smart-Cut: A new silicon on insulator material technology based on hydrogen implantation and wafer bonding, *Japanese Journal of Applied Physics*, **36**, 1636–1641 (1997).
32. Brendel R, A novel process for ultrathin monocrystalline silicon solar cells on glass, *Proc. 14th European Photovoltaic Solar Energy Conference*, Barcelona, 1997, p. 1354 (Stephens, Bedford, 1997).
33. Brendel R, Review of layer transfer processes for crystalline thin-film silicon solar cells, *Japanese Journal of Applied Physics*, **40**, 4431–4439 (2001).
34. Brendel R, Auer R, Artmann H, Textured monocrystalline thin-film Si cells from the porous silicon (PSI) process, *Progress in Photovoltaics*, **9**, 217–221 (2001).
35. Tayanaka H, Yamauchi K, Matsushita T, Thin-film crystalline silicon solar cells obtained by separation of a porous silicon sacrificial layer, *Proc. 2nd World Conference on Photovoltaic Solar Energy Conversion*, Vienna, 1998, p. 1272 (European Commission, Ispra, 1998).
36. Berge C, Bergmann RB, Rinke TJ, Werner JH, Monocrystalline silicon thin film solar cells by layer transfer, *Proc. 17th European Photovoltaic Solar Energy Conference*, Munich, 2001, 1277–1281 (WIP, Munich, 2001).
37. Feldrapp K, Horbelt R, Auer R, Brendel R, Thin-film (25.5 μm) solar cells from layer transfer using porous silicon with 32.7mA/cm^2 short-circuit current density, *Progress in Photovoltaics*, **11**, 105–112 (2003).

38. Morikawa H, Nishimoto Y, Naomoto H, Kawama Y, Takami A, Arimoto S, Ishihara T, Namba K, 16.0% efficiency of large area (10 cm × 10 cm) thin film polycrystalline silicon solar cell, *Solar Energy Materials and Solar Cells*, **53**, 23–28 (1998).
39. Reber S, Schillinger N, Bau S, Waldenmayer B, Progress in high-temperature silicon epitaxy using the RTCVD160 processor, *Proc. 19th European Photovoltaic Solar Energy Conference*, Paris, 2004, 471–474 (WIP, Munich, 2004).
40. Bai Y, Ford DH, Rand JA, Hall RB, Barnett AM, 16.6% efficient Silicon-Film™ polycrystalline silicon solar cells, *Proc. 26th IEEE Photovoltaic Specialists Conference*, Anaheim, 1997, 35–38 (IEEE, New York, 1997).
41. Lüdemann R, Schaefer S, Schüle C, Hebling C, Dry processing of mc-silicon thin-film solar cells on foreign substrates leading to 11% efficiency, *Proc. 26th IEEE Photovoltaic Specialists Conference*, Anaheim, 1997, 159–162 (IEEE, New York, 1997).
42. Reber S, Faller FR, Hebling C, Lüdemann R, Crystalline silicon thin-film solar cells on SiC based ceramics, *Proc. 2nd World Conference Photovoltaic Solar Energy Conversion*, Vienna, 1998, 1782–1785 (European Commission, Ispra, 1998).
43. Gordon I, Carnel L, Van Gestel D, Beaucarne G., Poortmans J, 8% efficient thin-film polycrystalline-silicon solar cells based on aluminum-induced crystallization and thermal CVD, *Progress in Photovoltaics*, **15**, 575–586 (2007).
44. Gordon I, Van Gestel D, Carnel L, Beaucarne G, J. Poortmans J, Pinckney L, Mayolet A, Thin-film polycrystalline-silicon solar cells on high-temperature substrates by aluminium-induced crystallization, *Proc. 21st European Photovoltaic Solar Energy Conference*, Dresden, 2006, 992–995 (WIP, Munich, 2006).
45. Beaucarne G, Gordon I, Van Gestel D, Carnel L, Poortmans J, Thin-film polycrystalline silicon solar cells: An emerging photovoltaic technology, *Proc. 21st European Photovoltaic Solar Energy Conference*, Dresden, 2006, pp. 721–725 (WIP, Munich, 2006).
46. Taguchi M, Kawamoto K, Tsuge S, Baba T, Sakata H, Morizane M, Uchihashi K, Nakamura N, Kiyama S, Oota O, HIT™ Cells – High-efficiency crystalline Si cells with novel structure, *Progress in Photovoltaics*, **8**, 503–513 (2000).
47. Van Gestel D, Romero MJ, Gordon I, Carnel L, D'Haen J, Beaucarne G, Al-Jassim M, Poortmans J, Electrical activity of intragrain defects in polycrystalline silicon layers obtained by aluminum-induced crystallization and epitaxy, *Applied Physics Letters*, **90**, 0921103 (2007).
48. Yamamoto K, Very thin film crystalline silicon solar cells on glass substrate fabricated at low temperature, *IEEE Transactions on Electron Devices*, **46**, 2041–2047 (1999).
49. Harbeke G, Krausbauer L, Steigmeier EF, Widmer AE, Kappert HF, Neugebauer G, High quality polysilicon by amorphous low pressure chemical vapor deposition, *Applied Physics Letters*, **42**, 249–251 (1983).
50. Spinella C, Lombardo S, Crystal grain nucleation in amorphous silicon, *Journal of Applied Physics*, **84**, 5383–5414 (1998).
51. Im JS, Sposili RS, Crystalline Si films for integrated active-matrix liquid-crystal displays, *MRS Bulletin*, March 1996, 39–48 (1996).
52. Matsuyama T, Wakisaka K, Kameda M, Tanaka M, Matsuoka T, Tsuda S, Nakano S, Kishi Y, Kuwano Y, Preparation of high-quality n-type poly-Si films by the solid phase crystallisation (SPC) method, *Japanese Journal of Applied Physics*, **29**, 2327–2331 (1990).
53. Matsuyama T, Baba T, Takahama T, Tsuda S, Nakano S, Polycrystalline Si thin-film solar cell prepared by solid phase crystallization (SPC) method, *Solar Energy Materials and Solar Cells*, **34**, 285–289 (1994).
54. Matsuyama T, Terada N, Baba T, Sawada T, Tsuge S, Wakisaka K, Tsuda S, High-quality polycrystalline silicon thin film prepared by a solid phase crystallisation method, *Journal of Non-Crystalline Solids*, **198–200**, 940–944 (1996).

55. Matsuyama T, Tanaka M, Tsuda S, Nakano S, Kuwano Y, Improvements of n-type poly-Si film properties by solid phase crystallization method, *Japanese Journal of Applied Physics*, **32**, 3720–3728 (1993).
56. Shi Z, Wenham SR, Green MA, Basore PA, Ji JJ, Thin films with light trapping, *USA patent* US6538195 (2003).
57. Ji JJ, Shi Z, Texturing of glass by SiO_2, *European patent EP1142031* (2000).
58. Keevers MJ, Young TL, Schubert U, Green, MA 10% efficient CSG minimodules, *Proc. 22nd European Photovoltaic Solar Energy Conference*, Milan, 2007, pp 1783–1790 (WIP, Munich, 2007).
59. Tomandl G, Determination of light-scattering properties of glass surfaces, *Journal of Non-Crystalline Solids*, **19**, 105–113 (1975).
60. Bradshaw JM, Gelder R, Method of producing a surface microstructure on glass, *USA patent* US005090982 (1992).
61. Campbell P, Enhancement of absorption in silicon films using a pressed glass substrate texture, *Glass Technology*, **43**, 107–111 (2002).
62. Deckman HW, Dunsmuir JH, Natural lithography, *Applied Physics Letters*, **41**, 377–379 (1982).
63. Ruby D, Zaidi SH, Metal catalyst technique for texturing silicon solar cells, International PCT patent application WO 0213279 A2 (2002).
64. Niira K, Senta H, Hakuma H, Komoda M, Okui H, Fukui K, Arimune H, Shirasawa K, Thin film poly-Si solar cells using PECVD and Cat-CVD with light confinement structure by RIE, *Solar Energy Materials and Solar Cells*, **74**, 247–253 (2002).
65. Gandon C, Marzolin C, Rogier B, and Royer E, Transparent textured substrate and methods for obtaining same, International PCT patent application WO 0202472 A1 (2002).
66. Aberle AG, Widenborg PI, Chuangsuwanich N, Glass texturing, International PCT patent application WO 04089841 A1 (2004).
67. Jin G, Widenborg PI, Campbell P, Varlamov S, Enhanced light trapping in SPC poly-Si thin film solar cells on aluminium induced textured glass superstrates, *Proc. 18th International Photovoltaic Science and Engineering Conference*, Kolkata, 2009, paper 4-2o-012 (Editors: Swati Ray and Parsathi Chatterjee, Indian Association for the Cultivation of Science, Kolkata, 2009).
68. Pankove JI, Optical Processes in Semiconductors (Englewood Cliffs,Prentice Hall, 1971).
69. Weast RC, Astle MJ, Beyer WH, CRC Handbook of chemistry and physics, 67th edn (CRC Press Inc, Boca Raton, 1987).
70. Widenborg PI, Aberle AG, Polycrystalline silicon thin-film solar cells on AIT-textured glass superstrates, *Advances in OptoElectronics*, **vol. 2007**, article ID 24584, 7 pages, doi:10.1155/2007/24584. Article freely available from publisher at www.hindawi.com/journals/aoe/.
71. Berger O, Inns D, Aberle AG, Commercial white paint as back surface reflector for thin-film solar cells, *Solar Energy Materials and Solar Cells*, **91**, 1215–1221 (2007).
72. Müller J, Rech B, Springer J, Vanecek V, TCO and light trapping in silicon thin film solar cells, *Solar Energy*, **77**, 917–930 (2004).
73. Widenborg PI, Chan SV, Walsh T, Aberle AG, Thin-film poly-Si solar cells on AIT-textured glass – Importance of the rear reflector, *Proc. 33rd IEEE Photovoltaic Specialists Conference*, San Diego, 2008, paper 34-05 (IEEE, New York, 2008).
74. Sands S, Defining luminous effects, *Just Paint*, **12**, 8–9 (2004).
75. Basore BA, Simplified processing and improved efficiency of crystalline silicon on glass modules, *Proc. 19th European Photovoltaic Solar Energy Conference*, Paris, 2004, pp. 455–458 (WIP, Munich, 2004).

76. Shah A, Meier J, Buechel A, Kroll U, Steinhauser J, Meillaud F, Schade H Dominé D, Towards very low-cost mass production of thin-film silicon photovoltaic (PV) solar modules on glass, *Thin Solid Films*, **502**, 292–299 (2006).
77. Springer J, Rech J, W. Reetz W, Müller J, Vanecek M, Light trapping and optical losses in microcrystalline silicon pin solar cells deposited on surface-textured glass/ZnO substrates, *Solar Energy Materials and Solar Cells*, **85**, 1–11 (2005).
78. Kuwano K, Tsuda S, Onishi M, Nishikawa H, Nakano S, Imai T, A new integrated type amorphous Si solar cell, *Japanese Journal of Applied Physics*, **20**, 213–218 (1981).
79. Aberle AG, Widenborg PI, Campbell P, Sproul A, Griffin M, Weber JW, Beilby B, Inns D, Terry M, Walsh T, Kunz O, He S, Tsao CY, Ouyang Z, Wong J, Hoex B, Shi L, Sakano T, Wolf M, Huang J,. Jin G, Huang L, Peng S, Lang M, Schmunk D, Bamberg F, Chan SV, Han J, Ruof T, Berger O, Di D, Fattal A, Gress P, Pelletier M, Mitchell E, Zhou Y, Fecker F, Pohlner S, Poly-Si on glass thin-film PV research at UNSW, *Proc. 22nd European Photovoltaic Solar Energy Conference*, Milan, 2007, pp. 1884–1889 (WIP, Munich, 2007).
80. Takayasu K, Method of and apparatus for application of liquid, International patent application, WO94/27737 (1994).
81. Basore PA, Pilot production of thin-film crystalline silicon on glass modules, *Proc. 29th IEEE Photovoltaic Specialists Conference*, New Orleans, 2002, pp. 49–52 (IEEE, New York, 2002).
82. Green MA, Basore, PA, Chang N, Clugston D, Egan R, Evans R, Hogg D, Jarnason S, Keevers M, Lasswell P, O'Sullivan J, Schubert U, Turner A, Wenham SR, Young T, Crystalline silicon on glass (CSG) thin-film solar cell modules, *Solar Energy*, **77**, 857–863 (2004).
83. Quinn LJ, Mitchell SJN, Armstrong BM, Gamble HS, Plasma-enhanced silicon nitride deposition for thin film transistor applications, *Journal of Non-Crystalline Solids*, **187**, 347–352 (1995).
84. Egan RJ, Young TL, Evans R, Schubert U, Keevers M, Basore PA, Wenham SR, Green MA, Silicon deposition optimization for peak efficiency of CSG modules, *Proc 21st European Photovoltaic Solar Energy Conference*, Dresden, 2006, pp. 874–876.
85. Basore PA, CSG-1: Manufacturing a new polycrystalline silicon PV technology, *Proc. 4th World Conference on Photovoltaic Energy Conversion*, Hawaii, 2006, pp. 2089–2093.
86. See www.csgsolar.com.
87. Strobel C, Zimmermann T, Albert M, Bartha JW, Beyer W, Kuske J, Dynamic high-rate-deposition of silicon thin film layers for photovoltaic devices, *Proc. 23rd European Photovoltaic Solar Energy Conference*, Valencia, 2008, pp. 2497–2504 (WIP, Munich, 2008).
88. Widenborg PI, Aberle AG, Surface morphology of poly-Si films made by aluminium-induced crystallisation on glass substrates, *Journal of Crystal Growth*, **242**, 270–282 (2002).
89. Goldschmidt JC, Roth K, Chuangsuwanich N, Sproul AB, Vogl B, Aberle AG, Electrical and optical properties of polycrystalline silicon seed layers made on glass by solid-phase crystallisation, *Proc. 3rd World Conference on Photovoltaic Energy Conversion*, Osaka, 2003, pp. 1206–1209.
90. Blum NA, Feldman C, The crystallization of amorphous silicon films, *Journal of Non-Crystalline Solids*, **11**, 242–246 (1972).
91. Aberle AG, Progress with polycrystalline silicon thin-film solar cells on glass at UNSW, *Journal of Crystal Growth*, **287**, 386–390 (2006).
92. Kunz O, Ouyang Z, Varlamov S, Aberle AG, 5% efficient evaporated solid-phase crystallised polycrystalline silicon thin-film solar cells, *Progress in Photovoltaics*, **17**, 567–573 (2009).
93. Ouyang Z, Kunz O, Wolf M, Widenborg P, Jin G, Varlamov S, Challenges of evaporated solid-phase-crystallised poly-Si thin-film solar cells on textured glass, *Proc. 18th International Photovoltaic Science and Engineering Conference*, Kolkata, 2009, paper 1-3p-023 (Swati Ray and Parsathi Chatterjee, eds, Indian Association for the Cultivation of Science, Kolkata, 2009).

94. Young DL, Stradins P, Xu Y, Gedvilas L, Reedy B, Mahan AH, Branz HM, Wang Q, and Williamson DL, Rapid solid-phase crystallization of high-rate, hot-wire chemical-vapor-deposited hydrogenated amorphous silicon, *Applied Physics Letters*, **89**, 161910–1 (2006).
95. Dyer TE, Marshall JM, Davies JF, Optoelectronic properties of polycrystalline silicon produced by low-temperature (600°C) solid-phase crystallization of hydrogenated amorphous silicon, *Philosophical Magazine B*, **69**, 509–523 (1994).
96. Morita T, Tsuchimoto S, Hashizume N, The low temperature polysilicon TFT technology for manufacturing of active matrix liquid crystal displays, *Material Research Society Symposium Proceedings*, **345**, 71–80 (1994).
97. Girginoudi S, Girginoudi D, Thanailakis A, Georgoulas N, Papaioannou V, Electrical and structural properties of poly-Si films grown by furnace and rapid thermal annealing of amorphous Si, *Journal of Applied Physics*, **84**, 1968–1972 (1998).
98. Keevers MJ, Turner A, Schubert U, Basore PA, Green MA, Remarkably effective hydrogenation of crystalline silicon on glass modules, *Proc. 20th European Photovoltaic Solar Energy Conference*, Barcelona, 2005, pp. 1305–1308 (WIP, Munich, 2005).
99. Andrä G, Bochmann A, Falk F, Gawlik A, Ose E, Plentz, Diode laser crystallized multicrystalline silicon thin film solar cells on glass, *Proc. 21st European Photovoltaic Solar Energy Conference*, Dresden, 2006, pp. 972–975 (WIP, Munich, 2006).
100. Falk F, Andrä G, Laser crystallization – a way to produce crystalline silicon films on glass or on polymer substrates, *Journal of Crystal Growth*, **287**, 397–401 (2006).

12
Amorphous Silicon-based Solar Cells

Eric A. Schiff[1], Steven Hegedus[2] and Xunming Deng[3]

[1]*Department of Physics, Syracuse University, Syracuse, New York, USA,* [2]*Institute of Energy Conservation, University of Delaware, Newark, Delaware, USA,* [3]*Department of Physics and Astronomy, University of Toledo, Toledo, Ohio, USA*

12.1 OVERVIEW

12.1.1 Amorphous Silicon: The First Dopable Amorphous Semiconductor

Crystalline semiconductors are very well known, including silicon (Si) (the basis of the integrated circuits used in modern electronics), Ge (the material of the first transistor), GaAs and the other III–V compounds (the basis for many light emitters), and CdS (often used as a light sensor). In crystals, the atoms are arranged in near-perfect, regular arrays or *lattices*. Of course the lattice must be consistent with the underlying chemical bonding properties of the atoms. For example, a silicon atom forms four covalent bonds to neighboring atoms arranged symmetrically about it. This "tetrahedral" configuration is perfectly maintained in the "diamond" lattice of crystal silicon.

There are also many *noncrystalline* semiconductors. In these materials the chemical bonding of atoms is nearly unchanged from crystals. Nonetheless, a fairly small, disorderly variation in the angles between bonds eliminates the regular lattice structure. Such noncrystalline semiconductors can have fairly good electronic properties – sufficient for many applications. Xerography is the most notable early example [1, 2]. It exploited the *photoconductivity* of noncrystalline selenium. Selenium's unusual properties when illuminated had been known since around 1870, and had already been used to make low-efficiency solar cells that were supplanted by crystalline silicon in the 1950s [3]. As do all semiconductors, selenium absorbs those photons from an incident light beam that have photon energies exceeding a threshold, which is the "bandgap energy". In solar cells, the photon that is absorbed generates one mobile electron and one positively charged "hole" that can be swept away by an electric field.

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

In Dundee, Scotland, Walter Spear and Peter LeComber discovered around 1973 that amorphous silicon prepared using a "glow discharge" in silane (SiH_4) gas had unusually good electronic properties; they were building on earlier work by Chittick, Sterling, and Alexander [4]. Glow discharges are the basis for the familiar "neon" light; under certain conditions, an electric voltage applied across a gas can induce a significant electrical current through the gas, and the molecules of the gas often emit light when excited by the current. Amorphous silicon was deposited as a thin film on substrates inserted into the silane gas discharge; the term "amorphous" is commonly applied to noncrystalline materials prepared by deposition from gases. Spear and LeComber reported in 1975 [5] that amorphous silicon's conductivity could be increased enormously, either by mixing some phosphine (PH_3) gas or some diborane (B_2H_6) gas with the silane. Just as it did for crystal silicon, the phosphorus-doping of the amorphous silicon had induced a conductivity associated with mobile electrons (the material was "*n*-type"), and the boron-doping had induced a conductivity associated with mobile holes (the material was "*p*-type").

In 1974, at the Radio Corporation of America (RCA) Research Laboratory in Princeton, David Carlson discovered that he could make fairly efficient solar cells using a silane glow discharge to deposit films. In 1976 he and Christopher Wronski reported a solar cell based on amorphous silicon [6] with a solar conversion efficiency of about 2.4% (for historical discussion see ref. [3, 7]).

Carlson and Wronski's report of the current density versus output voltage is presented in Figure 12.1, along with curves from far more efficient, contemporary cells [8, 9] based on multijunction stacks of cells. As these scientists had discovered, the optoelectronic properties of amorphous silicon made by glow discharge (or "plasma deposition") are very much superior to the amorphous silicon thin films prepared, for example, by simply evaporating silicon. After several years of uncertainty, it emerged that plasma-deposited amorphous silicon contained a significant percentage of hydrogen atoms bonded into the amorphous silicon structure, and that these hydrogen atoms were essential to the good electronic properties [10]. As a consequence, the improved form of amorphous silicon has generally been known as "hydrogenated amorphous silicon" (or, more briefly, a-Si:H). In recent years, many authors have used the term "amorphous silicon" to refer to the hydrogenated form, which acknowledges that the unhydrogenated forms of amorphous silicon are only infrequently studied today.

Why was there so much excitement about the amorphous silicon solar cells fabricated by Carlson and Wronski? First, the technology involved is relatively simple and inexpensive compared

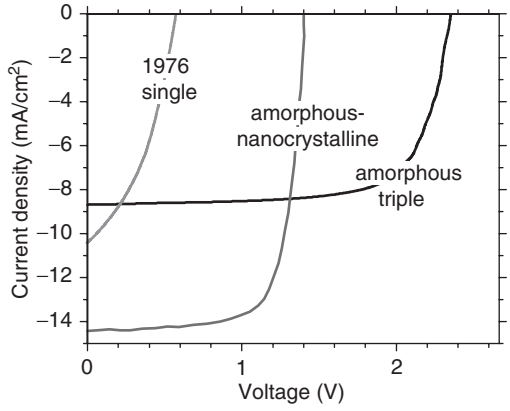


Figure 12.1 Current density vs. voltage under solar illumination for a very early single-junction amorphous silicon solar cell [6], from a tandem amorphous–nanocrystalline cell [9], and from an all-amorphous triple-junction [8]

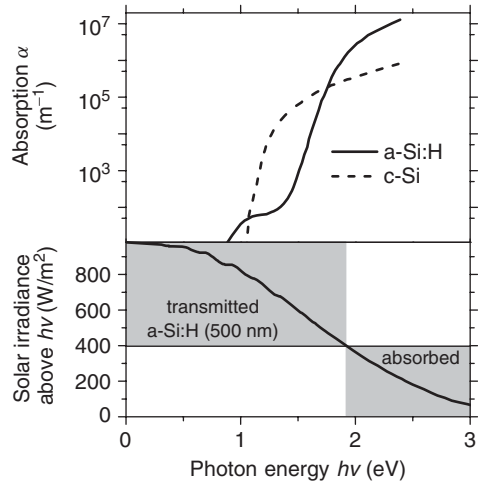


Figure 12.2 (upper panel) Spectra of the optical absorption coefficient $\alpha(h\nu)$ as a function of photon energy $h\nu$ for crystalline silicon (c-Si) and for hydrogenated amorphous silicon (a-Si:H). After reference [11]. (lower panel) Solar irradiance due to photons with energies greater than $h\nu$. The gray areas indicate the irradiance that is transmitted through, or absorbed by, a 500 nm a-Si:H film

with the technologies for growing crystals. Additionally, the optical properties of amorphous silicon are very promising for collecting solar energy, as we now explain. In Figure 12.2, the upper panel shows the spectrum for the optical absorption coefficients $\alpha(h\nu)$ for amorphous silicon and for crystalline silicon [11].[1] In the lower panel of the figure, we show the spectrum of the "integrated solar irradiance"; this is the intensity (in W/m^2) of the solar energy carried by photons that have an energy larger than $h\nu$ [12].

We use these spectra to find out how much solar energy is absorbed by layers of varying thickness. The example used in the figure is an a-Si:H layer with a thickness $d = 500$ nm. Such a layer absorbs essentially all photons with energies greater than 1.9 eV (the energy at which $\alpha = 1/d$). We then look up how much solar irradiance lies above 1.9 eV. Assuming that reflection of sunlight has been minimized, we find that about 420 W/m^2 is absorbed by the layer (the gray area labeled "absorbed"). 580 W/m^2 of energy is transmitted through such a layer. These energies may be compared with the results for c-Si, for which a 500-nm-thick layer absorbs less than 200 W/m^2.

To absorb the same energy as the 500 nm a-Si:H layer, a slab of crystal silicon must be much thicker.[2] It isn't necessarily a competition; a very interesting cell can be made by combining a-Si:H with a thin film of nanocrystalline silicon, with the results shown as the middle curve of Figure 12.1.

In the remainder of this section we first describe how amorphous silicon solar cells are realized in practice, and we then briefly summarize some important aspects of their electrical characteristics.

[1] We assume familiarity with the concept of a photon energy $h\nu$ and of an optical absorption coefficient α; see Chapter 3, Section 3.2 in this volume.

[2] The very different optical properties of c-Si and a-Si reflect the completely different nature of their electronic states. In solid-state physics textbooks, one learns about the "selection rules" that greatly reduce optical absorption in c-Si, which is an "indirect bandgap" semiconductor. Such selections rules do not apply in a-Si. Additionally, the "bandgap" of a-Si is considerably larger than for c-Si.

12.1.2 Designs for Amorphous Silicon Solar Cells: A Guided Tour

Figure 12.1 illustrates the tremendous progress over the last 25 years in improving the efficiency of amorphous silicon-based solar cells. In this section we briefly introduce the two basic ideas of contemporary, high-efficiency devices: the *pin* photodiode structure, and the stacking of these structures to form multijunction photodiode structures. A good deal of this chapter is devoted to more detailed reviews of the implementation and importance of these concepts.

12.1.2.1 pin *photodiodes*

The fundamental photodiode inside an amorphous silicon based solar cell has three layers deposited in either the *p*-*i*-*n* or the *n*-*i*-*p* sequence. The three layers are a very thin (typically 10 nm) *p*-type layer, a much thicker (typically hundreds of nanometers) undoped *intrinsic i*-layer, and another thin *n*-type layer. As illustrated in Figure 12.3, in this structure excess electrons are actually donated from the *n*-type layer to the *p*-type layer, leaving the layers positively and negatively charged (respectively), and creating a sizable "built-in" electric field (typically more than 10^4 V/cm).

Sunlight enters the solar cell as a stream of photons that pass through the *p*-type layer, which is a nearly transparent "window" layer. The solar photons are mostly absorbed in the much thicker intrinsic layer; each photon that is absorbed will generate one electron and one hole photocarrier [13, 14]. The photocarriers are swept away by the built-in electric field to the *n*-type and *p*-type layers, respectively – thus generating solar electricity!

The use of a *pin* structure for a-Si:H-based solar cells is something of a departure from solar cell designs for other materials, which are often based on simpler *pn* structures. For doped a-Si:H, it turns out that minority photocarriers (holes in *n*-type a-Si:H, electrons in *p*-type a-Si:H) don't move very far, and so a *pn* structure would only collect photocarriers from photons generated in an extremely thin layer of doped a-Si:H. Indeed, in analyzing the performance of a-Si:H based solar cells, one normally considers any photons absorbed by the doped layers to be "wasted." Thus a critical goal is keeping the doping atoms out of the absorber layer, which enables this layer to be pure enough so that it can be thick enough to capture most of the sunlight.

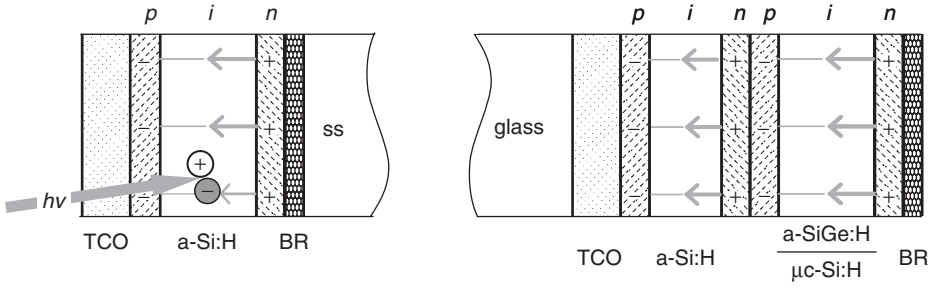


Figure 12.3 (left) Single-junction *pin* solar cell with a stainless-steel (ss) substrate. In a *pin* structure in the dark, excess electrons are donated from the *n*-type doped layer to the *p*-type layer, establishing the space charges and electric fields as shown. Under illumination, each photon absorbed in the undoped, intrinsic layer generates one electron and one hole photocarrier. The electric field causes these carriers to drift in the directions shown. (right) Tandem *pin* solar cell with a glass "superstrate". The second intrinsic layer typically has a lower bandgap, and absorbs photons that are transmitted through the first layer. For amorphous silicon based cells, photons invariably enter through the *p*-type window layer as shown here

In Section 12.4 you will find a more detailed description of the device-physics of the *pin* solar cell; the description explains why the window layer is the *p*-type one, and also explains the design trade-offs which determine the thickness of the absorber layer.

12.1.2.2 Multijunction solar cells

The conversion efficiency of the relatively simple, amorphous silicon *pin* photodiode structure just described can be significantly improved by depositing two or three such photodiodes, one on top of another, to create a "multijunction" device. We illustrate a "tandem" device on the right of Figure 12.3.[3] The main advantage to the tandem design over the simpler single-junction one is due to "spectrum splitting" of the solar illumination [15]. Since the absorption coefficient for light rises rapidly with the photon energy, the front layer of a tandem cell acts as a "low-pass" optical filter. This effect is illustrated in Figure 12.2, which shows that a 0.5 μm layer of a-Si:H absorbs photons with energies larger than 1.9 eV, and passes photons with smaller energies. The "wasted" low-energy photons can be efficiently harvested by using a material with a lower bandgap than amorphous silicon as the back layer; amorphous silicon–germanium alloys, and also nanocrystalline silicon, both have larger absorption coefficient below 1.9 eV than does a-Si:H. Overall, the advantages of the multijunction design are sufficiently compelling that they usually overcome the additional complexity and cost of the deposition facility. As illustrated in Figure 12.1, both tandem and triple-junction devices are in use today. We discuss multijunction solar cells in detail in Section 12.5.

12.1.3 Staebler–Wronski Effect

Amorphous silicon-based solar cells exhibit a significant decline in their efficiency during their first few hundred hours of illumination; nanocrystalline silicon cells do not exhibit a comparable effect. Figure 12.4 illustrates this effect for a single-junction amorphous cell and for a triple-junction module made at United Solar Ovonic [16, 17]. The single-junction cell loses about 30% of its initial efficiency after about 1000 hours; the triple-junction module loses about 15% of its initial efficiency.

All amorphous silicon-based solar cells exhibit this type of initial behavior under illumination; the behavior is mostly due to the Staebler–Wronski effect [18], which is the light-induced change in hydrogenated amorphous silicon (a-Si:H) and related materials used in the cell. Although we have not illustrated it here, the Staebler–Wronski effect anneals nearly completely within a few minutes at temperatures of about 160 °C, and anneals substantially at summer operating temperatures of 60 °C.

The Staebler–Wronski effect contributes to noticeable seasonal variations in the conversion efficiency of a-Si:H based modules in the field. In Figure 12.5 we illustrate the daily average conversion efficiency and ambient temperature for a triple-junction module installation in Switzerland. The module performed best in hot weather. The relative increase in efficiency with temperature is about $+5 \times 10^{-3}$/K. It is noteworthy that there was no discernible, permanent degradation of this module over the three-year extent of this test; one study over ten years does suggest degradation of about 0.7% per year [19], which is comparable to the long-term degradation of c-Si cells.

[3] It is worth noting that the adjoining *p*-type and *n*-type layers do *not* form a *p-n* junction diode, but rather a simple ohmic contact. We discuss the interesting physics underlying this fact in Section 12.5.3.

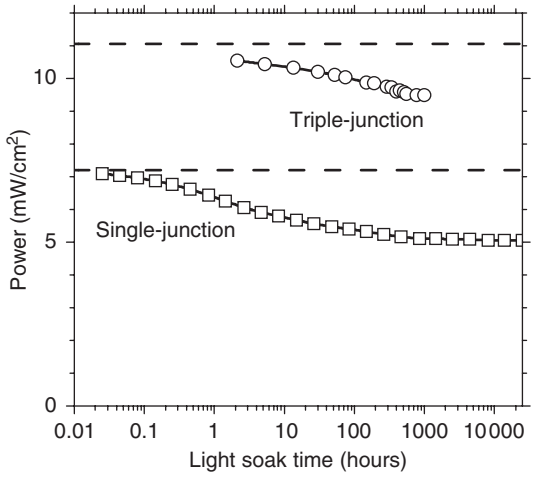


Figure 12.4 The conversion efficiency in a-Si:H-based solar cells declines noticeably upon the first exposure to sunlight. The figure illustrates this decline under a solar simulator (100 mW/cm^2) for a single-junction cell (260 nm *i*-layer thickness) and for a triple junction module made at United Solar Ovonic [16, 17]; the dashed lines indicate the initial power measured for each device

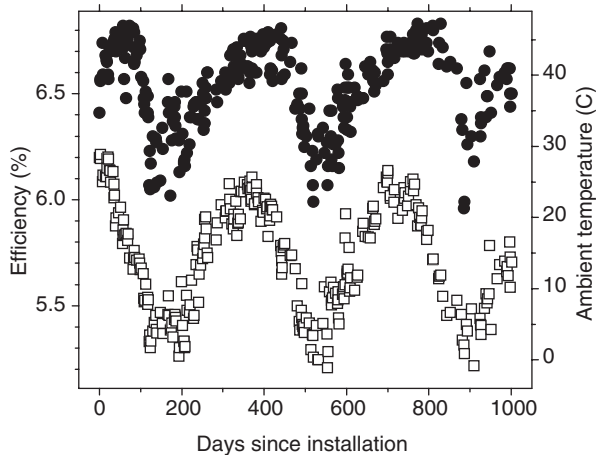


Figure 12.5 Seasonal variations in the average conversion efficiency (solid symbols) of an amorphous silicon triple-junction module [20], along with the daily mean temperature (open symbols)

This positive trend of efficiency with temperature is atypical of solar cells made with other materials; for example, the temperature coefficient for crystal silicon solar cells is about -4×10^{-3}/K [21, 22]. Interestingly, if the temperature-dependence of a-Si:H solar cells is measured quickly – so that there is no time for the Staebler–Wronski effect to set in, the temperature coefficient is also negative (about -1×10^{-3}/K) [21]. The behavior of a module in the field may be understood as a competition of slow annealing of the Staebler–Wronski effect (which yields the positive temperature coefficient) and of a smaller, intrinsic negative coefficient [23, 24].

12.1.4 Synopsis

The remainder of this chapter is organized as follows. In Section 12.2 we introduce some of the fundamental physical concepts required to interpret the scientific literature about amorphous silicon and related materials including amorphous silicon-germanium alloys and nanocrystalline silicon. Section 12.3 surveys the principal methods such as plasma deposition that are used to make amorphous silicon based solar cells. Section 12.4 describes how the simplest, single-junction solar cell "works," by which we mean how the photoelectric behavior of the cell is related to the fundamental concepts. High-efficiency solar cells based on amorphous silicon technology are multijunction devices, and in Section 12.5 we discuss how these are made and how their performance can be understood and optimized. Section 12.6 describes some of the issues involved in manufacturing modules. To conclude this chapter, Section 12.7 presents some of the directions which we consider to be important for future progress in the field.

There have been several excellent monographs and review chapters on amorphous silicon and amorphous silicon-based solar cells in recent years. In the body of the chapter, we direct the reader to these works where we feel that they may be useful for expanded or complementary discussion.

12.2 ATOMIC AND ELECTRONIC STRUCTURE OF HYDROGENATED AMORPHOUS SILICON

12.2.1 Atomic Structure

Silicon atoms in amorphous silicon mostly have the same basic structure that they have in crystal silicon: each silicon atom is connected by covalent bonds to four other silicon atoms arranged as a tetrahedron around it. This understanding emerges from measurements of the scattering ("diffraction") of X-rays by the two materials [25] as well as from theoretical and computational studies of the two materials.

If you build a noncrystalline silicon structure with wooden sticks (to represent covalent bonds) and wooden balls drilled with four small holes for the sticks (to represent the silicon atoms), you will have some trouble in making a noncrystalline structure. To avoid a crystalline structure, you will need to bend the sticks. Quite soon, you will have to give up on the fourth stick on some atom, and you will have created an imperfect noncrystalline structure with a "dangling bond." Your problem is related to tetrahedral bonding: there are too many constraints on the positions of atoms to keep *all* bond lengths and angles reasonably close to the values demanded by silicon's chemistry in any noncrystalline structure. The same conclusion is reached by mathematical and computational methods [26, 27]. Alloys such as As_2Se_3, which easily form noncrystalline glasses by cooling from a liquid, have an average number of bonds per atom of about 2.7 or less.

For hydrogenated amorphous silicon (a-Si:H), silicon–hydrogen bonds partially resolve this structural problem. Several percent of the silicon atoms make covalent bonds with only three silicon neighbors; the fourth valence electron of the silicon bonds to a hydrogen atom. This crucial hydrogen is essentially invisible to X-rays, but is quite evident in proton magnetic resonance [28] and infrared spectroscopy [29], secondary ion mass spectroscopy [30], and hydrogen evolution during annealing [31].

There are quite a few distinct atomic configurations for the hydrogen in a-Si:H. The two principal "phases" of hydrogen evidenced by proton magnetic resonance are termed the "dilute"

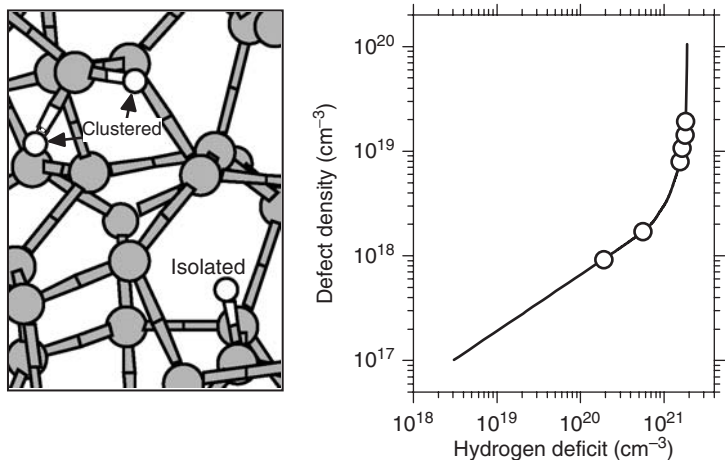


Figure 12.6 (*Left*) computer model of the chemical bonding for hydrogenated amorphous silicon. The larger, gray spheres indicate Si atoms; the smaller, white spheres indicate hydrogen atoms, which are found in clustered and isolated configurations as indicated; (*right*) correlation of the defect (dangling bond) density in a-Si:H with the density of hydrogen removed from the material by heating (the hydrogen deficit). The data points are derived from deuterium and defect profiles by Jackson, *et al*. [35] (350 °C deuteration). The curve is a fit to a model proposed by Zafar and Schiff [34]

or isolated and "clustered" phases [28]. In the dilute phase a particular hydrogen atom is about 1 nm away from any other hydrogen atom; in the clustered phase there are two or more hydrogen atoms in close proximity. A computer calculation of a particular instance of this structure [32] is presented in Figure 12.6. The densities of hydrogen in each of the individual phases, as well as the total density of hydrogen, depend upon the conditions under which the material is made.

12.2.2 Defects and Metastability

While the underlying structure illustrated in Figure 12.6 is noncrystalline, it is a chemically ideal structure: each atom forms the normal number of chemical bonds (four for silicon, one for hydrogen). This noncrystalline atomic structure largely determines the overall electronic and optical properties of the material, as we discuss shortly. However, many electronic properties in a-Si:H are also strongly affected by gross defects of chemical bonding. The atomic structure of the bonding defects in a-Si:H has been extensively studied using electron spin resonance. A single type of defect, the D-center, dominates most measurements in undoped a-Si:H [25]. The D-center is generally identified as a silicon dangling bond [33].

A dangling bond may be envisioned using Figure 12.6: just imagine that the hydrogen atom is removed from the dilute-phase site in the lower right-hand corner of the figure, leaving behind a single unbonded electron (the "dangling bond"). This simple picture is consistent with the following observation: the density of dangling bonds increases when hydrogen is removed from a-Si:H by heating. We present a comparison of a model for this relationship together with measurements illustrating the effect in Figure 12.6 [34, 35]. Note that the density of dangling bonds is generally much lower than the density of hydrogen lost from the structure; this effect has been attributed to the evolution of hydrogen from clustered-phase sites, which presumably does not create dangling bonds.

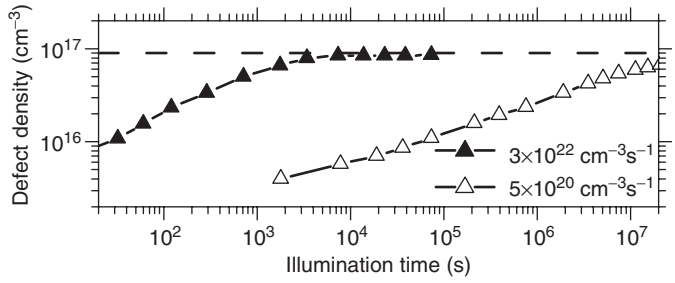


Figure 12.7 Plot of the defect (dangling bond) density during extended illumination of an a-Si:H film as measured by Park, Liu, and Wagner [38]. Data are given for high- and low-intensity illumination; the legend indicates the photocarrier generation rates at the two intensities

The most intense defect research in a-Si:H has not been focused on the direct hydrogen-defect relation, but rather on the light-soaking effects. We illustrated how light-soaking degrades the solar conversion efficiency in Figure 12.4, and Figure 12.7 below we illustrate how it increases the defect density. For the high-intensity illumination, the defect density reaches a steady-state at about $10^{17}\,\mathrm{cm}^{-3}$. It is very important in specifying solar modules that a-Si:H reach such a "stabilized" condition after extended light-soaking.

Although the defect density is not the only property of a-Si:H modified following light-soaking [36], most workers believe that the principal cause of the Staebler–Wronski effect is this increase in dangling bond density after light-soaking. The close connection between hydrogen and defects in a-Si:H has led to several efforts to understand the defect creation in terms of metastable configurations of hydrogen atoms [34, 37]. The idea is that illumination provides the energy required to shift hydrogen atoms away from their dilute-phase sites, thus creating dangling bonds. The technological importance of establishing the atomic mechanism underlying the Staebler–Wronski effect lies in the possibility that this effect can be mitigated in a-Si:H by changing its preparation conditions.

An essential feature of the light-soaking effects on a-Si:H cells and films is that most of the effects are "metastable," and can be removed nearly completely by annealing of a light-soaked sample at a temperature above 150 °C. More generally, the stabilized condition of a-Si:H cells and films is quite temperature-dependent. For example, Figure 12.5 showed that the module efficiency is substantially affected by the seasons, and is highest following the hottest days. The measurement may be understood by considering that the stabilized condition is due to competition between two rates: the creation of metastable defects by light, and a thermally activated process which anneals them away.

12.2.3 Electronic Density-of-States

The most important concept used in understanding the optical and electronic properties of semiconductors is the electronic density-of-states, $g(E)$. The idea is a simple approximation: if a single electron is added to a solid, it may be viewed as occupying a well-defined state (or molecular "orbital") at a particular level energy E. In a range of energies ΔE, the number of such states per unit volume of the solid is $g(E)\Delta E$.

In Figure 12.8 we have illustrated the density-of-states for hydrogenated amorphous silicon as it has emerged primarily from measurements of electron photoemission [39, 40], optical absorption [41], and electron and hole drift mobilities [42]. In the dark at low temperatures, the

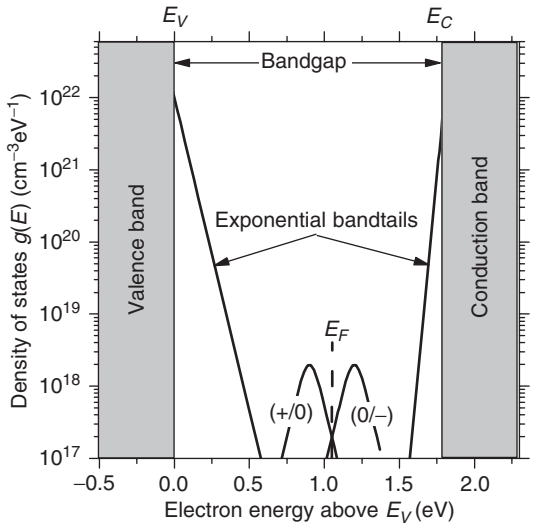


Figure 12.8 Density of electronic states $g(E)$ in hydrogenated amorphous silicon. The shaded areas indicate delocalized states in the bands; these bands themselves have tails of localized states with an exponential distribution. Midway between the bands are levels belonging to gross defects such as dangling Si bonds

states with energies below the Fermi energy E_F are filled by electrons; above the Fermi energy the states are empty. There are two strong bands of states illustrated: an occupied valence band ($E < E_V$), originating with the Si–Si and Si–H bonding orbitals, and an unoccupied conduction band ($E > E_C$), originating with "antibonding" orbitals.

12.2.4 Band Tails, Band Edges, and Bandgaps

Between the conduction and valence bands lies an "energy gap" where the density-of-states is very low. Any functional semiconductor, crystalline or noncrystalline, must have such an energy gap. For perfect crystals, the valence and conduction band edge energies E_V and E_C are well defined, as is the *bandgap* $E_G = E_C - E_V$. In disordered semiconductors there are exponential distributions of band tail states near these band edges. For the valence band tail we write $g(E) = g_V \exp(-(E - E_V)/\Delta E_V)$. The width ΔE_V of this exponential distribution is important in interpreting optical absorption experiments, where it is usually identified with the exponential "Urbach" tail of the spectrum apparent in Figure 12.2. For a-Si:H, a typical value $\Delta E_V = 50 \times 10^{-3}$ eV. ΔE_V is also used to account for the very slow drift of holes in an electric field (i.e. the hole *drift mobility*); [42, 43]. The conduction band tail width ΔE_C is much narrower; for the best a-Si:H materials it is about 22×10^{-3} eV, but increases markedly for amorphous silicon–germanium [44].

Given the presence of exponential band tails, the very existence of a band edge energy can reasonably be questioned. Remarkably, detailed analysis of drift-mobility measurements supports the concept of a well-defined band edge [42, 45]. Most workers consider the band edge to be the energy that separates electron orbitals that are localized (i.e. have well defined locations in space) from orbitals that are delocalized. The band edges are correspondingly termed the conduction and valence band *mobility edges* [46].

Unfortunately, for noncrystalline semiconductors there is no single, conclusively established procedure for locating the band edges within the density-of-states. The bandgap is thus difficult to determine without some ambiguity. Since amorphous silicon based materials with varying bandgaps are used in solar cells, it is nonetheless very important to establish conventional procedures for comparing bandgaps. There are several optical methods to determine an effective bandgap in a-Si [47, 48]. The most common approach is to analyze measurements of the optical absorption coefficient $\alpha(h\nu)$ similar to those in Figure 12.2; one typical analysis yields an "optical" or "Tauc" bandgap E_T [49]

$$\alpha(h\nu) = (A/h\nu)(h\nu - E_T)^2. \tag{12.1}$$

The Tauc bandgap E_T is found from the intercept of a graph of $(h\nu\alpha)^{1/2}$ versus $(h\nu)$. The proportionality constant A incorporates several effects, and is not usually studied separately.

The bandgap obtained using this procedure is typically about 1.75 eV in a-Si:H, but does vary with deposition conditions. Alloying with germanium or carbon causes sizable changes. A simpler procedure than that of Tauc is to define the bandgap to be the photon energy corresponding to a particular optical absorption coefficient α; using $\alpha = 3 \times 10^3\,\mathrm{cm}^{-1}$ yields values (denoted as $E_{3.5}$) similar to the Tauc procedure. Finally, there is undoubtedly a difference between these optical estimates of the bandgap and the true, "electrical" bandgap $E_G = E_C - E_V$ (see Figure 12.8). Internal photoemission measurements [50] have suggested that the electrical bandgap is 50–100 meV larger than the Tauc bandgap.

12.2.5 Defects and Gap States

Between the band tails lie defect levels; in undoped a-Si:H, these levels appear to be due primarily to the dangling bonds (D-centers) measured by electron spin resonance. For example, infrared absorption at photon energies around 1.2 eV is sensitive to the optical processes which detach an electron from a defect and promote it to the conduction band, or which transfer an electron from the valence band to a defect. This infrared signal is visible in Figure 12.2; for samples of varying electronic properties, the infrared absorption coefficient is proportional to the D-center density over a range of at least 100 [51].

The next issue to be resolved is the positions of the corresponding levels, as illustrated in Figure 12.8. The D-center is "amphoteric:" there are three charge states (with $+e$, 0, and $-e$ charges), leading to two levels (transitions between the 0/+ and −/0 charge states). A *rough* guide to level positions estimated under near-dark conditions is the following. The (+/0) level is about 0.6 eV below E_C in low-defect-density, undoped a-Si:H [52]. The (+/0) level lies about 0.3 eV below the (−/0) levels; the difference between the two levels is usually termed the "correlation energy" of the D-center [52].

The actual level positions apparently vary between doped and intrinsic a-Si:H [25], between intrinsic samples with varying densities of D-centers [53], and possibly between dark and illuminated states [54]. These effects are usually explained using a "defect pool" model [25].

12.2.6 Doping

Doped layers are integral to *pin* solar cells. Doping, which is the intentional incorporation of atoms such as phosphorus and boron in order to shift the Fermi energy of a material, works very differently in amorphous silicon than in crystals. For example, in crystalline silicon (c-Si), phosphorus (P) atoms substitute for silicon atoms in the crystal lattice. P has five valence electrons, so in the "fourfold coordinated" sites of the Si lattice, four electrons participate in bonding to

neighboring silicon atoms. The fifth excess electron occupies a state just below the bottom of the conduction band, and the dopants raise the Fermi energy to roughly this level.

In a-Si, most phosphorus atoms bond to only three silicon neighbors; they are in "three-fold coordinated" sites. This configuration is actually advantageous chemically; phosphorus atoms normally form only three bonds (involving the three valence electrons in p atomic orbitals). The final two electrons are paired in s atomic orbitals, do not participate in bonding, and remain tightly attached to the P atom. The reason that this more favorable bonding occurs in a-Si, but not in c-Si, is the absence of a rigid lattice. As a thin film of a-Si grows, the network of bonds adjusts to incorporate impurity atoms in nearly ideal chemical arrangement. In c-Si, it would be necessary to grossly rearrange several Si atoms in the lattice, and to leave a number of dangling Si bonds, in order to accommodate the P atom in this configuration. The extra energy for this rearrangement is larger than would be gained from more ideal bonding of P, and so substitutional doping is favored.

Thus phosphorus doping is a paradox in amorphous silicon. It is, at first, unclear why it occurs at all, since doping involves fourfold coordinated P, and P atoms are generally threefold coordinated in a-Si. This puzzle was first solved in 1982 by Street, who realized that independent formation of both a positively charged, fourfold coordinated P_4^+ and a negatively charged dangling bond D^- can occur occasionally instead of the more ideal threefold coordination [25]. This understanding leads to two important consequences. First, doping is inefficient in a-Si; most dopant atoms do not contribute a "free" electron, and do not raise the Fermi energy. Second, for each dopant atom which does contribute an electron, there is a balancing, Si dangling bond to receive it. These defect levels lie well below the conduction band, so the fourfold-coordinated phosphorus atoms are less effective in raising the Fermi energy than in c-Si. Additionally, the negatively charged dangling bonds induced by doping are very effective traps for holes. Since transport of both electrons and holes is essential to photovoltaic energy conversion, photons absorbed in doped layers don't contribute to the power generated by solar cells.

12.2.7 Alloying and Optical Properties

The structural and optical properties we have described can be varied substantially by changes in deposition conditions. For example, changing the substrate temperature or the dilution of silane by hydrogen (in plasma deposition) causes a change in the the optical bandgap for a-Si:H films over at least the range 1.6–1.8 eV [55]; these changes can be ascribed to changes in the hydrogen microstructure of the films. Even larger changes can be effected by alloying with additional elements such as Ge, C, O, and N; alloying is readily accomplished by mixing the silane (SiH_4) source gas with gases such as GeH_4, CH_4, O_2 or NO_2, and NH_3, respectively. The resulting alloys have very wide ranges of bandgaps, as we illustrate for a-$Si_{1-x}Ge_x$:H below. For simplicity, we shall usually refer to these alloys using the abbreviated notation: a-SiGe for a-$Si_{1-x}Ge_x$:H, etc.

Only some of these materials have proven useful in devices. In particular, a-SiGe alloys with optical gaps down to about 1.45 eV are employed as absorber layer in multijunction *pin* cells; the narrower bandgap of a-SiGe compared to a-Si allows for increased absorption of photons with lower energies [56]. Figure 12.9 (left panel) illustrates how the spectrum of the absorption coefficient $\alpha(h\nu)$ changes for a-SiGe alloys with different atomic percentages x; the different optical bandgaps are indicated as labels. Two features of these data should be noted. First, the Urbach slopes remain constant (at about 50 meV) over the entire range of bandgaps. Second, the plateau in the absorption coefficient at the lowest photon energies are indicative of the defect density, which indicates that this density increases steadily as the bandgap diminishes. The role of Ge-induced alloying on defects and optoelectronic properties was comprehensively studied at Harvard University in the 1980s [57]. It was determined that structural inhomogeneity resulting in two-phase material was responsible for the degradation in alloy properties with increasing Ge.

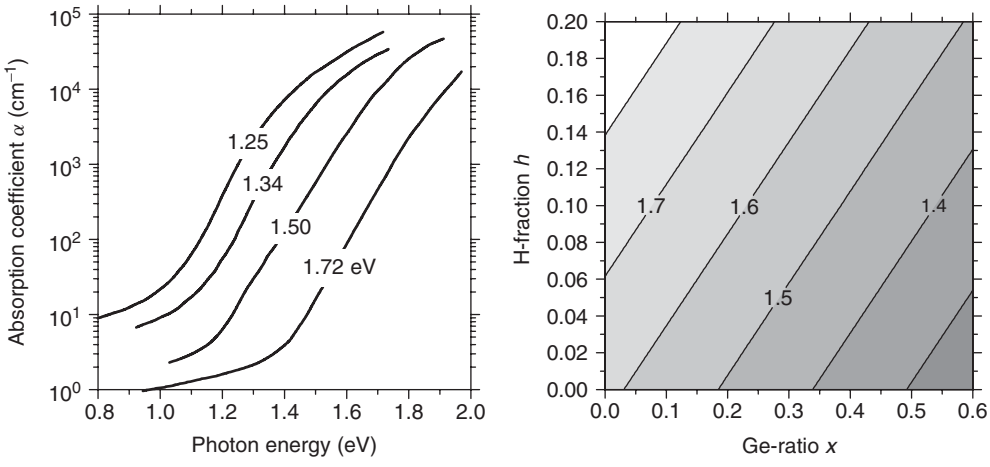


Figure 12.9 (*Left*) absorption coefficient spectra for a-SiGe alloys; the optical bandgaps and corresponding Ge fractions x are: 1.25–0.58, 1.34–0.48, 1.50–0.30, 1.72–0.0 [55]; (*right*) typical optical bandgaps for a-$Si_{1-x}Ge_x$:H alloys for varying Ge-ratio x and atomic fraction h of hydrogen

Figure 12.9 (right panel) is a contour plot showing how the optical bandgap of a-$Si_{1-x}Ge_x$:H varies with the Ge-ratio x and with atomic fraction h of hydrogen. The figure reflects experimental results for a-Si:H alloys of varying H-fraction [55] and for a-SiGe:H alloys for which both x and h were reported [58].[4] Note that, for constant fraction h, the bandgap decreases about 0.7 eV as the Ge-ratio x increases from 0 to 1. The bandgap increases with atomic fraction of hydrogen h. Figure 12.9 should be viewed as a useful approximation; in particular, the atomic fraction h is only one aspect of the hydrogen microstructures in a-SiGe alloys, and quantitative deviations from the contour plot are likely. Additionally, only some of the materials represented in the figure are useful as absorber layers. In particular, as the Ge ratio x rises to about 0.5, the optoelectronic properties become so poor that these alloys are no longer useful in solar cells [59]. Similarly, only limited ranges of the atomic fraction of hydrogen h yield useful absorber layers.

It might be thought that a-SiC would be equally useful as a wider-bandgap absorber; despite some promising research [60], this material is not being used as an absorber layer by manufacturers. B-doped a-SiC is used extensively as a p-type, window layer [61]. a-SiO and a-SiN are used as insulators in thin film transistors [62], but are not major components in solar cells.

12.2.8 Briefing: Nanocrystalline Silicon

The same deposition processes that are used to make amorphous silicon can also be used to make hydrogenated "nanocrystalline" silicon (nc-Si:H). For the solar cell applications described in this chapter, nanocrystalline silicon may be viewed as something of a silicon "clafouti" (a provincial French dessert). Fine silicon crystallites with diameters of several nanometers are bound with a hydrogenated amorphous silicon batter. It should also be noted that the material presently known as nanocrystalline silicon was called "microcrystalline" silicon for many years.

[4] Figure 12.9 is based on the function $E_G = 1.62 + 1.3h - 0.7x$ that we obtained by fitting to experimental results reported by Hama, *et al*. [55] and Middya, *et al*. [58].

For photon energies below about 1.6 eV (i.e. wavelengths larger than 750 nm), nanocrystalline silicon has optical absorption that is much stronger than a-Si:H, and is similar to crystal silicon (Figure 12.2). This optical absorption makes nc-Si:H an interesting alternative to a-SiGe:H alloys as the infrared absorber in multijunction solar cells; tandem cells of this type are sometimes called "micromorph" solar cells, which is a word constructed by combining "microcrystalline" and "amorphous".

The experimental tool which has proven most useful in evaluating the relative volumes of crystalline and amorphous phases in nc-Si:H is Raman scattering, which indicates that the mixed-phase nanocrystalline materials incorporated in micromorph solar cells generally have amorphous phase fractions that are 10% or larger. A complete introduction to the properties of nc-Si:H is beyond the scope of this chapter, and we refer the interested reader to a recent review [63]. We do note two of the distinctions between nc-Si:H and a-Si:H. While both materials can be described by a density-of-states picture such as Figure 12.8, the parameters differ significantly; we return to this subject in Section 12.4. While dangling bonds are important in nc-Si:H as well as a-Si:H, the Staebler–Wronski degradation is much weaker in nc-Si:H than a-Si:H, despite the presence of a fraction of a-Si:H.

12.3 DEPOSITING AMORPHOUS SILICON

12.3.1 Survey of Deposition Techniques

The first preparations of a-Si:H by Chittick *et al*. [64] and by Spear and LeComber [65] used a silane-based glow discharge induced by radio frequency (RF) voltages; the method is now often termed plasma enhanced chemical vapor deposition (PECVD). Since this pioneering work, many deposition methods have been explored with the intention of improving material quality and deposition rate. Among these methods, RF-PECVD using 13.56-MHz excitation is still the most widely used today in research and manufacturing of a-Si-based materials. However, emerging film deposition methods, mostly directed toward higher deposition rate or toward making improved nanocrystalline silicon films, have been extensively explored in recent years. Table 12.1 summarizes the most extensively studied deposition processes used as well as some of their advantages and disadvantages. Among these, PECVD with very high frequency (VHF) and hot-wire (HW) catalytic deposition process will be further discussed in this section because of their potential for use in future high-throughput solar cell manufacturing.

12.3.2 RF Plasma-Enhanced Chemical Vapor Deposition (RF-PECVD) at 13.56 MHz

The growth of a-Si and nc-Si by PECVD is a complex process determined by the following factors: plasma properties such as electron density and energy distribution; gas phase reaction chemistry; precursor transport to the growth surface; and surface reactions. Figure 12.10 shows a schematic of a typical RF PECVD chamber and related parts. A silicon-containing gas such as a mixture of SiH_4 and H_2 flows into a vacuum chamber that is evacuated by a pump. Power from an external RF supply is applied between two electrode plates installed inside; sometimes one is grounded. For a given RF voltage across the plates, there is usually a range of gas pressures for which plasmas occur. The plasma excites and decomposes the gas and generates radicals and ions in the chamber. Substrates may be mounted on one or both of the electrodes, and thin hydrogenated silicon films grow on the substrates as these radicals diffuse into them. The substrates are heated (150–300 °C)

Table 12.1 Various deposition processes used for depositing amorphous silicon-based materials

Processes	Maximum rate[a] [Å/s]	Advantages	Disadvantages	Manufacturers	Reference
RF PECVD	3	High quality, uniform	Slow	Many	[77–79]
DC PECVD	3	High quality, uniform	Slow	BP Solar	[66, 67]
VHF PECVD	20	High quality, fast especially for nc-Si	Poor uniformity	Oerlikon Solar, Sharp	[68, 94]
Microwave PECVD	100	Very fast	Film quality not as good	Canon	[70]
Hot-wire	50	Very fast	Poor uniformity	None	[71, 72]
Photo-CVD	1	High quality	Slow	None	[73, 74]
Sputtering	3		Poor quality, slow	None	[75, 76]

[a]Maximum deposition rate: the deposition rate beyond which the film quality deteriorates rapidly; these numbers are empirical, not fundamental limits, and represent current results at the time of publication.

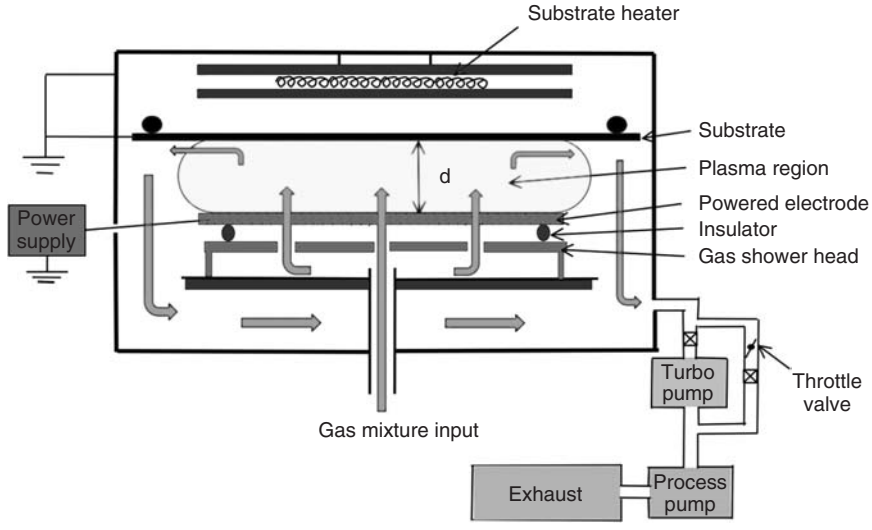


Figure 12.10 Schematic of a typical RF PECVD deposition chamber. The electrode-to-substrate spacing is 'd'

to achieve optimum film quality; this effect is attributed to thermally activated surface diffusion of adatoms on the growing film. Note that a-Si and nc-Si films are deposited at much lower temperature than other polycrystalline thin films for solar cells such as CdTe or CIGS.

A PECVD system usually consists of several major parts: (1) a gas delivery system (gas cylinders, pressure regulators, mass flow controllers, and various gas valves to direct gas flows); (2) a deposition chamber (electrodes, substrate mounts, substrate heaters, and the RF power feed through) capable of high vacuum; (3) a pumping system usually consisting of 3 pumps: a turbomolecular pump and mechanical backing pump for establishing high-purity vacuum, and a separate process

pump for removing the gases and other by-products only used during deposition; (4) a pressure control system (capacitance manometer, ionization gauges, thermocouple gauges, and/or throttle valve to monitor and control the chamber pressure); (5) high-frequency power supply (RF or VHF supply with impedance matching network); and (6) an exhaust system for the process gases (typically either a chemical scrubber to neutralize the gases or a "burn box" to pyrolyze them). In multichamber systems there is a transfer system to move substrates inside the vacuum system between various deposition chambers through appropriate gate valves. Large-volume manufacturing systems typically do not have turbomolecular pumps due to their cost and maintenance.

The film growth in a PECVD process consists of several steps: source gas diffusion, electron impact dissociation, gas-phase chemical reaction, radical diffusion, and deposition [77–80]. To deposit "device"-quality a-Si films, the deposition conditions need to be controlled within certain ranges desirable for high-quality a-Si growth [81]. Typical ranges of parameters for a-Si are summarized in Table 12.2. Note that device quality nc-Si films require different conditions, as will be discussed in Section 12.5.4.

The pressure range is usually between 0.05 and 2 Torr. Lower pressure is desirable for making uniform deposition, and higher pressure is more desirable for obtaining higher growth rates and for preparing nc-Si films. Most researchers use a pressure between 0.5 and 1 Torr for a-Si deposition. The RF power should be set at around 10–100 mW/cm^2 for a capacitively coupled reactor. Below 10 mW/cm^2, it is difficult to maintain a plasma. Higher power is desirable for higher deposition rate. However, above 100 mW/cm^2, the rapid reactions in the gas can create a silicon polyhydride powder that contaminates the growing Si film. This problem can be mitigated by using very low pressure or strong hydrogen dilution, both of which result in reduced growth rate.

The substrate temperature is usually set between 150 and 350 °C. At lower substrate temperature, more H is incorporated in the film. As expected from Figure 12.9, this increases the band gap of a-Si:H slightly [81]. However, lower substrate temperature (<150 °C) exacerbates silicon polyhydride powder formation unless high hydrogen dilution is used. At higher substrate temperature (>300 °C), less hydrogen is incorporated and the bandgap is somewhat reduced. These effects are attributed to the thermal enhancement of the surface diffusivity of adatoms during growth; presumably at higher temperatures the silicon network is more ideal and binds less hydrogen. Researchers exploit the substrate temperature effect on the bandgap in device making. Wider-bandgap (lower substrate temperature) materials are useful in the top component cell of a triple-junction solar cell [83]. Narrower-bandgap (higher substrate temperature) a-Si is useful as the top cell *i*-layer of an a-Si/a-SiGe tandem cell. Proper hydrogen dilution is critical to maintain high-quality a-Si at high or low temperatures. However, at temperatures higher than 350 °C the quality of the material degrades due to loss of hydrogen passivation of dangling bonds.

Table 12.2 Ranges of RF-PECVD (13.56 MHz) deposition conditions for a-Si:H films with optimal properties. These numbers are empirical, not fundamental limits, and represent current results at the time of publication

Range	Pressure [Torr]	RF power density [mW/cm^2]	Substrate temperature [C]	Electrode spacing [cm]	Active gas flow[a] [sccm/cm^2]	H_2 dilution R[b]
Upper	2	100	350	5	0.02	100
Medium	0.5	20	250	3	0.01	10
Lower	0.05	10	150	1	0.002	0

[a]Flows of active gases, such as SiH_4, GeH_4, or Si_2H_6, for each unit area of the deposition area (electrode + substrate + chamber walls)

[b]Hydrogen dilution R, defined here as the ratio of hydrogen and active gas flows (e.g. H_2/SiH_4)

Optimum spacing (d) between the RF powered electrode and the substrate in an RF PECVD reactor is usually between 1 and 5 cm for a-Si deposition. Smaller spacing is more desirable for a uniform deposition, while a larger spacing is easier to maintain a plasma. Some of the silicon atoms in the gases directed into the chamber are deposited onto the substrates or the chamber walls; the remainder is pumped to the exhaust. The flow rate of each gas and the pressure determine the residence time of molecules for that species, which then determines the deposition rate for a given RF power. Manufacturers may prefer conditions that lead to higher gas utilization (lower gas flows and higher RF power) to minimize SiH_4 cost. But this compromises the quality of a-Si films deposited near the downstream area when a linear flow scheme is used. For an R&D type deposition system with a 200-cm^2 electrode area and for the deposition of a-Si at the rate of 1 Å/s, a few sccm (cubic centimeters per minute at atmospheric pressure) of SiH_4 flow is typical. As one may easily calculate, for such a chamber with an electrode diameter of 16 cm and an electrode gap of 2.54 cm, 1 sccm of SiH_4 (or 0.005 sccm/cm^2 for this chamber) for a 1-Å/s deposition rate corresponds to a gas utilization of 11%. Commercial scale PECVD systems are rumored to have under 20% utilization; i.e. 80% of the incident Si atoms are not converted into film. Fortunately, silane is only a few percent of the cost of materials used in making a module. For the deposition of high-quality, stable a-Si material, a hydrogen dilution at appropriate level is usually used, as will be discussed in Section 12.3.6.

Another important aspect of the growth of high-quality a-Si films is the reduction of contaminants, such as oxygen, carbon, nitrogen, or metal elements. Fortunately, because of the flexibility of the bonding network in an amorphous solid, the tolerance level for contaminants in a-Si is much higher than that of its crystalline counterpart. Contamination can occur from the atmosphere (CO_2, H_2O, N_2) due to vacuum leaks, impurities in the gases (H_2O, chlorosilanes and siloxanes), pump oil backstreaming (hydrocarbons) or from adsorption on the reactor surfaces (H_2O) [85, 86]. Although there are no standards, it is generally accepted that device-quality intrinsic a-Si films should have O, C, and N impurities below the following concentrations: $O < 10^{19}$, $C < 10^{18}$, and $N < 10^{17}/cm^3$. When the contaminants in the *i*-layer exceed these values, the device performance, particularly the fill factor, will suffer as a result of the reduced lifetime of photogenerated carriers [87].

To understand and monitor the film growth in a PECVD process, various spectroscopic tools, including optical emission spectroscopy [88], optical absorption spectroscopy [89], and residual gas analyzer [90] have been used to characterize the plasma and the concentration of various species inside the reactor. It is believed that the SiH_3 radical is mostly responsible for the growth of high-quality a-Si film [91]. Such spectroscopic tools could be useful in studying and monitoring the active species and contaminants during growth, especially for process control in manufacturing.

RF PECVD systems may be designed with different geometries. In R&D, the substrates and electrode are usually placed horizontally, while in manufacturing the substrates are often installed vertically for higher throughput production and a smaller footprint.

12.3.3 PECVD at Different Frequencies

Standard RF-PECVD uses a frequency f of 13.56 MHz, which is allotted for industrial processes by federal and international authorities. A much larger frequency range has been explored, including DC ($f = 0$), low frequency ($f \sim$ kHz), very high frequency (VHF) ($f \sim 20$–150 MHz), and microwave frequency (MW) ($f = 2.45$ GHz). DC glow discharges [92] were used in the early days of a-Si material and device research at RCA Laboratories [82], and were used in manufacturing at BP Solar, Inc. [93]. AC glow discharge, including RF, VHF, and MW PECVD, are much more widely used than DC because of the relative ease in maintaining the plasma and because of more efficient ionization. VHF and MW deposition have been intensively studied because of the higher deposition rate both for amorphous and nanocrystalline silicon films. The effect of DC, RF, and

VHF discharge on growth rate, bandgap, hydrogen content and stability found that that substrate temperature and hydrogen dilution had roughly similar effects on a-Si films made from all three types of plasmas [94].

12.3.3.1 VHF PECVD

The primary advantage of VHF plasma is that higher excitation frequency ($f = 40–100$ MHz) enables deposition of a-Si and nc-Si films that are more stable at high rates (>10 Å/s) without making polyhydride powder. In contrast, film and device quality and stability suffer when deposition rates are increased >3 Å/s by just increasing RF power at 13.56 MHz. VHF plasma was pioneered by the group at Université de Neuchâtel [95] as a route to higher deposition rates for nc-Si films although its benefit for a-Si:H at high rates and better stability was also recognized. High-quality devices have been obtained using VHF deposition by a number of groups [96–99] and it is being implemented in production at Oerlikon [100], Canon [101], Mitsubishi Heavy Industry [105], and Sharp [106].

With increasing plasma excitation frequency, simulations indicate that the average electron energy decreases and the total electron density increases [102, 103]. This was confirmed experimentally using Langmuir probe measurement and optical emission spectroscopy [104].

Figure 12.11 shows the trends in electron energy, electron density and quarter-wavelength size for a wave as a function of plasma excitation frequency for a specific set of H_2 plasma conditions given in the caption. The higher electron concentration in the plasma increases the beneficial radical density (atomic hydrogen and film growth precursor SiH_x). These radicals provide both selective etching of disordered amorphous phase, reducing defects and voids, and fast growth of crystalline grains. Both are beneficial to nc-Si growth.

Table 12.3 compares four single-junction solar cells with a-Si intrinsic layers fabricated using low and high frequencies and low and high RF power; otherwise the deposition conditions were identical. While for low-power deposition the cell performances are similar, at high deposition rate, the VHF-produced devices are superior in both initial efficiency and stability. The ability to

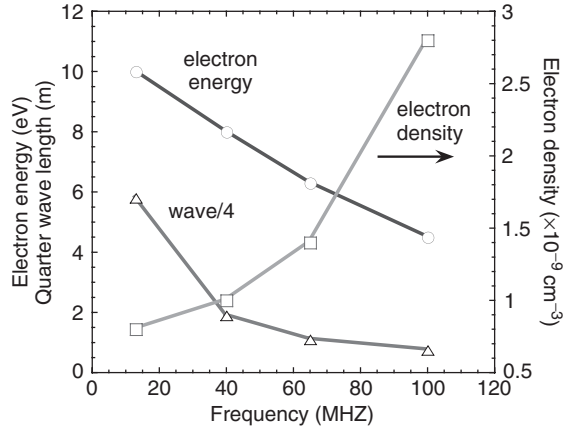


Figure 12.11 Electron energy, electron density, and quarter wavelength as a function of the PECVD frequency (conditions: H_2 flow rate: 50 sccm, VHF power: 150 W, pressure: 8 Pa as adopted from Figure 3 of reference [106]). The plasma wavelength is dependent only on the frequency, not the plasma gas conditions

Table 12.3 Comparison of solar cell properties for cells with *i*-layers deposited using RF and VHF frequencies and different deposition rates. Single-junction (1J) a-Si devices [115] and triple junction (3J) a-Si/a-SiGe/a-SiGe devices [97] are shown. In both cases, the VHF-deposited devices are superior at high deposition rate

Excitation frequency [MHz]	Device structure	Deposition rate [Å/s]	Initial power [mW/cm^2]	Degradation [%]
RF (13.56)	1J	0.6	6.6	14
VHF (70)	1J	10	6.5	10
RF (13.56)	1J	16	5.3	36
VHF (70)	1J	25	6.0	22
RF (13.56)	3J	1	~12.0	~16
RF (13.56)	3J	4-6	~10.7	~21
VHF	3J	6-8	~11.0	~8

make high-quality a-Si material at high rate using VHF could be very important for high-throughput manufacturing, especially when making thick nc-Si *i*-layers.

Although the advantages of using VHF deposition for high-rate growth have been clearly demonstrated, there are two principal challenges to applying VHF deposition in manufacturing: (1) Nonuniform deposition occurs on a large, production-scale substrate due to RF standing waves formed when the electrode size is comparable to a quarter of the wavelength of the RF wave. This is seen in Figure 12.11 where the wavelength of the plasma at 50 MHz is approaching ~1 m; e.g. the dimension of the reactor and glass substrate, leading to standing waves and spatially nonuniform plasma due to interference nodes. The manufacturer Oerlikon uses VHF-PECVD at 40 MHz, thus avoiding the problem for 1 meter sized substrates. Mitsubishi Heavy Industry has developed a "ladder" electrode allowing injection of small differences in phase or frequency between adjacent electrode elements to eliminate nodes yielding initial cell efficiencies over 11% at 2.5 nm/s with $f = 60$ MHz [105]. Sharp uses pulsed VHF where the rapid on–off nature of the plasma avoids establishment of standing waves [106]. (2) Uniformity also requires great care when coupling VHF power from the generator to large electrodes [107, 108] and from the electrodes to the substrate.

12.3.3.2 Microwave glow discharge deposition

Glow discharge deposition at a microwave frequency of 2.45 GHz produces very high deposition rates [109, 110], as expected from Figure 12.11. When the MW plasma is in direct contact with the substrate, the deposited films show poor optoelectronic properties compared with RF-deposited films, and are not suitable as intrinsic layers for high-efficiency solar cells. *Remote* MW excitation has yielded high-quality films [111, 112]. Substrates are placed outside the plasma region. The MW plasma is used to excite or decompose a carrier gas such as He, Ar, or H_2 that passes through the MW zone toward the substrates. The excited carrier gas then excites SiH_4 or Si_2H_6 directed into the chamber near the substrates. Using such an indirect excitation process, the concentration of SiH_3 radicals can be maintained, while the concentrations of other radicals (SiH_2, SiH, etc.) can be minimized, thus reducing the high deposition rate of the direct plasma. MW plasma deposition has been studied at USO [113] and Canon [114], and was used by Canon in their 10 MW_p triple-junction production line. Generally, the structural and optoelectronic properties of MW-deposited a-Si-based films are poorer than RF-deposited films. However, at a very high deposition rate, for example 50 Å/s, the MW-deposited films will be relatively superior to films made using RF and VHF deposition though all will be inferior to those made <10 Å/s.

12.3.4 Hot-wire Chemical Vapor Deposition

Several years after hot-wire chemical vapor deposition (HWCVD) was introduced [116, 117], Mahan *et al*. [118] improved the deposition process and produced "device quality" a-Si films. Since then, HWCVD has been studied and used on an experimental scale worldwide for depositing high-quality a-Si- and μc-Si-based films at high rate. The set-up for a HWCVD system is similar to the schematic shown in Figure 12.10 for RF-PECVD except that the RF electrode is replaced with a heated filament. In a HW process, SiH_4 gas or a mixture of SiH_4 and other gases such as H_2 or He is directed into the chamber. The gas is catalytically excited or decomposed into radicals or ions by a metal filament (Pt, W, Ta) heated to a high temperature (around 1800–2000 °C). The silicon radicals then diffuse inside the chamber and deposit onto a substrate placed a few centimeters away which is heated to 150–450 °C. Mahan *et al*. demonstrated that HWCVD a-Si materials show relatively lower H content in the film and improved stability against light-induced degradation compared with RF PECVD films [118]. The improved HWCVD a-Si has been incorporated in an *n-i-p* solar cell as the intrinsic layer and solar cells with ~10% initial efficiency have been demonstrated [119].

HWCVD is considered very promising. Although it has not yet been incorporated into any of today's large-scale manufacturing facilities, the ability to deposit a-Si and a-SiGe films at very high rate (~up to 150–300 Å/s) [120] has attracted tremendous interest. Another reason researchers are interested in HW CVD is its effectiveness in making nc-Si and polycrystalline silicon films. The group at Utrecht has made single junction nc-Si, triple-junction a-Si/a-SiGe/nc-Si and a-Si/nc-Si/nc-Si (no Ge) devices on flexible stainless steel having *stabilized* efficiencies of 8.6, 10.6, and 10.6%, respectively [121].

There are several concerns about incorporating HW processes in manufacturing. First, the uniformity of HW films is still poorer than that of RF PECVD films, although some companies have worked on this and made significant improvement [122]. Second, the filament needs to be improved to reduce the maintenance time in production. Third, HW-deposited solar cells have not yet achieved the same performance as cells prepared using low deposition rate, RF PECVD, although they have not enjoyed anywhere near the same level of R&D effort.

12.3.5 Other Deposition Methods

Beside PECVD and HW deposition methods, other deposition processes have been explored for depositing a-Si films. We only list those for which solar cell results were reported. These include (1) reactive sputter deposition from silicon targets using a mixture of hydrogen and argon [123]; (2) photo-CVD using ultraviolet excitation and mercury sensitization [124, 125]; (3) remote plasma chemical vapor deposition [126]; (4) electron cyclotron resonance (ECR) microwave deposition [127, 128] and; (5) gas jet deposition [129]. These deposition methods either yielded poorer a-Si films or solar cells compared with RF PECVD deposited films and devices, or they could not be easily scaled to provide large-area uniform films, and therefore, are not (or not yet) used in large-scale a-Si PV production.

12.3.6 Hydrogen Dilution

Hydrogen dilution of the silane gas mixture during a-Si deposition has been found to reduce the density of defect states and improve the stability of the material against light-soaking effects. Solar cells with *i*-layers deposited using high H_2 dilution ratios $R = [H_2]/[SiH_4]$ (brackets indicate the flow rate of the gas) showed improved performance and stability [130, 131]. There are two other

important effects of hydrogen dilution. As R is increased, the deposition rate decreases. When R is increased sufficiently, the deposited silicon films become nanocrystalline.

Real-time spectroscopic ellipsometry (RTSE) has been applied to examine the microstructure evolution for silicon films deposited under varying levels of hydrogen dilution of silane, leading to the very useful concept of "phase diagrams" [132]. Results based on *in situ* RTSE of the growing film are presented as Figure 12.12 [133]; this diagram pertains to a particular RF power level, substrate (c-Si), and substrate temperature. For lower dilutions ($R < 10$), films are invariably amorphous, but there is a transition to a "roughened" surface beyond a critical thickness. This roughening transition is suppressed as dilution is increased. For higher dilutions, the growing thin film initially adopts an amorphous structure, called the "protocrystalline" regime. As the film thickens, crystallites form in the amorphous matrix (creating a "mixed phase"). The crystallites are linked to improved stability [134]. Ultimately, the film becomes entirely nanocrystalline. The details of the phase diagram depend strongly upon the details of deposition, in particular upon power, thickness and substrate conditions, but the structure of the phase diagram is thought to be universal.

These effects of hydrogen dilution during growth are likely due to the following effects either singly or in combination. (1) Atomic hydrogen "etches" a growing film, removing strained weaker bonds that are in energetically unfavorable locations; (2) a high flux of atomic hydrogen promotes the surface diffusivity of adatoms so that they can move around to more energetically stable positions and form stronger bonds; (3) atomic hydrogen diffuses into the network, restructuring it and promoting a more stable structure. For the same reasons, sufficiently large hydrogen dilution induces the formation of nanocrystalline Si. The enhancement of short-range and long-range order through hydrogen dilution has been observed in many deposition techniques, including PECVD (DC, RF, VHF, and MW) and HW CVD; of course, the transitions from amorphous to nanocrystalline structures occur at different dilution levels for different deposition techniques. It is well established now that amorphous silicon has better stability against light-induced defects when

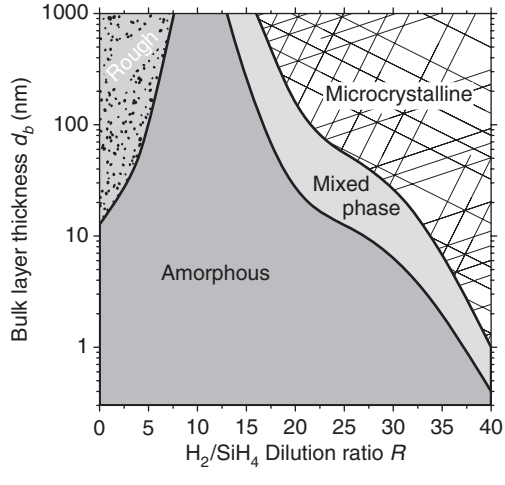


Figure 12.12 Phase diagram for the structure of plasma-deposited silicon thin films for varying dilution ratios R of silane in hydrogen and film thickness d_b; thin films were deposited onto a single-crystal Si substrate. For lower dilutions ($R < 10$) the films remain amorphous, but undergo a roughening transition in thicker films. For high dilutions, films start out as amorphous, develop and silicon crystallites, and ultimately become entirely microcrystalline. Based upon the phase diagram proposed by Ferlauto *et al*. [133] on the basis of *in situ* spectroscopic ellipsometry measurements

deposited under the conditions that are close to the nanocrystalline formation, sometimes called "near-the-edge" material [135].

The hydrogen dilution level for the transition from amorphous to nanocrystalline silicon thin films depends on other deposition conditions as well. At higher substrate temperatures (above 300 °C), the transition from amorphous to nanocrystalline state occurs at a higher R, likely due to the low sticking coefficient of hydrogen on the surface. At the low-temperature side (below 250 °C), it again takes a higher hydrogen dilution to reach the transition between amorphous to nanocrystalline; this effect is likely due to the low surface diffusivity of hydrogen during growth. When a-Si is deposited at a lower temperature with higher H dilution, more H is incorporated and the material has a wider bandgap. By staying within the protocrystalline region while reducing the deposition temperature, single-junction a-Si *n-i-p* cells with a wide-gap a-Si absorber were deposited having 1.04 V open-circuit voltage [83]. It was also observed that materials deposited in the transition region ("the edge") between protocrystalline (but still amorphous) and nanocrystalline formation show medium-range structural order [136]. The evolution in grain structure with R has been visualized via micrographs and diagrams [137].

12.3.7 High-rate Deposition of Nanocrystalline Si (nc-Si)

Due to their lower absorption (indirect bandgap ~1.1 eV), nc-Si films must be 5–10 times thicker than a-Si films when incorporated in the bottom cell in a multijunction cell structure, thus necessitating approximately a 5–10 times increase in deposition rate to maintain production throughput. Two conditions required for high-quality and high-rate nc-Si film growth are a sufficient flux of atomic hydrogen to the film growth surface to induce crystallinity, and a high flux of Si-containing radicals transported to the film growth surface to enhance the growth rate. Just increasing R leads to nc-Si growth, but at much reduced growth rate. (The lower deposition rate aids crystalline bond formation since it allows more time for the arriving Si atoms to form Si–Si bonds before being buried by the next wave of arriving Si.) Just increasing the RF power increases the growth rate, but also the ion bombardment of the film surface, leading to higher defects and breakdown of crystallinity. So far, two techniques have been widely studied and applied in production to enhance the growth rate of nc-Si films by PECVD while reducing the ion bombardment; (1) increasing the plasma excitation frequency from RF (13.5 MHz) to VHF (40–100 MHz); and (2) increasing the deposition pressure from <1 T to 5–10 T (called high pressure depletion condition or HPD). The group at AIST in Japan have studied these two processes in considerable detail and their effect on nc-Si growth rate, properties, and device performance; an excellent review is found in reference [96].

Figure 12.13 shows the empirical dependence of growth rate on RF power density for "device-grade" nc-Si for the four combinations of RF and VHF with LPD and HPD. The highest rate (>2 nm/s) requires high power and VHF in the HPD regime.

The VHF PECVD can increase the growth rate of nc-Si films without deterioration of the material quality [138] because the higher electron concentration in the plasma (Figure 12.11) increases the beneficial radical density (atomic hydrogen and SiH_x precursors) while decreasing the harmful ion energy. These radicals provide both selective etching of disordered amorphous phase and fast growth of crystalline grains.

Regarding the standard $f = 13.56$ MHz plasma, increasing the RF power at pressures less than 1 Torr leads an increased concentration of high-energy positive ions bombarding the growth surface which disrupts crystallinity and increases the concentration of negative ions which coagulate to create dust particles in the plasma. However, higher pressure (>1 T) plasma reduces the ion energy and bombardment to the growth surface due to collision losses. In the HPD growth regime,

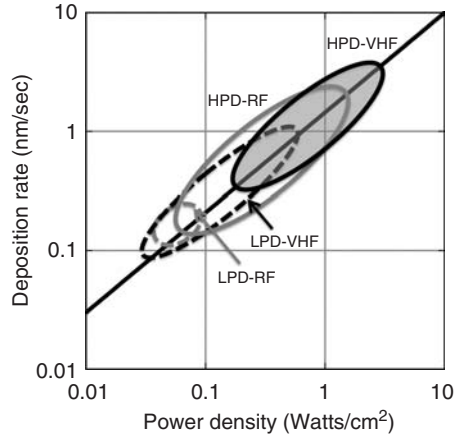


Figure 12.13 Trends of deposition rate of "device-quality" nc-Si vs the RF or VHF power density. LPD and HPD refer to low- and high-pressure depletion regimes, respectively. LPD data-dashed line, HPD data-solid line, RF-gray, VHF-black. The VHF-HPD data has shaded fill. Approximate range of values of data for the four combinations based on data from over 30 published reports. The highest rates, over 2 nm/s, occur with both high frequency and high pressure. The solid line is an empirically derived fit showing a power-law dependence with an 0.85 exponent. Based on figure 1 of reference [96]

the saturation of the deposition rate with increasing pressure was attributed to a depletion of Si source gas silane. This enhances the flux of atomic hydrogen to the film growth surface by reducing the hydrogen annihilation reaction in the gaseous phase, $H + SiH_4 \rightarrow H_2 + SiH_3$ due to lack of available SiH_4 molecules. The electron density increases and energy decreases with increasing pressure, becoming qualitatively similar to the VHF plasma for pressures above 1 Torr. The growth of nc-Si at high growth rate in standard RF plasma CVD has been demonstrated using high-power and high-pressure deposition conditions [139] which have been scaled up to $\sim 1\,m^2$ area [140]. However, there tends to be a very narrow process window for achieving high growth rate and high efficiency, especially over large areas where uniformity is critical.

12.3.8 Alloys and Doping

As was discussed in Section 12.2.7, a-Si-based alloys can be deposited using a gas mixture of SiH_4 with other gases such as GeH_4, CH_4, O_2 (or NO_2), and NH_3 for obtaining a-$SiGe_x$, a-SiC_x, a-SiO_x and a-SiN_x, respectively. Among these alloy materials, only a-SiC, as a wide bandgap *p*-layer, and a-SiGe, as a low-bandgap absorber layer, have been explored extensively for PV applications. a-SiGe was widely studied in the 1980s and has been in production at United Solar Ovonic, and previously Canon and BP Solar, as the narrow bandgap absorber for multijunction devices for 20 years. As we see from Figure 12.9, the bandgap E_G decreases with increasing Ge content. When E_G is decreased to below 1.4 eV, the defect density becomes so high that the materials are no longer practical as the intrinsic layer for solar cells. Despite tremendous progress, device quality a-SiGe with low bandgap (below 1.3 eV) has not been demonstrated, largely due to structural defects leading to low mobility and electronic defects [57, 141] and the intensity of research on low bandgap a-SiGe has declined.

A critical aspect for a-SiGe is the deposition uniformity. Because of the different dissociation rates of germane (GeH_4) and of silane (SiH_4) in an RF plasma, the film deposited near the gas inlet

side of the chamber has higher Ge content than the film near the exhaust. This nonuniformity makes it difficult to implement the process over large areas in manufacturing. By taking advantage of the approximately similar dissociation rate of GeH_4 and disilane (Si_2H_6), some research groups use a mixture of GeH_4 and Si_2H_6 for the fabrication of a-SiGe alloy and successfully obtain uniform films [142].

As discussed in Section 12.2.6, a-Si can be doped *n*-type by mixing phosphine (PH_3) with the gas mixture or doped *p*-type by mixing diborane (B_2H_6), BF_3, or trimethylboron [TMB, $B(CH_3)_3$] with the gas mixture during deposition. Because of the need for transparency in *p*-layers, which act as the "window" layer for sunlight, most cells have either μc-Si or a-SiC as the uppermost *p*-layer. Amorphous SiC *p*-layers in *p-i-n* devices are usually deposited using a mixture of SiH_4 and CH_4 diluted in hydrogen [143], leading to bandgaps of 1.85–1.95 eV. The μc-Si *p*-layer for *n-i-p* devices is generally made in a PECVD process using high H dilution with high RF power at relatively low temperature. The optimum *p*-layer for a-Si *n-i-p* solar cells is a mixed phase of amorphous and nanocrystalline [144]. The *p*-layer is typically as thin as possible, ~50–150 Å, in order to minimize parasitic absorption. Note that properties of such thin layers may differ significantly from those of >1000 Å films often used for characterization and depend strongly on the substrate [145].

12.4 UNDERSTANDING A-SI PIN CELLS

12.4.1 Electronic Structure of a *pin* Device

Figure 12.14 illustrates the profiles of the bandedge levels E_C and E_V for an a-Si:H-based *pin* solar cell in the dark and under illumination. These spatial variation of the levels is due to the presence of a "built-in" electric field $F(x)$ within the device, which causes all electron level energies such as E_C and E_V to vary in space in the same way. For E_C the expression is $eF(x) = \frac{\partial E_C(x)}{\partial x}$, where e is the magnitude of the charge on an electron. The figure is based on computer calculations that we describe shortly; direct measurements of these profiles aren't readily accomplished with current techniques.

Where do these built-in electric fields come from? In isolation, *p*-type and *n*-type materials have very different Fermi energies; in the calculation of Figure 12.14 we assumed that E_F was 0.2 eV above E_V for the *p*-layer, and 0.1 eV below E_C for the *n*-layer. When the *pin* device is assembled, these Fermi energies must be equalized to create thermal equilibrium. Electrons are donated from the *n*-layer to the *p*-layer, which generates a built-in electric field; while the level positions such as E_C and E_V now vary across the device, the Fermi energy itself is constant. The original difference in Fermi energies becomes the "built-in potential" eV_{BI} across the device, which is illustrated in Figure 12.14a. The built-in field is illustrated in Figure 12.14c (labeled "dark"). Electrons and holes that are generated by photon absorption will drift in the built-in electric field in the directions illustrated in Figure 12.14a.

There are several computer programs that are convenient for solar cell calculations with a-Si:H and nc-Si:H [146]. The calculations used to prepared the figures for this section were done with the AMPS-1D© computer program [147] and a minimal set of parameters that doesn't include defects [148]; in amorphous semiconductors, band tail states are always present, and may dominate defects in determining some optoelectronic properties. We shall shortly give references to more complex calculations including defects, which must certainly be included in modeling of light-soaked cells. One good introduction to modeling of these cells is the monograph of Schropp and Zeman [149].

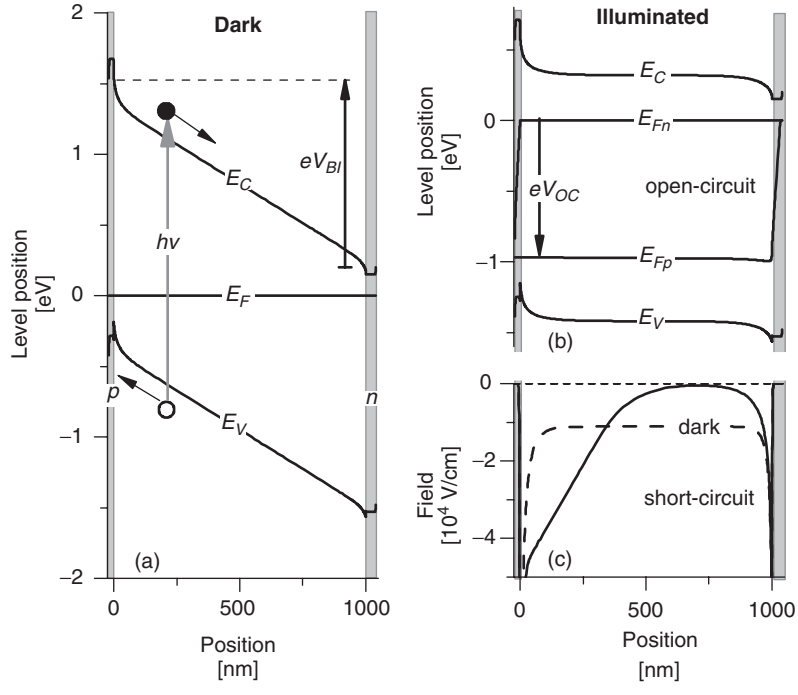


Figure 12.14 Calculated profiles for electron level positions and for electric fields in a *pin* solar cell. (a) Levels in the dark; note the built-in potential V_{bi} and that the p-layer has a slightly (0.2 eV) larger bandgap than the i-layer. (b) Open-circuit levels under illumination (uniform photogeneration $G = 3 \times 10^{21}\,\mathrm{cm^{-3}s^{-1}}$); note the open-circuit voltage V_{OC}. (c) Short-circuit electric field, which collapses under illumination from near-uniformity in the dark to a width $d_C \sim 400\,\mathrm{nm}$

12.4.2 Voltage Depends Weakly on Absorber-layer Thickness

In Figure 12.15 we show experimental results for the power and the open-circuit voltage of a series of a-Si:H solar cells with varying thickness [150]; measurements are shown both for the initial state of the cells, and after light-soaking for 800 hours. For the moment, we emphasize the initial state of the cells, for which the figure illustrates two important features: (i) the output voltage V_{OC} hardly varies with the thickness, and (ii) the power output rises steeply for thinner cells, and then becomes nearly independent of thickness beyond some "useful thickness".

We first explain how V_{OC} is related to the profiles of Figure 12.14, and why V_{OC} depends only weakly on thickness. Figure 12.14b presents calculated open-circuit profiles of the band edge levels E_C and E_V for a cell with uniformly absorbed illumination, which turns out to be a useful simplification compared with solar illumination. No Fermi energy is shown in this panel of the figure because the cell is not in thermal equilibrium – it is exposed to light. Instead, "quasi-Fermi energies" E_{Fn} and E_{Fp} are illustrated that describe the photogenerated populations of electrons and holes separately (see Section 3.3.1 for explanation of quasi-Fermi levels). Notice that these quasi-Fermi energies merge together at the left edge of the p-layer, and again at the right edge of the $n-$layer; merging means that an ordinary Fermi energy applies despite the presence of light. The product eV_{OC} is the difference in these two quasi-Fermi levels, as illustrated in the figure. The center region of the cell is acting essentially as a solar battery!

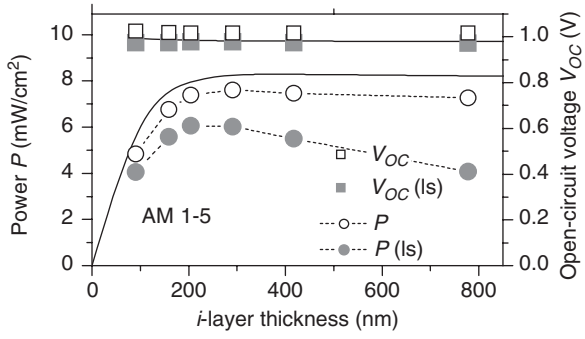


Figure 12.15 Power (circles) and open-circuit voltage (squares) recorded for each of a series of as-deposited a-Si:H solar cells with varying intrinsic layer thickness [150]. The open symbols are initial measurements, and the solid symbols are measurements after 800 hours of light-soaking. The solid lines are calculations described in the text

The electron quasi-Fermi energy E_{Fn} [151, 152] is defined by the following equation:

$$E_{Fn} \equiv E_C + k_B T \ln\left(\frac{n}{N_C}\right), \tag{12.2}$$

where n is the density of mobile electrons in the conduction band (i.e. in the shaded region of the conduction band in Figure 12.8). N_C is the effective density (cm^{-3}) of these conduction band states, and $k_B T$ is the product of Boltzmann's constant k_B and the temperature T (in K). A similar expression defines the quasi-Fermi energy for holes E_{Fp} in terms of the density of holes p and the effective density N_V of valence band states.

In Figure 12.14 the hole quasi-Fermi level E_{Fp} is nearly constant across the cell, showing sizable variation only where it catches up to E_{Fn} in the n-layer. Similarly, the electron quasi-Fermi level is constant except near the p-layer. These behaviors mean that the quasi-Fermi levels in the middle of the cell largely determine V_{OC}.

It is possible to derive a formula for the separation of E_{Fn} and E_{Fp} in the i-layer based on the density-of-states of Figure 12.8 and on the approximations that the conduction band tail traps and the deep levels can be neglected. The result is [162]:

$$eV_{OC} = E_G - \frac{k_B T}{2}\left\{\ln\left(\frac{b_R N_C^2}{G}\right) + 2\ln\left(\frac{b_T N_V^2}{G}\right)\right\} + \frac{(k_B T)^2}{2\Delta E_V}\ln\left[\frac{b_T}{b_R}\left(\frac{b_T N_V^2}{G}\right)\right]. \tag{12.3}$$

The photogeneration rate G is the rate per unit volume ($\text{s}^{-1}\,\text{cm}^{-3}$) at which light creates pairs of electrons and holes; the calculation assumes that it is uniform in the intrinsic layer of the cell. E_G is the bandgap energy; and ΔE_V is the width of the valence bandtail. The coefficient b_R describes the rate at which electrons and holes "recombine" and annihilate each other; specifically, the calculation assumes that the recombination rate R per unit volume is $R = b_R n P$, where n is the density of electrons and P is the total density of holes, both mobile and trapped in the valence bandtail. b_T describes the rate at which the mobile holes are trapped onto valence bandtail states.

Three properties of V_{OC} in actual cells can be understood from this equation. (i) V_{OC} doesn't depend on thickness; indeed, thickness isn't a variable in the equation. (ii) V_{OC} increases with increasing bandgap E_G. If no other properties change with bandgap, the change should be

1:1 – a 0.1 eV increase in bandgap yields a 0.1 V increase in V_{OC}. (iii) V_{OC} decreases roughly linearly with increasing temperature T.

Equation (12.3) gives a very good account of the temperature-dependence of V_{OC} in a-Si:H cells when it is measured using uniformly absorbed light [153], and when the decline of E_G with increasing temperature is incorporated. This equation does neglect defects. Defects are undoubtedly important in light-soaked cells [154–156], although they barely affect V_{OC}. For as-deposited cells, neglecting them in calculations of P and V_{OC} appears to be a good approximation at solar illumination intensities [157].

12.4.3 What is the Useful Thickness for Power Generation?

We now turn to the useful thickness of a cell, beyond which its power no longer increases. In Figure 12.15 this thickness is about 200 nm. There are two factors that determine this thickness. First, the blue components of the solar spectrum are absorbed very close to the interface; in a 1000 nm absorber, more than 70% of the light is absorbed within the first 200 nm [158], so the additional 800 nm doesn't absorb much more.

The more important factor is the range of travel of the photogenerated photocarriers; a hole photocarrier generated near the p-layer is more likely to be collected than one generated far away from it. The range over which photocarriers can be collected in a-Si:H is mostly determined by the "drift mobility" μ_D of the holes, and we take a detour to explain this property. μ_D describes the displacement $x(t)$ of a carrier following its photogeneration at time $t = 0$. If the field is uniform, then μ_D can be defined using $x(t) = \mu_D F t$. Some graphs of the displacement–time curves for carriers in a-Si:H and nc-Si:H are illustrated in Figure 12.16 for $F = 10^4$ v/cm [159]. The electrons in a-Si:H have a mobility of about 2 cm^2/V s, which is about 500 times smaller than for crystalline silicon; the disorder of the amorphous network obviously interferes with electron motion.

The holes in a-Si:H are even more sluggish; it takes them about 1 µs to move 200 nm, which is 1000 times longer than it takes electrons. This low drift-mobility is due to trapping of the holes by the valence bandtail states [160, 161]. Holes in nc-Si:H are much more mobile than holes in a-Si:H, but their mobility is still within the range characteristic of amorphous, rather than crystalline, silicon.

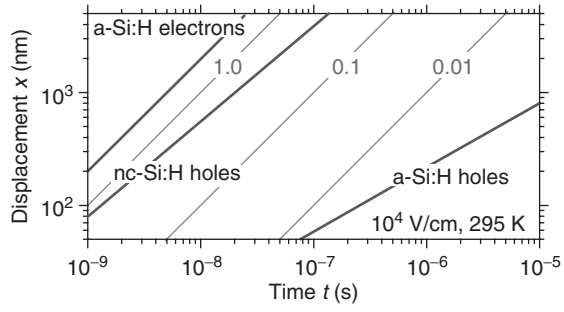


Figure 12.16 Photocarrier displacements as a function of time following photogeneration at $t = 0$. The bolder lines for electrons and holes in a-Si:H and for holes in nc-Si:H are based on "multiple-trapping" fitting parameters to experiments [159]. The fine, gray reference lines indicate displacements for drift mobilities 1.0, 0.1, and 0.01 cm^2/V s

Low hole drift mobilities affect the functioning of a solar cell markedly [162]. When photogeneration rates are large enough, the density of positive electrical charge due to the slowly drifting holes builds up enough to rebuild the built-in electric field profile of the cell. For short-circuit conditions, this effect can be seen in Figure 12.14c: the electric field that is nearly constant across the device in the dark is changed by illumination, and becomes stronger near the p-layer and becomes negligible near the n-layer.

Figure 12.17 shows calculations of these effects at the maximum power point of the cell. In Figure 12.17a, the photocarrier recombination profile $R(x)$ is illustrated for three values of the hole's band mobility μ_p; the hole drift-mobility, while lower than μ_p, is proportional to it. There is a "collection zone" with $R \sim 0$ that stretches out from the p-layer interface; carriers generated in the zone generally contribute to the photocurrent (and power) from the cell. Beyond the collection zone is a recombination region, where carriers are likely to recombine instead of contributing to the cell's power.

The width of the collection zone, which we define from the depth at which $R/G = 0.5$, increases from about 200 to 400 nm as the hole band mobility increases from 0.1 to 1.0 $cm^2/V\,s$. The lower panel shows that the collection zone is associated with positive space-charge ρ near the p-layer. Because photogenerated electrons have much larger drift mobilities than holes, the electrons generated in the collection zone are quickly swept towards the n-layer and into the recombination region. The slowly drifting holes that are left behind create a space-charge. In turn, the space-charge causes the magnitude of the electric field to fall – faster for larger space-charges. Larger hole mobilities reduce the space-charge, allowing the field to change more slowly and to stretch out the collection zone [163]. Although recombination is the fate of carriers that aren't collected, the recombination parameters in a-Si:H play a minor role in determining d_C compared with the mobility parameters.

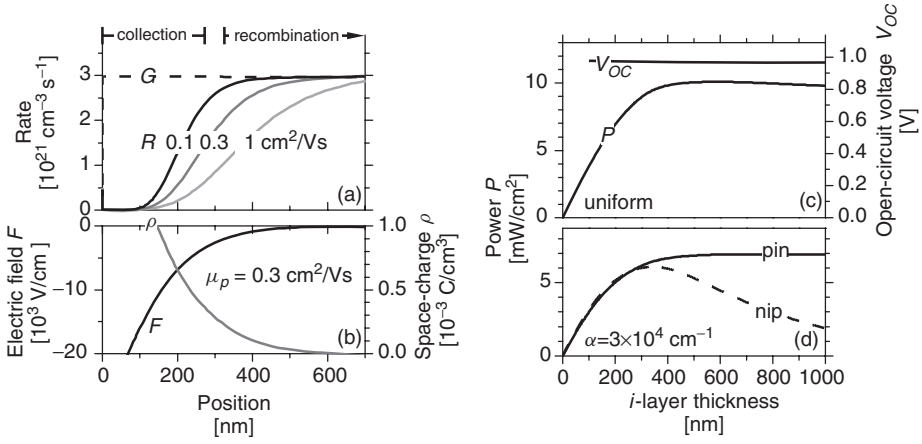


Figure 12.17 (a,b) Calculated profiles of the recombination rate R, the electric field F, and the space-charge density ρ calculated for an a-Si:H cell (1000 nm i-layer thickness) with uniform photogeneration G. The three profiles for R correspond to the three hole band mobilities μ_p indicated; the associated voltages and powers are (0.69 V, 7.3 mW/cm^2), (0.71, 9.8), (0.72, 13.9). (c) Calculated power and open-circuit voltage for varying i-layer thickness and $G = 3 \times 10^{21}\,s^{-1}\,cm^{-3}$; note that the power saturates at thicknesses larger than the collection zone width shown in panels (a) and (b). (d) Power for varying thicknesses using photogeneration G that decays exponentially from a value 3×10^{21} at the p/i interface (conventional illumination) and at the n/i interface (reversed illumination)

Figure 12.17c shows the corresponding power calculated for cells with varying i-layer thicknesses. The power saturates for i-layer thicknesses larger than the width of the collection zone, but V_{OC} is nearly unaffected by the cell thickness, as we've discussed previously. This figure agrees fairly well with the experimental measurements under solar illumination in Figure 12.15.

Panels (a–c) of Figure 12.17 showed calculations with uniform photogeneration G. Panel (d) shows the power for the nonuniform photogeneration associated with light that is absorbed with an absorption coefficient of $\alpha = 3 \times 10^4\,\mathrm{cm}^{-1}$. The photogeneration G at the p/i interface is the same for all the panels, but G falls exponentially with depth for (d). The curve labeled *pin* in panel (d) corresponds to the standard situation of illumination through the p-layer; as expected from (c), the power rises monotonically. Very thin cells have nearly uniform photogeneration, and they agree well between (c) and (d). The power for thicker cells saturates at nearly the same thickness as for (c), but at a somewhat smaller power than for uniform photogeneration.

The curve labeled *nip* corresponds to "incorrect" illumination through the n-layer, and it explains why a-Si:H solar cells are illuminated through their p-layers. Very thin cells ($d < 200\,\mathrm{nm}$) still have uniform photogeneration, and don't notice the change. For thicker cells, the power actually declines with thickness, which is very different than the case of uniform photogeneration. In these cells, holes are mostly generated near the n-layer, and then must travel across the entire cell to be collected. The space-charge buildup problem is even worse than it is for uniform photogeneration, and the cell's power falls. The fact that a-Si:H cells need to be illuminated through their p-layers is a consequence of the poorer drift-mobility for holes than for electrons in a-Si:H. In a (hypothetical) material in which an electron's drift-mobility was worse than the hole's, we would want to illuminate a cell through its n-layer. In general, you want to design a solar cell so that light is incident through the contact which collects the "limiting carrier" (the one with the shorter diffusion or drift length) so they have less distance to travel before being collected. The hole drift length limitations are responsible for the voltage dependent collection loss which impacts the p-i-n solar cell fill factor and photocurrent [84].

12.4.4 Doped Layers and Interfaces

Our discussion so far has neglected the details of the doped layers in these *pin* solar cells. For the models that were shown, the p and n layer properties were adjusted until they were essentially ideal, and had little effect on the calculations. Achieving such ideal layers and interfaces in actual cells is more challenging; generally speaking, the p/i interface is the more difficult one, and poor p/i interface regions lead to reduced values of the open-circuit voltage.

Two important innovations from the late 1980s improve V_{OC} and are in widespread use. In substrate solar cells, generally made on stainless steel in the *nip* growth sequence, the final p-layer that yields the best open-circuit voltage is a boron-doped "nanocrystalline" silicon film [164]. In superstrate cells, usually made on TCO-coated glass, the best open-circuit voltages have been achieved using boron-doped amorphous silicon–carbon alloys (a-SiC:H:B); an indication of the subtlety required to achieve high open-circuit voltages is that cells using a-SiC:H:B p-layers also include a thin ($<10\,\mathrm{nm}$) "buffer layer" of undoped a-SiC:H between the p-layer and the intrinsic-layer of the cell [165–167].

The mechanisms underlying these technologies have been studied by several groups [168, 169]. Lowering the Fermi energy of the p-layer to increase V_{BI} is plainly desirable, but wouldn't account for the success of buffer layers. Their success probably involves a barrier to the diffusion of photogenerated electrons from the i-layer into the p-layer.

An important question is whether these strategies have allowed V_{OC} to reach the "intrinsic limit" that is determined by the intrinsic layer, and not by the doped layers. For two reasons, we

believe that the best V_{OC} cells are close to this limit. First, the temperature dependence of V_{OC} in high V_{OC} cells is consistent with calculations of the intrinsic limit [157]. Second, the dependence of V_{OC} upon the bandgap of the absorber layer is also consistent with this limit, as we discuss in a subsequent section.

12.4.5 Light-soaking Effects

In Figure 12.15 we illustrated the power output for a series of a-Si:H cells of varying thickness prepared at United Solar Ovonic in 2009 [150]. Results were shown both for the initial state of the cells and after 800 hours of light-soaking under open-circuit conditions; it is worth noting that light-soaking effects on nc-Si:H cells are much smaller than for a-Si:H cells. We think that the initial cell performance of both a-Si:H and nc-Si:H cells can often be understood without considering defects (the D-centers) at all. As light-soaking proceeds, the defect density steadily rises as was illustrated in Figure 12.7; ultimately, the rising defect density does reduce the cell efficiency. It is remarkable that light-soaking reaches its steady-state just as its effects on the cell efficiency become significant, but there is no consensus about the reason for this apparent coincidence.

For some cells that were light-soaked with uniformly absorbed illumination, reasonably satisfactory calculations of cell properties have been done by simply adding a uniform density of light-induced defects to the bandtail [157]. This approach did not give a good account for the light-soaked measurements in Figure 12.15. One expects that a light-soaked cell could have a defect density that is not constant across the intrinsic layer thickness. In addition, researchers have often invoked the principle that the properties of a defect depend on the details of its creation, so the energy level within the bandgap as well as the density of defects may vary with spatial location within the cell. While the complexity of such modeling is considerable, several researchers have done it; see [172] and the references there for more information.

12.4.6 Alloy and Nanocrystalline Cells

In Figure 12.18 we present a summary of the open-circuit voltages (V_{OC}) for amorphous silicon-based solar cells as a function of the bandgap of the intrinsic absorber layer [173]; the measurements include results from silicon–germanium, silicon–carbon alloy cells, and (separately) nc-Si:H cells. The measurements were done under standard solar illumination conditions. For amorphous cells, they show that the voltage output can be related to the bandgap of the cells by $V_{OC} = (E_G/e) - 0.80$; we shall refer to the difference of 0.8 V as the "V_{OC}-deficit". The fitting (and deficit) are consistent with the V_{OC} (predicted by Equation 12.3) [174].

In the figure, we assigned nc-Si:H a bandgap of 1.12 eV, which is the same as that of c-Si [175]. The deficit of 0.80 V between E_G/e and V_{OC} in a-SiGe:H alloys shrinks to about 0.55 V for the highest efficiency nc-Si:H cells (about 10%). The reduced deficit in nc-Si:H compared with a-SiGe:H alloys reflects smaller values both for the effective band densities-of-states N_C and N_V and for the band tail widths [176].

The useful thickness of cells made with the Ge-alloys of a-Si:H is apparently not greatly different from that of a-Si:H. This result is consistent with the limited measurements, indicating that hole drift mobilities do not change greatly with alloying [159]; electron drift mobilities are substantially affected by alloying.

The thickness of optimized nc-Si:H cells is nearly ten times larger than for a-Si:H. This is fortunate, because nc-Si:H has a lower absorption coefficient than a-Si:H throughout most of the visible spectrum. The increased thickness is enabled by the large increase in hole drift mobility (see

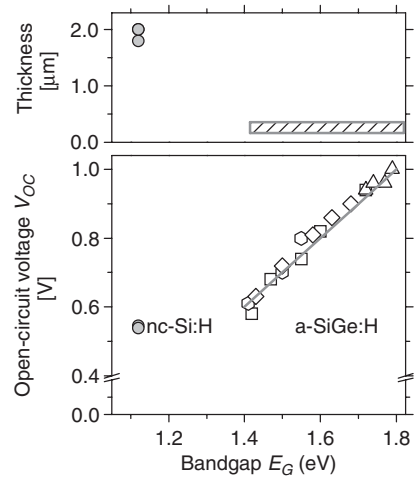


Figure 12.18 Open-circuit voltages and thicknesses for nc-Si:H, a-SiGe:H, and a-SiC:H single-junction solar cells as a function of absorber layer bandgap [170, 171, 173]

Figure 12.16) in nc-Si:H compared to a-Si:H [176], which allows photocarrier collection from a larger thickness of material. The "micromorph" cell design we describe shortly takes advantage of these complementary properties to give larger efficiency than could be obtained with either a-Si:H or nc-Si:H alone.

12.4.7 Optical Design of a-Si:H and nc-Si:H Solar Cells

In this section we briefly review the use of back reflectors and substrate texturing, which are optical design principles that are used to improve the power output of most thin film solar cells. Figure 12.19a shows the "quantum efficiencies" measured for a 250-nm-thick a-Si:H solar cell with varying texturing and back reflectors [177], along with absorptances calculated from the optical absorption coefficient spectrum for a-Si:H. The dashed absorptance curve labeled "no light trapping" corresponds to $A(\lambda) = 1 - \exp(-\alpha(\lambda)d)$, and is the fraction of light that is absorbed by a 250 nm layer of a-Si:H assuming normal incidence and neglecting reflection from both the front and rear interfaces. This absorptance falls with increasing wavelength, which just corresponds to the absorption coefficient spectrum of Figure 12.2 and the thickness of 250 nm.

The quantum efficiency (QE) of the solar cell is defined as the ratio, at a specific wavelength, of the photocurrent density j (A/cm^2) to the incident photon flux f(cm^{-2}s^{-1}):

$$QE(\lambda) = j(\lambda)/[ef(\lambda)], \tag{12.4}$$

where e is the electronic charge. The QE measurements for an a-Si:H *pin* cell deposited on a smooth, TCO-coated substrate without a back reflector are indicated with the diamond symbols. For wavelengths longer than 550 nm, they are fairly close to the absorptance calculated with minimal light-trapping. This result indicates that each photon absorbed by the cell in this range contributes to the photocurrent. At shorter wavelengths, the quantum efficiency falls away from the absorptance curve. This "blue dip" is mainly due to the absorption of light in the p-layer, which doesn't contribute significantly to power generation. The absorption coefficients of a-Si:H and of the p-layer are larger than 10^5 cm^{-1} at these wavelengths, so even a thin p-layer absorbs a significant fraction of the incident light. Photocarriers that are generated in the p-layer (or the n-layer) are

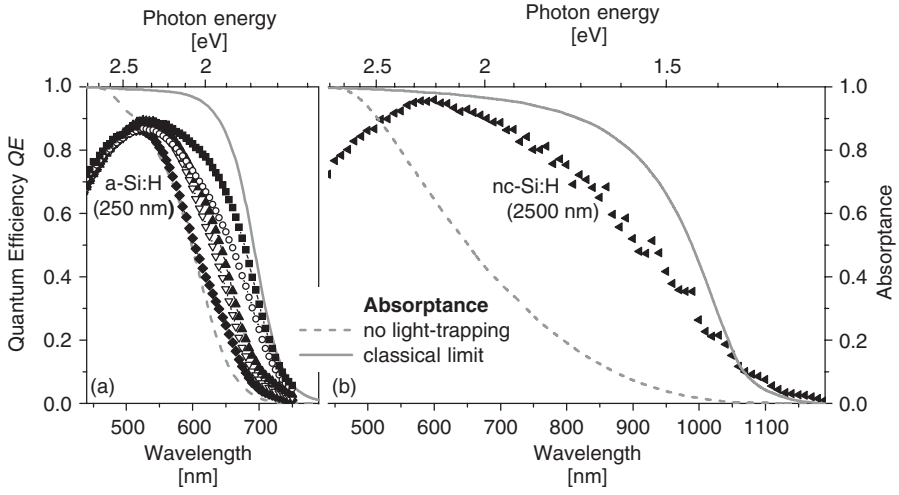


Figure 12.19 (a) Absorptance spectra calculated for a 250 nm a-Si:H, and quantum efficiency spectra measured for a series of a-Si:H *nip* solar cells with a 250 nm thick *i*-layer and varying substrates and back reflectors [177]: ♦: (untextured, none), ∇: (untextured, untext.), ▲ (untextured, textured), ○ (textured, none), ■ (textured, untextured). (b) Calculated absorptance spectra and quantum efficiency measured for a nc-Si:H solar cell [184]

generally neglected when estimating the photocurrent. Electrons generated in a *p*-layer are much more likely to be captured by a dopant or a dangling bond than to escape to the *i*-layer, and similarly for holes generated in an *n*-layer.

As can be seen in Figure 12.19, the introduction of a smooth back reflector roughly doubles the QE for the longer wavelengths, which are weakly absorbed. The back reflector permits the incident beam to travel twice through the film. The remaining curves show that "texturing" of the substrate and of the back reflector further increases the QE from the cells, by allowing light to traverse more than twice the thickness of the film before leaving the film. The substrate texturing also leads to a modest improvement of the QE in the blue spectral region (beyond 2.5 eV), which is due to a reduction in the front-surface reflectance of the cell. For a-Si:H, the best texturing schemes increase the short-circuit current of the working cell by about 25% [178, 179].

There are several perspectives that can be offered to explain the success of texturing for increasing the QE. One simple viewpoint is that texturing or patterning causes the incident light to be refracted at angles that increase the path length of the light inside the film, and that give rise to the possibility of total internal reflection. This view is correct for solar cells such as crystalline silicon cells that are very thick compared with the wavelength of light, but it doesn't apply directly to fairly thin a-Si:H solar cells. A deeper analysis was given in 1984 by Yablonovitch [180]. He noted that the density of electromagnetic modes is larger in a dielectric film by the factor n^2 than in vacuum, and that randomization of the incident light by texturing thus leads to a corresponding increase in the intensity of light inside the cell. For a cell with a perfect back reflector, the increase in the absorptance of the cell due to this "stochastic light trapping" is $4n^2$ for weakly absorbed light.

We shall refer to the stochastic limit as the "classical limit" for light-trapping, following Zhou and Biswas [181]. Tiedje, *et al*. [182] proposed an approximate formula for the classical limit to absorptance that is usable at all wavelengths, including those where the absorption is not weak. We have used their formula to calculate the classical limit curves in Figure 12.19; note that the classical limit absorptance depends fairly strongly on the thickness of the absorber.

For the a-Si:H cells, the best results approach fairly closely to the classical limit at longer wavelengths. Several authors have studied the QE from a-Si:H cells using combined optical and electrical modeling in order to better understand origins of the measured QEs; as one example, Krc and his collaborators were able to account for the measured QE spectra for a-Si:H cells of several thicknesses by using measurements of the scattering by a textured substrate [183].

Figure 12.19b shows absorptances and a QE measurement for a nc-Si:H cell of 2500 nm thickness [184]. The absorptances were calculated using the optical absorption coefficient spectrum of c-Si, which is a decent approximation to nc-Si:H for this purpose [185]. It is important that the effects of light-trapping extend over a much broader region of the spectrum for nc-Si:H than for a-Si:H. This effect is due to the weaker wavelength dependence of the absorption coefficient α for nc-Si:H compared with a-Si:H, and it also implies that light-trapping strategies can contribute more to the efficiency of nc-Si:H cells than to a-Si:H cells.

The implementations of texturing and back reflectors, as well as of a front "antireflection" coating to reduce the reflection, vary dramatically between superstrate and substrate cell designs; this subject is discussed at more length in Chapter 17 of this Handbook. Superstrate cells usually incorporate a textured, transparent conductive oxide (TCO) coating on the transparent substrate (usually glass). There are many technologies for producing TCO layers from varying materials (typically SnO_2 or ZnO for a-Si based cells) and with varying texture and electrical properties. The semiconductor layers are then deposited onto the textured TCO. Plasma deposition of the *p*-layer onto a textured TCO can lead to difficulties: the oxide layer may be chemically "reduced," and achieving ideal properties for a thin *p*-layer can be difficult. Finally, the back reflector deposited on top of the semiconductor layers is often a two-layer structure: a thin TCO layer, followed by the reflective metal (typically Ag, for best reflectivity, or Al, for improved yield in production).

In substrate cells, the semiconductor layers are actually deposited onto the back reflector, which is again a two-layer structure starting with a textured silver or aluminum metallization and then a textured TCO [186]. Following deposition of the semiconductor layers, a top TCO layer is applied.

While most workers have studied light-trapping effects in the context of textured dielectrics and imperfect textured back-reflectors, the "plasmonic" properties of metal films are also important, and can lead both to losses and to improved light-trapping. Indeed, recent calculations have shown that structured back reflectors can lead to light-trapping that exceeds the classical limit [181]. There have been numerous studies of plasmonic light-trapping configurations for cells based on amorphous silicon and other materials, although as of 2009 there have been no reports of cell quantum efficiencies exceeding the classical limit [187].

12.5 MULTIJUNCTION SOLAR CELLS

12.5.1 Advantages of Multijunction Solar Cells

Amorphous silicon solar cells can be fabricated in a stacked structure to form multijunction solar cells. This strategy is particularly successful for amorphous materials, both because there is no need for lattice matching, as is required for crystalline heterojunctions, and also because the bandgap is readily adjusted by alloying with Ge or by forming nc-Si. Figure 12.3 illustrated the structure of a tandem cell with two junctions (i.e. two *pin* photodiodes) in series. Multijunction, a-Si-based solar cells can be fabricated with higher solar conversion efficiency than single-junction cells (Figure 12.4) and are presently used in most commercial modules. Until the late 1990s the most

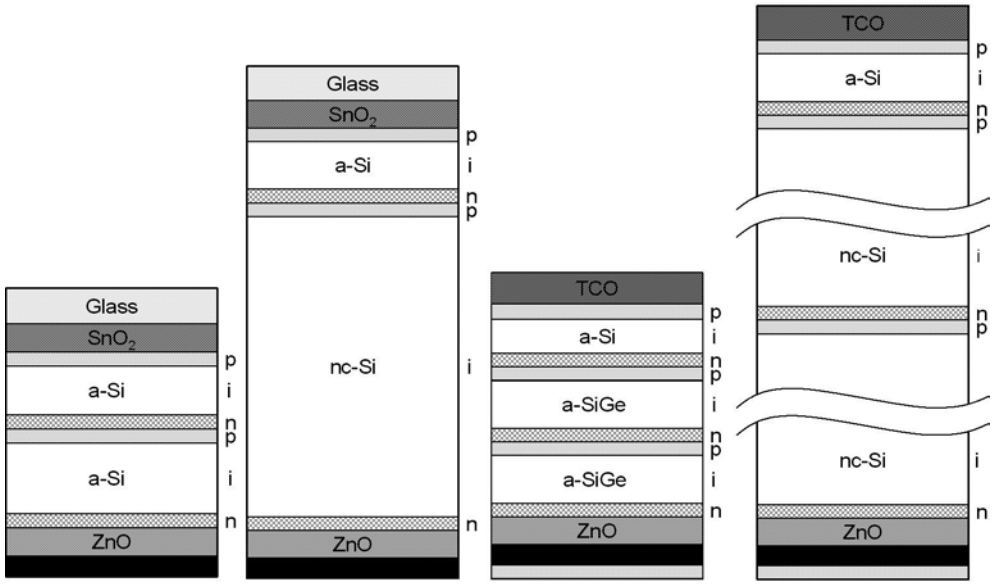


Figure 12.20 Multijunction device structures currently in production or under serious investigation: (a) a-Si/a-Si superstrate tandem; (b) a-Si/nc-Si "micromorph" superstrate tandem; (c) a-Si/a-SiGe/a-SiGe substrate triple; and (d) a-Si/nc-Si/nc-Si substrate triple. The light enters from the top in each device. The solid black region is the rear metal contact. TCO can be ITO or ZnO. Relative difference in thickness of the *i*-layers is suggested

common multijunction a-Si based solar cells were tandem or triple junctions with a-SiGe low band gap absorbers. Since then, tandem "micromorph" cells with a-Si wide-bandgap top cells and nc-Si low-bandgap bottom cells have been widely studied and several large manufacturers have begun to commercialize such a-Si/nc-Si multiple-bandgap tandem modules. Figure 12.20 shows several commonly studied multijunction solar cell structures, where the *i*-layer thicknesses are shown "somewhat" proportional to reality to give a visual indication of the significant difference between an a-SiGe and nc-Si bottom cell.

One fundamental concept underlying multijunction solar cells is "spectrum splitting." Consider two *pin* junctions, one deposited on top of the second. The top junction "filters" the sunlight to the bottom junction: photons absorbed in the top cell are removed from the light that reaches the bottom cell. We illustrated this filtering effect in Figure 12.2, which shows that 500 nm of a-Si:H absorbs essentially all incident photons with energies greater than 2 eV, and passes photons with smaller energies. In practice, the thickness of the top *pin* junction is adjusted so that it filters out about half of the photons that would otherwise have been absorbed in the bottom *pin* junction.[5] Since the photons that are absorbed in the top junction have relatively large energies, we can use a material with a relatively large bandgap as the absorber for this junction, and we shall obtain a larger open-circuit voltage across the top junction than that across the bottom junction. This is the

[5] In this chapter, we discuss only "two-terminal" multijunction cells in which the same electrical current flows through the series-connected cells. See Chapter 8 for further discussion of two-, three- and four-terminal multijunction cell operation.

"spectrum-splitting" effect. Note that a multiple-bandgap tandem cell does not absorb more photons than the bottom cell would on its own; rather, the top cell absorbs its share photons and converts them to electrons with a higher energy. Thus, there is more incident ***power*** converted, not more incident photons absorbed. The physics of multiple junction devices is also discussed in Chapters 4 and 8 of this Handbook.

For specificity, consider a tandem cell with a 1.55 eV bandgap bottom junction and a 1.80 eV bandgap top junction. The bottom cell will have $V_{OC} = 0.65$ V while the top cell will have $V_{OC} = 0.90$ V. In the absence of the top junction, an optimized 1.55-eV junction might deliver about $J_{SC} = 20\,\text{mA/cm}^2$. Assuming a fill factor (FF) of 0.7, the power output will be $9.1\,\text{mW/m}^2$. When assembled in tandem, the current through each junction is about half, or $10\,\text{mA/cm}^2$. However, the open-circuit voltage will more than double ($V_{OC} = 0.65 + 0.90 = 1.55$ V) so the power output rises to $11.2\,\text{mW/m}^2$ – for a 19% improvement over the single-junction 1.55 eV device.

We can distinguish three reasons for improved efficiency in a-Si-based multijunction cells over single-junction cells. The first is the spectrum-splitting effect we have just described. The second is a "junction thinning" effect: the uppermost junction in a well-optimized multijunction cell is thinner than in single-junction cells, which means that their fill factors are better. This effect is exploited in "same bandgap" tandem cells that use identical materials for the first and second junctions [149]. Third, a multijunction cell delivers its power at a higher operating voltage and lower operating current than a single-junction cell; the lower current reduces resistive losses as the current flows away from the junctions and into its load.

On the other hand, it is more challenging to fabricate multiple-junction solar cells than single-junction cells for at least 3 reasons: (1) The performance of a multijunction cell is more sensitive to the spectrum of the incident light due to the spectrum-splitting feature. This makes it even more critical to control the band gaps and thicknesses of the individual layers. (2) Deposition of the low-bandgap bottom cell(s) are more problematic whether they are a-SiGe or nc-Si absorbers. a-SiGe alloys require germane (GeH_4) gas which is several times more expensive than silane and highly toxic. Manufacturers need to implement even stricter safety procedures to handle GeH_4 than the other hydride gases used for a-Si deposition such as SiH_4 and PH_3. Alternatively, nc-Si absorbers need to be 5–10 times thicker than a a-SiGe absorber, so deposition time and rate become through-put limiting factors. (3) Multijunction cells require formation of an ohmic yet reverse-biased transparent n/p interconnect 'tunnel' junction between them.

One potential issue with multi-bandgap modules is their sensitivity to variation of incident spectra as occurs over the day and from day to day as air mass, turbidity and solar angle changes. Since the bandgaps and thicknesses are generally optimized for air mass 1.5 global (AM1.5G) illumination spectra, changing the relative ratio of blue to green to red light over the day or seasons would change the relative balance between the photogeneration in each junction. However, both theoretical [188]) and empirical studies [189], the latter conducted in two very different climates, indicate that multibandgap multijunction a-Si based cells perform comparable to or better than single junction a-Si modules under varying spectra and intensity. A recent study of the commercially important micromorph tandem modules found that while the annual output was affected by the spectrum, the performance peaked in the same average photon energy as was experimentally observed over the year and which was different from AM1.5 [190].

Overall, the advantages and benefits of higher stabilized output power for multiple-junction cells do outweigh the difficulties in the fabrication and sensitivity to spectral irradiance. We first discuss the a-SiGe based multijunction cells, and then we discuss micromorph cells.

12.5.2 Using Alloys to Vary the Band Gap

The bandgap of a-SiGe alloys can be continuously adjusted between 1.7 and 1.1 eV by varying the percentage of Ge. Calculations for idealized materials indicate that a bandgap near 1.2 eV would provide the maximum efficiency (see Figure 8.9c); unfortunately, the optoelectronic quality (i.e. defects and carrier transport) of a-SiGe degrades rapidly when the a-SiGe bandgap is reduced below about 1.4 eV, and these materials have not proven useful for PV application.

Figure 12.21 shows the $J-V$ characteristics of a series of a-SiGe solar cells with different Ge concentrations in the i-layer (of constant thickness, and without a back reflector) [191]; the bandgaps are indicated in the legend. As the bandgap is reduced by incorporating more Ge in an i-layer, V_{OC} goes down, as was previously illustrated in Figure 12.18. Since the narrower-bandgap materials catch more sunlight, J_{SC} increases. The fill factors of the cells also decrease as the bandgap decreases; this could reflect either decline in the hole drift-mobility or an increase in defect density; of course, the greatest interest is in these effects for the "light-soaked" state.

Similar to the deposition of a-Si, a-SiGe devices made with high hydrogen dilution show improved quality and light stability [192]. Optoelectronic properties (mobility, lifetime, collection length) of narrow-bandgap a-SiGe material are nonetheless inferior to those of a-Si.

12.5.2.1 Bandgap grading of a-SiGe i-layers

To enhance the fill factor of cells with a-SiGe i-layers, band gap *grading* is used to enhance the collection of holes [193, 194]. An asymmetric V-shaped bandgap profile is created by adjusting the Ge content across the i-layer. Wider-bandgap material is deposited closest to the n- and p-layers. The plane of narrowest bandgap lies much closer to the p-layer (through which the photons enter the device) than the n-layer. Such a grading scheme allows more light to be absorbed near the p-layer so that "slower" holes do not have to travel far to get collected (see Figure 12.16). Also, the

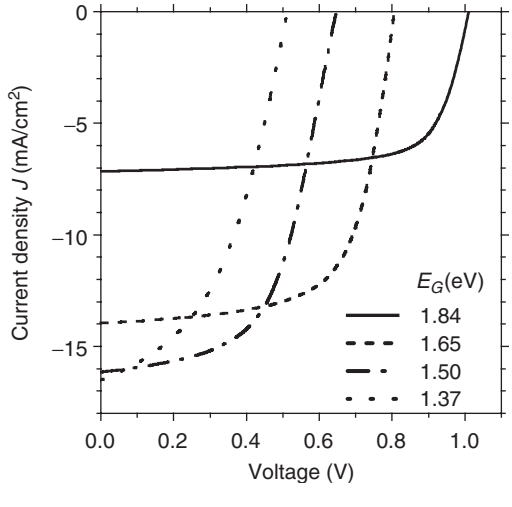


Figure 12.21 Performance of a-Si and a-SiGe *nip* solar cells with different Ge concentrations in the i-layer; the i-layer bandgaps are indicated in the legend. The fill factors for these cells are 0.70, 0.62, 0.55, and 0.43 for the cells with i-layer bandgaps of 1.84, 1.65, 1.50, and 1.37 eV, respectively [191]

tilting of the valence band creates a potential gradient or electric field that assists holes generated in the middle or near the *n*-side of the *i*-layer to move toward the *p*-layer. With appropriate hydrogen dilution during growth and bandgap grading, a-SiGe cells can generate up to 27 mA/cm^2 under standard AM1.5 light although it is more typical to evaluate them under filtered light to duplicate the illumination they see in the middle or bottom of a triple junction.

12.5.2.2 a-SiC alloys for high-bandgap absorbers

The bandgap of a-SiC can be adjusted between 1.7 and 2.2 eV, depending mainly on the C concentration [195]. After extensive research in the early 1990s, most workers decided that a-SiC is not suitable for use as the *i*-layer of the uppermost cell in a multijunction structure. a-SiC material with enough C incorporated to increase the bandgap appreciably over a-Si suffered significant decrease in electron drift mobility [43] and was highly defective especially after light-soaking [196].

12.5.3 a-Si/a-SiGe Tandem and a-Si/a-SiGe/a-SiGe Triple-junction Solar Cells

Dual-junction a-Si/a-Si (same bandgap tandem) solar cells have lower material cost (no GeH_4) than tandem cells using a-SiGe, and have slightly higher efficiencies (0.5–1% absolute) than a-Si single junction cells. Same bandgap a-Si/a-Si tandems have been in production for decades [197]. Multibandgap dual junction (a-Si/a-SiGe [92] or a-Si/nc-Si [201, 256]) and triple-junction (a-Si/a-SiGe/a-SiGe or a-Si/a-SiGe/nc-Si) solar cells use a spectrum-splitting approach to collect the sunlight. They achieve higher conversion efficiencies, commonly over 10% stabilized for small area <1 cm^2 cells [92, 100, 106, 120]. Some additional details and references may be found later in Table 12.4. While all-amorphous a-Si(1.8 eV)/a-SiGe(1.6 eV)/a-SiGe(1.4 eV) triple-junction solar cells have produced the most efficient a-Si-based cells today [8], triple junctions based on a-Si/a-SiGe/nc-Si or a-Si/nc-Si/nc-Si (no Ge) have produced nearly equivalent stabilized efficiencies [100, 120, 198]. Figure 12.20c,d show the two flavors of triple-junction substrate cells grown on SS foil, and an indication of the relative difference in thickness. In all cases, light enters from the *p*-layer so that holes need to travel less distance to get collected than electrons. In the following, we will briefly describe the two designs and typical deposition processes that are most broadly used today.

For *nip* triple junction cells deposited on an SS substrate, a reflective metal layer is deposited first on the substrate by sputtering or evaporation, followed by the sputter deposition of a ZnO buffer layer. Usually, silver is used as the reflective layer for research cells because of its high reflectivity, whereas aluminum is used in production because of difficulties with production yield for silver. The metal layer is deposited at high temperature (300–400 °C); self-segregation in the metal film forms the texture needed for light trapping. The sample is then moved into a RF PECVD deposition system for the deposition of semiconductor layers. The bottom *nip* with an a-SiGe *i*-layer (1.4–1.5 eV bandgap) is deposited first. A second a-SiGe middle cell (1.6–1.65 eV *i*-layer bandgap) is then added. Finally, the top a-Si cell (1.8–1.85 eV *i*-layer band gap) is added; the intrinsic layer is made using high H dilution at relatively low temperature. An indium–tin-oxide (ITO) layer is deposited on top via evaporation or sputtering. This layer is approximately 70 nm thick and serves as both the top electrode and an antireflection coating. Metal grids are evaporated or sputter-deposited on top of ITO to further reduce contact resistance.

For *pin* superstrate double junction cells deposited on glass, the glass substrate is first coated with a textured transparent conducting oxide, usually SnO_2 or ZnO, using one of the several methods such as atmospheric pressure (APCVD) [199, 200] or low-pressure (LPCVD) chemical vapor deposition. More details about textured TCO's are given in Chapter 17 of this Handbook. A

Table 12.4 Efficiency of small-area solar cells fabricated in different laboratories. Different stabilization approaches and time have been used. While not all results have been confirmed by independent testing, we think they are reliable. The first two are single junction; all others are multi-junction. USO = United Solar Ovonic

Structure	Deposition rate (Å/s)	Initial η [%]	Stable η [%]	Organization	Reference
a-Si			10.0	Oerlikon	[209]
		11.2	9.5	U. Neuchâtel	[210]
nc-Si		10.1	10.1	Kaneka	[211]
			9.9	U. Neuchâtel	[212]
a-Si/a-SiGe	1	14.4	12.4	USO	[213]
		11.6	10.6	Sanyo	[214]
		11.6	10.6	BP Solar	[215]
a-Si/nc-Si		13.0	12.0	Kaneka	[216]
		13.5	11.8	USO	[217]
	7	>12.8	11.9	Oerlikon	[218]
		12.6	11.1	U. Neuchâtel	[219]
	4.5	12.0	11.4	Utrecht	[220]
a-Si/a-SiGe/a-SiGe	1	15.2	13.0	USO	[8]
	1	11.7	11.0	Fuji	[221]
		12.5	10.7	U. Toledo	[222]
a-SiC/a-SiGe/a-SiGe		11.4	10.2	Sharp	[223]
a-Si/nc-Si/nc-Si	5	14.1	13.2	USO	[224]
	20	13.0	11.5	Canon	[225]
		12.0	12.0	Kaneka	[226]
		10.9	10.6	Utrecht	[117]
a-Si/a-SiGe/nc-Si		14.3	13.3	USO	[227]
		12.4	11.0	U Toledo	[228]
		11.4	10.7	ECD	[229]

pin top cell having an a-Si *i*-layer is then deposited, followed by either the a-SiGe or nc-Si bottom bottom cell. The structure is finished with the deposition of a ZnO buffer layer and metal reflector in the back. More details of manufacturing *pin* and *nip* cells and modules are given in Section 12.6.

Note that for the past 15 years, United Solar Ovonic/ECD has been almost exclusively responsible for the evolution of the a-Si/a-SiGe/a-SiGe triple-junction cell. Utrecht University and Canon have been sole developers of the a-Si/nc-Si/nc-Si triple-junction cell. But there have been numerous groups worldwide working on the a-Si/nc-Si double-junction.

12.5.3.1 Current matching

In a triple-junction cell, the three component cells are stacked monolithically. Since these component cells are connected in series to form a two-terminal device, the cell with minimum current density

during operation will limit the total current of the triple-junction stack. Therefore, the current densities of each of the component cells need to be *matched* (made the same) at the maximum power point (J_{MP}) for each cell in sunlight.

The short-circuit currents J_{SC} of the component cells are only a rough guide to this matching. For an a-Si/a-SiGe/a-SiGe triple-junction cell, the bottom a-SiGe cell usually has the lowest *FF* and the top a-Si cell usually has the highest *FF*. So, each cell has a different ratio of J_{MP}/J_{SC} Therefore, the J_{SC} of the bottom cell needs to be slightly greater than the J_{SC} of the middle cell, which in turn needs to be slightly greater than the J_{SC} of the top cell. For an optimized triple-junction cell, the differences in J_{SC} between the bottom and the middle and between the middle and the top cells are each about $0.5-1\,mA/cm^2$. This intentional "mismatch" in the J_{SC} values which allows matching the J_{MP} values at cells at the operating point is accomplished by adjusting the bandgaps and thicknesses of the component cell *i*-layers. While adjusting for current matching, one must consider that the bottom cell benefits from the light enhancement from the back reflector, while the middle and top cells receive little benefit from the back reflector.

Current-matching in tandem micromorph cells is complicated by the fact that the bottom nc-Si cell absorbs much more light than the top a-Si cell. This can be alleviated somewhat by increasing the thickness of the top cell however this decreases the stabilized FF. A better approach is to deposit a highly transparent dielectric reflector such as ZnO or SiO_2 between the top and bottom cell to reflect more light back into the top junction hence reduce light in the bottom [201, 202]. This may allow the tandem to be "bottom cell" limited (the current from the bottom cell will be smaller) which improves stability dramatically since the bottom nc-Si cell is naturally stable with regards to light-soaking, but spectral sensitivity, manufacturability and cost of the additional sputtered ZnO chamber must be evaluated [203].

12.5.3.2 Tunnel junctions

Multijunction solar cell require *tunnel junctions* (also called shorting junctions) at the interfaces between adjacent *pin* cells. These n/p junctions connect one cell to another and are reverse-biased under normal operation; they must generate negligible V_{OC} and have negligible resistance and optical absorption [204].

One might think that they would have electrical properties like classic *pn* junction diodes. However, as was discussed in Section 12.2.6, dangling bonds are generated when doping is increased. Carriers that are trapped on defects on one side of the interface can move to traps on the other side simply by quantum mechanical tunneling. This process "short-circuits" electrical transport involving the conduction band and valence band states. For this reason, the doped layers at the tunnel junction, particularly the sublayers near the interface, are made with very high doping and even intentional defects. The large density of dangling bonds permits the efficient neutralization[6] (by tunneling) of holes from the cell below and electrons from the cell above. Doped nc-Si layers or other defective layers have been useful to enhance the recombination without adding absorption. Tunnel or shorting junctions can be studied by making *pnip* or *npin* devices [205] or by QE measurements in functioning tandems using red, blue or no bias light [206].

12.5.3.3 Multijunction I–V measurement and interpretation

Measuring the $I-V$ performance of a multiple-junction, spectrum-splitting solar cell requires paying close attention to the illumination spectrum (see Chapter 18 for a more detailed discussion).

[6] One can consider this as a neutralization process.

A triple-junction cell for which the *pin* component cells are current-matched under the standard AM1.5 global spectrum may show poorer performance under a different light source, for example, a tungsten lamp or a cloudy day. The triple-cell J_{SC} is usually close to the J_{SC} of the limiting component cell except when there is a large mismatch and the limiting cell has a very low fill factor. The V_{OC} of the triple cell is the sum of the V_{OC} values of the component cells (reduced by any photovoltages at the tunnel junctions). It should be noted here that the bottom component cell in a triple-stack generates only about one-third of the photocurrent that it would under full sunlight; therefore, its V_{OC} is slightly smaller (usually by $\sim$20 mV) than when it is exposed to the full sunlight. The middle cell will have about half the current that it would under the full sunlight. The fill factor of the triple cell depends sensitively on the fill factor of the limiting component cell and on the current mismatch among the component cells. A large mismatch leads to a higher triple-cell *FF*, while on the other hand it also leads to a lower triple-cell current.

12.5.3.4 Quantum efficiency measurements in multijunction cells

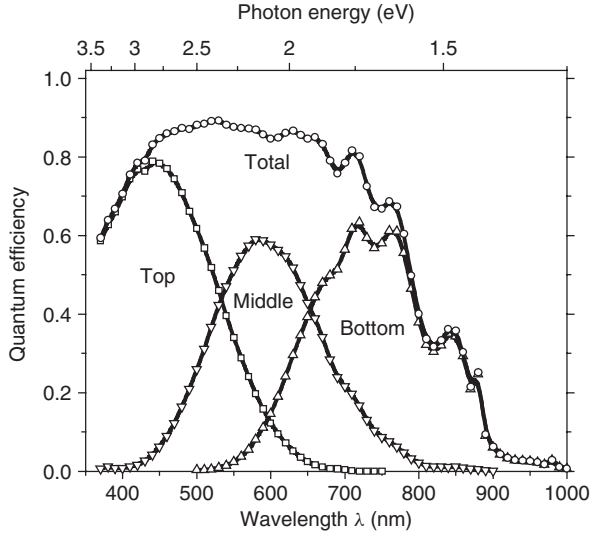
In measuring the *QE* of a triple-junction solar cell, appropriate light bias and electrical bias need to be applied during the *QE* measurements [207, 208] (see also Chapter 17). A simple *QE* measurement without these optical and electrical biases would yield a Λ-shaped curve, with response only from photons in the middle of the spectrum that were absorbed in all three junctions because a current can flow through the cell only if all of the component cells are illuminated simultaneously. This can be a useful diagnostic test [190].

When the *QE* of a specific component cell needs to be measured, say the middle cell, a DC bias light illuminates the triple cell through a filter that transmits only light to be absorbed in the top (blue light) and bottom (red light) cells. Under this condition, the middle cell current is limiting regarding the AC monochromatic light. Therefore, the AC photocurrent is due only to the the monochromatic light that is absorbed in the middle cell. This AC photocurrent, modulated by an optical chopper, can be easily detected using a lock-in amplifier (see Section 18.4).

The other two component cells can be measured in the same way, except with different optical filters for the bias light (blue–green bias light to measure the bottom cell and orange–red bias light to measure the top cell). When the cells are measured without externally applied electrical voltage bias, the component cell being measured is actually under reverse bias, equal to the sum of the open-circuit voltages V_{OC} of the other two component cells. The *QE* under reverse bias will be close to the *QE* under short-circuit condition only when the component cell *FF* is high. To measure the *QE* under short-circuit current condition, an electrical voltage must be externally applied to cancel out the voltage generated by the other component cells under the optical bias light (i.e. the sum of their V_{OC}'s). Figure 12.22 shows the *QE* curves of a triple-junction solar cell measured using this method [158]. The short-circuit current of component cells can be calculated by integrating the *QE* values with the AM1.5 light spectrum. The outer profile in Figure 12.22 is obtained by adding the three component cell *QE* curves.

The long-wavelength QE behavior for each of the component cells is determined by its *i*-layer thickness, hole collection length and bandgap. Additionally, for the bottom cell the back reflector is significant. However, the short-wavelength QEs for the middle and bottom component cells are largely determined by the thicknesses and bandgaps of the top and middle cells, respectively, since these component cells act as filters for the shorter-wavelength (higher-energy) photons. The short-wavelength behavior of *QE* for the top cell is sensitive to the absorption of ITO and top cell *p*-layer as well as the loss of electrons that are diffused back to the *p*-layer and get trapped.[7]

[7] More recent triple junction devices from USO have large cell currents (8.8/9.2/8.8 mA/cm^2) and integrated total currents of $>$26 mA/cm^2.



Short-Circuit Currents J_{SC} (mA/cm^2)	AM1.5	Xenon
Top	6.87	7.19
Middle	7.84	7.98
Bottom	8.62	7.34
Total	23.03	22.51

Figure 12.22 Quantum efficiency curves of component cells of a typical triple-junction solar cell. The table indicates the short-circuit current densities J_{SC} for the component cells measured for AM1.5 illumination and with a xenon illuminator [222]

12.5.3.5 High-efficiency multiple-junction solar cells

Table 12.4 lists some properties of multiple-junction solar cells ($<1\,cm^2$) fabricated in selected laboratories around the world. The degradation of multiple-junction solar cells containing only amorphous layers is usually in the range 10–20%, while the degradation of single-junction a-Si solar cells is usually in the range 20–40% (see Figure 12.15). Those with one or two nc-Si bottom cells degrade less (<5%) especially if they are bottom cell limited due to the inherent stability of the nc-Si. These percentages apply to the cell's properties after 1000 h of light soaking under 1 sun light intensity at 50 °C, which is the standard protocol used for gauging light degradation today. As we showed in Figure 12.4, the degradation of triple cells is smaller than in single junctions. As one can see from Table 12.4, the highest stabilized cell efficiency is 13.0% for a triple-junction device structure made at United Solar Ovonic. Table 12.4 also includes the best solar cells made using nc-Si as a component cell. The highest stable efficiency so far using a-Si/μc-Si micromorph tandem structure is 12% for a cell made at Kaneka.

12.5.4 Nanocrystalline Silicon (nc-Si) Solar Cells

Nanocrystalline Si (also known as microcrystalline silicon (μc-Si)) has been studied extensively for three decades and has been used for doped layers in a-Si [164] and nc-Si [230] solar cells. Because of the difficulties in passivating the defects located at the grain boundaries, nc-Si was not actively considered as an intrinsic layer in the *pin* or *nip* type solar cells until 1992, when Faraji *et al*. [231] and Meier *et al*. [232] reported the fabrication of nc-Si-based *pin* solar cells using VHF PECVD. Since then, nc-Si and even larger grain poly-Si thin film *pin* solar cells have been fabricated by a number of research groups [233].

A *QE* curve for such a nc-Si-based cell was presented in Figure 12.22. One can see that the QE of the nc-Si extends to longer wavelengths (>850 nm) than the a-Si and a-SiGe cells.

Table 12.5 Stabilized efficiency of large-area a-Si based prototype modules. These have not necessarily been verified by outside organizations and may not represent a standard commercial product. Modules larger than 0.7 m^2 are typically standard products

Structure	Stable η [%]	Size [m^2]	Company	Reference
a-Si	5.9	1.4	Ersol	[258]
a-Si	6.6	5.7	Signet	[259]
a-Si	6.3	1.5	Mitsubishi HI	[260]
a-Si/a-Si	6.9	1.4	Schott	[261]
a-Si/a-Si	6.0	0.94	EPV	[262]
a-Si/a-SiGe	9.3	0.52	Sanyo	[263]
a-Si/a-SiGe/a-SiGe	9.0	0.32	Fuji	[160]
a-Si/a-SiGe/a-SiGe	6.7	2.2	USO	[264]
a-Si/a-Si/a-SiGe	9.2	0.81	USO	[265]
a-Si/nc-Si	11.0	0.52	Sharp Solar	[106]
a-Si/nc-Si	10	1.4	Oerlikon	[266]
a-Si/nc-Si	8.0	1.4	Ersol	[267]
a-Si/nc-Si	9.0	1.4	Sharp	[268]
a-Si/nc-Si	9.2	1.4	Astronergy	[269]
a-Si/nc-Si	9.8	0.20	MHI	[105]
a-Si/nc-Si	10.0	0.37	Kaneka	[195]
a-Si/nc-Si	8.6	1.2	Kaneka	[270]
a-Si/nc-Si	8.0*	5.7	Appl. Mat.	[255]
a-Si/nc-Si/nc-Si	12.2	0.80	Canon	[100]

*At 444 W, the largest thin film module ever reported

Remarkably, the photocurrents generated from a ~1.5–2.0 μm thick nc-Si cell have exceeded 25 mA/cm^2 with suitable back reflectors [234]. Therefore, such cells are suitable for use as the bottom cell of a multijunction cell with a-Si-based cells as the top cell. The advantages of using nc-Si as the narrow-bandgap cell instead of a-SiGe are: (1) the higher *QE* in the long-wavelength region; (2) negligible light-induced degradation, even for *i*-layers with 30–50% a-Si fraction (the generation occurs in the c-Si which does not suffer from the Staebler–Wronski effect); (3) reduced materials cost, since nc-Si is deposited using SiH_4, which is a relatively low-cost gas compared with GeH_4; and (4) μc-Si cells can be made with higher *FF* than a-SiGe. On the other hand, the concerns associated with nc-Si compared with a-SiGe for the bottom cell are: (1) nc-Si cells require much thicker *i*-layers (1.0–2.0 μm vs 0.2–0.3 μm for a-SiGe) to absorb the sunlight; this is a consequence of the lower interband absorption coefficients in (indirect band gap) crystals compared with amorphous semiconductors; (2) much longer time is needed to complete the deposition of a thick μc-Si layer unless the deposition rate for μc-Si is increased by 10× compared to a-SiGe; and (3) nc-Si solar cells have lower V_{OC} (around 0.53 V) than a-SiGe cells yielding the same J_{SC}.

The three most common deposition methods for nc-Si are RF-PECVD (13.56 MHz), VHF-PECVD (40–100 MHz), and HWCVD as discussed in Section 12.3. It is widely understood that RF PECVD deposition of nc-Si requires high-power/high-pressure regime [96]. A typical deposition

rate for an a-SiGe *i*-layer is 3 Å/s (much beyond that the quality degrades); to complete a nc-Si cell with comparable deposition time, one would need to deposit μc-Si with at least ~20 to 30 Å/s deposition rate so that it would not be rate-limiting during production. Single junction *p-i-n* nc-Si devices with >9% efficiency have been made at >20 Å/s [235].

There are several differences in the deposition of optimized nc-Si compared with a-Si solar cells. First, the influence of substrate texture is significant [236]. The "gold standard" texture for a-Si is the Asahi Type U which is highly faceted, with sharp peaks and valleys. While a-Si is deposited very conformally over these sharp features, nc-Si tends not to form in the sharp valleys, leading to voids and defects. Thus, a more rounded "dimpled" texture is ideal for nc-Si [237]. This can be achieved by acid etching of specular sputtered ZnO [237, 238] or deposition of textured LPCVD ZnO [239]. A second difference is that nc-Si benefits from a graded H dilution profile. Low dilution ($R < 5$) initially gives an a-Si film, but will eventually lead to nc-Si formation if the film is thick enough, meaning the initial "incubation" layer will be several tens or hundreds of nanometers of a-Si. But the growth rate will be sufficiently high (see Figure 12.12). A high dilution ($R > 5$) will initiate nc-Si growth more quickly than low R, but the growth rate is much slower. So a two-step or graded profile is often used: first, a highly diluted plasma eliminates an initial a-Si "incubation" region, and after a few tens or hundreds of nanometers of nc-Si, the dilution is decreased such that the bulk of the nc-Si layer is grown at higher growth rate [240, 241]. This greatly increases the need for careful control and optimization of growth; the optimum profile may depend on the substrate material and texture, temperature, and total nc-Si film thickness.

12.5.5 Micromorph and Other nc-Si-Based Multijunction Cells

The Neuchâtel group [95, 99, 202, 203, 232] first used an a-Si *pin* junction as the top component cell and a nc-Si *pin* as the bottom component cell for a-Si/nc-Si tandem cells; they named these cells *micromorph* devices. Modeling by several groups for a range of PV technologies shows that 1.7 eV/1.1 eV bandgaps for the top/bottom cell provide a nearly ideal bandgap pair for tandem cells, regardless of their lifetime or crystallinity (see contour diagram in Figure 8.9b which is for high efficiency high-efficiency III–V tandems, or reference [242] for polycrystalline thin film tandems. By happy coincidence, this is exactly the bandgap relation for a-Si and nc-Si.

In order for an a-Si/nc-Si tandem cell to have comparable performance to an a-Si/a-SiGe cell, the bottom cell nc-Si must generate at least 26 mA/cm^2. (But cells which are bottom-cell limited have better stability since the nc-Si does not degrade.) Such a high current requires the nc-Si layer to be 1.0–2.0 μm thick and utilize advanced light enhancement schemes. In order to maintain current matching in a micromorph cell, the top a-Si component cell must generate 13 mA/cm^2 (i.e. half the current for a stand-alone μc-Si). In addition, this a-Si cell needs to be relatively stable under light exposure so that the tandem cell could be stable.

Two approaches have been taken. First, the a-Si *i*-layer can be deposited at a relatively higher temperature where the lower H concentration will lead to a reduced bandgap, ~1.65–1.70 eV. Secondly, a highly transparent but semireflective dielectric layer (like ZnO or SiO_2) can be inserted at the tunnel junction between the top and the bottom cell [202]. This semi-reflective layer permits current matching without increasing the thickness of the top cell, i.e. enhancing the top component cell current at the expense of the bottom cell. With these two approaches, 13 mA/cm^2 J_{SC} was obtained from the top cell with a 3000-Å-thick a-Si layer. Innovative approaches are needed to further increase the current beyond the present level. With the micromorph tandem design, solar cells with 11–12% stable efficiency have been fabricated [201, 203, 243].

One can also fabricate a-Si/μc-Si/μc-Si triple cells. Such a design would relax the stringent requirement on the a-Si top cell due to current matching since it now only needs to generate

one-third of the bottom cell current. However, the middle cell thickness will be much greater than the a-SiGe cell it replaced, leading to a total thickness exceeding 6 μm compared with less than 1 μm for an all-amorphous triple [120].

Still another approach to a triple-junction cell design is to combine a 1.8 eV a-Si top cell, a 1.6 eV a-SiGe middle cell, and a 1.1 eV μc-Si bottom cell [198]. Such a cell design would have the advantages of a thinner and more stable top cell than for a micromorph tandem cell, would have better long-wavelength collection, and would reduce consumption of (expensive) GeH_4 gas compared with an all-amorphous, a-Si/a-SiGe/a-SiGe triple-junction cell. Issues relating to photocurrent matching of the junctions via thickness or bandgap, stability, and deposition time have been discussed (Unisolar [198, 244], Utrecht [121]). United Solar Ovonic has achieved nearly identical *stabilized* efficiencies, 12.6–13.0%, for a-Si/a-SiGe/μc-Si triple cells as well as a-Si/μc-Si/μc-Si triple cells deposited at relatively high rates but the a-Si/a-SiGe/a-SiGe triples remain more manufacturable [198, 244]. However, Canon has gone into production with the a-Si/nc-Si/nc-Si triple [101] in which each cell generates 10–11 mA/cm^2, leading to a total current for triple stack of over 31 mA/cm^2. This is ~4 mA/cm^2 more current than for a-Si/a-SiGe/a-SiGe triple cells, although the total thickness is likely an order of magnitude larger as well. Amazingly, they found similar stabilized performance ~12.5% (estimated from 13.4% initial and 6% degradation) for nc-Si deposition at 10, 20 and 30 Å/s.

12.6 MODULE MANUFACTURING

During the past 10 years, there has been a rapid increase in the production of a-Si power modules. Production (not capacity) increased from 29 MW in 2003 when the first edition of this book was published to 270 MW in 2008, a compound annual growth of 56%. Still, the a-Si share of the world PV market only increased from 4 to 5% during that time. In 2008, United Solar Ovonic shipped 112 MW of triple-junction flexible modules while MHI, Kaneka, Sharp each shipped 40–60 MW of single-junction a-Si or micromorph glass modules [245]. Sharp has just opened a manufacturing facility for 1.1 × 1.4 m^2 tandem (a-Si/nc-Si) and/or triple junction (a-Si/nc-Si/nc-Si) modules, with existing 160 MW capacity and an additional 640 MW plant under construction, which would make it the largest a-Si PV plant in the world. One must be careful to distinguish capacity from actual production. Quite often facilities do not operate at full capacity, either because there is insufficient market, or because operating at full throughput has a negative influence on yield. Thus, we tend to avoid quoting present or future capacity figures from specific companies.

An exciting recent development is the number of complete "turnkey" fabrication lines which are in the process of being installed and having their micromorph modules qualified by outside organizations. Complete micromorph production lines in the 50–100 MW range are available from companies such as Oerlikon Solar, Applied Materials, Ulvac Solar and Leybold Optics. A commercially-based survey of over a dozen turnkey single-junction and micromorph production line manufacturers listing their production line capabilities (MW/yr), module size, throughput, performance, yield, capital cost, etc has been published [246]. Oerlikon claims 95 and 140 W in production for their 1.4 m^2 a-Si and a-Si/nc-Si modules, or 6.7 and 10% efficiency, respectively. Applied Materials has announced a stabilized efficiency exceeding 8% on a "champion" 5.7 m^2 module [255]. Both Oerlikon and Applied Materials report to have sold over ten of their production lines, mostly in Asia and Europe. This concept sets up a model similar to the IC fab lines, and is quite different from CdTe or CIGS where each company's process equipment is unique. It remains to be seen if this model is economically valid for PV.

To transform small-area R&D developments into any type of large-scale manufacturing, key issues including uniform deposition over large areas, process gas utilization, deposition rate,

production throughput, process reproducibility, machine maintainability and serviceability, process automation, and production yield must be addressed [247].

For a large-scale production line, both in-line and batch plasma processing have been used by major manufacturers. In the following, we use the production process at United Solar Ovonic as the example for the substrate-type *nip* cell in-line process and that at EPV as the example for the superstrate-type *pin* cell batch process.

12.6.1 Continuous Roll-to-roll Manufacturing on Stainless Steel Substrates[8]

The continuous, "roll-to-roll" a-Si PV manufacturing process was developed by Energy Conversion Devices, Inc. (ECD) [248] and has been used by ECD's PV joint ventures and partners (United Solar Ovonic, Sovlux, and Canon) [249] as well as Power Film. Roll-to-roll refers to the process whereby a roll (also called "web") of flexible SS or polymer is unrolled and fed into the manufacturing process, and then rolled up after a manufacturing step has been completed.

The front-end process consists of four continuous, roll-to-roll steps in separate machines: (1) substrate washing; (2) sputter deposition of the back reflector; (3) a-Si semiconductor deposition; and (4) ITO top electrode deposition. Rolls of magnetic SS web, typically 125 μm thick, 0.35 m wide, and currently 1500 m long, are unwound from a modular "pay-off" chamber on one side and wound up in a modular "take-up" chamber on the other side, guided by magnetic rollers.

In the roll-to-roll washing machine, the SS web is guided through ultrasonic detergent cleaning stations with spinning brushes rubbing the surface, multiple deionized water rinse baths, and an infrared drying chamber. An oil-free, particle-free, clean SS roll is then wound up with protective interleaf.

The roll is then loaded into the pay-off chamber of the back-reflector sputter machine then pulled through several DC magnetron sputter deposition zones with metal targets (Al, Ag, or other alloys) for the reflective layer and ZnO targets for the deposition of ZnO buffer layer. The substrate is maintained at elevated temperature during sputtering so that the metal films develop a texture useful for optical enhancement [250].

The roll is then loaded into the RF PECVD machine for the continuous deposition of nine a-Si based semiconductor layers (*nip/nip/nip*) as well as all of the buffer layers on both sides of a-SiGe absorber layers. The deposition of the different layers occurs sequentially, but in a single pass. Innovative laminar flow "gas gates" isolate the feedstock gases in different chambers to prevent cross-contamination, while at the same time allowing the web to pass through the sequence of chambers continuously.

After the semiconductor deposition, the roll is then loaded into the TCO deposition machine, which uses either reactive evaporation of indium in oxygen ambient or sputtering from an ITO target in an argon atmosphere. The thickness of the ITO is carefully monitored to achieve optimum antireflection properties.

It would be difficult to integrate the four roll-to-roll steps into one machine due to the different pressure ranges for the four machines: atmospheric pressure for the washing, a few mTorr for back-reflector sputtering, around 1 Torr for PECVD, and a few mTorr for TCO sputtering.

[8] The first continuous roll-to-roll solar cell deposition process was for the Cu_2S cell on Cu foil as demonstrated in the early 1980s at the Institute of Energy Conversion at University of Delaware (see US Patent 4,318,938 issued March 9, 1982).

At this point, the SS roll is a giant solar cell, 1500 m long, which needs to be converted into many smaller series connected cells to get higher voltage for the modules. Thus, the roll of TCO coated a-Si is cut into cells or "slabs" of selected sizes. The standard slabs then go through a shunt passivation process in an electrolyte bath to remove and isolate small shunts by converting the TCO at the shunting point into an insulator [188, 251]. The grids, either carbon paste or copper wire coated with carbon paste, are then applied to the slab to complete the large area triple junction cell, which generates ~2.3 V voltage and ~3.2 A current for a 400 cm^2 module. The cells are then connected together in series with the grids/bus bar of one cell connected to SS substrate (the opposite electrode) of the neighboring strip cell, like roofing shingles. Bypass diodes are also installed for protection. The connected cells are then encapsulated with transparent ethylene vinyl acetate (EVA) and Tefzel layers and cured in an oven. After framing, if needed, the module is tested under standard test conditions (STC).

Alternatively, Power Film (formerly Iowa Thin Films, Inc.) deposits an a-Si tandem on rolls of flexible plastic (Kapton^{TM}) substrate. The cell interconnect is achieved by laser scribing.

12.6.2 a-Si Module Production on Glass Superstrates

The manufacturing of a-Si solar panels on glass superstrates is being developed by several companies including Energy Photovoltaics Solar (EPV) [252], Mitsubishi Heavy Industry (MHI) [106], Sharp [106], and Schott Solar [253]. New turnkey "fab lines" for production of single junction a-Si or tandem micromorph modules are being sold by Applied Materials (AMAT) [254, 255] and Oerlikon (previously developed by Uniaxis) [99, 256]. They all utilize monolithic integration via three laser scribes to produce a module containing ~100 series-connected cells.

The process begins with large sheets of "float" glass, 3 mm thick, with sizes of 1.2 × 0.6, 1.4 × 1.0, 2.0 × 2.2, or 2.2 × 2.6 m. A textured tin oxide TCO layer is deposited using an APCVD process either at the glass supplier's plant or at the PV plant. Alternatively, smooth ZnO can be sputtered and wet-etched for texture, or textured ZnO deposited directly by LPCVD to provide an textured TCO on glass commonly used for micromorph modules [238]. The substrate is edge-polished and cleaned then silver frits are applied as busbars and cured in a belt furnace. The textured TCO layer is "scribed" by a laser into strips about 9 mm wide. The substrates are then loaded in the PECVD machine for the deposition of the six semiconductor layers for an a-Si/a-Si or a-Si/nc-Si *pin*/*pin* tandem structure. Some systems hold many substrates in parallel (batch) and some feed them (vertically) through a series of load-locked plasma chambers. Etching the systems with a reactive gas such as NF_3 is critical to avoid build-up and cross-contamination [254]. The semiconductor deposition is followed by the sputter or LPCVD deposition of a much thinner ZnO buffer layer. The next laser scribe line is applied adjacent to the first scribe line. This second scribing is done at a lower laser power so that the ZnO and a-Si layers are scribed while the underlying textured TCO layer remains intact. An aluminum layer is sputter-deposited as the back reflector and back contact. A third scribing of the Al adjacent to the second completes the interconnection of neighboring cells in series, as shown in Figure 12.23. A fourth, high-power laser scribing around the perimeter of the solar panel isolates the active area from the edges. The modules are then encapsulated with EVA and either a second piece of glass or a protective polymer backing sheet.

12.6.3 Manufacturing Cost, Safety, and Other Issues

An important aspect of any manufacturing process is cost, which consists mostly of raw materials, labor, capital depreciation of the machines, and administration. The overall production cost per unit of product is reduced as the production volume goes up. At time of publication, capital expense costs

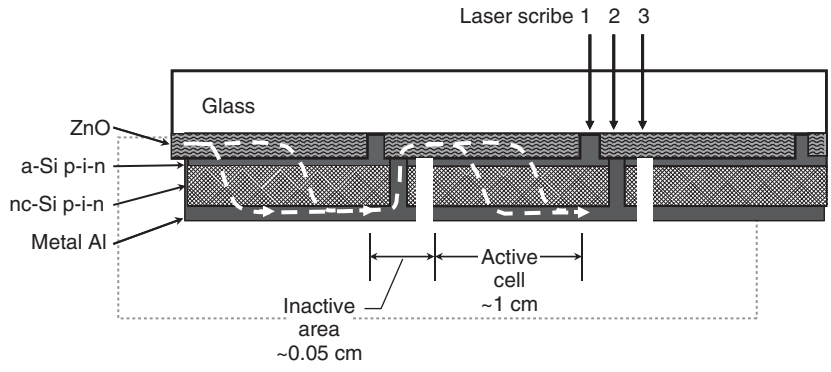


Figure 12.23 Cell interconnection (monolithic integration) of superstrate-type solar cells showing the three laser scribes. See Plate 3 for the colour figure

for these 50–100 MW production lines is $2.00–3.00 (USD)/W [245]. While actual production costs are closely guarded, an industry analysis group has estimated production costs of $US 1.10–1.50/W, but Oerlikon claims it is already under 1.0 (USD)/W and will be at 0.70 (USD)/W in 2011 (Greentech Media on-line September 7, 2010). Several studies have predicted that at a high production volume, perhaps 100 MW_p/year, the cost is expected to be lower than $1/$W_p$ [247]. Currently, the major materials costs are the module framing, encapsulation, and the substrates (glass or SS).

Another important aspect with regard to a-Si PV manufacturing is the plant safety. Although there is no toxic material in the final product, the manufacturing processes involve toxic and/or pyrophoric gases such as germane, phosphine, trimethylboron, silane, hydrogen, and so on. Amorphous Si PV manufacturers adopted many of the safety procedures developed by the integrated circuit industry to improve workplace safety [257]. Toxic gases are diluted in hydrogen or silane from 1 to 20%. Gas cylinders are installed outside the building or in fireproof gas cabinets. Toxic gas monitors are installed throughout the plants. Automatic gas isolation and operation shutdowns are implemented.

12.6.4 Module Performance and Reliability

The solar conversion efficiency of all thin film production modules is much lower than for the research and development (R&D) scale, small-area solar cells since production processes are constrained by cost and throughput considerations [247]. The differences in efficiency are mostly from the TCO and glass quality (cost-driven), semiconductor material quality, deposition uniformity (throughput driven), encapsulation loss, busbar shadow loss and electrical loss, and small shunts. The efficiency differences between R&D and manufacturing have been analyzed [271, 272]. The two biggest loss factors are TCO substrate optical quality and need for a simpler, more robust, but less optimum process for manufacturing.

Two aspects of PV modules are generally evaluated and verified by third parties: maximum solar conversion efficiency or output power under STC, and long-term reliability under specific environmental conditions. Actually, instead of maximum Watts under STC, it is the annualized energy yield in kW h/kW that is of more importance, but this requires more extensive testing and monitoring. Table 12.5 lists the efficiency of a-Si based PV modules produced by selected organizations around the world. Despite significant differences in device structure (single-junction a-Si, tandem junction a-Si, or triple-junction a-Si/a-SiGe/a-SiGe) and in substrate (glass or stainless steel), the commercial modules typically have stabilized efficiencies in the range 5.5–6.5%.

The exception is tandem a-Si/nc-Si modules which have around 8–10%, but they have only become available as of 2009 (Sharp, Kaneka, and Astronergy). In 2008, there were about 15 manufacturers with a-Si modules on the market with over 400 models, although many are just sourced from the original manufacturer and re-branded [273].

Although the stabilized conversion efficiency values of a-Si-based modules are lower than those of most other commercially available solar cells, they are less temperature sensitive due to a smaller temperature coefficient; i.e. −0.2%/C for a-Si versus −0.4–0.5%/C for c-Si. Consequently, a-Si PV products have demonstrated a 5–15% higher annual energy yield (kW h produced per kW installed) than c-Si PV products in many locations around the world [247].

Like all PV products, a-Si modules must pass reliability and qualification testing as described in Chapter 18 of this Handbook. IEC 61646 [274] documents the relevant qualification testing for thin film modules. These tests include repeated thermal cycling between −40 and 90 °C, humidity freeze cycles, hail impact, high voltage bias under wet spray, and others. Long-term large-scale outdoor deployment of a-Si modules has been successfully demonstrated in several locations [201, 275, 276]. Degradation rates of ∼1% relative per year are similar to those of c-Si modules. Thin film Si modules are typically warranted to maintain 80% of the rated output after 20 or even 25 years, the same as c-Si, multi-Si and CdTe modules.

12.7 CONCLUSIONS AND FUTURE PROJECTIONS

12.7.1 Advantages of a-Si-Based Photovoltaics

Amorphous Si absorbs sunlight more strongly than c-Si; a thin layer (0.3 μm) of a-Si is sufficient to absorb >90% of the available sunlight. Its bandgap can be varied with incorporation of Ge or by forming nc-Si, thus allowing straightforward fabrication of multijunction devices with a single deposition process. Amorphous Si PV products can be deposited at a low temperature on inexpensive substrates with a low-cost continuous or batch process, are environmentally friendly, contain no heavy metals (like Cd) or rare elements (like In and Te), and use a small fraction of the Si required for c-Si solar cells. No material-supply limitations have been identified with the TW scale implementation of a-Si [277]. The energy payback time (the time required for an a-Si module to generate the energy used in its production) is among the very shortest for any PV technology, variously estimated at 1–1.5 years. Even though it has been in commercial production for approximately 20 years and the production process is more mature and proven, the cost is expected to decline as the production volume is increased.

When deposited on glass, a-Si modules can be used for power production or building-integrated applications. When deposited on lightweight and flexible substrates such as stainless steel or plastic, they are highly desirable for portable power and other building-integrated applications. They have a high radiation tolerance which is important for space power applications.

12.7.2 Status and Competitiveness of a-Si Photovoltaics

Over the last quarter of a century, significant progress has been made in the understanding of properties and of deposition processes for a-Si-based materials and solar cells. In 1997, a-Si-based solar cells with 15.2% initial efficiency and 13% stable efficiency were demonstrated [8]. The manufacturing volume of a-Si solar modules has increased more than tenfold over the past 10 years, and capacity is presently more than 270 MW_p/year. Presently there are manufacturers with

decades of experience, such as United Solar Ovonic, Sharp, Mitsubishi, and EPV as well as new ones emerging with huge new automated module fabrication lines made by Oerlikon Solar, Applied Materials, Ulvac Solar and Leybold Optics. Applied Materials "fab" lines make modules on 5.7 m^2 sheets of glass with 45 MW (single junction a-Si) or 80 MW (micromorph tandem) expected annual output. Despite its low efficiency, some industry experts predict that a-Si will surpass CdTe as the lowest-cost thin film PV technology partially on the strength of these (unproven) new fab lines [278]. But all such market predictions are speculation at this point. As Figure 1.7 in Chapter 1 points out, the decrease in LCOE is rather steep for low-efficiency modules, and the LCOE is much more sensitive to efficiency than cost. Thus, most experts agree that modules under 8% efficiency have no future regardless of how low their cost. Hence, a-Si manufacturing must achieve higher module performance. Multijunction devices with nc-Si would seem to be the *sine qua non* of future production.

Significant progress has been made in the development of faster deposition processes (>5 Å/s) that achieve essentially the same quality as the present slow processes, as discussed in Section 12.3. As rapid-deposition and high-gas-utilization processes are incorporated into production, further cost reduction will be achieved.

Additionally, the use of nanocrystalline silicon as the narrow-bandgap absorber layer in an a-Si-based tandem solar cell has been demonstrated, with cells exceeding 12% conversion efficiency (stabilized) reported by several labs and large-area modules with stabilized efficiencies exceeding 8% reported by several manufacturers, including 1.4 m2 commercial scale modules with 10% stabilized efficiency claimed by Sharp and Oerlikon. Cells incorporating nc-Si show superior light stability over extended light-soaking; in fact they have negligible degradation if they are bottom-cell limited. Finally, there have been analyses showing different pathways to reach a stable 15% tandem device [279, 280].

12.7.3 Critical Issues for Further Enhancement and Future Potential

To increase application of a-Si-based PV significantly beyond today's level, the following critical issues must be addressed.

1. Light-induced degradation must be better understood. Approaches for reducing or controlling the degradation need to be further developed. At this moment, there are many engineering compromises in the device design, such as the use of thin *i*-layers to limit the degradation. If the materials can be made more stable under light, these compromises can be relaxed and the device can be made with much higher efficiency.
2. As the gross defects associated with light soaking are minimized, the next performance limitation to be addressed is improving the drift mobility of holes.
3. We need to improve a-SiGe so that narrower-bandgap materials can be incorporated into cells and more of the infrared region of the solar spectrum can be exploited.
4. Faster deposition processes need to be developed that (at least) preserve the conversion efficiencies achieved by present processes. This is critical for low-cost and high-throughput manufacturing. In addition, these high-rate processes must also achieve high gas utilization.
5. Nanocrystalline Si-based solar cells need to be fully explored as narrow-bandgap component cells in tandem or triple-junction cells as an alternative to a-SiGe. But much faster deposition processes, >20 Å/s, will be required. Methods for improving the open-circuit voltage to >0.60 V needs to be better understood.
6. Transparent conductors are needed with better light trapping and better templating for nc-Si growth. Optical performance includes less internal parasitic absorption and texture that allows high J_{SC} without sacrificing V_{OC} due to shunting. Resistance to plasma damage for high power H-diluted nc-Si deposition is valuable. Plasmonic back reflectors are promising.

7. Module design needs to be further improved and the costs associated with framing and encapsulation need to be further reduced. At the same time the durability of modules in standard environmental tests must be preserved or improved.
8. We need to find new applications for a-Si PV products in all of its present markets, including building-integrated PV, space power, and consumer electronics as well as grid-connected, large-scale power generation.

ACKNOWLEDGEMENTS

We thank Rana Biswas (Iowa State University), Nerio Cereghetti (LEEE), Gautam Ganguly (BP Solar, Inc.), Subhendu Guha, Baojie Yan, Jeff Yang, and Guozhen Yue (United Solar Ovonic LLC), Scott Jones and Stanford Ovshinsky (Energy Conversion Devices), Bolko von Roedern (National Renewable Energy Laboratory), Chris Wronski (Pennsylvania State University), Brent Zhu, Brad Culver and Ujjwal Das (Institute for Energy Conversion) for their generous help with this article. Additionally, all three co-authors benefitted greatly from the open discussion and collaborations with our colleagues from the US National Renewable Energy Laboratory supported a-Si Industry/University/National Lab Research Teams from 1992 to 2004.

REFERENCES

1. Williams E M, *The Physics and Technology of Xerographic Processes*, John Wiley & Sons, Inc., New York, (1984).
2. Mort J, *The Anatomy of Xerography: Its Invention and Evolution*, McFarland, Jefferson, N.C. (1989).
3. Perlin J, *Space to Earth: The Story of Solar Electricity* (aatec publications, Ann Arbor, 1999).
4. An historical discussion is given by Chittick R C and Sterling H F, in *Tetrahedrally Bonded Amorphous Semiconductors*, edited by Adler D and Fritzsche H, Plenum, p. 1 New York (1985).
5. Spear W E, LeComber P G, *Solid State Comm*. **17**, 1193 (1975).
6. Carlson D E, Wronski C R, *Appl. Phys. Lett*. **28**, 671–673 (1976).
7. Wronski C R, Carlson D E, Amorphous Silicon Solar Cells, in *Clean Electricity from Photovoltaics*, Archer M D and Hill R (eds), World Scientific (2001).
8. Yang J, Banerjee A, Guha S, *Appl. Phys. Lett*. **70**, 2977–2979 (1997).
9. Yoshimi M, Sasaki T, Sawada T, Suezaki T, Meguro T, Matsuda T, Santo K, Wadano K, Ichikawa M, Nakajima A, Yamamoto K, *Conf. Record, 3rd World Conference on Photovoltaic Energy Conversion*, pp 2789–2792 (Osaka 2003).
10. Fritzsche H, in *Amorphous and Heterogeneous Silicon Thin Films*, Collins R W, *et al*. (eds), Materials Research Society, Symposium Proceedings Vol. 609, p. A17.1 Warrendale (2001).
11. Vaneček M A, Poruba A, Remeš Z, Beck N, Nesládek M, *J. Non-Cryst. Solids* **227–230**, 967 (1998).
12. The figure was calculated based on the hemispherical irradiance (37 ° south-facing) American Society for Testing and Materials (ASTM) Table G159-98 Standard Tables for References Solar Spectral Irradiance at Air Mass 1.5: Direct Normal and Hemispherical for a 37 ° Tilted Surface.
13. Carasco F and Spear W E, *Phil. Mag. B* **47**, 495 (1983). Near room temperature, these authors reported that a-Si:H has a "quantum efficiency" of essentially 1.00 for generating photocarriers when a photon is absorbed. This ideal value is rather surprising. Many other noncrystalline materials have "geminate recombination" of the electron and hole immediately after their generation, which would of course lead to a loss of conversion efficiency; see reference [14].
14. Schiff E A, *J. Non-Cryst. Solids* **190**, 1 (1995).

15. The spectrum splitting multi-bandgap a-Si/a-SiGe tandem was proposed by Vik Dalal and Ed Fagen of the Institute of Energy Conversion: Dalal V, Fagen E, *Proc 14th IEEE Photovoltaic Spec Conf* (1980) 1066; and also patented by Dalal #4,387,265 (1983).
16. Guha S, in *Technology and Applications of Amorphous Silicon*, 252–305, Street R A (ed), Springer, Berlin (1999). Fig. 6.10 of this paper is a valuable compilation of power measurements for varying cell thicknesses and light-soaking histories.
17. Guha S, Yang J, Banerjee A, Glatfelter T, Hoffman K, Xu X, *Technical Digest - 7th International Photovoltaic Science and Engineering Conference PVSEC-7*, p. 43 (Nagoya 1993).
18. Staebler D L and Wronski C R, *Appl. Phys. Lett*. **31**, 292 (1977).
19. Gregg A, Blieden R, Chang A, Ng Herman, *Proceedings 31st IEEE Photovoltaic Specialists Conference*, 1615, (Orlando, 2005).
20. Measurements furnished through the courtesy of N. Cereghetti, Laboratory of Energy, Ecology and Economy (LEEE), Scuola Universitaria Professionale della Svizzera Italiana. These data apply to the 0.5 kW array, and are described in more detail by Cereghetti N, Chianese D, Rezzonico S, Travaglini G, *Proceedings of the 16th European Photovoltaic Solar Energy Conference*, James and James, London (2001).
21. Emery K, Burdick J, Calyem Y, Dunlavy D, Field H, Kroposki B, Moriary T, Ottoson L, Rummel S, Strand T, Wanlass M W, *Proceedings 25th IEEE Photovoltaic Specialists Conference*, p. 1275, (Washington DC, 1996).
22. Kameda M, Sakai S, Isomura M, Sayama K, Hishikawa Y, Matsumi S, Haku H, Wakisaka K, Tanaka K, Kiyama S, Tsuda S, Nakano S, *Proceedings 25th IEEE Photovoltaic Specialists Conference*, p. 1049 (Washington 1996).
23. del Cueto J A and von Roedern B, *Prog. Photovolt: Res. Appl*. **7**, 101 (1999).
24. Carlson D E, Lin G, and Ganguly G, in *Proceedings of the 28th Photovoltaic Specialists Conference*, p. 707, IEEE (2000).
25. Street R A, *Hydrogenated Amorphous Silicon*, Cambridge University Press, Cambridge (1991).
26. Phillips J C, *J. Non-Cryst. Solids* **34**, 153 (1979).
27. Boolchand P, Lucovsky G, Phillips J C, and Thorpe M F, *Phil. Mag*. **B85**, 3823 (2005).
28. Reimer J A and Petrich M A, in *Amorphous Silicon and Related Materials*, Vol. A, 3, Fritzsche H (ed.), World Scientific, Singapore (1989).
29. Zhao Y, Zhang D L, Kong G L, Pan G, Liao X B, *Phys. Rev. Lett*. **74**, 558 (1995).
30. Santos P V, Johnson N M, and Street R A, *Phys. Rev. Lett.* **67**, 2686 (1991).
31. Beyer W, Herion J, Wagner H, Zastrow U, *Phil. Mag*. **B63**, 269 (1991).
32. Figure courtesy of R. Biswas; for information on the calculations, see Biswas R and Li Y P, *Phys. Rev. Lett*. **82**, 2512 (1999).
33. The assignment of the D-center observed in electron paramagnetic resonance measurements with a dangling bond has been challenged in favor of "floating bonds" (Stathis J H, Pantelides S T, *Phys. Rev. B* **37**, 6579–6582 (1988)).
34. Zafar S, Schiff E A, *Phys. Rev. Lett*. **66**, 1493 (1991).
35. Jackson W B, Tsai C C, Thompson R, *Phys. Rev. Lett*. **64**, 56 (1990).
36. See the review by Fritzsche H, *Annu. Rev. of Mater. Res*. **31**, 47 (2001).
37. Branz H M, *Phys. Rev. B* **59**, 5498 (1999).
38. Park H R, Liu J Z, Wagner S, *Appl. Phys. Lett*. **55**, 2658 (1989).
39. Ley L, *J. Non-Cryst. Solids* **114**, 238 (1989).
40. Jackson W B, Kelso S M, Tsai C C, Allen J W, Oh S J, *Phys. Rev. B* **31**, 5187 (1985).
41. Cody G, Tiedje T, Abeles B, Brooks B, Goldstein Y, *Phys. Rev. Lett*. **47**, 1480 (1981).
42. Tiedje T, in *Hydrogenated Amorphous Silicon II*, 261–300, Joannopoulos J D and Lucovsky G (eds), Springer-Verlag, New York (1984).
43. Gu Q, Wang Q, Schiff E A, Li Y-M, Malone C T, *J. Appl. Phys.* **76**, 2310 (1994).
44. Wang Q, Antoniadis H, Schiff E A, Guha S, *Phys. Rev. B* **47**, 9435 (1993).

45. Gu Q, Schiff E A, Chevrier J-B, Equer B, *Phys. Rev. B* **52**, 5695 (1995).
46. Mott N V, *Conduction in Non-Crystalline Solids*, Oxford University Press, Oxford (1987).
47. Dawson R, Li Y, Gunes M, Heller D, Nag S, Collins R, Wronski C, *Amorphous Silicon Technology – 1992*, Thompson M J *et al*. (eds), Materials Research Society, Symposium Proceedings Vol. 258, pp 595–600, Pittsburgh (1993).
48. Hishikawa Y, Nakamura N, Tsuda S, Nakano S, Kishi Y, Kuwano Y, *Japan J Applied Physics* **30**, 1008–1014 (1991).
49. Tauc J, in *Optical Properties of Solids*, 277, Abeles F (ed.), North Holland, Amsterdam (1972).
50. Chen I S, Wronski C R, *J. Non-Cryst. Solids* **190**, 58 (1995).
51. Jackson W B, Amer N, *Phys. Rev. B* **25**, 5559 (1982).
52. Lee J-K, Schiff E A, *Phys. Rev. Lett.* **68**, 2972 (1992).
53. Antoniadis H, Schiff E A, *Phys. Rev. B* **46**, 9482–9492 (1992).
54. Han D, Melcher D C, Schiff E A, Silver M, *Phys. Rev. B* **48**, 8658 (1993).
55. Hama S T, Okamoto H, Hamakawa Y, Matsubara T, *J. Non-Cryst. Solids* **59–60**, 333 (1983).
56. Guha S, Payson J S, Agarwal S C, Ovshinsky S R, *J. Non-Cryst. Solids* **97, 98**, 1455 (1987).
57. Mackenzie K, Burnett J, Eggert J, Li Y, Paul W, *Physical Review B* **38**, 6120–6136 (1998).
58. Middya A R, Ray S, Jones S J, Williamson D L, *J.Appl. Phys.* **78**, 4966 (1995).
59. Stutzmann M, Street R A, Tsai C C, Boyce J B, Ready S E, *J. Appl. Phys.* **66**, 569 (1989).
60. Li Y-M, Catalano A, Fieselmann B F, in *Amorphous Silicon Technology – 1992*, Thompson M J *et al*. (eds), 923, Materials Research Society, Symposium Proceedings Vol. 258, Pittsburgh (1993).
61. Arya R R, Catalano A, Oswald R S, *Appl. Phys. Lett.* **49**, 1089 (1986).
62. Tsukada T, in *Technology and Applications of Amorphous Silicon*, Street R A (ed.), Springer, Berlin (2000).
63. Vallat-Sauvain E, Shah A, Ballat J, in *Thin Film Solar Cells: Fabrication, Characterization, and Applications*, Poortmans J, Arkhipov A (eds), John Wiley & Sons, Ltd (2006).
64. Chittick R, Alexander J, Sterling H, *J. Electrochem. Soc.* **116**, 77–81 (1969).
65. Spear W, LeComber P, *J. Non-Cryst. Solids* **8–10**, 727–738 (1972).
66. Arya R, Carlson D, *Prog. Photovoltaics* **10**, 69–76 (2002).
67. Carlson D, *US Patent 4,317,844* (1982).
68. Curtins H, Wyrsch N, Shah A, *Electron. Lett.* **23**, 228–230 (1987).
69. Chatham H, Bhat P, Benson A, Matovich C, *J. Non-Cryst. Solids* **115**, 201–203 (1989).
70. Saito K, Sano M, Matsuyama J, Higasikawa M, Ogawa K, Kajita I, *Tech. Digest PVSEC-9*, 579 (1996).
71. Matsumura H, *Jpn. J. Appl. Phys.* **25**, L949–L951 (1986).
72. Mahan A, Carapella J, Nelson B, Crandall R, Balberg I, *J. Appl. Phys.* **69**, 6728–6730 (1991).
73. Takei T, Tanaka T, Kim W, Konagai M, Takahashi K, *J. Appl. Phys.* **58**, 3664–3668 (1991).
74. Rocheleau R, Hegedus S, Buchanan W, Jackson S, *Appl. Phys. Lett.* **51**, 133–135 (1987).
75. Paul W, Lewis A, Connel G, Moustakas T, *Solid State Commun.* **20**, 969–972 (1976).
76. Moustakas T, Wronski C, Tiedje T, *Appl. Phys. Lett.* **39**, 721–723 (1981).
77. Chapman B, *Glow Discharge Processes*, John Wiley & Sons, Inc., New York (1980).
78. Luft W, Tsuo Y, *Hydrogenated Amorphous Silicon Alloy Deposition Processes*, Marcel Dekker, New York (1993).
79. Kushner M, *J Appl Phys* **63**, 2532–2552 (1988).
80. Matsuda A, *J Vac Sci Tech A* **16**, 365–368 (1998).
81. Y Hishikawa, S Tsuda, K Wakisaka, Y Kuwano, *J Appl. Phys.* **73**, 4227–4231 (1993).
82. Zanzucchi P, Wronski C, Carlson D, *J Appl Phys* **48**, 5227–5236 (1977).
83. Deng X, Narasimhan K, Evans J, Izu M, Ovshinsky S, *Proc. 1st World Conf. on Photovoltaic Energy Conversion* p. 678 (Hawai'i 1994).
84. Hegedus S, *Progress in Photovoltaics* **5**, 151–168 (1997).

85. Dickson C, Fieselmann B, Oswald R, *J Crystal Growth* **89**, 49–61 (1988).
86. Kamei T, Matsuda A, *J Vac Sci Tech A* **17**, 113–119 (1999).
87. Kinoshita T, Isomura M, Hishakawa Y, Tsuda S, *Jap J Appl Phys* **35**, 3819–3824 (1996).
88. Kampas F, *J. Appl. Phys.* **54**, 2276–2280 (1983).
89. Jasinski, J, Whittaker, E, Bjorklunk G, Dreyfus R, Estes R, Walkup R, *Appl. Phys. Lett.* **44**, 1155–1157 (1984).
90. Robertson R, Gallagher A, *J. Chem. Phys.* **85**, 3623–3630 (1986).
91. Gallagher A, *J. Appl. Phys.* **63**, 2406–2413 (1988).
92. Doughty D, Gallagher A, *Phys Rev A* **42**, 6166–6170 (1990).
93. Arya R, Carlson D, *Prog. Photovoltaics* **10**, 69–76 (2002).
94. Platz R, Wagner S, Hof C, Shah A, Wieder S, Rech R, *J Appl Phys* **84**, 3949–3953 (1998).
95. Shah A, *et al*., *Solar Energy Matl Solar Cells* **78**, 469–491 (2003).
96. Smets A, Matsui T, Kondo M, *J Appl Physics* **104**, 034508 (2008).
97. Yue G, Yan B, Yang, J, Guha, *Proc 33rd IEEE Photovoltaic Specialist Conf,* paper 260 (San Diego, 2008).
98. Schropp R, *et al*., *Proc 31st IEEE Photovoltaic Specialist Conf* (Orlando, 2005) pp 1371–1374.
99. Vetterl O, *et al*., *Solar Energy Materials Solar Cells* **62**, 97–108 (2000).
100. Kroll U *et al*., *Proc 22nd Euro PVSEC* (Milan, 2007).
101. Saito K, Sano M, Otoshi H, Sakai A, Okabe S, Ogawa K. *Proc 3rd World Conf on PV Energy Conversion* (Osaka, 2003).
102. C. Ferreira, J. Loureiro, *J. Phys. D* **16**, 2471–2483 (1983).
103. M. Wertheimer, M. Moisan, *J. Vac. Sci. Tech.* A3, 2643–2649 (1985).
104. S. Oda, J. Noda, M. Matsumura, J, *Jpn. J. Appl. Phys.* **29**, 1889 (1990).
105. H Takatsuka, M Noda, Y Yonekura, Y Takeuchi, Y Yamauchi, *Solar Energy* **77**, 951–960 (2004).
106. Fujioka, Y, Shimuzu A, Fukuda H, Oouchida T, Tachibana S, Tanamura H Nomoto K, Okamoto K, Abe M, *Solar Energy Materials and Solar Cells* **90**, 3416–3421 (2006).
107. Ito N, Kondo M, Matsuda A, *Proc. 28th Photovoltaic Specialists Conference*, pp 900–903 (Anchorage, 2000).
108. H Meiling, W van Sark, J Bezemer, W van der Weg *J Appl Phys* **80**, 3546–3551 (1996).
109. Kato I, Wakana S, Hara S, Kezuka H, *Jpn. J. Appl. Phys.* **21**, L470 (1982).
110. Hudgens S, Johncock A, Ovshinsky S, *J. Non-Cryst. Solids* **77–78**, 809 (1985).
111. Watanabe T, Azuma K, Nakatani M, Suzuki K, Sonobe T, Shimada T, *Jpn. J. Appl. Phys.* **25**, 1805 (1986).
112. Saito K, Sano M, Matsuyama J, Higasikawa M, Ogawa K, Kajita I, *Tech. Digest PVSEC-9*, p. 579 (1996).
113. Guha S, Xu X, Yang J, Banerjee A, *Appl. Phys. Lett.* **66**, 595–597 (1995).
114. Saito K, Sano M, Ogawa K, Kajita I, *J. Non-Cryst. Solids* **164–166**, 689 (1993).
115. Deng X, Jones S, Liu T, Izu M, Ovshinsky S, *Proc. 26th IEEE Photovoltaic Specialists Conference*, p. 591, (Anaheim, 1997).
116. Matsumura H, *Jpn. J. Appl. Phys.* **25**, L949–L951 (1986).
117. Wiesmann H, Ghosh A, McMahon T, Strongin M, *J. Appl. Phys.* **50**, 3752 (1979).
118. Mahan A, Carapella J, Nelson B, Crandall R, Balberg I, *J. Appl. Phys.* **69**, 6728–6730.
119. Bauer S, Herbst W, Schroder B, Oechsner H, *Proc 26th IEEE Photovoltaic Spec Conf* 719–722 (Anaheim, 1997).
120. Mahan A, Xu Y, Nelson B, Crandall R, Cohen J, Palinginis K, Gallagher A, *Appl. Phys. Lett.* **78**, 3788 (2001).
121. Schropp R, Li H, Franken R, Rath J, van der Wert C, Scuttauf J, Stolk R, *Thin Solid Films* **516** (2008) 6818–6823; or *Solar Energy Materials and Solar Cells* 2009 (on-line 20-Mar-2009).
122. Osono S, Kitazoe M, Tsuboi H, Asari S, Saito K, *Thin Solid Films* **501**, 601 (2006).

123. Moustakas T, Maruska H, Friedman R, *J. Appl. Phys.* **58**, 983–986 (1985).
124. Konagai M, Kim W, Tasaki H, Hallerdt M, Takahashi K, *AIP Conf. Proc.* **157**, 142–149 (1987).
125. Rocheleau R, Hegedus S, Buchanan W, Jackson S, *Appl. Phys. Lett.* **51**, 133–135 (1987).
126. Parsons G, Tsu D, Lucovsky G, *J. Vac. Sci. Technol., A* **6**, 1912–1916 (1988).
127. Sakamoto Y, *Jpn. J. Appl. Phys.* **16**, 1993–1998 (1977).
128. Dalal V, Maxson T, Girvan R, Haroon S, *Mater. Res. Soc. Symp. Proc.* **467**, 813–817 (1997).
129. Jones S, Crucet R, Deng X, Izu M, *Mater. Res. Soc. Symp. Proc.* **609**, A4.5 (2000).
130. Guha S, Narasimhan K, Pietruszko S, *J. Appl. Phys.* **52**, 859 (1981).
131. Okamoto S, Hishikawa Y, Tsuda S, *Japan J Appl Physics* **35**, 26–33 (1996).
132. Collins R, *et al*., *Solar Energy Materials Solar Cells* **78**, 143–180 (2003).
133. Ferlauto A, Koval R, Wronski C, Collins R, *Appl. Phys. Lett.* **80**, 2666 (2002).
134. Kamei T, Stradins P, Matsuda A, *Appl Phys Lett* **74**, 1707–1709 (1999).
135. Guha S, *Solar Energy* **77**, 887–892 (2004).
136. Guha S, Yang J, Williamson D, Lubianiker Y, Cohen D, Mahan H, *Appl Phys Lett* **74**, 1860–1862 (1999).
137. Vallat-Sauvain E, Kroll U, Meier J, Shah A, Pohl J, *J Appl Physics* **87**, 3137–3142 (2000).
138. Shah A V, Meier J, Vallat-Sauvain E, Wyrsch N, Kroll U, Droz C, Graf U, Material and solar cell research in microcrystalline silicon, *Sol. Energy Mat. Sol. Cells* **78**, 469–491 (2003).
139. Guo L, Kondo M, Fukawa M, Saitoh K, Matsuda A, *Jpn. J. App. Phys.* **37**, L1116–L1118 (1998).
140. Rech B, Repmann T, van den Donker M N, Berginski M, Kilper T, Hupkes J, Calnan S, Stiebig H, Wieder S *Thin Solid Films* **511–512**, 548–555 (2006).
141. Paul W, Street R, Wagner S, *J Electronic Mat* **22**, 39–48 (1993).
142. Guha S, Payson J, Agarwal S, Ovshinsky S, *J. Non-Cryst. Solids* **97–98**, 1455 (1987).
143. Tawada Y, Tsuge T, Kondo M, Okamoto H, Hamakawa Y, *J Appl Physics* **53**, 5273–5281 (1982).
144. Pearce J, Podraza N, Collins R, Al-Jassim M, Jones K, Wronski C, *J Applied Physics* **101**, 114301 (2007).
145. Rath J, Schropp, *Solar Energy Materials Solar Cells* **53**, 189–203 (1998).
146. Burgelman M, Verschraegen J, Degrave S, Nolletet P, *Prog. Photovolt: Res. Appl.* **12**, 143–153 (2004).
147. Zhu & AMPS-1D H, Kalkan AK, Hou J, Fonash SJ, *AIP Conf. Proceedings* **462**, 309 (1999). -1D is a copyright of Pennsylvania State University.
148. Welcome, expert. The intrinsic layer parameters for most of the calculations in this section were published in Jianjun Liang, Schiff EA, Guha S, Baojie Yan, Yang J, *Appl. Phys. Lett.* **88** 063512 (2006). Idealized doped layer parameters were chosen; the details have no significant effect on the calculations.
149. Schropp R E I, Zeman M, *Amorphous and Microcrystalline Silicon Solar Cells: Modeling Materials, and Device Technology*, Kluwer, Boston (1998).
150. Measurements are provided through the courtesy of United Solar Ovonic LLC. The untextured cells were made on uncoated stainless steel, which has a reflectivity of about 30%.
151. Rose A, *Concepts in Photoconductivity and Allied Topics* (Krieger, 1978).
152. Fonash S, *Solar Cell Device Physics* (Academic Press, New York, 1981).
153. Zhu K, Yang J, Wang W, Schiff EA, Liang J, Guha S, in *Amorphous and Nanocrystalline Silicon Based Films - 2003*, Abelson JR, Ganguly G, Matsumura H, Robertson J, Schiff EA (eds.) (Materials Research Society Symposium Proceedings Vol. 762, Pittsburgh, 2003), pp 297–302.
154. Zeman M, in *Thin Film Solar Cells: Fabrication, Characterization, and Applications*, Poortmans J, Arkhipov A (eds), John Wiley & Sons, Ltd (2006).

155. Dutta U, Chatterjee P, Tchakarov S, Uszpolewicz M, Roca I Cabarrocas P, *J. Appl. Phys.* **98**, 044511 (2005).
156. Eray A and Nobile G, The optimization of a-Si:H *pin* solar cells: more insight from simulation, in *Recent Developments in Solar Energy*, Tom P. Hough (ed.) (Nova Science, 2007).
157. Liang J, Schiff E A, Guha S, Yan B, Yang J, *Appl. Phys. Lett.* **88**, 063512 (2006).
158. For optically simple, untextured cells such as those in Figure 12.15, this fraction can be calculated from the optical absorption spectrum and the solar illumination spectrum (Figure 12.2).
159. Schiff E A, *J. Phys.: Condens. Matter* **16**, S5265 (2004). Figure 12.16 was generated using the multiple-trapping equations and fitting parameters presented in this paper.
160. In Figure 12.16, the displacement of electrons is proportional to time, which is unsurprising. However, the displacement of holes increases only as (roughly) the square-root of time. This peculiar behavior is called "anomalous dispersion" [161]. In a-Si:H and related materials, anomalous dispersion is attributed to trapping by states in the bandtail. Thus a free hole – one in a level below E_V – has a band mobility μ_p around 0.5 cm^2/Vs. However, a hole is trapped and released by valence bandtail states many times in a microsecond, which reduces its travel in a microsecond by more than 100 times. For more details, see [159]. The conduction band also has a bandtail; it is about 22 meV wide in a-Si:H, which is about half the width of 45–50 meV width for the valence bandtail. The narrower width means it has little effect near room-temperature, which is why an electron's drift-mobility is so much larger than that of a holes.
161. Scher H, Shlesinger M F, Bendler J T, *Physics Today* **44** (1), 26 (1991).
162. Schiff E A, *Solar Energy Materials and Solar Cells* **78**, 567 (2003).
163. It should be noted that the collection zone's width is not the ambipolar diffusion length. The latter is a well-known concept for high-mobility semiconductors such as c-Si. Ambipolar diffusion applies near the edge of what we've termed the recombination region, and its minor effect in a-Si:H can be just barely discerned in Figure 12.17. Instead, the concept of a limiting carrier collection, or drift, length is the more useful and appropriate for analysis of a-Si solar cells [84].
164. Guha S, Yang J, Nath P, Hack M, *Appl. Phys. Lett.* **49**, 218 (1986).
165. Arya RR, Catalano A, Oswald R S, *Appl. Phys. Lett.* **49**, 1089 (1986).
166. Hegedus SS, Rocheleau R E, Tullman R N, Albright D E, Saxena N, Buchanan W A, Schubert K E, Dozier R, *Proceedings 20th IEEE Photovoltaic Specialists Conference*, p. 129 (Las Vegas 1988).
167. Sakai H, Yoshida T, Fujikake S, Hama T, Ichikawa Y, *J. Appl. Phys.* **67**, 3494 (1990).
168. Munyeme G, Zeman M, Schropp REI, van der Weg WF, *Phys. Stat. Sol. (c)* **1**, 2298 (2004).
169. Williams EL, Jabbour GE, Wang Q, Shaheen SE, Ginley DS, Schiff EA, *Appl. Phys. Lett.* **87**, 223504 (2005).
170. Bailat J., Domine D., Schluchter R, Steinhauser J, Fay S, Freitas F, Bucher C, Feitknecht L, Niquille X, Tscharner T, Shah A, Ballif C, *Proceedings IEEE 4th World Conference on Photovoltaic Energy Conversion*, p. 1533 (Hawai'i 2006).
171. Mai Y, Klein S, Carius R, Wolff J, Lambertz A, Finger F, *J. Appl. Phys.* **97**, 114913 (2005).
172. Klaver A and van Swaaij R.A.C.M.M., *Solar Energy Materials and Solar Cells* **92**, 50 (2008).
173. Crandall R S and Schiff E A, *13th NREL Photovoltaics Program Review*, Ullal H S, Witt C E (eds), *Amer. Inst. of Phys., Conf. Proc*. **353**, pp. 101–106 (1996).
174. Two effects cancel to yield the net slope of 1.0 v/eV in Figure 12.18. As the bandgap decreases, the photogeneration rate increases and increases V_{OC} slightly. Concurrently, the bandtails widen, which decreases V_{OC}.
175. B. Pieters B, Stiebig H, Zeman M, van Swaaij R. A. C. M. M., *J. Appl. Phys.* **105**, 044502 (2009).

176. Schiff E A, in *Amorphous and Polycrystalline Thin-Film Silicon Science and Technology-2009*, Flewitt A, Wang Q, Hou J, Uchikoga S, Nathan A (eds) (*Mater. Res. Soc. Symp. Proc*. Volume 1153, Warrendale, PA, 2009), 1153-A15-01.
177. Hegedus S and Deng X, *Proceedings 25th IEEE Photovoltaic Specialists Conference*, p. 1061, (Washington DC 1996).
178. Hegedus S, Buchanan W, Liu X, Gordon R, *Proceedings 25th IEEE Photovoltaic Specialists Conference*, p. 1129, Institute of Electrical and Electronics Engineers (1996).
179. Lechner P, Geyer R, Schade H, Rech B, Müller J, *Proceedings 28th IEEE Photovoltaic Specialists Conference*, p. 861, Institute of Electrical and Electronics Engineers (2000).
180. Yablonovitch E, *J. Opt. Soc. Am.* **72**, 899 (1982).
181. Zhou D, Biswas R, *J. Appl. Phys.* **103**, 093102 (2008).
182. Tiedje T, Yablonovitch E, Cody G D, Brooks B G, *IEEE Trans. on Elect. Devices* **ED-31**, 711 (1984).
183. Krc J, Zeman M, Smole F, Topic M, *J. Appl. Phys.* **92**, 749–755 (2002).
184. Yan B, Yue G, Xu X, Yang J, Guha S, *Phys. Status Solidi A* **207**, 671 (2010).
185. Poruba A, Fejfar A, Remes Z, Springer J, Vanecek M, Kocka J, Meier J, Torres P, Shah A, *J. Appl. Phys.* **88**, 148.
186. Banerjee A and Guha S, *J. Appl. Phys.* **69**, 1030 (1991).
187. Atwater H, Polman A, *Nature Materials* **9**, 205–213 (2010).
188. Smith Z, Wagner S *Proceedings 19th IEEE Photovoltaic Specialists Conference*, pp 204–209 (New Orleans 1987).
189. Jardine C, Conibeer G, Lane K, *Proceedings of the 17th European Photovoltaic Solar Energy Conference*, pp 724–727 (Munich 2001).
190. Minemoto T *et al*., *Solar Energy Material Solar Cells* **91**, 120–122 (2007).
191. Agarwal P, Povolny H, Han S, Deng X, *J. Non-Cryst. Solids* **299–302**, 1213–1218 (2002).
192. Yang L, Chen L, Catalano A, *Mater. Res. Soc. Symp. Proc.* **219**, 259–264 (1991).
193. Guha S, Yang J, Pawlikiewicz A, Glatfelter T, Ross R, Ovshinsky S, *Appl. Phys. Lett.* **54**, 2330 (1989).
194. Zimmer J, Stiebig H, Wagner H, *J. Appl. Phys.* **84**, 611–617 (1998).
195. Bullot J, Schmidt M, *Physica Status Solidi (B)* **143**, 345–416 (1987).
196. Li Y-M, *Proc. Mater. Res. Soc. Symp.*, **297**, p. 803 (1993).
197. For example at Energy Photovoltaics (EPV, previously Chronar and Advanced Photovoltaic Systems); Power Film (previously Iowa Thin Film); and Sinar Solar.
198. Yang J, Yan B, Yue G, Guha S, *Proc 31st IEEE Photovoltaic Spec Conf,* pp 1359–1362 (Orlando, 2005).
199. Iida H, Shiba N, Mishuka T, Karasawa H, Ito A, Yamanaka M, Hayashi Y, *IEEE Electron Device Lett.* **EDL-4**, 157–159 (1983).
200. Gordon R, Proscia J, Ellis F, Delahoy A, *Sol. Energy Mater.* **18**, 263–281 (1989).
201. Yamamoto K *et al*. *Solar Energy* **77**, 939–949 (2004).
202. Fischer D, *et al*. *Proc 25th IEEE Photovoltaic Specialist Conf,* pp 1053–1056 (Washington DC 1996).
203. Meier J, Kroll U, Vallat-Sauvain E, Spritznagel J, Graf U, Shah A, *Solar Energy Matl Solar Cells* **77**, 983 (2004).
204. Hegedus S, Kampas F, Xi J, *Appl. Phys. Lett.* **67**, 813 (1995).
205. Banerjee A, Yang J, Glatfelter T, Hoffman K, Guha S, *Applied Physics Lett* **64**, 1517–1519 (1994).
206. Loffler J, Gordijn A, Stolk R, Li H, Rath J, Schropp R, *Solar Energy Materials and Solar Cells* **87**, 251–259 (2005).
207. Burdick J, Glatfelter T, *Sol. Cells* 9 (classic reference but no longer available) **18**, 310–314 (1986).
208. Mueller R, *Sol. Energy Mater. Sol. Cells* **30**, 37–45 (1993).

209. Benagli *et al*., *Proc. 24th EU PVSEC* paper 3BO.9.3, pp 2293–2298 (Hamburg 2009).
210. Meier, J. Spitznagel, U. Kroll, C. Bucher, S. Faÿ, T. Moriarty, A. Shah, *Thin Solid Films*, **451–452**, pp 518–524 (2004).
211. Yamamoto K, Yoshimi M, Suzuki T, Tawada Y, Okamoto Y, Nakajima A, Thin Film Poly-Si Solar Cell on Glass Substrate Fabricated at Low Temperature, in *Amorphous and Microcrystalline Silicon Technology – 1998*, pp 131–138 (Mat. Res. Soc. Symp. Proc. Vol. 507, 1998).
212. Bailat, D. Dominé, R. Schlüchter, J. Steinhauser, S. Faÿ, F. Freitas, C. Bücher, L. Feitknecht, X. Niquille, T. Tscharner, A. Shah, C. Ballif, *Proc 4th WCPEC Conference*, 1533–1436 (Waikoloa 2006).
213. Yang J, Banerjee A, Lord K, Guha S, *2nd World Conf. On Photovoltaic Energy Conversion*, p. 387 (1998).
214. Hishikawa Y, Ninomiya K, Maryyama E, Kuroda S, Terakawa A, Sayama K, Tarui H, Sasaki M, Tsuda S, Nakano S, *Conference Record 1st IEEE World Conf*. Photovoltaic Solar Energy Conversion, pp 386–393 (1994).
215. Arya R, Oswald R, Li Y, Maley N, Jansen K, Yang L, Chen L, Willing F, Bennett M, Morris J, Carlson D, *Conference Record 1st IEEE World Conf. Photovoltaic Solar Energy Conversion*, p. 394 (1994).
216. Yamamoto K *et al*. *Proceedings 31st IEEE Photovoltaic Specialists Conference* 1468–1471 (Orlando 2005).
217. Yang J, Guha S, *SPIE* 2009.
218. Bailat J, *et al*., *Proc 26th EU-PVSEC*, paper 3BO-11-5 (Valencia 2010) .
219. Domine D, Buehlmann, Bailat J, Billet A, Feltrin A, Ballif C, *Proc. 23rd EU-PVSEC*, Valencia, Spain, p. 2096, 2008.
220. Rath J, *Proceedings 4th IEEE WCPEC* (2006 Waikoloa) pp 1473–1476.
221. Yoshida T, Tabuchi K, Takano A, Tanda M, Sasaki T, Sato H, Fijikake S, Ichikawa Y, Harashima K, *Proceedings 28th IEEE Photovoltaic Specialists Conference*, 762–765 (2000).
222. Wang W, Povolny H, Du W, Liao X, Deng X, *Proceedings 29th IEEE Photovoltaic Specialists Conference*, 1082–1085 (2002).
223. Nomoto K, Saitoh H, Chida A, Sannomiya H, Itoh M, Yamamoto Y, *Solar Energy Materials and Solar Cells* **34**, 339–346 (1994).
224. Yue G, Yan B, Ganguly G, Yang J, Guha S, *Appl. Phys. Lett.* **88**, 263507 (2006).
225. Saito K, Sano M, Matuda K, Kondo Takaharu, Nishimoto T, Ogawa K, Kajita I, *Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, pp 351–354 (1998).
226. Yamamoto K *et al*. *Solar Energy* **77**, 939–949 (2004).
227. Yan B, Yue G, Owens JM, Yang J, Guha S, *Conference Record 4th WCPEC*, p. 1477 (2006, Waikoloa).
228. Deng X *et al*., *Proceedings 4th WCPEC* pp 1461–14 (Waikoloa 2006).
229. Jones S, Crucet R, Capangpangan R, Izu M, Banerjee A, *Mater. Res. Soc. Symp. Proc.* **664**, A15.1 (2001).
230. Hamma S, Roca I Cabarrocas P, *Solar Energy Matl and Solar Cells* **69**, 217–239 (2001).
231. Faraji M, Gokhale S, Ghoudhari S, Takwake M, Ghaisas S, *Appl. Phys. Lett.* **60**, 3289–3291 (1992).
232. Meier J, Fluckiger R, Keppner H, Shah A, *Appl. Phys. Lett.* **65**, 860–862 (1994).
233. Yamamoto K, Yoshimi M, Suzuki T, Okamoto Y, Tawada Y, *Proc 26th IEEE Photovoltaic Spec Conf*, 575–580 (Anaheim 1997).
234. Yang J, Yan B, Yue G, Guha S, *Proc 31st IEEE Photovoltaic Spec Conf*, pp 1359–1362 (Orlando 2005).
235. Matsui T, Matsuda A, Kondo M, *Solar Energy Materials and Solar Cells* **90**, 3199–3204 (2006).

236. Bailat J, Vallat-Sauvain E, Feitknecht L., Droz C., Shah A., *Journal of Non-Crystalline Solids* **299–302**, 1219–1223 (2002).
237. Kluth O, Zahren C, Steibig H, Rech B, Schade H, *Proc 19th EU PVSEC*, 3DV1.56 (Paris, 2004).
238. Kluth O, Rech B, Houben L, Wieder S, Schöpe G, Beneking C, Wagner H, Löffl A, Schock H W, *Thin Solid Films* **351**, 247–253 (1999).
239. Rech B, Kluth O, Repmann T, Roschek T, Springer J, Muller J, Finger F, Stiebig H, Wagner H, *Solar Energy Matl Solar Cells* **74**, 439 (2002).
240. Yan B, Yue G, Yang J, Guha S, Wiilliamson D, Han D Jiang C, *Appl Phys Lett* **85**, 1955–1957 (2004).
241. van den Donker, Rech B, Finger F, Houben L, Kessels W, van de Sanden, *Progress in Photovoltaics* **15**, 291–301 (2007).
242. Coutts T, Emery K, Ward S, *Progress in Photovoltaics* **10**, 195–203 (2002).
243. Ganguly G, Yue G, Yan B, Yang J, Guha S, Proc 4th World Conf Photovoltaic Energy Conversion, pp 1712–1715 (Waikaloa, 2006).
244. Yan B, Yue G, Guha S, *Material Res Soc Symp Proc* **989**, 0989-A15-01 (2007).
245. *Photon International 2008 Module Production Survey* pp 195–202 (March 2009).
246. Richard D, *Photon International*, pp 190–207 (November 2009).
247. Hegedus S, *Progress in Photovoltaics* **14**, 393–411 (2006).
248. Izu M, Ellison T, *Solar Energy Materials and Solar Cells* **78**, 613–626 (2003).
249. Guha S, Yang J, *Proceedings 29th IEEE Photovoltaic Specialists Conference*, 1070–1075 (New Orleans, 2002).
250. Deng X, Narasimhan K, *Proceedings 1st World Conf. on Photovoltaic Energy Conversion*, p. 555 (Hawai'i, 1994).
251. Nath P, Hoffman K, Vogeli C, Ovshinsky S, *Appl. Phys. Lett.* **53**, 986–988 (1988).
252. Jansen K, Varvar A Groelinger J, *Proceedings 33rd IEEE PVSC* (San Diego, 2008).
253. H. Maurus, M. Schmid, B. Blersch, P. Lechner, H. Schade. *Proceedings of 3rd World Conference on Photovoltaic Solar Energy Conversion*, pp 2375–2378 (Osaka, 2003).
254. Klein S, Repmann T, Wieder S, Muller J, Buschbaum S, Rhode M, *Proc 22nd EU PVSEC* (Milan, 2007).
255. Fan Y, *et al*., *Proc 34th IEEE Photovoltaic Spec Conf* (2009, Philadelphia).
256. Meier J, *et al*., *Material Research Soc Symp Proc* Vol 889, 0989-A24-01 (2007).
257. Brookhaven National Photovoltaic Environmental Research Center www.bnl.gov/pv.
258. Ersol Solar Nova T-85 product specification sheet.
259. Signet Solar SI S1-380Ax product specification sheet.
260. Mitsubishi Heavy Industry MA-100-T2 product specification sheet.
261. Schott Solar ASI-TM-100 product specification sheet.
262. Energy Photovoltaic EPV-53 product specification sheet.
263. Okamoto S, Terakawa A, Maruyama E, Shinohara W, Hishikawa Y, Kiyama S, *Proc 28th IEEE Photovoltaic Specialists Conference*, pp 736–741 (Anchorage, 2000).
264. United Solar Ovonic PVL-144 product specification sheet.
265. Xu X, *et al*., *Proceedings 34th IEEE Photovoltaic Specialists Conference* (2009, Philadelphia) 002159–002164.
266. Fecioru-Morariu M, *et al*., *Proc 26th EU-PVSEC*, paper 3AV.1.23 (Valencia 2010)
267. Ersol Vega-T Module Data Sheet, downloaded 13 August, 2009.
268. Sharp Solar Press Release, 14 October, 2008.
269. Astronergy CEO Liyou Yang, Greentech Media Webinar (September 29, 2010).
270. Kaneka U-EA105 product specification sheet.
271. Rech B *et al*., *Proc. 2nd World Conf. on Photovoltaic Solar Energy Conversion*, pp 391–396 (1998).
272. Shah A, *et al*., *Progress in Photovoltaics* **12**, 113–142 (2004).

273. 2008 Module Survey, *Photon International* (February 2009) 136–209.
274. International Electrotechnical Commission, *Thin-film terrestrial photovoltaic (PV) modules – Design qualificiation and type approval* (2008).
275. Osburn D, *Proc Amer Solar Energy Soc Conf* (2003 Austin). Also available at www.nrel .gov/ncpv/thin_film/pn_techbased_amorphous_silicon.html.
276. Adelstein J, Sekulic B, *Proc 31st IEEE Photovoltaic Spec Conf,* pp 1627–1630 (Orlando) 2005.
277. http://www.nrel.gov/pv/thin_film/docs/035098_pvfaq_materials.pdf.
278. PV News Thin Film Forecast, published by Prometheus News June 2007.
279. Vanecek M, Springer J, Poruba A, Kluth O, Repmann T, Rech B, Meier J, Shah A, *Proc 3rd WCPEC* (Osaka 2003) 5PLD1-01.
280. Kondo M, Nagasaki S, Miyahara H, Matsui T, Fujiyabashi T, Matsuda A, *Proc 3rd WCPEC* (Osaka 2003) S20B9-01.

13

Cu(InGa)Se_2 Solar Cells

William N. Shafarman[1], Susanne Siebentritt[2] and Lars Stolt[3]

[1]*University of Delaware, Newark, DE, USA,* [2]*University of Luxembourg, Luxembourg,* [3]*Solibro Research AB, Sweden*

13.1 INTRODUCTION

Cu(InGa)Se_2-based solar cells have often been touted as being among the most promising of solar cell technologies for cost-effective power generation. In the past decade this has led to a surge in manufacturing development by many start-up and large companies worldwide. One company, Solar Frontier is currently setting up a 1 GW production facility. The promise of Cu(InGa)Se_2 is partly due to the advantages of thin films for low-cost, high-rate semiconductor deposition over large areas using layers only a few micrometers thick and for fabrication of monolithically interconnected modules. Perhaps more importantly, very high efficiencies have been demonstrated with Cu(InGa)Se_2 at both the cell and the module levels. A solar cell efficiency of 20.0% with 0.5 cm^2 total area has been fabricated by the National Renewable Energy Laboratory (NREL) [1, 2].[1] Furthermore, several companies have demonstrated large-area, production-scale modules with efficiencies of 12–14% including a confirmed 13.5% efficiency on a monolithically interconnected 3459 cm^2 module [3] and pilot-scale modules with areas ~1000 cm^2 and efficiencies >15%. Finally, Cu(InGa)Se_2 solar cells and modules have shown excellent long-term stability [4] in outdoor testing. In addition to their potential advantages for large-area terrestrial applications, Cu(InGa)Se_2 photovoltaics can be made very lightweight and flexible which is desirable for building integrated and portable applications and have also shown high radiation resistance, compared to crystalline silicon and III-V solar cells [5, 6] so they are also promising for space applications.

The history of $CuInSe_2$ solar cells starts with work done at Bell Laboratories in the early 1970s, following synthesis and characterization of the material first reported by Hahn in 1953 [7] and, along with other ternary chalcopyrite materials, characterization by several groups [8]. The Bell

[1] Note added in proof: a Cu(InGa)Se_2 solar cell with efficiency 20.3% was reported by the Zentrum für Solarenergie- und Wasserstoff-Forschung (ZSW) [Powalla M *et al*., presented at the European Materials Research Society Spring Meeting, Strassbourg (2010), Green, M. A., Emery, K. A., Hishikawa, Y., Warta, W., Prog. Photovolt. 18, 346 (2010)]

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

Labs group grew crystals of a wide selection of these materials and characterized their structural, electronic, and optical properties [8–10]. The first $CuInSe_2$ solar cells were made by evaporating *n*-type CdS onto *p*-type single crystals [11]. These devices were initially recognized for their potential as near-infrared photodetectors since their spectral response was broader and more uniform than Si photodetectors. Optimization for solar cells increased the efficiency to 12% as measured under outdoor illumination "on a clear day in New Jersey" [12].

There has been relatively little effort devoted to devices on single-crystal $CuInSe_2$ since this early work, in part because of the difficulty in growing high-quality crystals [13]. Instead, nearly all the focus has gone to thin film solar cells because of their inherent advantages. The first thin film $CuInSe_2$/CdS devices were fabricated by Kazmerski *et al.* using films deposited by evaporation of $CuInSe_2$ powder along with excess Se [14]. However, $CuInSe_2$ solar cells began to receive a lot of attention when the first high-efficiency, 9.4%, thin film cells were demonstrated by Boeing [15]. At the same time, interest in Cu_2S/CdS thin film solar cells waned owing to problems related to electrochemical instabilities and many of these researchers turned their focus to $CuInSe_2$.

The Boeing devices were fabricated using $CuInSe_2$ deposited by coevaporation, that is, evaporation from separate elemental sources [16], onto ceramic substrates coated with a Mo back electrode. Devices were completed with evaporated CdS or (CdZn)S deposited in two layers with undoped CdS followed by an In-doped CdS layer that served as the main current-carrying material [16]. Throughout the 1980s, Boeing and ARCO Solar began to address the difficult manufacturing issues related to scale-up, yield, and throughput, leading to many advances in $CuInSe_2$ solar cell technology. Boeing focused on depositing $Cu(InGa)Se_2$ by coevaporation, while ARCO Solar focused on a precursor reaction process with Cu and In deposition at a low temperature followed by a reactive anneal in H_2Se. These two processing approaches, coevaporation and precursor reaction, today remain the most common deposition methods and produce the highest device and module efficiencies.

The basic solar cell configuration implemented by Boeing provided the basis for a series of improvements that have led to the high-efficiency device technology of today. The most important of these improvements to the technology include the following:

- The absorber-layer bandgap was increased from 1.04 eV for $CuInSe_2$ to 1.1–1.2 eV by the partial substitution of In with Ga, leading to a substantial increase in efficiency [17].
- The 1- to 2-μm-thick doped (CdZn)S layer was replaced with a thin, $\leq$50 nm, undoped CdS and a conductive ZnO current-carrying layer [18]. This increased the cell current by increasing the short-wavelength (blue) response.
- Soda-lime glass replaced ceramic or borosilicate glass substrates. Initially, this change was made for the lower costs of the soda-lime glass and its good thermal expansion match to $CuInSe_2$. However, it soon became clear that an increase in device performance and processing tolerance resulted primarily from the beneficial indiffusion of sodium from the glass [19].
- Cell were deposited on flexible substrates – polyimide or metal foils – demonstrating possibilities for the manufacturing advantages of roll-to-roll processing and very lightweight flexible modules for space or mobile applications [20].
- Advanced absorber fabrication processes were developed that incorporate bandgap gradients and other modifications that improve the operating voltage and current collection [21, 22].

From its earliest development, $CuInSe_2$ was considered promising for solar cells because of its favorable electronic and optical properties, including its direct bandgap with high absorption coefficient and inherent *p*-type conductivity. As science and technology developed, it also became apparent that it is a very forgiving material since: (1) high-efficiency devices can be made with a wide tolerance to variations in $Cu(InGa)Se_2$ composition [23, 24]; (2) grain boundaries are inherently passive so even films with grain sizes less than 1 μm can be used; and (3) device behavior is insensitive to defects at the junction caused by a lattice mismatch or impurities between the

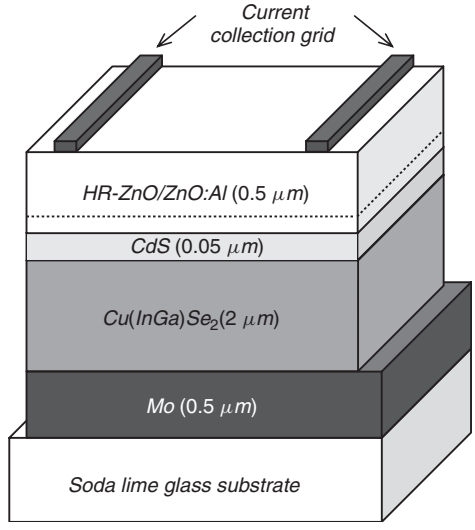


Figure 13.1 Schematic cross-section of a typical Cu(InGa)Se_2 solar cell

Cu(InGa)Se_2 and CdS. The latter enables high-efficiency devices to be processed despite exposure of the Cu(InGa)Se_2 to air prior to junction formation.

High-efficiency $CuInSe_2$-based solar cells with efficiency 18% or greater have been fabricated by at least ten groups around the world. While these groups employ a variety of processing technologies, all the solar cells have the same basic cell structure built around a Cu(InGa)Se_2/CdS junction in a substrate configuration with a Mo back contact. However, some of the record cells and modules also contain buffers other than CdS (Section 13.4.4). Figure 13.1 shows a cross-sectional schematic of a standard device. This structure utilizes a soda-lime glass substrate, coated with a sputtered Mo layer as a back contact. After Cu(InGa)Se_2 deposition, the junction is formed by chemical bath-deposited CdS with thickness ~50 nm. Then a high-resistance (HR) ZnO layer and a transparent conducting oxide, typically doped ZnO or indium tin oxide (ITO), are deposited, usually by sputtering or chemical vapor deposition. Either a current-collecting grid or monolithic series interconnection completes the device or module, respectively. A transmission electron microscopy (TEM) micrograph of the same structure, shown in Figure 13.2, clearly demonstrates the polycrystalline nature of these materials and the conformal coverage of the CdS layer.

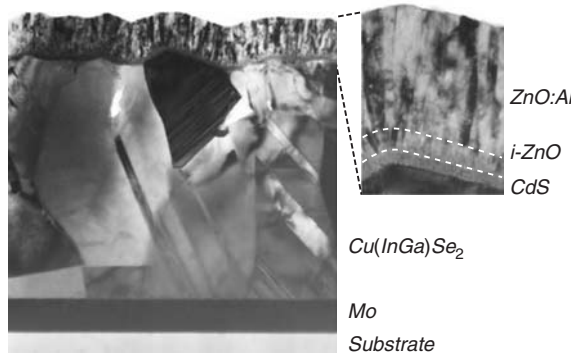


Figure 13.2 TEM cross-section of a Cu(InGa)Se_2 solar cell

With the quantity and diversity of worldwide commercial ventures into development of $Cu(InGa)Se_2$ module manufacturing, it is not possible to provide a list of companies and approaches that wouldn't be incomplete almost immediately. Approaches include deposition by coevaporation, precursor reaction and a wide variety of novel approaches that build on these processes; different materials choices for the window layers including cadmium-free materials; glass-based and flexible modules; in-line and batch processes; and a wide range of cell and module interconnection and packaging configurations.

Despite the level of effort on developing manufacturing processes, there remain large differences in efficiency between laboratory-scale solar cells, minimodules, and the best full-scale modules. This is, in part, inherently true of any photovoltaic technology as areas get larger, but may be exacerbated in this case by the necessity for developing completely new processes and equipment for the large-area, high-throughput deposition needed for manufacturing thin film PV. It is compounded by a still incomplete scientific base for $Cu(InGa)Se_2$ materials and devices, due partly to the fact that it has not attracted a broader interest for other applications. This lack of a fundamental base has been perhaps the biggest hindrance to the maturation of $Cu(InGa)Se_2$ solar cell technology as, historically, much of the progress has been empirical. Still, in many areas a deeper understanding has emerged in the recent years.

In this chapter we will review the current status and the understanding of thin film $Cu(InGa)Se_2$ solar cells from a technology perspective. For deeper scientific discussion of some aspects, we refer to suitable references. In order of presentation, this review covers (Section 13.2) structural, optical, and electrical properties of $Cu(InGa)Se_2$ including a discussion of the influence of Na and O impurities; (Section 13.3) methods used to deposit $Cu(InGa)Se_2$ thin films, the most common of which can be divided into two general types, multisource coevaporation and two-stage processes of precursor deposition followed by Se annealing; (Section 13.4) junction and device formation, which typically is done with chemical bath CdS deposition and a conductive ZnO layer; (Section 13.5) device operation with emphasis on the optical, current-collection, and recombination-loss mechanisms; (Section 13.6) module-manufacturing issues, including process and performance issues and a discussion of environmental concerns; and finally, (Section 13.7) a discussion of the outlook for $CuInSe_2$-based solar cells and critical issues for the future.

13.2 MATERIAL PROPERTIES

The understanding of $Cu(InGa)Se_2$ thin films, as used in photovoltaic (PV) devices, is primarily based on studies of its base material, pure $CuInSe_2$. Reviews of early work on $CuInSe_2$ can be found in references [25–27]. However, the material used for making today's solar cells is $Cu(InGa)Se_2$ containing not only Ga but also significant amounts (of the order of 0.1%) of Na [28]. Even though the behavior of $CuInSe_2$ provides a good basis for the understanding of device-quality material, there are pronounced differences when Ga and Na are present in the films.

In this section the structural, optical, and electrical properties of $CuInSe_2$ are reviewed, along with information about the surface and grain boundaries and the effect of the substrate. In each case, as appropriate, the effect of the alloying with $CuGaSe_2$ to form $Cu(InGa)Se_2$ and the impact of Na and O on the material properties will be discussed. A number of basic material properties is summarized in Table 13.1.

13.2.1 Structure and Composition

$CuInSe_2$ and $CuGaSe_2$ have the chalcopyrite lattice structure. This is a diamond-like structure similar to the sphalerite structure, but with an ordered substitution of the group I (Cu) and group III

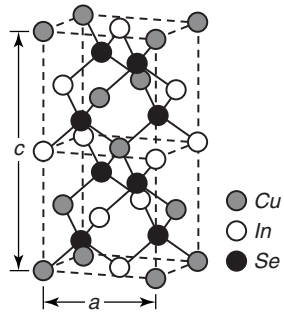


Figure 13.3 The unit cell of the chalcopyrite lattice structure

(In or Ga) elements on the group II (Zn) sites of sphalerite. This gives a tetragonal unit cell depicted in Figure 13.3, with a ratio of the tetragonal lattice parameters c/a close to 2 (see Table 13.1). The deviation from $c/a = 2$ is called the tetragonal distortion and stems from different strengths of the Cu–Se and the In–Se or Ga–Se bonds.

The possible phases in the Cu–In–Se system are indicated in the ternary phase diagram in Figure 13.4. Thin films of Cu–In–Se prepared under an excess supply of Se, that is, normal conditions for thin film growth of Cu(InGa)Se_2, have compositions that fall on, or close to, the tie-line between Cu_2Se and In_2Se_3. Chalcopyrite $CuInSe_2$ is located on this line. So are a number of phases called ordered defect compounds (ODC), because they have a lattice structure described by the chalcopyrite structure with a structurally ordered insertion of intrinsic defects. A comprehensive

Table 13.1 Selected properties of $CuInSe_2$

Property		Value	Units	Reference
Lattice constants	a	5.78	Å	[29]
	c	11.62	Å	
Density		5.75	g/cm^3	[29]
Melting temperature		986	°C	[30]
Thermal expansion coefficients at room temperature	(a axis)	11.23×10^{-6}	1/K	[31]
	(c axis)	7.90×10^{-6}	1/K	
Thermal conductivity at 273 K		0.086	W/(cm·K)	[32]
Dielectric constant	Low frequency	13.6		[33]
	High frequency	7.3–7.75		[34]
Effective mass [m_e]	Electrons – exp.	0.08	m_{e0}	[35]
	theory (||c-axis)	0.08	m_{e0}	[36]
	(||a-axis)	0.09	m_{e0}	
	Holes (heavy) – exp.	0.72	m_{e0}	[34]
	Holes (light)-exp	0.09	m_{e0}	[37]
	theory (||c-axis)	0.66, 0.12	m_{e0}	[36]
	(||a-axis)	0.14, 0.25	m_{e0}	
Energy gap [E_g]		1.04	eV	[7]
Energy gap temperature coefficient		-1.1×10^{-4}	eV/K	[38]

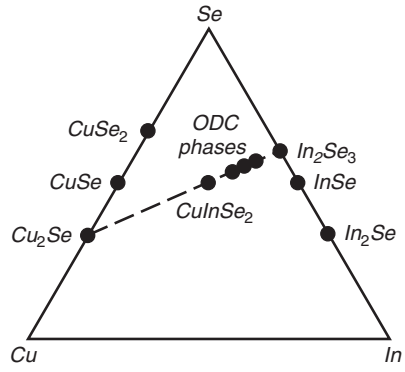


Figure 13.4 Ternary phase diagram of the Cu–In–Se system. Thin film composition is usually near the pseudobinary Cu_2Se–In_2Se_3 tie-line

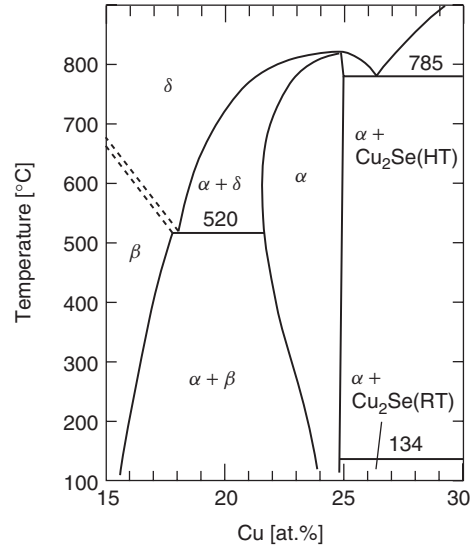


Figure 13.5 Pseudobinary In_2Se_3–Cu_2 Se equilibrium phase diagram for compositions around the $CuInSe_2$ chalcopyrite phase, denoted α. The δ phase is the high-temperature sphalerite phase, and the β phase is an ordered defect phase. Cu_2Se exists as a room-temperature (RT) or high-temperature (HT) phase. (After Gödecke T, Haalboom T, Ernst F, *Z. Metallkd*. **91**, 622–634 (2000) [39])

study of the Cu–In–Se phase diagram has been completed by Gödecke *et al*. [39]. A detail of the Cu_2Se–In_2Se_3 tie-line near $CuInSe_2$ is described by the pseudobinary phase diagram reproduced in Figure 13.5 [39]. Here α is the chalcopyrite $CuInSe_2$ phase, δ is a high-temperature (HT) phase with the sphalerite structure, and β is an ODC phase. It is interesting to note that the single phase field for $CuInSe_2$ at low temperatures is relatively narrow compared with earlier beliefs, and does not contain the composition 25% Cu. At higher temperatures, around 500 °C, where thin films are typically grown, the phase field widens toward the In-rich side. Typical average compositions of device-quality films have 22–24 at.% Cu, which fall within the single-phase region at growth temperature.

$CuInSe_2$ can be alloyed in any proportion with $CuGaSe_2$, thus forming $Cu(InGa)Se_2$. Similarly, the binary phase In_2Se_3 at the end point of the pseudobinary tie-line can be alloyed to

form $(InGa)_2Se_3$, although it undergoes a structural change at Ga/(In + Ga) = 0.6 [40]. In high-performance devices, Ga/(In + Ga) ratios are typically 0.2–0.3. In this chapter, Cu(InGa)Se_2 will be assumed to have composition in this range if not otherwise specified.

One of the central characteristics of Cu(InGa)Se_2 is its ability to accommodate large variations in composition without appreciable differences in optoelectronic properties. This tolerance is one of the cornerstones of the potential of Cu(InGa)Se_2 as a material for efficient low-cost PV modules. Solar cells with high performance can be fabricated with Cu/(In + Ga) ratios from 0.7 to nearly 1.0. This property can be understood from theoretical calculations that show that the defect complex $2V_{Cu} + In_{Cu}$, that is, two Cu vacancies with an In on Cu antisite defect, has very low formation energy, and also that it is expected to be electrically inactive [41]. Thus, the creation of such defect complexes can accommodate the Cu deficiency in Cu-poor/In-rich compositions of $CuInSe_2$ without adverse effects on the photovoltaic performance. Furthermore, crystallographic ordering of this defect complex is predicted [41], which explains the observed ODC phases $Cu_2In_4Se_7$, $CuIn_3Se_5$, $CuIn_5Se_8$, and so on.

The chalcopyrite phase field is increased by the addition of Ga or Na [42]. This can be explained by a reduced tendency to form the ordered defect compounds owing to higher formation energy for Ga_{Cu} (in $CuGaSe_2$) than for In_{Cu} (in $CuInSe_2$). This leads to destabilization of the $2V_{Cu} + In/Ga_{Cu}$ defect cluster related to the ODC phases [43, 44]. The effect of Na in the $CuInSe_2$ structure has been calculated by Wei *et al.* [45], with the result that Na replaces In_{Cu} antisite defects, reducing the density of compensating donors. In measurements of polycrystalline and epitaxial Cu(InGa)Se_2 films, Na is found to strongly reduce the concentration of compensating donors [44] and also to increase the net acceptor density [46]. Together with a tendency for Na to occupy Cu vacancies, the reduced tendency to form antisite defects also suppresses the formation of the ordered defect compounds. The calculated effect of Na is therefore consistent with the experimental observations of increased compositional range in which single-phase chalcopyrite exists and increased conductivity [45, 47].

13.2.2 Optical Properties and Electronic Structure

The absorption coefficient α for $CuInSe_2$ is very high, larger than 3×10^4/cm for 1.3 eV and higher photon energies [48–50]. Figure 13.6 shows the absorption coefficient of thin film Cu(InGa)Se_2 with $x \equiv Ga/(In+Ga) = 0$ and 0.3 and the total fraction of photons in the solar spectrum absorbed as a function of the film thickness. The high absorption coefficient means that 95% of the incident solar illumination is absorbed in films of only 1 μm thickness. In many studies it was found that the energy (E) dependence at the fundamental absorption edge near the bandgap E_g can be approximately described by

$$\alpha = A\sqrt{E - E_g}/E \tag{13.1}$$

as for a typical direct bandgap semiconductor, where the prefactor A depends on the density of states.

A rather complete picture of the optical properties, including bandgap values and other critical points of $CuInSe_2$, $CuGaSe_2$ and other Cu-ternary chalcopyrites is given in reference [48]. Spectroscopic ellipsometry measurements of single crystal samples were carried out and the dielectric functions were obtained. Similar studies of Cu(InGa)Se_2 having compositions from $0 \le x \le 1$ have been made on bulk polycrystalline ingots [49] and on polycrystalline thin films [50].

As for other semiconductor alloys the composition dependence of the bandgap can be fit to an empirical equation with a quadratic dependence on x. For the case of thin films [50] this is

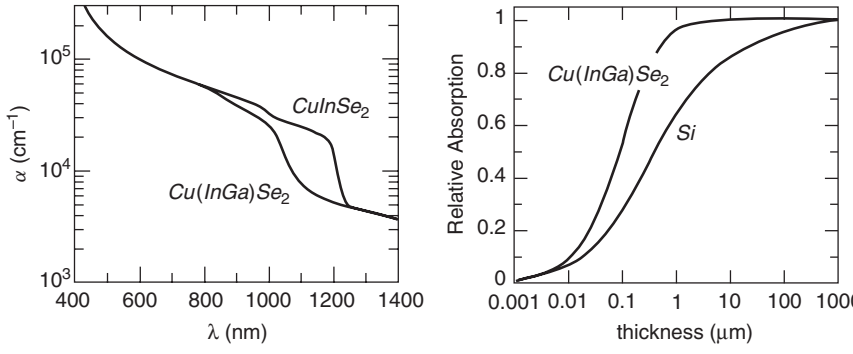


Figure 13.6 Absorption coefficient of $CuInSe_2$ and $Cu(InGa)Se_2$ with $x = 0.3$ and $E_g = 1.18$ eV [50] (left), and the total fraction of photons with $E > E_g$ in the AM1.5 solar spectrum absorbed in $Cu(InGa)Se_2$ as a function of thickness, compared to crystalline Si (right)

written as:

$$E_g = 1.04 + 0.65x - 0.26x(1 - x) \tag{13.2}$$

where the so-called bowing parameter b is 0.264, the bandgap of $CuInSe_2$ is 1.035 eV and $CuGaSe_2$ is 1.68 eV. A value of $b = 0.21$ was obtained by theoretical calculations compared with values in the range of 0.11–0.26 determined in various experiments [51].

The electronic structure of $CuInSe_2$ and other chalcopyrite semiconductors has been studied by optical measurements and by *ab initio* calculations. In chalcopyrite materials, the degeneracy of the valence band maximum (the Γ point) is generally removed. This is unlike the situation in more common semiconductors such as Si, III–V or II–VI compounds, where the heavy and the light hole bands are degenerate at the center of the Brillouin zone and only the third band is split off by spin–orbit coupling. Chalcopyrites have an additional crystal field splitting due to the tetragonal distortion [8]. One manifestation of this is different absorption edges for light polarized parallel or perpendicular to the c-axis of the crystal [36, 52].

A remarkable feature of all Cu chalcopyrites that makes them relevant for photovoltaics is that their bandgaps are much lower than their binary analogs such as ZnSe. This is due to the contribution of the Cu d-states and their hybridization with the p-states of the anion. This effect was implied by comparing experimental spin orbit and crystal field values obtained from the valence band splittings with the predictions of the pseudo-cubic model with calculated spin–orbit and crystal field splitting [8]. *Ab initio* calculations of the band structure confirmed that the bandgap lowering is mostly due to the contribution of Cu d-electron states to the valence band maximum [53].

A critical parameter of the electronic structure that affects transport properties is the effective mass, which describes the curvature of the bands at the Γ-point. A number of measurements have been made for the effective masses of $CuInSe_2$. Values of $0.1m_0$ are reported for the electron effective masses [54, 55] and $0.7m_0$ for the hole effective masses [47, 44]. These values are reproduced by recent refined local density approximation calculations, which further show a strong anisotropy of the hole effective masses with a factor of 4 difference between parallel to the c-axis and perpendicular to the c-axis [46]. $CuInSe_2$ has the lower effective hole mass perpendicular to the c-axis, whereas $CuGaSe_2$ has lower effective hole mass parallel to the c-axis. Polarized photoluminescence measurements have shown that only very little Ga changes the symmetry of the valence band from $CuInSe_2$-like to $CuGaSe_2$-like [52]. Solar cells are made from $Cu(InGa)Se_2$ with around 30% Ga in the III-sites. Therefore the effective hole mass in solar cells absorbers should be lower parallel to the c-axis.

13.2.3 Electronic Properties

$CuInSe_2$ with excess Cu is always *p*-type, but In-rich films can be made *p*-type or *n*-type depending on the Se content [56]. By annealing in a selenium overpressure, *n*-type material can be converted to *p*-type, and conversely, by annealing in a low selenium pressure, *p*-type material becomes *n*-type [57]. This contrasts with the behavior of $CuGaSe_2$, which is always p-type [58]. Device-quality Cu(InGa)Se_2 films, grown with the excess Se available, are *p*-type with a carrier concentration of about $10^{15}-10^{16}/cm^3$ [59]. There is a large spread in mobility values reported for $CuInSe_2$. The highest values of hole mobilities have been obtained for epitaxial films, where $200\,cm^2/Vs$ has been measured for Cu(InGa)Se_2 with about $10^{17}/cm^3$ in hole concentration [44], up to $250\,cm^2/Vs$ was found in pure $CuGaSe_2$ [60]. Bulk crystals have yielded values in the range $15 - 150\,cm^2/Vs$ with electron mobilities ranging from 90 to $900\,cm^2/Vs$ [57]. Conductivity and Hall effect measurements of polycrystalline thin film samples are made cross-grain, but for device operation through-the-grain values are more relevant, since individual grains may extend from the back contact to the interface of the junction. Thus a capacitance-based technique was used to determine mobilities in working solar cells, giving values of $5-20\,cm^2/Vs$ [61].

A large number of intrinsic defects are possible in the chalcopyrite structure. Accordingly, a number of electronic transitions have been observed by methods such as photoluminescence, temperature dependent Hall measurements, photoconductivity, photovoltage, optical absorption, and capacitance measurements. From the theoretical side there are local density approximation calculations available on the defect formation enthalpies and the defect energy levels. The comparison of photoluminescence measurements on bulk crystals, epitaxial films and polycrystalline films reveals more or less the same spectral features, indicating identical shallow defects in materials prepared under very different conditions. This observation suggests that $CuInSe_2$ and Cu(InGa)Se_2 are doped by native defects. Of the twelve native defects possible (three vacancies, three interstitials, and six antisites), those with the lowest formation energies are, depending on the composition of material, the Cu vacancy, the In or Ga vacancy, the Cu_{III} or III_{Cu} antisite and the Se vacancy [43, 45]. The Se vacancy has recently been identified as an amphoteric defect [62], which can describe the metastable effects observed in solar cells (see Section 13.5.2). The III_{Cu} antisite defect has been shown to be involved in a DX center [63], which could be responsible for doping limits in wide band gap Cu(InGa)Se_2.

The interpretation of numerous photoluminescence and Hall measurements was enabled by results on epitaxial films, showing that there are four dominating shallow defects including three acceptors and one donor [64–66]. The defect ionization energies are shown in Table 13.2. An important observation is that defect spectroscopy by photoluminescence is most useful in material grown under Cu-excess since the photoluminescence spectra of Cu-poor material are dominated by fluctuating potentials [67, 64] due to the high degree of compensation [68].

Many speculations have been proposed to correlate the observed defect energy levels with certain defect structures. Comparison with theoretical results cannot be used, because the experimental defect level energies are much smaller than those calculated and the trend between $CuInSe_2$ and $CuGaSe_2$ is reversed in some cases. Annealing experiments cannot be used in the case of ternary compounds, because this would require control of two of the constituents [69]. One would expect that electron spin resonance measurements allow a correlation between the electronic defect levels and defect structures, but only a broad resonance was found related to Cu^{2+} and a single completely isotropic peak that did not allow any correlation with a defect structure [70].

While photoluminescence and Hall measurements give only information about the shallow, doping defects, various methods of capacitance measurements have been applied to study the deep defects, further from the band edges. Two dominant deep defects appear in all Cu(InGa)Se_2 samples: one, sometimes labeled N2, at an energy of 250–300 meV from the valence band [71] and

Table 13.2 Experimentally observed defect levels in $CuInSe_2$ and $CuGaSe_2$

Defect	Energy position	Material
Acceptor 1	$E_V + 0.04$ eV	$CuInSe_2$
	$E_V + 0.06$ eV	$CuGaSe_2$
Acceptor 2	$E_V + 0.06$ eV	$CuInSe_2$
	$E_V + 0.10$ eV	$CuGaSe_2$
Acceptor 3	$E_V + 0.08$–0.09 eV	$CuInSe_2$
	$E_V + 0.13$–0.15 eV	$CuGaSe_2$
Donor 1	$E_C - 0.01$ eV	$CuInSe_2$
	$E_C - 0.01$ eV	$CuGaSe_2$
Deep 1	$E_V + 0.3$ eV	$CuInSe_2$
	$E_V + 0.3$ eV	$CuGaSe_2$
Deep 2	$E_V + 0.8$ eV	$CuInSe_2$
	$E_V + 0.8$ eV	$CuGaSe_2$

a deeper one at 800 meV from the valence band [72]. Because the N2 defect is far enough away from the midgap energy for all Ga/In ratios it does not present a serious problem in terms of a recombination center. However the 800 meV defect becomes a midgap defect for high Ga contents and therefore may play a major role as a recombination center in alloys with Ga contents greater than the standard absorber.

13.2.4 The Surface and Grain Boundaries

The surface morphology and grain structure are most commonly characterized by scanning electron microscopy (SEM), but transmission electron microscopy (TEM) and atomic force microscopy have also proved valuable. A typical SEM image is shown in Figure 13.7 and TEM cross-sectional image in Figure 13.2. In general, the films used in devices have grain diameters on the order of 1 μm, but the grain size and morphology can vary greatly, depending on fabrication method and conditions. A variety of structural defects including twins, dislocations, voids and stacking faults have been observed [73–76].

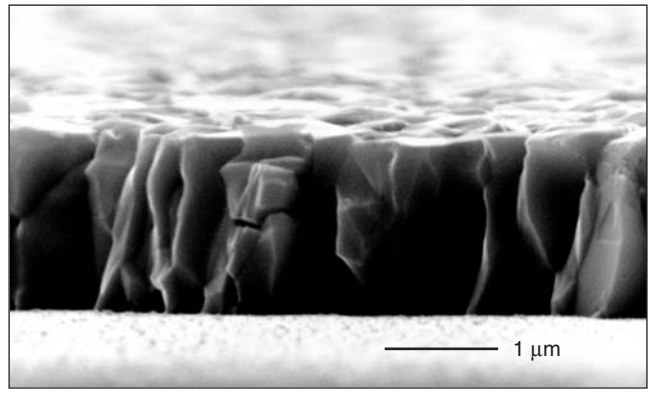


Figure 13.7 Scanning electron microscopy image of a typical $Cu(InGa)Se_2$ film deposited on a Mo-coated glass substrate by coevaporation

The Cu(InGa)Se_2 surface tends to facet to {112} growth planes [77]. It has been shown by X-ray photoelectron spectroscopy (XPS) that the free surfaces of $CuInSe_2$ films with slightly Cu-poor composition have a composition close to $CuIn_3Se_5$ [78], corresponding to one of the ordered defect phases. Attempts to identify such a layer on top of the films have been inconclusive. Instead, it seems that the composition gradually changes from the bulk to the surface of the films. Cu migration has been attributed to Fermi-level-dependent defect formation [79]. The surface Fermi level moves towards the conduction band due to surface defects, adsorbates or junction formation. Reduced Cu concentration has then been interpreted as the formation of Cu vacancies, which becomes more favorable with higher Fermi level [41]. Alternatively, it was proposed that band bending induced by surface charges drives electromigrating Cu into the bulk, leaving the surface depleted of Cu [42]. This depletion is stopped when the composition is that of $CuIn_3Se_5$, since further depletion requires a structural change of the material. Electromigration of Cu in $CuInSe_2$ has been demonstrated and also correlated with type conversion of the chalcopyrite material [80].

The band bending as well as the $CuIn_3Se_5$ composition of the $CuInSe_2$ surfaces disappears as oxides form on the surface when the material is exposed to atmosphere for some time. The surface oxidation is enhanced by the presence of Na [47]. Surface compounds that have been identified after oxidation include In_2O_3, Ga_2O_3, SeO_x, and Na_2CO_3 [81].

It has been common practice to post-treat completed Cu(InGa)Se_2 devices in air at typically 200 °C. When devices were fabricated using vacuum-evaporated CdS or (CdZn)S to form the junction, these anneals were often done for several hours to optimize the device performance [15, 82]. The main effect associated with oxygen is explained as passivation of selenium "surface" vacancies on the grains [83]. This model assumes that the donor-type V_{Se} at the grain boundaries can act as a recombination center. The positive charge associated with these donor-type defects reduces the effective hole concentration at the same time that the intergrain carrier transport is impeded. When oxygen substitutes for the missing selenium, these negative effects are canceled.

The overall noted beneficial effect of the presence of Na on the PV performance of Cu(InGa)Se_2 thin films lacks a complete explanation. It was proposed that the catalytic effect of Na on oxidation, by enhanced dissociation of molecular oxygen into atomic oxygen, makes the passivation of V_{Se} on grain surfaces more effective [84]. This model is consistent with the observation that Na and O are predominantly found at the grain boundaries rather than in the bulk of the grains in $CuInSe_2$ thin films [85]. However other investigations do not find any composition difference between grain interior and grain boundary [98, 99].

One of the remarkable properties of Cu(InGa)Se_2 solar cells is their insensitivity to grain size and morphology that leads to the conclusion that there is no significant recombination loss at grain boundaries. Since solar cells with polycrystalline absorbers outperform single-crystalline solar cells, it has even been argued that grain boundaries are actually beneficial for solar cells. However, there is still no definitive picture of the grain boundary behavior.

Two basic explanations for the lack of electronic losses at grain boundaries can be considered. In one case, the grain boundaries are inherently passivated with no electronic defects in the bandgap, possibly because of the effects of Na or O or because the ternary nature of $CuInSe_2$ shifts the defects correlated with grain boundary disorder into the bands [86]. In the other case, variations of the valence or conduction band edge create a barrier to transport of at least one carrier type so that there is no recombination. Theoretical calculations of surface formation have predicted a decrease in the valence band maximum associated with Cu vacancies [87]. However, attempts to measure compositional changes at the grain boundaries [88–90] have given mixed results. Alternatively, defects at the grain boundaries can trap majority carriers and lead to band bending and a space-charge region in the grain [91, 92].

A range of experimental techniques have been applied to study the band bending associated with grain boundaries in films with a wide variety of compositions, and deposition methods and are reviewed in Reference [93]. This includes transport measurements (temperature-dependent conductivity and Hall effect) in which the activation energy for cross-grain hole transport is related to the barrier, scanning tunneling techniques such as Kelvin probe force microscopy which can measure the work function variation across a single grain boundary, and scanning tunneling luminescence or cathodoluminescence. These measurements all generally indicate relatively small barriers, in the range 20–100 meV [93]. Thus the highest-efficiency devices may require both a transport barrier at the boundary and low density of electronically active defects at the grain surfaces.

13.2.5 Substrate Effects

The effects of the substrate on the properties of thin film polycrystalline $Cu(InGa)Se_2$ can be classified into three categories: (1) thermal expansion, (2) chemical effects, and (3) surface influence on nucleation. These each depend on the substrate material: glass or a flexible metal foil or plastic web.

It can be assumed that after growth, when the substrate and film are still at the growth or reaction temperature, the stress in the $Cu(InGa)Se_2$ film is low. Cooling down after growth imposes a temperature change of about 500 °C, and if the thermal expansion of the substrate and $Cu(InGa)Se_2$ is different stress will be built up in the film. The thermal expansion coefficient for $Cu(InGa)Se_2$ is around 9×10^{-6}/K in the temperature interval of interest, which is similar to that of soda-lime glass. A $CuInSe_2$ film deposited on a substrate with a lower thermal expansion coefficient, such as borosilicate glass, will be under increasing tensile stress during cool-down. Typically, such films exhibit voids and microcracks [74]. When the thermal expansion coefficient of the substrate is higher than that of the film material, like for polyimide, it will result in compressive stress in the thin film material, which may lead to adhesion failures.

The most important effect of the soda-lime glass substrate on $Cu(InGa)Se_2$ film growth is that it supplies sodium to the growing chalcopyrite material. This effect is distinct from the thermal expansion match of soda-lime glass [94]. The sodium diffuses through the Mo back contact, which also means that it is important to control the properties of the Mo [95]. The resulting microstructure of $Cu(InGa)Se_2$ is influenced by the presence of Na with larger grains and a higher degree of preferred orientation, with the (112) plane parallel to the substrate. An explanation for this effect when high concentrations of Na are present has been proposed by Wei *et al.* [45].

The preferred orientation of the $Cu(InGa)Se_2$ can vary greatly with different growth processes or substrates. While the highest efficiency devices are reported to have a (220)/(204) preferred orientation [96], controlled experiments to show a direct link between orientation and device behavior are inconclusive. The variation is controlled by the surface on which the chalcopyrite material nucleates. A comparison between $Cu(InGa)Se_2$ grown on normal Mo-coated substrates and directly on soda-lime glass shows a much more pronounced (112) orientation on glass, yet no difference in the Na concentration measured in the films on the two substrates [19]. Further, the preferred orientation of the $Cu(InGa)Se_2$ film has been shown to be directly correlated to the orientation of the Mo film [97].

13.3 DEPOSITION METHODS

A wide variety of deposition methods has been used to prepare $Cu(InGa)Se_2$ thin films. To determine the most promising technique for the commercial manufacture of modules, the overriding criteria are that the deposition can be completed at low cost while maintaining high deposition or processing rate

with high yield and reproducibility. Compositional uniformity over large areas is necessary for high yield. Device considerations dictate that the Cu(InGa)Se_2 layer should be at least 1 μm thick to fully absorb incident light, as shown in Figure 13.6, and that the relative compositions of the constituents are kept within the bounds determined by the phase diagram, as discussed in Section 13.2.1. For solar cell or module fabrication, the Cu(InGa)Se_2 is deposited on a molybdenum-coated glass or flexible substrate.

The most promising deposition methods for the commercial manufacture of modules can be divided into two general approaches that have both been used to demonstrate high device efficiencies and are now being used for manufacturing. The first approach is physical vapor deposition in which all the constituents – Cu, In, Ga, and Se – can be simultaneously delivered to a substrate heated to 450–600 °C and the Cu(InGa)Se_2 film is formed in a single growth process. This is usually achieved by thermal coevaporation from elemental sources at temperatures greater than 1000 °C for Cu, In, and Ga.

The second approach is a two-step process that separates the delivery of the metals from the reaction to form device-quality films. Typically a precursor containing the Cu, Ga, and In is deposited using low-cost and low-temperature methods that facilitate uniform composition. Then the films are annealed in a Se and/or S atmosphere, also at 450–600 °C. The reaction and anneal step may take longer time than formation of films by coevaporation due to reaction and diffusion kinetics, but is amenable to batch processing. High process rate can be achieved by moving continuously through sequential process steps or with a batch process whereby longer deposition or reaction steps can be implemented by handling many substrates in parallel.

13.3.1 Substrates and Sodium Addition

Soda-lime glass, which is used in conventional windows, is a common substrate material used for Cu(InGa)Se_2 since it is available in large quantities at low cost and has been used to make the highest-efficiency devices. Cu(InGa)Se_2 deposition requires a substrate temperature T_{SS} of at least 350 °C and the highest-efficiency cells have been fabricated using films deposited at the maximum temperature, $T_{SS} \approx 550$ °C, which the glass substrate can withstand without softening too much [98]. Soda-lime glass has a thermal expansion coefficient of 9×10^{-6}/K [98], which provides a good match to the Cu(InGa)Se_2 films. The glass composition typically includes various oxides such as Na_2O, K_2O, and CaO. These provide sources of alkali impurities that diffuse into the Mo and Cu(InGa)Se_2 films during processing [19], producing the beneficial effects discussed in Section 13.2.

Cu(InGa)Se_2 can also be deposited on flexible substrates that may provide manufacturing advantages by allowing continuous roll-to-roll processes and enable fabrication of flexible, lightweight PV products [99, 100]. Light weight is particularly important for space applications. High-temperature polymer is an appealing substrate choice [20] which provides the lightest weight and has the advantage of being electrically insulating to facilitate monolithic interconnection for module fabrication (see Section 13.6.2). Polyimide web currently available commercially can be processed at temperatures up to ~425 °C so the cell efficiency is typically lower than on glass. Metal foils can withstand higher temperatures, but are electrically conductive and in some cases are reactive with the Se in the Cu(InGa)Se_2 process. While a variety of metal foil substrates have been used [99], stainless steel is the most common and has produced the highest efficiency flexible Cu(InGa)Se_2 cells [96].

With any flexible substrate Na must be actively provided to the Cu(InGa)Se_2 since there is no diffusion from the substrate. Even with soda-lime glass, a process that provides a more controllable supply of Na than diffusion from the substrate may be preferred to improve uniformity

and manufacturing yield. This can be achieved by blocking sodium from the substrate with a diffusion barrier such as SiO_x, Al_2O_3, or SiN [101, 102]. Then on the barrier-coated glass or flexible substrate Na can be directly provided to the Cu(InGa)Se_2 growth process by depositing a sodium-containing precursor layer, typically NaF with thickness of the order of 10 nm, onto the Mo film [103]. Alternatively, Na can be co-deposited with the Cu(InGa)Se_2 [104]. Even a post-deposition Na treatment gave the same increased cell performance as Na provided during deposition [105], indicating that the benefits of Na are related to material or surface properties and not film growth.

13.3.2 Back Contact

The Mo back contact, used for high-efficiency devices, is typically deposited by direct current (dc) sputtering. The thickness is determined by the resistance requirements that depend on the specific cell or module configuration. A film with thickness 1 μm will typically have a sheet resistance of 0.1–0.2 Ω/□, a factor of 2–4 higher resistivity than bulk Mo. Sputter deposition of the Mo layer requires careful control of the pressure to control stress in the film [106] and to prevent problems such as poor adhesion that stress might cause. During Cu(InGa)Se_2 deposition, a $MoSe_2$ layer forms at the interface [107]. Its properties are influenced by the Mo film: less $MoSe_2$ forms on dense Mo, sputter-deposited under low pressures [75]. This interfacial layer does not degrade device performance and may even promote formation of an ohmic contact. Metals other than Mo have been investigated and comparable cell performance attained with Ta and W [108].

13.3.3 Coevaporation of Cu(InGa)Se_2

The highest-efficiency devices have been deposited by thermal coevaporation from elemental sources. An illustration of a laboratory system for Cu(InGa)Se_2 coevaporation is shown in Figure 13.8. The process uses line-of-sight delivery of the Cu, In, Ga, and Se from Knudsen-type effusion cells or open-boat sources to the heated substrate [109]. While the evaporation temperatures for each metal will depend on the specific source design, typical ranges are 1300–1400 °C for Cu, 1000–1100 °C for In, 1150–1250 °C for Ga, and 250–350 °C for Se evaporation.

The sticking coefficients of Cu, In, and Ga are very high, so the film composition and growth rate are determined simply by the flux distribution and effusion rate from each source.

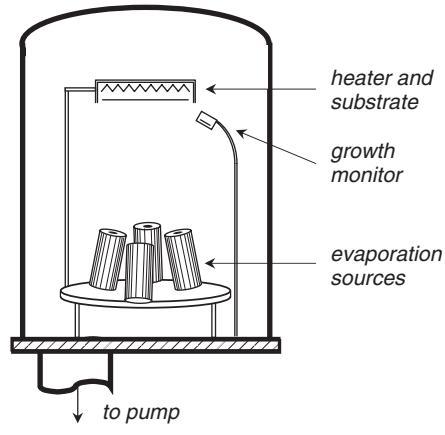


Figure 13.8 Laboratory-scale configuration for multisource elemental coevaporation

As long as there is sufficient Se availability, the composition of the final film tends to follow the pseudo-binary tie-line between $(InGa)_2Se_3$ and Cu_2Se (see Figure 13.4) according to the relative concentration of Cu compared with In and Ga. The relative concentrations of In and Ga determine the bandgap of the film, according to Equation (13.2), and the effusion rates can be varied over the course of a deposition to change the film composition through its thickness. Se has a much higher vapor pressure and lower sticking coefficient, so it is always evaporated in excess of that needed in the final film. Insufficient Se can result in a loss of In and Ga in the form of In_2Se or Ga_2Se [110].

Different deposition variations, using elemental fluxes deliberately varied over time, have been explored using coevaporation. Four different sequences that have been used to fabricate devices with efficiencies greater than 16% are shown in Figure 13.9. In each case, the targeted final composition is Cu-deficient with $Cu/(In + Ga) = 0.8–0.9$. Typical deposition rates vary from 20 to 200 nm/min, depending on the effusion rates from the sources. So, for a film thickness of 2 μm, the total deposition time may vary from 10 to 90 min.

The first process is the simplest homogeneous process in which all fluxes are constant throughout the deposition process [111]. In most cases, however, the fluxes are varied so that the bulk of the film is grown with Cu-rich overall composition and contains a Cu_xSe phase in addition to Cu(InGa)Se_2 [16]. The evaporation sources are then adjusted to finish the deposition with In- and Ga-rich flux so that the final film composition has the desired Cu-deficient composition. One modification of this is the second process shown in Figure 13.9. This process was first implemented with $CuInSe_2$ films deposited on non-Na-containing substrates at $T_{SS} = 450\,^{\circ}C$, producing films with increased grain size and improved device performance. The effect of Cu_xSe as a flux for enhanced grain growth at higher T_{SS} was proposed by Klenk *et al*. [112]. However, in devices containing Na and Ga and with $T_{SS} > 500\,^{\circ}C$, no difference was found in the device performance using films with Cu-rich or uniform growth processes [111].

The third process shown in Figure 13.9 is a sequential process in which the In and Ga are deposited separately from the Cu. This was first proposed by Kessler *et al.* [113] with the deposition of an $(InGa)_xSe_y$ compound, followed by the deposition of Cu and Se until the growing film reaches the desired composition. The layers interdiffuse to form the Cu(InGa)Se_2 film. A modification by Gabor *et al*. [114] allows the Cu delivery to continue until the film has an overall Cu-rich composition. Then a third step is added to the process in which In and Ga, again in the presence of excess Se, are evaporated to bring the composition back to Cu-deficient. The metals interdiffuse, forming the ternary chalcopyrite film. This process has been used to produce the highest-efficiency devices [1]. The improved device performance has been attributed to a bandgap gradient, which results from the Ga concentration decreasing from the Mo back contact toward the film's free surface and then increasing again in the top few tenths of a micrometer [21]. Improved performance has also been attributed to larger grain size of the films [115].

The last process shown in Figure 13.9 is an in-line process in which the flux distribution results from the substrate moving sequentially over the constantly effusing Cu, Ga, and In sources. This was first simulated in a stationary evaporation system [116] and has subsequently been implemented by several companies in manufacturing systems.

Various modifications to the evaporation process that may provide some advantage under certain deposition conditions include modified surface termination with lower Ga content [1], addition of H_2O vapor [117] or optimized control of the Se flux [118] during deposition.

A reproducible coevaporation process requires good control of the elemental fluxes from each evaporation source and development of reliable diagnostic tools is critical [119]. While the evaporation rates from each source can be controlled simply by the source temperature, this may not give good reproducibility, especially for the Cu source that is at the highest temperature. Open-boat sources in particular will not give reproducible evaporation rates, since they depend critically on the

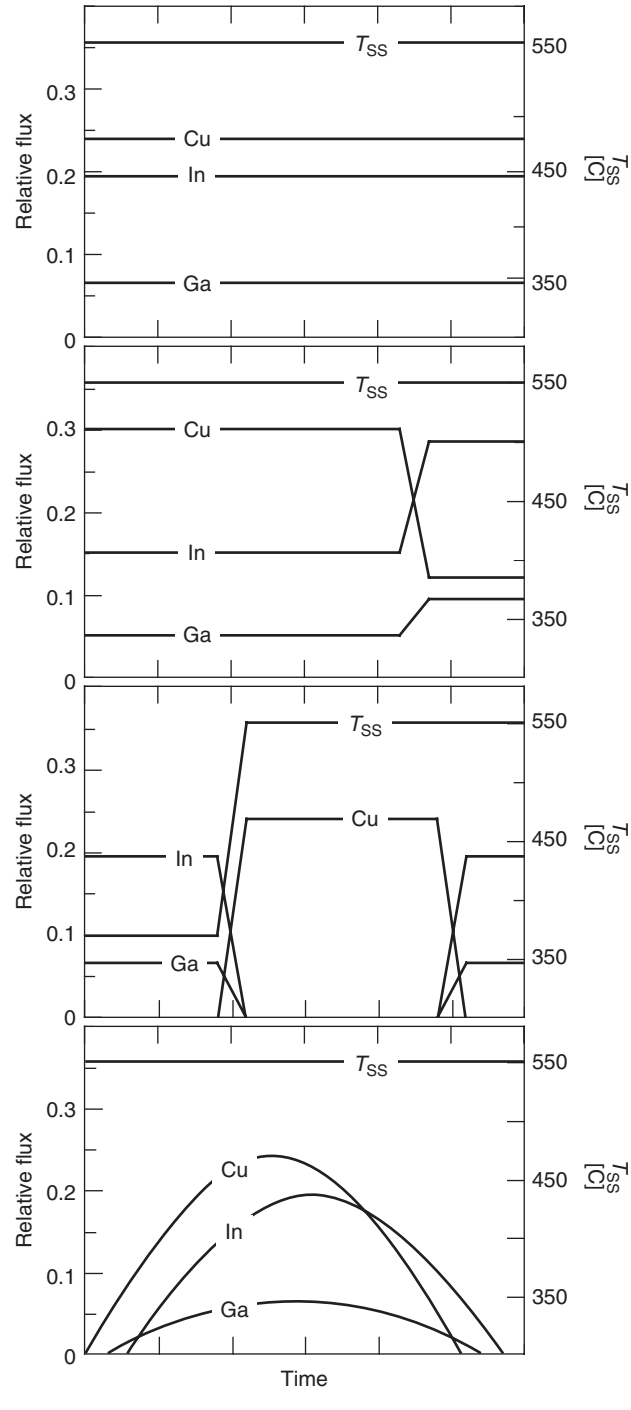


Figure 13.9 Relative metal fluxes and substrate temperature for different coevaporation processes. In all cases, a constant Se flux is also delivered

filling level and are not radiatively insulated from the rest of the deposition system. Consequently, direct *in situ* measurement of the fluxes is often used to control the evaporation sources. Electron impact spectroscopy [16], mass spectroscopy [120], and atomic absorption spectroscopy [121] have all been successfully implemented. Direct flux measurement may be more valuable in a manufacturing scale process, particularly if source depletion over long run times causes the relation between source temperature and effusion rate to vary over time. In addition, the process can be monitored by *in situ* film thickness measurement using a quartz crystal monitor, or optical spectroscopy or X-ray fluorescence of the growing film [122]. The latter can also be used to measure composition. When the process includes a transition from Cu-rich to Cu-poor composition near the end of the deposition, it can be monitored by a change in the film structure using laser light scattering [123] a change in the emissivity or temperature of the film [124] or by the infrared transmission [125].

The primary advantage of elemental coevaporation for depositing Cu(InGa)Se_2 films is its considerable flexibility to choose the process specifics and to control film composition and bandgap. As proof of this flexibility, high-efficiency devices have been demonstrated using many process variations. The primary disadvantage results from the difficulty in control, particularly of the Cu-evaporation source, and the resulting need for improved deposition, diagnostic, and control technology. There has been a lack of commercially available equipment for large-area thermal evaporation, especially the evaporation sources for Cu, Ga, and In, but recently several equipment companies are offering systems specifically designed for Cu(InGa)Se_2 deposition.

13.3.4 Precursor Reaction Processes

The second common approach to Cu(InGa)Se_2 film formation is a two-step process in which a precursor film containing Cu, In, and Ga is deposited by any of several methods and then reacted at high temperature to form Cu(InGa)Se_2. This is sometimes referred to selenization though S is also added in many cases as described below. This approach was first demonstrated by Grindle *et al*. [126] who sputtered Cu/In layers and reacted them in hydrogen sulfide to form $CuInS_2$. This was adapted to $CuInSe_2$ by Chu *et al*. [127]. The highest-efficiency Cu(InGa)Se_2 cell reported using precursor reaction is 16.5%, [128], but there has been far less effort at optimizing laboratory-scale cell efficiencies than with coevaporated Cu(InGa)Se_2. The highest reported efficiency for any monolithically-interconnected large-area Cu(InGa)Se_2 modules has been produced using the reaction of sputtered precursors by Showa Shell [129].

The precursor, which contains the Cu, Ga, In, and in some cases Se or S, determines the final composition of the film. Different processes are chosen for their low cost, including equipment costs and materials utilization, spatial uniformity, and application or deposition rate. Sputtering is an attractive process because it is easily scalable using commercially available deposition equipment. It can provide good uniformity over large areas with high deposition rates, but materials utilization can be a concern. Electrodeposition can provide very high materials utilization at low cost. Application of particles in an ink or spray can also provide high utilization and uniformity. All of these methods are being developed for commercial manufacturing.

Reaction of the precursor films to form Cu(InGa)Se_2 can be done in H_2Se at 400–500 °C for times up to 60 min. Poor adhesion [130] and excessive formation of a $MoSe_2$ layer [131] at the Mo/Cu(InGa)Se_2 interface may limit the reaction time and temperature. Reaction in H_2Se has the advantage that it can be done at atmospheric pressure and can be precisely controlled, but the gas is highly toxic and requires special precautions for its use. The precursor films can also be reacted in a Se vapor, which might be obtained by thermal evaporation, to form the Cu(InGa)Se_2 film [132]. A third reaction approach is rapid thermal processing (RTP) of either elemental layers, including Se, [133, 134] or amorphous evaporated Cu–In–Ga–Se layers [135]. Reaction using

diethylselenium, which provides the advantage of atmospheric reaction with a less toxic gas than H_2Se, has also been demonstrated [136].

The reaction chemistry and kinetics for the conversion of Cu–In precursors to $CuInSe_2$ has been characterized by X-ray diffraction (XRD) of time-progressive reactions [137] and by *in situ* differential scanning calorimetry [138]. The reaction path was shown to be the same for the reaction of Cu/In layers in either H_2Se or elemental Se [139].

In situ XRD measurements were used to study the reaction chemistry and kinetics of stacked Cu-In-Se layers in an RTP process [140]. In these experiments, $CuInSe_2$ formation follows a sequence of reactions starting with the formation of $CuIn_2$ and $Cu_{11}In_9$ intermetallic compounds. These react with Se as the temperature increases to form a series of binary selenide compounds. The formation of $CuInSe_2$ then follows from two reactions:

$$CuSe + InSe \rightarrow CuInSe_2$$
$$\tfrac{1}{2}Cu_2Se + InSe + \tfrac{1}{2}Se \rightarrow CuInSe_2.$$

The reaction is completed in ~10 minutes at 400 °C, with the rates dependent on the Se concentration and the availability of Na. When Ga was added to the stacked metal precursor a third reaction to form a chalcopyrite phase was noted:

$$\tfrac{1}{2}Cu_2Se + \tfrac{1}{2}Ga_2Se_3 \rightarrow CuGaSe_2,$$

although the formation of $CuGaSe_2$ is slower than $CuInSe_2$. Finally, formation of $Cu(InGa)Se_2$ in this case occurs by the intermixing of $CuInSe_2$ and $CuGaSe_2$.

The slower formation of $CuGaSe_2$, which is attributed to the greater stability of gallium selenide compounds compared to indium compound [141], leads to the formation of a composition gradient in the reacted film. Ga in the reacted film commonly accumulates near the Mo forming a $CuInSe_2/CuGaSe_2$ structure, so the resulting device behaves like $CuInSe_2$ [142] and lacks the increased operating voltage and other benefits of a wider bandgap discussed in Section 13.5.4. Nevertheless, Ga inclusion provides improved adhesion of the $CuInSe_2$ film to the Mo back contact and greater device performance, possibly owing to an improved structure with fewer defects. The Ga and In can be effectively interdiffused, converting the films to uniform bandgap and increasing V_{OC} in devices, by annealing in an inert atmosphere for 1 h at 600 °C [143], but this anneal may be impractical for commercial processing.

Films in the best devices formed by precursor reaction have the bandgap increased by the incorporation of S near the front surface, forming a graded $Cu(InGa)(SeS)_2$ layer [22, 144] that can give enhanced V_{OC}. A two-stage reaction process reported by Showa Shell includes a partial reaction of a Cu–Ga–In precursor in H_2Se at 450 °C for ~20 minutes followed by reaction in H_2S at 480 °C for ~15 minutes [144]. It was further shown that, by varying the time and temperature profiles of the two stages with the second stage temperature increased above 480 °C, that the Ga can be homogeneously distributed through the film [145]. A critical aspect of this process is that the H_2Se reaction stage does not fully convert the film to a chalcopyrite phase and the film still contains a Ga rich intermetallic phase [146]. Figure 13.10 shows compositional depth profiles measured by Auger electron spectroscopy of Cu–Ga–In precursor films reacted in H_2Se at 450 °C for 30 or 15 minutes then H_2S at 550 °C for 15 minutes [147]. With the longer H_2Se reaction, the precursor is almost fully reacted and the Ga is segregated to the back of the film (near the Mo) similar to films that are only reacted in H_2Se. With the shorter, incomplete H_2Se reaction the Ga is distributed throughout the film after the high-temperature H_2S reaction. In both cases, there is also a S gradient near the free surface of the film due to diffusion during the H_2S reaction. The higher temperature for the H_2S reaction stage also enables the reaction to be completed more quickly

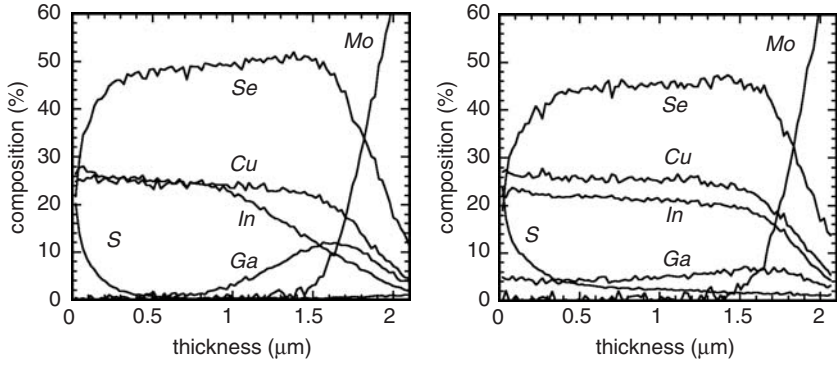


Figure 13.10 Compositional depth profiles of Cu–Ga–In films reacted in H_2Se at 450 °C for 30 minutes (left) or 15 minutes (right) and then reacted in H_2S for 15 minutes at 550 °C. (After Hanket G, Shafarman W, Birkmire R, *Proc. 4th World Conf. Photovoltaic Solar Energy Conversion*, pp 560–563 (2006))

and to produce better quality material, but without adhesion problems that would likely occur with higher-temperature H_2Se reaction. However, large-area substrates in the manufacturing scale used by Showa Shell and others may limit the reaction temperature.

The primary advantages of precursor reaction processes for Cu(InGa)Se_2 deposition are the ability to utilize more standard and well-established techniques for the metal deposition and reaction and anneal steps and to compensate for long reaction times with a batch processing mode or RTP of Se-containing precursors. Overall composition and uniformity are controlled by the precursor deposition and can be measured between the two steps as a process control. The biggest drawbacks to these processes, the limited ability to control through-film composition and poor adhesion, can be largely overcome by a two-stage reaction process. This also enables a controlled S-profile near the front of the device which widens the bandgap and increases V_{OC}. H_2Se and, to a lesser extent, H_2S used in many of these processes are hazardous gases that may be costly to handle.

13.3.5 Other Deposition Approaches

$CuInSe_2$-based films have been deposited using a wide range of thin film deposition methods, in addition to those discussed above, which have been proposed as potential low-cost alternatives for manufacturing. These include reactive sputtering [148], hybrid sputtering in which Cu, In, and Ga are sputtered while Se is evaporated [149], closed-space sublimation [150], chemical bath deposition (CBD) [151], laser evaporation [152], and spray pyrolysis [153]. Great effort was made to explore different thin film deposition techniques before coevaporation and the two-step processes above became dominant. These early methods are reviewed in reference [27].

13.4 JUNCTION AND DEVICE FORMATION

The first experimental device that indicated the potential for $CuInSe_2$ in high-performance solar cells was a heterojunction between a *p*-type single-crystal of $CuInSe_2$ and a thin film of *n*-type CdS [11, 12]. Consequently, from the earliest thin film work the junction was formed by depositing CdS on the $CuInSe_2$ films [154]. The device was further developed to contain an undoped layer of

CdS, followed by CdS doped with In, both deposited by vacuum evaporation [15]. This defined the device structure (see Figure 13.1), which is basically the same as is commonly used today since the doped CdS is functionally a transparent conductor. Recognizing that there was a significant loss in current because light absorbed in the CdS was not collected, a performance gain was achieved by alloying the CdS with ZnS to widen the bandgap [16]. Further improvement of the performance was achieved when the doped CdS layer was replaced with doped ZnO [155, 156]. The undoped CdS layer adjacent to the Cu(InGa)Se_2 film was reduced in thickness in order to maximize the optical transmission. Since ZnO has a wider bandgap than CdS, more light is transmitted into the active part of the device, resulting in a current gain. A conformal and pinhole-free coating of this thin CdS layer is obtained by using chemical bath deposition to make the CdS buffer layer.

13.4.1 Chemical Bath Deposition

Chemical bath deposition (CBD) of thin film materials, also referred to as solution growth, has been used in particular for chalcogenide materials such as PbS [157], CdS [158], and CdSe [159]. A variety of precursor compounds or ions can be used to deposit a specific compound. CBD of CdS buffer layers on Cu(InGa)Se_2 is generally made in an alkaline aqueous solution (pH > 9) of the following three constituents:

1. a cadmium salt; for example, $CdSO_4$, $CdCl_2$, CdI_2, $Cd(CH_3COO)_2$.
2. a complexing agent; commonly NH_3 (ammonia).
3. a sulfur precursor; commonly $SC(NH_2)_2$ (thiourea).

The concentrations of the various components of the solution can be varied over a range and each laboratory tends to use its own specific recipe. One example of a recipe that is being used to fabricate state-of-the-art Cu(InGa)Se_2 solar cells is:

1. $1.4 \times 10^{-3} M$ CdI_2 or $CdSO_4$.
2. $1\,M$ NH_3.
3. $0.14\,M$ $SC(NH_2)_2$.

The Cu(InGa)Se_2 film is immersed in a bath containing the solution and the deposition takes place in a few minutes at a temperature of 60–80 °C. This can be done either by immersion in a room-temperature bath that subsequently is heated to the desired temperature or by preheating the solution. The reaction proceeds according to the formula

$$Cd(NH_3)_4{}^{2+} + SC(NH_2)_2 + 2OH^- \rightarrow CdS + H_2NCN + 4\,NH_3 + 2\,H_2O$$

In practice, the chemical bath deposition is typically done in the laboratory with a very simple apparatus consisting of a hot plate with magnetic stirring, a beaker holding the solutions into which the substrate is immersed, and a thermocouple to measure bath temperature. A typical arrangement, incorporating a water bath for more uniform temperature, is shown in Figure 13.11. Scale-up of the CBD process for manufacturing is discussed in Section 13.6.1.

The growth of CdS thin films by CBD occurs from ion-by-ion reaction or by clustering of colloidal particles. Depending on the bath condition, the resulting CdS lattice structure may be cubic, hexagonal, or a mixture [160]. Under typical conditions used for Cu(InGa)Se_2 solar cells, the relatively thin CdS layers grow ion-by-ion, resulting in dense homogeneous films [161] with mixed cubic/hexagonal or predominantly hexagonal lattice structure [75, 162, 163]. The films consist of crystallites with a grain size of the order of tens of nanometers [162].

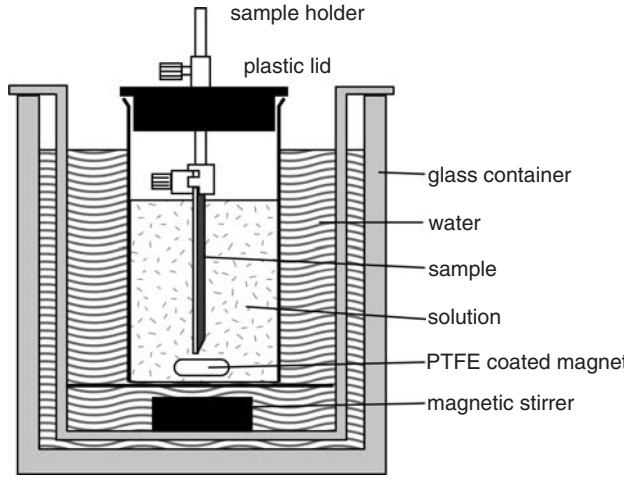


Figure 13.11 Typical laboratory apparatus for chemical bath deposition of CdS

Compositional deviation from stoichiometry is commonly observed. In particular, films tend to be sulfur-deficient and contain substantial amounts of oxygen [164, 165]. In addition, significant concentrations of hydrogen, carbon, and nitrogen have also been detected in device-quality films [166]. The concentration of these impurities has been correlated to a reduction of the optical bandgap and the amount of cubic CdS in relation to hexagonal CdS [167].

13.4.2 Interface Effects

The interface between the Cu(InGa)Se_2 and the CdS is characterized by pseudo-epitaxial growth of the CdS and intermixing of the chemical species. Electronic band alignment will be discussed in Section 13.5.3. Transmission electron microscopy has shown that chemical bath deposited CdS layers on Cu(InGa)Se_2 films exhibit an epitaxial relationship at the interface with (112) chalcopyrite Cu(InGa)Se_2 planes parallel to the (111) cubic or (002) hexagonal CdS planes [75, 163]. The lattice mismatch is very small for pure $CuInSe_2$ with a (112) spacing of 0.334 nm compared with a spacing of 0.336 nm for (111) cubic and (002) hexagonal CdS. The lattice spacing between the planes relates directly to the lattice spacing within the planes in these tetrahedrally bonded materials. In Cu(InGa)Se_2 the lattice mismatch increases with the Ga content. $CuIn_{0.7}Ga_{0.3}Se_2$ and $CuIn_{0.5}Ga_{0.5}Se_2$ have (112) spacings of 0.331 and 0.328 nm, respectively.

When Cu(InGa)Se_2 films are immersed in the chemical bath for deposition of CdS, they are also subjected to chemical etching of the surface. In particular, native oxides are removed by the ammonia [168]. Thus, the CBD process cleans the Cu(InGa)Se_2 surface and enables the epitaxial growth of the CdS buffer layer.

In early single-crystal work, p–n homojunction diodes were fabricated by indiffusion of Cd or Zn into p-type $CuInSe_2$ [169, 170] at 200–450 °C. Investigations of $CuInSe_2$/CdS interfaces did show interdiffusion of S and Se above 150 °C and rapid Cd diffusion into $CuInSe_2$ above 350 °C [171]. More recently, intermixing of the constituents of the Cu(InGa)Se_2/CdS heterojunction has been observed, even when the relatively low-temperature CBD process is used for growth of

the CdS layer [172]. Investigations of the effect of a chemical bath without the thiourea showed an accumulation of Cd on the Cu(InGa)Se_2 surface, possibly as CdSe [168]. Accumulation of Cd on the Cu(InGa)Se_2 surface was also observed in the initial stage of CdS growth in the complete chemical bath [173]. The results were not conclusive as to whether any interfacial compound is formed, but TEM investigations showed the presence of Cd up to 10 nm into the Cu-deficient surface region of the Cu(InGa)Se_2 layer [163]. At the same time, a reduction of the Cu concentration was noted. An interpretation in which Cu^+ is replaced with Cd^{2+} is proposed, on the basis of the very close ion radii of these species, 0.96 and 0.97 Å, respectively. XPS and secondary ion mass spectrometry (SIMS) profiles of Cu(InGa)Se_2 films and $CuInSe_2$ single-crystals exposed to chemical baths without thiourea also show evidence of indiffusion or electromigration of Cd [174].

13.4.3 Other Deposition Methods

Vacuum evaporation of 2 to 3-μm-thick CdS was commonly used in early development but it is difficult to nucleate and grow very thin continuous CdS layers such as those normally used in current state-of-the-art Cu(InGa)Se_2 devices. Substrate temperatures of 150 to 200 °C used to obtain good optical and electrical properties of evaporated CdS films are substantially higher than the substrate temperature used for chemical bath deposition.

Sputter deposition leads to more conformal coating of the relatively rough Cu(InGa)Se_2 films compared to evaporation. The general success of sputtering for industrial large-area deposition motivated the exploration of sputter-deposited CdS buffer layers. Using optical emission spectroscopy to control the sputtering process, Cu(InGa)Se_2 devices with efficiencies up to 12.1% were fabricated, compared with 12.9% for reference cells with chemical bath-deposited CdS [175]. Both evaporation and sputtering are vacuum processes, which can be incorporated in-line with other vacuum processing steps and do not create any liquid wastes. Still, CBD remains the preferred process for the CdS layer owing to its advantages in forming thin conformal coatings.

Atomic layer deposition (ALD) is a variation of chemical vapor deposition (CVD) that also allows accurate control of the growth of thin conformal layers [176]. The method is being industrially used for deposition of another II–VI compound, ZnS. Inorganic precursors for deposition of CdS require the substrate temperature to be excessively high (>300 °C) and work with organic precursors has been limited. The strong driving force for replacement of the environmentally undesirable cadmium has focused the development of ALD on materials other than CdS. This is also valid for more conventional CVD, although some metal–organic CVD (MOCVD) work has been reported. The full potential for chemical vapor deposited CdS has therefore not been explored.

Electrodeposition can be used to deposit CdS films but its use has not been reported in Cu(InGa)Se_2 devices.

13.4.4 Alternative Buffer Layers

The cadmium content in Cu(InGa)Se_2 PV modules with CBD CdS buffer layers is low. Investigations show that the cadmium in Cu(InGa)Se_2 modules can be handled safely, both with respect to environmental concerns and hazards during manufacturing (see Section 13.6.5). However, regulation of Cd usage in electronic products is becoming stricter so interest in Cd-free devices is rising. Another reason to use alternative buffers is the gain in short-circuit current, that can be attained

when using buffers with a wider bandgap than CdS. There are in principle two approaches to Cd-free devices: (1) finding a buffer material that replaces CdS and (2) omitting the CdS layer and depositing ZnO directly onto the Cu(InGa)Se_2 film. In practice, the two approaches tend to merge when the chemical bath deposition of CdS is replaced with a surface treatment of the Cu(InGa)Se_2 with no or negligible film deposition before the subsequent deposition of the ZnO. Generally, very good efficiencies have been achieved with simple omission of the buffer although poor reproducibility may result [174, 177].

A number of approaches and materials have been tried for the deposition of alternative buffers, and a selection of promising results is presented in Table 13.3. CBD, used for CdS, has also been applied to various other materials. Other methods based on chemical reactions, typically provide excellent coverage for very thin layers, including ALD, MOCVD, and ion layer gas reaction (ILGAR). For some In-containing compounds however physical vapor deposition methods such as evaporation or sputtering have also provided good results. The materials used usually have a larger bandgap than CdS and thus allow for higher transmission of light to the absorber. Care is taken to choose a material with smaller electron affinity than the absorber, to avoid the formation of a cliff (i.e. buffer CB lower than absorber CB) in the conduction band (see Section 13.5.3).

When evaluating the values in Table 13.3, one must keep in mind that the quality of the Cu(InGa)Se_2 layer varies significantly between the experiments, some are used for record cells with

Table 13.3 Performance of Cu(InGa)Se_2 thin film solar cells with various buffer layers and junction-formation methods alternative to chemical bath deposition of CdS

			Solar cell results					
Material	E_g (eV)	Deposition method	η (%)	J_{SC} (mA/cm^2)	V_{OC} (V)	FF (%)	area (cm^2)	Reference and comments
CdS	2.4	CBD	20.0	35.5	0.690	81	<1	[1]
Zn(S,O,OH)	3.0–3.8	CBD	18.6	36.1	0.661	78	<1	[96]
			15.2	36.2	0.601*	70	900	[129] module
		ILGAR	14.2	35.9	0.559	71	<1	[178]
		ALD	16.0	32.0	0.684	73	<1	[179]
Zn(Se,O,OH)	2.0–2.7	CBD	14.4	33.9	0.583	73	<1	[180]
			11.7	36.5	0.508*	63	20	[181] minimodule
		MOCVD	13.4	34.3	0.551	71	<1	[182]
		ALD	11.6	35.2	0.502	65	<1	[183]
(Zn,In)Se	2.0	Evaporation	15.1	30.4	0.652	76	<1	[184]
$In(OH)_3$	5.1	CBD	14.0	32.1	0.575	76	<1	[185]
In(OH,S)	2.0–3.7	CBD	15.7	35.5	0.594	75	<1	[186]
In_2S_3	2.7	ALD	16.4	31.5	0.665	78	<1	[187]
			10.8	29.5	0.592*	62	900	[188] module
		ILGAR	14.7	37.4	0.574	68	<1	[189]
		Evaporation	14.8	31.3	0.665	71	<1	[190]
		Sputter	12.2	27.6	0.620	71	<1	[191]
ZnMgO	3.6	ALD	13.7	30.8	0.610	73	<1	[192]
Buffer-free		Partial electrolyte	15.7	34.6	0.636	72	<1	[193]
		ILGAR-*i*-ZnO	14.5	34.9	0.581	71	<1	[194]
		Zn-treated *i*-ZnMgO	16.2	40.2	0.587	69	<1	[195]
		Untreated *i*-ZnMgO	12.5	33.2	0.544	69	<1	[196]

*V_{OC} per cell for modules

CdS buffers, others come from industrial pilot lines. Also the evaluation methods are different, some are in-house measurements, others are certified. In order to evaluate the various Cd-free junction-formation methods in that respect, the efficiency resulting from a number of experiments is displayed in Figure 13.12 together with its reference, or estimation thereof. Generally, the Cd-free device is comparable to the CBD–CdS device within typical variations.

Altogether, there are several possibilities for obtaining high efficiency without Cd. All the listed methods include one or more of the elements Zn, In, and S. Zn is directly included in most of the buffer materials or indirectly as ZnO transparent contact with In_xSe_y, $In(OH,S)_x$, and In_2Se_3. Indications that *n*-type doping with Zn occurs similarly to that with Cd have been found by the treatment of $Cu(InGa)Se_2$ in Cd and Zn solutions [174], and are consistent with junction formation by solid-state diffusion into single-crystals [170].

In Figure 13.12 a slight tendency can be noted toward larger difference between Cd-free and CdS reference cells for the direct ZnO (no buffer) approaches. It appears as if a buffer layer between the $Cu(InGa)Se_2$ and the ZnO is beneficial. Such a layer will help invert the absorber surface (see Section 13.5.3), and possibly also serve to protect the junction and near-surface region during subsequent deposition of the transparent contact materials.

13.4.5 Transparent Contacts

The early $Cu(InGa)Se_2$ devices used CdS bilayers: CdS doped with In or Ga as front-contact layers in addition to the undoped CdS buffer contacting the $Cu(InGa)Se_2$. Short-wavelength light (<520 nm) was absorbed near the surface in the thick CdS layer and did not generate any photocurrent. When chemical bath deposition allowed CdS buffer layers to be thin enough such that it no longer limited the short-wavelength collection in the $Cu(InGa)Se_2$, photocurrent could be gained by

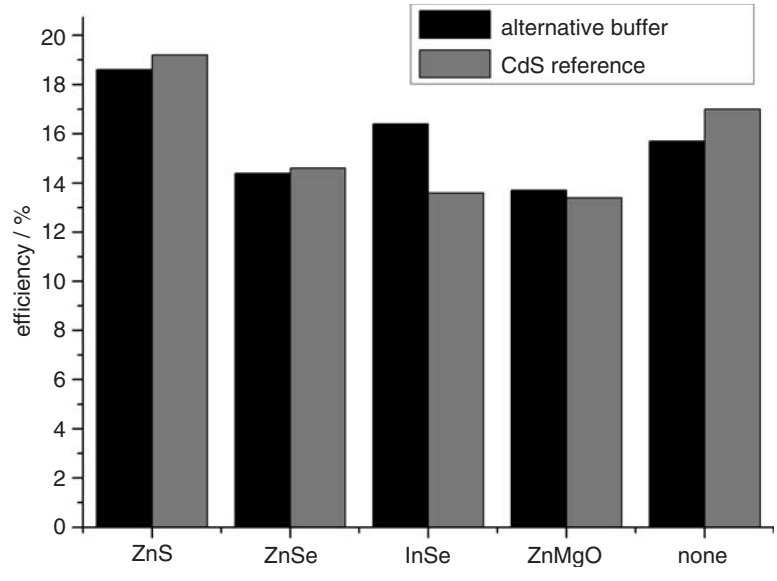


Figure 13.12 The efficiency of $Cu(InGa)Se_2$ solar cells with a selection of Cd-free buffer layers together with corresponding values of $Cu(InGa)Se_2$ cells with chemical bath-deposited CdS

increasing the bandgap of the contact layer. Since the contact layer must also have high conductivity for lateral current collection, the obvious choice is a transparent conducting oxide (TCO), a class of materials used in such devices as displays and low-emission coatings on window glass panes. There are three main materials in this class: doped SnO_2, In_2O_3:Sn (ITO), and doped ZnO. SnO_2 requires relatively high deposition temperatures that restrict the potential in Cu(InGa)Se_2 devices that cannot withstand temperatures greater than 200–250 °C after CdS is deposited. ITO and ZnO can both be used, but the most common material is ZnO, favored by potentially lower material costs. An overview of TCO thin film materials can be found in Chapter 17.

The most commonly used low-temperature deposition method for TCO films is sputtering. ITO layers are routinely fabricated on an industrial scale using dc sputtering. Industrial practice is to use ceramic ITO targets and to sputter in an Ar:O_2 mixture. Typical sputter rates range between 0.1 and 10 nm/s, depending on the application [197].

Sputtering of doped ZnO films is not as common in other applications such as displays as is sputtering of ITO. Nevertheless, it is the preferred method for depositing the transparent front contact on Cu(InGa)Se_2 devices, with and without CBD CdS, by the majority of the R&D groups. Typically, ZnO:Al films are deposited by radio frequency (rf) magnetron sputtering from ceramic ZnO:Al_2O_3 targets with 1 or 2 weight% Al_2O_3. In large-scale manufacturing, dc sputtering from ceramic targets is favored since it requires simpler equipment and offers higher deposition rates [198].

Reactive dc sputtering from Al/Zn alloy targets has also been used in the fabrication of Cu(InGa)Se_2/CdS devices with the same performance as with rf sputtered ZnO:Al [199]. The use of Zn/Al alloy targets allows lower costs than ceramic ZnO:Al_2O_3 targets, but reactive sputtering requires very precise process control owing to the so-called hysteresis effect [200] so that optimal optoelectronic properties are achieved only within a very narrow process window. Deposition rates in the 5–10 nm/s range have been achieved.

Chemical vapor deposition provides another deposition option and is used by at least one commercial manufacturer of Cu(InGa)Se_2 modules to deposit ZnO [129]. The reaction occurs at atmospheric pressure between water vapor and diethylzinc and the films are doped with fluorine or boron.

As with the Mo back contact, the requirements for sheet resistance of the transparent contact layer will depend on the specific cell or module design and the sheet resistance is usually controlled by the layer thickness. Typically, small-area cells use layers with 20–50 Ω/□ and thicknesses of 100 – 500 nm. Commercial modules may require 5–10 Ω/□ with correspondingly greater thicknesses because the module geometry requires transport over greater distances and therefore causes higher series resistance. With increasing thickness and lower resistance there is an increasing loss in transparency due to free carrier absorption that reduces the light which can be absorbed by the Cu(InGa)Se_2. Alternative TCO materials with higher doping efficiency and thus higher mobility, which achieve the same conductivity with a lower free carrier concentration, have been suggested to reduce free carrier absorption [201].

13.4.6 High-resistance Window Layers

It is common practice to use a layer of undoped high-resistivity (HR) ZnO before sputter deposition of the TCO layer. Depending on the deposition method and conditions, this layer may have a resistivity of 1–100 Ω cm compared with the transparent contact with 10^{-4}–10^{-3} Ω cm. Typically, 50 nm of HR ZnO is deposited by rf magnetron sputtering from an oxide target.

The gain in performance by using an HR ZnO HR layer in ordinary devices with CBD–CdS is related to the CdS thickness [199, 202, 203]. One explanation of the role of an HR ZnO layer is given by [202] as resulting from locally nonuniform electronic quality of the $Cu(InGa)Se_2$ layer that can be modeled by a parallel diode with high recombination current. The influence of these regions on the overall performance is reduced by the series resistance of the HR ZnO layer. This series resistance has a negligible effect on the performance of the dominant parts of the device area. A related explanation would attribute the local areas with poor diode characteristics to have been caused by pinholes in the CdS layer, which create parallel diodes with a $Cu(InGa)Se_2$/ZnO junction. In this case improved diode quality due to the HR layer would improve overall performance. Either case is consistent with the observation that a beneficial effect from the HR buffer, typically an increase in *Voc*, is not observed when the CBD–CdS layer is thick enough [203].

Another potential reason for using an HR ZnO buffer layer is to add protection of the interface region from sputter damage induced during deposition of the TCO layer which typically requires more harsh conditions. This seems to be particularly important for some alternative Cd-free buffer layers or with dc magnetron-sputtered TCO layers [204].

13.4.7 Device Completion

In order to contact laboratory test cells, a metal contact is deposited onto the TCO layer. It is shaped as a grid with minimum shadow area in order to allow as much light as possible into the device. Solar cell measurement standards recommend a minimum cell area of $1\,cm^2$, but many labs routinely use cells on the order of $0.5\,cm^2$. The metal grid contact can be made by first depositing some tens of nanometers of Ni to prevent the formation of a high-resistance oxide layer, and subsequently depositing of the order of a micrometer thickness of Al. Evaporation through an aperture mask is a suitable deposition method, though it has been argued that for the highest efficiencies, photolithographically defined grids are preferable [205].

After deposition of the metal grid, the total cell area is defined by removing the layers on top of the Mo outside the cell area by mechanical scribing or laser patterning. Alternatively, just the layers on top of the $Cu(InGa)Se_2$ can be removed, by photolithography and etching, since the lateral resistance of the $Cu(InGa)Se_2$ prevents collection outside the cell area. Or the TCO layer can be deposited through an aperture mask to define the cell area.

Finally, on the highest-efficiency devices an antireflection layer may be deposited to minimize optical losses. Typically this is an evaporated MgF_2 layer with thickness $\sim$100 nm. However, though this is not relevant for a module in which a cover glass or encapsulant is required.

The only significant difference in the device layers between lab cells and modules is the thickness of the TCO. Integrated modules (see Section 13.6.2) normally do not have any grid that assists in current collection over the cell area, so a substantially thicker TCO layer, that is, higher sheet conductivity, is needed in order to keep resistive losses low. A TCO layer with higher sheet conductivity may also have lower optical transmission in the infrared due to increased free-carrier absorption resulting in a decreased photocurrent.

13.5 DEVICE OPERATION

$Cu(InGa)Se_2$ solar cells have achieved more than 20% efficiency in laboratory-scale devices, largely by empirical processing improvements and despite an incomplete understanding of the underlying mechanisms and electronic defects that control the device behavior. In recent years considerable effort has gone to developing models of the effects of interfaces, grain boundaries, point defects,

etc. to enable both a better understanding of the devices and identification of pathways to further improvements.

The operation of Cu(InGa)Se_2/CdS solar cells is characterized by high quantum efficiency (QE) and short-circuit current. The open-circuit voltage increases with the bandgap of the absorber layer and is largely insensitive to grain boundaries and defects at the Cu(InGa)Se_2/CdS interface. The device operation can be described by identifying loss mechanisms. These can be divided into three categories. The first are optical losses that limit generation of carriers and therefore the short-circuit photocurrent (J_{SC}). The second are recombination losses that limit the open-circuit voltage (V_{OC}) and fill factor (FF). Finally, there are parasitic losses, such as series resistance, shunt conductance, and voltage-dependent current collection, which are most evident by their effect on FF but can also reduce J_{SC} and V_{OC}.

A basic device model has emerged from over two decades of investigations in which the voltage is limited by recombination through bulk trap states in the space-charge region of the Cu(InGa)Se_2 absorber layer. Recombination at the Cu(InGa)Se_2/CdS interface is minimized by an effective inversion of the Cu(InGa)Se_2 surface. Processing-dependent exceptions to this behavior, for example cases in which the voltage is limited by interface recombination, help to confirm this basic understanding. Similarly, characterization of the metastable response to light exposure or current injection further increase understanding of the device physics.

13.5.1 Light-generated Current

The highest-efficiency Cu(InGa)Se_2 devices have $J_{SC} = 36\,\text{mA/cm}^2$ [1] out of a possible $42.8\,\text{mA/cm}^2$ available for a bandgap of 1.12 eV under AM1.5 global illumination. Quantum efficiency is a valuable tool to characterize the losses responsible for this difference in current. The light-generated current is the integral of the product of the external quantum efficiency (QE_{ext}) and the illumination spectrum. QE_{ext} is controlled by the bandgap of the Cu(InGa)Se_2 absorber layer, the CdS and ZnO window layers, and a series of loss mechanisms. These losses are illustrated in Figure 13.13 where typical QE curves at two different voltage biases, 0 and

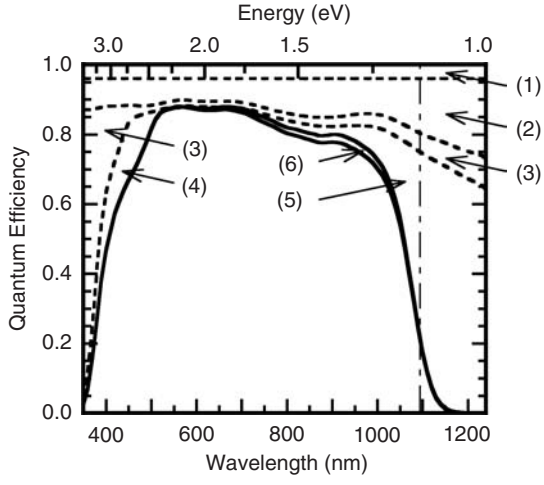


Figure 13.13 Quantum efficiency (solid lines) at 0 and – 1 V and optical losses for a Cu(InGa)Se_2/CdS/ZnO solar cell in which the Cu(InGa)Se_2 has $E_g = 1.12$ eV

Table 13.4 Current loss, ΔJ, for $E > 1.12$ eV due to the optical and collection losses illustrated in Figure 13.13 for a typical Cu(InGa)Se_2/CdS/ZnO solar cell

Region in Figure 13.13	Optical loss mechanism	ΔJ [mA/cm^2]
(1)	Shading from grid with 4% area coverage	1.7
(2)	Reflection from Cu(InGa)Se_2/CdS/ZnO	3.8
(3)	Absorption in ZnO	1.8
(4)	Absorption in CdS	0.8
(5)	Incomplete generation in Cu(InGa)Se_2	1.9
(6)	Incomplete collection in Cu(InGa)Se_2	0.4

-1 V, are shown. The QE curve at -1 V is slightly higher at longer wavelengths, because of the larger space charge width under reverse bias, which increases the effective collection length. The current loss under 100 mW/cm^2 illumination is listed in Table 13.4 for each of these mechanisms. Losses 1–5 are optical and 6 is electronic. In practice, the magnitude of each of these losses will depend on the details of the device design and optical properties of the specific layers. The losses include the following:

1. Shading from a collection grid used for most devices. In an interconnected module this will be replaced by the area used for the interconnect, as discussed in Section 13.6.2.
2. Front surface reflection. On the highest-efficiency devices this is minimized with an antireflection layer.
3. Absorption in the ZnO layer. Typically, there is 1–3% absorption through the visible wavelengths, which increases in the near IR region, $\lambda > 900$ nm, where free-carrier absorption becomes significant, and for $\lambda < 400$ nm near the ZnO bandgap.
4. Absorption in the CdS layer. This becomes appreciable at wavelengths below ~520 nm corresponding to the CdS bandgap 2.42 eV. The loss in QE for $\lambda < 500$ nm is proportional to the CdS thickness since it is commonly observed that electron–hole pairs generated in the CdS are not collected. Figure 13.13 shows a device with a ~30 nm-thick CdS layer. In practice, the CdS layer is often thicker and the absorption loss greater.
5. Incomplete absorption in the Cu(InGa)Se_2 layer near the Cu(InGa)Se_2 band gap. Bandgap gradients, resulting from composition gradients in many Cu(InGa)Se_2 films, also affect the steepness of the long-wavelength part of the QE curve.
6. Incomplete collection of photogenerated carriers in the Cu(InGa)Se_2, discussed below.

QE_{ext} is then given by

$$QE_{\text{ext}}(\lambda, V) = [1 - R(\lambda)][1 - A_{\text{ZnO}}(\lambda)][1 - A_{\text{CdS}}(\lambda)]QE_{\text{int}}(\lambda, V) \tag{13.3}$$

where R is the total reflection, including the grid shading, A_{ZnO} is the absorption in the ZnO layer and A_{CdS} is the absorption in the CdS layer. QE_{int}, the internal quantum efficiency, is the ratio of photogenerated carriers collected to the photon flux that arrives at the absorber layer and can be approximated by [206]

$$QE_{\text{int}}(\lambda, V) = 1 - \exp[-\alpha(\lambda)(W(V) + L_{diff})] \tag{13.4}$$

where α is the Cu(InGa)Se_2 absorption coefficient, W is the space-charge width in the Cu(InGa)Se_2, and L_{diff} is the minority-carrier diffusion length. This approximation assumes that all carriers generated in the space-charge region are collected without recombination loss. Since W is a function

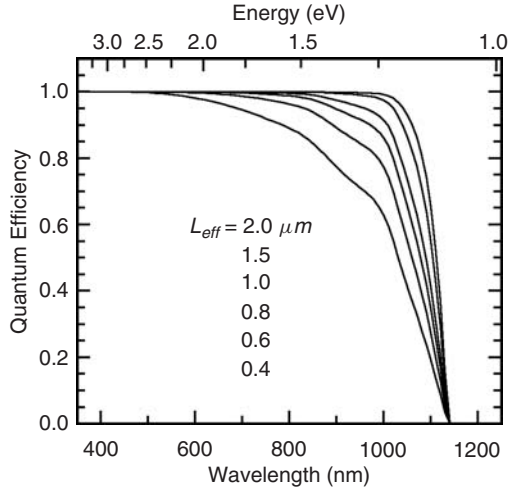


Figure 13.14 Calculated internal quantum efficiency $QE = 1 - \exp[-\alpha(\lambda)L_{eff}]$ for Cu(InGa)Se_2 with $E_g = 1.12$ eV and collection lengths from 2.0 μm (top) to 0.4 μm (bottom), using absorption coefficient data from reference [50]

of the applied voltage bias, QE_{int} and total light-generated current are, in general, voltage-dependent, so the latter can be written as $J_{\text{L}}(V)$. Values of W in the range 0.1–0.5 μm have been reported for typical cells at 0 V.

Internal quantum efficiencies, calculated from Equation (13.4) with values of total effective collection length $L_{eff} \equiv W(V) + L_{diff}$ from 2.0–0.4 μm, are shown in Figure 13.14. This calculation is for Cu(InGa)Se_2 with $E_g = 1.12$ eV and assumes that no carriers generated by light with $E < E_g$ are collected. If L_{eff} is smaller than 1 μm, an increasingly significant fraction of electrons is generated deeper in the Cu(InGa)Se_2 layer than the collection length due to the decreasing absorption coefficient. Thus carriers are not collected and the QE decreases at longer wavelengths. This incomplete collection can be a significant loss mechanism for Cu(InGa)Se_2 devices [156, 207]. The effect of $J_{\text{L}}(V)$ on current–voltage behavior increases as the collection length decreases with forward voltage bias and therefore has its largest effect on the fill factor and V_{OC} [208, 209]. This voltage-dependent current collection can decrease J_{SC} as illustrated in Figure 13.13 by the increase in QE measured at −1 V applied voltage bias compared to that measured at 0 V.

This analysis of current collection is for the case with the Cu(InGa)Se_2 sufficiently thick that all light is absorbed. While a typical Cu(InGa)Se_2 absorber layer thickness is ∼2 μm in laboratory scale devices and 1.2–1.5 μm in commercial modules, there is incentive to reduce d as much as possible in manufacturing in order to increase throughput and reduce material costs. But if the Cu(InGa)Se_2 is thinner than ∼1.0 μm, current loss due to incomplete absorption at long wavelengths becomes significant. The effect of reduced thickness is similar to that of reduced collection length if reflection at the back contact is neglected so Figure 13.14 with $d = L_{eff}$ could be taken as a limiting case for the expected QE_{int} with thin absorbers. Device performance with sub-micrometer-thick absorber layers has been characterized using films deposited by coevaporation [210, 211]. The decrease in J_{SC} for $d < 1$ μm was greater than expected due to the incomplete optical absorption and could not be explained by device simulation [212]. With thinner absorber layers the back contact can increasingly affect the device behavior. Light passing through the absorber layer and reaching the back contact means that low optical reflection at the Mo/Cu(InGa)Se_2 back contact is relevant [213]. Results for alternative back contacts with high reflection have been reported [108,

213, 214]. For example, with a TiN contact, an improvement in J_{SC} of 0.8 mA/cm^2 was reported with 0.45-μm-thick Cu(InGa)Se_2, in reasonable agreement with model results for an increase in reflectivity [212].

V_{OC} and FF, on the other hand, can be maintained at the same values as with 2-μm-thick films, even with thicknesses as low as 0.5 μm [210, 211]. This is surprising since more minority carriers should reach and recombine at the back contact as the thickness decreases, thereby reducing V_{OC}. The $MoSe_2$ interlayer which forms at the Mo/Cu(InGa)Se_2 interface [107] may act as a back surface field to reflect minority carriers – as has been suggested based on the band alignment [215] – thus preventing back surface recombination. This idea was supported by results using alternative back contacts. For example, using 0.6-μm-thick Cu(InGa)Se_2 layers V_{OC} was reduced with a ZrN/Cu(InGa)Se_2 contact, but was recovered with a thin deposited $MoSe_2$ interlayer [213]. A back-surface field formed with a Ga gradient can also prevent the loss in V_{OC} by reducing recombination at the back contact with thin absorbers [210].

Even though only some of the loss in J_{SC} can be easily explained by incomplete optical absorption, a potential approach to increasing current is to use light trapping to increase the optical path length in the absorber layer. This is commonly used in amorphous Si-based solar cells (see Chapter 12) where, similar to the case with sub-micrometer Cu(InGa)Se_2, the absorber layer thickness is smaller than the optical absorption depth. In Cu(InGa)Se_2 solar cells light trapping might be implemented with a textured TCO layer to scatter the incident light or with a textured back contact if the texture and morphology is propagated through the conformal coating of subsequent layers so that the top TCO layer again scatters the incident light. Approaches to form textured TCO films are discussed in Chapter 17. In either case, light scattering from the rough surface of the as-deposited Cu(InGa)Se_2 must also be taken into account.

13.5.2 Recombination

The current–voltage ($J-V$) behavior of Cu(InGa)Se_2/CdS devices can be described by a general diode equation:

$$J = J_D - J_L = J_O \exp\left[\frac{q}{AkT}(V - R_S J)\right] + G(V - R_S J) - J_L \tag{13.5}$$

with the diode current J_O given by:

$$J_O = J_{OO} \exp\left(-\frac{\Phi_b}{AkT}\right) \tag{13.6}$$

The ideality factor A, barrier height Φ_b, and prefactor J_{OO} depend on the specific recombination mechanism that dominates J_O, while the series resistance R_S and shunt conductance G are losses that occur in series or parallel with the primary diode. General expressions for A, Φ_b, and J_{OO} in the cases of recombination through the interface, space-charge region, or bulk of the absorber layer can be found in various textbooks.

To understand the specific diode behavior of Cu(InGa)Se_2/CdS solar cells, it is instructive to look at the effects of the Cu(InGa)Se_2 bandgap, varied by changing $x \equiv Ga/(In + Ga)$, and temperature. Figures 13.15 and 13.16 show $J-V$ and QE curves for three devices with $x = 0$, 0.24, and 0.61, corresponding to $E_g = 1.04$, 1.14, and 1.36 eV, respectively [207]. V_{OC} increases and the position of the long-wavelength QE edge shifts to shorter wavelength as E_g increases. Figure 13.17 shows the temperature dependence of V_{OC} for these devices. In each case, as

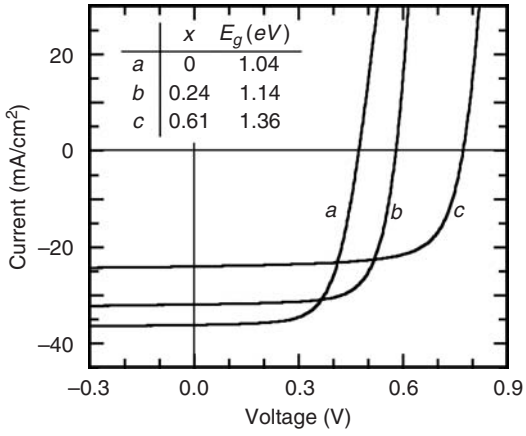


Figure 13.15 Current–voltage curves for Cu(InGa)Se_2/CdS solar cells with different relative Ga content giving (a) $E_g = 1.04$, (b) 1.14, and (c) 1.36 eV

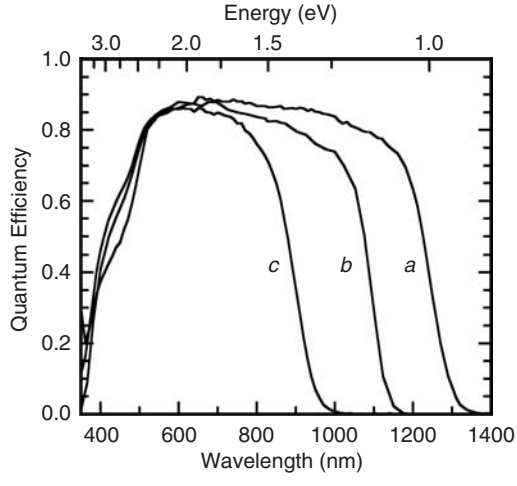


Figure 13.16 Quantum efficiency curves for the devices shown in Figure 13.15

$T \to 0$, $V_{OC} \to E_g/q$. Thus with the barrier height $\Phi_b = E_g$ combining Equations (13.5) and (13.6) and assuming $G \ll J_L/V_{OC}$, the open-circuit voltage becomes

$$V_{\mathrm{OC}} = \frac{E_{\mathrm{g}}}{q} - \frac{AkT}{q} \ln\left(\frac{J_{\mathrm{OO}}}{J_{\mathrm{L}}}\right). \tag{13.7}$$

Different recombination paths are effectively connected in parallel so that V_{OC} will be controlled by the single dominant mechanism with the highest recombination current, J_{REC}. The values of Φ_b and A can be used to distinguish between recombination in the bulk absorber, in the space charge region of the Cu(InGa)Se_2, or at the Cu(InGa)Se_2/CdS interface [216–218]. Each of the curves in Figure 13.15 can be fit to Equation (13.5) with $A = 1.5 \pm 0.3$. For a wide range of thin film solar cells, it has been demonstrated that $V_{OC}(T \to 0) = E_g/\mathrm{q}$ and $1 < A < 2$ similar to the data above. Measured values of $\Phi_b = E_g$ have been reported for $CuInSe_2$ [156,

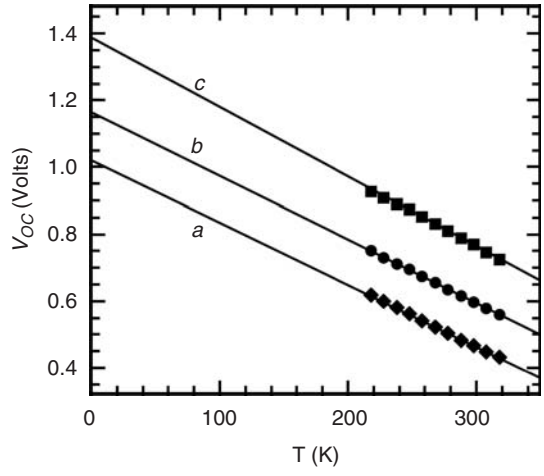


Figure 13.17 Temperature dependence of V_{OC} for the devices shown in Figure 13.15

219] and many Cu(InGa)$(SeS)_2$ [217] devices, independent of the bandgap of a (CdZn)S buffer-layer [219], and for a variety of different absorber-layer deposition processes [220]. These results for Φ_b and A indicate that Cu(InGa)Se_2/CdS solar cells operate with the diode current controlled by Shockley–Read–Hall-type recombination in the Cu(InGa)Se_2 layer. This recombination is greatest through deep trap states in the space-charge region of the Cu(InGa)Se_2 where there are comparable supplies of electrons and holes available, that is, where $p \approx n$. The variation in A between 1 and 2 depends on the energies of the deep defects that act as dominant trap states [221]. As these states move toward the band edges, $A \rightarrow 1$ and the recombination becomes closer to band-to-band bulk recombination. One characteristic of some very high efficiency devices is their relatively low values of $A \approx 1.1 - 1.3$ [1].

It may seem surprising that recombination at the Cu(InGa)Se_2/CdS interface does not limit V_{OC} since, in processing Cu(InGa)Se_2 solar cells, no special efforts are made to match lattice constants or reduce interface defects, and the devices are typically exposed to air between the Cu(InGa)Se_2 and CdS depositions. This can be explained by effective type inversion of the near-junction region of the Cu(InGa)Se_2 [222, 223] induced by band bending and interface Fermi-level pinning [216, 218, 224]. The Cu(InGa)Se_2/CdS band diagram shown in Figure 13.18 shows the Fermi level at the interface close to the conduction band so that electrons in the near surface region of the Cu(InGa)Se_2 are effectively majority carriers and there is an insufficient supply of holes available for recombination through the interface states. As explained by Turcu *et al*. [217] formation of the ordered defect compound at the surface of Cu-poor Cu(InGa)Se_2 plays a critical role in preventing interface recombination over the range of compositions and band alignments created by alloying the absorber and/or buffer layers. This surface has a wider bandgap [78, 225] and lower valence band energy, shown in Figure 13.18, creating a barrier to prevent holes from reaching and recombining at the absorber/buffer interface. This explanation is strongly supported by an important exception – when the absorber layer is processed without Cu-vacancies as is often the case for wide-bandgap devices (see Section 13.5.4), especially those using Cu(InGa)S_2 absorbers [226]. Regardless of the bandgap, such devices have lower V_{OC} and $\Phi_b < E_g$ which is indicative of interface recombination [227, 217] enabled by the absence of a barrier to hole transport.

It has alternatively been proposed that doping due to Cd diffusion during the chemical bath deposition of CdS results in the formation of an n-type emitter and a p–n homojunction in the Cu(InGa)Se_2 [174]. However, when simulating a buried p-n junction in a strongly absorbing

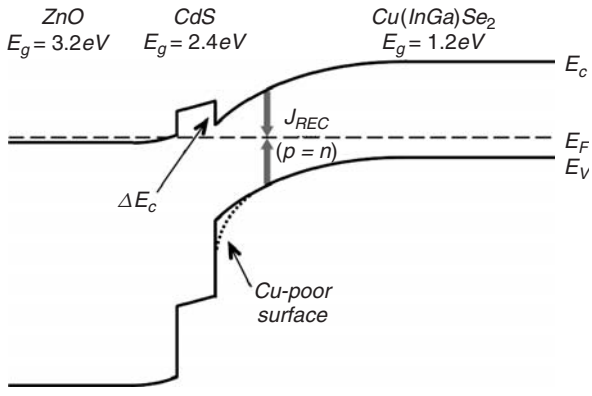


Figure 13.18 Band diagram of a ZnO/CdS/Cu(InGa)Se_2 device at 0 V in the dark. The recombination current J_{REC} is greatest where $p = n$ in the space charge region of the Cu(InGa)Se_2 and not at the interface. A Cu-poor surface layer is shown by a dashed line

material such as Cu(InGa)Se_2, it becomes immediately clear that this leads to a considerable loss in quantum efficiency in the blue spectral region, because charge carriers created by highly absorbed photons recombine at the absorber surface and never reach the actual *p-n* junction [218]. This is different in Si solar cells which absorb solar light much more weakly. Thus, the surface of the Cu(InGa)Se_2 absorber must be inverted, but it does not form a homojunction.

The primary effect of sodium on device performance is on V_{OC} [228] and may be explained by passivation of interface recombination. Cu(InGa)Se_2 devices with a diffusion barrier that restricted Na diffusion from the glass substrate had V_{OC} reduced by 120 mV, Φ_b less than the absorber bandgap [229], and a high density of defects near the Cu(InGa)Se_2/CdS junction indicated by capacitance measurements [230].

One other potential source of recombination that might reduce V_{OC} is back-surface recombination at the Mo/Cu(InGa)Se_2 interface. This will be negligible so long as the minority-carrier diffusion length is small compared with the total Cu(InGa)Se_2 thickness. As with the case of thin absorber layers, a back-surface field may be implemented by increasing the Ga content near the Mo to give a bandgap gradient [213].

In real Cu(InGa)Se_2 materials with imperfect structures, trap defects will not exist at discrete energies, but form defect bands or tails at the valence and conduction bands. Then the total recombination current can be determined by integrating over the defect spectrum. While the specific point defects that create the recombination traps are not identified, it should be noted that the defect at $E_V + 0.8$ eV identified by transient photocapacitance measurements may be important in this regard [72]. Recombination through an exponential bandtail was used to explain the temperature dependence in A observed in some devices [231]. Analysis of the temperature dependence of A was further explained by a tunneling enhancement of the recombination current, particularly at reduced temperatures [232]. Admittance spectroscopy is a useful tool to characterize the distribution of electronic defects in Cu(InGa)Se_2/CdS solar cells [233] and the density of the acceptor state $\sim$0.3 eV from the valence band has been correlated to V_{OC} [234]. The minority-carrier lifetime is another valuable parameter to characterize Cu(InGa)Se_2/CdS devices. Transient photocurrent [235] and time-resolved photoluminescence [209, 236] measurements each were used to determine lifetimes in the range 10–250 ns for high-efficiency devices. Still, a critical problem that remains is to identify which of the defects discussed in Section 13.2 provides for the recombination that limits voltage in the devices.

Metastable effects, i.e. reversible changes in the performance of solar cells, have been well documented and provide additional insight into the role of defects in the device behavior. These were first reported as a light-soaking effect that increases the open-circuit voltage and the fill factor under illumination [237]. Persistent photoconductivity has been related to the same defect changes as the light-soaking effect [238]. It was recently proposed that this metastable behavior arises from the complex of a Se vacancy and a Cu vacancy [239]. Upon trapping of a photogenerated or injected electron the Se divacancy changes from donor-like to acceptor-like. A one-to-one correlation between increases in deep acceptor and hole carrier densities under a variety of light soaking and injection conditions was observed [240] as predicted by the Se divacancy model. Experimentally observed activation energies agree well with predicted energy barriers for transitions in the divacancy model [241]. Because of the various metastable effects, light soaking under white light has been introduced into the standard efficiency measurement procedures. $J-V$ curves are measured after light soaking and sweeping the voltage from forward to the reverse bias.

In practice, analysis of $J-V$ data is commonly used to determine the diode parameters J_O, A, and Φ_b [242]. This requires that R_S and G are negligible, or else suitable corrections need be made to the data, and that J_L is independent of V over the range of analysis. Failure to account for $J_L(V)$ can lead to errors in analysis of current–voltage data [208] including erroneously high values of A. In many cases the fundamental diode parameters cannot be reliably determined except from $J-V$ data measured in the dark [207]. In addition, it must be verified that there are no non-ohmic effects at any contacts or junctions, which cause the appearance of a second diode for which Equation (13.5) does not account. Such non-ohmic behavior is often observed at reduced temperatures [219, 220]. Once it has been demonstrated that all these parasitic effects are negligible, or corrections have been made, then J_O can be determined by a linear fit to a semilogarithmic plot of $(J + J_L)$ versus $(V - R_S J)$ and A can be determined from the slope of the derivative dV/dJ versus $1/J$ in forward bias [243], or both J_O and A can be obtained by a least-squares fit to Equation (13.5). Finally, Φ_b can be determined from the temperature dependence of V_{OC}, as in Figure 13.17, or of J_O.

It must be noted that most descriptions of transport and recombination ignore the effect of grain boundaries, implicitly assuming that grains are columnar and all transport can proceed without crossing grain boundaries. As noted above, this is rarely, if ever, strictly true, so a comprehensive description of Cu(InGa)Se_2 solar cells must account for the possibility of recombination at grain boundaries reducing current collection or voltage. Numerical simulations of different device models concluded that a valence band barrier greater than 200 meV, consistent with the measurements described in Section 13.2.4, could best prevent losses at the grain boundaries and enable high-efficiency devices [244, 245]. Proposals that the current collection was aided by transport along the grain boundaries could not be simulated without a concurrent loss in voltage.

13.5.3 The Cu(InGa)Se_2/CdS Interface

The Cu(InGa)Se_2/CdS band diagram in Figure 13.18 shows bandgap widening caused by the Cu vacancies at the surface of the Cu(InGa)Se_2. There is experimental evidence that only the surface is Cu depleted and a thick ODC layer is not formed [246]. The band diagram shows that the conduction-band offset ΔE_C between the CdS and the Cu(InGa)Se_2 is also important for creating the type inversion in the Cu(InGa)Se_2. In this diagram, the bulk Cu(InGa)Se_2 layer is p-type with E_g depending on the relative Ga concentration, the CdS layer is n-type with $E_g = 2.4$ eV and is totally depleted, and the bulk ZnO n^+-layer has $E_g = 3.2$ eV. The HR ZnO layer between the n^+-ZnO layer and the CdS is also assumed to be depleted. Positive ΔE_C indicates a spike in the conduction band, that is, the conduction-band minimum in the CdS is at higher energy than the conduction-band minimum of the Cu(InGa)Se_2. Figure 13.18 shows the case with $\Delta E_C = 0.3$ eV

and a -0.3 eV conduction-band offset between the ZnO and the CdS [78]. Simulation of current transport and recombination has considered the effect of ΔE_C [247–249]. These models show that if ΔE_C is greater than about 0.5 eV, collection of photogenerated electrons in the Cu(InGa)Se_2 is impeded and J_{SC} or FF is reduced sharply. With a smaller spike, electrons can be transported across the interface assisted by thermionic emission [247]. On the other hand, for sufficiently negative ΔE_C the induced type inversion of the Cu(InGa)Se_2 of the absorber surface may be eliminated, the electron concentration at the interface on the buffer side increased, and interface state recombination will limit V_{OC}.

Owing to its importance in the electronic behavior of Cu(InGa)Se_2/CdS devices, several efforts have been made to calculate or measure ΔE_C with varying results. Band-structure calculations gave $\Delta E_C = 0.3$ eV [250] for pure $CuInSe_2$. Since the conduction-band minimum in $CuGaSe_2$ is higher in energy [251], a smaller offset is expected for Cu(InGa)Se_2. On the other hand the conduction-band minimum of the ODC compound has been calculated to be lower than the chalcopyrite [252].

Experimentally accessible are the valence band offsets via photoemission spectroscopy (XPS, UPS or synchrotron based). For unoxidized Cu(InGa)Se_2 interfaces with CdS a valence band offset of -0.9 eV was found which lies between the theoretical values for the chalcopyrite and the ODC [253]. The same valence-band offset was found for pure $CuInSe_2$ polycrystalline [78] or epitaxial [254] films, or single-crystals [255], independent of the surface orientation or the deposition method of the CdS film [177]. Assuming an ODC surface with a band gap of 1.4 eV, these measurements result in a conduction-band offset of 0.3 eV. All these measurements were done at *in situ* prepared surfaces of the CIGS absorber, whereas in solar cell preparation, the absorbers surface is exposed to air before depositing the CdS buffer layer. Additionally the determination of conduction-band offsets from valence-band offsets assumes that the bulk band gaps are valid for the surfaces. In particular, interdiffusion between the buffer and absorber can lead to changes in the bandgaps and in this case the conduction-band offset cannot be determined from the valence-band offset. The conduction-band offset is directly accessible by inverse photoemission spectroscopy (IPES) and has been found to be zero for the $CuInSe_2$/CdS interface [256]. This was attributed to intermixing at the interface. For the suppression of interface recombination the absence of a cliff appears necessary, thus the finding of zero conduction-band offset would be compatible with the $J-V$ measurements that find recombination in the space-charge region as the dominating path.

13.5.4 Wide and Graded Bandgap Devices

While the highest efficiency devices generally have $Ga/(In + Ga) \approx 0.1-0.3$ giving $E_g \approx 1.1-1.2$ eV, significant effort has been made to develop high-efficiency solar cells based on wider-bandgap alloys. This is driven primarily by the expectation that wider-bandgap alloys will yield higher module efficiencies due to reduced losses related to the trade-off between higher voltage and lower current at maximum power. The resulting reduction in power loss, proportional to I^2R, can be used to either: (1) increase the module's active area by increasing the spacing between interconnects or grid lines, or (2) decrease the optical absorption in the TCO layers since they can tolerate greater resistance. Wider bandgap gives a lower coefficient of temperature for the device or module output power [257], which will improve performance at the elevated temperatures experienced in most real terrestrial applications. Wide-bandgap devices could also be used as the top cell in a tandem or multijunction cell structure.

The wider-bandgap materials that have attracted the most attention for devices are Cu(InGa)Se_2 and $CuInS_2$. $CuGaSe_2$ has $E_g = 1.68$ eV, which is well suited for the wide-bandgap cell in tandem structures. $CuInS_2$ has $E_g = 1.53$ eV, which could be nearly optimum for a single-junction device. The highest-efficiency devices based on $CuInS_2$ are deposited with Cu-rich

overall composition and then the excess Cu, in the form of a Cu_xS second phase, is etched away before CdS deposition [226]. Consequently, as explained above, the interface is more critical in these devices and improved V_{OC} has resulted from careful control of the surface, including Ga incorporation [258].

Cu(InAl)Se_2 solar cells have also been considered and 17% device efficiency has been demonstrated with $E_g = 1.15$ eV [259]. Since $CuAlSe_2$ has $E_g = 2.7$ eV, the alloy requires smaller changes in relative alloy concentration and lattice parameter from $CuInSe_2$ than the Ga alloys to achieve comparable bandgap. Finally promising results for wide-bandgap devices, including high V_{OC}, have recently been demonstrated with (AgCu)(InGa)Se_2 films [260, 261] where replacement of Cu with Ag enables the bandgap to be increased while lowering the alloy melting temperature which could enable films with lower defect densities. The highest-efficiency wide-bandgap devices using different alloys are listed in Table 13.5.

There are a variety of effects of Ga incorporation on film characteristics and device behavior as the bandgap is increased. The addition of a small amount of Ga to $CuInSe_2$ increased the open-circuit voltage, even when the Ga was confined to the back of the absorber and did not increase the bandgap in the space-charge region [142]. Improved adhesion with Ga addition was similarly observed with S [266] and Al [259] alloys. The effect of increasing bandgap in Cu(InGa)Se_2/CdS solar cells on V_{OC} and efficiency, which has been observed by several groups using various processes, is shown in Figure 13.19. Efficiency is roughly independent of bandgap for $E_g < 1.3$ eV or Ga/(In + Ga) < 0.4 [207] while V_{OC} increases approximately linearly with E_g. With even wider-bandgap absorbers, V_{OC} increases to greater than 0.8 V, but the efficiency decreases due to two effects: increased recombination reduces V_{OC} below that expected from Equation (13.7) [215, 216]; and voltage-dependent current collection [207] causes the fill factor to decrease. The dashed line in Figure 13.19 shows a line with slope $\Delta V_{OC}/\Delta E_g = 1$. Ideally, the increase in V_{OC} would have only a slightly smaller slope due to the dependence on J_L in the second term of Equation (13.7). Several changes have been observed with increasing Ga content. Admittance spectroscopy showed a correlation between the recombination and the density of a defect with an activation energy ~0.3 eV, which increases with E_g [267]. The defect band centered at 0.8 eV from the valence band independent of the relative Ga concentration moves closer to midgap for increasing bandgap and therefore becomes more efficient as a recombination trap [72]. The ideality factor A increases toward $A = 2$ with increasing Ga content [267], consistent with the traps moving closer to midgap [221]. The band alignment between the absorber and the CdS buffer changes from a spike to a cliff configuration [268], which may affect the recombination mechanism [218]. Additionally, Cu(InGa)Se_2 shows greater carrier concentration with increasing E_g which makes for smaller space-charge widths and thus shorter collection lengths [269]. Finally, a decrease in the

Table 13.5 Highest total area efficiencies for wide-bandgap devices with different alloy absorber layers. The record-efficiency low-bandgap $CuInSe_2$ and Cu(InGa)Se_2 cells shown for comparison

Material	E_g [eV]	Efficiency [%]	V_{OC} [%]	J_{SC} [mA/cm^2]	FF [%]	Reference
$CuInSe_2$	1.02	14.5	491	41.1	71.9	[262]
Cu(InGa)Se_2	1.12	20.0	692	35.7	81.0	[263]
$CuGaSe_2$	1.68	9.5	905	14.9	70.8	[264]
Cu(InGa)S_2	1.53	12.9	832	22.9	67.0	[258]
Cu(InAl)Se_2	1.51	9.9	750	20.1	65.8	[265]
(AgCu)(InGa)Se_2	1.6	13.0	890	20.5	71.3	[261]
Ag(InGa)Se_2	1.7	9.3	949	17.0	58	[260]

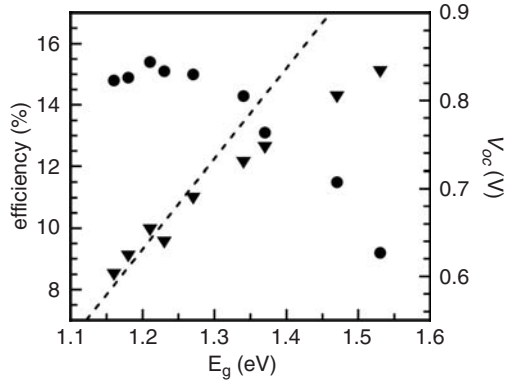


Figure 13.19 Efficiency (●) and V_{OC} (▼) as a function of Cu(InGa)Se_2 bandgap, varied by increasing the relative Ga content, (From Shafarman W, Klenk R, McCandless B, *Proc. 25th IEEE Photovoltaic Specialist Conf*., pp 763–768 (1996) [274]). The dashed line has slope $\Delta V_{OC}/\Delta E_g = 1$

grain size of coevaporated Cu(InGa)Se_2 films with increasing Ga content has been reported [270] but there is no evidence that this effects device performance and pure $CuGaSe_2$ absorbers have been demonstrated with grain sizes in the same range as Cu(InGa)Se_2 absorbers with low Ga [264].

Bandgap gradients formed by controlled incorporation of Ga or S have been proposed as a means to increase device efficiency by separately reducing recombination and collection losses [21, 271–273]. A gradient in the conduction band from wider at the Cu(InGa)Se_2/Mo interface to narrower near the space charge region has been used to effectively enhance the electric field thereby enhancing minority-carrier collection [273, 274] and to reduce back-surface recombination when the diffusion length is comparable to the film thickness [275]. Alternatively, the opposite gradient, i.e. from wider at the Cu(InGa)Se_2/CdS interface to narrower at the edge of the space charge region, could reduce recombination and increase V_{OC}. In this case, the smaller bandgap in the bulk portion of the device can still enable high optical absorption and J_{SC} [21, 273]. However care needs to be taken to make sure that the minimum is well inside the space-charge region, otherwise a barrier is formed that hampers the charge carrier transport [272]. The most effective implementation of a surface bandgap gradient may be the incorporation of S near the front surface [22] since the main effect is in lowering the valence band, instead of raising the conduction band as with Ga, and there should be less impact on collection of light-generated electrons.

An additional motivation for the development of wide-bandgap alloys is their potential incorporation into tandem cells which may provide an approach to higher-efficiency thin film modules. A proposed structure for a monolithic tandem device structure, in which a wide bandgap I–III–VI_2-based cell with $E_g \geq 1.5$ eV is directly connected via a shorting junction to a narrow-bandgap cell, is shown in Figure 13.20. $CuInSe_2$ with $E_g = 1.04$ eV is well suited for the bottom cell. Simulation of a series connected tandem using polycrystalline thin films found the optimum bandgap to be 1.72 eV for the top cell and 1.14 eV for the bottom cell [276]. A 1.04 eV bottom cell using $CuInSe_2$ would be ideally matched by a top cell with 1.65 eV, so $CuGaSe_2$ is well suited. Inclusion of real optical losses generally necessitates even higher bandgap for the top cell although it can be broadened using modified tandem structures, for example using a reduced thickness or area of the top cell. [277].

In either the two-terminal monolithic configuration shown or in a four-terminal, mechanically stacked configuration [278], the wide-bandgap cell is fully illuminated under the complete solar spectrum while illumination to the narrow-bandgap cell is filtered by the wide-bandgap cell. Thus the wide-bandgap cell makes a much greater contribution to the tandem cell performance so

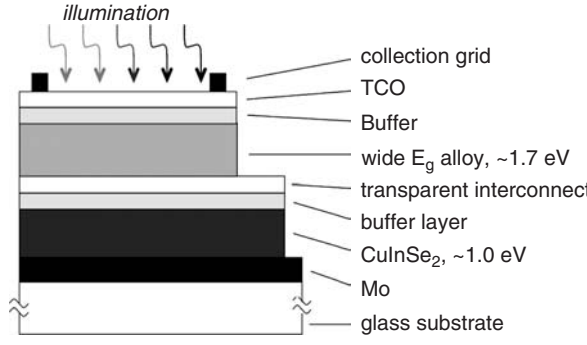


Figure 13.20 Substrate configuration for a two-terminal monolithic tandem solar cell using $CuInSe_2$-based alloys for top and bottom cells

its high efficiency is necessary. Several other difficult problems need to be solved to make this approach viable [279]. The top cell will need high optical transparency to the IR light with energy less than that of the absorber layer bandgap. This means that the Mo contact must be replaced by a transparent back contact that does not react with Se at the film deposition temperatures. Some promising results have used highly doped oxide films [280] and may include a very thin $MoSe_2$ layer [281]. Next, a shorting junction must be developed that allows transport of holes from the top cell to the bottom cell and electrons in the other direction without parasitic optical absorption. Approaches to this may follow from knowledge learned with III–V or a-Si multijunction devices (see Chapters 8 or 12) or the TCO back contacts used could just provide the needed junction [279]. Finally, the monolithic configuration will require the top, wide-bandgap, cell to be processed on the bottom cell without degrading its performance.

13.6 MANUFACTURING ISSUES

The competitiveness of a PV technology will primarily be governed by its performance, reliability, and costs. The best $Cu(InGa)Se_2$ cells and modules have demonstrated efficiency on a par with many commercial crystalline silicon products. Long-term stability doesn't appear to be a fundamental problem, as shown in field tests of prototype modules, but low-cost large volume production remains to be demonstrated in practice.

It is evident that thin films have the potential to be produced at very low costs. At the low end of production cost, moisture barriers of aluminum films that are deposited on plastic foils for food packaging cost less than 0.01 $\$/m^2$ to produce. More advanced functional coatings are more expensive to manufacture. For example, thin film coatings on architectural glass cost of the order of 1 $\$/m^2$. Thus PV modules constructed from thin film materials have the possibility for very low manufacturing costs. Whether $Cu(InGa)Se_2$ module production will be able to achieve this low-cost potential will depend on how well the process technology fulfills the requirements for material costs, throughput, and yield.

13.6.1 Processes and Equipment

Deposition processes can be either batch-type, in which a number of substrate plates are processed in parallel, or in-line, in which one substrate plate immediately follows the preceding one. In batch

processing, a process step is completely finished before the next batch is started, whereas a substrate plate may enter an in-line process step before the previous substrate is finished and the process keeps running continuously.

One common view on volume production is that in-line continuous processing is a prerequisite for low costs. Ideally, all the steps in a module manufacturing process could be synchronized so that modules could move through the entire process line continuously. While this requires the process time of each step to be the same, some flexibility is afforded by the option to run slower steps with multiple pieces of equipment in parallel. Fabrication of large-area thin film products with physical vapor deposition is often made in a continuous or quasi-continuous in-line system. However, the cost of a batch process can be equally low, provided the throughput is large enough. This may be the case, for example, with reaction of precursors where the reaction time may be long. For manufacturing of $Cu(InGa)Se_2$ modules, this means that the CdS chemical bath deposition can well fulfill low-cost production criteria, even though it is normally a batch process. Similarly, growth of the $Cu(InGa)Se_2$ layer by batch selenization does not necessarily need to be associated with higher costs than $Cu(InGa)Se_2$ fabricated by in-line coevaporation, provided the cycle time is short enough or batch size large enough.

Roll-to-roll processing, originally demonstrated for semiconductor thin films with evaporation of CdS for solar cells [282], may be implemented in a combined mode. A single process step can be operated continuously to coat rolls of substrate material that may be of the order of 1000 m long. Then the coated roll can be moved to the next process step as in a batch mode. It is also possible to have a single roll move continuously through multiple processes.

Sputtering is a mature large-scale deposition process which has been widely used for fabrication of large-area thin film coatings of various kinds, for example, in the glass industry. Similar processes are used in the fabrication of most $Cu(InGa)Se_2$ modules for the Mo back contact and the TCO front contact, so the same type of equipment, available from a number of suppliers, can be used. Manufacturing scale roll-to-roll sputter deposition equipment is also commercially available. Sputtering is also one option for deposition of the metal precursor films for fabrication of the $Cu(InGa)Se_2$ layer by a precursor reaction process. Other options, including electrodeposition or ink coating, are being developed for their potential advantages in process costs though equipment may still require custom development. The selenization step also requires specific custom-made process equipment. This could be furnaces in which batches of plates with the precursor layers are exposed to a selenium-containing atmosphere or an in-line reaction chamber in which the plates or rolls are continuously transported through an environment with selenium and substrate temperature control [283].

Elemental coevaporation of the $Cu(InGa)Se_2$ layer requires custom-made equipment including specially designed evaporation sources for uniform deposition of large-area substrates with accurate control over long deposition times (many hours). In-line evaporation using linear sources is a straightforward approach that is being developed at several laboratories and companies. A sketch of such a piece of equipment is illustrated in Figure 13.21. The configuration shown has sources designed for downward evaporation which requires a more complicated source design than sources which evaporate upward. But the downward evaporation is preferable for glass substrates because support underneath the glass as it moves through the process makes higher substrate temperature possible than if the glass were suspended above the sources.

The chemical bath deposition of CdS or Cd-free buffer layers is suitable for low-cost batch processing, in that it is a surface-controlled process that requires a limited solution volume. The equipment for dipping batches of $Cu(InGa)Se_2$-coated substrate plates is relatively simple, even for large areas. It is now commercially available or can be custom-made. However the generation of liquid hazardous waste is an added expense in this process. High utilization or continuous approaches for chemical bath deposition have been proposed [284], as well as recycling procedures [285].

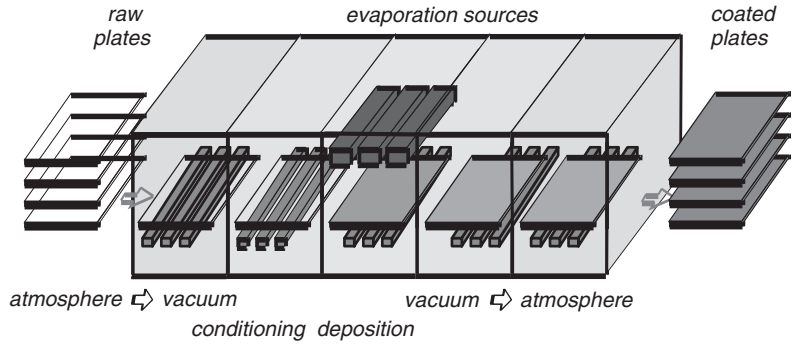


Figure 13.21 In-line coevaporation system for Cu(InGa)Se_2 with linear evaporation sources above the substrate plates and heaters below them [Courtesy of Zentrum für Sonnenenergie- und Wasserstoff-Forschung (ZSW)]. Reproduced by permission of Michael Powalla, ZSW Stuttgart, 2001

Alternative buffer layer processes such as ALD and ILGAR have potential manufacturing advantages due to the avoidance of wet processing and waste generation, but large-scale, high-throughput manufacturing equipment is still in development.

Chemical vapor-deposited doped ZnO as an alternative to sputtering is typically done as a batch process with a relatively small number of substrate plates deposited per run. Throughput will eventually become an issue. However, in-line CVD processes have been developed, for example, in the manufacture of amorphous silicon solar modules.

13.6.2 Module Fabrication

Two principle approaches to module fabrication are used in industry and are discussed below: (1) monolithic integration, where the films are deposited on the large area of the module and individual cells are prepared by scribing and interconnection; and (2) preparation of smaller area cells (typically a few hundred cm^2), which are equipped with a grid for current collection and which are soldered together to form strings and then modules, much as in Si wafer technology.

Soda-lime float glass is the substrate material that so far has given the best results in terms of both performance and reproducibility. It fulfills criteria on cost (3–4 $\$/m^2$ in large volumes), smoothness, and stability, so it is well suited for commercial production. One limitation that needs to be addressed in the development of production processes is that soda-lime glass starts to soften above 500 °C. At the same time, the best PV properties of Cu(InGa)Se_2 are achieved at growth temperatures above 500 °C. Plastic deformation due to glass softening is not acceptable in a module production process and careful optimization of the time–temperature profile is needed to minimize the deformation. Development of glass with higher softening temperature, compatibility with Cu(InGa)Se_2 processing in terms of thermal expansion and Na, and competitive cost, could provide a big benefit for Cu(InGa)Se_2 manufacturing.

An essential opportunity for cost advantage with thin film PV modules compared with silicon wafer–based PV modules is the possibility of monolithic interconnection to greatly simplify module fabrication. This allows modules to be fabricated directly, instead of first making cells followed by tabbing and stringing to make the series interconnection as required for Si-wafer solar cells. A typical monolithic interconnection is illustrated schematically in Figure 13.22. The most common way to make the patterning is by using laser ablation for the Mo patterning (P1) and mechanical

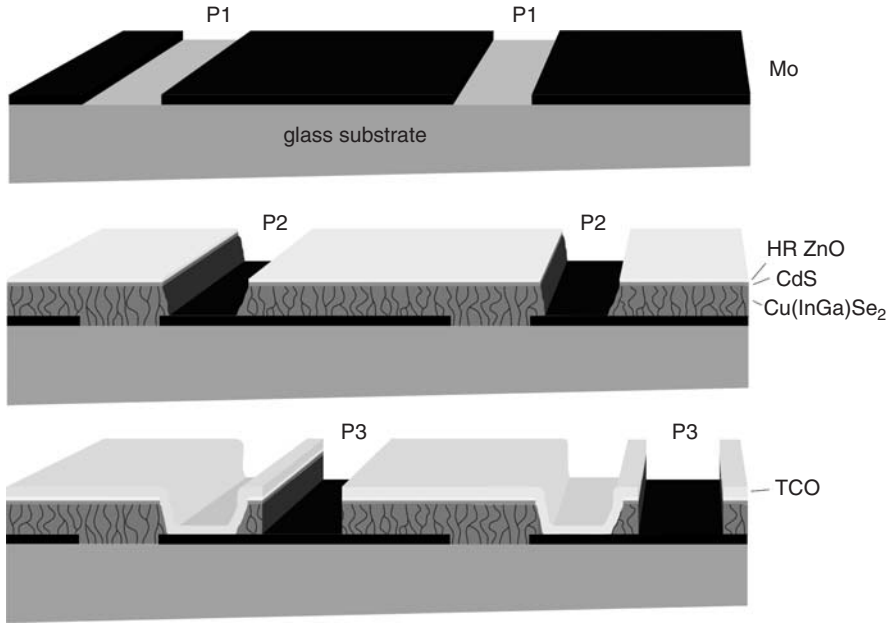


Figure 13.22 Schematic description of the manufacturing steps to make monolithic interconnections for thin film Cu(InGa)Se_2 PV modules

scribing for the two subsequent patterning steps (P2 and P3). With careful optimization, the total width of the interconnect, which is dead area with regards to performance, has been reduced to less than 200 μm [129]. The cell width depends on the current density (and hence the bandgap or relative Ga content) of the cell and the sheet resistance of the TCO layer (since the Mo sheet resistance is typically much smaller). A procedure for optimization is discussed in Chapter 17. The scribed cell width determines the current, and the number of cells in series (limited by the overall module dimensions) then determine the module voltage.

A typical schematic flow for module processing is shown in Figure 13.23, although many variations may be used for deposition and interconnection. The final fabrication steps include attachment of electrical wires and busbars. These are metal strips that can be soldered, welded, or glued to contact areas near the edges of the substrate plates. Before lamination with a front coverglass, the thin film layers are removed from the outer rim of the substrate plate in order to improve the adhesion to the lamination material, which is usually ethylene vinyl acetate (EVA).

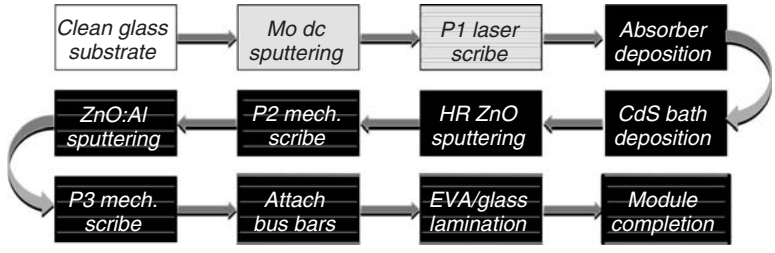


Figure 13.23 Process sequence for manufacturing Cu(InGa)Se_2 modules

Edge sealing, attachment of leads or a junction box, and framing finish the product, although a frame can be omitted for some applications.

Flexible substrate materials are attractive, both to make a lightweight flexible product with advantages for certain applications and to deposit the thin film materials in roll-to-roll processes, which are potentially cost-advantageous. Substrate materials that have shown promising results include polyimide, titanium, and stainless steel [99]. The drawbacks of polyimide are low temperature tolerance, since the best polyimide films readily available can only withstand 400–450 °C, and high thermal expansion. A limitation of titanium and steel is their conductivity, which means that an electrically isolating layer between the foil and the Mo back contact [101] is needed to enable monolithic series-interconnection of the cells. Again, commercial development of improved substrates, in this case either a higher temperature polymer or insulator-coated foil, is desirable.

With a conductive foil substrate, since monolithic integration is not possible, the substrate is cut into individual cells onto which collection grids are applied, typically by printing, similar to Si wafer cells. The cells are mechanically interconnected by tabbing and stringing or in a shingle-lapping configuration in which the back of one cell at its edge overlaps the top of the next cell, connected with a solder, conductive epoxy or tape. The extra handling of individual cells increases the number of processing steps and adds to the manufacturing cost. However, the ability to measure each cell and bin them according to their performance can relax constraints on production yield.

For a flexible module, of course no front coverglass is used. Instead the module needs an encapsulation layer whose requirements are very demanding. This encapsulant must maintain flexibility, transparency, UV-resistance and keep moisture ingress to a very low level for the full rated lifetime of the module. For a lifetime of 20–30 years no such material is currently available.

13.6.3 Module Performance and Stability

The efficiencies of the best cells, minimodules and full-scale production modules decreases as the area increases. There are inherent losses associated with making modules instead of cells, both from series resistance and from inactive device area. In an optimized, series-interconnected thin film module design, these kinds of losses correspond to about 1% absolute unit of efficiency. With a more advanced design using metal grids for interconnection, interconnect losses can be reduced [286]. Module performance is also critically dependent on the homogeneity of the material and device properties, including $Cu(InGa)Se_2$ composition, over large areas. The fact that modules need thicker TCOs to avoid excessive series resistances, which leads to higher free carrier absorption, has been discussed above (Section 13.4.5) Another kind of difference between modules and record cells is associated with the freedom to use higher process temperatures with small substrates that are less sensitive to glass deformation. Similarly, laboratory processes may use very low deposition rates to achieve optimum qualities while manufacturing processes must require high throughput which can necessitate much higher deposition rates. Thus results on small-area cells are not necessarily relevant to commercial module fabrication, but indicate the potential of the materials.

In a product the initial efficiency is of little interest if it deteriorates after some time in operation. $Cu(InGa)(SeS)_2$ modules fabricated by ARCO Solar and later Siemens Solar showed stable performance in field tests over more than 12 years [4] and this has been further extended to two decades [287]. Several manufacturers have since shown similar stability in outdoor tests over shorter durations. On the other hand, degradation of unencapsulated cells has been observed after exposure to 85% relative humidity at 85 °C for 1000 h [288], the so-called damp heat test, which is one of the certification tests in the IEC 61646 protocol. A loss of power was due to decreased fill factor caused by increased series resistance, attributed to degradation of ZnO, and to a lesser extent, a degradation of the junction diode which reduced V_{OC} and fill factor.

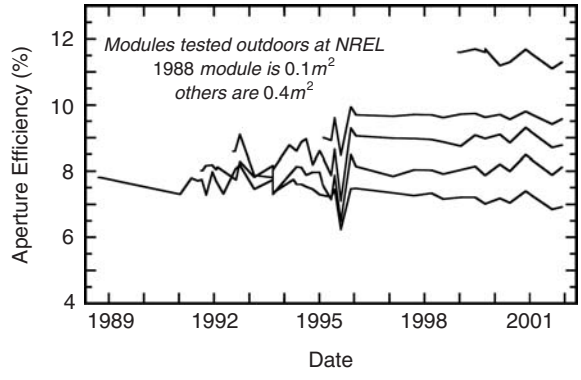


Figure 13.24 Examples of outdoor testing results at NREL of Cu(InGa)Se_2 modules showing stability over 12 years. Fluctuations in years 1992–1996 are due to changes in testing conditions. (Data courtesy of Shell Solar Industries)

While the damp heat test is severe for an exposed cell, it shows the need for encapsulation techniques that minimize the exposure of the thin film materials to moisture. Loss in output of modules after damp heat exposure could be completely recovered after light exposure so a modified damp heat test which includes illumination was proposed for Cu(InGa)Se_2 [289].

The outdoor module performance demonstrated in Figure 13.24 shows that Cu(InGa)Se_2 PV modules have the stability and performance to compete in any power application, be it stand-alone or grid-connected. Thin film modules have a great advantage over silicon-wafer PV for consumer applications in which the power needed may be relatively small. The large substrate plates, which have a power of 80 W_P or more, can easily be cut into smaller pieces, to essentially any power specification. This is much less costly than making small crystalline-silicon modules in which each cell has to be cut into pieces before assembling the modules. Additionally, the patterning structure of the interconnects can be designed to fit a large variety of shape and voltage requirements. For example, one manufacturer has made a number of 80 W modules using the same size module, but having a range of current and voltage outputs, determined by trading off the scribed cell width against number of cells. Aesthetically, the solid black appearance of Cu(InGa)Se_2 and other thin film modules may be preferred to the nonuniform bluish appearance of most silicon-wafer modules in building-integrated applications.

Finally, for space applications, Cu(InGa)Se_2 thin film solar cells offer potential advantages since the radiation tolerance is higher than in crystalline-silicon solar cells [5, 6]. The potential to use a lightweight plastic or foil substrate could lead to solar cells with very high specific power, that is, power divided by mass, which is critical for some space applications (see Chapter 9). However, Cu(InGa)Se_2 space solar cell technology has not yet reached a commercial stage.

13.6.4 Production Costs

Material costs have direct and indirect components, and depend on the material yield of the deposition processes. The direct material costs, that is, the cost of the feedstock, will not be reduced by an increased volume of the production, depending only on the feedstock market price and how much material is needed in the film. Indirect costs, including preparation of sputtering targets or other source materials, will be reduced when production volumes are sufficiently large. The material yield, or fraction of the source material that ends up in the film, can vary from less than 50% for some

thin film processes to greater than 90%, for example for electrodeposition or ink-based processes. While in general, the direct material costs are very small, uncertainty in the future availability of some materials, in particular In and Se, may lead to increased costs. However, the greatest materials costs are, and will likely remain, in the non-PV components including substrates, encapsulation, wiring, etc. [290].

In addition to materials, the other main production cost for thin film modules is the capital cost of the equipment. To first order, any large-scale automated deposition equipment will have comparable price. Therefore, the throughput or production capacity will be very important for determining the capital cost. The commercial availability of large-area deposition systems depends on the specific process, although in all cases equipment manufacturers are developing and beginning to offer manufacturing-scale deposition equipment and even turnkey production lines specifically for thin film PV. The biggest cost advantage related to equipment will come in the replication of process equipment, and eventually the full process line, once the engineering, optimization, and equipment building have all been finalized [290].

Costs around 20 $\$/m^2$ for each thin film deposition or process step may be acceptable in pilot production, but clear pathways toward costs in the range 1–5 $\$/m^2$ for large-volume production need to be identified. Throughput has a direct effect on cost. In an in-line process, this will depend on the substrate width and linear speed, which fundamentally depends on the deposition rate and desired thickness of the layer. If the deposition rate is relatively low, it can be compensated by having a long deposition zone in the system, for example, by having multiple targets in a sputtering system with only a relatively small increase in capital cost.

All cost advantages for thin films are lost if the production is not completed with high yield. The overall manufacturing yield can be broken down into electrical yield and mechanical yield. The electrical yield reflects the module reproducibility since it is the fraction of the modules produced which fulfill minimum performance criteria. The mechanical yield is the fraction of the substrates entering the production line that make it to the end. Mechanical losses result from broken glass substrates or malfunctioning equipment. In general, the overall yield should be well over 90%.

Another manufacturing cost is the energy usage. One measure is the energy payback time, defined as the time for a module to recover the primary energy consumption throughout its life-cycle by its own energy production. One recent comparison of large desert-based systems reported an energy payback time for $Cu(InGa)Se_2$ with 11% rated efficiency of 1.6 years – as good as or better than other PV technologies [291] (for further discussion see Chapter 1).

Production-cost analyses result in a range of projected manufacturing costs and depend on many assumptions and different processes. Various scenarios show that manufacturing costs can be lower than $\$1/W_p$ and as low as 0.4–0.6 $\$/W_p$ for sufficiently large-volume (>100 MW scale) manufacturing [290].

13.6.5 Environmental Concerns

One of the environmental issues related to the materials in $Cu(InGa)Se_2$ modules is the availability of less common elements. The content of the critical materials in grams per kW_P has been calculated assuming 12% module efficiency and the result is compared with the amount refined annually in Table 13.6 [292]. The fourth column expresses how much module power could be obtained from the amount refined annually and the last column shows a similar calculation based on the reserves of the various elements. Owing to uncertainties in estimates of reserves, or maximum resources, Table 13.6 gives only an indication of where, and at what level, potential problems in material supply may occur. It is clear that In supply is the potential bottleneck as regards primary material supply. In is primarily obtained as a by-product of zinc mining so its availability is tied to known Zn deposits

Table 13.6 Critical materials in Cu(InGa)Se_2 modules with respect to primary supply (after Andersson B, Azar C, Holmberg J, Karlsson S, *Energy* **5**, 407–411 (1998) [292]

Element	Material content[g/kW_P]	Amount refined[kton/y]	Amount refined/content[GW_P/y]	Reserves/ content[TW_P]
Mo	42	110	2600	130
Cu	17	9000	529 000	30 000
In	23	0.13	5.7	0.1
Ga	5	0.06	12	2.2
Se	43	2	46	1.9
Cd	1.6	20	12 500	330
Zn	37	7400	200 000	4100

since there has been little exploration for primary sources of In. Based on assumptions of improved extraction rates, recycling, utilization in manufacturing as well as thickness and performance of the modules, it was estimated that sustainable Cu(InGa)Se_2 manufacturing rates of ~100 GW/yr could be achieved by 2050 with known supplies of In [293]. Still, In availability will ultimately limit the capacity for Cu(InGa)Se_2 module manufacturing. This can be mitigating by developing In-free alloys of which the kesterite family of materials, including Cu_2ZnSnS_4 and $Cu_2ZnSnSe_4$ are most promising with efficiencies close to 10% [294].

$CuInSe_2$ toxicity has been studied by administering it to rats [295]. Even at high doses negligible effects were detected. A lowest observed adverse effect level (LOAEL) of 8.3 μg/kg/day for humans was derived from these studies.

The other substances that constitute Cu(InGa)Se_2 modules are largely nontoxic, except for Cd. Many health aspects of its use in PV manufacturing have been studied by Fthenakis and Moskowitz [296]. Chemical bath deposition of CdS is the process step that presents the greatest health concerns due to the use of Cd, thiourea, and the generation of waste solutions. In electrodeposition of CdTe, which also is a wet process using Cd precursors, it was found that the greatest health hazards from Cd are from dust generated during feedstock preparation and from fine particles near the baths [296]. Biological monitoring at a process station showed that exposures can be maintained at a level that presents no risk to workers. Thiourea is a toxic and carcinogenic substance that also presents an exposure risk. Rinse water and dilute solutions of acids and Cd-compounds can be treated by a two-stage precipitation/ion exchange process. The Cd can be removed, and recycled, down to 1–10 ppb levels [296].

Most Cu(InGa)Se_2 processes use elemental Se, but the forms that are handled are solid shots or pellets that give off very little dust that could be inhaled. Elemental Se is considered to have a relatively low biological activity, but many compounds are very active and highly toxic. In particular, hydrogen selenide, a gas used in some selenization processes, is extremely toxic with an "immediately dangerous to life and health" (IDLH) value of only 2 ppm [296].

There are also environmental concerns for the hazards during the operation of Cu(InGa)Se_2 modules with one potential risk being the leaching of critical materials into rainwater. This only happens if a module is broken or crushed, exposing the normally well-encapsulated active layers. An experimental study of the emissions of toxic elements into rainwater from crushed $CuInSe_2$ modules and into soil exposed to the water concluded that no acute danger to humans or the environment is likely to occur [297]. The main hazard during the active life of the $CuInSe_2$ modules is related to fire accidents. A study of the potential risks associated with fires in PV power plants shows that they are very limited [298]. A fire in a commercial-size system could result in harmful concentrations up to 300 m downwind of the fire if most of the $CuInSe_2$ materials are released. With release of 10% of

the $CuInSe_2$ materials, concentrations were not harmful even under worst-possible meteorological conditions. The study concluded that there are no immediate risks to the public from fires in sites with $CuInSe_2$ modules.

Concerns for disposal of Cu(InGa)Se_2 have also been tested with respect to leachability. Zn, Mo, and Se are eluted in the highest amounts. On the basis of landfill criteria, $CuInSe_2$ modules will pass requirements in both Germany and the United States [295]. Because of the low volume and leaching rates of critical elements from $CuInSe_2$ modules, they will not be classified as hazardous waste according to most US regulations [299].

The evolution of environmental regulations, disposal options, and economics makes recycling increasingly important. In large-scale use of Cu(InGa)Se_2 modules, the supply of rare elements, in particular indium, but also selenium and gallium, provides a further motivation for recycling. The cost of recycling may be favorably offset if module materials can be reclaimed. In particular, if the glass sheets can be salvaged and reused, there will be a net gain associated with the recycling procedure. Thus, recycling may be an important consideration in the choice of encapsulation method. Double-glass structures are functional and may reduce the release of $CuInSe_2$ materials during fires, but may increase the costs for recovering metals and reusing glass plates [299].

13.7 THE Cu(InGa)Se_2 OUTLOOK

Clearly, there has been tremendous progress in Cu(InGa)Se_2 solar cells as evidenced by the high module and cell efficiencies fabricated by many groups and companies, the range of deposition and device options that have been developed, and the growing base of science and engineering knowledge of these materials and processes. There is good reason to be optimistic that cell efficiencies and module performance as well yield will continue to improve. Still, there is an incomplete understanding of many of the critical problems associated with the semiconductor processing and an ongoing need to devote time and research focus, both at the laboratory scale (to address fundamental issues) and on the pilot line (to address equipment and scale-up problems) and to validate processes and improve module efficiencies.

From their earliest development, $CuInSe_2$-based solar cells, along with other thin film PV materials including CdTe, and amorphous Si, attracted interest because of their potential to be manufactured at a lower cost than Si wafer-based PV and this has now been emphatically demonstrated with the large scale CdTe manufacturing by First Solar which now makes the lowest-priced modules on the market (see Chapter 14). However, after more than 30 years of research and development, Cu(InGa)Se_2, manufacturing is only now moving into the large-volume production where similar economies of scale will enable real cost advantages to be tested. One of the significant developments in recent years has been the investment of auxiliary and support industries that are developing deposition equipment, diagnostic tools, and new materials such as substrates and encapsulation materials that are designed specifically for thin film PV and Cu(InGa)Se_2 in particular. These companies apparently now recognize the vast commercial potential of thin film PV.

A critical question must be asked: what needs to be done to ensure that Cu(InGa)Se_2 solar cell technology reaches its potential for large-scale power generation?

Part of the answer is to address the critical need for the accelerated development of mature manufacturing technology and standardized deposition equipment and processes based on well-developed engineering models. Also, new diagnostic and process-control tools will have to be developed. This requires fundamental materials and device knowledge to determine what properties can be measured in a cell or module fabrication process that can act as reliable predictors of final

performance. Better processes, equipment, and control based on a more solid knowledge base can directly translate to higher throughput, yield, and performance.

Despite many advances, there is still a need for improvement in the fundamental science of the materials and devices. Significant improvements in efficiency will only come from increased V_{OC} so the chemical and electronic nature of the defects that limit it, and their origin, must be understood. This can contribute to a comprehensive model for the growth of Cu(InGa)Se_2, relating processing parameters to defect formation, junction formation, and device limitations. In addition, the fundamental understanding of the role of sodium and the nature of the grain boundaries and free surface remains incomplete. A greater understanding of the role of the CdS layer and the chemical bath process might enable alternative materials that do not contain cadmium and have wider bandgap to be utilized with greater efficiency and reproducibility.

Finally, the potential of Cu(InGa)Se_2 will be achieved most quickly by fully exploiting applications that take advantage of the unique benefits that arise, for example, from relatively high efficiency with flexibility, light weight, or radiation resistance. Also Cu(InGa)Se_2 in either flexible or rigid products may be preferred in many cases for building-integrated PV [300]. This will increase demand to spur increased commercial development and hasten achieving the very large economies of scale needed to achieve very-low-cost manufacturing.

A second critical question to be asked is: what might be the breakthroughs that could lead to the next generation of thin film Cu(InGa)Se_2-based solar cells?

Further development of wide-bandgap alloys to enable cells to be made with $E_g \geq 1.5\,\text{eV}$ without any decrease in performance will have several benefits for module fabrication and performance, as discussed in Section 13.5.4. In addition, development of a high-efficiency cell with $E_g \approx 1.7$ eV is a prerequisite for tandem cells based on the polycrystalline thin films to be developed. A monolithic tandem cell, while technically very challenging, has the potential to attain efficiencies of 25% or more.

Low-temperature processing of the Cu(InGa)Se_2 layer without loss of efficiency in the final solar cell can have significant additional benefits. With lower substrate temperature, alternative substrate materials, such as a flexible polymer web, can be utilized. In addition, lower T_{SS} can reduce thermally induced stress on the substrate, allowing faster heat-up and cool-down, and decrease the heat load and stress on the entire deposition system. Similarly, there would be cost and processing advantages to a cell structure that enables the use of a Cu(InGa)Se_2 layer much less than 1 μm. Reduced thickness obviously reduces the quantity of materials used and may increase throughput rates. Development of In-free alloys can also have significant benefit.

With all these challenges to improve the fundamental knowledge behind Cu(InGa)Se_2 materials and devices and to develop new manufacturing technology and breakthrough advancements, research and development on Cu(InGa)Se_2 and related materials remains exciting and promising. All of the reasons for the initial excitement over the potential for thin film Cu(InGa)Se_2 remain valid. The high efficiency, demonstrated stability, and tolerance to material and process variations give great hope that it will be a major contributor to our solar electric future.

REFERENCES

1. Repins I *et al,. Prog. Photovolt*. **16**, 235–239 (2008).
2. Repins I *et al., Proc. SPIE 2009 Solar Energy + Technology Conf.* (2009).
3. Tanaka T *et al., Proc. 17th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 989–994 (2001).
4. Wieting R, *AIP Conf. Proc.* **462**, 3–8 (1999).
5. Burgess R *et al., Proc. 20th IEEE Photovoltaic Specialist Conf.*, pp 909–912 (1988).

6. Jasenek A *et al., Thin Solid Films* **387**, 228–230 (2001).
7. Hahn H *et al., Z. Anorg. Allg. Chem.* **271**, 153–170 (1953).
8. Shay J, Wernick J, *Ternary Chalcopyrite Semiconductors: Growth, Electronic Properties, and Application*, Pergamon Press, Oxford (1974).
9. Tell B, Shay J, Kasper H, *Phys. Rev.* **B4**, 4455–4459 (1971).
10. Tell B, Shay J, Kasper H, *J. Appl. Phys.* **43**, 2469–2470 (1972).
11. Wagner S, Shay J, Migliorato P, Kasper H, *Appl. Phys. Lett.* **25**, 434–435 (1974).
12. Shay J, Wagner S, Kasper H, *Appl. Phys. Lett.* **27**, 89–90 (1975).
13. Meakin J, *Proc. SPIE Conf. 543: Photovoltaics*, 108–118 (1985).
14. Kazmerski L, White F, Morgan G, *Appl. Phys. Lett.* **29**, 268–269 (1976).
15. Mickelsen R, Chen W, *Proc. 15th IEEE Photovoltaic Specialist Conf.*, pp 800–804 (1981).
16. Mickelsen R, Chen W, *Proc. 16th IEEE Photovoltaic Specialist Conf.*, pp 781–785 (1982).
17. Chen W *et al., Proc. 19th IEEE Photovoltaic Specialist Conf.*, pp 1445–1447 (1987).
18. Potter R, *Sol. Cells* **16**, 521–527 (1986).
19. Hedström J *et al., Proc. 23rd IEEE Photovoltaic Specialist Conf.*, pp 364–371 (1993).
20. Basol B *et al., Sol. Energy Mater. Sol. Cells* **43**, 93–98 (1996).
21. Gabor A *et al., Sol. Energy Mater. Sol. Cells* **4**, 247–260 (1996).
22. Tarrant D, Ermer J, *Proc. 23rd IEEE Photovoltaic Specialist Conf.*, pp 372–375 (1993).
23. Rocheleau R, Meakin J, Birkmire R, *Proc. 19th IEEE Photovoltaic Specialist Conf.*, pp 972–976 (1987).
24. Mitchell K *et al., IEEE Trans. Electron. Devices* **37**, 410–417 (1990).
25. Kazmerski L, Wagner S, Cu-Ternary Chalcopyrite Solar Cells, in Coutts T, Meakin J, (eds), *Current Topics in Photovoltaics*, pp 41–109, Academic Press, London (1985).
26. Haneman D, *Crit. Rev. Solid State Mater. Sci.* **14**, 377–413 (1988).
27. Rockett A, Birkmire R, *J. Appl. Phys.* **70**, R81–R97 (1991).
28. Rockett A, Bodegård M, Granath K, Stolt L, *Proc. 25th IEEE Photovoltaic Specialist Conf.*, pp 985–987 (1996).
29. Suri D, Nagpal K, Chadha G, *J. Appl. Cystallogr.* **22**, 578–83 (1989) (JCPDS 40-1487).
30. Ciszek T F, *J. Cryst. Growth.* **70**, 405–410 (1984).
31. Bondar I, Orlova N, *Inorg. Mater.* **21**, 967–970 (1985).
32. Neumann H, *Sol. Cells* **16**, 399–418 (1986).
33. Li P, Anderson R, Plovnick R, *J. Phys. Chem. Solids* **40**, 333–334 (1979).
34. Chattopadhyay K, Sanyal I, Chaudhuri S, Pal A, *Vacuum* **42**, 915–918 (1991).
35. Arushanov E, *et al., Physica B* **184**, 229–31 (1993).
36. Persson C, *Appl. Phys. Lett.* **93**, 072106 1–3 (2008).
37. Neumann H, *et al., phys. stat. sol. b* **108**, 483–87 (1981).
38. Nakanishi H Y, Endo S, Irie T, Chang B H, *Proc. Int. Conf. Ternary and Multinary Compounds.* 99–104 (1987).
39. Gödecke T, Haalboom T, Ernst F, *Z. Metallkd.* **91**, 622–634 (2000).
40. Ye J, Yoshida T, Nakamura Y, Nittono O, *Jpn. J. Appl. Phys.* 35, 395–400 (1996).
41. Zhang S, Wei S, Zunger A, *Phys. Rev. Lett.* **78**, 4059–4062 (1997).
42. Herberholz R *et al., Eur. Phys. J.* **6**, 131–139 (1999).
43. Wei S, Zhang S, Zunger A, *Appl. Phys. Lett.* **72**, 3199–3201 (1998).
44. Schroeder D, Rockett A, *J. Appl. Phys*. **82**, 4982–4985 (1997).
45. Wei S, Zhang S, Zunger A, *J. Appl. Phys.* **85**, 7214–7218 (1999).
46. Schuler S *et al., Phys. Rev. B* **69**, 045210 (2004).
47. Ruckh M *et al., Sol. Energy Mater. Sol. Cells* **41/42**, 335–343 (1996).
48. Alonso M *et al., Phys. Rev. B* **63**, 075203 1–13 (2001).
49. Alonso M *et al., Appl. Phys. A* **74**, 659–664 (2002).
50. Paulson P, Birkmire R, Shafarman W, *J. Appl. Phys.* **94**, 879–888 (2003).
51. Wei S, Zunger A, *Appl. Phys. Lett.* **72**, 2011–2013 (1998).

52. Hönes K, Eickenberg M, Siebentritt S, Persson C, *Appl Phys Lett*, **93** 092102 1–3 (2008).
53. Jaffe J, Zunger A, *Phys. Rev. B* **29**, 1882–1906 (1984).
54. Weinert H *et al*, *phys. stat. sol.* b, **81**, K59–61 (1977).
55. Arushanov E *et al*, *Physica B*, **184**, 229–31 (1993).
56. Noufi R, Axton R, Herrington C, Deb S, *Appl. Phys. Lett.* **45**, 668–670 (1984).
57. Neumann H, Tomlinson R, *Sol. Cells* **28**, 301–313 (1990).
58. Siebentritt S, *Thin Solid Films* **403–404** 1–8 (2002).
59. Heath J, Cohen J, Shafarman W, *J. Appl. Phys* **95**, 1000–1010 (2004).
60. Siebentritt S, *Thin Solid Films* **480–81**, 312–317 (2005).
61. Lee J, Cohen J, Shafarman W, *Thin Solid Films* **480–481**, 336–340 (2005).
62. Lany S, Zunger A, *Phys. Rev. Lett.* **93**, 156404 1–4, (2004).
63. Lany S, Zunger A, *Phys. Rev. Lett.* **100**, 016401 1–4 (2008).
64. Bauknecht A, Siebentritt S, Albert J, Lux-Steiner M, *J. Appl. Phys.* **89**, 4391–4400 (2001).
65. Siebentritt S, Rega N, Zajogin A, Lux-Steiner M, *phys. stat. sol. C* **1**, 2304–10 (2004).
66. Siebentritt S *et al., Appl. Phys. Lett.* **86**, 091909 1–3 (2005).
67. Dirnstorfer I *et al., phys. stat. sol. a* **168**, 163–175 (1998).
68. Shklovskii B, Efros A, *Electronic Properties of Doped Semiconductors*. 1984, Berlin: Springer-Verlag.
69. Bardeleben, H.J.v., *Solar Cells* **16**, 381–90, (1986).
70. Aubin, V., Binet, L., Guillemoles, J. F., *Thin Solid Films*, **431-2**, 167–71 (2003).
71. Turcu M, Kötschau I, Rau U, *J. Appl. Phys.* **91**, 1391–99 (2002).
72. Heath J *et al., Appl. Phys. Lett.* **80**, 4540 (2002).
73. Kiely C, Pond R, Kenshole G, Rockett A, *Philos. Mag. A* **63**, 2149–2173 (1991).
74. Chen J *et al., Thin Solid Films* **219**, 183–192 (1992).
75. Wada T, *Sol. Energy Mater. Sol. Cells* **49**, 249–260 (1997).
76. Lei C *et al*, *J. Appl. Phys.* 100, 073518 (2006).
77. Liao D, Rockett A, *J. Appl. Phys.* **91**, 1978–1983 (2002).
78. Schmid D, Ruckh M, Grunwald F, Schock H, *J. Appl. Phys.* **73**, 2902–2909 (1993).
79. Klein A, Jaegermann W, *Appl. Phys. Lett.* **74**, 2283–2285 (1999).
80. Gartsman K *et al., J. Appl. Phys.* **82**, 4282–4285 (1997).
81. Kylner A, *J. Electrochem. Soc.* **146**, 1816–1823 (1999).
82. Damaskinos S, Meakin J, Phillips J, *Proc. 19th IEEE Photovoltaic Specialist Conf.*, pp 1299–1304 (1987).
83. Cahen D, Noufi R, *Appl. Phys. Lett.* **54**, 558–560 (1989).
84. Kronik L, Cahen D, Schock H, *Adv. Mater.* **10**, 31–36 (1998).
85. Niles D, Al-Jassim M, Ramanathan K, *J. Vac. Sci. Technol., A* **17**, 291–296 (1999).
86. Yan Y *et al., Phys. Rev. Lett.*, **99**, 235504 (2007).
87. Persson C, Zunger A, *Phys. Rev. Lett.* **91**, 266401 1–4 (2003).
88. Hetzer M *et al., Appl. Phys. Lett.* **86**, 162105 1–3 (2005).
89. Lei C *et al*, *J. Appl. Phys.* **101**, 024909 1–5 (2007).
90. Yan Y, Noufi R, Al-Jassim M, *Phys. Rev. Lett.* **96**, 205501 1–4 (2006)
91. Seto J, *J. Appl. Phys.* **46**, 5247–54 (1975).
92. Siebentritt S, Schuler S, *J. Phys. Chem. Solids* **64**, 1621–26 (2003).
93. Rau U, Taretto K, Siebentritt S, *Appl. Phys. A* **96**, 221–34 (2009).
94. Bodegård M, Stolt L, Hedström J, *Proc. 12th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1743–1746 (1994).
95. Bodegård M, Granath K, Rockett A, Stolt L, *Sol. Energy Mater. Sol. Cells* **58**, 199–208 (1999).
96. Contreras M *et al., Prog. Photovolt.* **7**, 311–316 (1999).
97. Schlenker T, Laptev V, Schock H, Werner J, *Thin Solid Films* **480–481**, 29–32 (2005).

98. Boyd D, Thompson D, Kirk-Othmer *Encyclopaedia of Chemical Technology*, Vol. 11, 3rd Edition, 807–880, John Wiley & Sons, Inc. (1980).
99. Kessler F, Herrmann D, Powalla M, *Thin Solid Films* **480–481**, 491–498 (2005).
100. Birkmire R, Eser E, Fields S, Shafarman W, *Prog. Photovolt.* **13**, 141–148 (2005).
101. Herz K *et al., Thin Solid Films* **431–432**, 392–397 (2003).
102. Palm J *et al., Thin Solid Films* **431–432**, 514–522 (2003).
103. Probst V *et al., Proc. 1st World Conf. Photovoltaic Solar Energy Conversion*, pp 144–147 (1994).
104. Bodegård M, Granath K, Stolt L, *Thin Solid Films* **361–362**, 9–16 (2000).
105. Rudmann D, Brémaud D, Zogg H, Tiwari A, *J. Appl. Phys.* **97**, 084903 1–5 (2005).
106. Vink T, Somers M, Daams J, Dirks A, *J. Appl. Phys.* **70**, 4301–4308 (1991).
107. Wada T, Kohara N, Nishiwaki S, Negami T, *Thin Solid Films* **387**, 118–122 (2001).
108. Orgassa K, Schock H, Werner J, *Thin Solid Films* **431–432**, 387–391 (2003).
109. Mattox D, *Handbook of Physical Vapor Deposition (PVD) Processing*, Noyes Publ., Park Ridge, NJ (1998).
110. Jackson S, Baron B, Rocheleau R, Russell T, *Am. Inst. Chem. Eng. J.* **33**, 711–720 (1987).
111. Shafarman W, Zhu J, *Thin Solid Films* **361–2**, 473–477 (2000).
112. Klenk R, Walter T, Schock H, Cahen D, *Adv. Mater.* **5**, 114–119 (1993).
113. Kessler J *et al., Proc. 12th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 648–652 (1994).
114. Gabor A *et al., Appl. Phys. Lett.* **65**, 198–200 (1994).
115. Hasoon F *et al., Thin Solid Films* **387**, 1–5 (2001).
116. Stolt L *et al., Appl. Phys. Lett.* **62**, 597–599 (1993).
117. Ishizuka S *et al., J. Appl. Phys.* **100**, 096106 1-3 (2006).
118. Hanna G *et al. Thin Solid Films* **431–432** 31–36 (2003).
119. Sakurai K *et al., Prog. Photovolt: Res. Appl.* **12**, 219–234 (2004).
120. Stolt L, Hedström J, Sigurd D, *J. Vac. Sci. Technol.* **A3**, 403–407 (1985).
121. Powalla M, Voorwinden G, Dimmler B, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1270–1273 (1997).
122. Eisgruber I *et al., Thin Solid Films* **408**, 64–72 (2002).
123. Scheer R *et al., Appl. Phys. Lett.* **82**, 2091–2093 (2003).
124. Nishitani M, Negami T, Wada T, *Thin Solid Films* **258**, 313–316 (1995).
125. Negami T *et al., Mater. Res. Soc. Symp.* **426**, 267–278 (1996).
126. Grindle S, Smith C, Mittleman S, *Appl. Phys. Lett.* **35**, 24–26 (1979).
127. Chu T, Chu S, Lin S, Yue J, *J. Electrochem. Soc.* **131**, 2182–2185 (1984).
128. Alberts V, *Semicond. Sci. Technol.* **22**, 585–592 (2007).
129. Kushiya K *et al. Thin Solid Films* ***517***, 2108–2110 (2009).
130. Kapur V, Basol B, Tseng E, *Sol. Cells* **21**, 65–70 (1987).
131. Sato H *et al., Proc. 23rd IEEE Photovoltaic Specialist Conf.*, pp 521–526 (1993).
132. Kessler J, Dittrich H, Grunwald F, Schock H, *Proc. 10th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 879–882 (1991).
133. Oumous H *et al., Proc. 9th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 153–156 (1992).
134. Palm J, Probst V, Karg F, *Solar Energy* **77**, 757–765 (2004).
135. Mooney G *et al., Appl. Phys. Lett.* **58**, 2678–2680 (1991).
136. Sugiyama M, *et al., J. Crystal Growth* **294**, 214–217 (2006).
137. Verma S, Orbey N, Birkmire R, Russell T, *Prog. Photovolt.* **4**, 341–353 (1996).
138. Wolf D, Müller G, *Thin Solid Films* **361–2**, 155–161 (2000).
139. Orbey N, Norsworthy G, Birkmire R, Russell T, *Prog. Photovolt.* **6**, 79–86 (1998).
140. Hergert F *et al., Journal of Physics and Chemistry of Solids* **66**, 1903–1907 (2005).

141. Dittrich H, Prinz U, Szot J, Schock H, *Proc. 9th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 163–166 (1989).
142. Jensen C, Tarrant D, Ermer J, Pollock G, *Proc. 23rd IEEE Photovoltaic Specialist Conf.*, pp 577–580 (1993).
143. Marudachalam M *et al., Appl. Phys. Lett.* **67**, 3978–3980 (1995).
144. Nagoya Y, Kushiya K, Tachiyuki M, Yamase O, *Solar En. Mat. Solar Cells* **67**, 247–253 (2001).
145. Alberts V, *Mat. Science Eng.* **B107** 139–147 (2004).
146. Hanket G, Shafarman W, McCandless B, Birkmire R, *J. Appl. Phys.* **102**, 074922 (2007).
147. Hanket G, Shafarman W, Birkmire R, *Proc. 4th World Conf. Photovoltaic Solar Energy Conversion*, pp 560–563 (2006).
148. Thornton J, Lomasson T, Talieh H, Tseng B, *Sol. Cells* **24**, 1–9 (1988).
149. Talieh H, Rockett A, *Sol. Cells* **27**, 321–329 (1989).
150. Guenoun K, Djessas K, Massé G, *J. Appl. Phys.* **84**, 589–595 (1998).
151. Murali K, *Thin Solid Films* **167**, L19–L22 (1988).
152. Galindo H *et al., Thin Solid Films* **170**, 227–234 (1989).
153. Abernathy C *et al., Appl. Phys. Lett.* **45**, 890 (1984).
154. Kazmerski L, Ireland P, White F, Cooper R, *Proc. 13th IEEE Photovoltaic Specialist Conf.*, 184–189 (1978).
155. Potter R, Eberspacher C, Fabick L, *Proc. 18th IEEE Photovoltaic Specialist Conf.*, 1659–1664 (1985).
156. Mitchell K, Liu H, *Proc. 20th IEEE Photovoltaic Specialist Conf.*, 1461–1468 (1988).
157. Cashman R, *J. Opt. Soc. Am.* **36**, 356 (1946).
158. Kitaev G, Uritskaya A, Mokrushin S, *Sov. J. Phys. Chem.* **39**, 1101 (1965).
159. Kainthla R, Pandya D, Chopra K, *J. Electrochem. Soc.* **127**, 277–283 (1980).
160. Kaur I, Pandya D, Chopra K, *J. Electrochem. Soc.* **127**, 943–948 (1980).
161. Lincot D, Ortega-Borges R, *J. Electrochem. Soc.* **139**, 1880–1889 (1992).
162. Lincot D, Ortega-Borges R, Froment M, *Philos. Mag. B* **68**, 185–194 (1993).
163. Nakada T, Kunioka A, *Appl. Phys. Lett.* **74**, 2444–2446 (1999).
164. Kylner A, Rockett A, Stolt L, *Solid State Phen.* **51–52**, 533–539 (1996).
165. Hashimoto Y *et al., Sol. Energy Mater. Sol. Cells* **50**, 71–77 (1998).
166. Kylner A, Lindgren J, Stolt L, *J. Electrochem. Soc.* **143**, 2662–2669 (1996).
167. Kylner A, Niemi E, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1321–1326 (1997).
168. Kessler J *et al., Tech. Digest PVSEC-6*, pp 1005–1010 (1992).
169. Yu P, Faile S, Park Y, *Appl. Phys. Lett.* **26**, 384–385 (1975).
170. Tell B, Wagner S, Bridenbaugh P, *Appl. Phys. Lett.* **28**, 454–455 (1976).
171. Kazmerski L, Jamjoum O, Ireland P, *J. Vac. Sci. Technol.* **21**, 486–490 (1982).
172. Heske C *et al., Appl. Phys. Lett.* **74**, 1451–1453 (1999).
173. Kylner A, *J. Electrochem. Soc.* **143**, 1816–1823 (1999).
174. Ramanathan K *et al., Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, pp 477–482 (1998).
175. Wang L *et al., MRS Symp.* **569**, 127–132 (1999).
176. Leskelä M, Ritala M, *Thin Solid Films* **409**, 138–146 (2002).
177. Hunger R *et al., Thin Solid Films* **515**, 6112–18 (2007).
178. Muffler, M *et al., Proc. 28th IEEE Photovoltaic Specialist Conf.*, 610–613 (2000).
179. Platzer-Björkman C, Kessler J, Stolt L, *Proc. 3rd World Conf. Photovoltaic Energy Conversion*, pp 461–64 (2003).
180. Eisele W *et al., Sol. Energy Mater. Sol. Cells* **75**, 17–26 (2003).
181. Ennaoui A *et al., Sol. Energy Mater. Sol. Cells* **67**, 31–40 (2001).
182. Siebentritt S *et al., Prog. Photovolt.* **12**, 333–38 (2004).

183. Ohtake Y *et al., Japanese J. Appl. Phys.* **34**, 5949–55 (1995).
184. Yamada A, Chaisitsak S, Ohtake Y, Konagai M *Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, pp 1177–1180 (1998).
185. Tokita Y, Chaisitsak S, Yamada A, Konagai M, *Sol. Energy Mater. Sol. Cells* **75**, 9–15 (2003).
186. Hariskos D *et al., Sol. Energy Mater. Sol. Cells* **41/42**, 345–53 (1996).
187. Naghavi N *et al., Mat. Res. Soc. Symp. Proc.* **763**, 465–70 (2003).
188. Spiering S *et al., Thin Solid Films* **431–2**, 359–63 (2003).
189. Allsop N *et al., Prog. Photovolt.* **13**, 607–616 (2005).
190. Strohm A *et al., Thin Solid Films* **480**, 162–67 (2005).
191. Hariskos D *et al., Proc. 19th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1894–97 (2004).
192. Törndahl T, Platzer-Björkman C, Kessler J, Edoff M, *Prog. Photovolt.* **15**, 225–235 (2007).
193. Ramanathan K *et al., Proc. 29th IEEE Photovoltaic Specialist Conf.*, pp 523–26 (2003).
194. Bär M *et al., Sol. Energy Mater. Sol. Cells* **75**, 101–07 (2003).
195. Negami T *et al., Proc. 29th IEEE Photovoltaic Specialist Conf.*, pp 656–59 (2002).
196. Glatzel T *et al., Proc. 14th Int. Photovoltaic Science Engineering Conf.*, (2004).
197. Lewis B, Paine D, *MRS Bull.* **25**, 22–27 (2000).
198. Menner R, Schäffler R, Sprecher B, Dimmler B, *Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, pp 660–663 (1998).
199. Ruckh M *et al., Proc. 25th IEEE Photovoltaic Specialist Conf.*, pp 825–828 (1996).
200. Westwood W, Reactive Sputter Deposition, in Rossnagel S, Cuomo J, Westwood W, (eds), *Handbook of Plasma Processing Technology*, Chap. 9, Noyes Publ., Park Ridge, NJ (1990).
201. Hagiwara Y, Nakada T, Kunioka A, *Solar Energy Mat. Solar Cells* **67**, 267–71 (2001).
202. Rau U, Schmidt M, *Thin Solid Films* **387**, 141–146 (2001).
203. Kessler J *et al., Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 775–778 (2000).
204. Cooray N, Kushiya K, Fujimaki A, Okumura D, *Jpn. J. Appl. Phys.* **38**, 6213–6218 (1999).
205. Jackson P *et al., Prog. Photovolt.*, **15** 507–519 (2007).
206. Klenk R, Schock H, Bloss W, *12th Euro. Conf. Photovoltaic Solar Energy Conversion*, 1588–1591 (1994).
207. Shafarman W, Klenk R, McCandless B, *J. Appl. Phys.* **79**, 7324–7328 (1996).
208. Eron M, Rothwarf A, *Appl. Phys. Lett.* **44**, 131–33 (1984).
209. Ohnesorge B *et al., Appl. Phys. Lett.* **73**, 1224–1227 (1998).
210. Lundberg O, Bodegard M, Malmstrom J, Stolt L, *Prog. Photovolt* **11**, 77–88 (2003).
211. Shafarman W, Huang R, Stephens S, *Proc. 4th World Conf. Photovoltaic Solar Energy Conversion*, 420–423 (2006).
212. Gloeckler M. Sites J, *J. Appl. Phys.* **98**, 103713–1-7 (2005).
213. Malmstrom J, Schleussner S, Stolt L, *App. Phys. Lett.* **85**, 2635–2637 (2004).
214. Guo S, Shafarman W, Delahoy A, *J. Vac. Sci. Tech.* **24**, 1524–1529 (2006).
215. Rau U, Schock H, Cu(InGa)Se_2 Solar Cells, in *Clean Electricity from Photovoltaics*, MD Archer, R. Hill (eds), Imperial College Press, London, UK, pp 277–343 (2001).
216. Phillips J *et al., Phys. Status Solidi B* **194**, 31–39 (1996).
217. Turcu M, Pakma O, Rau U, *Appl. Phys. Lett.* **80**, 2598–2600 (2002).
218. Klenk R, *Thin Solid Films* **387**, 135–140 (2001).
219. Roy M, Damaskinos S, Phillips J, *Proc. 20th IEEE Photovoltaic Specialist Conf.*, pp 1618–1623 (1988).
220. Shafarman W, Phillips J, *Proc. 23rd IEEE Photovoltaic Specialist Conf.*, pp 364–369 (1993).
221. Sah C, Noyce R, Shockley W, *Proc. Inst. Radio Engrs.* **45**, 1228–1243 (1957).
222. Turner G, Schwartz R, Gray J, *Proc. 20th IEEE Photovoltaic Specialist Conf.*, pp 1457–1460 (1988).

223. Schwartz R, Gray J, Lee Y, *Proc. 22nd IEEE Photovoltaic Specialist Conf.*, pp 920–923 (1991).
224. Turcu M, Rau U, *J. Phys, Chem. Solids* **64**, 1591–1595 (2003).
225. Kashiwabara H *et al. Mater. Res. Soc. Symp. Proc.* **1012**,. 89–95 (2007).
226. Scheer R *et al., Appl. Phys. Lett.* **63**, 3294–3296 (1993).
227. Eron M, Rothwarf A, *J. Appl. Phys.* **57**, 2275–2279 (1985).
228. Rudmann D *et al., Appl. Phys. Lett.*, **84**, 1129–1131 (2004).
229. Thompson C, Hegedus S, Shafarman W, Desai D, *Proc. 33rd IEEE Photovoltaic Specialist Conf.* (2008).
230. Erslev P, Halverson A, Shafarman W, Cohen J, *Mater. Res. Soc. Symp. Proc.* **1012**, -Y12-30, (2007).
231. Walter T, Herberholz R, Schock H, *Solid State Phen.* **51**, 301–316 (1996).
232. Rau U, *Appl. Phys. Lett.* **74**, 111–113 (1999).
233. Walter T, Herberholz R, Müller C, Schock H, *J. Appl. Phys.* **80**, 4411–4420 (1996).
234. Herberholz R *et al., Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1246–1249 (1997).
235. Nishitani M, Negami T, Kohara N, Wada T, *J. Appl. Phys.* **82**, 3572–3575 (1997).
236. Metzger W, Repins I, Contreras M, *Appl. Phys. Lett.* **93**, 022110 (2008).
237. Ruberto M, Rothwarf A, *J. Appl. Phys.* **61**, 4662–69 (1987).
238. Rau U *et al., Appl. Phys. Lett.* **73**, 223–5 (1998).
239. Lany S, Zunger A, *J. Appl. Phys.* **100**, 113725 (2006).
240. Lee J, Heath J, Cohen J, Shafarman W, *Mat. Res. Soc. Symp. Proc.* **865**, 373–378 (2005).
241. Igalson M, *Mat. Res. Soc. Symp. Proc.* **1012**, 211–216 (2007).
242. Hegedus S, Shafarman W, *Prog. Photovolt* **12**, 155–76 (2004).
243. Sites J, Mauk P, *Sol. Cells* **27**, 411–417 (1987).
244. Gloeckler M, Sites J, Metzger W, *J. Appl. Phys.* **98**, 113704 (2005).
245. Taretto K, Rau U, *J. Appl. Phys.* **103**, 094523 (2008).
246. Rockett A *et al., Thin Solid Films* **431-32**, 301–06 (2003).
247. Niemegeers A, Burgelman M, De Vos A, *Appl. Phys. Lett.* **67**, 843–845 (1995).
248. Liu X, Sites J, *AIP Conf. Proc.* **353**, 444–453 (1996).
249. Minemoto T *et al., Thin Solid Films* **67**, 83–88 (2001).
250. Wei S, Zunger A, *Appl. Phys. Lett.* **63**, 2549–2551 (1993).
251. Wei S, Zunger A, *J. Appl. Phys.* **78**, 3846–56 (1995).
252. Zhang S, Wei S, Zunger A, *J. Appl. Phys.* **83**, 3192–96 (1998).
253. Schulmeyer T, *et al., Proc. 3rd World Conference on Photovoltaic Energy Conversion*, pp 364–67 (2003).
254. Schulmeyer T *et al., Appl. Phys. Lett.* **84**, 3067–9 (2004).
255. Löher T, Jaegermann W, Pettenkofer C, *J. Appl. Phys.* **77**, 731–38 (1995).
256. Morkel M, *et al., Appl. Phys. Lett.* **79**, 4482–4 (2001).
257. Kniese R *et al.,* in *Wide-Gap Chalcopyrites*, S Siebentritt, and U Rau (eds), Springer, Berlin, Heidelberg, 2006, pp 235–254.
258. Merdes S, *et al., Appl. Phys. Lett.* **95**, 213502 (2009).
259. Marsillac S *et al., Appl. Phys. Lett.* **81**, 1350–1352 (2002).
260. Nakada *et al., Mater. Res. Soc. Symp. Proc.* **865**, 327–334 (2005).
261. Hanket G, Boyle J, Shafarman W, *Proc. 34th IEEE Photovoltaic Specialist Conf* (2009).
262. AbuShama J *et al., Prog. Photovolt.* **12**, 39–45 (2004).
263. Green M, Emery K, Hishikawa Y, Warta W, *Prog. Photovolt.* **17**, 320–326 (2009).
264. Young *et al., Prog. Photovolt.* **11**, 535–541 (2003).
265. Shafarman W *et al. Proc. 29th IEEE Photovoltaic Specialist Conf*, pp 519–522 (2002).
266. Ohashi T, Hashimoto Y, Ito K, *Sol. Energy Mater. Sol. Cells* **67** 225–230 (2001).
267. Hanna G, Jasenek A, Rau U, Schock H, *Thin Solid Films* **387**, 71–73 (2001).

268. Schulmeyer T *et al., Thin Solid Films* **451–52**, 420–423 (2004).
269. Schuler S *et al., Mat. Res. Soc. Symp. Proc.* **668**, H5.14.1 (2001).
270. Abou-Ras D *et al., phys. stat. sol. (RRL)* **2**, 135–137 (2008).
271. Gray J, Lee Y, *Proc. 1st World Conf. Photovoltaic Solar Energy Conversion*, pp 123–126 (1994).
272. Topic M, Smole F, Furlan J, *J. Appl. Phys.* **79**, 8537–8540 (1996).
273. Dullweber T, Hanna G, Rau U, Schock H, *Sol. Energy Mater. Sol. Cells* **67**, 145–150 (2001).
274. Shafarman W, Klenk R, McCandless B, *Proc. 25th IEEE Photovoltaic Specialist Conf.*, pp 763–768 (1996).
275. Dullweber T *et al., Thin Solid Films* **387**, 11–13 (2001).
276. Coutts T, *et al. Progress in Photov*. **11**, 359–375 (2003).
277. Schmid M, Klenk R, Lux-Steiner M, *Sol. Energy Mater. Sol. Cells* **93**, 874–78 (2009).
278. Nishiwaki S, Siebentritt S, Walk P, Lux-Steiner M, *Prog. Photovolt.* **11**, 243–248 (2003).
279. Shafarman W, Paulson P, *Proc. 31st IEEE Photovoltaic Specialist Conf*, pp 231–235 (2005).
280. Nakada T, *et al., Solar Energy* **77**, 739–747 (2004).
281. Abou-Ras D, *et al., Thin Solid Films* **480–481**, 433–438 (2005).
282. Russell T *et al., Proc. 15th IEEE Photovoltaic Specialist Conf.*, pp 743–748 (1982).
283. Probst V *et al., Thin Solid Films* **387**, 262–267 (2001).
284. McCandless B, Shafarman W, *Proc. 3rd World Conf. Photovoltaic Solar Energy Conversion*, pp 562–565 (2003).
285. Malinowska B, Rakib M, Durand G, *Prog. Photovolt.* **10**, 215 (2002).
286. Kessler J, Wennerberg J, Bodegård M, Stolt L, *Sol. Energy Mater. Sol. Cells* **67**, 59–65 (2001).
287. del Cueto J, *et al., Proc. 33rd IEEE Photovoltaic Specialist Conf.* (2008).
288. Wennerberg J, Kessler J, Stolt L, *Sol. Energy Mater. Sol. Cells* **75**, 47–55 (2003).
289. Kushiya K, *et al., Proc. 4th World Conf. Photovoltaic Solar Energy Conversion*, pp 348–351 (2006).
290. Hegedus S, *Prog. Photovolt.* **14**, 393–411 (2006).
291. Ito M *et al., Prog. Photovolt.* **16**, 17–30 (2008).
292. Andersson B, Azar C, Holmberg J, Karlsson S, *Energy* **23**, 407–411 (1998).
293. Fthenakis V, *Ren. Sust. Energy Reviews* **13**, 2746–2750 (2009).
294. Todorov T, Reuter K, Mitzi D, *Adv. Mater.* **22**, (2010).
295. Thumm W *et al., Proc. 1st World Conf. Photovoltaic Solar Energy Conversion*, pp 262–265 (1994).
296. Fthenakis V, Moskowitz P, *Prog. Photovolt.* **3**, 295–306 (1995).
297. Steinberger H, *Prog. Photovolt.* **6**, 99–103 (1998).
298. Moskowitz P, Fthenakis V, *Sol. Cells* **29**, 63–71 (1990).
299. Eberspacher C, Fthenakis V, *Proc. 26th IEEE Photovoltaic Specialist Conf.*, pp 1067–1072 (1997).
300. Pagliaro M, Ciriminna R, Palmisano G, *Prog. Photovolt.* **18**, 61–72 (2010).

14

Cadmium Telluride Solar Cells

Brian E. McCandless[1] and James R. Sites[2]

[1] *University of Delaware, Newark, Delaware, USA*
[2] *Colorado State University, Fort Collins, Colorado, USA*

14.1 INTRODUCTION

Thin film cadmium telluride (CdTe) solar cells are the basis of a rapidly expanding technology with major commercial impact on solar energy production. As the leading technology for US thin film module shipments for 2006–2010, CdTe thin film modules have demonstrated long-term stability, competitive performance, and continue to attract production-scale capital investments. This chapter reviews the status of CdTe thin film solar cells, with emphasis on the properties that make CdTe a favorable material for terrestrial photovoltaic solar energy conversion, the historical development of CdTe solar cells, methods for device fabrication, analysis of device operation, fabrication strategies, and the fundamental technical challenges associated with present and future development of thin film CdTe cells and modules.

Calculations of the dependence of ideal solar cell conversion efficiency on energy bandgap (E_G) show that CdTe is an excellent match to our sun, a G2 spectral-class star with an effective black-body photosphere surface temperature of 5700 K, and a total luminosity of 3.9×10^{33} erg/s. CdTe is a group II^B-VI^A compound semiconductor with a direct optical bandgap that is nearly optimally matched to the solar spectrum for photovoltaic energy conversion. The direct bandgap, $E_G \approx 1.5$ eV, and high CdTe absorption coefficient, $>5 \times 10^5$/cm, for photons with $E > E_G$, means that high quantum yield can be expected from the ultraviolet to the CdTe bandgap wavelength, $\lambda \approx 825$ nm. The high CdTe absorption coefficient for photons with $E > E_G$ translates into 99% absorption of AM1.5 photons with energy above the bandgap within 2 μm of the CdTe surface. The theoretical solar cell efficiency versus bandgap and the optical absorption coefficient versus energy for CdTe and other selected photovoltaic materials are compared in Figure 14.1 [1, 2].

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

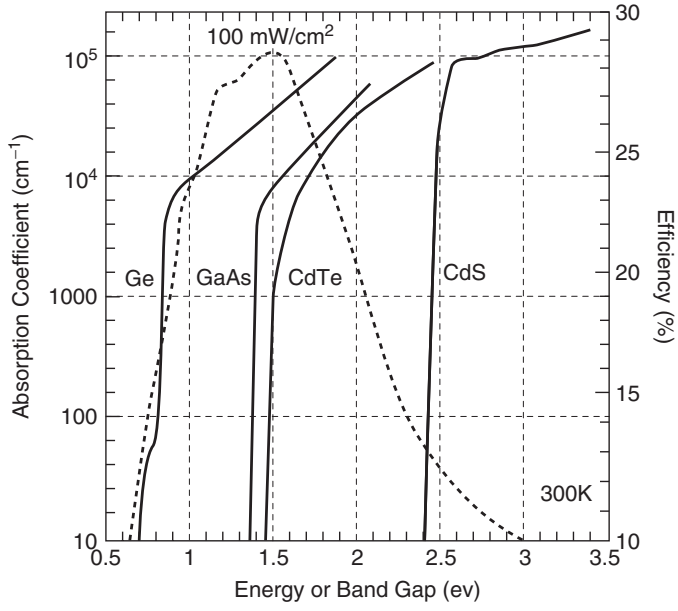


Figure 14.1 Theoretical solar cell efficiency (dotted) for AM1.5 spectral irradiance versus bandgap and absorption coefficient (solid) versus energy. Common absorber materials are highlighted

14.2 HISTORICAL DEVELOPMENT

CdTe was first chemically synthesized by the French chemist Margottet in 1879 [3], but emerged as a viable electronic material in 1947 when Frerichs synthesized CdTe crystals by the reaction of Cd and Te vapors in a hydrogen atmosphere and measured the photoconductivity of the crystal [4]. The early foundation for understanding the electronic nature of CdTe emerged from subsequent studies of single crystals purified by zone refinement. In 1954, Jenny and Bube [5] first reported that p-type and n-type conductivity could be obtained in CdTe by doping with foreign impurities. Shortly thereafter, Krüger and de Nobel [6] showed that the conductivity type could also be controlled by varying the Cd–Te stoichiometry. Cd excess yields n-type and Te excess yields p-type conductivity, analogous to PbS, PbSe, and PbTe. In 1959, the $p-T-x$ diagram of the Cd–Te system and its relationship to intrinsic conduction and extrinsic conductivity via foreign-atom incorporation was established by de Nobel [7], who proposed the existence of two electronic levels associated with Cd vacancies and one with interstitial Cd to account for the measured changes in conductivity at different temperature and Cd partial pressures. Furthermore, the electronic levels associated with In as an n-type dopant and Au as a p-type dopant were estimated.

Loferski at RCA first proposed using CdTe for photovoltaic solar energy conversion in 1956 [1]. Although methods for controlling n and p-type conductivity in CdTe crystals were established by 1960, limited research was directed at the development of p/n homojunctions. In 1959, Rappaport, also at RCA, demonstrated single-crystal homojunction CdTe cells with conversion efficiency ~2% fabricated by diffusion of In into p-type CdTe crystals, yielding $V_{OC} = 600\,\text{mV}$, $J_{SC} \approx 4.5\,\text{mA/cm}^2$ ($73\,\text{mW/cm}^2$ irradiance), and fill factor (FF) = 55% [8]. In 1979, the CNRS group in France achieved >7% conversion efficiency for a device made by close-space vapor transport deposition (VTD) of p-type arsenic-doped CdTe films onto n-type crystals, with $V_{OC} = 723\,\text{mV}$, $J_{SC} \approx 12\,\text{mA/cm}^2$ (AM1 irradiance) and $FF = 63\%$ [9]. Later they reported cells with

efficiency >10.5%, with $V_{OC} = 820\,mV$, $J_{SC} = 21\,mA/cm^2$, and $FF = 62\%$ [10]. Little subsequent work on p/n CdTe homojunctions has been reported.

In contrast to p/n homojunction development, CdTe heterojunction solar cells have been widely investigated since 1960, proceeding along two paths, according to CdTe conductivity type. For n-type CdTe single crystals and polycrystalline films, extensive work was carried out on heterojunctions with p-type Cu_2Te. In the early 1960s, n-type CdTe/p-type Cu_2Te devices having a structure analogous to the CdS/Cu_2S solar cell [11] were fabricated by surface reaction of n-type single crystals or polycrystalline films in acidic aqueous solutions containing Cu salts for topotaxial conversion of CdTe to p-type Cu_2Te [12–16]. By the early 1970s, the best thin film CdTe/Cu_2Te cells achieved cell efficiencies >7%, with $V_{OC} = 550\,mV$, $J_{SC} \approx 16\,mA/cm^2$ ($60\,mW/cm^2$ irradiance), and $FF = 50\%$, as reported by Justi *et al*. [16]. Interestingly, these cells utilized an underlying 5-μm-thick n-type CdS layer to improve adhesion and electrical contact of the 20-μm-thick CdTe film on molybdenum substrates. Difficulty in controlling the Cu_2Te formation process, poor device stability in CdTe/Cu_2Te cells, and lack of a transparent p-type conductor ultimately shifted research emphasis to heterojunction structures employing p-type CdTe. Other work with n-type CdTe utilized Schottky barrier devices, formed by heating Pt or Au grids in contact with n-type CdTe single crystals [17] or electrodeposited CdTe thin films, with efficiencies approaching 9% [18].

For solar cells with single-crystal p-type CdTe, heterojunctions using stable oxides, such as In_2O_3:Sn (ITO), ZnO, SnO_2, and CdO have been more widely investigated. In these devices, the short-wavelength spectral response is influenced primarily by the transmission of the heteropartner and low-resistance contact, collectively referred to as the window layer. Solar cells based on p-type CdTe single crystals with electron-beam-evaporated indium–tin oxide (ITO) window layers with efficiencies = 10.5% were developed by the Stanford group, 1977, with $V_{OC} = 810\,mV$, $J_{SC} \approx 20\,mA/cm^2$, and $FF = 65\%$ [19]. In 1987, cells made by the reactive deposition of indium oxide, In_2O_3, on p-type CdTe single crystals yielded total area efficiencies = 13.4%, with $V_{OC} = 892\,mV$, $J_{SC} = 20.1\,mA/cm^2$, and $FF = 74.5\%$ [20]. In this device, the CdTe crystal had a hole concentration of $6 \times 10^{15}/cm^3$ and the CdTe (111) face was etched in bromine methanol prior to loading into vacuum for In_2O_3 deposition. The V_{OC} of this cell remains the highest ever reported for a CdTe photovoltaic device. Solar cells with ZnO window layers on p-type CdTe single crystals yielded poorer junction behavior, with efficiency <9% and $V_{OC} = 540\,mV$ [21].

Cells made by evaporating n-type CdS films onto single-crystal p-type CdTe were first prepared by Muller *et al*. in the mid-1960s [22, 23], yielding conversion efficiencies less than 5%. In 1977, Mitchell *et al*. reported a conversion efficiency of 7.9% with $V_{OC} = 630\,mV$ for a cell with 1-μm-thick CdS and an ITO transparent electrode [24]. The highest efficiency for a cell fabricated with thin film CdS on p-type CdTe single crystal was reported by Yamaguchi *et al*. in 1977. Their cell utilized 0.5-μm-thick CdS deposited by chemical vapor deposition onto the (111) face of phosphorous-doped CdTe single crystals and gave 11.7% efficiency with $V_{OC} = 670\,mV$ [25].

Thin film CdTe/CdS heterojunction solar cells have been fabricated in two different configurations, referred to as *substrate* and *superstrate*. In both configurations, light enters the cell through the transparent conducting oxide (TCO) and CdS films. However, in the superstrate cell, the TCO, CdS, and CdTe layers are sequentially deposited onto a glass superstrate, which also serves as the mechanical support for the cell, and light must pass through the supporting glass before reaching the CdS/CdTe junction. In the substrate configuration, the CdTe film is deposited first onto a suitable substrate, followed by sequential deposition of CdS and the TCO. Novel schemes have also been demonstrated for fabricating *substrate* configuration cells by transferring the entire cell from a disposable superstrate to a substrate (see, for example, reference [26]).

Superstrate polycrystalline CdTe/CdS heterojunction thin film solar cells were first demonstrated in 1969 by Adirovich *et al*. with evaporated CdTe on a CdS/SnO_2/glass superstrate,

yielding an efficiency >2% [27]. This was followed in 1972 by Bonnet and Rabenhorst, who, in their paper for the 9th European Photovoltaic Specialists Conference, described a 5–6% efficient substrate design CdS/CdTe/Mo made by chemical vapor-deposited CdTe and vacuum-evaporated CdS films [28]. This paper delineated the fundamental issues that still influence the development of highly efficient CdTe/CdS thin film solar cells: (1) the role of Cu in *p*-type doping of CdTe; (2) the controlling role of doping efficiency in CdTe; (3) the effects of abrupt versus graded CdTe–CdS junctions; (4) the effects of active versus passive grain boundaries; and (5) the formation of low-resistance contacts to *p*-type CdTe.

Development of thin film CdTe/CdS solar cell fabrication processes during the 1980s and 1990s, almost always in the superstrate configuration, was advanced by refinements in device design, post-deposition treatments, and formation of low-resistance contacts rather than by refinements in specific deposition methods. This is primarily due to the relatively high chemical stability of CdTe compared with the elemental and compound precursors used to prepare it. Thus, numerous film-fabrication techniques have been used to deposit CdTe for moderate- to high-efficiency solar cells, and eight of these are reviewed in this chapter. The photovoltaic behavior of CdTe/CdS solar cells having conversion efficiency from ~10 to ~16% has been remarkably independent of the CdTe deposition technique.

In spite of tolerance to the deposition technique, two enigmatic aspects of processing high-efficiency thin film CdTe/CdS solar cells persist, that is, the use of superstrate device configuration, with CdTe deposited onto CdS, and the need for processing step(s) that expose the CdTe and CdS films to Cl and O. During the 1980s, significant gains in performance were obtained by empirical optimization of superstrate fabrication processes with respect to processing variables such as the CdTe deposition temperature, post-deposition heat treatment, growth or treatment chemical environment, and CdTe contact formation. For example, the Matsushita Battery Industrial Company reported that for screen-print/sintered CdTe cells, it was critical to control the $CdCl_2$, O, and Cu concentrations in the structure by adjusting the slurries and the temperature–time sequences of the sintering step [29]. The Monosolar electrodeposition process was optimized to the 10%

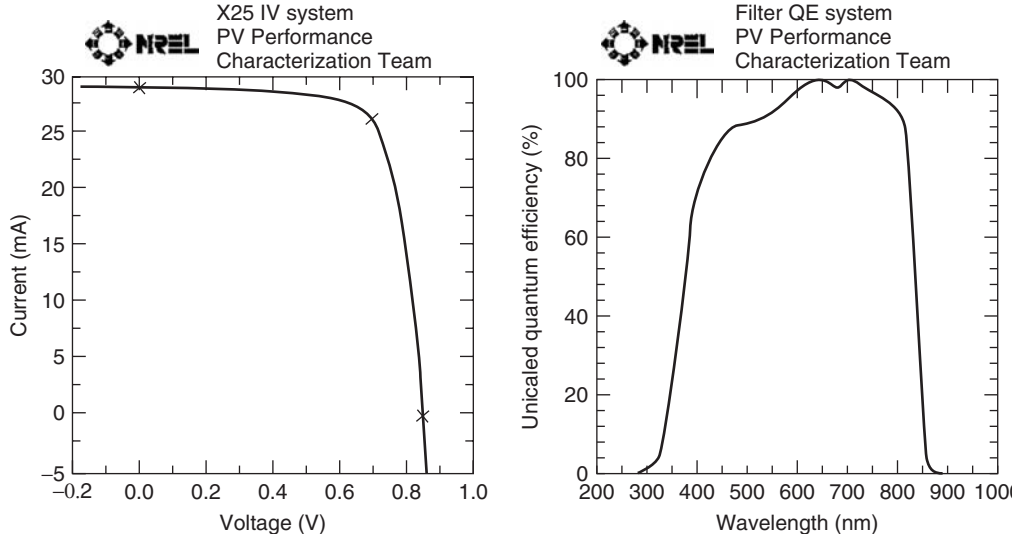


Figure 14.2 Current–voltage and relative quantum efficiency curves for 16.4%-efficient CdTe/CdS thin film solar cell [37]

efficiency level by addition of Cl to the CdTe plating bath and the use of a "type-conversion junction formation" post-deposition treatment to electrically activate the cell [30]. The group at Kodak achieved the 10% efficiency level with close-space sublimation-deposited CdTe by optimizing the CdTe deposition temperature and the oxygen content in the deposition ambient [31].

A turning point for thin film CdTe cell performance, with a collateral benefit for processing tolerance, was the application of a post-deposition air-heat treatment of CdTe/CdS structures coated with $CdCl_2$, by groups at Ametek and at the Institute of Energy Conversion [32, 33]. Combining the "$CdCl_2$ treatment" with advancements in low-resistance contact formation led to the achievement by the group at the University of South Florida in 1993 of a >15%-efficient cell with CdTe deposited by close-space sublimation [34]. Refinements in window-layer processing [35] and employing vapor $CdCl_2$ treatments [36] have led to additional improvements. The record efficiency to date remains at 16.5%, by the group at the National Renewable Energy Laboratory, with $V_{OC} = 845\,\text{mV}$, $J_{SC} = 25.9\,\text{mA/cm}^2$, and $FF = 75.5\%$ [37]. The $J-V$ and quantum efficiency (QE) characteristics of this cell are shown in Figure 14.2.

Since 2000, the development of superstrate polycrystalline CdTe/CdS solar cells has translated to a source of highly successful PV power. Commercial CdTe modules, as discussed in Section 14.6 below accounted for approximately half of all PV product produced in the United States in 2008.

14.3 CdTe PROPERTIES

This section summarizes the fundamental properties of CdTe and describes methods for depositing polycrystalline CdTe thin films. CdTe is unique among the II^B-VI^A compounds, such as ZnS, CdSe, and HgTe, in that it exhibits the highest average atomic number, least negative formation enthalpy, lowest melting temperature, largest lattice parameter, and highest ionicity. Electronically, CdTe exhibits amphoteric semiconducting behavior, making it possible both intrinsically and extrinsically to dope CdTe n and p-type. All these factors complement its nearly ideal optical bandgap and absorption coefficient for terrestrial photovoltaic devices, making it a forgiving material to deposit and control in thin film form. Table 14.1 presents pertinent physical and optoelectronic data for CdTe.

The synthesis of II^B-VI^A compounds is facilitated by the large negative formation enthalpies (ΔH_f) and correspondingly low vapor pressures (p_{sat}) of the compounds compared with their constituent elements: for CdTe, $\Delta H_f = -22.4\,\text{kcal/mol}$ and $p_{sat}(400\,^\circ\text{C}) = 10^{-5}$ Torr and for CdS, $\Delta H_f = -30\,\text{kcal/mol}$ and $p_{sat}(400\,^\circ\text{C}) = 10^{-7}$ Torr [46]. The equilibrium reaction for CdTe solid and Cd and Te vapors is

$$\text{Cd} + 1/2\text{Te}_2 \Leftrightarrow \text{CdTe}$$

The phase diagram for the Cd-Te system is shown for atmospheric pressure in Figure 14.3a. The individual vapor–solid equilibria for CdTe, CdS, Cd, Te, and $CdCl_2$ are shown in Figure 14.3b over the 100–600 °C temperature range typically employed for fabricating solar cells. Congruent evaporation of CdTe facilitates vapor-deposition techniques, and the comparatively high sublimation pressures for Cd and Te ensure single-phase composition when deposited in vacuum at temperatures above ~300 °C. CdTe is also the stable product of cathodic reduction from solutions containing Cd and Te ions due to the similar reduction potentials for Cd and Te and the low solubility product of CdTe.

The $T-x$ phase equilibrium of the CdTe system at atmospheric pressure is characterized by the Cd ($x = 0$) and Te ($x = 1$) endpoints and by the compound CdTe (Figure 14.3a). Note that the

Table 14.1 CdTe optoelectronic and physiochemical properties

Property	Value or range	Reference
CdTe optical bandgap E_g (300 K)	1.50 eV ± 0.01 eV	Single crystal [38] Polycrystalline film [39]
$CdTe_{0.95}S_{0.05}$ alloy optical bandgap	1.47 eV ± 0.01 eV	Polycrystalline film [159]
Temperature dependence dE_g/dT	−0.4 meV/K	[40]
Electron affinity χ_e	4.28 eV	[41]
Absorption coefficient (600 nm)	$>5 \times 10^5$/cm	[39]
Refractive index (600 nm)	~3	[42]
Static dielectric constant ε (θ)	9.4, 10.0	[41, 43]
High-frequency dielectric constant ε (∞)	7.1	[43]
m_e^*	0.096 m_0	[44]
m_h^*	0.35 m_0	[44]
μ_e	500–1000 cm^2/V s	[44]
μ_h	50–80 cm^2/V s	[44]
Space group	F-43 m	[45]
Lattice parameter a_o (300 K)	6.481 Å	[45]
Cd–Te bond length	2.806 Å	Calculated from a_o
Density	6.2 g/cm^3	Calculated from structure
Heat of fusion ΔH_f° (300 K)	−24 kcal/mol	[46]
Entropy S^o (300 K)	23 cal/deg-mol	[46]
Sublimation reaction	$CdTe \rightarrow Cd + 1/2Te_2$	[46]
Sublimation pressure p_{sat}	$\log(P_s/\text{bar}) = -10\,650/T(\text{K})$ $2.56 \log(T) + 15.80$	[46]
Melting point	1365 K	[44]
Thermal expansion coefficient (300 K)	5.9×10^{-6}/K	[47]

CdTe melt temperature, $T_m = 1092\,^\circ$C, is significantly higher than for either Cd, $T_m = 321\,^\circ$C, or Te, $T_m = 450\,^\circ$C [49]. A detailed examination of the $T-x$ projection around the CdTe stoichiometric composition indicates a very narrow, $\sim 10^{-6}$ at.%, symmetrical existence region at $T < 500\,^\circ$C. At higher temperatures, the existence region widens and is asymmetrical on the Cd-rich side up to 700 °C, becoming Te-rich at higher temperatures [44]. The existence region and intrinsic defect structure are related by the preparation conditions of the bulk material and have been the subject of intensive investigation from the time since de Nobel (see for example [50]). Krüger published a comprehensive review of the defect chemistry in 1977 [51], and recently, theoretical treatments of defect levels in CdTe have extended this basis [52]. A critical topic of study is how the bulk properties transfer to thin film CdTe.

The solid-state properties of CdTe are derived from the ionic character of the CdTe bond. Among the II^B-VI^A compounds, CdTe has the highest value on the Phillips ionicity scale = 0.717, which is below the Phillips threshold value of 0.785 for octahedral coordination [53]. Geometrical considerations show that tetrahedral coordination is favored in ionic binary compounds having cation/anion radius ratio between 0.225 and 0.732, while octahedral coordination is favored for a ratio greater than 0.732 [54]. In CdTe, the cation/anion radius ratio is $r(Cd^{2+})/r(Te^{2-}) = 0.444$, thus favoring tetrahedral coordination.

Tetrahedral atomic coordination, with the four nearest neighbors of the other element and the twelve next-nearest neighbors, leads to diamond structure in monatomic solids and zincblende and wurtzite structures in binary solids. Solid CdTe at atmospheric pressure exists in a face-centered cubic zincblende structure with unit cell dimension of 6.481 Å and CdTe bond length of 2.806 Å.

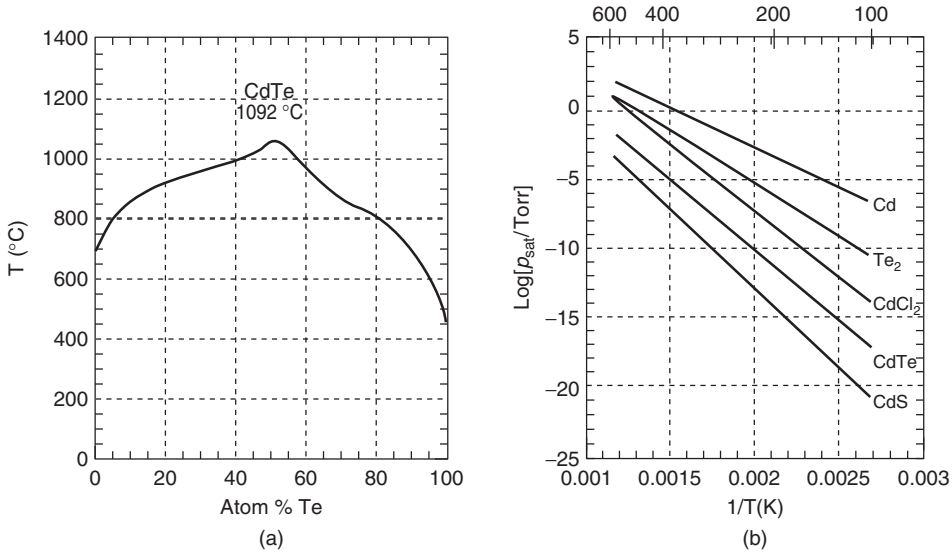


Figure 14.3 (a) CdTe $T-x$ phase diagram (1 atm) [48]; (b) vapor–solid p_{sat} versus $1/T$ diagram for CdTe, CdS, $CdCl_2$, Cd, and Te [46]

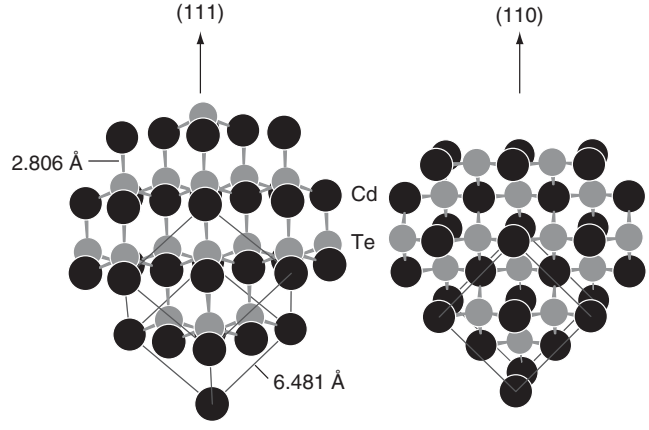


Figure 14.4 Zincblende CdTe crystal structure showing orientation with respect to (111) and (110) planes. The Cd atoms are black and Te atoms are gray. The Cd–Te bonds and fcc unit cells are indicated for each view

Figure 14.4 depicts two views of the CdTe zincblende structure viewed across the closest-packing (111) plane, with alternating anion and cation planes, and viewed across the (110) plane, with equal numbers of anions and cations in each plane. These are the predominant orientations encountered in CdTe thin films.

CdTe polytypes may also be formed, depending on the formation pressure [55]. The hexagonal wurtzite structure, typically associated with tetrahedral coordination in predominantly covalent solids, is occasionally found in CdTe deposited in vacuum. However, no pure wurtzite bulk specimens have ever been reported [56]. Octagonal coordination, leading to the halite NaCl structure, can be induced in CdTe by subjecting single crystals to high pressure, above 35 kbar. Among

II–VI compounds, only CdO, with ionicity = 0.785, occurs with the octahedrally coordinated halite structure at standard pressure and temperature.

The bulk optical and electronic properties of CdTe arise from the electronic band structure near the valence-band maximum (VBM) and the conduction-band minimum (CBM). For zincblende CdTe, the VBM and CBM occur at the same momentum position Γ within the first Brillouin zone, giving rise to a direct bandgap of 1.5 eV at 300 K. The temperature coefficient of the CdTe bandgap is about -0.4 meV/K, which implies minimal variation over normal solar cell operating temepratures. The band curvature about the extrema determines the effective mass of electrons at the CBM and of holes at the VBM and controls carrier-transport properties and interband density-of-states (see Table 14.1).

Qualitatively, the band structure of CdTe can be understood from its relatively high ionicity, since the parts of the Bloch functions having the same periodicity as the lattice are related to the Cd and Te atomic orbitals. The conduction band arises from the first unoccupied level of the cation, namely, the 5*s* level of Cd. The uppermost valence band consists of the highest occupied level of the anion, namely, the 5*p* level of Te. Detailed calculations of the $E-k$ band structures in cubic CdTe and other II–VI compounds were originally carried out by the local pseudopotential method in the mid-1960s [57] and more recently using the linearized augmented plane-wave method, considering all electrons and relativistic kinematics [58]. These calculations have been used to determine the fundamental electronic and thermodynamic properties at interfaces between CdTe and other II–VI compounds [59].

Imperfections or defects in CdTe disrupt the periodic structure, producing localized electronic states within the bandgap, and hence alter the electronic and optical properties. The types of defects controlling electronic properties include native defects, chemical impurities, and complexes thereof; native defects and impurities can occur substitutionally or interstitially. For example, cadmium vacancy, V_{Cd}, gives rise to shallow acceptor states, while cadmium substitution on a tellurium site, Cd_{Te}, gives rise to shallow acceptor states. Interstitial cadmium, Cd_i, gives rise to a relatively shallow donor state, while interstitial tellurium, Te_i, gives rise to deep states. Although shallow states can readily produce acceptors and donors [60], the net bulk doping obtained depends on both formation probability and degree of ionization, which in turn govern the degree to which a desirable acceptor state is compensated.

An on-going critical issue for CdTe solar cell processing has been to obtain acceptor concentrations greater than 10^{14} cm^{-3}. During processing with thermal treatment steps, the semiconductor system is pushed towards equilibrium, and first-principles calculations have shown that a self-compensation mechanism occurs during *p*-type doping [61]. As the Fermi shifts towards the VBM, the chemical potential for donor formation increases, compensating further acceptor formation and essentially pinning the maximum acceptor concentration. Deep states, on the other hand, can act as traps which reduce carrier lifetime, with the effect of increasing carrier recombination. A summary of deep electronic states measured in CdTe and CdZnTe is given in reference [62]. A selected group of shallow and deep defects encompassing native, impurity, and complex levels in CdTe is shown in Figure 14.5. The desired electrical properties are obtained by activation treatments that incorporate specific impurities into the CdTe and CdS layers. These post-deposition treatments may introduce $CdCl_2$, O_2, and Cu into CdTe, which in turn activate or passivate native defects [63]. A comprehensive review of bulk diffusivities of group I, II, and III impurities in CdTe is given in reference [64].

The electronic properties of polycrystalline CdTe used in all present-generation cells, however, can vary from the bulk properties in several important regards. In particular, the grain boundary defect states will have different energies within the bandgap and will have different energetics of formation. It is a reasonable assumption, as will be discussed later, that the various post-deposition strategies to improve cell performance are primarily altering grain-boundary states and have less

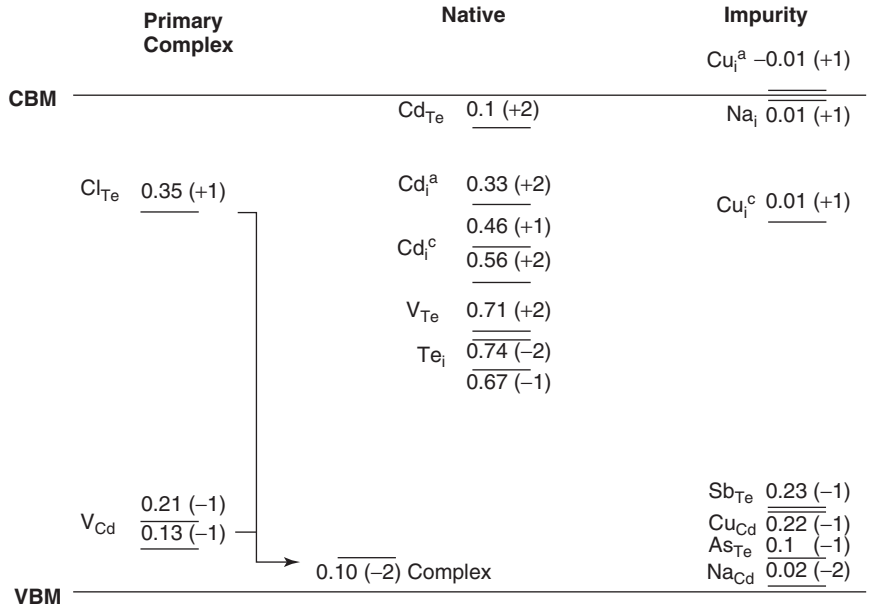


Figure 14.5 CdTe band structure with predicted doping and defect levels. Charge states are in parentheses; energy is in electron volts measured from the conduction band for donor (positive) states and valence band for acceptor (negative) states. The superscripts a and c represent alternative interstitial sites. (Adapted from Wei S, *Mtg. Record, National CdTe R&D Team Meeting* (2001) Appendix 9 [65])

effect on the bulk CdTe states. For example, in one straightforward experiment, devices were fabricated with electrodeposited CdTe films having fibrous grains 2 μm in length and 0.15 μm in width. Although both samples received the same $CdCl_2$ treatment, one sample received a brief oxidation treatment prior to the $CdCl_2$ treatment. The cells produced similar V_{OC}, but different photocurrent [66]. The CdTe film on the cell with the oxidation step retained its as-deposited grain structure and CdS film thickness. The cell exhibited excellent collection, with IQE > 90%, indicating improved lifetime of photogenerated carriers compared with the sample without the oxidation step, which exhibited grain coalescence, CdS film loss, and IQE < 60%.

This result highlights connections between physical, chemical and electronic properties in the CdTe/CdS thin film solar cell. The polycrystalline aspect of cell fabrication raises critical issues affecting development of thin film photovolatics: (1) separating intragrain from grain-boundary effects; (2) identifying the effect of grain boundaries on device operation; and (3) controlling film properties over large area, with $>10^{12}$ grains per square meter in a CdTe module having 1-μm-wide grains. Over the course of the CdTe solar cell development, these issues have been addressed by advancing characterization techniques in parallel with empirical optimization of film deposition and post-deposition treatments. A detailed review of the analytical techniques employed to probe the microstructure, microchemistry, and electronic properties of thin film CdTe/CdS solar cells is beyond the scope of this chapter. However, several powerful methods have emerged to provide a quantitative assessment of film properties and are discussed in references [67, 68], and specifically for CdTe/CdS solar cells in references [69–72]. Some of these methods may also find application as diagnostic sensors for in-line process-control feedback during module manufacture. These are listed

below, accompanied by one or more references in which the techniques are applied to CdTe/CdS thin film solar cells.

Morphology and structure:

Scanning electron microscopy (SEM) [73]
Transmission electron microscopy (TEM) [74]
Atomic force microscopy (AFM) [75]
X-ray diffraction (XRD) [76]

Bulk chemical composition:

Energy dispersive X-ray spectroscopy (EDS) [77]
X-ray diffraction (XRD) [78]
Auger electron spectroscopy (AES) [79]
Secondary ion mass spectroscopy (SIMS) [80, 81]

Surface chemical composition:

X-ray photoemission spectroscopy (XPS) [82]
Glancing incidence X-ray diffraction (GIXRD) [83]

Optoelectronic properties:

Optical absorption [84]
Ellipsometry [85–87]
Raman scattering [88]
Photoluminescence (PL) [89, 90]

Junction analysis:

Current–voltage versus illumination and temperature ($J-V-T$) [91–93]
Spectral response [24]
Capacitance–voltage ($C-V$) [94]
Optical beam induced current (OBIC) [95]
Electron beam induced current (EBIC) [96]
Cathode luminescence (CL) [97]

14.4 CdTe FILM DEPOSITION

Numerous methods have been employed to deposit CdTe thin films for solar cells, as detailed in the special issue of the *International Journal of Solar Energy* [98] and other review articles [99–101]. We will review here eight methods that have demonstrated viability for the commercial manufacture of CdTe solar cells and modules over the past decade. Figure 14.6 presents schematic views of each fabrication procedure, including nominal temperature and pressure conditions, film thickness, and growth rate. The following discussion of the eight methods is organized by three chemical concepts: (1) condensation/reaction of Cd and Te_2 vapors on a surface (PVD, VTD, CSS, and sputter deposition), (2) galvanic reduction of Cd and Te ions at a surface (electrodeposition), (3) reaction of precursors at a surface [metal–organic chemical vapor deposition (MOCVD), screen-print deposition, and spray deposition].

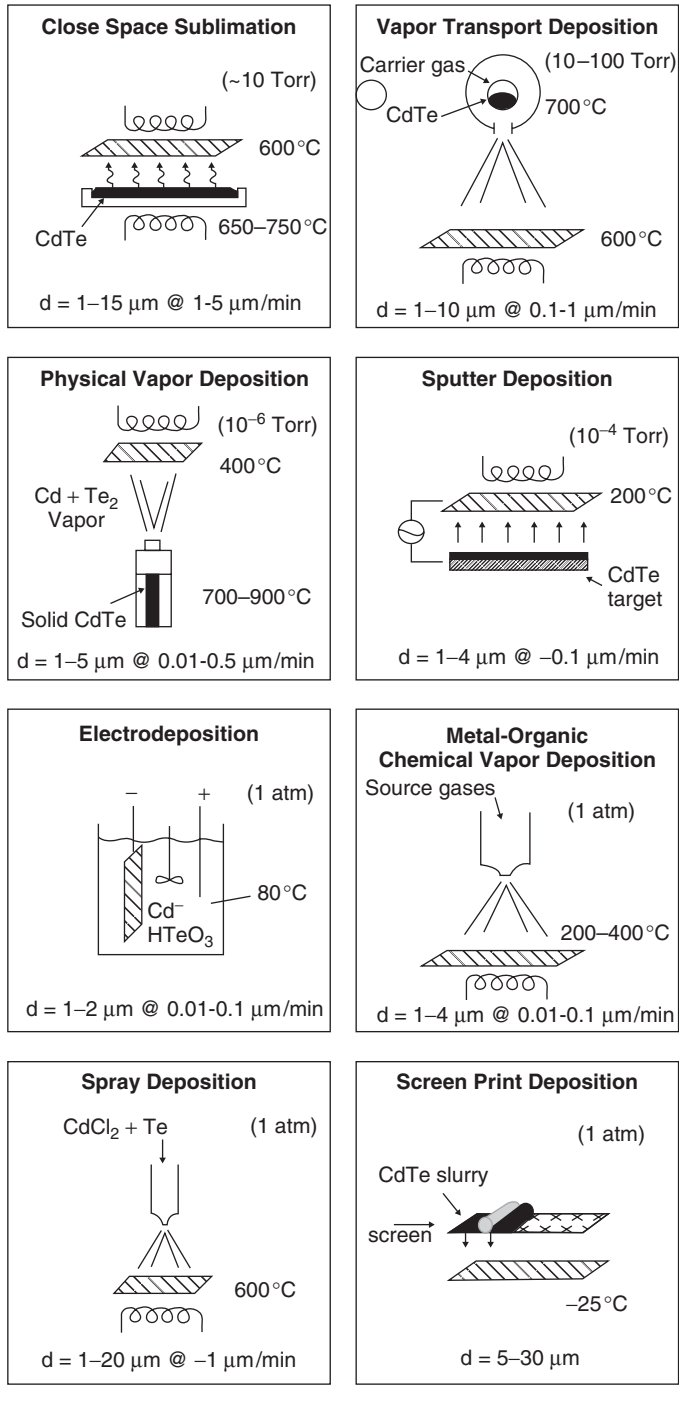


Figure 14.6 Schematic representations of eight CdTe thin film deposition techniques. The substrate in each view is the cross-hatched rectangle. Film thickness d and growth rate are shown at the bottom of each panel

14.4.1 Condensation/Reaction of Cd and Te_2 Vapors on a Surface

14.4.1.1 Physical vapor deposition (PVD)

The basis for vapor deposition of CdTe is the equilibrium between Cd and Te_2 vapors and CdTe solid, $Cd + 1/2Te_2 \Leftrightarrow CdTe$. As a consequence, CdTe can be deposited by coevaporation from elemental sources, by direct sublimation from a CdTe source or by vapor transport using a carrier gas to entrain and deliver Cd and Te_2 vapors from either elemental or CdTe sources. Congruent sublimation of the CdTe compound fixes the gas-phase composition for deposition from a CdTe source, and the relatively low vapor pressure of CdTe compared with elemental Cd and Te facilitates the deposition of single-phase solid films over a wide range of substrate temperatures (see Figure 14.3b). Similar considerations allow coevaporation from multiple II–VI binary sources to deposit alloys in pseudobinary systems such as $Cd_{1-x}Zn_xTe$ and $CdTe_{1-x}S_x$.

Evaporation can be carried out from open crucibles or from Knudsen-type effusion cells, with the latter providing superior control over beam distribution and utilization. For effusion-cell evaporation, the deposition rate and uniformity of the species arriving at the substrate are controlled by source temperature, effusion-cell geometry, source to substrate distance, and total pressure [102, 103]. Within the effusion cell, mass transport to the nozzle exit occurs in a transitional flow regime, between free molecular flow and diffusion-limited flow. Effusion cells are typically constructed of boron nitride or graphite and are radiatively heated. For deposition in moderate vacuum, $\sim 10^{-6}$ Torr, with a CdTe source effusion cell with 0.5-cm-diameter orifice and a temperature of 800 °C, at a source to substrate distance of 20 cm, a commercially viable deposition rate of ~ 1 μm/min is obtained on a substrate at a sufficiently low temperature (~ 100 °C) for Cd and Te sticking coefficients to approach unity. At higher substrate temperatures, the sticking coefficients of impinging Cd and Te decrease, lowering deposition rate and imposing a practical substrate temperature limit of less than 400 °C for modest CdTe utilization. As-deposited films typically exhibit (111) preferred orientation and normal grain-size distribution with a mean grain diameter that depends on film thickness and substrate temperature; for 2-μm-thick films, the mean grain diameter ranges from ~ 100 nm at 100 °C to ~ 1 μm at 350 °C. The physical vapor deposition (PVD) process has been investigated by Stanford University [104], the Institute of Energy Conversion at University of Delaware [105] and industrial (Canrom and Central Research Laboratory at Japan Energy Corporation [106]) groups.

14.4.1.2 Close-space sublimation (CSS)

To evaporate CdTe films onto substrates at temperatures above 400 °C, re-evaporation of Cd and Te from the growing CdTe surface limits the deposition rate and utilization. This can be mitigated by depositing at higher total pressure, ~ 1 Torr, but mass transfer from the source to the substrate becomes diffusion-limited, so the source and substrate must be brought into close proximity. For close-space sublimation (CSS), also known as close-space vapor transport (CSVT), the CdTe source material is supported in a holder having the same area as the substrate; the source holder and substrate cover serve as susceptors for radiative heating and conduct heat to the CdTe source and the substrate, respectively. A solid insulating spacer allows thermal isolation of the source from the substrate, so that a temperature differential can be sustained throughout the duration of the deposition. The ambient for deposition typically contains a nonreactive gas such as N_2, Ar, or He. A small partial pressure of O_2 appears to be crucial for obtaining good film density and solar cell junction quality. As-deposited CSS films deposited above 550 °C exhibit nearly random orientation and normal grain size distribution with mean grain size that is comparable to film thickness. The CSS process can also achieve >1 μm/min deposition rate and has been intensively investigated by

groups at Kodak [107], USF NREL [108, 109], NREL [110], Matsushita [111] and Antec [112, 113]. It has yielded the highest small-area cell performance of any process shown in Figure 14.6 and has been used commercially by Abound Solar (formerly AVA Solar) [114], Primestar Solar [115], and Antec, GmbH.

14.4.1.3 Vapor transport deposition (VTD)

VTD allows high-rate deposition at high substrate temperature at pressures approaching 0.1 atm onto moving substrates. While CSS is diffusion-limited, VTD works by convective transfer of a vapor stream saturated with Cd and Te to the substrate. The substrate is at a lower temperature ($<600\,^{\circ}C$) relative to the source ($>800\,^{\circ}C$). Supersaturation of the Cd and Te vapors results in condensation on the substrate and reaction to form CdTe. The CdTe source consists of a heated chamber containing solid CdTe in which the carrier gas mixes with the Cd and Te vapors and is exhausted through a slit over or under the moving substrate at a distance of the order of $\sim$1 cm. The geometrical configuration of the source influences the uniformity and utilization of the vapors in the carrier gas. The carrier-gas composition can be varied, as with CSS, to include N_2, Ar, He, and O_2. As-deposited VTD films are similar to CSS films, again with nearly random orientation and normal grain size distribution with mean grain size that is comparable to film thickness [116]. The VTD process can provide a very high deposition rate onto moving substrates, as demonstrated by the Institute of Energy Conversion [117], and has led to the extremely successful commercial development led by First Solar, LLC [118].

14.4.1.4 Sputter deposition

CdTe films have also been deposited by radio-frequency magnetron sputtering from compound targets. This deposition method differs from the previous ones in that the overall film formation process may occur quite far from thermal equilibrium. Mass transfer of Cd and Te occurs via ablation of the CdTe target by Ar^+, followed by diffusion to the substrate and condensation. Typically, sputter deposition is carried out at a substrate temperature less than 300 °C and at pressures $\sim$10 mTorr. As-deposited films 2-μm-thick deposited at 200 °C exhibit mean grain diameter $\sim$300 nm and nearly random orientation. The sputter-deposition technique has been investigated by groups at the University of Toledo [119] and at NREL [120].

14.4.2 Galvanic Reduction of Cd and Te Ions at a Surface

14.4.2.1 Electrodeposition

Electrodeposition of CdTe consists of the galvanic reduction of Cd and Te from Cd^{+2} and $HTeO_2^+$ ions in an acidic aqueous electrolyte. The reduction of these ions utilizes six electrons in the following reactions:

$$HTeO_2{}^+ + 3H^+ + 4e^- \rightarrow Te^o + 2H_2O,\ E_o = +0.559\ V$$
$$Cd^{+2} + 2e^- \rightarrow Cd^o,\ E_o = -0.403\ V$$
$$Cd^o + Te^o \rightarrow CdTe$$

The large difference in reduction potential necessitates limiting the concentration of the more positive species, Te, to maintain stoichiometry in the deposit. In practice, the low Te species concentration ($10^{-4}\ M$) limits the CdTe growth rate due to Te depletion in the solution at the growing

surface and subsequent mass transport. To overcome this, the electrolyte is stirred vigorously, and different methods of Te replenishment are employed. Thickness and deposition area are limited by the ability to maintain a constant deposition potential over the entire surface of the growing film. As-deposited films can be fabricated as stoichiometric CdTe, Te-rich (by increasing Te species concentration in the bath) or Cd-rich (by depositing at low potentials with limited Te species concentration). The relation between deposition rate, stoichiometry, bath composition and deposition potential was quantitatively developed in the late 1970s by F. A. Kröger [121]. As-deposited electrodeposited CdTe films on CdS thin film substrates typically exhibit strong (111) orientation with columnar grains having a mean lateral diameter of 100–200 nm. Electrodeposition of CdTe has been intensively studied by the group at Monosolar [122], Ametek [123], and the University of Texas [124]. In the 1980s, the Monosolar process was transferred to SOHIO and later to BP Solar, where commercial development took place at the factory in Fairfield, California, where they produced a record efficiency 10.9%, 0.48 m^2 module. In the early 1990s, the Ametek electrodeposition process was transferred to the Colorado School of Mines in Golden Colorado.

14.4.3 Precursor Reaction at a Surface

14.4.3.1 Metal organic chemical vapor deposition (MOCVD)

MOCVD is a nonvacuum technique for depositing CdTe films at moderately low temperature from organic Cd and Te precursors such as dimethylcadmium and diisopropyltellurium in hydrogen carrier gas. The substrates are supported on graphite susceptors and can be heated radiatively or by coupling to a radio frequency generator. Deposition occurs by pyrolytic decomposition of the source gases and reaction of the Cd and Te species. As a consequence, the growth rate depends strongly on the substrate temperature, which typically ranges from 200 to 400 °C. As-deposited films 2-μm-thick deposited at 400 °C exhibit columnar grain structure with lateral grain diameter ~1 μm. The MOCVD process has been investigated by groups at SMU/USF [125] and Georgia Institute of Technology [126].

14.4.3.2 Spray deposition

Spray deposition is a nonvacuum technique for depositing CdTe from a slurry containing CdTe, $CdCl_2$, and a carrier such as propylene glycol. The slurry can be sprayed onto room-temperature or heated substrates, after which a reaction/recrystallization treatment is performed. The application of spray deposition to CdTe films was developed by the Photon Energy Corporation during the 1980s. The company was sold to Coors in 1995, and its name was changed to Golden Photon. Cells with >14% efficiency and modules with 9% efficiency were fabricated, but commercial development ceased in 1997. In the spray deposition process, the mixture was sprayed onto the substrates at room temperature and pre-baked at 200 °C to evaporate away the carrier medium, followed by a bake in the presence of O_2 at 350–550 °C, a mechanical densification step, and a final treatment at 550 °C. Films produced by this method vary in morphology, grain size, and porosity, but films used to make high-efficiency cells exhibited a 1- to 2-μm-thick dense region near the CdTe–CdS interface, a relatively porous back surface region, and random crystallographic orientation. A dramatic consequence of the fabrication process was the intermixing of the CdS and CdTe layers, leading to a nearly uniform $CdTe_{1-x}S_x$ alloy throughout the absorber, which reduced its bandgap to ~1.4 eV. In the highest-efficiency cells made by spray deposition, CdS diffusion and subsequent alloy formation consumed most of the CdS film, resulting in an enhanced blue-spectral response and correspondingly high short-circuit densities. The method was investigated extensively by the group at Golden Photon [127, 128].

14.4.3.3 Screen-print deposition and sintering

Screen-print and related deposition techniques using nanoparticle or microcystalline precursors are perhaps the simplest of the CdTe techniques. A viscous combination of Cd, Te and $CdCl_2$, or CdTe particles in a suitable binder is applied to the substrate through a screen or by doctor-blading to create a film layer. Following a drying step to remove the binder solvents, the layer is baked at temperatures up to 700 °C to recrystallize the film and activate the junction. Screen-printed films typically have a thickness ranging from 10 to 20 μm with lateral grain dimension of ~5 μm and random crystallographic orientation. Screen-printed CdTe can be traced back to the 1970s with the pioneering work of Matsushita [129], and subsequently by groups at the University of Seoul [130] and the University of Ghent [131]. Recent work using nanoparticle precursors with film thickness less than 10 μm has been carried out at the University of California at Berkely [132].

14.5 CdTe THIN FILM SOLAR CELLS

All high-efficiency CdTe solar cells to date have essentially the same *superstrate* structure, as demonstrated by Bonnet and Rabenhorst in 1972 [28]. This structure, in which the TCO and CdS are first deposited onto a suitably transparent material is depicted in Figure 14.7. The alternative *substrate* configuration in which the CdTe is deposited first onto a suitably conductive substrate, followed by deposition of CdS and TCO, has not yet attained high conversion efficiencies, primarily because of lower CdS/CdTe junction quality and difficulty in maintaining low-resistance electrical contact to the CdTe through cell processing.

The primary CdTe/CdS photodiode junction is designed to occur between the *p*-type CdTe absorber and the *n*-CdS window layer. There are, however, a number of complicating factors, such as the need for a high-resistance oxide layer when the CdS is thin, the need for a thermal treatment with $CdCl_2$ and oxygen to improve the CdTe quality, the interdiffusion of CdS and CdTe, and the secondary barrier associated with the back contact. The following sections address these complications.

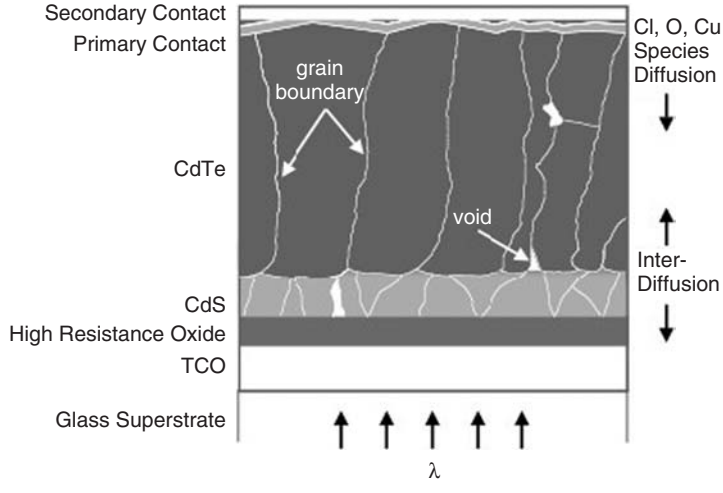


Figure 14.7 Basic CdTe solar cell structure. The polycrystalline nature of the CdS and CdTe layers are indicated schematically and are not to scale

14.5.1 Window Layers

The first step in the fabrication of a superstrate CdTe cell is to coat the glass superstrate with a transparent conducting oxide (TCO), such as SnO_2, indium–tin oxide, In_2O_3:Sn (referred to as ITO), or cadmium stannate, Cd_2SnO_4. The TCO serves as the front contact and lateral current-carrying conductor. To obtain high current density in the completed cell, the CdS layer needs to be sufficiently thin to transmit most of the blue photons. The integrity of ultra-thin CdS layers can be difficult to control during high temperature processing, raising the possibility of direct junctions between CdTe and TCO, producing local shunting or excessive forward current. It has been found that the deposition of a second, highly resistive, transparent oxide layer, referred to as the HRT layer, between the TCO and CdS can significantly ameliorate this problem by improving junction quality and uniformity as also reported for $CuInSe_2$/CdS and a-Si thin film solar cells [133]. Materials used for the resistive layer include undoped SnO_2 [134], Zn-doped SnO_2, In_2O_3 [135, 136], Ga_2O_3 [116], and Zn_2SnO_4 [137]. For cells fabricated on tin oxide-coated soda-lime glass, the HRT layer can also serve as a diffusion barrier protecting the CdS/CdTe junction from contamination by mobile impurities from the glass [138].

Most CdTe cells utilize *n*-type CdS for the window layer immediately adjacent to the CdTe. The processing possibilities for depositing good-quality CdS are nearly as varied as those shown in Figure 14.6 for CdTe, and include chemical bath deposition, sputter deposition, and physical vapor deposition. The choice is usually driven by compatibility with the other deposition processes in a fabrication line. While it is generally desirable to keep the CdS layer as thin as possible to transmit a high fraction of the short-wavelength photons ($<$520 nm) and thereby increase the short wavelength QE and J_{SC}, ultra-thin CdS can also lead to a decrease in V_{OC}. Thin-film CdTe cells with the CdS layer omitted altogether have not performed well as of this writing (see, for example, Table 2 in reference [133]), and it is a reasonable assumption that cells with ultra-thin CdS have regions where CdTe is in direct contact with the TCO.

In practice, as will be discussed in more detail later, cell-processing conditions promote inter-diffusion between CdTe and CdS in response to the thermodynamic driving force for alloy formation on each side of the interface. The resulting bandgap shift in CdS reduces the window-layer transmission and lowers the short-wavelength quantum efficiency [139, 140]. The CdS interdiffusion can be minimized either by heat treatment of the CdS with $CdCl_2$ to recrystallize the film or by judicious control of device processing to reduce the remaining CdS thickness [137] effectively to zero [128].

Another strategy to reduce window absorption has been to mix CdS with ZnS to increase the bandgap of the layer, and hence the photon transmission, but simple mixing has not produced net performance gains, and ZnS is chemically less stable than CdS during $CdCl_2$ treatment. The highest-efficiency CdTe cells to date have used bilayer structures consisting of a Cd_2SnO_4 TCO and Zn_2SnO_4 HRT layer between the CdS and the glass, which utilizes their wide optical bandgaps and inherent conductive properties. An additional dimension of this strategy is that the Zn_2SnO_4 HRT layer may contribute to CdS consumption during processing [141].

14.5.2 CdTe Absorber Layer and $CdCl_2$ Treatment

A number of successful techniques for depositing device-quality CdTe films were presented schematically in Figure 14.6. However, in cases of chlorine-free deposition techniques such as PVD, CSS, VT and sputter deposition, the actual deposition of the CdTe film has been found to be less critical than the post-deposition processing, which generally involves exposure to a moderately high-temperature treatment in conjunction with exposure to a chlorine-containing species such as $CdCl_2$, often referred to as the "$CdCl_2$ treatment." The most critical constituent of the treatment step is the chlorine, which can be delivered in a variety of ways, such as dipping the CdTe layer

in a $CdCl_2$:CH_3OH or $CdCl_2$:H_2O solution, followed by drying to precipitate a $CdCl_2$ film [142, 143], evaporating a layer of $CdCl_2$ onto the CdTe [144], exposure to $CdCl_2$ vapor [145–147] or $ZnCl_2$ vapor [148] or exposure to HCl [149] or Cl_2 gas [150]. Chlorine species may also be incorporated during CdTe film formation, in the form of Cl^- ions in an electrodeposition bath [151] or as a component of a screen-printing slurry [73]. The typical temperature–time range for post-deposition processing is from 380 to 450 °C for 15–30 min, depending on CdTe film thickness, with thinner films requiring shorter treatment time and reduced $CdCl_2$ concentration, as discussed in Section 14.4.3. Different vapor methods for controlling $CdCl_2$ concentration are presented in reference [152].

The use of chlorine and subsequent heat treatment can alter the CdTe cell in a variety of ways. For example, the treatment can promote recrystallization and grain growth in films having sub-micrometer initial CdTe crystallite size [153]. Table 14.2 more generally compares the changes in grain size, aspect ratio, and crystallite orientation for CdTe deposited by several different methods.

For CdTe films having sub-micrometer initial crystallite size, significant recrystallization occurs during the $CdCl_2$ treatment. This takes two forms: (1) intragrain, or primary, recrystallization changes grain orientation from typically (111) to random (see Figure 14.8); and (2) intergrain, or secondary, recrystallization, resulting in grain coalescence. Films deposited at high temperature or heat treated at high temperature in the presence of oxygen prior to $CdCl_2$ exposure exhibit undetectable grain growth (secondary recrystallization) after $CdCl_2$ treatment. Under these processing conditions, film randomization is the predominant effect, showing that $CdCl_2$ still exerts an intragrain influence on lattice arrangement. The effects on crystallographic properties and the close relation between CdTe defect chemistry and electronic properties suggests that the $CdCl_2$ treatment can affect defect levels which affect lifetime and doping.

Post-deposition heat treatment with $CdCl_2$ modifies both the CdTe film and the CdS/CdTe interface region through enhanced chemical reactivity and species mobility. Structural analysis reveals that CdTe film orientation is randomized by the treatment, accompanied by a decrease in lattice parameter, with the strongest effect located at the junction interface. Similarly, the CdS film has been found to be randomized with an increase in lattice parameter. The $CdCl_2$ behaves as a fluxing agent, breaking Cd-Te and Cd-S bonds, facilitating primary recrystallization and anion interdiffusion at significantly lower temperatures than required without $CdCl_2$. Chemically, the $CdCl_2$ treatment can modify the Cd/Te stoichiometry, forming cadmium vacancies (likely involving

Table 14.2 Structural changes due to $CdCl_2$ on CdTe deposited by different methods (data from films examined by McCandless). For sprayed and screen-printed cells, random film orientation is obtained as a result of the film formation process

Deposition method	Film thickness [μm]	Mean grain size: D Init → After $CdCl_2$ [μm]	Orientation Init → After $CdCl_2$
PVD	4	0.1 → 1	(111) → (220)
ED	2	0.1 → 0.3	(111) → (110)
Spray	10	10 → 10	Random → random
Screen	12	∼ 10	Random → random
VTD	4	4 → 4	Random → random
CSS	8	8 → 8	Random → random
Sputter	2	0.3 → 0.5	(111) → (?)
MOCVD	2	0.2 → 1	(111) → random

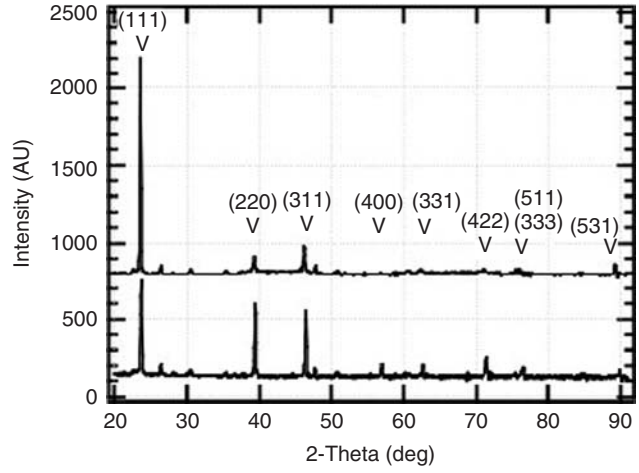


Figure 14.8 X-ray diffraction patterns of PVD CdTe/CdS thin film structures before (top) and after (bottom) vapor $CdCl_2$ heat treatment at 420 °C for 20 min, showing randomization of the film orientation. The primary peaks are due to the 1.5-μm-thick CdTe layer; the other peaks are due to the underlying CdS and ITO films

oxygen) and chlorine interstitial species. The $CdCl_2$ treatment reduces the CdTe cross-grain sheet resistance by up to three orders of magnitude [154], likely attributed to the $V_{Cd}-Cl_i$ acceptor complex, reduced donor defect density, and improved carrier mobility or lifetime. Oxygen in the ambient may aid the formation of V_{Cd} by preferential reaction at the surface to form CdO, reducing the local Cd concentration. A defect chemistry model for the doping process has been suggested by Marfaing [155] and a doping model based on electronic band offsets by Zhang, *et al.* [156]. The combination of Cl and O may thus adjust the relative cation and anion concentrations at electronic levels. Both the single-donor and the double-acceptor states are pushed closer to the band edges, resulting in a single, relatively shallow acceptor state. Although this complex is a more effective dopant than the Cd vacancies alone, excess Cl could eventually lead to compensating Cl_{Te} donors.

The impact of $CdCl_2$ treatment on cell operation is markedly increased photocurrent and uniformity, with improvement in open-circuit voltage and fill factor. Figure 14.9 compares the light $J-V$ behavior of three PVD cells having 4-μm-thick CdTe and 0.2-μm-thick CdS, processed with the same back contact but with: (1) no post-deposition treatment; (2) air-heat treatment at 550 °C; and (3) optimized $CdCl_2$ treatment at 420 °C for 20 min in air. With no treatment, the device exhibits very low photocurrent and high series resistance. The spectral response is low overall and exhibits a peak near the CdTe band edge, suggesting low carrier lifetime and *p-i-n* device operation [157]. With either air treatment or vapor $CdCl_2$ + air treatment, the $J-V$ and spectral response behavior correspond to device operation with higher lifetime and carrier density. For devices made by other methods, similar behavior is obtained, but the starting condition (Figure 14.9a) is generally improved by deposition at high temperature in oxygen-containing ambient.

The effect of the $CdCl_2$ treatment on photocurrent micro-uniformity is shown in Figure 14.10 for cells with CdTe deposited by CSS. The quantum efficiency (QE) map on the left was made on a cell following a typical $CdCl_2$ treatment, and shows spatially uniform collection. The map on the right was for a cell fabricated without the $CdCl_2$ treatment and exhibits considerable nonuniformity [158]. The large local reductions in photocurrent without $CdCl_2$ treatment occur in areas associated with high grain-boundary density. The QE of the sample with $CdCl_2$ treatment is ~0.82

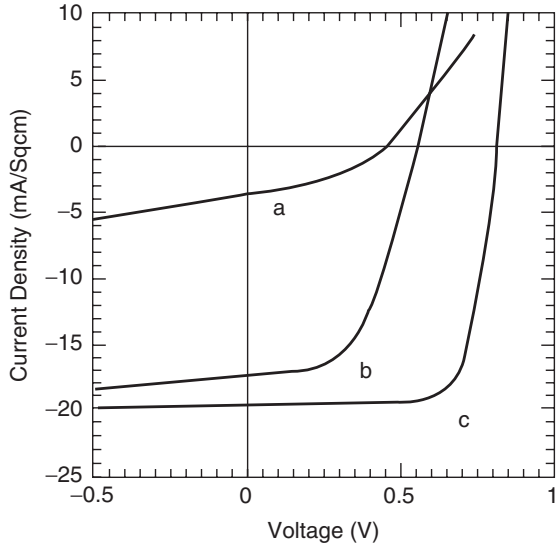


Figure 14.9 Light AM1.5 $J-V$ curves of PVD devices with Cu_2Te/C contacts and different postdeposition processing: (a) no heat treatment; (b) treatment at 550 °C for 5 min in air; (c) treatment in $CdCl_2$ vapor at 420 °C for 20 min in air [116]

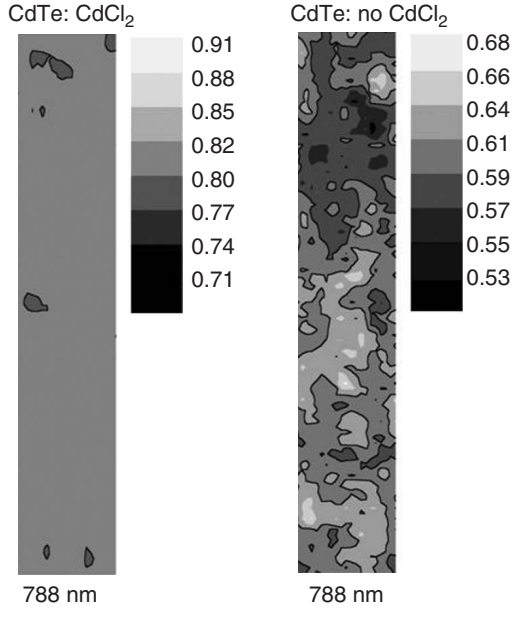


Figure 14.10 Local variations in quantum efficiency with 1-μm beam and $\lambda = 788$ nm. Areas shown are $50 \times 10\,\mu m$

over 95% of the measured area, while that of the sample without $CdCl_2$ treatment ranges from 0.50 to 0.68.

14.5.3 CdS/CdTe Intermixing

All CdS/CdTe cells are exposed to processing temperatures of at least 350 °C during $CdCl_2$ treatment. With some deposition techniques, such as spray pyrolysis, much higher temperatures are involved. Hence, a chemical reaction between CdTe and CdS can occur, and this reaction may be the driving force for bulk and grain-boundary interdiffusion of CdTe and CdS. It has been widely reported that a continuous series of CdTe–CdS solid-solution alloys, $CdTe_{1-x}S_x$, can be formed by co-deposition of CdTe and CdS at temperatures less than 200 °C or aqueous electrodeposition. The optical bandgap of these alloys varies with composition x, according to $E_g(x) = 2.40x + 1.51(1-x) - bx(1-x)$, with bowing parameter, $b \approx 1.8$, as shown in Figure 14.11 [159]. Thermal treatment of alloy films above 350 °C with $CdCl_2$ and above 500 °C without $CdCl_2$ induces phase segregation for the mid-range alloys.

For equilibrated CdTe–CdS mixed crystals at temperatures above 625 °C, numerous references have established the $T-x$ phase relations. The temperature range of the miscibility gap has been extended down to 360 °C, i.e. to the temperature range used for depositing and processing CdTe films, by lattice-parameter determination of equilibrated $CdTe_{1-x}S_x$ alloy films, as shown in Figure 14.12 [146, 166]. Thermodynamic analysis of the asymmetric phase boundaries using non-ideal solution thermodynamics reveals positive values of excess-mixing enthalpies $\Delta H^{EX} = 3.5$ kcal/mol for CdS into CdTe and $\Delta H^{EX} = 5.6$ kcal/mol for CdTe into CdS. For CdS dissolved in CdTe, the experimentally obtained excess-mixing enthalpy supports those calculated from first-principles band structure theory for the CdTe–CdS system [167].

The crystallographic forms of the solid alloys are generally zincblende (F-43 m) for $CdTe_{1-x}S_x$ and wurtzite ($P6_3mc$) for $CdS_{1-y}Te_y$. The zincblende to wurtzite transition in

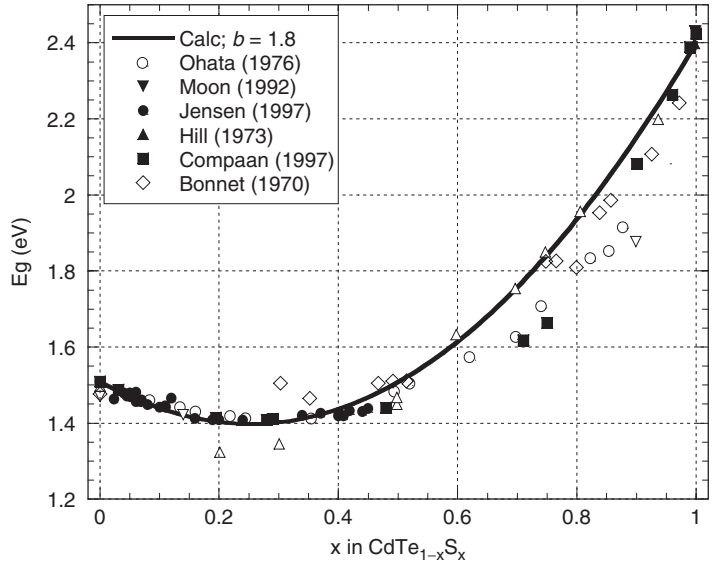


Figure 14.11 Optical bandgap of low-temperature deposited $CdTe_{1-x}S_x$ alloy thin films versus composition. (Data listed in order from References [160–165])

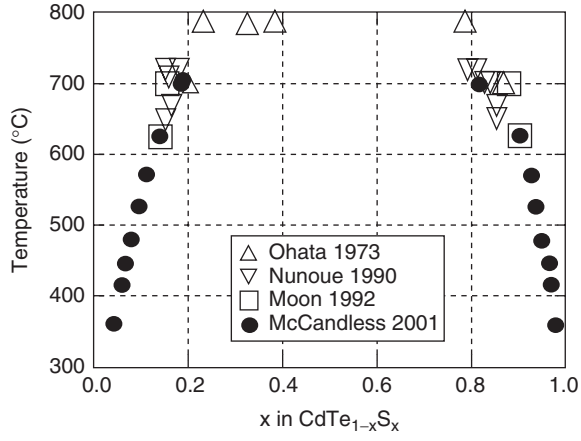


Figure 14.12 CdTe–CdS pseudobinary equilibrium phase diagram. (Data listed in order from References [160, 78, 161, 170])

metastable, low-temperature deposited films occurs at $x \approx 0.3$, and the lattice parameter within each structure type follows Vegard's rule. Metastable and equilibrated $CdTe_{1-x}S_x$ alloy films exhibit the same dependence of E_g, with minimum at 1.39 eV, corresponding to the zincblende–wurtzite transition. The $CdCl_2$ treatment accelerates the equilibration process, but does not change the endpoint compositions.

Formation of $CdTe_{1-x}S_x$ and $CdS_{1-y}Te_y$ alloys at the interface region occurs by interdiffusion between CdS and CdTe during $CdCl_2$ treatment in air and is enhanced by grain-boundary diffusion [153, 169]. The diffusion mechanism is limited by Cd self-diffusion. Diffusion of CdTe into CdS can reduce the transmissive properties of the window between 500 and 650 nm by reducing the bandgap. However, diffusion of CdS into CdTe is a faster process and is more difficult to control, especially for cell structures with ultra-thin, <100 nm, CdS films. The progressive diffusion, which creates a lattice parameter and optical constant distribution within the film, can be detected by X-ray diffraction line-profile analysis (Figure 14.13) [170] and by ellipsometry [171].

The bulk and grain-boundary diffusion coefficients are both thermally activated and are sensitive to ambient composition during the treatment. The thermal dependence of bulk and grain-boundary diffusion coefficients for bulk diffusion (D_B) and grain-boundary diffusion process for CdS in CdTe (D_{GB}) are [170]:

$$D_B = 2.4x10^7 e^{-\left(\frac{2.8eV}{kT}\right)} \text{ cm}^2/\text{s}$$
$$D_{GB} = 3.4x10^6 e^{-\left(\frac{2.0eV}{kT}\right)} \text{ cm}^2/\text{s}$$

for treatment in 9 mTorr $CdCl_2$ and 150 mTorr O_2. Figure 14.14 shows that only grain-boundary diffusion, via surface reactions, is sensitive to $CdCl_2$ and O_2 concentration. For treatment in air, the grain-boundary diffusion dependence on $CdCl_2$ partial pressure (P_{sat}) over the range 1–100 mTorr has been determined for cells with CdTe deposited by vapor transport [117]:

$$D_{\text{GB}} = 9.0\text{x}10^4 P_{\text{sat}}{}^{2.73}\text{e}^{-\left(\frac{2.0}{kT}\right)} \text{ cm}^2/\text{s}$$

Alloy formation has both beneficial and detrimental effects. The interdiffusion process narrows the absorber-layer bandgap, resulting in higher long-wavelength quantum efficiency. Although

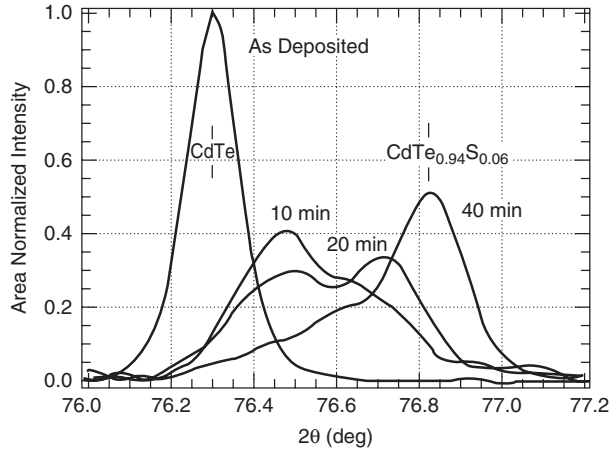


Figure 14.13 Time-progressive X-ray diffraction line profiles for the (511)/(333) reflection for PVD CdTe/CdS thin film structures deposited at 250 °C and treated in $CdCl_2$:Ar:O_2 vapor at 420 °C. Positions of pure CdTe and $CdTe_{1-x}S_x$ alloy with $x = 0.06$ are indicated (McCandless, unpublished results)

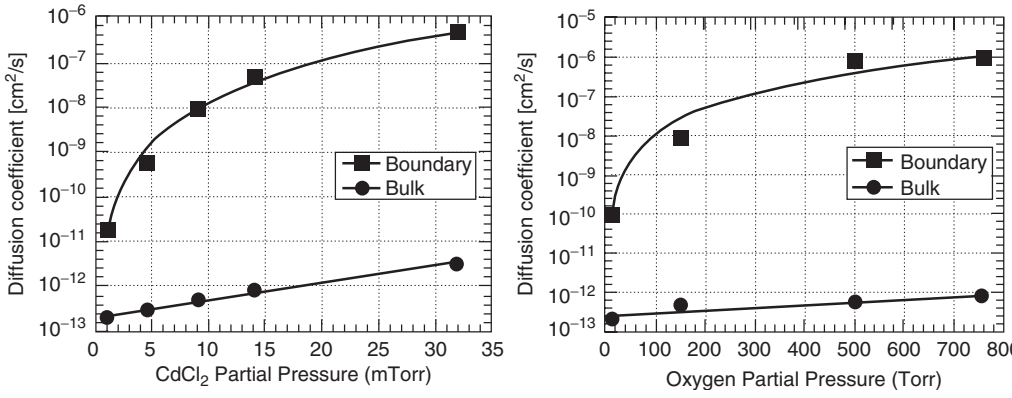


Figure 14.14 Sensitivity of bulk and grain-boundary diffusion coefficients (a) to $pCdCl_2$ at constant $pO_2 \sim 125$ Torr, at $T = 420$ °C; (b) to pO_2 at constant $pCdCl_2 = 9$ mTorr, at $T = 420$ °C (McCandless, unpublished results)

this gain is somewhat offset by a reduction in the built-in voltage, open-circuit voltages exceeding 820 mV have been obtained in cells having alloyed $CdTe_{1-x}S_x$ absorber layers with composition $x > 0.05$, made by spray pyrolysis [172]. Intermixing reduces interfacial strain [73] and may reduce the dark recombination current [83]. The CdS film thickness is reduced, which can be beneficial for window transmission, but nonuniform CdS consumption can result in lateral junction discontinuities and higher junction recombination current. Thus, alloy formation can be more pronounced in films deposited at low temperature having small grains and high grain-boundary density, such as by electrodeposition. In such films, a high partial pressure or concentration of $CdCl_2$ together with O_2 during $CdCl_2$ treatment will result in considerable alloy formation; optimization of the $CdCl_2$ treatment for cells with different CdTe thickness requires consideration of the grain structure. For cells with CdTe deposited at high temperature in the presence of Cl species during growth, such as

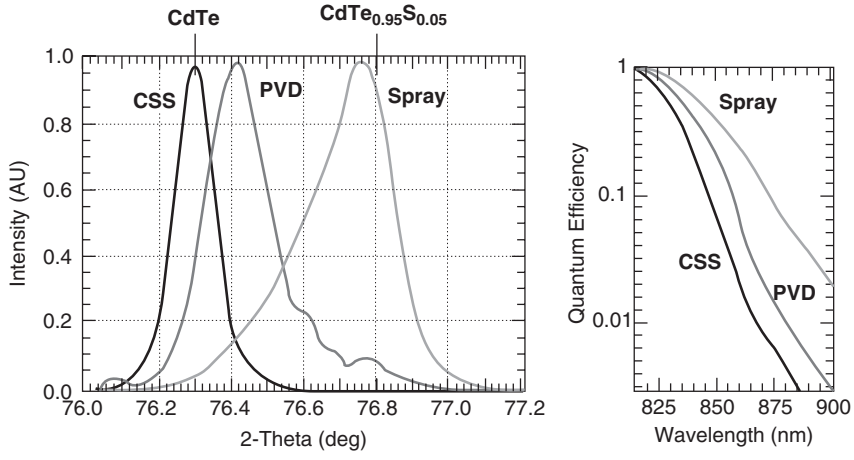


Figure 14.15 Correlation of XRD (511)/(333) reflection with long-wavelength quantum efficiency (McCandless, unpublished results)

sintering, alloy formation is dramatically enhanced and has resulted in devices with uniform alloy composition, with $x = 0.05$ for films formed at 550 °C.

Figure 14.15 shows the relationship between alloy formation in the absorber and the long-wavelength QE. For CSS cells, deposited at 600 °C in the absence of $CdCl_2$ and having large grains, the absorber layer exhibits a narrow XRD profile, as shown in Figure 14.15, at the CdTe position, indicating a negligible degree of alloy formation, and hence low CdS diffusion into CSS-deposited CdTe after $CdCl_2$ treatment. The long-wavelength QE edge occurs near that expected for pure CdTe wavelength. In contrast, cells made by spray pyrolysis, with the CdTe film formed at high temperature in the presence of $CdCl_2$, exhibit an asymmetrical XRD line profile with its peak corresponding to $CdTe_{0.96}S_{0.05}$ and a tail extending toward pure CdTe. The XRD line profile thus indicates nonuniform S distribution in the absorber layer, evident in the QE plot as a shallow decrease in long-wavelength response. The PVD case is intermediate to the CSS and spray cases, having smaller grains than CSS films and receiving less exposure to $CdCl_2$ than sprayed films. It appears, however, that apart from the differences in current generation, device operation is fundamentally similar for cells with differing amounts of $CdTe_{1-x}S_x$ alloy in the absorber layer.

14.5.4 Back Contact

The top region shown in Figure 14.7 above is the back contact, consisting of a primary contact to CdTe, which typically consists of a tellurium-rich p^+ surface, and a secondary contact, which is the current-carrying conductor. As with other wide-bandgap p-type semiconductors, there is a tendency to form a Schottky barrier with essentially all metals. Hence achieving a low-resistance ohmic contact has proven to be challenging. The most common strategy is to form a Te-rich surface by selective chemical etching and then to apply copper or a copper-containing material. Copper will react with Te to form a p^+-layer that can then be contacted with a metal or with graphite. The subtelluride Cu_2Te has been directly detected at the back surface using GIXRD methods [173]. Also, Cu acts as a relatively shallow donor in CdTe and can be diffused into CdTe from a doped contact material such as graphite paste [34] or ZnTe:Cu [174].

There are a variety of surface treatments that have been used, prior to formation of the copper layer, to reduce the back barrier. Table 14.3 summarizes representative back-contact formation

Table 14.3 Examples of back-contact formation methods (NP = nitric + phosphoric acid mixture, BDH = sequential reaction in bromine, acidic dichromate, and hydrazine)

CdTe deposition method	Surface treatment	Primary contact	Thermal treatment	Additional contact	Reference
PVD	Te + H_2	Cu	200 °C/Ar	C	[177]
ED	BDH	Cu	None	Ni or Au	[178]
Spray	Etch	C + dopant	None	None	[179]
Screen	None	C + Cu dopant	400 °C/N_2	None	[180]
VTD	BDH	Cu	200 °C/Ar	C	[152]
CSS	NP Etch	C + HgTe + Cu	200 °C/He	Ag paste	[34]
Sputter	Br Etch	ZnTe:N	In situ	Metal	[181]
MOCVD	Br Etch	ZnTe:Cu	In situ	Metal	[182]

methods used to form low-resistance contacts on devices of moderate and high efficiency. Although fabrication laboratories tend to utilize a single surface treatment and contact material, there is little evidence that the preferred contact process is dependent on the CdTe deposition method [175], but it is highly sensitive to film thickness [176, 117].

The high bulk-diffusion coefficient for Cu in CdTe, $\sim 3 \times 10^{-12}$ cm^2/s at 300 K, coupled with its multiple valence states and weak Cu–Te bond give rise to potential stability issues related to the use of Cu to reduce the contact barrier. Alternatives to Cu in the back contact process have been explored with moderate success, such as ZnTe:N [183] and Sb_2Te_3 [184], exhibiting reasonably low-resistance contact to the CdTe, forming only modest barriers when used in conjunction with an appropriate metal as the current-carrying conductor. To date, however, contacts demonstrably free of Cu have not shown significant promise for high performance.

One consequence of a back-contact barrier, and the value of keeping it small, is evident in calculated band diagrams (Figure 14.16) for a 2-μm-thick CdTe layer [185]. The back contact is essentially a second diode with opposite polarity and smaller barrier than the primary junction. For

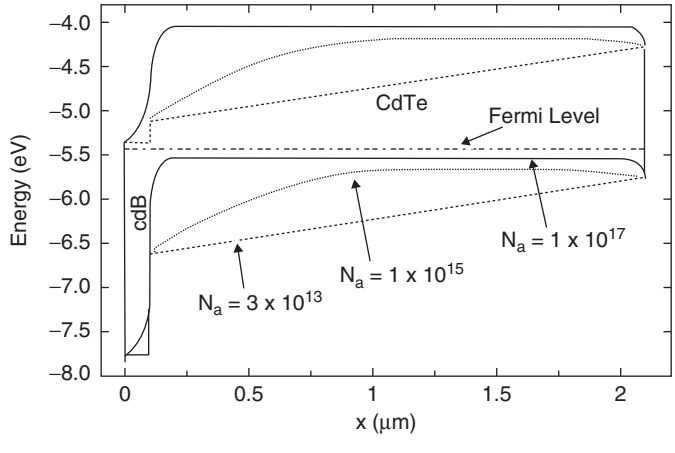


Figure 14.16 CdTe/CdS junction band diagrams at $V = 0$ for three values of CdTe acceptor density and a back-contact barrier

a thick absorber, or one with reasonably large carrier density (solid curve), the bands are flat over most of the absorber thickness, the primary junction effectively blocks the forward current, and a modest back barrier has only a minor effect on the current–voltage curves. For carrier densities more typical of CdTe (between dashed and dotted lines), however, the depletion widths of the primary and back-contact diodes overlap. The effective reduction in the barrier for electrons at the back contact means that forward current can flow more easily, reducing the open-circuit voltage. Even if the depletion regions do not overlap, large electron lifetimes can also lead to excessive forward current and reduce voltage [186].

14.5.5 Cell Characterization and Analysis

14.5.5.1 As-deposited cell operation

Considerable information about the electrical properties of CdTe solar cells has been deduced from straightforward measurements of current versus voltage ($J-V$), quantum efficiency (QE) and capacitance versus frequency ($C-f$) and voltage ($C-V$). More detailed information can be obtained from the temperature dependence or the time evolution of these curves.

The $J-V$ curve (Figure 14.2 above) for the record-efficiency cell follows a standard diode equation with additional factors included to take account of circuit resistance and forward-current mechanisms other than thermionic emission.

$$J = J_0 \exp[(V - JR)/AkT] - J_{SC} + V/r \tag{14.1}$$

For the illuminated $J-V$ curves of the record-efficiency CdTe cell, the prefactor J_0 is 1×10^{-9} A/cm^2, the series resistance R is 1.8 Ω cm^2, the diode quality factor A is 1.9, and the shunt resistance r is 2500 Ω cm^2. The values for R and A were found using the technique described in Reference [187].

The forward-current ($J + J_{SC}$) data for the record-efficiency CdTe cell is replotted in Figure 14.17 using a logarithmic scale, and the analogous data of a high-efficiency GaAs cell [188] is shown for comparison. Both curves have had small corrections for series and shunt resistances (R and r). Since the absorber materials have similar bandgaps, the CdTe and GaAs cells should have similar $J-V$ curves. The V_{OC} difference shown, however, is more that 200 mV. At maximum power (MP), the voltage difference is larger yet, approaching 300 mV, because the CdTe A-factor is 1.9 compared with 1.0 for the GaAs. The physical difference is due to the additional recombination-current paths for the CdTe junction. This excess forward current for the CdTe cell is roughly two orders of magnitude greater than that of the GaAs cell under normal operating conditions.

The implication from Figure 14.17 is that there is considerable room for improvement of V_{OC} in CdTe cells through reduction in such recombination, or equivalently, an increase in the commonly used lifetime parameter. The correlation between minority-carrier lifetime (τ) and open-circuit voltage is in fact quite strong. The lifetime can be measured reasonably reliably by time-resolved photoluminescence (TRPL), where the PL signal is tracked as a function of time following an excitation pulse. Figure 14.18 shows the $V_{OC}-\tau$ correlation for a large number of CdTe cells [189]. The record-efficiency cell is in the upper right with a lifetime value of about 2 ns. An obvious question posed by Figure 14.18 is whether further increases in the CdTe lifetime, basically a materials issue, will lead to significantly higher voltage, and hence efficiency. This question is addressed in Section 14.7, which deals with the future of CdTe cells.

The effect of a back-contact barrier on the device band diagram was shown earlier in Figure 14.16. The effect on measured $J-V$ curves is illustrated in Figure 14.19. Here the two

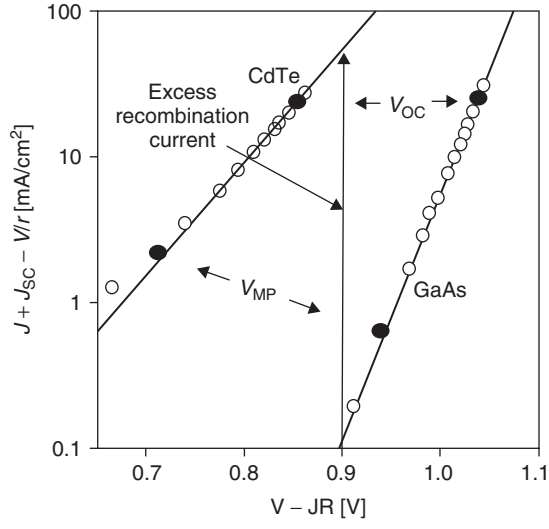


Figure 14.17 Comparison of high-efficiency CdTe and GaAs-illuminated $J-V$ data on a logarithmic plot, corrected for resistive effects

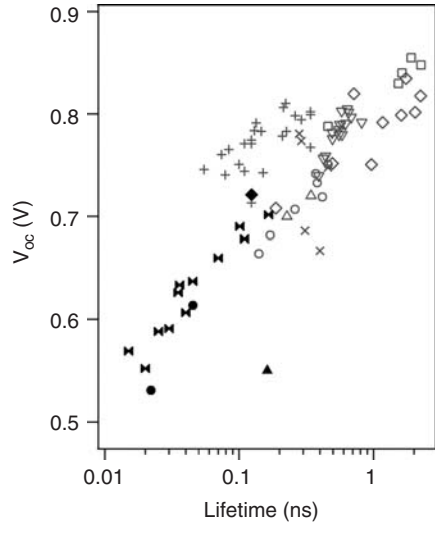


Figure 14.18 Open-circuit voltage versus lifetime from TRPL measurements from more than 80 CdTe cells. Figure courtesy W. Metzger 2010, used by permission

depletion regions did not overlap, the two diodes were treated as independent circuit elements. The calculated fit to the curves, assuming a back-diode barrier, is in fact quite good. The impact of the back barrier becomes larger as the temperature is lowered. This impact is very dramatic in the first quadrant, where the shape of the $J-V$ curves is commonly described as "rollover" [190–192]. The degree of "rollover" of $J-V$ curves such as those illustrated in Figure 14.19 has been shown to be related to the amount of copper used in the fabrication of the back contact [193]. With smaller amounts of copper, rollover is observed at higher temperatures, implying a larger back-contact

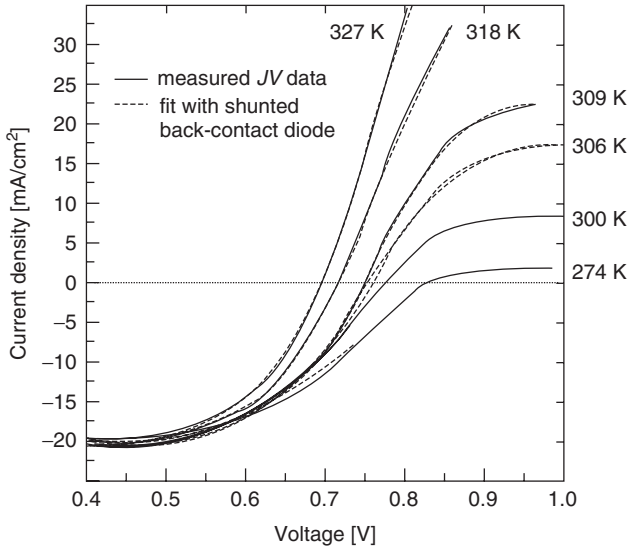


Figure 14.19 Measured and calculated curves for a CdTe cell with a back-contact barrier [190]

barrier, which has a greater impact on cell performance. Increased rollover leads to decreased FF, so there is in general an inverse correlation between Cu content and back barrier.

14.5.5.2 Accelerated life: "stress" testing

While the addition of Cu can reduce the back-contact barrier, and hence inhibit rollover, it can also lead to a decrease in CdTe cell stability, at least for cells held at elevated temperatures. Several authors have seen significant performance changes when CdTe cells were held at high temperatures (typically 60–110 °C) for extended periods of time [194–198]. Such studies, often referred to as "stress" tests, generally report an initial decrease in fill factor, followed by a decrease in V_{OC}. Only in extreme cases is J_{SC} affected. Figure 14.20 shows the illuminated $J-V$ curves for an NREL-manufactured CdTe cell soon after it was fabricated, and at different periods of time following light-soaking at 100 °C under open-circuit bias. Similar curves have been seen with cells from other manufacturers when they also were exposed to 60–110 °C. The dark $J-V$ curves for such cells also show a progressive increase in rollover with continuing temperature stress. Qualitatively similar $J-V$ characteristics are measured for the devices of Figures 14.19 and 14.20; for stressed cells, there is an increase in rollover as measurement temperature is decreased or as time at open-circuit 100 °C stress is increased. As the temperature of the $J-V$ measurement is reduced, the acceptor concentration decreases, resulting in the back-barrier height exerting a greater effect, as shown in Figure 14.16 above. The rollover in the stressed devices could likewise be due to a reduction in carrier density in the CdTe layer.

There is conflicting evidence of the dependence of CdTe cell performance changes with stress temperature, but one study [196] deduced an empirical activation energy near 1 eV, which predicts that $J-V$ changes at 100 °C are accelerated 500–1000 times compared with a typical annual range of outdoor temperatures for solar panels. Hence, one would not expect significant performance changes for modules until they had been deployed in the field for many years.

Additional "stress" studies have shown that reductions in cell efficiency are less when the voltage bias is held at short-circuit or maximum power, compared with open-circuit [196–198],

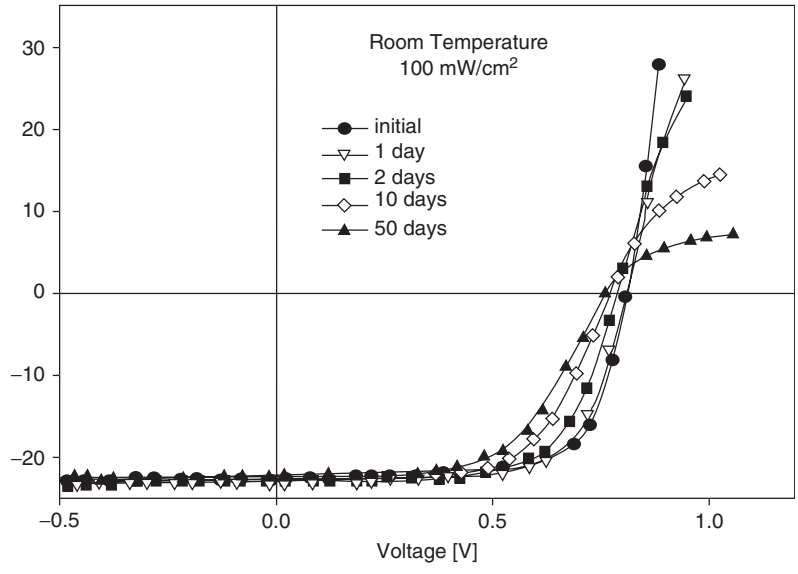


Figure 14.20 $J-V$ of CdTe cells following extended exposure to 100 °C at V_{OC} [196]

or when less Cu is used in the back contact [193]. Also, cells made with no intentional Cu can degrade, and these cells in their as-deposited state exhibit similar characteristics, Jo, $J_L(V)$, rollover to those of stressed cells containing Cu [199]. The movement of copper out of the back-contact region should be faster when forward bias reduces the electric field within the cell [196, 198, 200, 201]. Cu movement away from the back contact can have at least two effects on cell performance: (1) an increase in back-barrier height, similar to the copper-free case; and (2) a detrimental effect on cell performance due to Cu movement toward the front junction, which can increase CdTe recombination states and hence reduce the carrier lifetime.

14.5.5.3 Optical losses

Optical losses can be separated into those absorbed or scattered before reaching the CdTe absorber and those due to incomplete absorption of the photons penetrating through the absorber. The quantum efficiency (QE) of a solar cell, especially when combined with independent reflection and absorption measurements of the cell and the individual window layers, is a powerful tool to analyze these losses. The data shown in Figure 14.21 are for a CdTe cell with relatively large losses in J_{SC}. This cell, with 4-μm CdTe and 0.2-μm CdS was deliberately chosen to illustrate the analysis process and is not indicative of the cells currently being manufactured.

To quantify the photon losses in Figure 14.21, the measured QE is multiplied by the light spectrum, in units of photons/cm^2/nm, integrated over wavelength, and multiplied by unit electric charge to yield J_{SC}. For comparison, the maximum current density J_{max} is found to be 30.5 mA/cm^2 assuming unity QE up to the bandgap cutoff wavelength (820 nm).

The optical regions shown in Figure 14.21 are the reflection from the cell, the absorption of the glass superstrate, the absorption of the SnO_2 conductive contact, and the loss below 500 nm associated primarily with absorption of a ~250-nm-thick CdS window. The separation of individual absorption terms was deduced from the reflection and transmission of partially completed cell structures terminated after each layer. The remaining loss region approaching the bandgap is

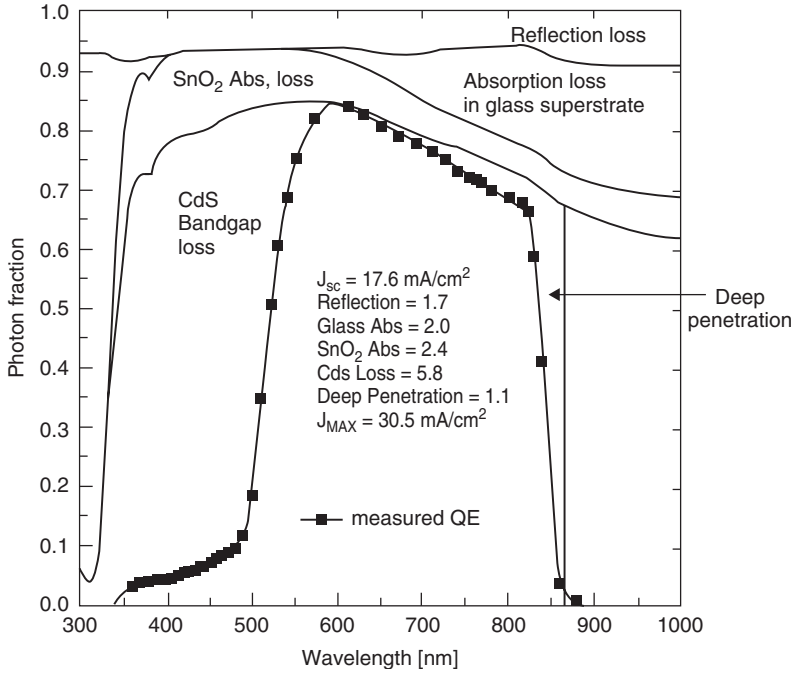


Figure 14.21 CdTe cell photon losses and quantum efficiency versus wavelength [202]

assumed to be due to the photons that penetrated too deeply for complete collection. Integration of these loss spectra weighted by the illumination spectrum, quantifies the loss in terms of the current density. These losses are listed in the inset to Figure 14.21 for the standard global AM1.5 spectrum [203], normalized to 100 mW/cm^2. The sum of the losses is the difference between the measured and the maximum current densities.

The inset tells us quantitatively that considerable current enhancement is possible in this case by employing a thinner CdS window, and to a lesser extent with a different glass or improved SnO_2 process. Potential for improvement in J_{SC} with an antireflection coating or from better collection of the deeply penetrating photons is relatively small. It is certainly possible, however, to reduce the larger losses. The record efficiency cell, for example, has a J_{SC} of 26 mA/cm^2. Its losses from glass, SnO_2, and CdS absorption are each more than a factor of five less than the cell shown in Figure 14.21, and no single loss factor contributes more than 1 mA/cm^2.

14.5.5.4 Capacitance analysis

The capacitance of a solar cell can yield information about extraneous states within the bandgap, and it can often give a credible profile of the carrier density within the absorber [204]. Figure 14.22(a) shows the measured capacitance of a CdTe cell as a function of frequency for three biases: 0, −1, and −3 V. The capacitance magnitude is relatively small, corresponding to a small carrier density and a large depletion width. The large depletion width implies that most photons are absorbed in a region with an electric field and hence photocurrent should not vary significantly with voltage [205], although significant voltage dependence of photocurrent collection can occur in cells with lower lifetimes [206]. The fact that the capacitance curves are relatively flat over nearly three decades in frequency strongly suggests that they are not significantly affected by extraneous states. The upturn

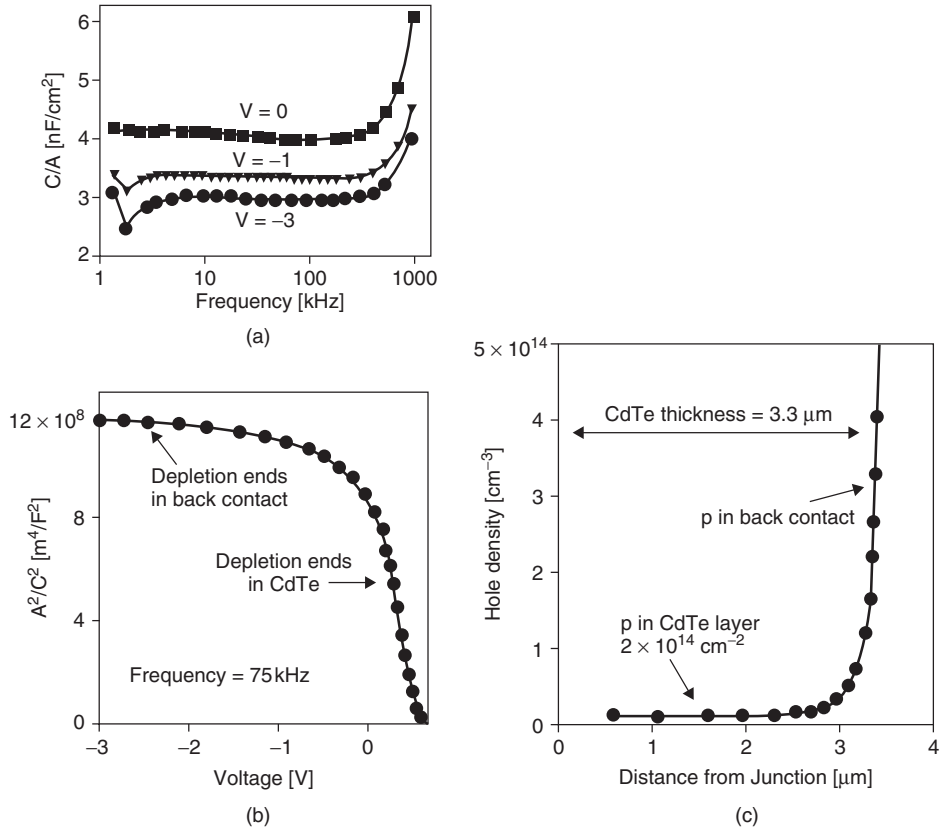


Figure 14.22 Capacitance measurements and hole-density determination for a CdTe solar cell

at high frequencies is due to circuit inductance [207], and the outlier points at low frequencies are due to an apparatus anomaly. Typically, a frequency in the middle of the flat region, in this case 75 kHz, would be chosen for subsequent $C-V$ measurements.

Capacitance versus voltage data at 75 kHz for the same cell is plotted in Figure 14.22(b) in the commonly used C^{-2} versus V format, in which the vertical axis is proportional to the square of the depletion width w for a given voltage, since $C/A = \varepsilon/w$. The slope is inversely proportional to the carrier density at the depletion edge. In this case, there are two distinct regions. In reverse bias, C^{-2}, and hence the depletion width, changes very little. Near zero bias and into forward bias, however, the depletion width narrows considerably.

The same data can be plotted (Figure 14.22c) as hole density ρ versus depletion width, which is referred to as the distance from the junction, since the depletion is assumed to be essentially all in the absorber. The two regions suggested by the C^{-2} versus V plot have become extremely clear. For the first 3 μm into the CdTe, the hole density is very low (mid-10^{14} range), but then it increases dramatically. The interpretation is that 3 μm is the thickness of the CdTe layer, and the rapid increase in hole density takes place as the depletion edge enters the back-contact region. In fact, the measured thickness of this particular CdTe was somewhat larger than 3 μm. The likely explanation is that there are local areas along grain boundaries where the back-contact material penetrates into the CdTe and effectively reduces its thickness. Following the elevated temperature cycles that led to Figure 14.20, there can also be a reduction in the electronic CdTe thickness.

CdTe solar cells fabricated by a variety of the techniques illustrated in Figure 14.6 have shown hole densities in the low-10^{14}/cm^3 or high-10^{13}/cm^3 range as determined from capacitance. A relatively low carrier density may be an impediment to higher-efficiency cells. It suggests that compensating native donors remain a problem for the CdTe cells made to date. The direct impact of the low hole density is that the Fermi level is 250–350 mV from the valence-band maximum, and hence limits the junction barrier, and therefore V_{OC}. A large problem is that the low densities are likely to be symptomatic of the excessive recombination states responsible for the voltage deficit seen in Figure 14.17.

14.6 CdTe MODULES

A CdTe photovoltaic module generally consists of electrically interconnected CdTe cells on a superstrate that serves as the mechanical support. The module's electrical output depends on individual cell output, interconnection scheme, and losses due to non-generating areas and interconnection-resistive losses. Obtaining high cell efficiency at the module scale depends on the successful transfer of small-area batch processing steps, such as the $CdCl_2$ treatment, to either large-area batch or continuous processing, and on minimizing resistive losses, optical losses due to the use of low-cost glass, and locally defective cell areas. In effect, the manufacturer's goal is to obtain a series-connected array of large-area CdTe/CdS diodes having spatially uniform physical and electrical properties in a high-throughput fabrication process.

Unlike cells based on strings of crystalline cells, thin film CdTe solar cells can be connected by monolithic integration, which has proven to be a robust technology and results in modules which are less sensitive to fractional shading. With monolithic integration, the cells on a single large-area substrate or superstrate are isolated and interconnected by judiciously scribing through the deposited layers at different stages of fabrication. Scribing can be achieved by mechanical means or, preferably, by laser scribing in which the stopping point is determined by matching the absorption properties of the different layers with the appropriate wavelength and power density [208]. A monolithically interconnected module is depicted in Figure 14.23. Generally three scribes are used: the first separates the TCO front contact to define adjacent cells, the second through the CdS and CdTe provides an electrical path from the TCO to the back contact of an adjacent cell, and the third isolates the back contact between cells. Collectively these scribes introduce a dead area on the module, which should be kept as small as practical. The photocurrent generated by each individual

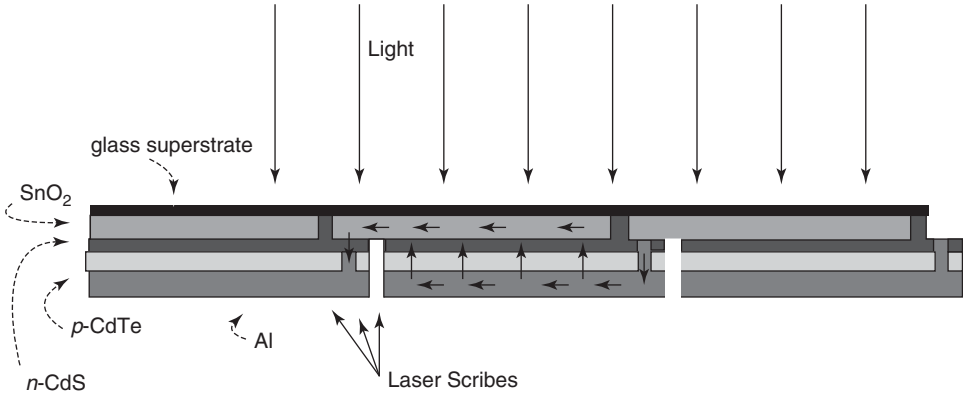


Figure 14.23 Schematic of a series-connected integrated CdTe module having three laser scribes

cell flows from one end of the module to the other. The module voltage is simply the sum of the voltages from the series connected individual cells. This monolithic structure and its three laser scribes are very similar to many amorphous silicon- and CIGS-based PV modules deposited on glass. Significant design analysis of module geometry and lateral-sheet resistances, to minimize resistive and dead-area losses, has been carried out for amorphous silicon PV modules, and is applicable to CdTe PV modules [209, 210].

The structure shown in Figure 14.23 utilizes a sheet of glass as the supporting superstrate. The choice of glass is based on the cost per watt, optical loss, and thermal tolerance of the glass. For example, the light-generated current in CdTe cells deposited on borosilicate glass is about 2 mA/cm^2 higher than on soda-lime glass, due to the higher optical absorption in the soda-lime glass at wavelengths beyond 600 nm. However, the high-transparency glasses, such as borosilicates, fused quartz, and Vycor, require more refining and in general are significantly more costly than the soda-lime glass commonly used for windowpanes. A review of the optical properties of these and other glasses is found in Reference [211]. An additional issue with low-cost soda-lime glass is that it exhibits a lower softening temperature, which restricts its compatibility with high-temperature processing. A useful compromise is to reduce the iron content of soda-lime glasses to improve their transparency at wavelengths greater than 600 nm and increase their melt temperature, but at a lower cost than borosilicate or other highly refined glasses.

Commercialization of CdTe PV modules relies on a constant supply of affordable raw materials, especially cadmium and tellurium. The material requirements for a facility with 1GW/year manufacturing capacity, with 100% utilization of Cd and Te to fabricate a 2-μm-thick absorber layer, are approximately 40 metric tons of cadmium and 60 metric tons of tellurium. Both elements are obtained as by-products in the smelting of ores; cadmium is obtained from zinc, copper, and lead refining, while tellurium is primarily obtained as a product of the electrolytic refining of copper and skimmings from lead production [212]. Tellurium is the scarcer and more costly component, however, tellurium availability is estimated to be ~1600 metric tons per year [213]. At present, the costs for ~95% pure cadmium and tellurium are ~$US 12/lb ($US 24 000/ton) and ~$US 20/lb ($US 40 000/ton), respectively. Thus, the total annual cadmium and tellurium costs for the 1 GW capacity plant would be ~$US 2.6M, which amounts to less than $US 0.03 per watt. This is significantly less than the price of the superstrate glass/TCO for the 100 km^2 required for 1 GW capacity, which is currently about $US 5.00/m^2 [214], or approximately $US 0.50 per watt. A detailed analysis of PV-manufacturing costs is presented in reference [215].

Present-generation CdTe PV modules are typically 0.72 m^2 in area and have achieved efficiencies above 10%, with peak power of the order of 75 W. At the time of this writing, the only large-scale commercial CdTe modules are manufactured by First Solar, LLC, in Toledo, Ohio, in Germany, and in Malaysia (see Figure 14.24) with an annual production of over 300 MW [216]. Three other CdTe module manufacturers have produced CdTe modules on a smaller scale: Matsushita Battery Company in Japan, BP Solar, in Fairfield, California; and Antec Solar GmbH, in Germany. An additional three companies, Abound Solar (formerly AVA), Calixo/Q-cells, and PrimeStar Solar are aggressively building their facilities and anticipate production capacity that may rival that of First Solar. Several companies are engaged in developing alternatives to the superstrate module. At least one company, such as Solexant, Inc., in San Jose, California, is developing roll-to-roll technology, while others are pursuing lift-off and transfer technology.

The First Solar module utilizes vapor transport deposition of CdTe onto moving substrates to achieve a high growth rate while maintaining high substrate temperature. The modules are formed on tin oxide-coated soda-lime TEC glass made by Pilkington. The production line is reported to be capable of 2.9 m^2 per minute throughput [217]. The success of First Solar validates the concept of high-throughput, large-area, robust manufacturing of the lowest-cost modules at present on the market. A photograph showing a First Solar module and module array is shown in Figure 14.24.

Figure 14.24 CdTe module and module array made by First Solar (photographs used by permission)

Issues confronting manufacturers of thin film CdTe solar cells include the transfer of small-area efficiencies to the module scale, control of CdTe uniformity during growth, reproducibility, and product certification with respect to expected lifetime. One example of a key trade-off is the CdS film thickness; thick CdS films improve processing latitude, but reduce light-generated current. With the ultra-thin CdS films used to obtain state-of-the-art performance in small-area cells, spatial variations in CdTe-film microstructure can affect the diffusion of CdS into CdTe and the resulting junction structure, highlighting the importance of area control over the film morphology and $CdCl_2$ treatment process. Incorporating buffer layers into the window-layer side of the structure, refining the post-deposition treatments, and developing in-line sensing diagnostic tools are viable pathways for widening the tolerance window needed to raise the current density in CdTe modules.

14.7 FUTURE OF CdTe-BASED SOLAR CELLS

During the past 30 years, there has been steady progress in refining the basic CdTe cell structure of Bonnet and Rabenhorst. The highest current densities achieved are similar to crystalline GaAs when adjusted for small differences in bandgap. Open-circuit voltage and fill factor are limited by excessive forward-current recombination, low bulk lifetime and low carrier density, but have nevertheless achieved values about 80% as large as GaAs, again adjusted for bandgap. There is some concern about the impact of diffusion of Cu atoms on stability, but degradation under normal operating conditions for well-fabricated cells is becoming negligible. Although the status of CdTe solar cells is healthy enough for expanded mainstream commercialization, no clear processing path to reach 20% efficiency has been identified, and significant cell-level basic research is thus required to overcome the factors limiting performance.

The CdTe thin film industry is now well-established, but its future depends on continued injection of creative ideas in the study of the material properties of CdTe devices, laboratory-scale device performance, and improvements to the photovoltaic modules. Translation of single-junction efficiency gains from the laboratory should additionally reduce module cost per watt. Achieving these goals, however, requires greater understanding of the relationship between processing conditions and critical material properties needed for high efficiency and long-term reliability.

Although the fundamental nature of polycrystalline CdTe is not fully understood, there are credible pathways towards the development of higher-efficiency single-junction thin film cells, demonstrating the practical merit of the empirical approach. In addition, several CdTe film deposition techniques yield similar device performance, by a suitable combination of post-deposition processing and back-contact formation. Reaching the 20% efficiency target and translating this to high module performance relies on determining and overcoming the mechanisms that limit the open-circuit voltage, and to a lesser extent fill factor, in present-generation cells. The current density in champion thin film CdTe cells has already reached 90% of its theoretical maximum for AM1.5 illumination and is primarily limited by well-understood optical losses in the glass/TCO/CdS structure.

Despite advances in commercial module development, the values of $V_{OC} \sim 850$ mV and $FF \sim 75\%$ have not changed for several years and fall short of the expectations based on bandgap. An increase in CdTe cell voltage of 200 mV would bring it very close to that achieved by similar-bandgap GaAs and would push efficiency up to 20%. There are basically two possibilities to increase voltage by this amount [218]. One is to simultaneously increase the CdTe carrier density by three orders of magnitude and its lifetime by a factor of ten. This strategy, however, requires major improvements in materials quality, and attempts to date to increase these parameters have not been successful. The second and probably more realistic possibility is to fabricate a fully depleted *n-i-p* structure with carrier density and lifetime comparable to that currently achieved, but to add a back electron-reflector layer between the absorber and back contact. An extended bandgap from an appropriate alloy, such as $Cd_{1-x}Zn_xTe$ or $Cd_{1-x}Mg_xTe$ discussed below with $x \sim 0.25$ or 0.15 respectively, should be sufficient. The caveat is that its deposition would need to be continuous with the CdTe absorber and not introduce significant interfacial states.

A second area ripe for future CdTe-cell exploration is the possibility of much thinner absorber layers. There would be clear advantages in terms of material utilization and deposition time, but possibly also for the most effective use of a back electron reflector. Recent work at the University of Toledo has achieved 10% efficiency for a 0.5-μm-thick CdTe absorber [219]. Part of that success was the result of adjusting the length of the CdTe treatment to be proportional to the thickness of the CdTe layer.

In addition to being a strong candidate for single-junction photovoltaic devices, CdTe can be alloyed with other II–VI compounds to alter its bandgap and allow the possibility that multijunction thin film cells can be fabricated [220]. The multijunction cell structures using CdTe-based wide-bandgap cells in monolithic structures must confront cell design, as well as both processing temperature and chemical stability. Materials based on alloys between CdTe and other group II–VI compounds will in principle allow a wide range of optoelectronic properties to be incorporated into devices by design. These semiconducting compounds provide a basis for the development of tunable materials, obtained by alloying different compounds in pseudo-binary configurations. For photovoltaic heterojunction devices, semiconductors using Cd, Zn, Hg cations and S, Se, Te anions exhibit a wide range of optical bandgaps, suggesting their potential for use in optimized device designs by tailoring material properties (Table 14.4). The generally high optical absorption coefficients, $\sim 10^5$/cm, and direct optical bandgaps of many II–VI semiconductors make them suitable for use in thin film photovoltaic devices. For the development of two-junction cells, top cells with an absorber bandgap near 1.7 eV and bottom cells near 1.1 eV are desirable [221, 222]. Another possibility, however, is to reduce the CdTe thickness to achieve current-matching.

The alloy systems shown in Table 14.4, separated by cation and anion substitution in pseudo-binary compounds, define a broad range of optical bandgap suitable as absorber layers in terrestrial photovoltaic converters. The isostructural systems $Cd_{1-x}Zn_xTe$ and $Hg_{1-x}Zn_xTe$ offer tunable systems with a wide range of bandgaps and controllable *p*-type conductivity. Thin film solar cells based on $Cd_{1-x}Zn_xTe$ were the subject of studies in the late 1980s by several laboratories, including the Georgia Institute of Technology (GIT) and International Solar Energy Technology

Table 14.4 Properties of pseudobinary II–VI alloys suitable for absorber layers

Compound	Single crystal optical E_g range 300 K [eV]	Optical bowing parameter	Stable endpoint structure	Miscibility gap?
Cation substitution				
$Cd_{1-x}Zn_xTe$	1.49–2.25	0.20	ZB–ZB	N
$Cd_{1-x}Mg_xTe$	1.50–3.00	0	ZB-W	?
$Hg_{1-x}Cd_xTe$	0.15–1.49	?	ZB-ZB	N
$Hg_{1-x}Cd_xSe$	0.10–1.73	?	ZB–W	N
$Hg_{1-x}Zn_xTe$	0.15–2.25	0.10	ZB–ZB	N
Anion substitution				
$CdTe_{1-x}S_x$	1.49–2.42	1.70	ZB–W	Y
$CdTe_{1-x}Se_x$	1.49–1.73	0.85	ZB–W	?
$CdSe_{1-x}S_x$	1.73–2.42	0.31	W–W	N
$HgTe_{1-x}S_x$	0.15–2.00	?	ZB–ZB	?
$HgSe_{1-x}S_x$	0.10–2.00	?	ZB–ZB	?

(ISET). Two approaches to depositing the $Cd_{1-x}Zn_xTe$ films had been considered in this work: synthesis by reaction of sequentially deposited metal layers (ISET) and metal organic chemical vapor deposition (GIT). CdS/$Cd_{1-x}Zn_xTe$ devices using $Cd_{1-x}Zn_xTe$ films made by the reaction of sequentially deposited metals with $x = 0.1$, corresponding to $E_g \sim 1.6$ eV, yielded 3.8% efficiency and suffered from low V_{OC} and FF [223]. Although little follow-up work was conducted to explain the low performance, for CdS/$Cd_{1-x}Zn_xTe$ devices made by MOCVD, it was found that the $CdCl_2$ + air treatment step reduced the bandgap from 1.7 to 1.55 eV, owing to chemical conversion of the zinc alloy to volatile $ZnCl_2$. The best cells made with the 1.55 eV bandgap yielded 4.4% conversion efficiency [224]. More recent work, with emphasis on control of oxidation and interface reactivity, has demonstrated the potential for ZnTe-based alloy thin film devices using VTD and CSS $Cd_{1-x}Zn_xTe$ absorbers with $ZnCl_2$ post-deposition treatments, yielding 12.4% efficiency for $Cd_{1-x}Zn_xTe$/CdS cells with $E_G = 1.58$ eV [225] and 2% efficiency for ZnTe/ZnSe cells with $E_G = 2.24$ eV [226].

Alloy systems in which a structural transformation occurs along the pseudo-binary tie-line, such as found with the $Cd_{1-x}Mg_xTe$ system, may restrict the usable alloy range to compositions away from the transition composition. In the $Cd_{1-x}Mg_xTe$ system, this occurs at x $\sim$ 0.7, and working heterojunction devices have been fabricated by groups at NREL at $E_G = 1.6$ eV [227] and the University of of Toledo [228].

CdTe-based thin film photovoltaic devices are also suited to applications beyond terrestrial power conversion, including space power generation, infrared detectors, and gamma radiation detectors. Using the current–voltage characteristics of state-of-the-art devices, AM0 operation at 60 °C can be determined by correcting for the temperature dependence of the bandgap and differences in the illumination spectrum. State-of-the-art cells with 16.5% AM1.5 efficiency at 25 °C translate to 13.9% AM0 efficiency at 60 °C, and typical cells having 12% AM1.5 efficiency at 25 °C translate to 10% AM0 efficiency at 60 °C. For cells on 0.05-mm-thick polyimide substrate at AM0 conditions, the 12% CdTe cells should yield a power-to-weight ratio of 1500 W/kg. CdTe research and development for space applications, in which AM0 power-to-weight ratio greater than 1000 W/kg is desired, has followed three approaches: (1) 6–7% AM1.5 efficiency for thin film

deposition in the *substrate* configuration on lightweight flexible substrates [26]; (2) 11% AM1.5 efficiency for transfer of completed superstrate cells from rigid superstrates to lightweight flexible substrates [229]; and (3) 11% efficiency for direct superstrate deposition onto 100-μm-thick glass foils [106]. Encouraging results of CdTe/CdS cell stability were obtained under 1-MeV electron bombardment at fluences of $10^{14}-10^{16}/cm^2$ [230].

Concerns about cadmium toxicity and the introduction of significant amounts of cadmium into the environment are essentially unfounded. The total life-cycle release of cadmium, normalized to GWh of electrical energy produced, has been shown to be much less than that from coal generation, and somewhat less than that from nuclear generation or silicon-based photovoltaic panels [231]. The management of cadmium in the manufacturing environment relies on a combination of appropriate engineering and chemical hygiene practices. Deployed modules are environmentally well sealed, which serves to both protect the cell from environmental deterioration and to contain the semiconducting materials, in the event of mechanical failure. By recycling modules at the end of life in a manner consistent with metal products, nearly all the cadmium in a module can be recycled at a cost of roughly 5 cents/W [232]. Alternatively, module deployment by a leasing arrangement or by confinement to industrially managed energy farms could facilitate total control over installed cadmium distribution. Part of the reason that the thin film CdTe technology is a negligible environmental issue is that the amount of cadmium used in thin film CdTe modules is quite small. A CdTe module 1 m^2 in area, producing approximately 100 W of power using a CdTe layer less than 2 μm thick, contains less than 10 g of cadmium, or about the same as a single nickel–cadmium flashlight battery [233].

This chapter has shown the development and present state of CdTe solar cells and modules. The cell-level research over the past three decades has now been translated into a large and impressively fast-growing CdTe industry. It is reasonable to expect that the rapid growth will continue and the manufacturing costs will be further reduced. As with all photovoltaic technologies, however, the ultimate success of the CdTe industry will depend on further improvements to the thin films, the device structures, and the large-scale manufacturing processes.

ACKNOWLEDGEMENTS

Brian McCandless wishes to acknowledge the contributions of the technical staff at the Institute of Energy Conversion, in particular Robert Birkmire, James Phillips, and Steven Hegedus for insightful discussions; Kevin Dobson and Joy Deep Dass for assistance in literature research; and Shannon Fields and Erten Eser for help with figure preparation. James Sites is deeply indebted to many colleagues in the CdTe community, but particularly to Dr. Alan Fahrenbruch and students who have done research with him in this area: Pete Mauk, Hossein Tavakolian, Rick Sasala, Ingrid Eisgruber, Gunther Stollwerck, Nancy Liu, Jennifer Granata, Jason Hiltner, Markus Gloeckler, Samuel Demtsu, and Jun Pan. Both authors gratefully acknowledge long-time research support from the United States National Renewable Energy Laboratory. The comparative nature of this work would not have been possible without access to samples and data generously provided by numerous CdTe research groups.

REFERENCES

1. Loferski J, *J. Appl. Phys*. **27**, 777–784 (1956).
2. Based on Rothwarf A, Boer K, *Prog. Solid State Chem*. **10**, 71–102 (1975).
3. Margottet J, *Annals Scientifiques de l'École Normale Supérieure*, 2nd edn, Vol 8, pp 247–298 (1879).
4. Frerichs R, *Phys. Rev*. **72**, 594–601 (1947).

5. Jenny D, Bube R, *Phys. Rev*. **96**, 1190–1191 (1954).
6. Krüger F, de Nobel D, *J. Electron*. **1**, 190–202 (1955).
7. de Nobel D, *Philips Res. Rpts* **14**, 361–399 and 430–492 (1959).
8. Rappaport P, *RCA Rev*. **20**, 373–397 (1959).
9. Mimilya-Arroyo J, Marfaing Y, Cohen-Solal G, Triboulet R, *Sol. Energy Mater*. **1**, 171 (1979).
10. Cohen-Solal G, Lincot D, Barbe M, *Conf. Rec. 4th ECPVSC*, pp 621–626 (Stresa, Italy, 1982).
11. Fahrenbruch A, Bube R, *Fundamentals of Solar Cells*, Academic Press, New York, 418–460 (1983).
12. Elliot J, Ed, *US Air Force ASD Technical* Report, pp 61–242 (1961).
13. Cusano D, *General Electric Res. Lab. Report*, No. 4582 (1963).
14. Bernard J, Lancon R, Paparoditis C, Rodot M, *Rev. Phys. Appl*. **1**, 211–217 (1966).
15. Lebrun *J, Conf. Rec. 8th IEEE Photovoltaic Specialist Conf*., pp 33–37 (1970).
16. Justi E, Schneider G, Seredynski J, *J. Energy Conversion* **13**, 53–56 (1973).
17. Ponpon J, Siffert P, *Rev. Phys. Appl*. **12**, 427–431 (1977).
18. Fulop G *et al., Appl. Phys. Lett*. **40**, 327–328 (1982).
19. Mitchell K, Fahrenbruch A, Bube R, *J. Appl. Phys*. **48**, 829–830 (1977).
20. Nakazawa T, Takamizawa K, Ito K, *Appl. Phys. Lett*. **50**, 279–280 (1987).
21. Aranovich J, Golmayo D, Fahrenbruch A, Bube R, *J. Appl. Phys*. **51**, 4260–4265 (1980).
22. Muller R, Zuleeg R, *J. Appl. Phys*. **35**, 1550–1556 (1964).
23. Dutton R, *Phys. Rev*. **112**, 785–792 (1958).
24. Mitchell K, Fahrenbruch A, Bube R, *J. Appl. Phys*. **48**, 4365–4371 (1977).
25. Yamaguchi K, Matsumoto H, Nakayama N, Ikegami S, *Jpn. J. Appl. Phys*. **16**, 1203–1211 (1977).
26. McClure J *et al., Sol. Energy Mater. Sol. Cells* **55**, 141–148 (1998).
27. Adirovich E, Yuabov Y, Yugadaev D, *Sov. Phys. Semicond*. **3**, 61–65 (1969).
28. Bonnet D, Rabenhorst H, *Conf. Rec. 9th IEEE Photovoltaic Specialist Conf*., pp 129–132 (1972).
29. Suyama N *et al., Conf. Rec. 21st IEEE Photovoltaic Specialist Conf*., pp 498–503 (1990).
30. Başol B, *Conf. Rec. 21st IEEE Photovoltaic Specialist Conf*., pp 588–594 (1990).
31. Tyan Y, Perez-Albuerne E, *Proc. 16th IEEE Photovoltaic Specialist Conf*., pp 794–800 (1982).
32. Meyers P, Liu C, Frey T, *U.S. Patent* 4,710,589 (1987).
33. Birkmire R, *Conf. Record NREL ARD Rev. Meeting*, pp 77–80 (1989).
34. Britt J, Ferekides C, *Appl. Phys. Lett*. **62**, 2851–2852 (1993).
35. Wu X *et al., J. Appl. Phys*. **89**, 4564–4569 (2001).
36. McCandless B, Hichri H, Hanket G, Birkmire R, *Conf. Rec. 25th IEEE Photovoltaic Specialist Conf*., pp 781–785 (1996).
37. Wu X *et al., Conf. Rec. 17th European Photovoltaic Solar Energy Conversion*, pp 995–1000 (2001).
38. Mitchell K, Fahrenbruch A, Bube R, *J. Appl. Phys*. **48**, 829–830 (1977).
39. Rakhshani A, *J. Appl. Lett.* **81**, 7988–7993 (1997).
40. Zanio K, in: *Semiconductors and Semimetals*, Vol. 13, *Cadmium Telluride*, Willardson R, Beer A (eds), Academic Press, New York (1978).
41. Fahrenbruch A, Colorado State University subcontract report (2000), unpublished.
42. Hartmann H, Mach R, Selle B, Wide Gap II-VI Compounds as Electronic Materials, in Kaldis E, (ed.), *Current Topics in Materials Science*, Vol. 9, North-Holland Publishing Company, New York (1982).
43. Madelung O, *Semiconductors Other than Group IV Elements and III-V Compounds*, Springer-Verlag, New York (1992).

44. Aven M, Prener J, (eds), *Physics and Chemistry of II-VI Compounds*, John Wiley & Sons, Inc., New York, pp 211–212 (1967).
45. International Committee for Diffraction Data, Card Number 15–770.
46. Knacke O, Kubaschewski O, Hesselmann K, *Thermochemical Properties of Inorganic Substances*, 2nd edn, Springer-Verlag, New York (1991).
47. Fonash S J, *Solar Cell Device Physics*, Academic Press, pp 78–79 (1981).
48. Data from *ASM Binary Phase Diagrams* (2000).
49. Hultgren R *et al., Selected Values of the Thermodynamic Properties of Binary Alloys*, American Society for Metals, Ohio, 627–630 (1971).
50. Muranevich A, Roitberg M, Finkman E, *J. Cryst Growth*, **64**, 285–290 (1983).
51. Krüger F, *Revue de Physique Appliquée* **12**, 205–208 (1977).
52. Wei S, Zhang S, Zunger A, *J. Appl. Phys*. **87**, 1304–1311 (2000).
53. Phillips J, *Bonds and Bands in Semiconductors*, 42, Academic Press, New York (1973).
54. Huheey J, Keiter E, Keiter R, *Inorganic Chemistry*, pp 113–126, Harper Collins, New York (1993).
55. Wu W, Gielisse P, *Mater. Res. Bull* **6**, 621–638 (1971).
56. Myers T, Edwards S, Schetzina J, *J. Appl. Phys*. **52**, 4231–4237 (1981).
57. Cohen M, Bergstresser T, *Phys. Rev*. **141**, 789–801 (1966).
58. Wei S, Zhang S, Zunger A, *J. Appl. Phys*. **87**, 1304–1310 (2000).
59. Wei S, Zunger A, *Appl. Phys. Lett*. **72**, 2011–2014 (1998).
60. Mathew X, Arizmendi J R, Campos J, *et. al*., *Solar Energy Materials and Solar Cells* **70**, 379–393 (2001).
61. Wei S, Zhang X, *Phys. Rev*. **B66**, 155211 (2002).
62. Castaldini A, Cavallini A, Fraboni B, *J. Appl. Phys*. **83**(4) 2121–2126 (1998).
63. Başol B, *Int. J. Sol. Energy* **12**, 25–35 (1992).
64. Capper P, Ed, *Properties of Narrow Gap Cadmium-Based Compounds*, INSPEC, London, pp 472–481 (1994).
65. Wei S, *Mtg. Record, National CdTe R&D Team Meeting* (2001) Appendix 9.
66. Unpublished work carried out by McCandless (2002).
67. Kazmerski L, *Sol. Cells* **24**, 387–418 (1988).
68. Mueller K, *Thin Solid Films* **174**, 117–132 (1989).
69. Levi D *et al., Sol. Energy Mater. Sol. Cells* **41/42**, 381–393 (1996).
70. Durose K, Edwards P, Halliday D, *J. Cryst. Growth* **197**, 733–742 (1999).
71. Dobson K, Visoly-Fisher I, Hodes G, Cahen D, *Sol. Energy Mater. Sol. Cells* **62**, 295–325 (2000).
72. Durose K, *et. al*., *Prog. Photovolt: Res. Appl*. **12**, 177–217 (2004).
73. Nakayama N *et al., Jpn. J. Appl. Phys*. **15**, 2281 (1976).
74. McCandless B, *Mat. Res. Soc. Symp. Proc*. **668**, H1.6.1–H1.6.12 (2001).
75. Ballif C, Moutinho H, Al-Jassim M, *J. Appl. Phys*. **89**, 1418–1424 (2001).
76. Rogers K *et al., Thin Solid Films*, **339**, 299–304 (1999).
77. Yan Y, Albin D, Al-Jassim M, *Appl. Phys. Lett*. **78**, 171–173 (2001).
78. Nunoue S, Hemmi T, Kato E, *J. Electrochem. Soc*. **137**, 1248–1251 (1990).
79. Martel A *et al., Phys. Status Solidi B* **220**, 261–267 (2000).
80. Dobson K, Visoly-Fisher I, Hodes G, Cahen D, *Adv. Mater*. **13**, 1495–1499 (2001).
81. Wu X *et al., J. Appl. Phys*. **89**, 4564–4569 (2001).
82. Waters D *et al., Conf. Rec. 2nd WCPVSEC*, pp 1031–1034 (1998).
83. Oman D *et al., Appl. Phys. Lett*. **67**, 1896–1898 (1995).
84. Rakhshani A, *J. Appl. Phys*. **81**, 7988–7993 (1997).
85. Li J, Podraza N J, Collins R W, phys. stat. sol. (a) **205** (4), 901–904 (2008).
86. Aspnes D, Arwin H, *J. Vac. Sci. Technol*. **A2**, 1309–1323 (1984).
87. Collins R *et al., 34th IEEE PVSC*, pp 389–392 (2009).

88. Fisher A *et al., Appl. Phys. Lett*. **70**, 3239–3241 (1997).
89. Grecu D *et al., J. Appl. Phys*. **88**, 2490–2496 (2000).
90. Okamoto T *et al., J. Appl. Phys*. **57**, 3894–3899 (1998).
91. Hegedus S, Shafarman W, *Prog. Photovolt.: Res. and Appl*. **12**(2–3), 155–176.
92. Rose D *et al., Prog. Photovolt*. **7**, 331–340 (1999).
93. Desai D, Hegedus S, McCandless B, Birkmire R, Dobson K, Ryan D, *Conf. Rec. 32nd IEEE PVSC and WCPEC-4*, pp 368–371 (2006).
94. Balcioglu A, Ahrenkiel R, Hasoon F, *J. Appl. Phys*. **88**, 7175–7178 (2000).
95. Dobson K *et al., Mat. Res. Soc. Symp. Proc*. **668**, H8.24.1–H8.24.6 (2001).
96. Galloway S, Edwards P, Durose K, *Sol. Energy Mater. Sol. Cells* **57**, 61–74 (1999).
97. Durose K, Edwards P, Halliday D, *J. Cryst. Growth* **197**, 733–740 (1999).
98. Bonnet D, (ed.), *Int. J. Solar Energy*, **12**, Harwood Academic Publishers, Reading, U.K. (1992).
99. Chu T, Chu S, *Prog. Photovolt*. **1**, 31–32 (1993).
100. Birkmire R, Eser E, *Annu. Rev. Mater. Sci*. **27**, 625–653 (1997).
101. Bonnet D, Meyers P, *J. Mater. Res*. **10**, 2740–2754 (1998).
102. Jackson S, Baron B, Rocheleau R, Russell T, *J. Vac. Sci. Technol. A* **3**, 1916–1920 (1985).
103. Jackson S, Baron B, Rocheleau R, Russell T, *AIChE J*. **33**, 711–721 (1987).
104. Fahrenbruch A, Bube R, Kim D, Lopez-Otero A, *Int. J. Sol. Energy* **12**, 197–222 (1992).
105. McCandless B, Youm I, Birkmire R, *Prog. Photovolt*. **7**, 21–30 (1999).
106. Takamoto T, Agui T, Kurita H, Ohmori M, *Sol. Energy Mater. Sol. Cells* **49**, 219–225 (1997).
107. Tyan Y, Perez-Albuerne E, Conf. Rec. 16^{th} IEEE Photovoltaic Specialist Conf., 794–800 (1982).
108. Chu T *et al., IEEE Electron. Dev. Lett*. **13**, 303–304 (1992).
109. Ferekides C *et al., Thin Solid Films* **361–362**, 520–526 (2000).
110. Wu X *et al., Conf. Rec. 28th IEEE Photovoltaic Specialist Conf*., pp 470–474 (2000).
111. Ohyama H *et al., Conf. Rec. 26th IEEE Photovoltaic Specialist Conf*., pp 343–346 (1997).
112. Bonnet D, Richter H, Jaeger K, *Conf. Rec. European 13th Photovoltaic Solar Energy Conversion*, pp 1456–1461 (1995).
113. Bonnet D, *Conf. Rec. 14th European Photovoltaic Solar Energy Conversion*, pp 2688–2693 (1997).
114. Barth K, *Conf. Rec. 34th IEEE Photovoltaic Specialists Conf*. pp 3–8 (2009).
115. Seymour F, *Proc. Mat. Rec. Soc*. pp 255–261 (2009).
116. McCandless B E, Buchanan W A, Birkmire R W, *Conf. Rec. 31st IEEE Photovoltaic Specialist Conf*., pp 295–298 (2005).
117. McCandless B, Buchanan W, *Conf. Rec. 33rd IEEE Photovoltaic Specialist Conf*. Pp 295–298 (2008).
118. Powell R *et al., U.S. Patent* 5,945,163 (1999).
119. Wendt R, Fischer A, Grecu D, Compaan A, *J. Appl. Phys*. **84**, 2920–2925 (1998).
120. Abou-Elfoutouh F, Coutts T, *Int. J. Sol. Energy* **12**, 223–232 (1992).
121. Kröger F, *J. Electrochem. Soc*. **125**, 2028–2034 (1978).
122. Başol B, *J. Appl. Phys*. **55**, 601–603 (1984).
123. Fulop G *et al., Appl. Phys. Lett*. **40**, 327–328 (1982).
124. Bhattacharya R, Rajeshwar K, *J. Electrochem. Soc*. **131**, 2032–2041 (1984).
125. Chu T, Chu S, *Int. J. Sol. Energy* **12**, 122–132 (1992).
126. Sudharsanan R, Rohatgi A, *Sol. Cells* **31**, 143–150 (1991).
127. Jordan J, *International Patent Application* WO93/14524 (1993).
128. Kester J *et al., AIP Conf. Ser*. **394**, 196 (1996).
129. Ikegami S, *Tech. Digest Int'l PVSEC-3*, pp 677–682 (1987).
130. Kim H, Im H, Moon J, *Thin Solid Films* **214**, 207–212 (1992).

131. Clemminck I, Burgelman M, Casteleyn M, Depuydt B, *Int. J. Sol. Energy* **12**, 67–78 (1992).
132. Gur I, Fromer N A, Geier M L, Alivisatos A P, *Science*, **310**, 462–465 (2005).
133. Bauer G, von Roedern B, *Conf. Rec. 16th European Photovoltaic Solar Energy Conversion*, pp 173–176 (2000).
134. Jordan J, Albright S, *U.S. Patent* 5,279,678 (1994).
135. McCandless B, Birkmire R, *Conf. Rec. 28th IEEE Photovoltaic Specialist Conf*., pp 491–494 (2000).
136. Takamoto T, Agui T, Kurita H, Ohmori M, *Sol. Energy Mater. Sol. Cells* **49**, 219–225 (1997).
137. Wu X *et al., Conf. Rec. 28th IEEE Photovoltaic Specialist Conf*., pp 470–474 (2000).
138. Romeo N, Bosio A, Tedeschi R, Canvari V, *Conf. Rec. 2nd WCPEC* (Vienna) pp 446–447 (1999).
139. McCandless B, Hegedus S, *Conf. Rec. 22nd IEEE Photovoltaic Specialist Conf*., pp 967–972 (1991).
140. Clemminck I *et al., Conf. Rec. 22nd IEEE Photovoltaic Specialist Conf*., pp 1114 (1991).
141. Wu X, *J. Appl. Phys*. **89**, 4564–4569 (2001).
142. Meyers P, Leng C, Frey T, *U.S. Patent* 4,710,589 (1987).
143. McCandless B, Birkmire R, *Sol. Cells* **31**, 527–535 (1990).
144. Barth K, Enzenroth R, Sampath W, *U.S. Patent* 6,423,565 (2002).
145. McCandless B, Hichri H, Hanket G, Birkmire R, *Conf. Rec. 25th IEEE Photovoltaic Specialist Conf*., pp 781–785 (1996).
146. Mahathongdy Y, Albin D, Wolden C, Baldwin R, *Conf. Rec. 15th NREL PV Rev. Meeting* pp 231–241 (1998).
147. McCandless B, *U.S. Patent* 6,251,701 (2001).
148. McCandless B E, Buchanan W A, Hanket G M, *Conf. Rec. 4th WCPEC* (New Orleans), pp 483–486 (2006).
149. Zhou T *et al., Conf. Rec. 1st WCPVSEC*, pp 103–106 (1994).
150. Qu Y, Meyers P, McCandless B, *Conf. Rec. 25th IEEE Photovoltaic Specialist Conf*., pp 1013–1016 (1996).
151. Başol B, Tseng E, Lo D, *U.S. Patent* 4,548,681 (1984).
152. McCandless B and Dobson K, *Solar Energy* **77**, 839 (2004).
153. McCandless B, Moulton L, Birkmire R, *Prog. Photovolt*. **5**, 249–260 (1997).
154. Birkmire R, McCandless B, Shafarman W, *Sol. Cells* **23**, 115–126 (1985).
155. Marfaing Y, *Thin Solid Films*, **387**, 123–128 (2001).
156. Zhang S, Wei S-H and Zunger A, *J. Appl. Phys*. **83** (6), 3192–3196 (1998).
157. Birkmire R, McCandless B, Hegedus S, *Int. J. Sol. Energy* **12**, 145–154 (1992).
158. Hiltner J, Sites J, *Mat. Res. Soc. Proc*. **668**, H 9.8, 1–7 (2001).
159. Jensen G, *PhD Dissertation*, Stanford University, Department of Physics (1997).
160. Ohata K, Saraie J, Tanaka T, *Jpn. J. Appl. Phys*. **12**, 1641–1642 (1973).
161. Moon D, Im H, *Powder Metall*. **35**, 53–58 (1992).
162. Jensen D, McCandless B, Birkmire R, *Mat. Res. Soc. Symp. Proc*. **428**, 325–330 (1996).
163. Hill R, Richardson D, *Thin Solid Films* **18**, 25–28 (1973).
164. Compaan A *et al., Mat. Res. Soc. Symp. Proc*. **428**, 367–371 (1996).
165. Bonnet D, *Phys. Stat. Sol*. **A3**, 913–919 (1970).
166. McCandless B, Hanket G, Jensen D, Birkmire R, *J. Vac. Sci. Technol. A* **20**(4), 1462–1467 (2002).
167. Wei S, Zhang S, Zunger A, *J. Appl. Phys*. **87**, 1304–1311 (2000).
168. Ohata K, Saraie J, Tanaka T, *Jpn. J. Appl. Phys*. **12**, 1198–1204 (1973).
169. Herndon M, Gupta A, Kaydanov V, Collins R, *Appl. Phys. Lett*. **75**, 3503–3506 (1999).
170. McCandless B, Engelmann M, Birkmire R, *J. Appl. Phys*. **89**(2), 988–994. (2001).
171. Collins R, *Conf. Rec. 4th WCPEC*, pp 392–395 (2006).

172. Albright S *et al., AIP Conf. Ser*. **268**, 17–32 (1992).
173. McCandless B, Phillips J, Titus J, *Conf. Rec. 2nd WCPVEC*, pp 448–452 (1998).
174. Gessert T, Duda A, Asher S, Narayanswamy C, Rose D, *Conf. Rec. 28th IEEE Photovoltaic Specialiss Conf*., pp 654–657 (2000).
175. McCandless B, Qu Y and Birkmire R, *Conf. Rec. 1st WCPVSEC* (Hawaii), pp 107–110 (1994).
176. Wu X, *et al*., *Prog. in Photovolt: Res. and Appl*., **14**, 471–483 (2006).
177. McCandless B, Qu Y, Birkmire R, *Conf. Rec 1st WCPVSEC*, pp 107–110 (1994).
178. Szabo L, Biter W, *U. S. Patent* 4,735,662 (1988).
179. Albright S, Ackerman B, Jordan J, *IEEE Trans. Elec. Dev*. **37**, 434–437 (1990).
180. Matsumoto H *et al., Sol. Cells* **11**, 367–373 (1984).
181. Lyubormisky I, Rabinal M, Cahen D, *J. Appl. Phys*., **81**, 6684–6691 (1997).
182. Ringel S, Smith A, MacDougal M, Rohatgi A, *J. Appl. Phys*. **70**, 881–889 (1991).
183. Drayton J *et al., Presented at Spring MRS* (San Francisco, April 2001).
184. Romeo N *et al., Sol. Energy Mater. Sol. Cells* **58**, 209–218 (1999).
185. McMahon T, Fahrenbruch A, *Conf. Rec. 28th IEEE Photovoltaic Specialist Conf*., pp 539–542 (2001).
186. Pan J, Gloeckler M and Sites J, *J. Appl. Phys*. **100**, 125405 (2006).
187. Sites J, Mauk P, *Sol. Cells* **27**, 411–417 (1989).
188. Kurtz S, Olson J, Kibler A, *Conf. Rec 23rd IEEE Photovoltaic Specialist Conf*., pp 138–141 (1990).
189. Metzger W, Albin D, Levi D, Sheldon P, Li X, Keyes B, Ahrenkiel R, *J. Appl. Phys*. **94**, 3549–3555 (2003).
190. Stollwerck G, Sites J, *Conf. Rec. 13th European Photovoltaic Solar Energy Conversion*, pp 2020–2022 (1995).
191. Niemegeers A, Burgelman M, *J. Appl. Phys*. **81**, 2881–2886 (1997).
192. McCandless B, Phillips J, Titus J, *Conf. Rec. 2nd WCPVSEC*, pp 448–452 (1998).
193. Asher S *et al., Conf. Rec. 28th IEEE Photovoltaic Specialist Conf*., pp 479–482 (2000).
194. Fahrenbruch A, *Sol. Cells* **21**, 399–412 (1987).
195. Meyers P, Phillips J, *Conf. Rec. 25th IEEE Photovoltaic Specialist Conf*., pp 789–792 (1996).
196. Hiltner J, Sites J, *AIP Conf. Ser*. 462, 170–173 (1998).
197. Gupta A, Townsend S, Kaydanov V, Ohno T, *Conf. Rec. NCPV Rev. Mtg*, pp 271–272 (2000).
198. Hegedus S, McCandless B, Birkmire R, *Conf. Rec. 28th IEEE Photovoltaic Specialist Conf*., pp 535–538 (2000).
199. Hegedus S and McCandless B, *Sol. Energy Mater. Sol. Cells* **88**, 75–95 (2005).
200. Dobson K, Visoly-Fisher I, Hodes G, Cahen D, *Sol. Energy Mater. Sol. Cells* **62**, 145–154 (2000).
201. Greco D, Compaan A, *Appl. Phys. Lett*. **75**, 36–363 (1999).
202. Stollwerck G, *MS Thesis*, Colorado State University (1995).
203. Hulstrom R, Bird R, Riordan C, *Sol. Cells* **15**, 365 (1985).
204. Mauk P, Tavakolian H, Sites J, *IEEE Trans. Electron Dev*. **37**, 1065–1068 (1990).
205. Liu X, Sites J, *J. Appl. Phys*. **75**, 577–581 (1994).
206. Hegedus S, Desai D and Thompson C, *Prog. Photovolt: Res. Appl*. **15**, 587–592 (2007).
207. Scofield J, *Sol. Energy Mater. Sol. Cells* **37**, 217–233 (1995).
208. Matulionis I, Nakada S, Compaan A, *Conf. Rec. 26th IEEE Photovoltaic Specialists Conf*., pp 491–494 (1997).
209. Willing F *et al., Conf. Rec. 21st IEEE Photovoltaic Specialists Conf*., pp 1432–1436 (1990).
210. van den Berg R *et al., Sol. Energy Mater. Sol. Cells* **31**, 253–261 (1993).

211. *Kirk-Othmer Encyclopedia of Chemical Technology*, 3rd edn, Vol. 11, pp 807–880, John Wiley & Sons, Inc., New York (1980).
212. Brown R, *U.S. Geological Survey Minerals Yearbook*, U.S.G.S., 67.1–67.4 (2000).
213. Andersson B, *Prog. Photovolt. Res. Appl*. **8**, 61–76 (2000).
214. Gerhardinger P, McCurdy R, *Mat. Res. Soc. Symp. Proc*. **426**, 399–410 (1996).
215. Zweibel K, *Sol. Energy Mater. Sol. Cells* **59**, 1–18 (1999).
216. Maycock P (ed.), *Photovoltaic News*, **20** (Feb, 2001).
217. Rose D *et al., Conf. Rec. 28th IEEE Photovoltaic Specialist Conf*., pp 428–431 (2000).
218. Sites J, Pan J, *Thin Solid Films* **515**, 6099–6102 (2007).
219. Plotnikov V, Kwon D, Wieland K, Compaan A, *Conf. Rec. 34th IEEE Photovoltaic Specialists Conf*. pp 1435–1438 (Philadelphia, 2009).
220. Coutts T *et al., Prog. Photovolt: Res. Appl*. **11**, 359–375 (2003).
221. Fan J, Palm B, *Sol. Cells* **11**, 247–261 (1984).
222. Nell M, Barnett A, *IEEE Trans. Elec. Dev*. **ED-34**, 257–265 (1987).
223. Basol B, Kapur V, Kullberg R, *Conf. Rec. 20th IEEE Photovoltaic Specialist Conf*., 1500–1504 (1988).
224. Rohatgi A, Sudharsanan R, Ringel S, MacDougal M, *Sol. Cells* **30**, 109–122 (1991).
225. McCandless B, Buchanan W, Hanket G, *Conf. Rec. 4th WCPEC*, 483–486 (2006).
226. Fang F, McCandless B, Opila R, *Conf. Rec. 4th IEEE Photovoltaic Specialist Conf*., pp 547–550 (2009).
227. Dhere R, Ramanathan K, Scharf J, *et al., Mat. Res. Soc. Symp. Proc*. **1012**, Y02–02 (2007).
228. Mathew X, Drayton J, Parikh V, *et al., Semicon. Sci. Technol*. **24**, 015012 (9 pp) (2009).
229. Romeo A, Batzner D, Zogg H, Tiwari A, *Mat. Res. Soc. Symp. Proc*. **668**, H3.3.1–H3.3.6 (2001).
230. Zweibel K, *Conf. Rec. IECEC* (Denver, CO, 1988).
231. Fthenakis V, Kim H, Alsema E, *Environ. Sci. Technol*. **42**, 2168–2174 (2008).
232. Fthenakis V, Eberspacher C, Moskowitz P, *Prog. Photovolt.: Res. Appl*. **4**, 447–456 (1996).
233. Zweibel K, Moskowitz P, Fthenakis V, *NREL Technical Report 520–24057* (Feb. 1998).

15

Dye-sensitized Solar Cells

Kohjiro Hara[1] and Shogo Mori[2]

[1]*National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan,* [2]*Faculty of Textile Science and Technology, Shinshu University, Ueda, Japan*

15.1 INTRODUCTION

Photoelectrochemical solar cells (PSCs), consisting of a photoelectrode, a redox electrolyte, and a counter electrode, have been studied extensively. Several semiconductor materials, including single-crystal and poly-crystal of n- and p-Si, n- and p-GaAs, n- and p-InP, and n-CdS, have been used as photoelectrodes. These materials when used with a suitable redox electrolyte can produce solar light-to-electric power conversion efficiency of more than 10% [1]. However, under irradiation, photocorrosion of the electrode in the electrolyte solution frequently occurs, resulting in poor stability of the cell, so efforts have been made worldwide to develop more stable PSCs.

Metal oxide semiconductors having wide-band-gap are stable under irradiation in solution. However, these materials cannot absorb visible light. Sensitization of wide-band-gap semiconductors, such as TiO_2, ZnO, and SnO_2, by materials absorbing visible light, has been extensively studied in relation to the development of photography technology since the late nineteenth century. In the sensitization process, sensitizers adsorbed onto the semiconductor surface absorb visible light and excited electrons are injected into the conduction band of the semiconductor electrodes. Dye-sensitized oxide semiconductor photoelectrodes have been used for PSCs. Gerischer and Tributsch studied a ZnO electrode sensitized by organic dyes including rose bengal, fluorescein, and rhodamine B [2]. In early studies, however, single-crystal and poly-crystal materials, which cannot adsorb a large amount of dye, were used for the photoelectrode, which resulted in low light harvesting efficiency and, consequently, low photon-to-current conversion efficiencies. Additionally, the examined organic dyes have narrow absorption ranges in visible light, which also contributed to low solar cell performance. Thus, to improve light-harvesting efficiencies and solar-cell performance, researchers used two approaches: developing photoelectrodes with larger surface areas that can adsorb large amount of dye; and synthesizing dyes with broader absorption ranges. Along

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

the direction of large surface area, Tsubomura *et al*. employed a porous ZnO electrode, which can adsorb large amount of dyes, with an iodine redox electrolyte, improving the energy conversion efficiency in comparison to electrode having a flat surface [3]. However, the surface area was still not large enough to adsorb sufficient amount of sensitizers.

In 1991, O'Regan and Grätzel reported 7% efficiency of DSSCs by expanding the surface area by utilizing nanosized TiO_2 and by employing new Ru-complex sensitizers capable of absorbing in the wide visible and near-IR region from 400 to 800 nm [4]. Since then, by optimizing the structure of nanoporous electrode, that of Ru-complex dyes, and composition of electrolytes, the efficiency has been improved to more than 11% [5, 6].

Due to the high efficiency, the DSSC is an attractive and promising unconventional solar cell that has been intensively investigated [7–10]. Research has been conducted towards elucidation of working principles, improvement of conversion efficiency, and commercialization. The cost of commercially fabricating DSSCs is expected to be relatively low because the cells are made of low-cost materials and assembly is simple. In this chapter, we describe the DSSC including its structure, operating mechanism, component materials, characteristics, and long-term stability, and discuss improvement of its performance and commercialization.

15.2 OPERATING MECHANISM OF DSSC

Figure 15.1 shows a schematic of a DSSC and presents the operating mechanism of electric power generation in the DSSC. The main components of a DSSC are a nanocrystalline oxide semiconductor thin film electrode (e.g., TiO_2 and ZnO) formed onto a transparent conducting oxide (TCO)-coated electrode (e.g., F-SnO_2/glass), a sensitizer, an electrolyte containing redox ions, and a counter electrode. First, a sensitizer molecule adsorbed on the surface of a nanocrystalline TiO_2 electrode absorbs the incident photons and is excited from the ground state (S) to the excited state (S*). One type of photoexcitation, causes transfer of an electron from the highest occupied molecular orbital (HOMO) of the sensitizer to the lowest unoccupied molecular orbital (LUMO). Subsequent injection of the excited electron into the conduction band of the TiO_2 electrode results in oxidation of the sensitizer molecule (Equation 15.2).

$$S + hv \rightarrow S^* \tag{15.1}$$

$$S^* \rightarrow S^+ + e^-(TiO_2) \tag{15.2}$$

The injection can occur not only from singlet, but also from triplet states if their energy levels are sufficiently high. Then, the injected electron diffuses through the TiO_2 electrode toward the transparent conducting oxide (TCO)-coated electrode. The oxidized sensitizer is reduced by I^- ions in the electrolyte, regenerating the ground state of sensitizer, and the I^- ions are oxidized to I_3^- ions. The I_3^- ions diffuse toward the counter-electrode and are then reduced to I^- ions. Note that these reduction processes in detail have not been proven. A recent paper showed that an iodine atom was formed by oxidation of I^-, and subsequently the atom reacts with I^-, forming I_2^- [11]. Overall, electric power is generated without permanent chemical transformation.

In conventional p–n junction-type solar cells and classical PSCs using poly or single crystal photoelectrodes, electronic contact between the components that form the photovoltaically active junction, and equilibrium between the electronic charge carriers in them, leads to space charge formation. Photogenerated charges are separated by the electric field in the space-charge layer. In DSSC, however, the individual particle size, which is typically around 20 nm, is too small to form a space-charge layer. In addition, cations as the counter charges of I^-/I_3^- redox couples in the electrolytes screen the electrons in the electrode, resulting in no potential gradient in the

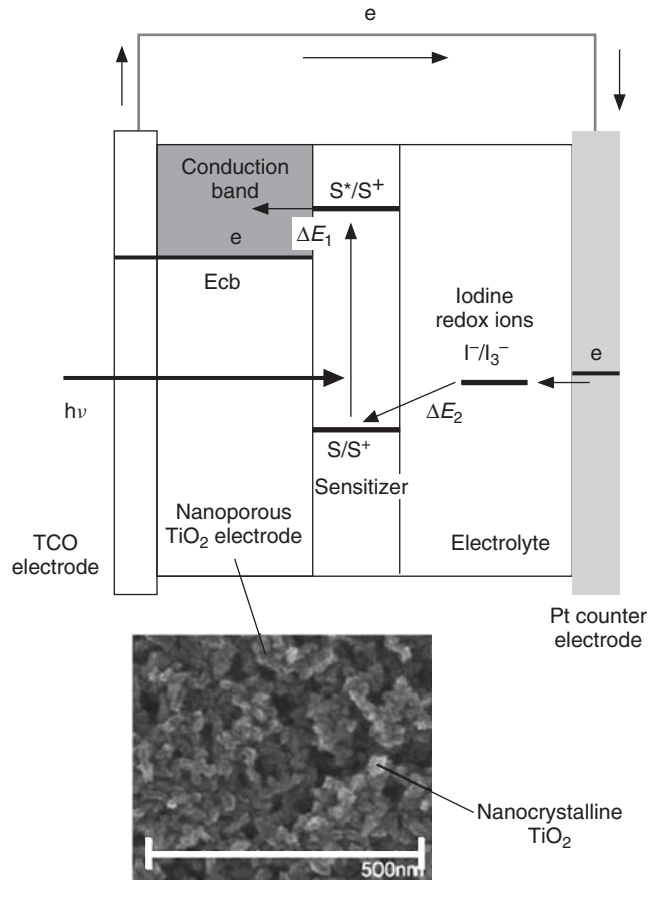


Figure 15.1 Schematic structure of nanocrystalline dye-sensitized solar cell (DSSC) presenting the mechanism of electric power generation and scanning electron microscope (SEM) photograph of a typical nanocrystalline TiO_2 electrode

electrode. Therefore, initial charge separation occurs just through the injection of electrons into semiconductors and subsequent injection of holes into electrolytes from the adsorbed sensitizers.

Since the photoelectrons and holes are not separated by electric field, like in a conventional $p-n$ junction-type solar cell, these are located close to each other but in different media. Thus, charge recombination in DSSCs occurs through interfacial charge transfer at semiconductor/dye/electrolyte interface. Effective charge separation is achieved when the recombination process is slow. This has been realized by using I^-/I_3^- redox couple. This charge separation process in DSSCs is similar to the mechanism for photosynthesis in nature, where chlorophyll functions as the sensitizer and charge transport occurs in the membrane.

The solar energy-to-electric power conversion efficiency, $\eta(\%)$, of solar cells, including DSSCs, is defined from the following equation:

$$\eta(\%) = \frac{J_{sc\times} V_{oc} \times FF}{I_0} \times 100 \tag{15.3}$$

where I_0 is the photon flux (ca. 100 mW cm^{-2} for AM 1.5 G), J_{sc} is the short-circuit current density under irradiation, V_{oc} is the open-circuit voltage, and FF is the fill factor. Basically, the V_{oc} value is determined by the energy gap between the Fermi level of the TiO_2 electrode, which is located near the conduction band edge potential (E_{cb}), and the redox potential of the I^-/I_3^- in the electrolyte (Figure 15.1). The E_{cb} value of the TiO_2 electrode and the redox potential of I^-/I_3^- are estimated to be -0.5 and 0.4 V versus, the normal hydrogen electrode (NHE), respectively [8]. Thus, for DSSCs with a TiO_2 electrode and the I^-/I_3^- redox mediator, the maximum V_{oc} is expected to be approximately 0.9 V. The E_{cb} of the TiO_2 electrode and the redox potential of I^-/I_3^- depends strongly on the type and concentration of electrolyte components. The adsorption of Li^+ has been known to shift the E_{cb} positively [12], while that of basic compounds such as 4-*tert*-butylpyridine (TBP) in the electrolyte shifts it negatively [13, 14]. The Fermi level of the TiO_2 scales with the electron density, which then depends on the recombination rate to dye cation and to I_3^-. Currently optimized DSSCs shows V_{oc} ranging from 0.75 to 0.85 V.

The J_{sc} value is directly determined by the product of light-harvesting efficiency (LHE), charge injection efficiency ϕ_{inj}, and charge collection efficiency of the injected electrons at the back contact η_c. The LHE is given by

$$\mathrm{LHE} = 1 - \mathrm{T} = 1 - 10^{-\mathrm{A}} \tag{15.4}$$

where T is transmittance, and A is absorbance. The LHE is determined by the effective absorption coefficient of dye-sensitized electrodes, which depend on the concentration of adsorbed dye, extinction coefficient of the sensitizers, and thickness of TiO_2 electrode. The energy gap between the HOMO and LUMO of the sensitizer (which corresponds to the bandgap E_g, for inorganic semiconductor materials) directly determines the photoresponse range of the DSSCs. In order to utilize wider spectrum of sun light, a small HOMO–LUMO energy gap is necessary, producing a large J_{sc}. In order to have high ϕ_{inj}, the LUMO must be sufficiently more negative than E_{cb} as shown in Figure 15.1; the energy gap between the two levels, ΔE_1, which can be regarded as the driving force for electron injection. After the injection, the resulting dye-cation must be regenerated by oxidizing I^-. In order to have fast electron transfer, the HOMO must be sufficiently more positive than the redox potential of I^-/I_3^- to accept electrons effectively; the difference between these two levels is given by ΔE_2. For the case of optimized DSSCs with TiO_2 and I^-/I_3^-, the ΔE_1 and ΔE_2 are approximately 0.2 and 0.5 eV, respectively [15, 16].

In order to have high LHE, incident photons must be absorbed by sensitizers. When porous electrode is prepared from TiO_2 nanoparticles (10–30 nm), the actual surface area of the TiO_2 electrode compared to its apparent surface area, roughness factor (rf), is >1000; that is, a 1 cm^2 TiO_2 film (10 μm thickness) has an actual surface area of 1000 cm^2. The dye is considered to be adsorbed on the TiO_2 surface in a monolayer. Thus, if the nanoporous TiO_2 film has a high rf, the amount of dye adsorbed is drastically increased (by an amount of the order of 10^{-7} mol cm^{-2}), resulting in nearly 100% absorption at the peak wavelength of the dye's absorption spectrum.

Charge collection efficiency is determined by electron diffusion length L, which is expressed by

$$L = \sqrt{D \cdot \tau} \tag{15.5}$$

where D is electron diffusion coefficient in nanoporous semiconductor electrodes and τ is electron lifetime in the electrode determined by electrode/dye/electrolyte interfacial recombination. In order to collect the all injected electrons, L must by larger than the thickness of dye-sensitized electrodes. For liquid-based DSSCs, L can be tens of micrometers, consisting of nearly 100% η_c for optimized DSSCs. However, typical solid state DSSCs have shorter L, requiring sensitizers having high absorption coefficients.

The *FF* is partially determined by the shunt and series resistances of the solar cell. The shunt resistance is not only related with the interfacial charge recombination but also, for some cases, with charge recombination at TCO/electrolyte interface. The series resistance can be reduced, for example, by reducing the electrolyte thickness and by the increase of Pt counter electrode surface area [17]. The FF seems also to be related to potential-dependent electron transfer/transport rate under light irradiation [18].

15.3 MATERIALS

15.3.1 TCO Electrode

Generally, TCO-coated glass is used as the substrate for the TiO_2 photoelectrode. For high solar cell performance, the TCO substrate must have low sheet resistance and high transparency. In addition, sheet resistance should be nearly independent of temperature up to 500 °C because sintering of the TiO_2 electrode is generally carried out at 450–550 °C. Indium–tin oxide (ITO) is one of most famous TCO materials. However, ITO has a low thermal stability of resistance, but it nevertheless has low resistance at ambient temperature. Usually, fluorine-doped SnO_2 is used as the TCO substrate for DSSCs (e.g. Nippon Sheet Glass Co., or Asahi Glass Co. Ltd., $R = 8 - 10\Omega/\square$). If TiO_2 electrode can be prepared at relatively low temperature below 200 °C, ITO-coated poly(ethylene terefutarate) (PET) or poly(ethylene naphthalate) (PEN) can be used as the substrate for plastic DSSCs.

15.3.2 Nanocrystalline TiO_2 Photoelectrode

15.3.2.1 TiO_2 nanoparticle

The primary role of porous electrode is to provide sufficient surface area for dye adsorption, and to convey all injected electrons to TCO. Thus, electrodes should be transparent for visible and infrared regions. The required surface area and thickness of the electrode are determined by the absorption coefficients of the sensitizers, but the thickness is limited by the electron diffusion length of the electrodes. For typical dyes, optimized cells results in 10–20 μm thickness of porous electrodes consisting of around 20 nm particles. The electrodes can be fabricated by commercial nanocrystalline TiO_2, such as P25 (Degussa) and ST-21 (Ishihara Sango kaisha Ltd.). To produce high-performance DSSCs, colloidal TiO_2 prepared by hydrolysis of Ti(IV) alkoxides, such as isopropoxide and butoxide, has usually been used. To obtain monodispersed particles of the desired size, the hydrolysis and condensation kinetics must be controlled. Titanium-alkoxides suitably modified with acetic acid or acetyl acetonate, yield colloids having a large surface area ($>200\,m^2\,g^{-1}$) and smaller particle diameter (5–7 nm) [9, 19]. Peptization results in segregation of the agglomerates to primary particles, after which the large agglomerates are removed by filtration. Autoclaving of the colloidal TiO_2 solution leads to growth of the primary particles to 10–25 nm and also to some extent increases the anatase crystallinity present. At higher autoclaving temperature, more growth of particles and rutile formation occur, particularly at temperatures above 240 °C. Generally, anatase rather than rutile TiO_2 is more suitable for the electrodes [20]. TiO_2 electrodes prepared from small nanoparticles (10–25 nm) are transparent. In addition, a layer consisting of large nanocrystalline TiO_2 particles (250–300 nm), which can scatter incident photons effectively, has been also applied on the top of the transparent layer to improve the light harvesting efficiency, as shown in next section.

15.3.2.2 Preparation of the TiO_2 electrode

The TiO_2 thin-film photoelectrode is prepared by a very simple process. TiO_2 colloidal solution (or paste) is coated on a TCO substrate (e.g., screen printing) and then sintered at 450–550 °C, producing a TiO_2 film around 10 μm in thickness. The film thickness can be controlled in screen printing by selection of past composition (i.e. wt % of TiO_2 nanoparticles in the paste), screen mesh size, and repetition of printing. In addition, one can increase thickness by multiple printing. The porosity of the film is also important because the electrolyte, which contains the redox ions, must be able to penetrate the film effectively to allow redox ions reaching all adsorbed dyes and diffusing back to the counter electrode. Appropriate porosity, 50–70%, is controlled in the sintering process by addition of a polymer such as polyethylene glycol (PEG) and ethyl cellulose (EC) into the TiO_2 colloidal solution or paste [21]. A scanning electron microscope (SEM) photograph of a typical nanocrystalline TiO_2 film is shown in Figure 15.1. Because the film is composed of TiO_2 nanoparticles and has nanoporous structure, the actual surface area of TiO_2 film compared with its apparent surface area, roughness factor (rf), is >1000; that is, a 1 cm^2 TiO_2 film (10 μm thickness) has an actual surface area of 1000 cm^2.

In order to absorb long wavelength light, where dyes have low absorption coefficient, thick TiO_2 is required. However, the practical thickness is limited by electron diffusion length. In order to trap such light in relatively thin TiO_2, a scattering layer on the top of the transparent TiO_2 film has been introduced. The scattering property of the film is important for improvement of the light-harvesting efficiency of the dye-sensitized film to increase incident photon-to-current conversion efficiency (IPCE). Optical enhancement due to the scattering in the TiO_2 film has been investigated in detail [9, 19, 22–24]. The path-length of the incident light and therefore the absorption due to the adsorbed dye can be controlled by the size of particles. A simulation of light scattering in the TiO_2 electrode of DSSC predict that a suitable mixture of small TiO_2 particles (e.g. 20 nm diameter) and of larger particles (250–300 nm diameter), which are effective light scattering centers, have the potential to enhance solar light absorption significantly [23]. The photocurrent of a DSSC can be increased by using a scattering film, compared with that for a transparent film [24]. Recent typical structure consists of 8 μm of a transparent layer and 4 μm of a scattering layer.

The improvement in photoresponse of DSSC due to the scattering effect is observed especially in the low-energy region (e.g. 650–900 nm) where the incident radiation penetrates the layer due to low absorption coefficient of the dye (Figure 15.3). Photons of 500–650 nm can be mainly absorbed at near the TCO/TiO_2 interface because of large absorption coefficient and therefore do not benefit from optical enhancement.

It has also been reported that $TiCl_4$ treatment of the film improves cell performance, especially the photocurrent [7]. After printing, the TiO_2 films are immersed in a 0.1–0.5 M $TiCl_4$ solutions and then sintered at 450 °C for 30 min. It has been reported that $TiCl_4$ treatment improves electron diffusion coefficients and lifetime, while the E_{cb} of the TiO_2 is positively shifted [25].

15.3.3 Ru-complex Photosensitizer

The Ru-complex sensitizer, which contributes the primary steps of photon absorption and the consequent electron injection, is adsorbed onto the TiO_2 surface. The chemical structure of typical Ru-complex sensitizers developed by Grätzel and co-workers are shown in Figure 15.2 while Figure 15.3 shows absorption properties of the complexes in solution. The *Y* axis is represented by absorbance and 1−T. N3 dye, *cis*-bis(4,4′-dicarboxy-2,2′-bipyridine)dithiocyanato ruthenium(II) [7], can absorb over a wide range of the visible region from 400 to 800 nm. Black dye, trithiocyanato 4,4′4″-tricarboxy-2,2′:6′,2″-terpyridine ruthenium(II) [26], absorbs in the near-IR region

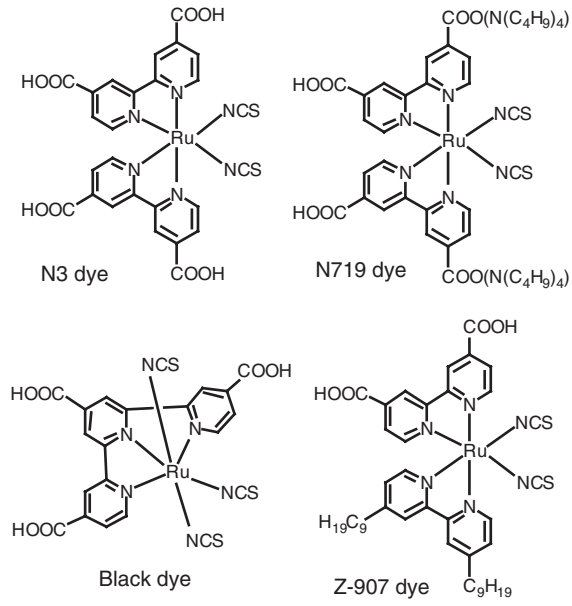


Figure 15.2 Molecular structures of typical Ru-complex photosensitizers for DSSC

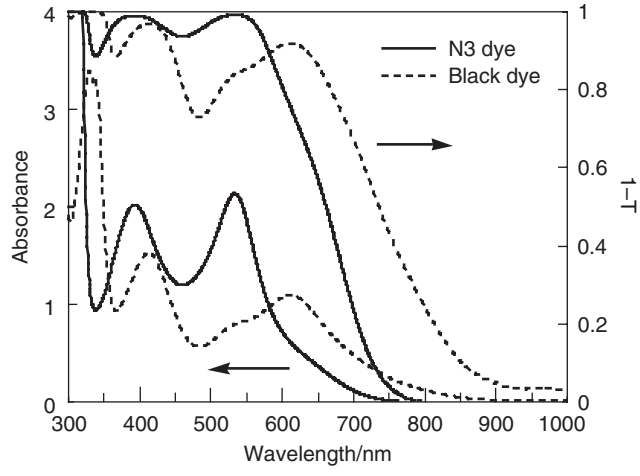


Figure 15.3 Absorption spectra of N3 dye and black dye represented by absorbance and light harvesting efficiency, 1–T (T: transmittance): (——) N3 dye, (- - -) black dye

up to 900 nm. Absorption by these dyes in the visible and near-IR regions is attributed to the metal-to-ligand charge-transfer (MLCT) transition. The HOMO and the LUMO are derived from the *d*-orbital of Ru metal and the π^* orbital of the ligand, respectively. The NCS ligand shifts the HOMO level negatively, leading to a red shift in the absorption property of the complex, and also contributes electron acceptance from iodide ions.

These Ru-complex sensitizers are dissolved at a concentration of 0.2–0.3 mM in ethanol or *tert*-butanol–acetonitrile, 1:1 mixed solution. The TiO_2 electrodes are immersed in the dye solution and then kept at 25 °C for more than 12 h to allow the dye to adsorb to the TiO_2 surface. Ru-complexes have carboxyl groups to anchor to the TiO_2 surface. Anchoring causes a large electronic interaction between the ligand and the conduction band of TiO_2, resulting in effective electron injection from the Ru-complex into the TiO_2. The Ru-complex is adsorbed on the TiO_2 surface via carboxylate bidentate coordination or ester bonding as measured by FT–IR absorption analysis [27–31]. The coverage of the TiO_2 surface with the N3 dye reaches near 100% as derived from the surface area of TiO_2 and the amount of the dye.

15.3.4 Redox Electrolyte

The electrolyte used in DSSCs contains I^-/I_3^- redox ions, which mediate electrons between the sensitizers and the counter electrode. Mixtures of iodides such as LiI, NaI, KI, tetraalkylammonium iodide (R_4NI), and imidazolium-derivative iodides with concentrations of 0.1–0.5 M (M: molar concentration) and 0.05–0.1 M I_2 dissolved in non-protonic organic solvents. Typical organic solvents used for DSSC are nitrile solvents having relative low viscosity, such as acetonitrile, propionitrile, methoxyacetonitrile, and 3-methoxypropionitrile. Viscosity of solvents directly affects ionic conductivity in the electrolyte, and consequently the solar-cell performance [32]. To improve solar-cell performance, especially Jsc, low-viscosity solvents are desired for high ionic conductivity. On the other hand, low-viscosity solvents typically have high vapor pressure, making it difficult to seal it for long term usage. The diffusion coefficient of I_3^- in methoxyacetonitrile is estimated as $5.4–6.2 \times 10^{-6}\,cm^2\,s^{-1}$ [12].

Solar cell performance of DSSCs also depends on counter-cations of iodides such as Li^+, Na^+, K^+, and R_4N^+ due to different ion conductivity in the electrolyte and adsorption property on the TiO_2 surface. The adsorption of cations shifts the conduction band level of the TiO_2 electrode positively [12, 33]. Basic compounds such as 4-*tert*-butylpyridine (TBP) have been added to the electrolyte solution, where TBP shifts the conduction band level of the TiO_2 electrode negatively, improving the open-circuit voltage [7, 13, 34]. Electrolytes also influence the charge recombination rate probably by changing the thickness of the double layer [13] and the free energy difference between the electrons in the electrode and I_3^- ions [35]. A typical electrolyte composition that produces high solar-cell performance for the Ru-complex sensitizers reported by Grätzel's group is a mixture of 0.5 M 1,2-dimethyl-3-hexylimidazolium iodide (DMHImI), 0.04 M LiI, 0.02 M I_2, and 0.5 M TBP in acetonitrile [19]. It has been reported that imidazolium derivatives, such as DMHImI and 1,2-dimethyl-3-propylimidazolium iodide (DMPImI), decrease the resistance of the electrolyte solution and improve photovoltaic performance [36, 37]. Br^-/Br_2 and hydroquinone have also been used as redox electrolyte for DSSC [3, 33, 38], but the iodine redox electrolyte gives the best performance.

15.3.5 Counter-electrode

Triiodide ions, I_3^-, formed by reduction of dye cations with I^- ion, are rereduced to I^- ions at the counter electrode. To reduce the triiodide ions, the counter-electrode must have high electrocatalytic activity for the reaction. Sputtered Pt coated on a TCO substrate (5–10 μg cm^{-2} or approximately 200 nm thickness) has been usually employed as a counter-electrode. In addition, the electrocatalytic activity of the Pt-sputtered TCO electrode for reduction of triiodide ions can be improved by formation of Pt colloids on the surface [39]. Small amounts of an alcoholic solution of H_2PtCl_6 are dropped on the surface of the Pt-sputtered TCO substrate, followed by drying and heating at 385 °C for 10 minutes, resulting in formation of Pt colloids on the surface. The resistance of the

Pt counter-electrode and electrolyte interface directly affect fill factor of the solar cell. A desirable exchange current density corresponding to electrocatalytic activity for the reduction of triiodide ions is 0.01–0.2 A cm^{-2} [9, 39]. Carbon materials and polymer materials such as PEDOT can also be used as the counter electrode instead of Pt [40–42].

15.3.6 Sealing Materials

A sealing material is needed to prevent leakage of the electrolyte and evaporation of the solvent. Chemical and photochemical stability of the sealing material against the electrolyte component, iodine, and solvent is required. Surlyn (Du Pont), a co-polymer of ethylene and acrylic acid has been used for the purpose.

15.4 PERFORMANCE OF HIGHLY EFFICIENT DSSCs

Figure 15.4 shows the external spectral response curve of the photocurrent for nanocrystalline TiO_2 solar cells sensitized by N3 dye and black dye with an I^-/I_3^- redox electrolyte, where the IPCE corresponding to external quantum efficiency is represented as a function of wavelength. IPCE is obtained by the following equation:

$$\text{IPCE}(\%) = \frac{1240(\text{eV}\cdot\text{nm}) \times J_{\text{ph}}}{\lambda \times \Phi} \times 100 \tag{15.6}$$

where J_{ph} (mA cm^{-2}) is the short-circuit photocurrent density obtained under monochromatic irradiation, and λ (nm) and Φ (mW cm^{-2}) are the wavelength and the intensity, respectively, of the external monochromatic light. Due to the reflection and absorption by the transparent conducting glass substrate, the highest IPCE is typically below 90%. The IPCE is also given by

$$\text{IPCE} = \text{LHE} \times \phi_{\text{inj}} \times \eta_{\text{c}} \tag{15.7}$$

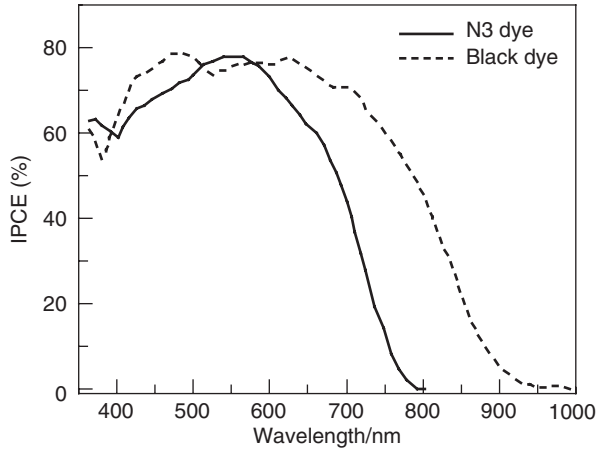


Figure 15.4 Spectral response curve of the photocurrent for the DSSC based on N3 and black dyes: (——) N3 dye, (- - -) black dye. The incident photon-to-current conversion efficiency (IPCE) is plotted as a function of wavelength

Table 15.1 Performance of DSSCs based on Ru-complex sensitizers

Dye	Area/cm^2	Jsc/mA cm^{-2}	V_{oc}/V	FF	η(%)	Reference
N719	0.16	17.7	0.85	0.75	11.2	[5]
Black dye	0.22	20.9	0.74	0.72	11.1	[6]
Ru terpyridyl	0.25	19.10	0.66	0.72	9.1	[45]
Z-910	0.16	17.2	0.78	0.76	10.2	[43]
K-73	0.16	17.2	0.75	0.69	9.5	[44]
CYC-B1	0.25	23.92	0.65	0.55	8.5	[47]

Light source: AM1.5 simulator, Electrode: nanocrystalline TiO_2 electrode, electrolyte: iodine redox in organic solvents.

For a nanocrystalline and nanoporous TiO_2 electrode (e.g. 10 μm thickness), which can adsorb a large amount of dye, LHE is almost equal to unity, as shown in Figure 15.3. Therefore, the IPCE value is predominantly determined by ϕ_{inj} and η_c.

As shown in Figure 15.4, DSSCs based on the Ru-complex photosensitizers can efficiently convert visible light to current. N3 dye responds to light from 400 to 800 nm, and black dye responds to the near-IR region up to 950 nm. The IPCE of the DSSC based on N3 dye reaches 80% at 550 nm and exceeds 70% in the region 400–650 nm. The DSSC based on black dye gives external IPCE values higher than 70% from 400 to 700 nm. Taking into consideration losses due to light reflection and absorption by the TCO substrate, the net photon-to-current conversion efficiency in this range exceeds 90%, which indicates a highly efficient DSSC with high ϕ_{inj} and η_c values.

High η values (>10%) have been attained with DSSCs based on Ru-complexes. Photovoltaic performance of DSSCs based on Ru-complex sensitizers are listed in Table 15.1. The high η values more than 11% were obtained with DSSCs, e.g. with N719 dye ($J_{sc} = 17.7\,\text{mA cm}^{-2}$, $V_{oc} = 0.846\,\text{V}$, $FF = 0.75$) [5], and with black dye ($J_{sc} = 20.9\,\text{mA cm}^{-2}$, $V_{oc} = 0.736\,\text{V}$, $FF = 0.72$) [6]. Additionally, high η values more than 9% have been reported [43–47].

15.5 ELECTRON-TRANSFER PROCESSES

15.5.1 Electron Injection from Dye to Metal Oxide

Electron-transfer processes in a DSSC based on N3 dye are schematically shown in Figure 15.5. Photoexcitation of a Ru-complex sensitizer results in an intramolecular MLCT transition. The HOMO and LUMO are derived from the d-orbital of Ru metal and the π^* orbital of the bipyridyl ligand, respectively [48]. The NCS ligands, which are strongly electron donating, shift the HOMO level negatively (thus red-shifting the absorption of the complex) and also accept electrons from iodide ions. After the photoexcitation, the excited electrons located in the bipyridyl ligands are efficiently injected into the conduction band of the TiO_2 electrode through the carboxyl groups anchored to the TiO_2 surface.

Electron injection processes from adsorbed dyes to the conduction band of metal oxides have been studied by transient absorption spectroscopy [49–54]. The observed time scale of optimized cells is femto- to picoseconds. The time constant of fluorescence decay, that is due to internal

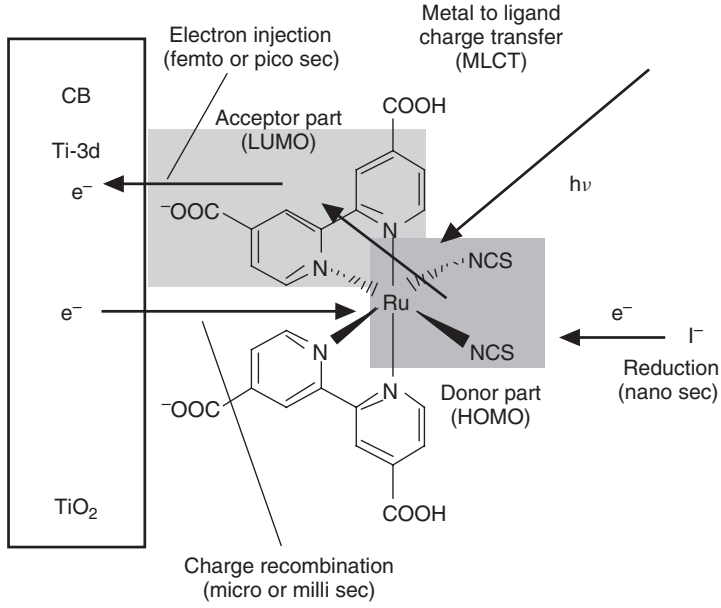


Figure 15.5 Schematic diagram of electron-transfer processes in DSSC. The dye is *cis*-dithiocyanato bis(4,4′-dicarboxy-2,2′-bipyridine)ruthenium(II) (N3 dye)

relaxation, is typically longer than a few nanoseconds [15], indicating high injection yield. The rate constant of the electron injection can be written as

$$k_{ET} = \frac{2\pi}{\hbar} J^2 \int_{-\infty}^{+\infty} \Psi_i(E)\Psi_f(E)\,\mathrm{d}E \tag{15.8}$$

where J is the transfer integral, $\Psi_i(E)$ and $\Psi_f(E)$ are the electron detachment and attachment spectrum respective [55]. In order to have sufficient overlap, the oxidation potential of excited state sensitizer is higher (more negative) than the E_{cb}, and the required difference has been observed to be larger than 0.2 V [15, 56]. The required ΔE_1 (Figure 15.1) was explained by assuming energetic heterogeniety of nanoporous metal oxide electrode surface [56]. For the case of ZnO, the injection efficiency is not dependent on the excitation wavelength, indicating the relaxation of the excited electrons in the dyes before the event of the injection [56], while direct injection from excited state of dye was seen with TiO_2 electrode [52]. Ru-complex dyes show fast intersystem crossing from the singlet to triplet excited state, and subsequent injection from the triplet excited state to the conduction band [52, 57].

Besides the potential difference, the spatial distance between the surface of metal oxides and LUMO of sensitizers also influences the transfer rate. Lian *et al*. [54] studied the influence of the distance on the injection kinetics by changing the non-conjugated methylene unit ($-CH_2-$) between carboxyl group and bipyridine moiety of Ru complex dyes. They found that the injection rate decreased with the increase of the distance. On the other hand, the decreased rate had little influence on the energy conversion efficiency of the solar cells. This was because the injection rate was already too fast in comparison with that of internal relaxation, and thus, the decrease of the rate did not affect on the injection yield. In other words, conjugation between the dye framework and linker sites is not the most important criterion [58], but the required distance for efficient electron injection seems to depend on the sensitizers [59].

The injection kinetics has been also studied for various metal oxides and organic dyes. For the case of ZnO, SnO_2, In_2O_3 with N3 dye, the injection yield can be unity [60]. On the other hand, the kinetics varied probably due to different density of states and electronic coupling [61]. For organic dyes having carboxyl group, the injection rates can be as fast as that of N3, consistent with high values of IPCE from the solar cells [62, 63].

15.5.2 Electron Transport in Nanoporous Electrode

Once electrons are injected into the conduction band, the electrons transport mostly by diffusion [64]. The values of diffusion coefficients D in nanoporous TiO_2 have been measured by various groups, mostly using current response against a small perturbation of incident light [65, 66]. The values obtained by analyzing the current response were increased from 10^{-8} to $10^{-4}\,cm^2/s$ with the increase of light intensity, while the value in crystal TiO_2 is of the order of $10^0\,cm^2/s$ [67]. The large difference and light intensity dependence have been modeled with intra-band charge traps, where electrons are captured and de-trapped during diffusion [68, 69]. Thus, the extracted values of D from current response are apparent values of D. By the model, with the increase of electron density, more traps are filled, and subsequent electrons would diffuse with less chance of trap capturing. The observed power-law dependence of the D was simulated by assuming exponential distribution of the traps [69]. In the model, trapped electrons are thermally de-trapped, and thus, a temperature-dependent D was expected. However, the dependence was not fully explained with the exponential trap distribution, suggesting some modification of the model is still needed [70].

Since the electrons in the nanoporous metal oxide are closely located to the electrolyte containing high concentration of cation, the electron transport is expected to occur by ambipolar diffusion [71]. In other words, the current of electrons is accompanied by the current of cations, and the observed current is limited by slower one. Thus, under low cation concentrations in electrolyte solution, electron diffusion is not limited by the trapping/de-trapping events, but also by the diffusion current of cations. The ambipolar diffusion coefficient is written by

$$D_{amb} = \frac{(n+p)}{(n/D_p)+(p/D_n)}$$

where n and p are the density of electrons and cations, respectively, and D_n and D_p are the diffusion coefficients of electrons and cations, respectively. Figure 15.6 shows the D with two different Li^+ concentrations [72]. With low $[Li^+]$, the apparent D was limited by the diffusion of Li^+ in the electrolyte, that is, the D did not increase with the increase of electron concentration in the TiO_2. The results suggest that electrolytes/hole transport layer should be also modified when electrodes having high electron diffusion coefficients are employed. The D_{amb} also depends on the species of electrolyte cations. This was correlated with the tendency of adsorption on the TiO_2 surface [73].

The origin and density of traps have not been elucidated yet. One can expect that they are mostly located at the surface of metal oxides due to high surface to volume ratio. To substantiate it, nanoporous TiO_2 electrodes were prepared from nanoparticles having various diameters [74, 75]. Above the diameter of 20 nm, the apparent D scales with the decrease of surface area, indicating large portion of the traps are on the surface. Under 20 nm, the D seems to be related with the number of boundary between particles, implying the traps are also located at grain boundaries.

Nanoporous electrodes can be prepared with and without annealing processes. With the increase of the temperature, the D was increased [76]. This can be interpreted with the increased contact area at inter-particle boundaries, but not with the elimination of grain boundaries [76]. The extreme case would be a nanowire where there is no boundary, which will be mentioned later. Interesting thing with the increase of the annealing temperature is that electron lifetime is also increased. This is also consistent with formation of traps at the boundary [77].

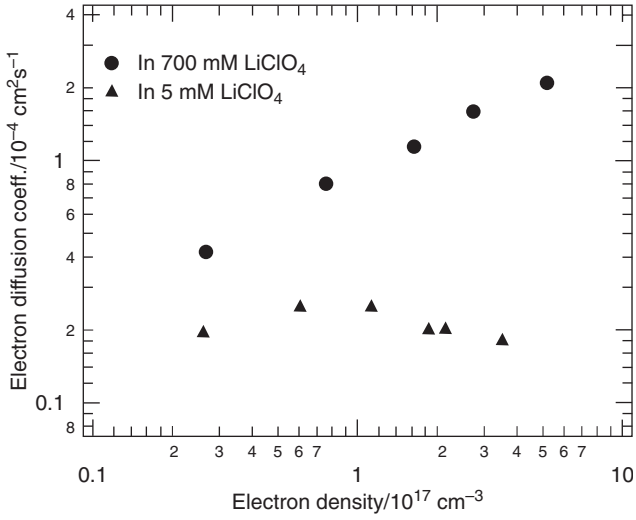


Figure 15.6 Electron diffusion coefficients in nanoporous TiO_2 electrode with different electrolyte compositions. Adapted from Figure 5 of reference [72]

15.5.3 Kinetic Competition of the Reduction of Dye Cation

The injected electrons in metal oxide can transfer to dye cation at the surface. This charge recombination process can be prevented by reducing the dye cation by I^- faster than the recombination. The transfer process from TiO_2 to the dye cation has been studied by transient absorption spectroscopy (Figure 15.7) [78]. When the transient time is characterized by a time giving the half of the initial optical density of dye cation, the half-time was decreased with the increase of electron density. The observation was modeled as trap-limited recombination [79], where the recombination rate is inversely related with transport rate [80]. Such correlation has been observed with various DSSCs [74]. Since the electrons in the nanoporous electrodes are trapped most of time, the rate is approximately related with the density of electrons in the conduction band [81]. The observed half-time varied between 10^{-3} and 10^{-9} s for the case of Ru-complex dyes. The half-time due to the reduction by I^- varies between 10^{-5} to 10^{-7} [82], which also depends on the species of counter-cations [83]. Thus, under optimized conditions, cationic state of Ru-complex dyes can be reduced by I^- faster than by the conduction band electrons. This is not always the case of various materials [84]. This will be addressed in the following sections.

The transfer rate from the conduction band to dye-cation is also influenced by the spatial distance between the surface of metal oxide and LUMO of the dyes, the free energy difference between them, and reorganization energy of the dyes [85], while other factors seem also to be involved [86]. For the case of I^-/I_3^- redox couple, the way of complex formation between dye^+ and I^- would be another important factor [87].

15.5.4 Charge Recombination between Electron and I_3^- Ion

Under optimized DSSCs, this process determines the electron lifetime in metal oxides. The electron lifetimes have been measured by various techniques such as photovoltage response [34], electrical impedance [88], and transient absorption measurements [82]. Similar to the transfer rate to dye

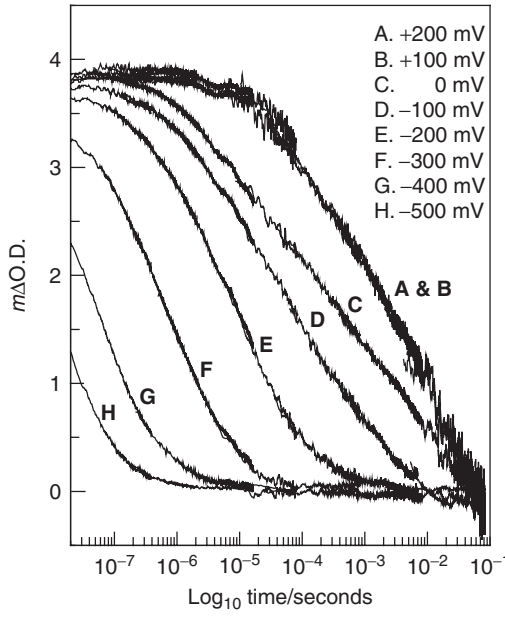


Figure 15.7 Transient absorption of dye-cation on TiO_2 electrode immersed in an electrolyte as a function of applied potential. Taken from Figure 2 of reference [78]

cation, the rate to I_3^- increased with the increase of light intensity. This can be modeled again with transport-limited recombination [89], and/or the recombination from surface traps, where the energy of electrons increases with the filling of the traps, increasing free energy difference between the trapped electrons and the ground state of dye [34].

Observed lifetime in typical nanoporous TiO_2 varies between 10^{-3} and 10^{1} s (Figure 15.8). Such long electron lifetimes balance the slow electron diffusion in nano-porous TiO_2, resulting in sufficient electron diffusion length [90]. When ferrocene/ferrocenium was employed as redox couple for DSSCs, the V_{oc} of the DSSCs was very low, suggesting a fast recombination rate [91]. This implies that the long electron lifetime with I^-/I_3^- is probably realized by the large reorganization energy of the I^-/I_3^- redox couple [92]. However, note that the drawback of the slow reduction rate of I_3^- is the large overpotential needed to reduce dye cation by I^-, resulting in large energy loss.

15.6 NEW MATERIALS

15.6.1 Photosensitizers

15.6.1.1 Metal-complex photosensitizers

In addition to Ru-complexes, other metal-complexes have also been synthesized, and their performance in DSSCs has been investigated. These include Fe-complexes [93, 94], Pt-complexes [95, 96], Os-complexes [97–99], and Re-complexes [100] (Figure 15.9). A DSSC based on a square-planar Pt complex shows an efficiency of 3% ($J_{sc} = 7.00\,mA\,cm^{-2}$, $V_{oc} = 0.60\,V$, $FF = 0.77$) under simulated AM 1.5 G solar irradiation [96]. A DSSC based on an Os-complex gives a wide photoresponse

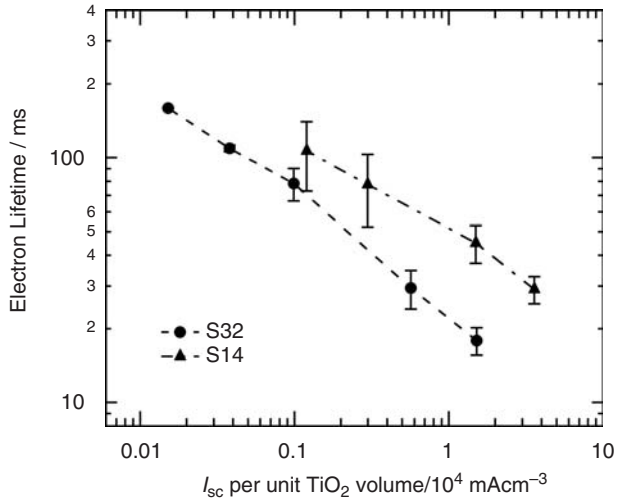


Figure 15.8 Electron lifetime in DSSCs as a function of short-circuit current per unit TiO_2 volume. S32 and S12 denote the samples prepared from TiO_2 nano-particles having average diameter of 14 and 32 nm, respectively. Adapted from Figure 4 of reference [74]

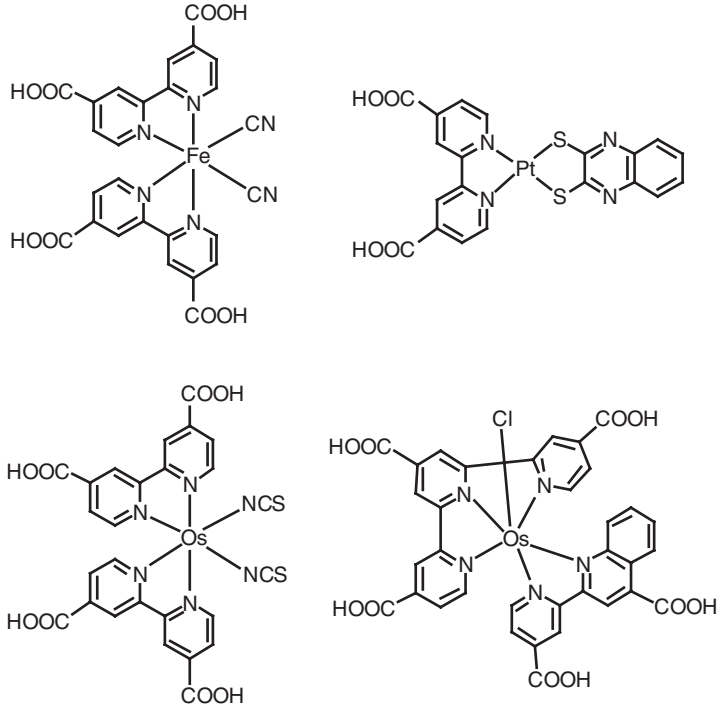


Figure 15.9 Molecular structures of Fe-, Pt-, and Os-complex photosensitizers

range (400–1100 nm) with a maximum IPCE of 65%, corresponding to $J_{sc} = 18.5\,mA\,cm^{-2}$ under AM 1.5 G irradiation [99]. However, solar cell performance with these metal complexes has not exceeded that of the Ru-complex sensitizers. One of the reasons of the better performance with Ru is that the energy levels of the HOMO and the LUMO of Ru-complexes are better matched to the iodine redox potential and the conduction band level of the TiO_2 electrode, respectively.

15.6.1.2 Phorphyrins and phthalocyanines

Porphyrin [101–105], phthalocyanine [106–108], and naphthalocyanine [109] derivatives have also been employed as sensitizers in DSSCs (Figure 15.10). A DSSC with a nanocrystalline TiO_2 electrode sensitized by Cu chlorophyllin shows an efficiency of 2.6% ($J_{sc} = 9.4\,mA\,cm^{-2}$, $V_{oc} = 0.52\,V$) under $100\,mW\,cm^{-2}$ irradiation [101]. For Zn-phthalocyanine, Mori and Kimura *et al*. [110] have achieved an efficiency of 4.6% by preventing the aggregation on TiO_2 electrode [110]. Campbell *et al*. [102] have studied various structures of porphyrin dyes, and have achieved a Zn-porphyrin that shows an efficiency of 7.1% ($J_{sc} = 14.0\,mA\,cm^{-2}$, $V_{oc} = 0.68\,V$, $FF = 0.74$) under AM 1.5 G irradiation [105]. Ultra-fast electron transfer (<100 fs to ~10 ps) from a Zn-porphyrin sensitizer to TiO_2 was observed by transient absorption spectroscopy [51, 111].

Co-adsorbates, such as cholic acid (CA) derivatives, usually improve the photovoltaic performance of solar cells based on porphyrins and phthalocyanines [101, 107], although the amount of the dye adsorbed on the TiO_2 surface is decreased by the presence of the co-adsorbate. Strong intermolecular interactions between porphyrin and phthalocyanine molecules lead to dye aggregation, which in turn leads to quenching processes by means of energy-transfer or charge-transfer reactions between the aggregated molecules and/or between aggregated molecules and monomers. Thus, the co-adsorbates would improve solar cell performance by suppressing dye aggregation on the surface. In addition, co-adsorbates would improve the V_{oc} by suppressing recombination between the injected electrons and I_3^- ions, while the effect seems to be dependent on the structure of sensitizers [112, 113].

15.6.1.3 Organic dyes

Metal-free organic dyes can also be utilized as sensitizers in DSSCs [114]. In early studies of dye-sensitized photoelectrochemical electrodes, 9-phenylxanthene dyes, such as rose bengal and rhodamine B, were used as sensitizers for ZnO electrodes [2, 3]. Organic dyes have several

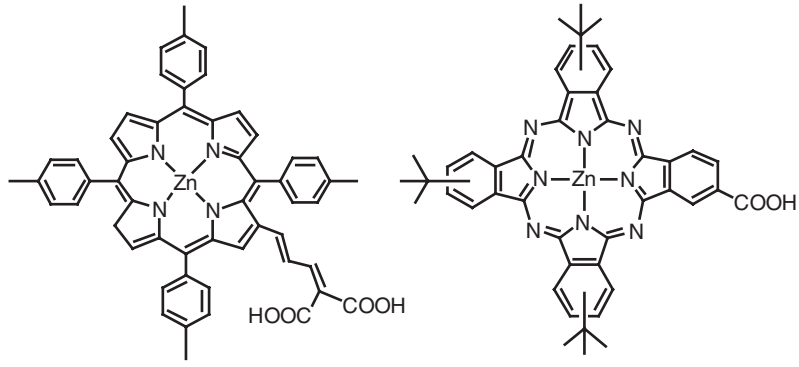


Figure 15.10 Molecular structures of porphyrin and phthalocyanine photosensitizers

advantages: a wide variety of structures can be obtained, structural modification is easy, and organic dyes have high molar absorption coefficients ($30\,000-100\,000\,M^{-1}\,cm^{-1}$) relative to those of Ru-complex sensitizers, owing to intramolecular $\pi-\pi^*$ transitions. In addition, dyes do not contain noble metals, such as Ru, Pt, and Os, which might limit the use of metal-containing sensitizers in large-scale commercial applications. The photovoltaic performance of DSSCs based on organic-dye sensitizers has been improved. Recently, high efficiencies (>9%) under AM 1.5 G irradiation have been obtained with DSSCs based on organic dyes [88, 115–124].

The molecular structures of some organic-dye sensitizers are shown in Figure 15.11. These molecules consist of a donor moiety (a part of molecule, e.g. an aniline, a coumarin, or a benzothiazol unit) and an acceptor moiety (e.g. carboxylic acid, an acrylic acid or a rhodanine ring) connected by a π-conjugated structure, such as a carbon–carbon double bond or an oligothiophene unit. This donor–acceptor structure gives a strong absorption with a large absorption coefficient in the visible region from 450 to 600 nm due to the intramolecular $\pi-\pi^*$ transition. The strong absorption in the visible region is desirable for harvesting the solar spectrum. In addition, these organic-dye sensitiers, like metal-complex sensitizers, have an anchoring group such as a carboxyl

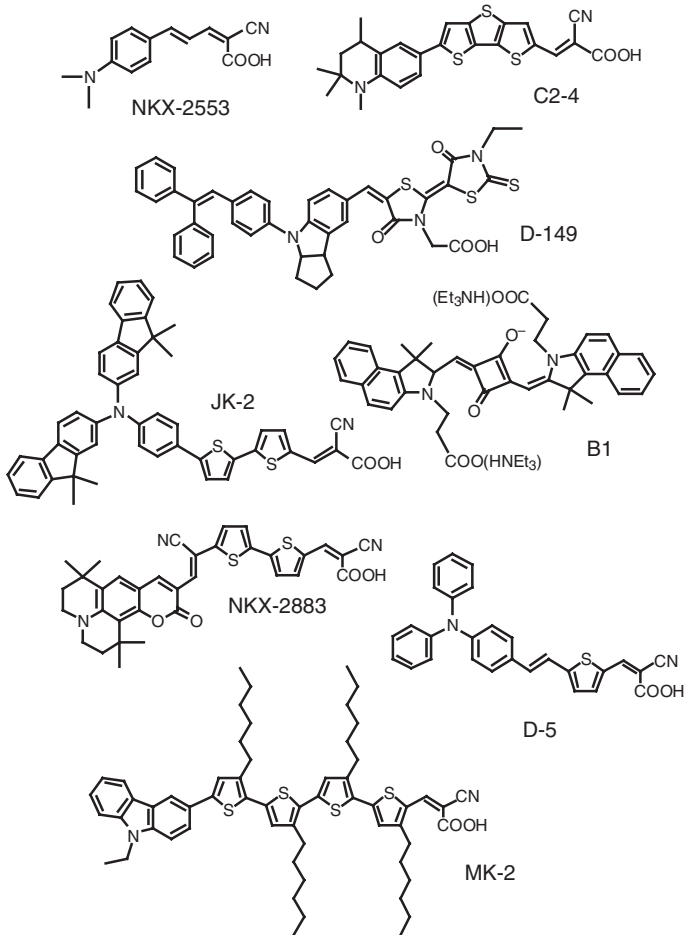


Figure 15.11 Molecular structures of organic dye photosensitizers

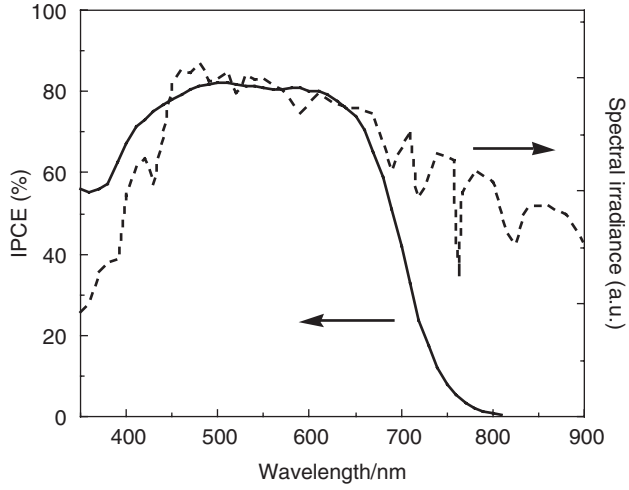


Figure 15.12 An IPCE spectrum for a DSSC based on a coumarin dye (NKX-2883), and the spectral irradiance of standard AM 1.5 G

group to adsorb to the TiO_2 surface. After the intramolecular charge transfer, the excited electron can be injected into the conduction band of the TiO_2 electrode via the anchoring carboxyl group.

The IPCE spectra for DSSCs composed of a nanocrystalline TiO_2 electrode, organic dye sensitizer NKX-2883, and an iodine redox (I^-/I_3^-) electrolyte are shown in Figure 15.12, along with the irradiance of the solar spectrum (under AM 1.5 G conditions). Photons with a wide range of wavelengths (350–800 nm) can be converted to current with the DSSCs based on these organic dyes. IPCE values higher than 70% were observed at 420–660 nm, where the irradiance of the solar spectrum is relatively strong; the maximum value was 81% at 490 nm.

The photovoltaic performances of DSSCs based on organic dye sensitizers are listed in Table 15.2. DSSCs based on indoline dyes (D149 and D205) produced high η values (up to 9.5%) under AM 1.5 G irradiation (100 mW cm^{-2}) [115, 125–127]. Kim *et al*. [37] designed some new organic dyes and reported an efficiency of 8.0% ($J_{sc} = 14.0$ mA cm^{-2}, $V_{oc} = 0.753$ V, $FF = 0.77$) with a DSSC based on JK-2 dye. We also obtained an η value of 8.3% ($J_{sc} = 15.2$ mA cm^{-2}, $V_{oc} = 0.73$ V, $FF = 0.75$) with a DSSC based on a carbazole dye MK-2 under simulated AM 1.5 G irradiation (100 mW cm^{-2}) with an aperture mask and without an antireflection (AR) film [123, 128]. This η value can be compared to the value of 9.2% ($J_{sc} = 15.9$ mA cm^{-2}, $V_{oc} = 0.78$ V, $FF = 0.74$) obtained with a DSSC based on Ru-complex N719 dye under similar conditions. In addition, a high J_{sc} value (18.8 mA cm^{-2}; with a mask and without an AR film) was obtained with a DSSC based on NKX−2883 [129]. These results indicate that organic dyes can perform nearly as well as Ru complexes in DSSCs.

15.6.1.4 Issues on sensitizing dyes

In comparison with the Ru-complex dye, N719, almost all organic sensitizing dyes suffer from low values of V_{oc}. The low V_{oc} can be due to positive shift of E_{cb} by dye adsorption and/or decreased electron lifetime. For the case of DSSCs employing coumarin, indoline, carbazole, porphyrin and phthalocyanine dyes, the low V_{oc} was due to short electron lifetime (Figure 15.13). [35, 128, 130–132]. Among the dyes, larger molecular size tends to show longer electron lifetime

Table 15.2 Performance of DSSCs based on organic-dye sensitizers

Dye	Area/cm^2	J_{sc}/mA cm^{-2}	V_{oc}/V	*FF*	η(%)	Reference
D-11	0.2	13.9	0.74	0.70	7.2	[120]
JK-2	0.16	14.0	0.75	0.77	8.0	[37]
JK-71	0.18	15.4	0.74	0.74	8.4	[116]
D-149	0.16	18.5	0.69	0.62	8.0	[126]
D-149	0.16	20.0	0.65	0.69	9.0	[115]
D-205	0.16	18.6	0.72	0.72	9.5	[118]
C217	0.16	16.1	0.80	0.76	9.8	[124]
NKX-2700	0.25	15.9	0.69	0.75	8.2	[88]
MK-2	0.25	15.2	0.73	0.75	8.3	[123]

Light source: AM1.5 simulator, Electrode: nanocrystalline TiO_2 electrode, electrolyte: iodine redox in organic solvents.

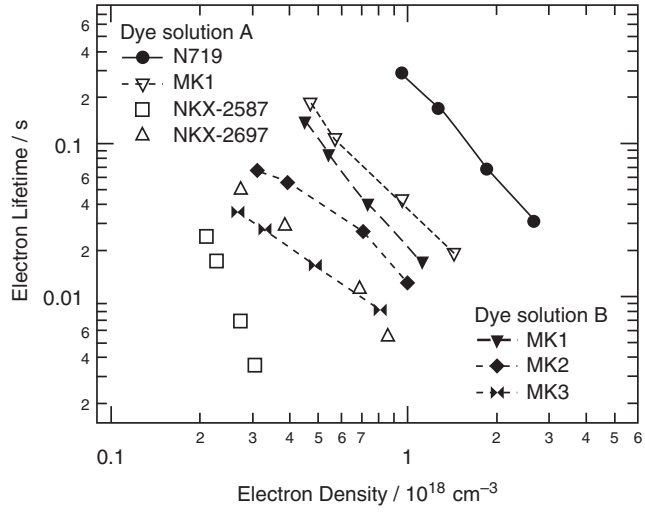


Figure 15.13 Electron lifetime in DSSCs with various sensitizing dyes. Solution A and B denote dye solutions prepared from AN/tBuOH/toluene and AN/tBuOH mixed solvents, respectively. Adapted from Figure 2(c) of reference [35]

[35]. When the adsorbed amount of the large molecule dyes was reduced, the lifetime decreased, indicating that the dyes physically prevent the approach of the I_3^- to the TiO_2 surface. However, even for the case of the larger metal-free organic dyes, the electron lifetime is still shorter than that with N719, implying that the dyes increase the concentration of I_3^- by forming weak dye–I_3^- complex. For the case of a metal-free organic dye, further decrease of dye loading turns to increase the lifetime, consisting with the formation of dye–I_3^- complex [35]. Moreover, the electron lifetime is influenced by the species of counter cation for I^-/I_3^- redox couple, and the way of influence differs among the examined dyes, implying that the cations are involved for the complex formation [35]. Note that the counter-cation would also influence on the [I_3^-] by changing the thickness of double layer on the TiO_2 surface [13]. In order to exceed the performance of N719, the complex formation must be avoided. While several series of dyes showed similar tendency mentioned above, opposite cases have also been reported For the case of oligoene dyes, the smallest size dyes showed high V_{oc} [63], and with the increase of the molecular size, which is to expand absorption spectrum

by increasing π conjugation, the values of V_{oc} were decreased [133]. This was interpreted with dispersion force, that is, longer conjugated molecules result in larger polarizability, and thus, larger intermolecular forces between the dye and acceptor species [134]. Since the force is weak, this could be avoided by adding obstacle units such as alkyl chains and twisted phenyl rings to the dyes without increasing the polarizability.

15.6.2 Semiconductor Materials

15.6.2.1 Nanowire/rod/tube

While nanoporous TiO_2 has been employed mostly for DSSCs, other materials and morphologies have been explored. One emerging area is employing metal oxide nanowire/rod/tube [135–139]. When solid state hole conducting materials are applied, one obstacle is its difficulty to penetrate into the nanosized pore of conventional porous electrodes. In addition, current trend of sensitizing dyes is to increase the molecular size to better absorb infrared light, which again require larger pore size. With these materials, vertically aligned nanowires allow easy penetration [140]. Another motivation to use nanotubes would be to reduce the number of boundaries, and to control charge traps.

Figure 15.14 shows SEM photographs of nanowires [135]. In comparison with nanoparticle-based electrodes, the nanowires showed superiority as a charge collector. In addition, nanowires showed only a fast component of charge injection, consistent with their higher crystallinity. Recombination kinetics are slower for nanorods compared with nanoparticles, Martinson [138] also reported slower charge recombination rate with faster electron transport in a rod array, in comparison with those with nanoporous electrodes. This is interesting because, as it was mentioned, faster diffusion often results in faster recombination. Recently, Quintana *et al*. [141] interpreted that their observed long electron lifetime was due to the formation of depletion layer. This is because ZnO has low dielectric constant, and thus, substantial potential drop can be formed in nanoparticles. The depletion layer would enhance charge separation efficiency, but also increase the barrier at grain boundary. For the case of nanorods, they can exploit only the increased separation efficiency

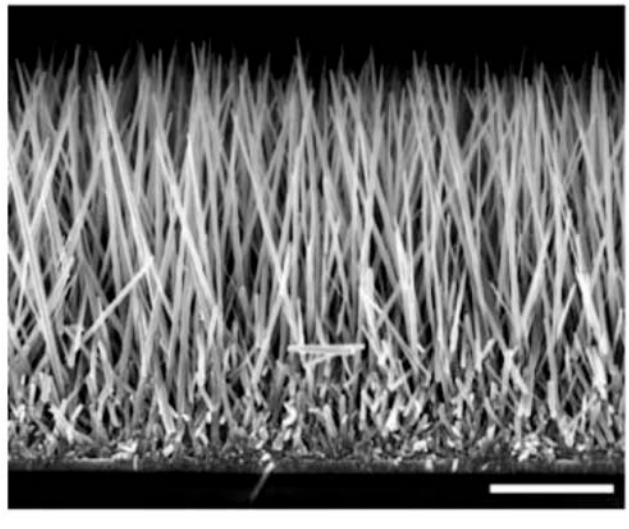


Figure 15.14 SEM photograph of ZnO wire. Scale bar is 5 μm. Taken from Figure 1(b) of reference [135]

without increasing the barrier for transport. The assist of depletion layer for charge separation was also mentioned by Law *et al*. from different point of view [135].

15.6.2.2 Other morphology and materials

In order to increase the range of the absorption spectrum for sensitizing dyes, either the decrease of LUMO level or the increase of HOMO level is required. For the case of the decrease of LUMO level, the E_{cb} of TiO_2 may be too high for charge injection of such dyes [142]. SnO_2 and In_2O_3 have much lower E_{cb} than that of TiO_2 and thus, may by suitable for the dyes. In comparison with TiO_2, the value of V_{oc} with these metal oxides is inherently decreased, while these materials would become a key material for the success of fabrication of tandem DSSCs, where more than two dyes having different LUMO levels are likely to be employed with more than two semiconductors having different E_{cb} connected in series.

The J_{sc} of dye-sensitized SnO_2 is typically lower than that of TiO_2-based DSSCs [143, 144]. This has been interpreted as due to fast charge recombination caused by higher electron mobility of the SnO_2 [144]. On the other hand, it has been shown that the electron transport properties of nanoporous TiO_2 are largely influenced by charge traps, whose density and energetic distribution depends on the fabrication condition of the electrodes [76]. Recently, a dye-sensitized SnO_2 solar cell showed comparable values of J_{sc} with that of TiO_2-based DSSCs [145]. The results were attributed to longer electron lifetime, suggesting the materials can be applied by utilizing charge traps.

Another approach is employing metal oxides having high E_{cb} potential. While not much progress has been done on such materials, doping of W to TiO_2 can shift the E_{cb} negatively. With the dye having high LUMO potential, 1 V of V_{oc} was achieved [146].

So far, most metal oxides are *n*-type semiconductors. On the other hand, sensitization of *p*-type semiconductors is also possible. NiO is one of materials behaving as a *p*-type semiconductor. Although the efficiency is low, hole injection from the dye to NiO has been observed [147, 148]. Due to the lack of appropriate dyes and redox couple, the efficiency of dye-sensitized NiO at this moment is low. However, the concept of tandem DSSCs by replacing Pt electrode in the conventional DSSCs has high potential toward the increase of energy conversion efficiency [149].

15.6.3 Electrolytes

15.6.3.1 Ionic liquid electrolytes

Room-temperature ionic liquids (molten salts) have been extensively studied as replacements for volatile organic solvents in electrochemical devices such as batteries because of their high ionic conductivity, electrochemical stability, and nonvolatility. Of these properties, nonvolatility is the most critical for ensuring the long-term stability of electrochemical devices. Such room-temperature ionic liquids have also been utilized and studied in DSSCs in place of liquid electrolytes [127, 150–153]. Ionic liquids used in DSSCs include imidazolium derivatives, such as 1-hexyl-3-methylimidazolium iodide (HMImI) [150] and 1-ethyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide (EMImTFSI) [151, 153]. Matsumoto *et al*. [151] reported that an N3 dye-sensitized TiO_2 solar cell using an EMIm salt having hydrofluoride anions, $H_2F_3^-$ or $H_3F_4^-$, as the electrolyte solvent produced 2.1% efficiency under AM 1.5 ($J_{sc} = 5.8\,mA\,cm^{-2}$, $V_{oc} = 0.65\,V$, and $FF = 0.56$) [151]. Kuang *et al*. [44, 127] reported a high η values more than 7% under AM 1.5 G irradiation with DSSCs based on a Ru-complex (K-19) and an organic dye (D205) with an electrolyte consisting of 0.2 M I_2,

0.5 M *N*-butylbenzimidazole (NBB), and 0.1 M guadinium thioisocyanate (GuNCS) in a mixture of 1-methyl-3-n-propylimidazolium (MPImI) and 1-ethyl-3-methylimidazolium tetracyanoborate ($EMIB(CN)_4$) [44, 127]. If the viscosity of these ionic liquids can be decreased similarly to that of organic solvents, the solar cell performance will be improved as a result of increased ionic mobility of the electrolyte. Besides the viscosity, employing large molecular size sensitizers with ionic liquids result in the increase of electron lifetime, giving high conversion efficiency [154].

15.6.3.2 New redox species

The required overpotential of I^-/I_3^- redox couple to reduce dye cation is around 0.5 V. The I^-/I_3^- redox couple has considerable absorption coefficient up to 500 nm. These properties of the redox couple decrease the amount of photon energy to be converted. On the other hand, the redox couple can reduce the dye cation before the reduction by the injected electrons, and can wait for electrons reach TCO before they transfer to I_3^-. Other redox couples may be too slow to reduce the dye cation [16], and/or too fast to accept the injected electrons [91]. Recently, several metal complexes designed for redox couples have been reported. Co-complexes showed smaller absorption coefficient with relatively high conversion efficiency [155]. Cu complexes showed lower (more positive) redox potential with relatively high V_{oc} [156].

15.6.3.3 Quasi-solid-state and solid-state electrolytes

Development of solid-state or quasi-solid-state DSSCs is essential for developing a cell with long-term stability and is critical for commercialization. Because liquid electrolytes using organic solvents are usually utilized in conventional DSSCs, techniques for sealing the cell must be perfectly established to prevent evaporation of components of the electrolyte, especially under high temperatures in outdoor applications. In addition, the solid-state DSSC would allow easier interconnection of a cell into a monolithic module.

Grätzel and co-workers studied DSSCs using a hole-transport material, 2,2′,7,7′-tetrakis(*N*, *N*-di-p-methoxyphenyl-amine)9,9′-spirobifluorene (OMeTAD), as a solid electrolyte [157–160]. OMeTAD is spin-coated on the surface of the dye-sensitized nanocrystalline TiO_2 electrode and then Au is deposited by vacuum evaporation as the counter-electrode, resulting in a solid-state DSSC. The η value of 3.2% ($J_{sc} = 4.6\,mA\,cm^{-2}$, $V_{oc} = 0.931\,V$, and $FF = 0.71$) was obtained under AM 1.5 G ($100\,mW\,cm^{-2}$) [160]. The electron injection from OMeTAD into cations of N3 dyes occurs in the range of <3 ps to >1 ns, which is faster than that from I^- ions [158], although the redox potential of OMeTAD is located at more positive potential than that of I^-/I_3^-. With an organic dye (D102) having high extinction coefficients enabling thin films, a DSSC based on OMeTAD has produced 4.1% ($J_{sc} = 7.7\,mA\,cm^{-2}$, $V_{oc} = 0.866\,V$, and $FF = 0.612$) under AM 1.5 G irradiation [159]. The high values of V_{oc} due to more positive redox potential is another advantages of the hole conductor, while further increase can be expected by reducing charge recombination [161].

Tennakone and co-workers [162–164] utilized a p-type inorganic semiconductor material, CuI (band gap, 3.1 eV), as a hole conductor and fabricated a solid–state DSSC. Acetonitrile solution of CuI is dropped on the surface of dye-coated TiO_2 film heated at approximately 60 °C and then is diffused into inside of the film. After evaporation of the acetonitrile, CuI can be deposited into nanoporous TiO_2 film. Au-coated TCO substrate as the counter-electrode is pressed onto the surface of the TiO_2/dye/CuI film. The efficiency reached 4.5% for the TiO_2/N3 dye/CuI/Au system, suggesting the possibility that a highly efficient solid-state DSSC could be produced [163]. In these systems, CuI is considered to be partially in contact with TiO_2, decreasing cell performance due to the recombination of injected electrons. To increase cell performance, there must be decreased

TiO_2/CuI contact. Solid-state DSSCs have been studied using other organic and inorganic hole conductor materials, *p*-type CuSCN [165, 166], polypyrrole [167], polyacrylonitrile [168], PEDOT [169, 170], poly(*N*-vinylcarbazole)[171], and P3HT [139].

Quasi-solidification of the electrolyte using a gelator is another method for replacing liquid electrolytes in DSSCs [172–175]. Gelation can be accomplished by adding gelator into the electrolyte without other changes in the components of the electrolyte. Yanagida and co-workers studied gelation of the electrolyte using L-valine derivatives as a gelator and measured solar cell performance of DSSCs using gel electrolytes [172]. Interestingly, the performance of DSSCs using a gel electrolyte was almost the same as that using a liquid electrolyte. Good long-term stability of the sealed cell using the gel electrolyte was obtained compared to that of a sealed cell using a liquid electrolyte.

Hayase and co-workers reported constructing a highly efficient DSSC using a gel electrolyte [173]. The gelator was dissolved in the electrolyte at high temperature, and consequently the gel solution was deposited on the dye-coated TiO_2 electrode surface and then cooled. Gelation is caused by polymerization between nitrogen and halogen compounds. A high efficiency, 7.3% ($J_{sc} = 17.6\,mA\,cm^{-2}$, $V_{oc} = 0.60\,V$, and $FF = 0.68$), was obtained for a DSSC based on N3 dye using the gel electrolyte under AM 1.5 G irradiation, compared with 7.8% for a solar cell based on a liquid electrolyte [173]. They concluded that the resistance of the electrolyte did not increase as a result of the gelation because no change in the fill factor was observed. The photocurrent increased linearly with increasing incident light intensity of up to $100\,mW\,cm^{-2}$, as well as a liquid DSSC. This suggests that gelation of the electrolyte does not suppress diffusion of I^- and I_3^- ions in the electrolyte.

15.7 STABILITY

15.7.1 Stability of Materials

The commercialization of DSSCs will require long-term stability of the dye molecules and of individual solar cells or large-area modules. The photostability and thermal stability of Ru complexes have been investigated in detail [7, 176–178]. For example, the thiocyanate (NCS) ligand of N3 dye is oxidized to a cyano group (–CN) under photoirradiation in methanol solution, as indicated by UV–vis absorption and NMR spectroscopy [7, 176]. Substitution of the NCS ligand of N719 dye by solvents (acetonitrile and 3-methoxypropionitrile) and TBP has been observed at elevated temperatures (80–110 °C) [178]. Decarboxylation of N3 dye occurs at temperatures higher than 290 °C under N_2 and higher than 250 °C in air [179]. When the dye is adsorbed onto the TiO_2 surface, the decarboxylation occurs at temperatures above 320 °C in air.

High dye stability can be obtained in a DSSC system by including I^- ions as the electron donor to dye cations. Degradation of the NCS ligand to a CN ligand due to an intramolecular electron transfer producing Ru^{2+} from Ru^{3+} occurs within 0.1–1 s [176], whereas the rate for reduction of Ru^{3+} to Ru^{2+} due to electron transfer from I^- ions to the dye cations is of the order of nanoseconds [49]. This comparison indicates that one molecule of N3 dye can contribute to the photon-to-current conversion process with a turnover number of at least 10^7–10^8 without degradation [176]. Taking this into consideration, N3 dye is considered to be sufficiently stable in the redox electrolyte under irradiation.

We investigated the thermal stability of coumarin dyes NKX-2311 and NKX-2677. In thermogravimetric analysis of the dyes, an 8% drop in mass was observed for NKX-2677 at around 278 °C under an O_2 atmosphere [62], and a similar mass drop was observed for NKX-2311 at around 230 °C under a N_2 atmosphere [180]. FT-IR absorption spectroscopy clearly indicated that

these mass drops are due to decomposition of the dye by decarboxylation [180]. From these results, we conclude that the thermal stability of organic dyes such as NKX-2677 is as good as that of N3 dye. For photostability, we found that employing oligothiophene unit has an advantages, probably due to the delocalization of holes in the unit [181].

We must also take into consideration the photoelectrochemical and chemical stability of the solvent in the electrolyte. Organic solvents employed in DSSCs are, for example, propylene carbonate, acetonitrile, propionitrile, methoxyacetonitrile, methoxypropionitrile, and their mixtures. It is known that carbonate solvents, such as propylene carbonate, decompose under illumination, resulting in the formation of a carbon dioxide bubble in the cell. Methoxyacetonitrile (CH_3O-CH_2CN) reacts with trace water in the electrolyte to produce corresponding amide (CH_3O-CH_2CONH_2), which decreases the conductivity of the electrolyte [182]. Acetonitrile and propionitrile are considered to be relatively stable, giving 2000 h stability under dark conditions at 60 °C [182].

The stability of vapor-deposited Pt electrocatalyst on a TCO substrate used as a counter-electrode has also been investigated. It was reported that the electrocatalytically active Pt layer did not seem to be chemically stable in an electrolyte consisting of LiI and I_2 dissolved in methoxypropionitrile [183].

15.7.2 Long-term Stability of Solar Cell Performance

The long-term stability of small DSSCs and large-area modules is currently being investigated for commercial applications [43, 44, 176, 182, 184, 185]. DSSCs based on Ru-complexes such as N3 and N719 dyes have good long-term stability under continuous irradiation. For example, Grätzel and co-workers achieved 7000 h of cell stability under 1000 W m^{-2} without UV light [176]. They concluded that the good long-term stability of the DSSC is due to effective photoinduced electron transfer from the Ru-complex sensitizer into the conduction band of the semiconductor, and from the iodine redox mediator to the sensitizer. The turnover number of one sensitizer molecule, which is defined by the number of electrons produced by one sensitizer molecule, reaches 500 million. Furthermore, Kern *et al*. [182] reported a long-term stability of more than 10 000 h with a cell in the absence of UV light at 17 °C at 2.5 sun [182].

We investigated the long-term stability of a DSSC based on coumarin dye NKX-2883 under continuous AM 1.5 G irradiation through a UV (<420 nm) cut-off filter (100 mW cm^{-2}, 50–55 °C) under open-circuit conditions. No signs of dye degradation and no decrease in solar cell performance were observed over a period of 1000 h [129]. This result suggests that coumarin dyes are also relatively stable under irradiation in solar cells, where redox ions are present and electron-transfer processes occur effectively. Taken together, these results strongly indicate that DSSCs show sufficient long-term stability during irradiation, although the stability of cell and module performance under UV irradiation and at high temperature and humidity must be investigated in further detail.

15.8 APPROACH TO COMMERCIALIZATION

15.8.1 Fabrication of Large-area DSSC Modules

Increased sheet resistance of the TCO substrate on scale-up of the DSSC leads to loss of efficiency, especially fill factor. Therefore, scale-up of the DSSC using a modular approach has been investigated [184–186]. A module consists of the basic cell with two TCO glass substrates coated with TiO_2 or platinum and inside the electrolyte plus interconnections. The electrolyte contains iodine and iodide, which dissolve metal materials, dissolved in an organic solvent. Therefore, standard

conductors such as silver will not work or has to be protected by a sealing materials. In addition, with an organic solvent present in such a system one must seal the system carefully also at the outside. To seal the modules, an inert material, glass frit even for interconnections have been used [187]. An efficiency of 7% was achieved using a module consisting of 12 interconnected cells with a total area of 112 cm^2 compared with 7.6% for a 3 cm^2 cell and 8% for a 1 cm^2 cell [187].

Figure 15.15 shows schematic structure of series connected DSSC modules [188]. In a Z-type module (Figure 15.15a), TiO_2 electrode and counter electrode are separately printed on each TCO substrate. The cells are separated from each other by scribing of TCO substrate and separator, and interconnection between each cell is necessary. In W-type module (Figure 15.15b), TiO_2 electrode and counter electrode are alternately printed on each TCO substrate and interconnection between each cell is not necessary. However, the counter-electrode must be transparent because light illumination is through the counter-electrode in this module. The third type is monolithic module as shown in Figure 15.15c. An advantage of the monolithic type is that one TCO substrate is only needed, having possibility of low-cost production due to decreasing the TCO substrate, which leads to high-cost DSSC production. In this type module, generally, carbon materials are used as counter-electrode material. Kay and Grätzel proposed a continuous process for the fabrication of monolithic series-connecting DSSC modules using laser scribing and printing techniques [189]. Figure 15.16 shows photographs of large-area DSSC module produced by Sharp Corporation, and Aisin Seiki Co. Ltd. and Toyota Central R&D Labs, Inc.

15.8.2 Flexible DSSC

Recently, polymer or metal foil substrates instead of glass have been utilized in constructing DSSCs, expanding possible commercial applications [41, 190–192]. Polymer substrates allow roll-to-roll

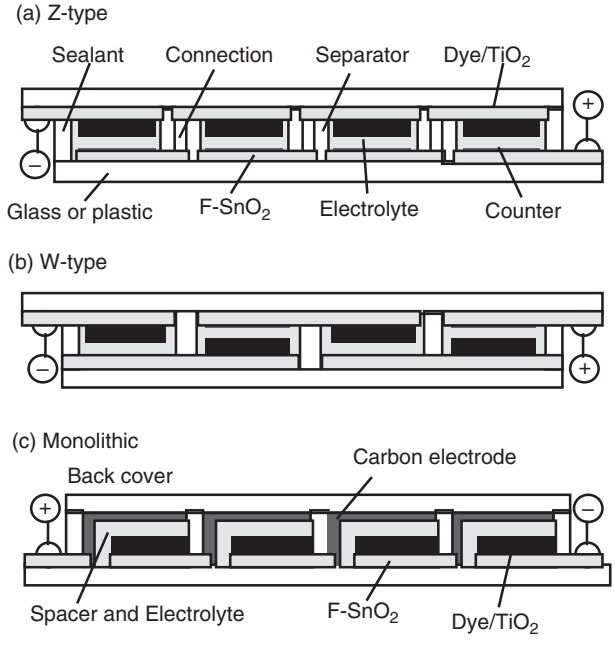


Figure 15.15 Schematic structures of series-connected DSSC modules

Figure 15.16 Large-area DSSC modules produced by: (a) Sharp Corporation; (b) Aisin Seiki Co. Ltd. and Toyota Central R&D Labs., Inc.

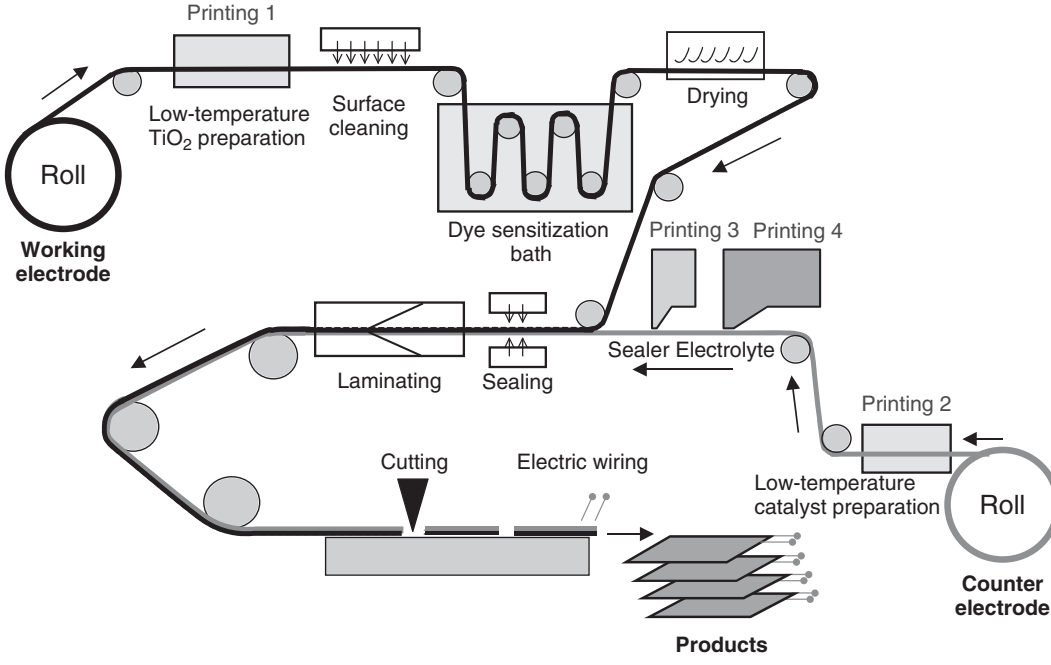


Figure 15.17 Schematic image of roll-to-roll production of plastic DSSC proposed by Professor Miyasaka (Toin University of Yokohama, Japan)

production, which can achieve high throughput. Figure 15.17 shows a schematic image of roll-to-roll production of plastic DSSCs proposed by Miyasaka at Toin University of Yokohama. When a polymer film is used as a substrate, aqueous TiO_2 paste without organic surfactants is sintered at relatively low temperatures, approximately 150 °C being sufficient to produce mechanically stable TiO_2 films. Miyasaka and co-workers reported high η of 5.5% at 1 sun condition with a DSSC based on plastic substrates and N719 dye [193]. They also fabricated large area plastic DSSC modules, as shown in Figure 15.18. Ito *el al*. attained a high η of 7.2% with a flexible DSSC based on Ti-metal foil substrate [192].

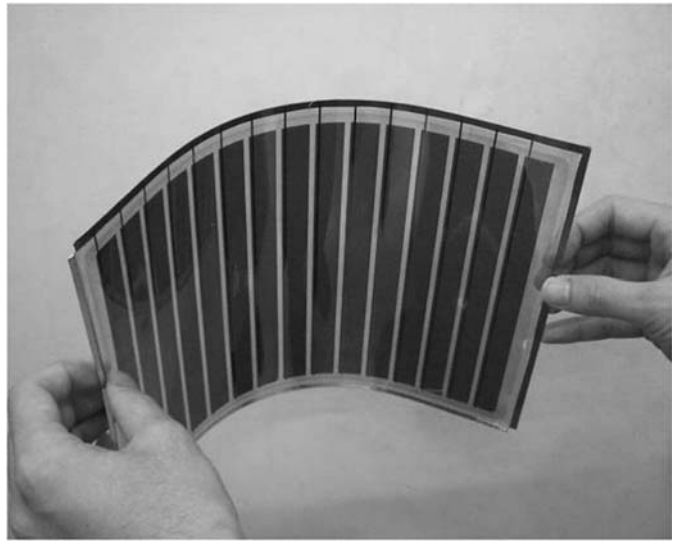


Figure 15.18 A large-area module of plastic DSSC produced by Professor Miyasaka and co-workers

15.8.3 Other Subjects for Commercialization

In view of appearance, colorful DSSCs can be made using various kind of dyes, depending on the use to be made of the cell. Aisin Seiki Co. Ltd. and Toyota Central R&D Labs., Inc. produced art objects using colorful DSSC (Figure 15.19). They chose several kinds of dye whose absorption property is different to fabricate colorful DSSCs.

DSSCs have been used for educational demonstration of solar energy-to-electricity conversion because of its simple fabrication [189]. One can purchase DSSC kits including all components, TCO-coated glass, TiO_2 electrodes, blackberries (i.e. dye), and electrolyte solution (http://www.solideas.com/) and easily demonstrate an artificial photosynthetic process. For detailed studies, one can purchase other materials, including Ru-complex sensitizers, TiO_2 paste, electrolytes, and sealing materials from Solaronix S. A. (http://solaronix.com/).

15.9 SUMMARY AND PROSPECTS

Since 1991, when Grätzel and coworkers reported the development of highly efficient, novel DSSCs, researchers throughout the world have intensively investigated DSSC mechanisms, new materials, and commercialization. Solar energy-to-electricity conversion efficiencies more than 11% (with small cells) have been attained with DSSCs based on Ru-complex sensitizers. Organic dye sensitizers, such as indoline dyes, coumarin dyes, and carbazole dyes, also show high performance as sensitizers for DSSCs. In addition, satisfactory long-term stability of sealed cells has been achieved under relatively mild test conditions (low temperatures and no UV exposure). It will be possible to achieve commercial DSSC production in the near future for indoor applications. For expanded commercial applications, however, there are several problems which confront us. Overcoming these problems greatly brings DSSCs close to expanded commercialization.

Figure 15.19 Art objects using colorful DSSC produced by Aisin Seiki Co. Ltd. and Toyota Central R&D Labs., Inc.

For commercial applications, cell efficiencies higher than 11% (e.g. 15%) and module efficiencies higher than 10% are desired. To attain 15% cell efficiency, both J_{sc} and V_{oc} must be further improved. Expanding the absorption of the sensitizers to the near-IR region is necessary for improved J_{sc}. The absorption threshold of black dye is close to 900 nm, and this value is expected to be the optimal threshold for single-junction solar cells, whose E_g is ~1.4 eV. A high J_{sc} value (20.9 mA cm^{-2}) was achieved with a DSSC based on black dye (with an aperture mask and an AR film). Development of new sensitizers that can absorb in the near-IR region with relatively large absorption coefficients is also desired. Improvement of the IPCE of DSSCs sensitized by black dye (or new sensitizers) in the 800–900 nm range (Figure 15.4), would increase the J_{sc} from 20.9 mA cm^{-2} to about 25 mA cm^{-2}. Improvement of V_{oc} from 0.75 V to the optimal value of ~0.9 V is also important for achievement of higher efficiencies. The fact that V_{oc} values have not yet reached the optimal value is mainly attributed to charge recombination between injected electrons and redox ions (I_3^-), and large overpotential for I^-/I_3^- redox couple. New sensitizers to suppress charge recombination and new redox couple/hole transport material are desired to improve the V_{oc}.

In addition, long-term stabilities of dye molecules and solar cell performance under more rigorous conditions (e.g. high temperature at near 80 °C, high humidity, and UV exposure) will be required for outdoor applications.

ACKNOWLEDGEMENTS

We gratefully acknowledge Dr Tatsuo Toyoda (Aisin Seiki Co., Ltd.), Professor Tsutomu Miyasaka (Toin University of Yokohama), and Dr Ryosuke Yamanaka (Sharp Corporation) for providing photographs of DSSC modules.

REFERENCES

1. Gibbons J F, Cogan G W, Gronet C M, Lewis N S, *Appl. Phys. Lett*. **45**, 1095–1097 (1984).
2. Gerischer H, Michel-Beyerle M E, Rebentrost F, Tributsch H, *Electrochimica Acta* **13**, 1509–1515 (1968).
3. Tsubomura H, Matsumura M, Nomura Y, Amamiya T, *Nature* **261**, 402–403 (1976).
4. O'Regan B, Grätzel M, *Nature* **353**, 737 (1991).
5. Nazeeruddin M K, *et al*., *J. Am. Chem. Soc*. **127**, 16835–16847 (2005).
6. Chiba Y, *et al*., *Jpn. J. Appl. Phys*. **45**, L638–L640 (2006).
7. Nazeeruddin M K, *et al*., *J. Am. Chem. Soc*. **115**, 6382–6390 (1993).
8. Hagfeld A, Grätzel M, *Chem. Rev*. **95**, 49–68 (1995).
9. Kalyanasundaram K, Gratzel M, *Coord. Chem. Rev*. **177**, 347–414 (1998).
10. Grätzel M, *J. Photochem. Photobiol. A* **164**, 3 (2004).
11. Gardner J M, Giaimuccio J M, Meyer G J, *J. Am. Chem. Soc*. **130**, 17252–17253 (2008).
12. Liu Y, Hagfeldt A, Xiao X R, Lindquist S E, *Sol. Energy Mater. Sol. Cells* **55**, 267–281 (1998).
13. Nakade S, *et al*., *J. Phys. Chem. B* **109**, 3480–3487 (2005).
14. Katoh R, *et al*., *J. Mater. Chem*. **17**, 3190–3196 (2007).
15. Hara K, *et al*., *J. Phys. Chem. B* **107**, 597–606 (2003).
16. Oskam G, Bergeron B V, Meyer G J, Searson P C, *J. Phys. Chem. B* **105**, 6867–6873 (2001).
17. Han L, *et al*., *Appl. Phys. Lett*. **86**, 213501 (2005).
18. Tachibana Y, *et al*., *Chem. Phys. Lett*. **364**, 297–302 (2002).
19. Barbe C J, *et al*., *J. Am. Ceram. Soc*. **80**, 3157–3171 (1997).
20. Park N G, van de Lagemaat J, Frank A J, *J. Phys. Chem. B* **104**, 8989–8994 (2000).
21. Saito Y, *et al*., *Sol. Energy Mater. Sol. Cells* **83**, 1–13 (2004).
22. Usami A, *Chem. Phys. Lett*. **277**, 105–108 (1997).
23. Ferber J, Luther J, *Sol. Energy Mater. Sol. Cells* **54**, 265–275 (1998).
24. Rothenberger G, Comte P, Grätzel M, *Sol. Energy Mater. Sol. Cells* **58**, 321–336 (1999).
25. O'Regan B C, Durrant J R, Sommeling P M, Bakker N J, *J. Phys. Chem. C* **111**, 14001–14010 (2007).
26. Nazeeruddin M K, *et al*., *J. Am. Chem. Soc*. **123**, 1613–1624 (2001).
27. Hara K, *et al*., *Sol. Energy Mater. Sol. Cells* **85**, 21–30 (2005).
28. Murakoshi K, *et al*., *Journal of Electroanalytical Chemistry* **396**, 27–34 (1995).
29. Zhang Q L, *et al*., *J. Phys. Chem. B* **108**, 15077–15083 (2004).
30. Hara K, *et al*., *Langmuir* **17**, 5992–5999 (2001).
31. Bauer C, Boschloo G, Mukhtar E, Hagfeldt A, *J. Phys. Chem. B* **106**, 12693–12704 (2002).
32. Nakade S, Kanzaki T, Wada Y, Yanagida S, *Langmuir* **21**, 10803–10807 (2005).
33. Hara K, *et al*., *Sol. Energy Mater. Sol. Cells* **70**, 151–161 (2001).
34. Schlichthörl G, Huang S Y, Sprague J, Frank A J, *J. Phys. Chem. B* **101**, 8141–8155 (1997).
35. Miyashita M, *et al*., *J. Am. Chem. Soc*. **130**, 17874–17881 (2008).
36. Bonhote P, *et al*., *Inorg. Chem*. **35**, 1168–1178 (1996).
37. Kim S, *et al*., *J. Am. Chem. Soc*. **128**, 16701–16707 (2006).
38. Vlachopoulos N, Liska P, Augustynski J, Grätzel M, *J. Am. Chem. Soc.* **110**, 1216–1220 (1988).
39. Papageorgiou N, *Coord. Chem. Rev*. **248**, 1421–1446 (2004).
40. Murakami T N, *et al*., *J. Electrochem. Soc*. **153**, A2255–A2261 (2006).
41. Ikeda N, Miyasaka T, *Chem. Lett*. **36**, 466–467 (2007).
42. Saito Y, Kitamura T, Wada Y, Yanagida S, *Chem. Lett*., 1060–1061 (2002).
43. Wang P, *et al*., *Adv. Mater*. **16**, 1806–1811 (2004).
44. Kuang D, *et al*., *J. Am. Chem. Soc*. **128**, 4146–4154 (2006).
45. Islam A, *et al*., *Chem. Mater*. **18**, 5178–5185 (2006).

46. Jiang K J, *et al*., *Chem. Commun*., 2460–2462 (2006).
47. Chen C Y, *et al*., *Angew. Chem., Int. Ed.* **45**, 5822–5825 (2006).
48. Hagfeldt A, Gratzel M, *Acc. Chem. Res*. **33**, 269–277 (2000).
49. Tachibana Y, *et al*., *J. Phys. Chem*. **100**, 20056–20062 (1996).
50. Haque S A, Tachibana Y, Klug D R, Durrant J R, *J. Phys. Chem. B* **102**, 1745–1749 (1998).
51. Tachibana Y, *et al*., *J. Phys. Chem. B* **104**, 1198–1205 (2000).
52. Benko G, *et al*., *J. Am. Chem. Soc*. **124**, 489–493 (2002).
53. Katoh R, *et al*., *Coord. Chem. Rev*. **248**, 1195–1213 (2004).
54. Asbury J B, *et al*., *J. Phys. Chem. B* **105**, 4545–4557 (2001).
55. Tachiya M, *Radiation Physics and Chemistry* **17**, 447–456 (1981).
56. Katoh R, *et al*., *J. Phys. Chem. B* **106**, 12957–12964 (2002).
57. Katoh R, *et al*., *Comptes Rendus Chimie* **9**, 639–644 (2006).
58. Sayama K, *et al*., *New J. Chem*. **25**, 200–202 (2001).
59. Matsui M, *et al*., *Dyes and Pigments* **80**, 233–238 (2009).
60. Katoh R, *et al*., *J. Phys. Chem. B* **108**, 4818–4822 (2004).
61. Furube A, *et al*., *J. Photochem. Photobiol., A: Chem*. **182**, 273–279 (2006).
62. Hara K, *et al*., *J. Phys. Chem. B* **109**, 15476–15482 (2005).
63. Kitamura T, *et al*., *Chem. Mater*. **16**, 1806–1812 (2004).
64. Södergren S, Hagfeldt A, Olsson J, Lindquist S E, *J. Phys. Chem*. **98**, 5552–5556 (1994).
65. Cao F, Oskam G, Meyer G J, Searson P C, *J. Phys. Chem*. **100**, 17021–17027 (1996).
66. Nakade S, *et al*., *J. Phys. Chem. B* **107**, 14244–14248 (2003).
67. Dloczik L, *et al*., *J. Phys. Chem. B* **101**, 10281–10289 (1997).
68. Nelson J, *Phys. Rev. B* **59**, 15374–15380 (1999).
69. Van de Lagemaat J, Frank A J, *J. Phys. Chem. B* **105**, 11194–11205 (2001).
70. Boschloo G, Hagfeldt A, *J. Phys. Chem. B* **109**, 12093–12098 (2005).
71. Kopidakis N, *et al*., *J. Phys. Chem. B* **104**, 3930–3936 (2000).
72. Nakade S, *et al*., *J. Phys. Chem. B* **105**, 9150–9152 (2001).
73. Kambe S, *et al*., *J. Phys. Chem. B* **106**, 2967–2972 (2002).
74. Nakade S, *et al*., *J. Phys. Chem. B* **107**, 8607–8611 (2003).
75. Kopidakis N, *et al*., *Appl. Phys. Lett*. **87**, 1–3 (2005).
76. Nakade S, *et al*., *J. Phys. Chem. B* **106**, 10004–10010 (2002).
77. Mori S, *et al*., *J. Phys. Chem. C* **112**, 20505–20509 (2008).
78. Haque S A, *et al*., *J. Phys. Chem. B* **104**, 538–547 (2000).
79. Nelson J, Haque S A, Klug D R, Durrant J R, *Phys. Rev. B* **63**, 2053211–2053219 (2001).
80. Bisquert J, Vikhrenko V S, *J. Phys. Chem. B* **108**, 2313–2322 (2004).
81. Bisquert J, Zaban A, Greenshtein M, Mora-Sero I, *J. Am. Chem. Soc*. **126**, 13550–13559 (2004).
82. Montanari I, Nelson J, Durrant J R, *J. Phys. Chem. B* **106**, 12203–12210 (2002).
83. Pelet S, Moser J E, Grätzel M, *J. Phys. Chem. B* **104**, 1791–1795 (2000).
84. Tatay S, *et al*., *J. Mater. Chem*. **17**, 3037–3044 (2007).
85. Clifford J N, *et al*., *J. Am. Chem. Soc*. **126**, 5225–5233 (2004).
86. Chang C-W, *et al. J. Phys. Chem. C* **113**, 11524–11531 (2009).
87. Clifford J N, *et al*., *J. Phys. Chem. C* **111**, 6561–6567 (2007).
88. Wang Z S, *et al*., *J. Phys. Chem. C* **111**, 7224–7230 (2007).
89. Kopidakis N, Benkstein K D, van de Lagemaat J, Frank A J, *J. Phys. Chem. B* **107**, 11307–11315 (2003).
90. Fisher A C, *et al*., *J. Phys. Chem. B* **104**, 949–958 (2000).
91. Gregg B A, Pichot F, Ferrere S, Fields C L, *J. Phys. Chem. B* **105**, 1422–1429 (2001).
92. Martinson A B F, Hamann T W, Pellin M J, Hupp J T, *Chemistry-a European Journal* **14**, 4458–4467 (2008).
93. Ferrere S, Gregg B A, *J. Am. Chem. Soc*. **120**, 843–844 (1998).

94. Yang M, Thompson D W, Meyer G J, *Inorg. Chem*. **39**, 3738 (2000).
95. Islam A, *et al*., *New J. Chem*. **24**, 343–345 (2000).
96. Islam A, *et al*., *Inorg. Chem*. **40**, 5371–5380 (2001).
97. Alebbi M, *et al*., *J. Phys. Chem. B* **102**, 7577–7581 (1998).
98. Kuciauskas D, *et al*., *J. Phys. Chem. B* **105**, 392–403 (2001).
99. Altobello S, *et al*., *J. Am. Chem. Soc*. **127**, 15342–15343 (2005).
100. Hasselmann G M, Meyer G J, *Zeitschrift Fur Physikalische Chemie-International Journal of Research in Physical Chemistry & Chemical Physics* **212**, 39–44 (1999).
101. Kay A, Grätzel M, *J. Phys. Chem*. **97**, 6272–6277 (1993).
102. Campbell W M, Burrell A K, Officer D L, Jolley K W, *Coord. Chem. Rev*. **248**, 1363–1379 (2004).
103. Imahori H, *et al*., *Langmuir* **22**, 11405–11411 (2006).
104. Hayashi S, *et al*., *J. Phys. Chem. C* **112**, 15576–15585 (2008).
105. Campbell W M, *et al*., *J. Phys. Chem. C* **111**, 11760–11762 (2007).
106. Nazeeruddin M K, Humphry-Baker R, Gratzel M, Murrer B A, *Chem. Commun*., 719–720 (1998).
107. He J, *et al*., *J. Am. Chem. Soc*. **124**, 4922–4932 (2002).
108. Cid J J, *et al*., *Angew. Chem., Int. Ed*. **46**, 8358–8362 (2007).
109. Li X Y, *et al*., *New J. Chem*. **26**, 1076–1080 (2002).
110. Mori S. *et al*., *J. Am. Chem. Soc.* **132**, 4045–4046 (2010)
111. Mozer A J, *et al*., *J. Am. Chem. Soc*. **131**, 15621 (2009).
112. Neale N R, *et al*., *J. Phys. Chem. B* **109**, 23183–23189 (2005).
113. Yum J H, *et al*., *Langmuir* **24**, 5636–5640 (2008).
114. Mishra A, Fischer M K R, Bauerle P, *Angew. Chem., Int. Ed*. **48**, 2474–2499 (2009).
115. Ito S, *et al*., *Adv. Mater*. **18**, 1202–1205 (2006).
116. Kim S, *et al*., *Chem. Commun*., 4951–4953 (2008).
117. Hwang S, *et al*., *Chem. Commun*., 4887–4889 (2007).
118. Ito S, *et al*., *Chem. Commun.*, 5194–5196 (2008).
119. Choi H, *et al*., *Angew. Chem., Int. Ed*. **47**, 327–330 (2008).
120. Hagberg D P, *et al*., *J. Am. Chem. Soc*. **130**, 6259–6266 (2008).
121. Wang M K, *et al*., *Adv. Mater*. **20**, 4460–4463 (2008).
122. Liu W H, *et al*., *Chem. Commun*., 5152–5154 (2008).
123. Wang Z S, *et al*., *Chem. Mater*. **20**, 3993–4003 (2008).
124. Zhang G L, *et al*., *Chem. Commun*., 2198–2200 (2009).
125. Horiuchi T, Miura H, Uchida S, *Chem. Commun*. **9**, 3036–3037 (2003).
126. Horiuchi T, Miura H, Sumioka K, Uchida S, *J. Am. Chem. Soc*. **126**, 12218–12219 (2004).
127. Kuang D, *et al*., *Angew. Chem., Int. Ed*. **47**, 1923–1927 (2008).
128. Koumura N, *et al*., *J. Am. Chem. Soc*. **128**, 14256–14257 (2006).
129. Wang Z S, *et al*., *Adv. Mater*. **19**, 1138–1141 (2007).
130. Hara K, Miyamoto K, Abe Y, Yanagida M, *J. Phys. Chem. B* **109**, 23776–23778 (2005).
131. O'Regan B C, *et al*., *J. Am. Chem. Soc*. **130**, 2906–2907 (2008).
132. Mozer A J, *et al*., *Chem. Commun*., 4741–4743 (2008).
133. Hagberg D P, *et al*., *Journal of Organic Chemistry* **72**, 9550–9556 (2007).
134. Marinado T, *et al*., *Langmuir* **26**, 2592–2598 (2010).
135. Law M, *et al*., *Nature Materials* **4**, 455–459 (2005).
136. Uchida S, *et al*., *Electrochemistry* **70**, 418–420 (2002).
137. Adachi M, *et al*., *Electrochemistry* **70**, 449–452 (2002).
138. Martinson A B F, McGarrah J E, Parpia M O K, Hupp J T, *Phys. Chem. Chem. Phys.* **8**, 4655–4659 (2006).
139. Ravirajan P, *et al*., *J. Phys. Chem. B* **110**, 7635–7639 (2006).
140. Hamann, T. W. *et al*. Energy & Enviromental Science 66–78 (2008).

141. Quintana M, Edvinsson T, Hagfeldt A, Boschloo G, *J. Phys. Chem. C* **111**, 1035–1041 (2007).
142. Islam A, *et al*., *Inorganica Chimica Acta* **322**, 7–16 (2001).
143. Kay A, Grätzel M, *Chem. Mater*. **14**, 2930–2935 (2002).
144. Green A N M, *et al*., *J. Phys. Chem. B* **109**, 12525–12533 (2005).
145. Fukai Y, Kondo Y, Mori S, Suzuki E, *Electrochem. Commun.* **9**, 1439–1443 (2007).
146. Iwamoto S, *et al*., *Chemsuschem* **1**, 401–403 (2008).
147. He J, Lindström H, Hagfeldt A, Lindquist S E, *J. Phys. Chem. B* **103**, 8940–8943 (1999).
148. Mori S, *et al*., *J. Phys. Chem. C* **112**, 16134–16139 (2008).
149. He J, Lindström H, Hagfeldt A, Lindquist S E, *Sol. Energy Mater. Sol. Cells* **62**, 265–273 (2000).
150. Papageorgiou N, *et al*., *J. Electrochem. Soc*. **143**, 3099–3108 (1996).
151. Matsumoto H, *et al*., *Chem. Lett*., 26–27 (2001).
152. Kubo W, *et al*., *J. Phys. Chem. B* **107**, 4374–4381 (2003).
153. Kawano R, *et al*., *J. Photochem. Photobiol., A: Chem*. **164**, 87–92 (2004).
154. Wang Z S, *et al*., *Chem. Mater*. **21**, 2810–2816 (2009).
155. Nusbaumer H, *et al*., *J. Phys. Chem. B* **105**, 10461–10464 (2001).
156. Hattori S, Wada Y, Yanagida S, Fukuzumi S, *J. Am. Chem. Soc*. **127**, 9648–9654 (2005).
157. Bach U, *et al*., *Nature* **395**, 583–585 (1998).
158. Bach U, *et al*., *J. Am. Chem. Soc*. **121**, 7445–7446 (1999).
159. Schmidt-Mende L, *et al*., *Adv. Mater*. **17**, 813–815 (2005).
160. Kruger J, Plass R, Gratzel M, Matthieu H J, *Appl. Phys. Lett*. **81**, 367–369 (2002).
161. Handa S, Haque S A, Durrant J R, *Adv. Funct. Mater.* **17**, 2878–2883 (2007).
162. Tennakone K, *et al*., *J. Photochem. Photobiol., A: Chem.* **117**, 137–142 (1998).
163. Tennakone K, *et al*., *Journal of Physics D-Applied Physics* **31**, 1492–1496 (1998).
164. Tennakone K, Perera V P S, Kottegoda I R M, Kumara G, *Journal of Physics D-Applied Physics* **32**, 374–379 (1999).
165. Oregan B, Schwartz D T, *Journal of Applied Physics* **80**, 4749–4754 (1996).
166. Kumara G R R A, *et al*., *Sol. Energy Mater. Sol. Cells* **69**, 195–199 (2001).
167. Murakoshi K, Kogure R, Wada Y, Yanagida S, *Sol. Energy Mater. Sol. Cells* **55**, 113–125 (1998).
168. Tennakone K, *et al*., *Chem. Mater*. **11**, 2474–2477 (1999).
169. Saito Y, Kitamura T, Wada Y, Yanagida S, *Synthetic Metals* **131**, 185–187 (2002).
170. Mozer A J, *et al*., *Appl. Phys. Lett*. **89**, 043509–043511 (2006).
171. Ikeda N, Miyasaka T, *Chem. Commun*., 1886–1888 (2005).
172. Kubo W, *et al*., *Chem. Lett*., 1241–1242 (1998).
173. Sakaguchi S, *et al*., *J. Photochem. Photobiol., A: Chem*. **164**, 117–122 (2004).
174. Usui H, Matsui H, Tanabe N, Yanagida S, *J. Photochem. Photobiol., A: Chem*. **164**, 97–101 (2004).
175. Komiya R, *et al*., *J. Photochem. Photobiol., A: Chem*. **164**, 123–127 (2004).
176. Kohle O, Gratzel M, Meyer A F, Meyer T B, *Adv. Mater*. **9**, 904 (1997).
177. Grunwald R, Tributsch H, *J. Phys. Chem. B* **101**, 2564–2575 (1997).
178. Nguyen H T, Ta H M, Lund T, *Sol. Energy Mater. Sol. Cells* **91**, 1934–1942 (2007).
179. Amirnasr M, Nazeeruddin M K, Gratzel M, *Thermochimica Acta* **348**, 105–114 (2000).
180. Hara K, *et al*., *Sol. Energy Mater. Sol. Cells* **77**, 89–103 (2003).
181. Katoh R, *et al*., *Energy & Environmental Science* **2**, 542–546 (2009).
182. Kern R, *et al*., *Opto-Electron. Rev*. **8**, 284–288 (2000).
183. Olsen E, Hagen G, Eric Lindquist S, *Sol. Energy Mater. Sol. Cells* **63**, 267–273 (2000).
184. Sommeling P M, *et al*., *J. Photochem. Photobiol., A: Chem*. **164**, 137–144 (2004).
185. Toyoda T, *et al*., *J. Photochem. Photobiol., A: Chem*. **164**, 203–207 (2004).
186. Okada K, *et al*., *J. Photochem. Photobiol., A: Chem*. **164**, 193–198 (2004).

187. Hanke K P Upscaling of the dye sensitized solar cell, *12th International Conference on Photochemical Conversion and Storage of Solar Energy*, Berlin, Germany, pp 1–9 (August 9–14 1998).
188. J. Kroon A H, *Dye-sensitized solar cells*. Springer: Heidelberg, pp 273–290 (2003).
189. Smestad G P, *Sol. Energy Mater. Sol. Cells* **55**, 157–178 (1998).
190. Lindström H, *et al*., *Nano Lett*. **1**, 97–100 (2001).
191. Uchida S, Tomiha M, Takizawa H, Kawaraya M, *J. Photochem. Photobiol., A: Chem*. **164**, 93–96 (2004).
192. Ito S, *et al*., *Chem. Commun*., 4004–4006 (2006).
193. Kijitori Y, Ikegami M, Miyasaka T, *Chem. Lett*. **36**, 190–191 (2007).

16

Sunlight Energy Conversion Via Organics

Sam-Shajing Sun[1] and Hugh O'Neill[2]

[1]*Chemistry Department and PhD Program in Materials Science & Engineering, Norfolk State University, Virginia, USA,* [2]*Center for Structural Molecular Biology, Chemical Sciences Division, Oak Ridge National Lab, Tennessee, USA*

16.1 PRINCIPLES OF ORGANIC AND POLYMERIC PHOTOVOLTAICS

16.1.1 Introduction

Inorganic crystalline semiconductor-based solar cell technology has become relatively mature, between 10–30% photoelectric power conversion efficiencies at AM 1.5 are readily available from commercial monolithic crystalline solar cell modules [1–4], over 40% efficiency has been demonstrated in crystalline multijunction tandem cells [4], and much higher theoretical efficiencies have been predicted [3]. That technology, however, suffers from a relatively high cost and a current shortage of the feedstock materials such as high quality silicon crystalline wafers. As the demand for renewable and clean solar energies rapidly grows, alternative materials or technologies that could reduce the solar cell costs and have sufficient feedstock supplies become attractive. While a variety of amorphous and thin film inorganic solar cells are being developed, industrial manufacturing cost and toxicity are still challenges [1, 2, 4]. Organic and polymeric photovoltaic materials and technology is another very attractive option [5–7].

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

Compared with their inorganic counterparts, recently developed organic and polymeric conjugated semiconducting materials appear very promising for photovoltaic applications due to several reasons:

- Lightweight, low consumption of materials (i.e. films can be very thin), flexible shape, versatile materials synthesis and device fabrications, and low cost for large scale industrial production.
- Almost continuous tunability of materials' energy levels and gaps via molecular design, synthesis, and processing.
- Integrability into other products such as textiles (that can be used for fabricating items such as clothing and solar cell tents), flexible packaging systems, lightweight consumer goods, and future "all-plastic" optoelectronic devices that may be compatible with biological tissues [8, 9].

16.1.2 Organic versus Inorganic Optoelectronics Processes

In organic π electron conjugated materials, the outer shell or valence π electrons are typically responsible for the electronic and optoelectronic properties (see chapter 3 in reference [8]). When a material rests at its lowest ground state, the highest occupied molecular orbital (HOMO) typically refers to a highest energy level and fully occupied electron bonding orbital, and the lowest unoccupied molecular orbital (LUMO) typically refers to a lowest energy level empty antibonding orbital, and a singly occupied molecular orbital is called a SOMO. HOMOs, LUMOs and SOMOs are also termed *frontier orbitals*. In typical organic semiconductors, including most organic crystalline semiconductors, the intermolecular electronic orbital coupling (also refers to orbital spatial and energetic overlap, or simply 'overlap', and mathematically expressed as an electronic coupling matrix element) are generally much poorer compared with their inorganic semiconductor counterparts. This is because typical organic semiconductors are molecular or amorphous conjugated materials with large intermolecular spacing and relatively random orientations, while most inorganic semiconductors are closely packed and ordered atomic crystals. Additionally, in inorganic semiconductors, strong orbital coupling or overlap is mostly on the atomic level and in the large periodic structural range forming electronic conduction bands (CBs) and valence bands (VBs) with substantial bandwidths (Bloch theorem) [1–4] (see chapter 3 in reference [8]). On the contrary, in most organic semiconductors, orbital overlap and coupling is mostly on the molecular levels, i.e. molecular shape or packing directly restricts or limits the intermolecular orbital coupling or band formation [7–9]. Therefore, stable conduction bands (CBs) and valence bands (VBs) with substantial bandwidth (i.e. over 0.1 eV) are rare in organic semiconductors (see Chapter 1 in [9]). The optical excitation energy gap E_g in organic semiconductors typically represent the smallest energy difference between discrete LUMO and HOMO orbitals, and free charge carriers will transfer or "hop" among different orbitals or sites instead of transporting in "bands". "Band like" organic semiconductors are rare [10]. In most organic semiconductors, when a photon with energy matching the E_g excites an organic molecule, an electron first transfers from the HOMO to the LUMO (i.e. interorbital electron transfer [5, 7, 8, 17, 18] and then quickly relaxes (reorganizes) with the hole to form a tightly bound electron–hole pair called a Frenkel exciton [11]). Figure 16.1 shows a schematic comparison between a Frenkel exciton versus a Wannier–Mott type exciton (common in inorganic semiconductors) [12] (see chapters 1 and 3 in reference [8]).

Figure 16.2 exhibits Coulombic potentials E_c versus exciton size or radius r. Figure 16.3 exhibits free energy surfaces of a ground state (S_0), first excited state (S_1), and the exciton dissociated state (S_1'). As Figure 16.1 shows, the Frenkel exciton typically has a size of less than 1 nm with binding energies typically larger than 0.1 eV [7, 11]. This is partly because the exciton binding energy E_B is proportional to the Coulombic attractive potential E_c($E_B = E_c + \lambda_2$, see Figure. 16.3, where λ_1 is the reorganization/relaxation energy of photoexcited exciton, λ_2 is the reorganization energy of exciton dissociation), and E_c is inversely proportional to the radius r between the electron

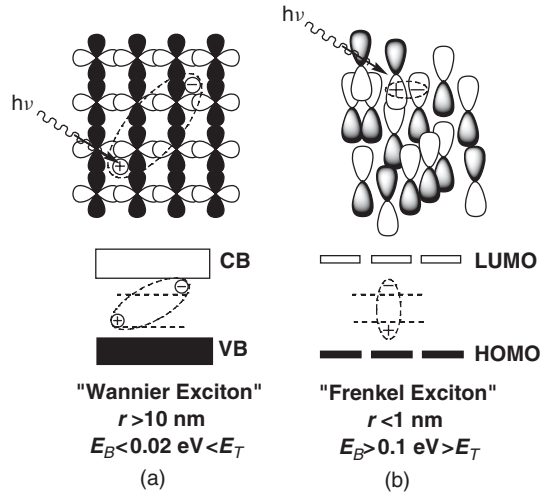


Figure 16.1 Inorganic Wannier–Mott (a) and organic Frenkel (b) excitons

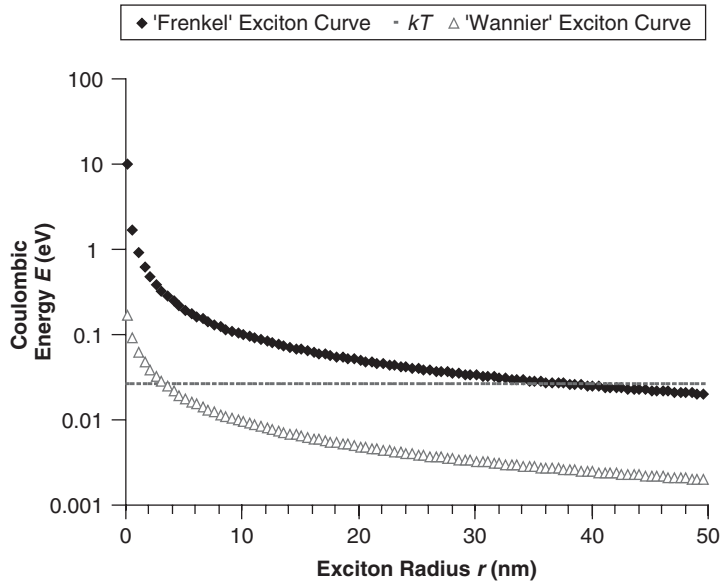


Figure 16.2 Schematic Coulombic potential vs size of Wannier (triangle) and Frenkel (diamond) type excitons. The horizontal dashed line represents room temperature thermal energy kT

and hole ($E_c = -e^2/4\pi\varepsilon\varepsilon_o r$) or the exciton size (see Figures 16.2 and 16.3). E_B is the minimum energy required to dissociate a neutral exciton particle into a free or uncorrelated hole particle, called a positive polaron, and a free or uncorrelated electron particle, called a negative polaron, both are also called charged carriers, or simply carriers (in organic or molecular materials, charge carriers are called polarons as they incur substantial electronic polarization or distortion in their surrounding lattice). Because of this, typical room-temperature thermal energy $E_T = kT$ (less than

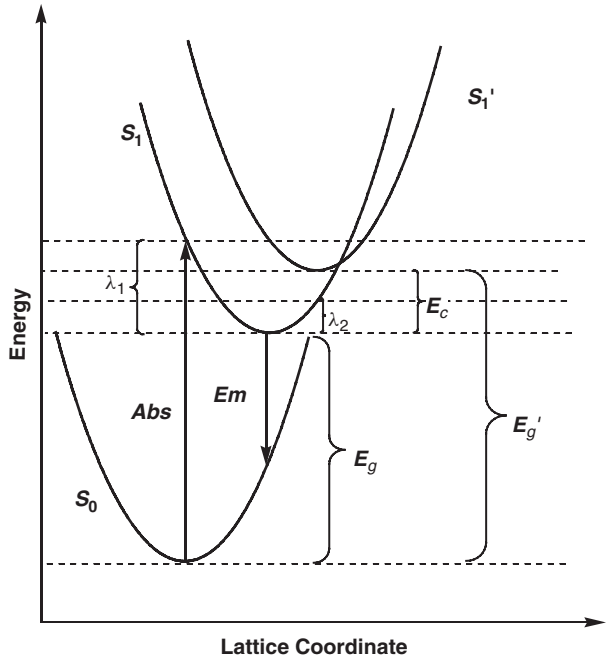


Figure 16.3 Schematic free energy surfaces of ground state (S_0), exciton state (S_1), and charge separated state (S_1')

0.03 eV) would not be sufficient to dissociate a Frenkel exciton. Hence, uncorrelated free charge carriers generated directly by an initial photoexcitation, i.e. valence band (VB) to conduction band (CB) free carrier photogeneration (also called "primary photo carrier generation mechanism" as in typical inorganic semiconductors) are rare in organic semiconductors [7–9]. The band model (or Bloch theorem) was derived based on a perfect or ideal periodic repeating potential structure with strong interatomic (or intermolecular) orbital coupling [7, 8] yet, even in perfect organic crystals, inter-molecular orbital coupling can be weak due to large inter-molecular distance or spacing [10].

In contrast, Wannier (or Wannier–Mott) excitons are loose electron–hole pairs with sizes typically greater than 10 nm and with binding energy typically less than 0.03 eV ([12], see chapter 1 in reference [8]), i.e., room temperature thermal energy ($kT \sim 0.03$ eV) would be sufficient for these excitons to dissociate into free charge carriers. Whether Frenkel or Wannier excitons are generated in a particular semiconductor depends on the frontier orbital coupling, structure periodicity, and dielectric constants of the materials. For instance, certain organic crystalline semiconductors exhibit band like transport or Wannier exciton characteristics with an increase of conductivity as temperature decreases in the low temperature regime, and a charge mobility of over 20 cm^2/Vs has been observed in organic crystal rubrene [10]. Since room temperature thermal energy kT is sufficient to dissociate the Wannier excitons, a photoexcitation at room temperature can easily generate a free (or uncorrelated) electron at the CB and a free (or uncorrelated) hole at the VB instantly, i.e, a "primary photocarrier generation" mechanism applies well to Wannier type excitons (Figures. 16.1–3). In the case of the Frenkel type of exciton, as in most organic semiconductors, room temperature thermal energy kT is not sufficient to dissociate the exciton, so extra energy (or a secondary driving force) is required in order to dissociate the Frenkel exciton. Such a secondary driving force can be simply an electric field formed at an organic donor/acceptor interface or junction due to their frontier orbital energy offsets. This is called the "secondary photocarrier generation" mechanism [7] (see chapter 3

in reference [8]). As a matter of fact, the initial sunlight harvesting steps of photosynthesis in most natural plants follows this secondary process [13] and will be further elaborated in Section 16.2.

16.1.3 Organic/Polymeric Photovoltaic Processes

The overall photovoltaic process of an organic/polymeric solar cell can be divided into at least following five critical steps:

1. Photon absorption and exciton generation
2. Exciton diffusion to a donor/acceptor interface
3. Exciton dissociation or charge carrier generation at donor/acceptor interface
4. Carrier transport toward respective electrodes
5. Carrier collection at the respective electrodes

For all currently reported organic/polymeric photovoltaic materials and devices, none of these five steps have been optimized. Therefore, it is not surprising that the reported power conversion efficiencies of organic or polymeric solar cells are relatively low (less than 7%) in comparison with their inorganic counterparts (typically more than 10%).

16.1.3.1 Photon absorption and exciton generation

A basic requirement of organic light harvesting system, which also applies to photosynthesis of natural plants (see Section 16.2), is that the material's optical excitation energy gap E_g must match the incident photon energy. As discussed earlier, in most organic materials, the energy gap defaults to the minimum energy difference between the HOMO and the LUMO. Since a Frenkel exciton is typically generated during an excitation of an organic semiconductor (i.e. an electron transfers from its HOMO to its LUMO), the optical energy gap E_g is therefore used instead of the conventional electronic energy gap (similar to E_g' as shown in Figure 16.3) that typically refers to the energy gap between the minimum or lowest point of the conduction band (CB) and the maximum or the highest point of valence band (VB) in inorganic semiconductors. If valence band VB is defined as containing "free" holes, and CB is defined as containing "free" electrons, then for a donor/acceptor binary organic system, the self-organized or well-coupled acceptor LUMO "bands" may then be called CB, and self-organized and well-coupled donor HOMO "bands" may then be called VB. Unfortunately, due to the poor or weak intermolecular orbital couplings, the "bands" are hard to form in organics, and the charge mobility is mainly due to "hopping" transport mechanism instead of "band like" transport.

For solar cell applications, solar radiation spans a broad range of light radiation with most intense photon flux between 600 and 1000 nm (1.3–1.8 eV at air mass 1.5 on surface of the earth) or 400–700 nm (1.8–3.0 eV, at air mass zero in space) [1–4]. For terrestrial applications, it is desirable that the energy gaps of a solar cell span a range from 1.3 to 1.8 eV. This may be achieved by an energy gap graded tandem style cell structure as will be discussed in a later section. So far, several popular conjugated semiconducting polymers widely used in organic solar cell studies have energy gaps higher than 1.8 eV [5–7]. For instance, several widely used alkyloxy derivatized poly-*p*-phenylenevinylenes (RO-PPV) has a typical energy gap of 2.3–2.5 eV, and several representative polythiophenes (including the most popular P3HT) have energy gaps between 1.8 and 2.0 eV with $1/\alpha$ values in a range of $1-2 \times 10^5\ \mathrm{cm}^{-1}$ (see chapter 23 in reference [5]), are all well above the maximum solar photon flux range of 1.3–1.8 eV. Although in certain inorganic PV materials, a high-energy photon may generate more than one exciton due to an Auger generation mechanism [14], so far no such mechanisms have been reported in organic PV systems to date. Also, due to the

generally low charge mobility or high resistance of amorphous organics, very thin films are generally desired. Therefore, a fraction of energy-matched photons may pass through the materials without being captured. This is why the photon absorption (or exciton generation) for PPV-based polymer solar cells are far from being optimized at AM 1.5. This 'photon loss' problem is in fact very common in most of the currently reported organic/polymeric photovoltaic materials and devices. The best way to capture most photons and at the same time not suffer the photovoltage loss appears to be a tandem cell structure which will be described in a later section. Also, one advantage of organic materials is the tunability of their energy levels via molecular design and synthesis [7–9]. Therefore, ample opportunity exists for improvement. Recent developments on lower-energy-gap conjugated polymers are such examples [15] (also see chapter 4 in reference [6]).

16.1.3.2 Exciton diffusion to donor/acceptor interfaces

Once a Frenkel exciton is photogenerated in either donor or acceptor phase, it typically diffuses (e.g. via intra- or intermolecular energy transfer or "hopping" processes, including Förster or Dexter energy transfer processes) to an adjacent or remote site [60, 61] (see chapter 3 in reference [8]). At the same time, the exciton will decay either radiatively (i.e. via photoluminescence (PL)), or nonradiatively to its ground state (i.e. an electron transfer from its LUMO back to HOMO) with typical lifetimes between picoseconds to nanoseconds [5–9]. Alternatively, in the solid state, some excitons may be trapped in defect or impurity sites to become a stable charged pair and then decay slowly or nonradiatively. Both exciton decay and trapping can contribute to the so-called exciton loss. The average distance an organic exciton can travel within its lifetime is called average exciton diffusion length (AEDL) [7]. In noncrystalline or amorphous materials, the AEDL depends heavily on the spatial property (i.e. morphology) of the materials. For most conjugated polymeric materials, the AEDL is typically in the range of 5–50 nm [5–9]. For instance, the AEDL of PPV is around 5–10 nm [9, 16]. Since the desired second step of photovoltaic process is that each photogenerated Frenkel exciton will be able to reach the donor/acceptor interface where exciton dissociation (charge separation) can occur, one strategy to minimize the exciton loss is to increase the AEDL by making a defect-free and large donor/acceptor interface (or easily accessible interface) morphology.

16.1.3.3 Exciton dissociation and charge carrier generation at donor/acceptor interfaces

When an exciton reaches a donor/acceptor interface, the interfacial potential field formed due to the donor/acceptor frontier orbital energy offsets, i.e. δE in Figure 16.4, will dissociate the exciton into a free electron at the acceptor LUMO and a free hole at the donor HOMO, provided the magnitude of this field or energy offset is such that the charge separation falls into the optimal electron transfer regime as defined by Marcus theory, and the electronic coupling matrix element for such interfacial electron transfer is sufficiently strong [7, 17, 18] (also see chapter 1 in reference [9]). This photoinduced interface charge separation process is also called 'photo-doping', because it is a photoinduced (in contrast to chemical or thermal induced) redox reaction between a donor and an acceptor [7–9]. It has been experimentally observed that such photoinduced charge separation process at a PPV/fullerene interface is orders of magnitude faster than either the PPV exciton decay or the separated charge recombination [19, 20]. This means the quantum efficiency at this interface is near unity, and a high efficiency organic photovoltaic system is feasible with improvement of other factors. As illustrated in Figure 16.4, process #1 is donor photo excitation or exciton formation, process #2 is donor exciton decay, process #3 is donor exciton dissociation (or charge transfer) at donor/acceptor interface, process #4 is separated charge recombination, process #5 is acceptor excitation, process #6 is acceptor exciton decay, process #7 is electron transfer at donor/acceptor interface after the acceptor excitation, and process #8 is hole transfer at the interface after the

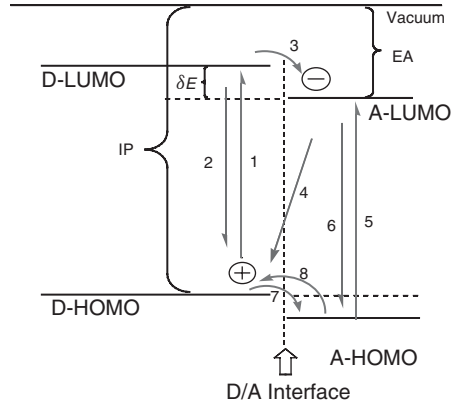


Figure 16.4 Scheme of molecular frontier orbitals and photoinduced electron and Dexter energy transfer processes in a donor/acceptor binary light harvesting system for various processes defined in the text

acceptor excitation (note process #8 is the same as process #7 in different presentations). Processes 1 and 5 are typical photo excitation processes and the efficiency can be near unity for all energy matched photons with materials of sufficient density or thickness. Processes 3, 7 (or 8) are charge transfers (or exciton dissociations) at the donor/acceptor interface, and their efficiency (in reference to exciton decay) can be near unity at optimal conditions, i.e. optimal energy offset and large electronic coupling of charge transfer [7, 17, 18, 20] (also see chapter 3 in reference [8]). Processes 2 and 6 are exciton decay and their efficiencies (in reference to exciton dissociation) are highly dependent upon morphology of the materials, e.g. the larger the domain size or more difficult for an exciton to reach a donor/acceptor interface, the more severe exciton decays. Process 4 is the recombination of separated charges and its efficiency is sensitive to the electronic coupling and energetic factors of charge recombination process, and they can be orders of magnitude slower than the exciton dissociation processes [20]. If both electron and hole are transferred simultaneously from a donor to an acceptor at the interface, then it is a Dexter energy transfer process instead of charge transfer [61] (see chapter 3 in reference [8]).

16.1.3.4 Carrier transport toward the electrodes

Once the carriers are generated at the donor/acceptor interface, the holes need to transport toward the large work function electrode (LWFE, positive electrode collecting holes), and electrons need to transport toward the small work function electrode (SWFE, negative electrode collecting electrons). The driving forces for the carrier transport mainly include the field created by the work function difference between the two electrodes, energetic hopping sites, and a "chemical potential" gradient [21]. The "Chemical potential" gradient driving force may be regarded as a particle density potential driving force, i.e. particles tend to diffuse from a higher density domain to a lower density domain due to thermal dynamics. In an organic donor/acceptor binary photovoltaic cell, the high-density electrons at the acceptor LUMO nearby the donor/acceptor interface tend to diffuse toward lower electron density region within the acceptor phase, and high-density holes at the donor HOMO nearby the donor/acceptor interface tend to diffuse toward the lower hole density region within the donor phase. For instance, in donor/acceptor double-layered Tang cell [22] as described in following sections, once an exciton is dissociated into an electron at acceptor side and a hole at donor side of the D/A interface, the electron would be "pushed" away from the interface

toward the SWFE negative electrode by both the "chemical potential" and by the field generated from the two electrode work functions. The holes would be "pushed" toward the LWFE positive electrode by the same forces but in the opposite direction. With this chemical potential driving force, even if the two electrodes are the same, asymmetric photovoltage could still be achieved (i.e. the donor HOMO would generate the positive and acceptor LUMO would generate the negative electrodes) [22]. Energetic hopping sites, such as many mid-gap state species including impurities, defects, or intentionally doped redox species, also facilitate the carrier transport or conductivities by providing 'shallow' hopping orbital sites for the electrons or holes to hop [7–9]. After the electron–hole pair is separated at the D/A interface, they can also recombine due to the Coulombic force (or potential drop of A-LUMO/D-HOMO) between the free electron and hole. Fortunately, the charge recombination rates in most reported cases are much slower (typically by orders of magnitudes) than the charge separation rates (typically in femto/pico seconds) [5–9, 19, 20], so there are sufficient opportunities for the carriers to reach the electrodes before they recombine. However, in most currently reported organic solar cells, the transport of electrons and holes to their respective electrodes is not smooth due to poor morphology or non-bicontinous nature of carrier transport pathways. If donor and acceptor phases are perfectly bicontinuous between the two electrodes, and all LUMO and HOMO orbitals are well aligned and overlapped to each other in both donor and acceptor phases, as in a highly self-assembled thin films or crystals, then the carriers would be able to transport smoothly toward their respective electrodes, and a very high efficiency solar cell could be achieved [7]. Currently, carrier thermal hopping is believed to be the dominant conductivity mechanism for most reported organic photovoltaic systems. Therefore, the "carrier loss" is believed to be another key contributing factor for the low power conversion efficiency of organic photovoltaics.

16.1.3.5 Carrier collection at the electrodes

It was proposed [23] that when the acceptor LUMO level matches the Fermi level of the SWFE, and the donor HOMO matches the Fermi level of the LWFE, an ideal ohmic contact would be established for efficient carrier collection at the electrodes. So far, there are no organic/polymeric photovoltaic cells that have realized this desired ohmic alignment due to the availability and limitations of materials and electrodes involved. It is also not clear such match would be the optimal situation, particularly when reorganization energies are involved [7, 17, 18]. There were a number of studies, however, focusing on the open-circuit voltage (V_{oc}) dependence on LUMO/HOMO level changes, electrode Fermi levels, and chemical potential gradients [21, 24]. The carrier collection mechanisms at electrodes are relatively less studied and are not well understood. It is possible that the carrier collection losses (including the carrier recombination at the electrodes) are other critical factors contributing to the low power conversion efficiency of most existing organic solar cells.

16.2 EVOLUTION AND TYPES OF ORGANIC AND POLYMERIC SOLAR CELLS

16.2.1 Single-layer Organic Solar Cells (Schottky Cells)

The first inorganic single-layer photovoltaic cell was developed by Charles Fritts in 1885 [25]. As illustrated in Figure 16.5a, the Fritts cell was composed of a semiconducting selenium thin layer sandwiched between two different metal electrodes, one very thin and semitransparent gold layer acting as a LWFE to collect photogenerated positive charges (holes), and the other copper layer acting as a SWFE to collect photogenerated negative charges (electrons). In this cell, when an energy

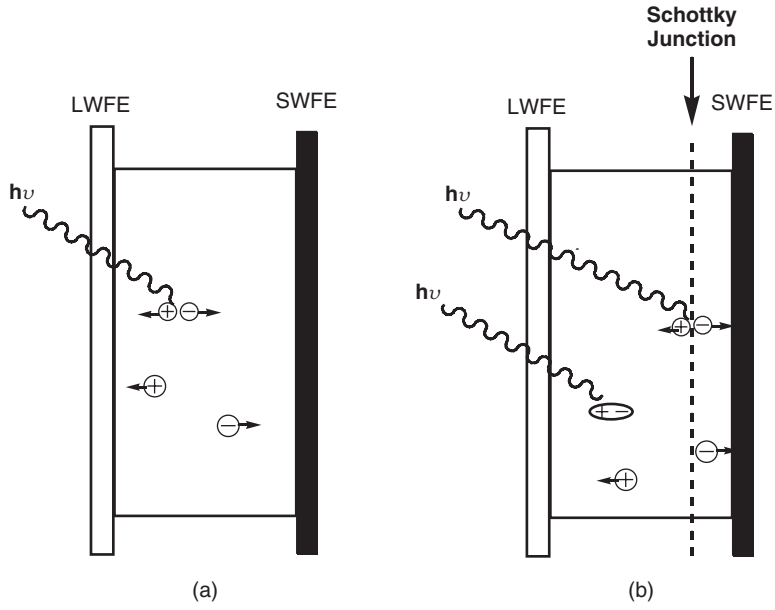


Figure 16.5 Schematic comparison of a classic (a) inorganic single layer and (b) organic single layer type photovoltaic cell

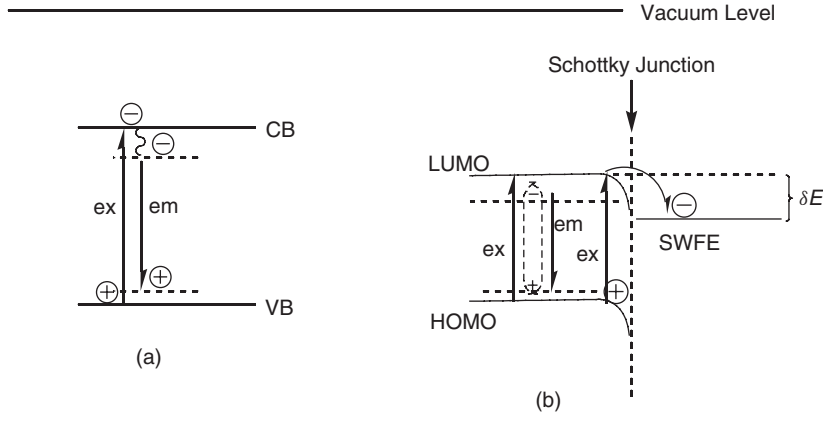


Figure 16.6 Energy level schemes for photon-induced excitation (ex) and emission (em) of (a) inorganic single layer cell (isolated layer, no band bending) and (b) organic single layer type photovoltaic cell with a small work function electrode (SWFE) metal, such as aluminum

matched photon strikes the selenium, a loosely bound Mott–Wannier exciton is generated first with subsequent dissociation of electron from the hole by room-temperature thermal energy kT, where the free electron moves in the conduction band (CB), and the free hole moves in the valence band (VB) of selenium, as illustrated in Figure 16.6a. The free electrons and holes can then diffuse to their respective electrodes driven by the field created between the two different work function metal electrodes. The Fritts cell had an overall photoelectric power conversion efficiency of about 1% [25].

In contrast, in the early organic single layer photovoltaic/photodiode cells such as Pochettino cell [26] as shown in Figures 16.5b and 16.6b, when an energy-matched photon strikes most part of the organic layer, only a strongly bound Frenkel exciton is generated that typically decays back to the ground state within its lifetime of nano-or picoseconds. Within such a lifetime regime, most Frenkel excitons can only travel 5–50 nm (average exciton diffusion length or AEDL) in most conjugated polymers, much less than the film thickness of typically over 100 nm, therefore, most excitons are lost. However, for those Frenkel excitons that are generated at or diffused into the Schottky junction area nearby the organic/SWFE interface (Figure 16.5b), the exciton can be dissociated into a free electron and a free hole by a field formed as a result of the orbital bending in the Schottky junction [1, 5]. It is believed that band bending is mainly due to impurity or defect sites near the organic/metal interface as even oxygen exposure enhances photovoltaic effects (see chapter 9 in reference [1]). Those single-layer organic solar cells had very small photoelectric power conversion efficiencies (typically less than 0.01%). They were also called Schottky cells because the carriers were mainly generated in the Schottky junction. The single-layer solar cells or Schottky cells may be categorized as the first generation organic solar cells.

16.2.2 Double-layer Donor/Acceptor Heterojunction Organic Solar Cells (Tang Cells)

From the spatial (geometry) structure point of view, the second generation solar cells were of the p/n junction or donor/acceptor double layer cells. In inorganic solar cells, it was represented by a double-layer p/n junction, as first demonstrated by Pearson *et al*. at Bell labs in 1954 [1–4]. As shown in Figures 16.7a and 16.8a, in a typical p/n junction inorganic solar cell, the photogenerated free electrons and holes were effectively separated into free electrons at n side and free holes at p side due to a p/n junction electric field. Most importantly, the separated electrons and holes are moving in separate domains (hole as majority carrier in the p-domain, and electron as majority carrier in the n-domain), and the likelihood of carrier recombination loss is much smaller. Also,

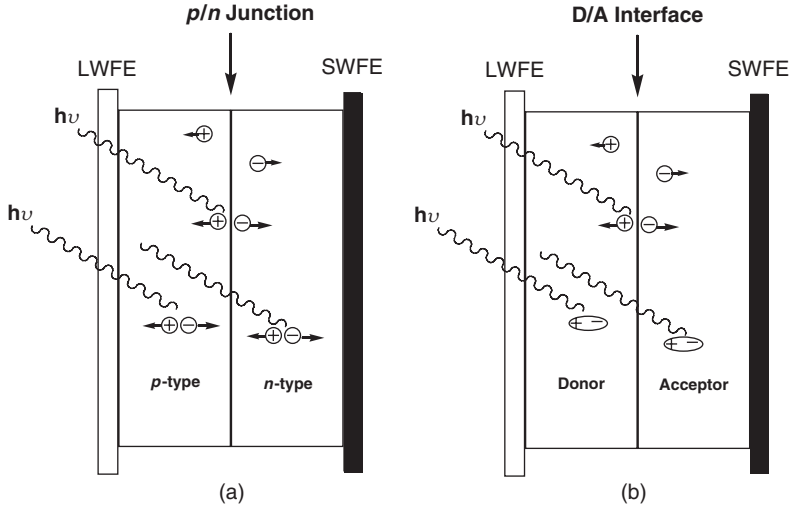


Figure 16.7 Schematic comparison of photocarrier generation processes of (a) classic inorganic p/n junction bilayer solar cells with free holes and electrons and (b) organic D/A junction bilayer type solar cells with excitons as (+-)

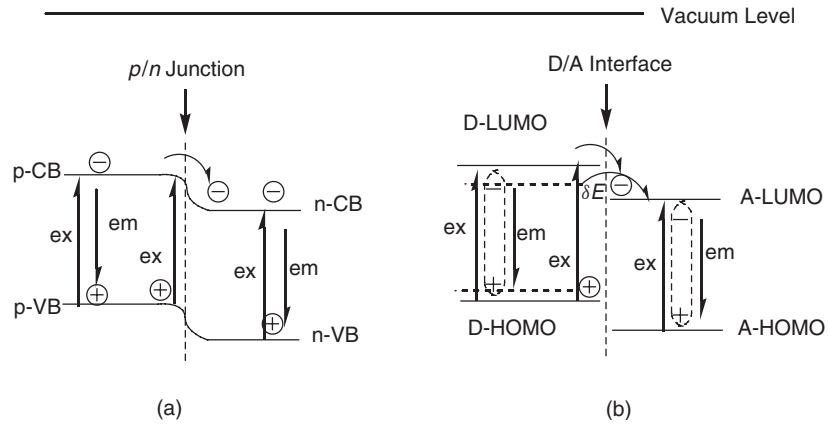


Figure 16.8 Energy level schemes of (a) inorganic p/n bilayer and (b) organic D/A bilayer type photovoltaic cells in open circuit voltage mode

due to the much higher density of charge carriers nearby the p/n junction than in the bulk, the asymmetric chemical potential force also "push" the carrier diffusion to their respective electrodes. Similarly, a major milestone in organic solar cell development was demonstrated by C. W. Tang at Kodak in the early 1980s using an organic electron donor/acceptor double-layer structure [22]. As shown in Figures 16.7b and 16.8b, once a photogenerated Frenkel exciton in either the donor or acceptor layer diffuses to the D/A interface, charge separation would occur where the electrons will transfer to or remain in the acceptor LUMO, and holes will transfer to or remain in the donor HOMO. Due to both the electrode's induced internal field and chemical potential driving forces, the electrons and holes would hop to their respective electrodes much more easily and quickly than in the single layered cells. The likelihood of charge recombination is much smaller than in the single-layer cells because electrons and holes move in two separate domain layers. Since the successful demonstration of the Tang cell, the organic and polymeric photovoltaic field started to grow rapidly as new organic/polymeric donors and acceptors were researched extensively. Figure 16.9 shows

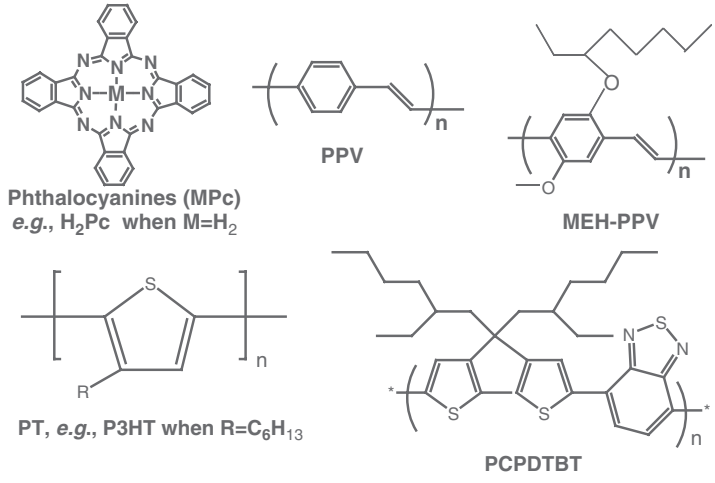


Figure 16.9 Representative organic/polymeric electron donors (p-type semiconductors)

chemical structures of some representative organic/polymeric donors, Figure 16.10 shows chemical structures of some representative organic/polymeric acceptors, and Figure 16.11 illustrates frontier orbital levels of some commonly used organic/polymeric electron donors and acceptors [7, 8]. The work functions of key electrode materials are also shown in Figure 16.11. Note that, donors and acceptors are relative. For instance, phthalocyanine H_2Pc is a donor in reference to perylene

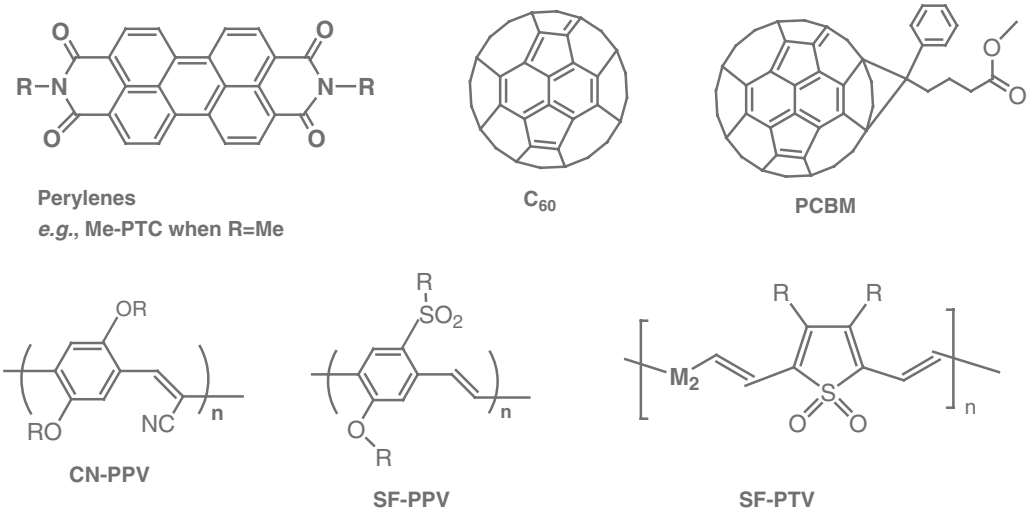


Figure 16.10 Representative organic/polymeric electron acceptors (*n*-type semiconductors)

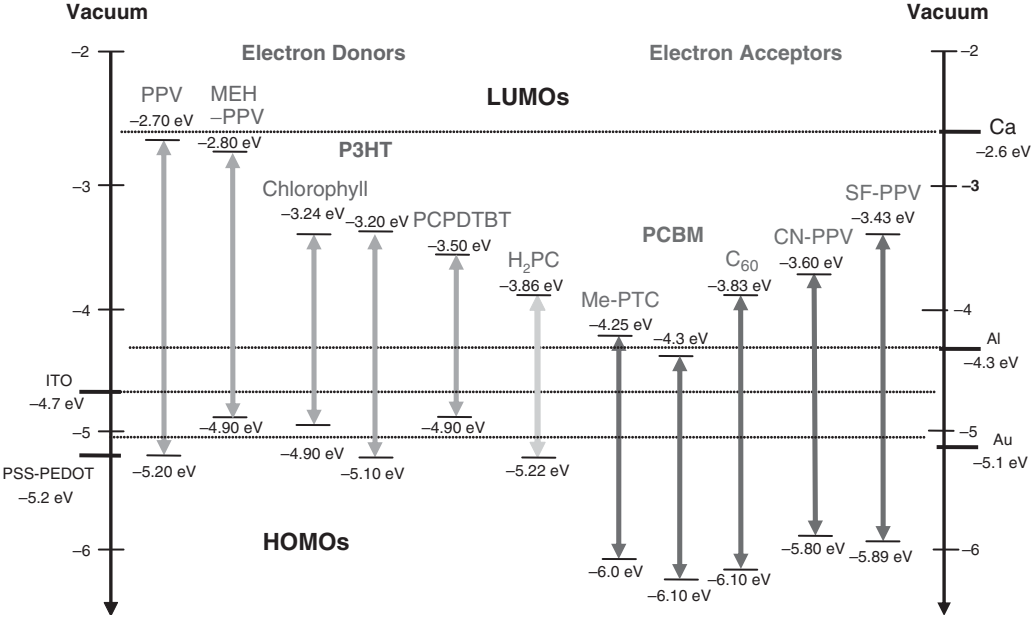


Figure 16.11 Frontier orbital levels of representative organic/polymeric electron donors and acceptors. The work functions of several representative electrodes are also shown on the sides

Me-PTC, but an acceptor in reference to MEH-PPV. However, a major limiting factor for the double-layer cell is still the relatively thick material layer (typically thicker than 100 nm) versus the relatively short average exciton diffusion length (AEDL, typically less than 50 nm). This causes many photogenerated excitons to be wasted before they reach the donor/acceptor interface. On the other hand, decreasing the film thickness results in decreased photon absorption.

16.2.3 Bulk Heterojunction Organic Solar Cells

The third generation organic/polymeric solar cells were categorized as "bulk heterojunction" or BHJ cells as shown in (Figure 16.12) (spatial profile) and Figure 16.13 (energetic profile) [23]. These cells are fabricated by intimately blending a donor (such as a conjugated donor type polymer) with an acceptor (such as a fullerene). In this way, the donor/acceptor interface (and thus the charge separation sites) are located randomly everywhere in the bulk, making it easier for an exciton to reach a nearby donor/acceptor interface and be dissociated into carriers (though some domain size might still be much larger than AEDL). For instance, it was found that a cell with D-i-A trilayer

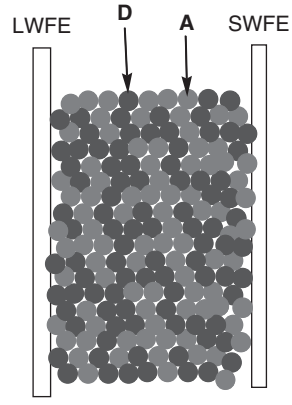


Figure 16.12 Scheme of a donor/acceptor blend type bulk heterojunction BHJ solar cell

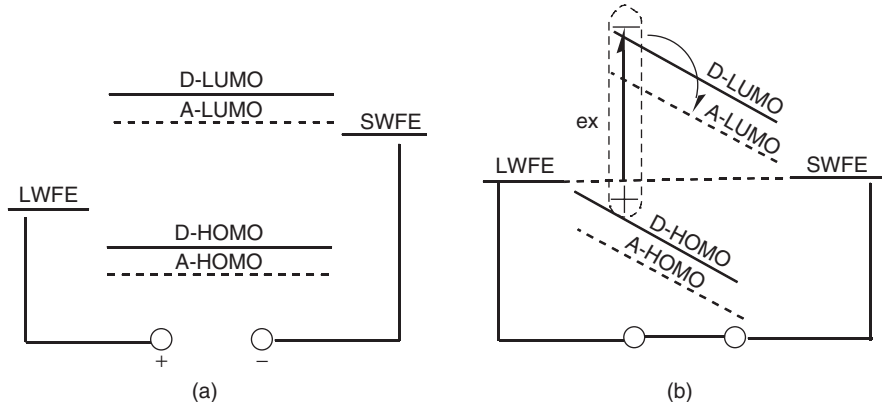


Figure 16.13 Energy level schemes of a bulk heterojunction photovoltaic cell in (a) open-circuit voltage mode and (b) short-circuit current mode

structure had nearly doubled photoelectric conversion efficiency over that of corresponding D-A bilayer cell under similar conditions, where D is a donor layer, A is an acceptor layer, and *i* represents a donor/acceptor blended layer [27]. A number of bulk heterojunction cells using conjugated polymers (e.g. MEH-PPV, P3HT) as the donor and fullerene derivatives (e.g. PCBM) as the acceptor have been intensely studied. They exhibited near unity photoinduced charge separation (internal quantum efficiency) at D/A interface and between 1 and 6% overall photoelectric power conversion efficiencies were reported under different radiation conditions [5, 6, 19, 20, 23, 24, 28, 29], with the best being P3HT/PCBM cells having about 5% efficiencies [28, 29]. The higher efficiencies over the previous generation double-layered cells can be attributed to the proximity of the photogenerated excitons to D/A interfaces, and the increased capture of photons as the films now can be made much thicker to harvest more light. However, even if the charge carrier generation is efficient, the carrier transport to the electrodes appears more problematic than the double-layered Tang cell because the donor and acceptor domains are not really bicontinuous between the two electrodes. The carriers may be stopped or trapped at any islands, or recombined frequently because of the poor or random phase morphology (Figure 16.12). Additionally, if both donor and acceptor are in direct contact with both electrodes, carrier recombination at the organic/electrode interface could be severe resulting in poor carrier collection efficiency at the electrodes. One interesting approach was to fabricate a D/A bilayer first and then enable donor and acceptor partially diffuse into each other to form a D-(D/A)-A concentration gradient type structure which is expected to increase the D/A interface and at the same time still keep the D/A spatial asymmetry between the two electrodes [30].

16.2.4 *N*-type Nanoparticles/Nanorods with *p*-type Polymer Blend Hybrid Solar Cells

A hybrid cell referred to here typically contains an *n*-type nanoparticle or nanorod blended with a *p*-type conjugated polymer such as PPV or polythiophene [5, 31]. Polythiophene has the advantages of better chemical stability and lower energy gap compared with PPV [6, 9]. The *n*-type nanoparticle/nanorods include a variety of inorganic semiconductors such as CdSe, CdTe, PbS, ZnO, or carbon nanotubes (SWNT) [5]. Advantages of these cells include the following: photons can be captured by either polymers or nanoparticles/nanorods; the energy gaps of the nanoparticles/nanorods can be tuned by their sizes; and the nanoparticles/nanorods are chemically robust [5]. Flexible thin film cells can be fabricated that are lighter weight and less expensive than the crystalline silicon cells. Disadvantages include discontinuous transportation pathways for charge carriers, poor control of phase morphology, relatively heavy inorganic materials used, and environmental concerns of lead or cadmium use.

16.2.5 Bicontinuous Ordered Nanostructure (BONS) Organic Solar Cells

A donor/acceptor (or p/n type) bicontinuous ordered nanostructure (BONS) solar cell was proposed and is shown in Figures 16.14–15 [32–36]. Figure 16.14 depicts the cell in the spatial regime, and Figure 16.15 depicts it in the energy regime in (a) the open-circuit situation and (b) the short-circuit situation. In BONS cells, a donor/acceptor bicontinuous two-phase separated and ordered nanostructure morphology can be in the form of an ordered column or cylinder morphology sandwiched and perpendicularly oriented between the two electrodes. Each column or cylinder cross-section diameter can be within the average exciton diffusion length of most organic semiconductors (5–50 nm). In comparison with the blend BHJ system, excitons can easily reach the donor/acceptor interface (at least on the column cross-section direction), and both charged carriers now have a continuous or uninterrupted transport pathway toward their respective electrodes. In this way, the active layer

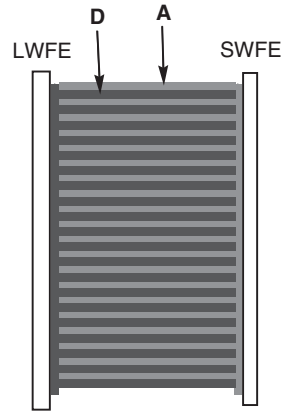


Figure 16.14 Schematic of a donor/acceptor (or p/n type) bicontinuous ordered nanostructure (BONS) type solar cell

thickness can be thicker than the average exciton diffusion length (the same as in the BHJ cell), but the charge transport in BONS appear much better then in BHJ. Since the BONS structural type cell was proposed [32–36], several approaches have been initiated in order to achieve BONS cells. These include, using block copolymers [32–40], n-type TiO_2 or other inorganic semiconducting porous or tunnel templated patterns filled with p-type conjugated polymers [31], n-type aligned ZnO or other semiconducting nanorods mixed with p-type conjugated polymers [41], aligned carbon nano tubes with p-type conjugated polymers [42], and discotic liquid crystalline molecules intended to be stacked in p/n bicontinuous column style structures [43, 44]. These kinds of BONS structure can be applied to all organic D/A, inorganic p/n, or any hybrid inorganic/organic p/n binary solar cells, as the two common features are that the nanodomain interfaces dissociate the excitons, while the nanotunnels provide smooth transport pathways for each carriers.

16.2.6 Tandem Structured Organic Solar Cells

In addition to the desired morphology in a single cell as described above, the optical excitation of the cell should match the energies of the photons to be harvested. In addition, the frontier orbital energy offsets between the donor and acceptor must also be such that the charge separation falls into the optimal electron transfer regime as defined by Marcus theory in order to separate the exciton most efficiently and at the same time minimize the charge recombination [5, 7, 17, 18, 37]. For this reason, optimizations at the energy regime are also critical. Since in each single cell (or subcell), either the donor or acceptor has only one energy gap and can only capture a very narrow range of energy matched photons, and sunlight radiation spans a broad range from UV all the way to IR, a tandem style stacked and serially connected cell is desirable [1–4, 7, 45–48] (see chapter 14 in reference [8]). An ideal multijunction tandem solar cell should have an energy gap gradient among the stacked cells with energy range spanning the whole solar spectrum, so that most of the sunlight can be captured (the cell therefore should appear dark!) [7]. The series of different gapped cells can be spatially stacked parallel to each other, with energy gaps in a descending order from large gap to small gap following the propagation direction of light radiation. In this way, the largest gapped cell at the front would capture highest energy photons first, but allow lower energy photons to pass through to the lower gapped cells behind, where the lower energy photos can be captured, and so on. In addition, even if some excitons in the front large gapped cells did not dissociate and relax to emit a smaller energy photon, that photon can still be captured by the next lower gapped cell [7].

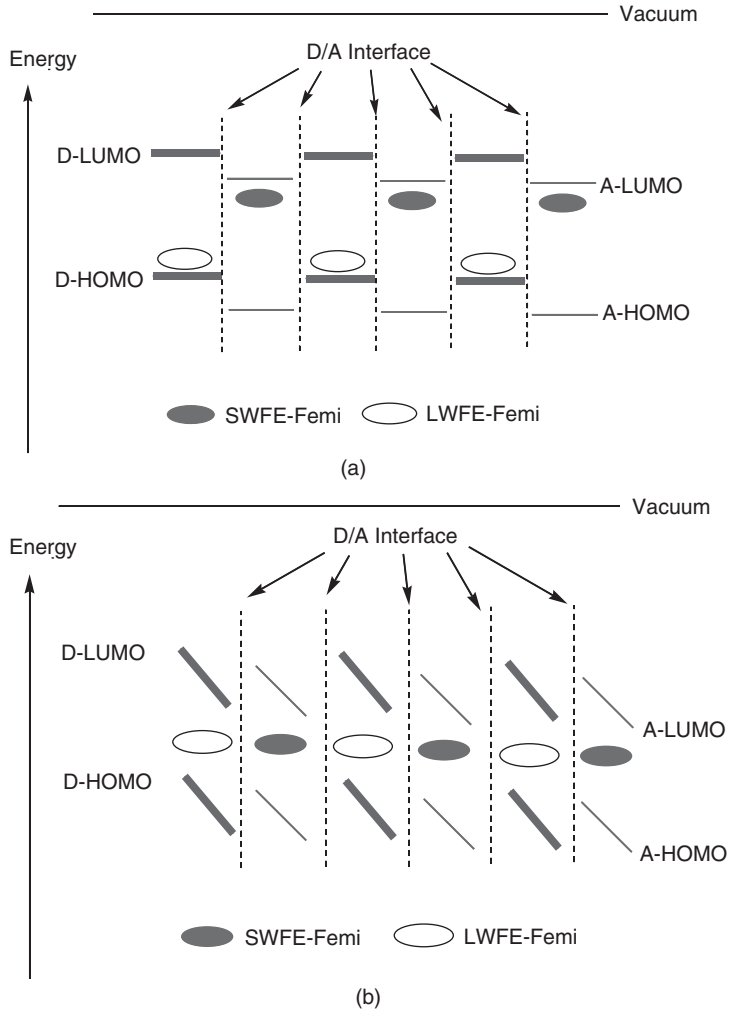


Figure 16.15 Energy level schemes of a donor/acceptor (or p/n type) bicontinuous ordered nanostructure (BONS) type solar cell in (a) open-circuit voltage mode and (b) short-circuit current modes. Each region between the dashed lines represent one column (either donor or acceptor) with diameter less than the typical exciton diffusion length of 5–50 nm

Since the open-circuit voltage of each cell is correlated to its donor HOMO and acceptor LUMO [24], it is therefore critical to connect the cells in a serial manner so the photovoltages can be added up. In this way, large solar cell voltages together with a high power conversion efficiency may be achieved. For instance, over 40% power conversion efficiencies have been reported in inorganic tandem type solar cells [4], and over 6% power conversion efficiency was reported in organic tandem type solar cells [46–48]. Figure 16.16 shows UV-VIS spectra of two organic donors (PCPDTBT and P3HT) and two organic acceptors (PCBM and $PC_{70}BM$) and their composites where a 6.5% power conversion efficiency tandem cell was achieved [47], and Figure 16.17 shows the incident photon-to-electron conversion efficiency (IPCE) and current–voltage ($I-V$) data of those cells [47]. The best device performance is summarized as follows: The PCPDTBT:PCBM

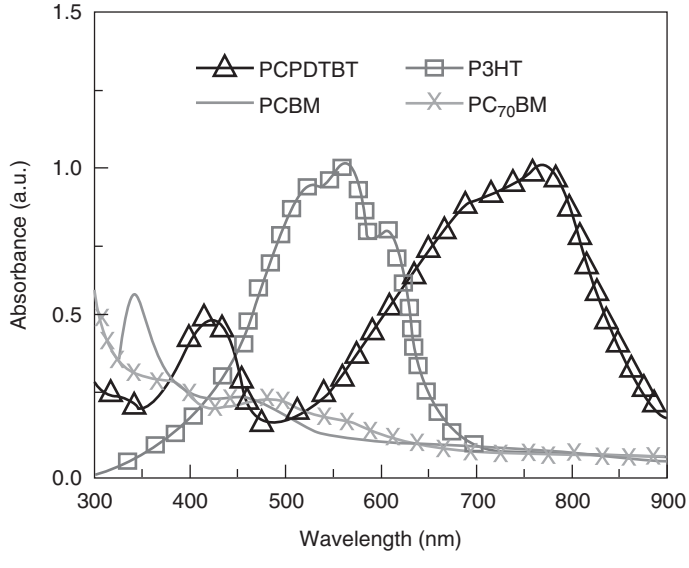


Figure 16.16 (a) Thin film absorption spectra of four representative OPV materials including two donors P3HT and PCPDTBT, and two acceptors PCBM and $PC_{70}BM$ in arbitrary units (b) absorption spectra of three thin film cells fabricated from PCPDTBT:PCBM composite (solid smooth curve); P3HT:$PC_{70}BM$ composite (middle curve with crosses), and a tandem device structure of the two cells (top curve with triangle) [47, with permission]

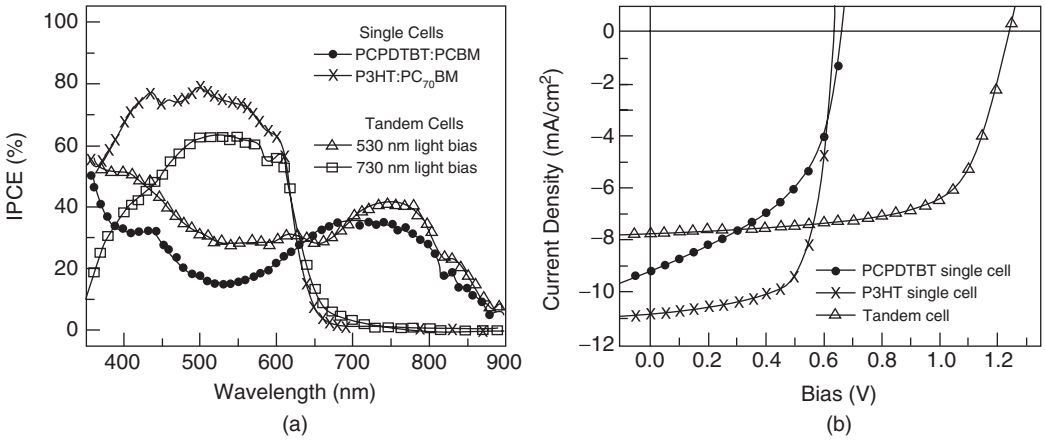


Figure 16.17 (a) IPCE spectra of two single cells and a tandem cell with bias light at two different wavelengths. The IPCE data were measured using modulation spectroscopy with a lock-in amplifier for the single cells and without light bias for the tandem cells. For the tandem cell measurements made with bias light, unmodulated monochromatic light with an intensity of $\sim 2\,\text{mW/cm}^2$ was used (b) $J-V$ characteristics of two single cells and a tandem cell with PCPDTBT:PCBM and P3HT: PC70BM composites under AM1.5G illumination from a calibrated solar simulator with irradiation intensity of $100\,\text{mW/cm}^2$ (about one sun) are presented [47, with permission]

single cell exhibits short circuit current $Jsc = 9.2\,\text{mA/cm}^2$, open circuit voltage $Voc = 0.66\,\text{V}$, fill factor $FF = 0.50$, and power conversion efficiency $\eta_e = 3.0\%$; the P3HT:PC_{70}BM single cell shows $Jsc = 10.8\,\text{mA/cm}^2$, $Voc = 0.63\,\text{V}$, $FF = 0.69$, and $\eta_e = 4.7\%$; and the tandem cell shows $Jsc = 7.8\,\text{mA/cm}^2$, $Voc = 1.24\,\text{V}$, $FF = 0.67$, and $\eta_e = 6.5\%$ [47]. The $I-V$ curves clearly show the photovoltage of the tandem cell is approximately the voltage sum of each subcells.

16.2.7 "Ideal" High-efficiency Organic Solar Cells

As discussed above, it appears that an ideal high-efficiency organic solar cell may be a tandem or multijunction style serially connected and parallel stacked multi layer cells (subcells) with energy gaps graded to cover most of the solar spectrum, and each subcell will have a donor/acceptor (or p/n) BONS structures as described earlier [7]. With further improvements in charge mobility via molecular and morphological engineering and optimizations, such a "dark colored tandem style plastic solar cell" may exhibit comparable photoelectric power conversion efficiencies yet with huge advantages of lower cost, flexibility, and lighter weight than the existing high-efficiency inorganic solar cells.

16.3 ORGANIC AND POLYMERIC SOLAR CELL FABRICATION AND CHARACTERIZATION

16.3.1 Organic and Polymeric Solar Cell Fabrication and Stability

In a schematic organic or polymeric solar cell device structure, as shown in Figure 16.18, the organic or polymeric semiconductor layer responsible for photovoltaic functions (called the active layer) is

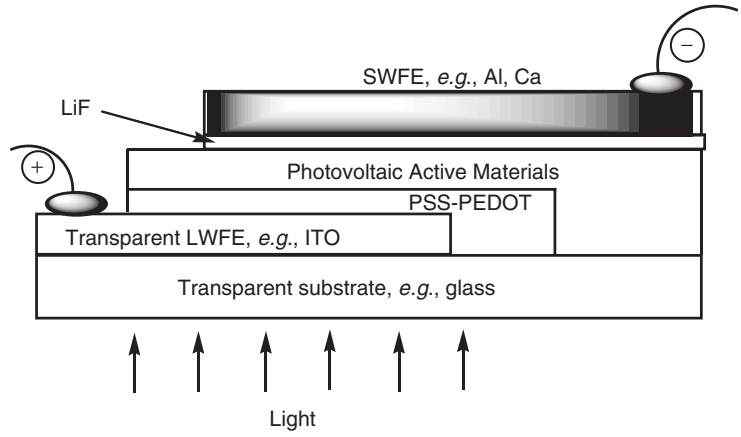


Figure 16.18 General cross-sectional scheme of an organic/polymeric solar cell

typically sandwiched between a transparent conducting electrode (TCE) (e.g., indium–tin–oxide (ITO) coated glass or polymer film as the LWFE, bottom) and a metal electrode (e.g. aluminum as the SWFE, top). However, it has recently been found that a poly(ethylene dioxythiophene):polystyrene sulfonic acid layer (PEDOT-PSS or PSS-PEDOT [5], chemical structure shown in Figure 16.19), and other similarly functionalized buffer layers would dramatically enhance the hole transfer between the active layer and the ITO electrode [49]. On the other hand, a thin layer LiF also appeared to enhance the electron transfer between the active layer and the metal electrode as will be further elaborated below. For small organic molecular solar cell fabrications, typically high vacuum (at least 10^{-6} Torr) vapor deposition and occasionally solution crystallization protocols are used to grow organic thin films on the TCE substrate, followed by the vacuum thermal evaporation deposition of metal electrode (mostly aluminum, sometimes calcium or silver) on top of the photovoltaic active layer. For research purposes, the cell size or shape can be made using a shadow mask on top of organic active layer when depositing metal electrodes, and it is common that several cells may be fabricated on one 2.5×7.6 cm ITO glass slide. Previously, organic solar cells are mostly made very small (e.g. 2×2 mm). Nowadays, it is strongly recommended or recognized in the community that at least 10×10 mm cell should be fabricated and tested for reliable comparisons or results. For polymer-based solar cell fabrications, typically solution spin coating (small devices), inkjet printing

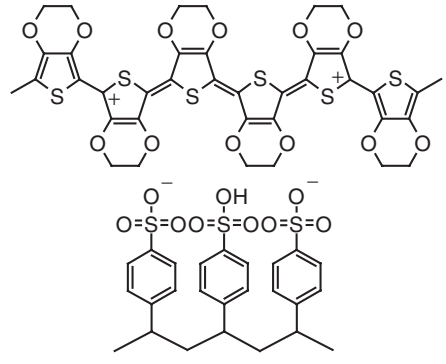


Figure 16.19 Chemical structure of PSS-PEDOT

(medium-size devices), or industrial scale roll-to-roll plastic thin film manufacturing (large-size sheets) protocols may be used for the active layer (chapter 19 in reference [6]). Solution processing offers advantage of low cost and convenience in large-scale industrial production [5, 6, 57].

The surface treatment of the transparent conducting electrode (TCO, such as ITO glass) is very critical. Different methods of cleaning ITO surface have been applied [49]. The aqueous solution processed PSS-PEDOT layer is coated on top of ITO glass electrode and below the active layer, and this has been observed to optimize the hole transfer between the active layer and the ITO electrode. The conductivity of PSS-PEDOT can reach 80 S/cm if produced by electropolymerization, and can be as low as 0.03 S/cm if produced by chemical polymerization. This is believed to be due to different compositions of the polymer [50]. Sheet resistance of PSS-PEDOT as low as 350–500 Ω/square inch has been reported [51]. It became an extremely attractive material when it was found that the performance and stability of polymer light-emitting diodes (LEDs) could be improved by inserting this material between the polymer active layer and ITO glass as a buffer layer [52, 53]. PSS-PEDOT is available in an aqueous dispersion, which can form uniform, transparent, conductive film by solution spin coating. The most important physical properties for its application in devices are the high work function (−5.2 eV) and the smooth surface. The improvement of the device performance after PSS-PEDOT layer application may be attributed to a number of factors in either spatial or energy regimes. In the spatial regime, for instance, commercial ITO glass surfaces have been found to be very rough and the conductivities were area sensitive [54]. The PSS-PEDOT layer has been found to smooth both the surface roughness and the conductivities at different spots. PSS-PEDOT may also block the electrons from reaching the ITO electrode, thus effectively preventing electron–hole recombination at the ITO surface. In the energy regime, the work function of PSS-PEDOT lies between the work function of the ITO (−4.7 eV) and the HOMO levels of most organic donor materials. This intermediate level appear to facilitate or optimize the hole transfer between the polymer and the ITO. However, PSS-PEDOT layer thickness should be optimized since too thin a layer may have many pinholes, while thicker layers cause a larger increase in the series resistance [54].

An inert gas environment for organic solar cell fabrication/testing is known to be crucial because many organic/polymeric materials react readily with oxygen or water, particularly during optical excitations. Such photo-oxidation may be attributed to photoinduced electron transfer from organic materials LUMO to the LUMO of oxygen (particularly singlet oxygen) or water. Therefore, organic solar cell encapsulation with good oxygen/water barriers is essential at the present time. For instance, Konarka demonstrated excellent stability of organic solar cells for thousands of hours under encapsulation [55]. Alternatively, materials frontier orbital level fine tuning (e.g., lowering the frontier orbitals of materials to weaken the electron transfer to oxygen and water) would also improve ‘the materials’ environmental stability (also chapter 18 in reference [6]).

16.3.2 Status and Challenges of OPV Manufacturing

Presently, there are at least two companies which have developed a manufacturing-scale process for OPV modules, Konarka and Plextronics. Both are in the USA. Konarka prints the various layers on a moving roll-to-roll web using inkjet “drop-on-demand” technology [56]. They have produced a 6.4% small-area (0.75 cm^2) cell and offer a series of ∼1.6% efficient series-connected modules for various consumer and remote power applications. They are constructing a module manufacturing facility in Boston (US) with a capacity of hundreds of megawatts. Plextronics makes small-area P3HT/PCBM cells (1 × 2 cm^2) using a traditional spin-coating method, then series interconnects 54 of them to make a 15 × 15 cm module [57]. They have achieved 1.1% total area or 3.4% active area efficiency (233 vs 108 cm^2, respectively). Their best small-area device is 5.4% (verified at NREL). Their business appears focused on selling the polymer inks, not modules.

Performance of the bulk heterojunction devices made from a polymer:fullerene blend depends on the solvent, annealing and drying conditions, chemical composition of donor, and choice of hole injecting layer (HIL). Production of modules at high throughput and high yield adds additional constraints on drying time, uniformity of thickness and composition, Additionally, the active polymer layers are formed at atmosphere pressure, perhaps in an inert atmosphere. Avoiding the vacuum systems common to all other thin films results in lower equipment cost, but greater sensitivity to atmospheric and particulate contamination.

We will briefly discuss issues relating to the inkjet printing process since it is clearly amenable to large-area, high-throughput coating. Konarka's inkjet printing process requires optimizing around very different conditions than for standard spin-coating or doctor blading application [56]. One of its unique features is that the monolithic series interconnection is achieved by careful registration of the printing of adjacent layers rather than some post-deposition laser scribing as commonly used in the other thin film module technologies (see Sections 12.6.2 or 14.6). The successive registration approach requires much wider "dead space" between adjacent cell strips, hence reduced active area, but simplifies the production. There are two types of problems encountered in printing OPV via a roll-to-roll process: those common to any high-quality, precision coating and packaging process and those that are unique to the handling of OPV materials. The challenges are numerous; some of those are described below.

A common serious concern in high-quality coating is web path cleanliness. Clean-room classifications of 1000 (ISO 6) or better are necessary to minimize airborne-source defects. For OPV manufacture, this type of contamination will turn a cosmetic issue into a catastrophic failure.

Defects resulting from coating fluid inhomogeneities such as gels, coagulants, foreign matter and bubbles will create voids or debris in the structure which will lead to functional failures in the cells. These can occur in any high-quality coating but, again, OPV systems are much more sensitive. Coating anomalies along the direction of the subtrate travel such as streaks and ribbing, cause uneven coating, and may result in lower cell performance.

Unique to OPV coating is the need for lateral registration of cell coatings and features to provide the cell-to-cell series interconnection. If any layer or artifact is misaligned as little as a few micrometers, the stacking in that region will not be complete and shorts or shunts will result. Likewise, the cells are defined by scoring the transparent conductive base underlayer. Incomplete or mis-registered scoring will destroy the cell.

OPV materials, prior to coating, require protection from certain environments and temperature extremes. After coating they require other drying and annealing conditions necessary to create high efficiencies. To determine and control these conditions, specific to a given coating line, requires theoretical knowledge, disciplined experimentation, and rigorous process control.

The packaging of the coated cell, presents another set of challenges since OPV cells are notorious for their sensitivity to moisture and oxygen [55, 58]. Unpackaged OPV cells are very sensitive to moisture and oxygen. Encapsulation webs are fragile and require processes that minimize stress prior to integration of the final packaged product.

16.4 NATURAL PHOTOSYNTHETIC SUNLIGHT ENERGY CONVERSION SYSTEMS

Photosynthesis is the biological process by which sunlight is captured and by a series of reactions is converted to the biochemical energy needed to support life. It provides all of the sustenance and the majority of the energy resources for our planet. The most common form of photosynthesis

is chlorophyll-based, and can be broadly separated into two different types of process. Oxygenic photosynthesis is carried out by plants, algae, and cyanobacteria and results in the conversion of carbon dioxide and water into carbohydrate and oxygen. Anoxygenic photosynthesis occurs in primitive bacteria such as purple and green bacteria and produces a carbohydrate, a reduced acid, and water from carbon dioxide and molecules such as hydrogen sulfide or reduced organic molecules. The photosynthetic process, like many other energy related biochemical processes, takes place in lipid bilayer membranes. The primary reactions are carried out by pigment-containing proteins that are integrally associated with the membrane. These membranes are either the cytoplasmic membrane of bacteria or the highly reticulated thylakoid membranes of cyanobacteria or chloroplasts in higher plants and algae. The chloroplast is a subcellular organelle that is ancestrally related to cyanobacteria.

The photosynthetic process can be divided into three phases: light harvesting by antenna systems, conversion of excitonic energy to chemical energy and synthesis and export of products. The individual reactions are performed by a series of protein complexes and have provided much inspiration for development of artificial systems that emulate their exquisite molecular architectures. This section will primarily deal with advances in bioinspired and biomimetic photosynthesis research related to the first two phases of photosynthesis that is very similar to organic photovoltaics.

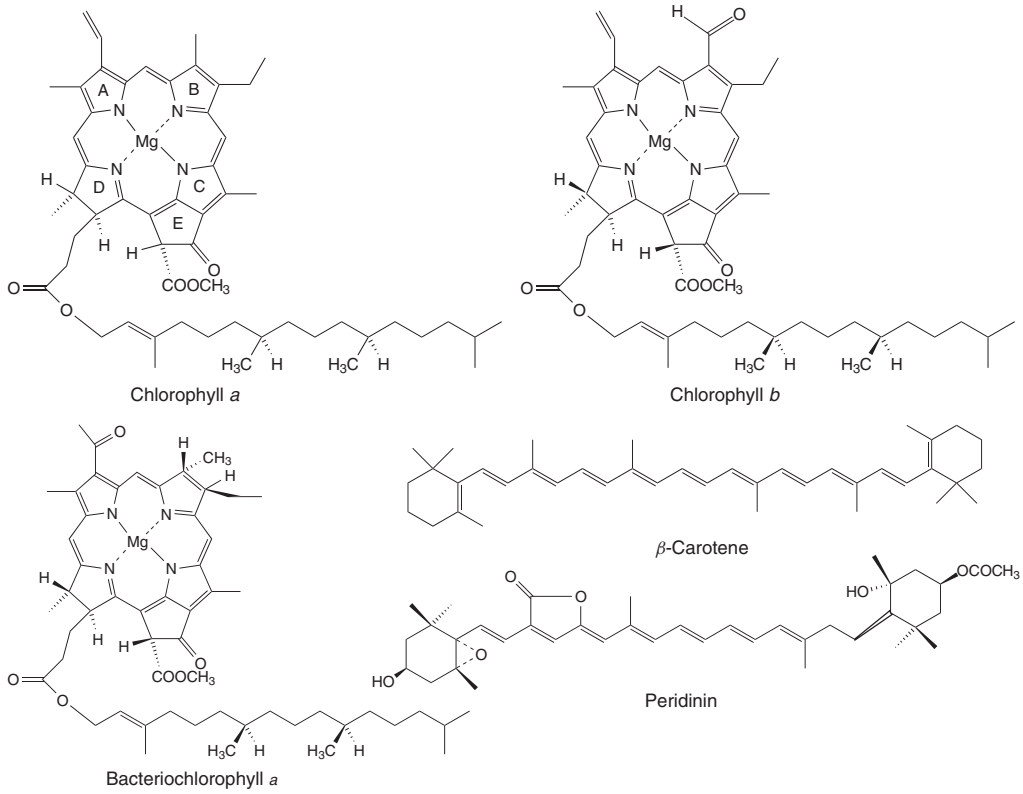
16.4.1 Photosynthetic Pigments

The first step in the photosynthetic process is absorption of sunlight photons by the pigments. All chlorophyll-based photosynthetic organisms contain several types of pigments each optimized for a different function [59]. Chlorins, such as chlorophyll a and b and the bacteriochlorophylls, are similar in structure to porphyrins except for the presence of an additional ring (E) called the isocylic ring (Figure 16.20). In addition, one or more of the rings are reduced decreasing the symmetry of the macrocycle causing an increase in the absorptivity of the pigment in the region corresponding to the photosynthetic active region of the solar spectrum. In Chlorophyll a and b, the primary pigments of green plants and algae, ring D is reduced. In bacteriochlorophyll of green and purple bacteria ring B is also reduced causing a red shift in the absorption spectrum of the pigment. The hydrocarbon tail, usually an isoprenoid group, attached to ring D serves to anchor the pigment in the protein environment. The nature of the central metal atom in the chlorin macrocycle plays a critical role in the photophysical properties of the pigments. It is most commonly a magnesium ion (Mg), although there are a few cases known where the Mg is replaced by zinc. The excited state lifetimes of pigments with these metals are relatively long, typically several nanoseconds. In contrast, substitution with iron, manganese, or copper decreases the excited state lifetimes by many orders of magnitude due to ultrafast internal conversion via the unfilled orbitals of the metal, and the pigments are not able to directly sensitize photosynthetic energy.

Carotenoids are another class of pigment common in photosynthetic organisms. They function as accessory antenna pigments absorbing light and transferring it to chlorophyll and are important for regulation of energy flow in photosynthetic systems. In addition, carotenoids also play a very important role in providing photoprotection to the organisms. They are very efficient at quenching undesirable excited states such as chlorophyll triplet states and singlet oxygen thus protecting the pigments from photooxidative damage.

16.4.2 Antenna Complexes

Most pigments in photosynthetic systems function as antenna systems that collect light and deliver the energy to the reaction center [59]. They greatly increase the amount of light available to the reaction center because collectively they can absorb more light than individual pigments, thus



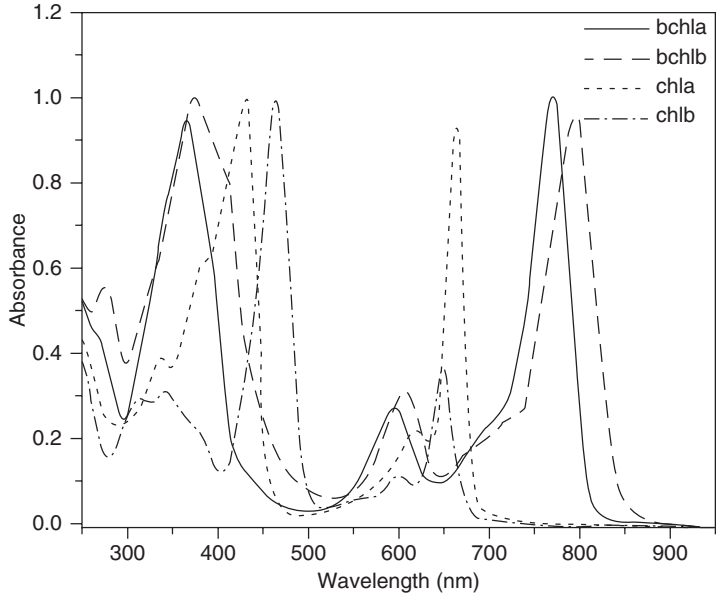
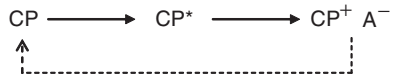


Figure 16.20 The chemical structures and absorption spectra of representative photosynthetic pigments. The rings are designated by letters. Chlorophyll (chl) a, bacteriochlorophyll (bchl) a, and carotenoids are found in both antenna and reaction center complexes while chlorophyll b is only found in antenna complexes. Peridinin is only found in one class of antenna complex (Reproduced by permission, [60])

increasing the overall efficiency of the photosynthetic process. An antenna system does not carry out any chemistry, but operates by intermolecular energy transfer of electronic excited states. This physical process is highly dependent on the energetic coupling of antenna pigments allowing energy transfer between chlorophylls to occur over several nanometers on a picosecond time scale. This is reflected in the exquisite organization and spatial arrangement of chromophores in the photosynthetic antenna protein complexes. Excitation energy transfer (EET), the process by which antennae complexes absorb light and funnel the resulting excited-state energy rapidly and efficiently to an acceptor moiety is the primary function of antenna complexes. The mechanism for EET can be either by coulombic interactions between transition dipole moments (Förester-type EET) in which the energy transfer rate depends on the sixth power of the distance between the pigments or via electron-exchange interaction through direct or indirect overlap of wavefunctions (Dexter-type EET) [61, 62]. The energy transfer rate depends on the relative orientation of the transition dipole moments of the energy donor and acceptor, and the spectral overlap of the pigments. It is sufficiently fast to deliver the energy to the reaction center complex in less than the nanosecond excited state lifetime of the pigment. In antenna complexes where the pigments are in close proximity and strongly energetically coupled, such as is found for the bacterial chlorosome and the LH1 and LH2 complexes, the excited state is considered delocalized over several molecules as an exciton [63]. A detailed description of individual natural antenna systems and their relationship to synthetic systems is presented later in the chapter.

16.4.3 Photosynthetic Reaction Centers

The final step in the excitonic diffusion by the antenna system is transfer of energy to a special chlorophyll dimer that is present in the photosynthetic reaction center. This represents the first step in the conversion of excitonic energy to redox or chemical energy. The photosynthetic reaction centers are large membrane bound multisubunit protein complexes that incorporate both chlorophylls and other electron transfer cofactors such as carotenoids, quinones, iron–sulfur clusters and metal centers. The basic mechanism of charge separation is the same in all the reaction centers. A special chlorophyll pair (CP) is promoted to an excited state by either direct photon absorption or more commonly by energy transfer from the antenna system. In its excited state, it is an extremely strong reducing species and rapidly transfers an electron to a nearby acceptor molecule (A) generating the ion pair (CP^+ A^-) according to Scheme 16.1.



Scheme 16.1 Photoinduced charge separation in a chlorophyll pair

The close proximity of the CP and A moieties make the system vulnerable to recombination and hence loss of energy as heat (dotted line in the scheme). However, this is avoided due to the positioning of secondary acceptors that successfully compete for electrons, preventing the recombination reaction. These reactions spatially separate the positive and negative charges and reduce the intrinsic recombination rate by orders of magnitude. For instance, in the Photosystem I reaction center of the oxygenic photosynthetic apparatus, after absorption of a photon, the photochemistry of the special chlorophyll pair (P700) induces a charge separation in the reaction center by rapid electron transfer to peripheral iron sulfur complexes (F_A and F_B) via the intermediates A_0, A_1 and F_x (see Figure 16.21). This results in the generation of a weak oxidant at the primary electron donor (P700) and a strong reductant at the terminal electron acceptor (4Fe-4S center). The photochemical reaction is completed within 100–150 ns [65] and generates approximately a 1 V

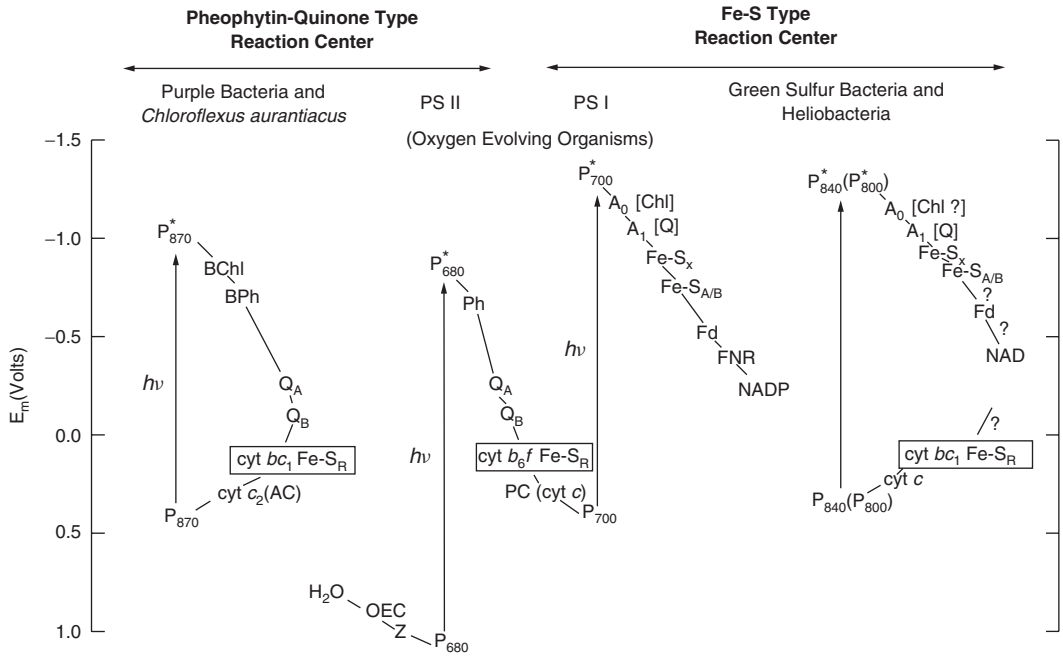


Figure 16.21 Schematic diagram of the electron transfer pathways of photosynthetic reaction centers. The cyclic electron transfer system of anoxygenic purple bacteria and anoxygenic green sulfur bacteria are shown on the far left and right, respectively. The so-called Z-scheme of oxygenic photosynthetic organisms carried out by Photosystem I (PS I) and Photosystem II (PS II) is shown in the center. The vertical axis is the midpoint redox potential of the electron carriers. The carriers in parentheses indicate alternative species in some organisms and question marks indicate electron transfer steps that are likely, but not unambiguously established. The abbreviations are listed elsewhere [64] (reproduced by permission)

potential difference [66] over a distance of 6 nm. The quantum yield of this photochemical reaction is close to 100% [67, 68]. However, the energy efficiency for conversion from photons to chemical energy that can be recovered during sugar metabolism is close to 20–25% [69].

The electron transfer pathways of the other major reaction centers are also presented in Figure 16.21, showing the redox midpoint potentials of the various cofactors that are part of the photosynthetic apparatus and their and secondary electron carriers. This schematic representation allows the energetics and the complete electron transfer pathways of the photosynthetic systems to be easily observed. The vertical arrows represent the excitation of the special chlorophyll pair to its excited state and their length is proportional to the energy of excitation of the primary electron donor pigment. The numeric designation (e.g. P700) represents the wavelength at which maximal absorption of the special chlorophyll occurs and the upper end of the arrow gives the redox potential of the chlorophyll excited state (P*700). The simplest and best understood reaction center, from purple bacteria, will be discussed later in the chapter.

16.5 ARTIFICIAL PHOTOSYNTHETIC SYSTEMS

The conversion of solar energy to electricity or chemical fuel by artificial photosynthesis is one of the most challenging goals in chemistry [70]. As described above, the supramolecular organization of

the natural photosynthetic process provides inspiration for development of a synthetic counterpart. In natural systems, efficient conversion of light to chemical energy is achieved by a very precise organization in the dimensions of: (a) space, by proximal location of molecular components; (b) energy, with respect to excited states and redox potentials of adjacent partners; and finally (c) time, concerning the rates of competing processes. Similar organization has been employed in molecularly engineered artificial systems by making use of covalent and noncovalent bonding strategies. There has been progress in the development of each aspect of artificial photosynthetic systems. However, integration into a working system has not been achieved.

16.5.1 Antenna Systems

The function of an antenna system is to increase the photon-capture cross-section area, leading to very high efficiency for energy trapping and transport to the reaction center. This in turn maximizes the turnover rate of the reaction center. In addition, interchromophore interactions as well as enrollment of different chromophoric species broadens the spectral range available for conversion of light into energy. Various synthetic strategies have been proposed that take their inspiration from natural systems.

16.5.2 Cyclic Porphyrin Arrays

In recent years covalently linked photosynthetic arrays have been investigated as artificial photosynthetic antennae [71]. These systems take their inspiration from the light harvesting antenna complex (LH2) present in purple bacteria such as *Rhodopseudomonas acidophila* [72, 73]. The LH2 complex is an $\alpha_a\beta_a$ circular nonamer that is comprised of two wheel-like assemblies of bacteriochlorophyll a (BChla) named B800 and B850, with diameters of 62 and 52–54 Å, respectively (Figure 16.22a). B800 contains 9 BChla molecules, one per $\alpha\beta$ apoprotein dimer, with a Mg–Mg interchromophore distance of 21.2 Å (Figure 16.22b). The BChla molecules are considered monomeric since there is little interchromophore electronic interaction. In contrast, the 18 BChla molecules in B850 form slipped-cofacial dimeric subunits, with a Mg–Mg interchromophore distance of 8.8 Å with each $\alpha\beta$ subunit and 9.5 Å between subunits (Figure 16.22b). The excitation energy transfer (EET) rate is primarily dictated by the center to center distance of neighboring porphyrins. In the natural system, the EET rate of the B850 assembly of *R. acidophila* LH2 was estimated at $270\,fs^{-1}$ by a Förester mechanism based on the large dipole interactions between the cofacial BChla [74]. The EET rate constant for the B800 assembly of *Rhodobacter sphaeroides* is estimated at $0.8–1.6\,ps^{-1}$, reflecting the longer distance between neighboring chromophores [75].

The primary motivation for the synthesis of cyclic porphyrin arrays is to duplicate the structure and function of the natural light harvesting antennae. These have been constructed by covalent, noncovalent, or metal coordination bonding [76–78]. Covalently bonded arrays exhibit the highest structural stability, but are difficult to make, requiring a template for the final macrocyclization steps to enable the precursor to take a favorably folded conformation for cyclization. In contrast, noncovalently assembled metal coordination bonded arrays are more susceptible to environmental effects that may induce dissociation as will be discussed later.

The synthetic schemes for formation of cyclic porphyrins have been recently reviewed, giving insight into the wide varieties of structures possible [71]. In general, two synthetic approaches have been employed for synthesis of cyclic porphyrins, either phenylene or acetylene bridged oligomeric structures or *meso-meso* linked arrays. While all cyclic porphyrins exhibit efficient EET rates [71], only very efficient EET processes have only been observed for *meso-meso* linked porphyrin arrays. These arrays are synthesized by reacting zinc porphyrin monomers possessing unsubstituted *meso* positions in the presence of a Ag(I) salt. The coupling reaction is highly

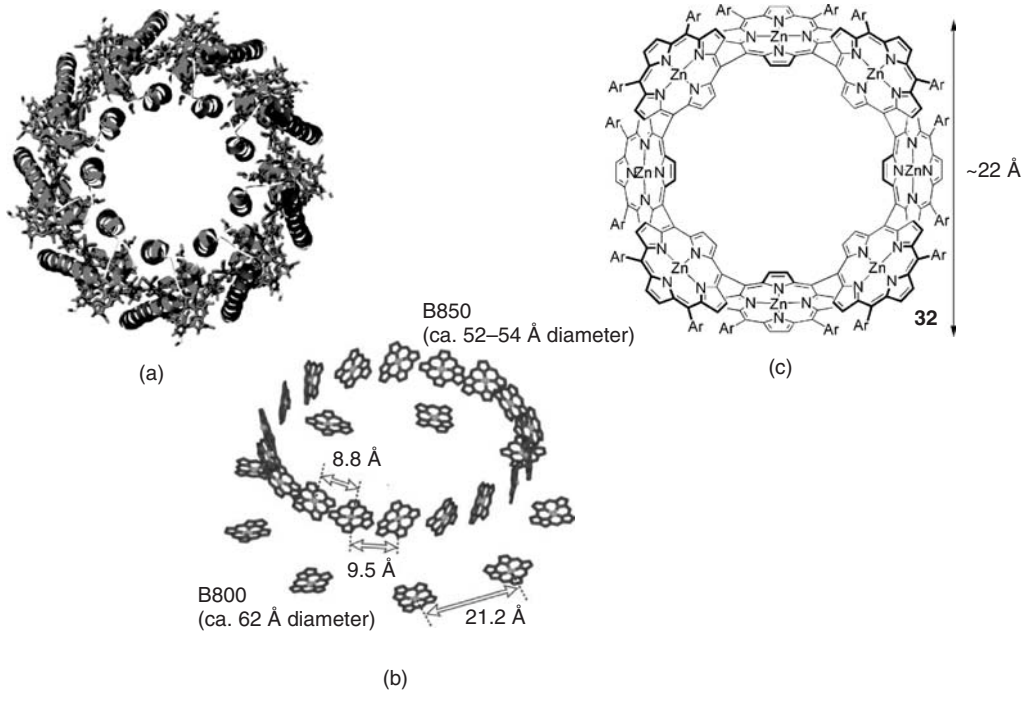


Figure 16.22 Comparison of the structures of the purple bacteria LH2 antenna complex and synthetic cyclic porphyrin. (a) Structure of purple bacteria LH2 antenna complex: (b) the arrangement of the bacteriochlorophyll in the protein complex; (c) the crystal structure of LH2 complex and spatial arrangement of bound chromophores. Structure of *meso-meso* linked 8 unit porphyrin array. Adapted from reference [71] reproduced by permission

regioselective, taking place only at the *meso* position on the porphyrin and with careful control of the reaction conditions, oligomers with the desired number of porphyrins can be obtained [79]. Directly linked cyclic porphyrin arrays with 4, 6, or 8 porphyrin units were synthesized from 5,10-diaryl zinc porphyrin as the starting monomer [80]. The structure of the largest array is shown in Figure 16.22c. These constructs exhibit EET rates of less than 1 ps, rivaling those observed in the natural LH2 antenna system. This reflects the very close interchromophore spatial arrangements leading to extremely large Förester-type interactions (Figure 16.22c). These studies give insight into the structural requirements for efficient EET and hold promise for the development of shape-persistent arrays as structural units for larger functional aggregates.

16.5.3 Dendrimers

Dendrimers are a class of well-defined macromolecules that exhibit extended tree-like nanometer scale architectures. Although there are no examples of a dendritic light harvesting antenna complex in nature, it is appropriate to discuss this approach as it highlights how chemistry can exploit naturally occurring molecules, such as chlorins, for the design and synthesis of novel architectures. Dendrimers are attractive for the construction of light harvesting antennae because repetitive and high-yielding reaction sequences allow incorporation of a large number of chromophores in a restricted space with topological control. Extended dendrimers can be designed in a rational manner with predictable photophysical and efficient light harvesting properties.

A large multiporphyrin array, intended to mimic the morphology of the chromophore array of the natural light harvesting antennae has been reported [81]. It is composed of four dendritic wedges of a zinc porphyrin heptamer ($7P_{Zn}$) anchored to a central free base porphyrin unit (P_{FB}) (Figure 16.23). The dendritic wedge acts as a donor to the acceptor focal porphyrin unit (P_{FB})with poly (benzyl ether) units at the periphery to make the arrays soluble in common organic solvents. The large number of chromophore units (28 in total) in the star shaped dendritic array ($(7P_{Zn})_4P_{FB}$) allows for efficient capture of photons.

The zinc porphyrin units cooperate with each other to facilitate long range energy migration in the absence of π-electronic conjugation. The EET rate constant for energy transfer from the excited singlet states of the dendrons to the central porphyrin of the star shaped polymer ($(7P_{Zn})_4P_{FB}$) is 1 ns^{-1} and yields 71% efficiency. In comparison, the EET rate for the heptamer conical shaped unit ($(7P_{Zn})_1P_{FB}$) is 0.1 ns^{-1}, an order of magnitude smaller than that for the star shaped dendrimer and the efficiency was 19%. The large differences observed in the rate and efficiency of energy transfer indicates that the morphology of the array can significantly influence energy transfer. The energy migration characteristics of the morphologically different ($(7P_{Zn})_4P_{FB}$) and ($(7P_{Zn})_1P_{FB}$) were also investigated by steady-state fluorescence depolarization in a viscous medium. Under these conditions molecular motions are suppressed and energy migration is expected to occur between randomly oriented chromophore units in the dendritic array. Excitation of ($(7P_{Zn})_4P_{FB}$) with 544 nm polarized light resulted in highly depolarized fluorescence from the Zn-porphyrin units before energy is transferred to the free-base core. The fluorescence anisotropy was evaluated to be 0.03, considerably smaller than 0.19 measured for a monomeric reference compound, indicating that efficient

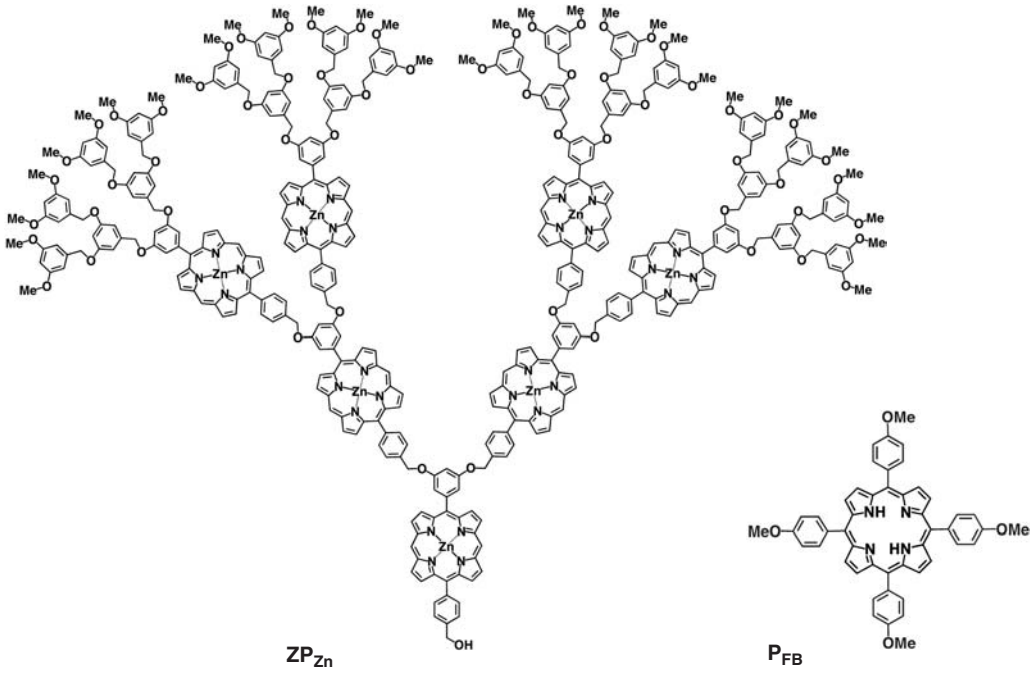


Figure 16.23 A large dendritic mulitporphyrin array as a mimic of bacterial light-harvesting antenna complex. The P_{FB} moiety functions as an electron acceptor and is the central molecule that anchors four $7P_{Zn}$ conical shaped dendritic wedges to make a star shaped dendrimer (reproduced from reference [81], by permission)

energy migration occurs among Zn-porphyrin units before energy is transferred to the free-base core. The fluorescence anisotropy factor for the conically shaped ($(7P_{Zn})_1P_{FB}$) was larger than that of its star-shaped counterpart at a value of 0.10. The results suggest a possible cooperation between the four dendritic wedges of ($(7P_{Zn})_4P_{FB}$) for facilitating energy migration among the porphyrin units. This approach for rational design of light harvesting antennae, incorporating 28 porphyrin units that cooperate with each other to facilitate long-range energy migration and transfer to a focal acceptor core mimics several catalytic properties of the natural purple bacteria light harvesting LH1 complex.

In addition to porphyrin dendritic light harvesting antennae, other approaches have also been investigated. Metal complex dendrimers have been synthesized that contain Ru (II) and Os (II) as the metal ion, and oligopyridine units such as 2,3- and 2,5-bis(2-pyridyl)pyrazine as bridging units and 2,2-bypyridine and 2,2-biquinoline as terminal ligands [82]. This has enabled construction of species containing up to 22 metal-based units. Dendrimers based on organic molecules have also been constructed. In polyphenylene dendrimers consisting of a terrylenedimide core with four perylenemonoimides attached to the scaffold and 8 naphthalenemonoimides at the rim, the antenna effect has been studied at the ensemble and single molecule level [83]. Efficient energy transfer from the perylenemonoimides and naphtahlenemonoimides to the core have been observed. Dendrimers based on host–guest systems take advantage of internal cavities in the dendrimers that can host ions or neutral molecules allowing the collected energy to be delivered by the same dendrimer to suitably tuned guests. Shape-persistent arrays comprised of a hexamine core surrounded by dimethoxybenzene and naphthalene units or poly(propylene amine) dendrimers with surface modified oligo(*p*-phenylene vinylene) units have been used to host dye molecules such as eosin [84]. Energy transfer from the dendrimer host to the dye guest has been shown to be very efficient by a Förestor-type mechanism due to the strong overlap between their respective absorption spectra.

16.5.4 Self-assembled Systems

The chlorosome, an organelle found in green sulfur photosynthetic bacteria, is one of the most efficient light harvesting apparatuses found in nature. Unlike the natural antenna systems discussed in an earlier section, it is the only known photosynthetic system where the majority of pigments (BChl c, d, e) are not organized in pigment–protein complexes, but instead as assemblies of pigment–pigment aggregates. The chlorosome is found attached to the inner side of the cytoplasmic membrane of the bacteria and its typical dimensions are $150 \times 50 \times 20$ nm, containing $\sim 10^5$ BChl molecules encapsulated in a lipid bilayer membrane [85]. Although the prevailing model of chlorosome structure describes the organization of the BChl aggregates into rod-like elements, a recent electron microscopy study complemented by solution small-angle and wide-angle X-ray scattering revealed the internal structure with a spacing of 20 Å that can be explained by a simple lamellar organization, in which the BChl molecules aggregate into semicrystalline lateral arrays of pigment molecules [86]. This photosynthetic structure has provided inspiration for the development of synthetic analoges that mimic the structural and functional properties of these exquisite structures. One approach to duplicate the self assembly process of naturally occurring BChls in chlorosomes has been appending onto porphyrins, strategically placed functional groups capable of inducing self-assembly. Interestingly, these studies in turn, have provided insight into the structure of the naturally occurring system.

Typically porphyrins with groups such as undecyl, octaethyl- or 3,5-di-tert-butylphenyl groups are used as starting materials [87]. These units are then further derivatized with functional groups, such as acetyl or hydroxyethyl groups, that promote self-assembly. The self-assembly process induces a red-shift and spectral broadening, similar to that observed in J-aggregates. The aggregates are dynamic systems, forming larger aggregates that can be dispersed by agitation or addition of suprastoichiometric reagents that compete for metal ligation. It was

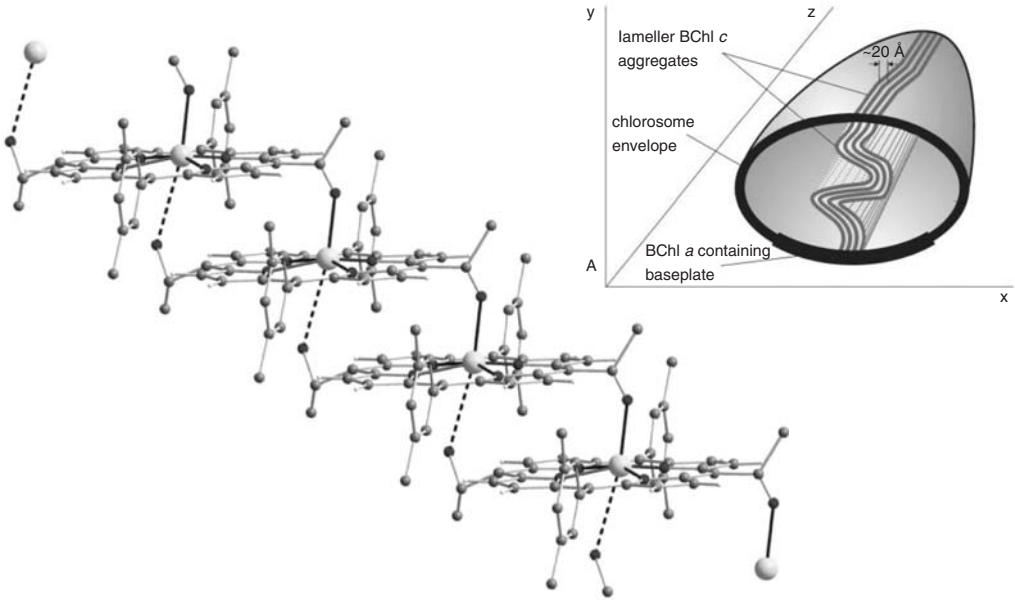


Figure 16.24 The crystal structure of a Zn porphyrin elucidated by X-ray diffraction analysis and schematic representation of BChl aggregates in the chlorosome. The Zn porphyrin has 5-hydroxyl and 15-acetyl groups. *Inset*: the arrangement of the lamellae inside the chlorosome. Each undulating line extends along the long axis of the chlorosome and through the height. (Adapted from [86] and [87], reproduced by permission)

determined that racemic mixtures of porphyrins self-assemble more easily than single enantiomers, suggesting that heterochiral self-assembly is more thermodynamically favorable than homochiral self-assembly [88]. This provides insight into why epimers of BChls are present in chlorosomes, even though in natural systems it is usually preferable to synthesize only one enantiomer of a molecule. The crystal structure of a Zn-porphyrin having 5-hydroxyl and 15-acetyl groups is presented in Figure 16.24. It revealed extended stacks of Zn-porphyrins where the Zn atom is ligated to a carbonyl O atom from one side of the porphyrin plane and an acetyl oxygen on the opposite side. At high concentrations of porphyrin the stacks form lamellar structure with a long crystallographic axis of 2 nm, identical to the Zn–Zn spacings that were independently observed in SAXS and TEM measurements [88]. There is no evidence of hydrogen bonding between the stacks. The overall morphology of the assembled porphyrins is very similar to the natural chlorosome. This provides support for the hypothesis that the chlorosome is a lamellar structure of antiparallel stacked Bchl dimers and not organized into rod-like elements.

16.6 ARTIFICIAL REACTION CENTERS

16.6.1 Bacterial Reaction Center

Two major taxonomic groups of photosynthetic bacteria, purple sulfur bacteria (*Chromatiaceae*) and purple non-sulfur bacteria (*Rhodospirillaceae*), contain photosynthetic reaction centers (PRC) of similar structure [89]. The purple bacterial PRC is the best-characterized of all naturally occurring

photosystems. These proteins are membrane spanning complexes that catalyze the first steps in the conversion of light energy to chemical energy during photosynthesis. The three-dimensional structures of *Rhodopseudomonas viridis* and *Rhodobacter sphaeroides* PRCs have been elucidated using X-ray crystallographic techniques [90, 91]. They consist of at least three polypeptide chains termed L (light), M (medium) and H (heavy) and a four-heme cytochrome c that is considered the fourth protein subunit. The polypeptide subunits L and M bind BChl, bacteriopheophytins (BP), quinones, a ferrous iron ion and a carotenoid as prosthetic groups. The L and M subunits each form five transmembrane helices, making up the central part of the PRC. Together with their prosthetic groups, they show a high degree of local twofold symmetry perpendicular to the membrane plane.

With the exception of the carotenoid, the prosthetic groups in the L–M complex are arranged in two symmetric branches, termed the A and B branch. These prosthetic groups carry out the primary reactions associated with charge separation and the subsequent secondary reactions associated with the conversion of excitonic energy to chemical energy. However, the PRC only utilizes the A-side prosthetic groups for electron transfer. The key molecular components are a BChl special pair (P), a BChl monomer (BC), a bacteriopheophytin (BP), a quinone (Q_A) and the four-heme c-type cytochrome (Cyt). The structural organization of the pigments in the PRC protein complex is shown in Figure 16.25 and schematically in Figure 16.21. They are held in a fixed and precise geometry that is optimized for efficient electron transfer, by the polypeptide chains. Since P belongs to both branches, its two BChl are denoted by subscripts L and M according to the subunit to which their Mg^{2+} is linked. The special pair P_L–P_M is located near the periplasmic outer membrane surface on the symmetry axis while the cytoplasmic inner membrane surface face lies at the level of Q_A (Figure 16.25). A ferrous ion is bound between the quinones close to the symmetry axis. It can be removed or exchanged with several divalent metals [92] without impairing the function of the PRC. The carotenoid is associated with the BChl of the B branch and is proposed to protect the PRC by quenching the triplet state of P before it can sensitize the formation of singlet oxygen, a powerful oxidizing agent. Crystallographic and spectroscopic data show that the carotenoid molecule is not in an all-*trans* conformation, but has a single *cis* bond near the center of the polyene chain [93].

The energy diagram for electron transfer in the purple bacteria PRC and the rate constants for each step are shown in Figure 16.25. Charge separation and energy transfer is initiated by

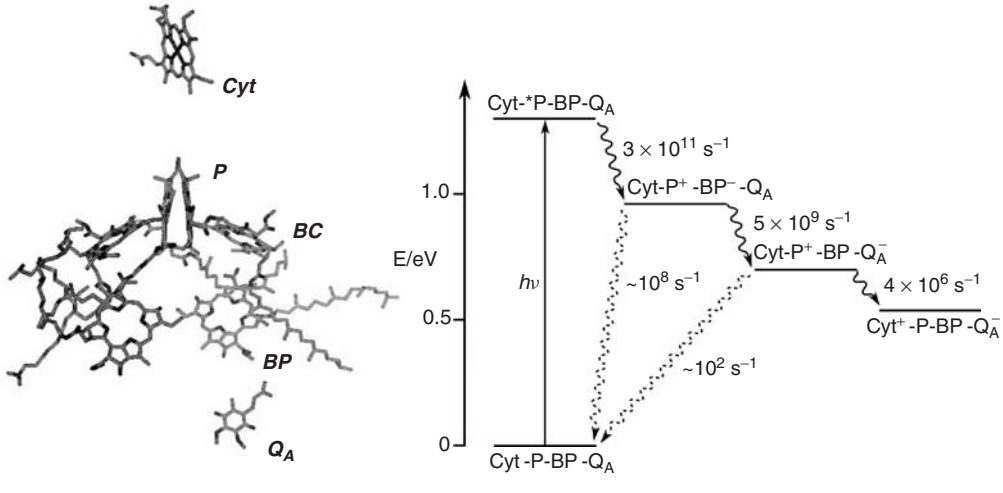


Figure 16.25 A simplified view of the organization of the electron transfer chain in the purple bacteria reaction center and the energy diagram showing the rate constants for each electron transfer step (Copyright Wiley-VCH Verlag GmbH & Co. KGaA.)

absorption of a photon or by exciton diffusion from the light harvesting antennae LH1 and LH2 to the "special chlorophyll pair" P producing a lowest excited state (P*). This is followed by a rapid electron transfer (~3 ps) to BP, the primary acceptor, leading to the formation of the P^+-BP^- intermediate. The electron is then transferred to from BP^- to the terminal electron acceptor, Q_A, in a 200 ps time frame. The P^+ moiety is reduced to its ground state via Cyt. The high efficiency of the charge separation process is due to the fact that the recombination steps are slower because they lie in the Marcus inverted region. The $Q_A{}^-$ group in the reaction center is reoxidized by electron transfer to a secondary quinone, Q_B. This ubiquinone accepts two photogenerated electrons and two protons, forming a hydroquinone, before dissociating from the reaction center and migrating to the cytochrome bc_1 complex in a process that takes ~200 μs. A multistep electron and proton transfer process via protein complexes ensues and results in formation of adenosine triphosphate (ATP), the universal cellular energy currency, filling the majority of the energy needs for the bacterium.

The features that lead to efficient absorption and translocation of photogenerated electrons and their subsequent conversion to chemical energy are the geometric organization of the molecular components coupled to the thermodynamic and kinetic properties of the various electron transport steps that favorably direct electrons in one direction and inhibit charge dissipation. This knowledge lays the foundation for the development of key criteria for designing synthetic systems for development of bio-inspired organic photovoltaics and solar fuel systems.

16.6.2 Artificial Reaction Centers

There have been many attempts to construct artificial systems that are capable of mimicking the function of the natural reaction center [93]. As described above, the natural systems utilize a series of short-range, fast and efficient, electron transfer reactions that lead to charge separation over long distances. A similar strategy has been employed for the development of synthetic systems formed by covalently linking various molecular components such that absorption of light is closely followed by a series of electron transfer steps that lead to a charge separated state. These systems have at a minimum single electron donor/acceptor pair called a dyad, but are usually more complex consisting of up to 6 molecular components (hexads). Mechanistic investigations of photoinduced charge separation in artificial systems have been performed on covalently linked organic compounds (for a recent review see reference [95]) in organic solvents. The excitation of the chromophore component is followed by a primary photoinduced electron transfer to a primary acceptor. This is followed by a thermal electron transfer process from a donor component to the oxidized chromophore, generating the charge separated state. The primary process competes with excited state deactivation while the secondary process competes with primary charge recombination. The final step is charge recombination between remote molecular components, leading to recovery of the ground state. An example of a triad is shown Figure 16.26. It consists of porphyrin as a light absorbing chromophore covalently linked to fullerene C_{60} and carotenoid as primary and secondary acceptors, respectively [96].

The electron transfer events leading to charge separation are shown schematically in the energy diagram in Figure 16.26. In 2-methyltetrahydrofuran, absorption of a photon leads to generation of a single excited state porphyrin ($C-{}^1P-C_{60}$) that rapidly and almost exclusively forms $C-P^+-C_{60}{}^-$ ($k_2 = 0.33\,ps^{-1}$) with a quantum yield of unity. A small fraction of $C-P-{}^1C_{60}$ excited states (step 3) are also formed but these decay to $C-P^+-C_{60}{}^-$ by electron transfer. The porphyrin ground state is recovered by electron transfer from the carotenoid to the porphyrin ($k_8 = 0.15\,ns^{-1}$), generating the $C^+-P-C_{60}{}^-$ charge separated state with a quantum yield of 0.88. It decays slowly by charge recombination to yield the carotenoid triplet state ($^3C-P-C_{60}$) with a rate constant of $2.9\,\mu s^{-1}$. For comparison, the triad geometry of the bacterial reaction center has a quantum yield of unity and a charge recombination constant of $\sim 10^2\,s^{-1}$. This triad has properties found in-the natural photosynthetic center, but not common in other artificial reaction centers. The $C-P^+-C_{60}{}^-$

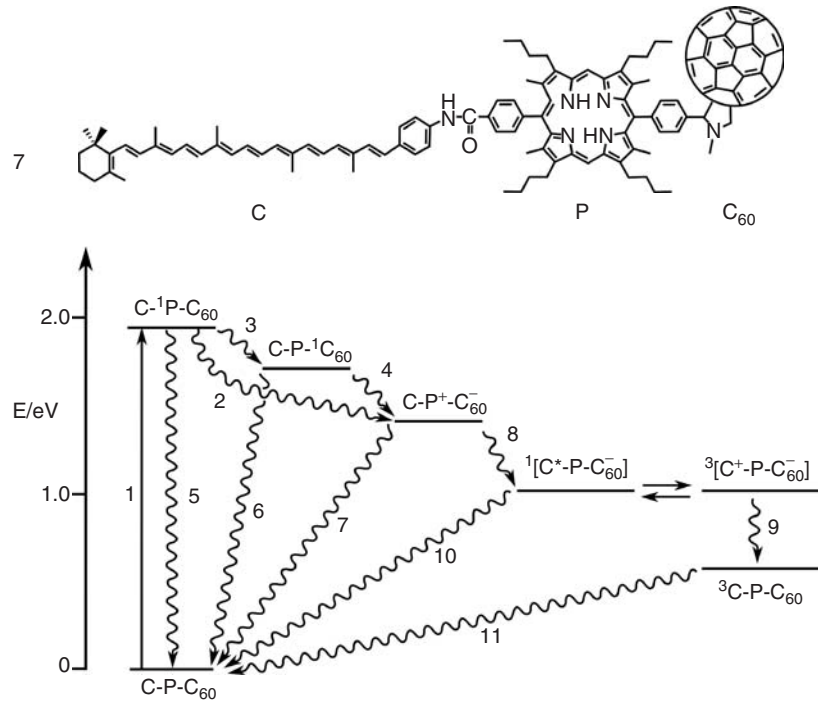


Figure 16.26 Structure of a triadic artificial reaction center and its energy-level diagram showing relevant transient species and interconversion pathways during charge separation processes. C, carotene; P, porphyrin; C_{60}, fullerene. Adapted from reference [96]

charge separated state occurs even at 8 K and recombination of $C{-}P^{+}{-}C_{60}{}^{-}$ yields $^3C{-}P{-}C_{60}$ with a unique EPR-detectable spin-depolarization pattern, two properties found in the natural reaction center, but not common in other artificial reaction centers. As described previously, in natural systems carotenoids quench chlorophyll triplets, thus providing protection against photodegradation caused by the formation of singlet oxygen. Interestingly, in toluene the $C{-}P{-}^1C_{60}$ excited state does not undergo charge separation but decays to a $^3C{-}P{-}C_{60}$ via a $C{-}P{-}^3C_{60}$ state. This highlights the effect of the local environment on the reactivity of the reaction center.

16.7 TOWARDS DEVICE ARCHITECTURES

Progress in integrating naturally occurring photosynthetic complexes with solid-state electronics provides insight into how these proteins and their synthetic bioinspired counterparts can be incorporated into device architectures.

The most common approach for incorporating photosynthetic reaction centers (PRCs) into photovolatics, electrochemical devices, and nano-optoelectronics is by their adsorption on electrode surfaces. The orientation of the adsorbed protein layers on an electrode surface is a critical concern because they behave like photodiodes, enabling unidirectional electron transfer from the special chlorophyll pair to the terminal acceptor site (See Section 16.6.1), once a photon of light is captured. Incorporation of a self-assembled monolayer (SAM) as an interlayer on the electrode provides a surface for physically or electrostatically orienting PRCs. Introduction of a specific tag such as a

polyhistidine tag by genetic mutation of the protein backbone has also been used to attach and orient the PRC on Ni-nitrilotriacetic acid (Ni-NTA)-terminated SAM. In most applications a monolayer of protein is desirable and studies have shown that the packing of the protein on the electrode surfaces can be increased by using vacuum deposition during fabrication.

The nature of the protein–electrode junction is also important due to the distance of the redox moieties in the protein from the electrode surface. Hence, the electronic coupling of between the protein and metal surface is important for the overall efficiency of the device. Das *et al*. [97] reported incorporation of the photosynthetic reaction center from purple bacteria into a solid-state device architecture. The PRCs were self-assembled on gold–indium–tin oxide surfaces and stabilized with peptide detergents before coating with fullerene (C_{60}) followed by silver. Direct communication with the electrodes was achieved using this approach and the short-circuit current density of the device was 0.12 mA/cm^2 under an excitation intensity of 10 W/cm^2 ($\lambda = 808$ nm). The internal quantum efficiency of the device was reported to be 12%.

Similar strategies have been employed for devices based on plant Photosystem I (PSI) (Figure 16.27). As described above, PSI is a transmembrane multisubunit protein–chlorophyll complex, isolated from plants and cyanobacteria that mediates vectorial light-induced electron transfer. Its nanosize dimensions, an internal energy yield of approximately 58% (23% of solar radiation), and its ability to generate a photovoltage of approximately 1 V with a quantum efficiency approaching 1 [99], make this molecule a promising unit for applications in molecular nanoelectronics and solar conversion applications. Dried monolayers of oriented PSI complexes, either directly attached to gold electrodes or assembled on mercaptoethanol amine SAMs generate photovoltages of 0.45V [100] and 1V [101] on illumination, respectively. The distance from the special chlorophyll pair to an electrode surface is 28 Å and distance from the terminal acceptor to the electrode is 17 Å. These relatively long distances make direct electron transfer to the electrode difficult. For this reason, only electrochemical cells that employ soluble redox mediators to facilitate communication between the protein the electrode surfaces have been reported to date. The light-induced current from these devices is currently less than 1 μA/cm^2 [98, 102].

This highlights a potential application where conductive polymers could play a role in realizing biohybrid photovoltaics. The energy levels in PSI are −4.58, −2.78, and −3.52 eV for the special chlorophyll pair (P700), P700 excited state, and the terminal FeS electron acceptor, respectively [100]. Therefore, a conductive polymer with bandgaps optimized for electron transfer (see

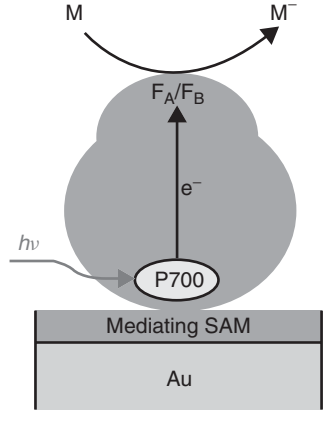


Figure 16.27 Schematic representation of attachment and orientation of PSI on an electrode surface. M. soluble mediator. Adapted from [98]

Figure 16.11) to and from PSI could increase the overall efficiency of electron transfer without compromising the photovoltaic properties on the protein complex. Devices based on the artificial light harvesting and reaction centers described in Sections 16.5 and 16.6 have not yet been demonstrated. However, these structures have the potential to surpass the efficiency and current density possible using natural systems because the relatively large size of PRCs (e.g. PSI is $17 \times 15 \times 9$ nm) will dictate the packing density of these complex structures in surfaces.

16.8 SUMMARY AND FUTURE PERSPECTIVES

The current relatively small (typically less than 6%) photoelectric power conversion efficiencies of organic and polymeric photovoltaic devices can be attributed mainly to the three severe losses [5, 34], including the "photon loss", the "exciton loss", and the "carrier loss" due to materials improper frontier orbital energy levels, energy gaps, energy offsets between the donors and the acceptors, poor material morphologies, and not yet optimized cell structures and fabrications. However, there is plenty of room for improvement. Optimizations in both the spatial and energy domains should be pursued simultaneously and systematically to achieve high-efficiency organic and polymeric photovoltaic devices [7].

In the spatial domain, a donor/acceptor bicontinous ordered nanostructure (BONS) morphology appears most promising to minimize both the exciton loss and the carrier loss [5, 7, 33–35], as the diameter of donor or acceptor phase could be controlled to be within the average exciton diffusion length (AEDL), depending upon the materials involved, and the uninterrupted charge transport pathway could be aligned perpendicular to the electrode planes. As discussed in an earlier section, several approaches are being pursued that may potentially achieve this structure. In the cases of carbon nanotubes and semiconducting nanorods, the main challenges are to fabricate cells with uniformly spaced and well-aligned tubes/rods perpendicular to the horizontal conducting substrates. In the cases of block copolymers, the main challenge lies in both synthetic chemistry and block copolymer processing. Figure 16.28 shows a potential "HEX"-style columnar tertiary morphology of a donor/acceptor type block copolymer [33–35].

In the energy domain, the optical excitation energy gaps in both donor and acceptor phases should match the intended photon energy. In addition, the donor/acceptor energy offset as well as the D-LUMO/A-LUMO electron transfer coupling element should be optimized to maximize photoinduced charge separation and minimize charge recombination (i.e. the A-LUMO/D-HOMO electron transfer has a relatively small coupling element, and energetically not balanced). Specifically, as shown in Figure 16.29, in electron transfer dynamic regime, there exists an optimal donor/acceptor LUMO (or HOMO) level offset where exciton dissociation is most efficient (or exciton quenching parameter (Y_{eq}) reaches its maximum [5, 7, 17, 18, 37]), and another optimal LUMO (or HOMO) level offset where charge recombination is relatively slow compared to charge separation (or recombination quenching parameter Y_{rq} become largest). Though the cell efficiency is mainly affected by the Y_{eq}, the molecules should be designed and developed such that the maximum Y_{rq} is close to or coincides with maximum Y_{eq} [5, 7, 17, 18, 37]. There also exists a third energy offset where the charge recombination (k_r) becomes most severe. The molecules should be designed and developed such that this worst charge recombination k_r is far away from maximum k_{eq}. These orbital offset values are critically important in molecular structure and energy level fine tuning. Finally, since sunlight is a very broad radiation with photon energy ranging from UV all the way to IR, a tandem style serially connected and parallel stacked cell structure with energy gaps gradually descending from UV to IR along the light radiation direction appears ideal. This would enable a broad capture of most solar photons and the summation of induced photovoltage.

As described in the second section of this chapter, the knowledge gained by investigation and elucidation of the molecular mechanisms of light harvesting, charge separation and energy

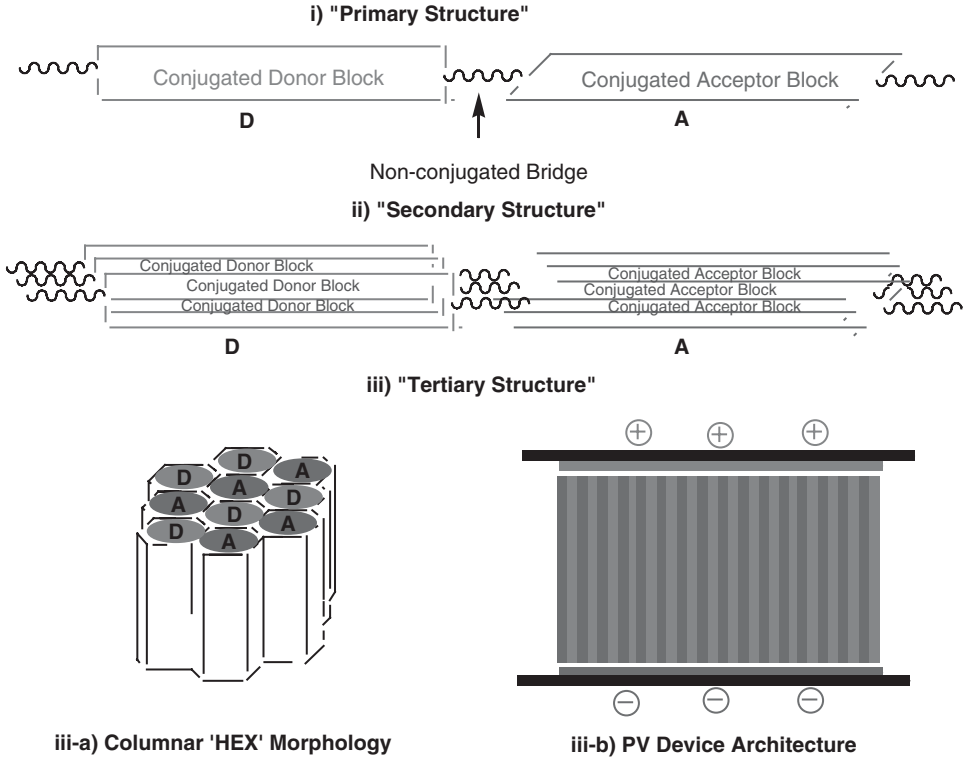


Figure 16.28 Scheme of a donor/acceptor block copolymer "primary structure", the "secondary structure", and the "tertiary structure" [33–35]

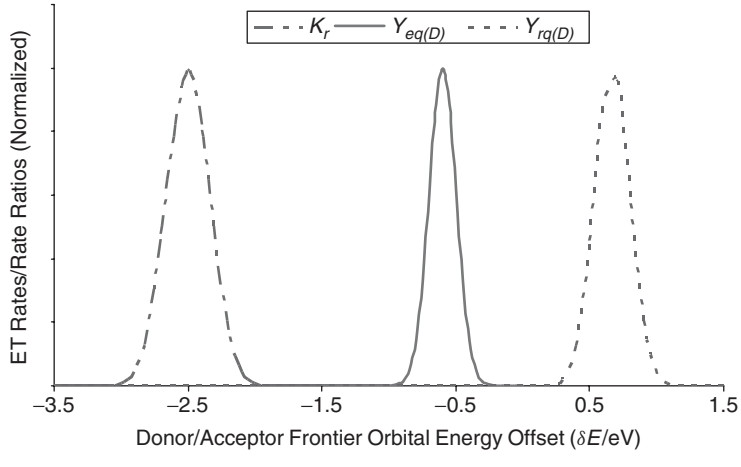


Figure 16.29 The RO-PPV exciton quenching parameter ($Y_{eq(D)} = k_{s(D)}/k_{d(D)}$, middle solid curve), charge recombination rate constant (K_r, left long dashed curve), and charge recombination quenching parameter ($Y_{rq(D)} = k_{s(D)}/k_{r(D)}$, right short dashed curve) versus LUMO offset of RO-PPV/SF-PPV-I pair [17, 18]

storage in natural photosynthetic systems has provided much inspiration for the development of synthetic analogs. Many of the strategies employed by natural systems could be adapted for development of organic photovoltaics. There have been many advances made on each aspect of this complex problem, but the goal of integrating functional components into a complete working system, that is both efficient and robust, has not yet been realized. The examples provided in the second section of this chapter are representative of the synthetic strategies that are currently under investigation. Although there has been much progress made in development of covalently linked functional building blocks that are capable of light harvesting and charge separation, chemistry based on covalent synthesis is highly inefficient and costly for synthesis of supramolecular integrated arrays. Self-assembly approaches will be required to achieve ordered complex architectures from functionalized building blocks. Fundamental concepts of how to achieve such structures, with the ability to self-assemble into complete artificial photosynthetic systems, are largely unknown [94]. In natural systems, supramolecular organization is achieved by supporting and orienting individual functional components on a polypeptide scaffold. There have been several reports that have taken inspiration from natural photosynthetic proteins for the development of synthetic analogues that employ a polypeptides to bind and orient chromophores and are capable of charge separation and light harvesting reactions [103–105].

Although developments in organic photovoltaics and bioinspired photosynthesis has, to date, developed on parallel paths, it is clear that there is much synergy between these two areas of research. With further advances in polymer synthesis, processing and characterizations, these approaches hold great promise for design and synthesis of polymers that can provide both structural and functional support for exciton diffusion, charge separation and charge transport in self-assembled systems. It is expected that high-efficiency photoelectric power conversion efficiency organic light harvesting systems, including organic photovoltaic cells, photodetectors, or any artificial photo-charge synthesizers/converters can be achieved. The dream of renewable, clean, inexpensive, portable or lightweight, and low-cost energy supply can be a reality.

ACKNOWLEDGEMENTS

The authors would like to thank the research/educational grant supports from a number of funding agencies including NASA, Department of Defense (MDA, AFOSR, ARO), National Science Foundation, Department of Education, and the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory (ORNL), and in particular, a faculty summer fellowship (awarded to Prof. Sun) from Department of Energy Oak Ridge National Lab. ORNL is managed by UT-Battelle, LLC for the Department of Energy under contract No.DE-AC05-00OR22725. This work has been coauthored by a contractor of the US Government under contract No. DE-AC05-00OR22725. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for US Government purposes.

REFERENCES

1. Archer MD, Hill R, (eds), *Clean Electricity From Photovoltaics*, Imperial College Press, London (2001).
2. Luque A, Hegedus S, (eds), *Handbook of Photovoltaic Science and Engineering*, 1st edn, John Wiley & Sons, Ltd, Chichester (2003).
3. Green MA, *Third Generation Photovoltaics: Advanced Solar Energy Conversion*, Springer, Berlin (2003).
4. Lazmerski L, *J. Elect. Spec*., **150**, 105–135 (2006).

5. Sun S, Sariciftci NS, (eds), *Organic Photovoltaics: Mechanisms, Materials and Devices*, CRC Press, Boca Raton, Florida (2005).
6. Brabec C, Dyakonov V, Scherf U, (eds), *Organic Photovoltaics: Materials, Device Physcis, and Manufacturing Technologies*, Wiley-VCH, Berlin (2008).
7. Sun S, Organic and Polymeric Solar Cells, in *Handbook of Organic Electronics and Photonics*, Nalwa HS, (ed), American Scientific Publishers, Los Angeles, California (2008), Vol. 3, Chapter 7, pp 313–350.
8. Sun S, Dalton L, (eds), *Introduction to Organic Electronic and Optoelectronic Materials and Devices*, CRC Press/Taylor Francis, Boca Raton, Florida (2008).
9. Skotheim TA, Reynolds JR, (eds), *Handbook of Conducting Polymers*, 3rd edn, CRC Press: Boca Raton, Florida (2007).
10. Podzorov V, Menard E, Rogers JA, Gershenson ME, *Phys. Rev. Lett*., **95**, 226601 (2005).
11. Frenkel YI, *Phys. Rev*., **37**, 1276 (1931).
12. Wannier NF, *Phys. Rev*., **52**, 191 (1937).
13. Blankenship R, in *Organic Photovoltaics: Mechanisms, Materials and Devices*, Sun S, Sariciftci NS, (eds), CRC Press, Boca Raton, Florida (2005), p. 37.
14. Kolodinski S, Werner J, Queisser H, *Sol. Energy Mater. Sol. Cells*, **33**, 275–285 (1994).
15. Kumar A, Ner Y, Sotzing G, in *Introduction to Organic Electronic and Optoelectronic Materials and Devices*, Sun S, Dalton L, (eds), CRC Press, Boca Raton, Florida (2008), Chapter 7, p. 211.
16. Knupfer M, *Appl. Phys. A*, **77**, 623–626 (2003).
17. Sun S, *Sol. Energy Mat. Sol. Cells*, **85**, 261–267 (2005).
18. Sun S, *Mater. Sci. Eng. B*., **116** (**3**), 251–256 (2005).
19. Sariciftci NS, Smilowitz L, Heeger AJ, Wudl F, *Science*, **258**, 1474 (1992).
20. Kraabel B, Hummelen J, Vacar D, Moses D, Sariciftci N, Heeger A, Wudl F, *J. Chem. Phys*., **104**, 4267–4273 (1996).
21. Gregg B, in *Organic Photovoltaics: Mechanisms, Materials and Devices*, Sun S, Sariciftci NS, (eds), CRC Press, Boca Raton, Florida (2005), p. 139.
22. Tang CW, *Appl. Phys. Lett*. **48**, 183–185 (1986).
23. Yu G, Gao J, Hummelen J, Wudl F, Heeger A., *Science* **270**, 1789–1791 (1995).
24. Brabec CJ, Cravino A, Meissner D, Sariciftci NS, Fromherz T, Minse M, Sanchez L, Hummelen JC, *Adv. Funct. Mater*., **11**, 374–380 (2001).
25. Perlin J, *From Space to Earth-The story of Solar Electricity*, AATEC Publications, Ann Arbor, Michigan (1999).
26. Pochettino, *Acad. Lincei Rendiconti*, **15**, 355–363 (1906).
27. Hiramoto M, Fujiwara H, Yokoyama M., *Appl. Phys. Lett*. **58**, 1062–64 (1991).
28. Ma W, Yang C, Gong X, Lee K, Heeger A, *Adv. Funct. Mater*., **15**, 1617–1622 (2005).
29. Li G, Shrotriya V, Huang J, Yao Y, Moriarty T, Emery K, Yang Y, *Nature Materials*, 4, 864–868 (2005).
30. Drees M, Davis R, Heflin R, in *Organic Photovoltaics: Mechanisms, Materials and Devices*, Sun S, Sariciftci NS, (eds), CRC Press, Boca Raton, Florida (2005), p. 559.
31. Coakley K, McGehee M, *Chem. Mater*., **16**, 4533–4542 (2004).
32. Fan Z, Wang Y, Haliburton J, Maaref S, Sun, S, NASA Tech Report NONP-NASA-CD-2002153469, NASA/STI Accession number: 20030016552, January, 2001.
33. Sun S, Photovoltaic Devices Based on a Novel Block Copolymer, *US Patent # 20040099307* (Appl.: November 14, 2002).
34. Sun S, *Sol. Energy Mat. Sol. Cells*, **79**, 257–264 (2003).
35. Sun S, *Poly. Mater. Sci. Eng*., **88**, 158 (2003).
36. Sun S, Fan Z, Wang Y, Haliburton J, Taft C, Seo K, Bonner C, *Syn. Met*., **137**, 883–884 (2003).
37. Sun S, Fan Z, Wang Y, Haliburton J, *J. Mater. Sci*., **40**, 1429–1443 (2005).

38. de Boer B, Stalmach U, van Hutten PF, Melzer C, Krasnikov VV, Hadziioannou G, *Polymer*, **42**, 9097 (2001).
39. Zhang C, Choi S, Haliburton J, Li R, Cleveland T, Sun S, Ledbetter A, Bonner C, *Macromolecules*, **39**, 4317 (2006).
40. Sun S, Zhang C, Choi S, Ledbetter A, Bonner C, Drees M, Sariciftci S, *Appl. Phys. Lett*., **90**, 043117 (2007).
41. Kang Y, Park N, Kim D, *Appl. Phys. Lett*., **86**, 113101 (2005).
42. Jin M, Dai L, in *Organic Photovoltaics: Mechanisms, Materials and Devices*, Sun S, Sariciftci NS, (eds), CRC Press, Boca Raton, Florida (2005), p. 579.
43. Schmidt-Mende L, Fechtenkötter A, Müllen K, Moons E, Friend RH, MacKenzie JD, *Science*, **293**, 1119 (2001).
44. Kippelen B, Yoo S, Haddock J, Domercq B, Barlow S, Minch B, Xia W, Marder S, Armstrong N, in *Organic Photovoltaics: Mechanisms, Materials and Devices*, Sun S, Sariciftci NS, (eds), CRC Press, Boca Raton, Florida (2005), p. 271.
45. Tsarenkov GV, *Sov. Phys. Semicond*. **9**(2), 166–171 (1975).
46. Xue J, Uchida S, Rand B, Forrest SR, *Appl. Phys. Lett*., **86**, 5757 (2005).
47. Kim J, Lee K, Coates NE, Moses D, Nguyen T, Dante M, Heeger AJ, *Science*, **317**, 222 (2007).
48. Dennler G, Prall H, Koeppe R, Egginger M, Autengruber R, Sariciftci NS, *Appl. Phys. Lett*., **89**, 073502 (2006).
49. Djuristic A, Kwong CY, in *Organic Photovoltaics: Mechanisms, Materials and Devices*, Sun S, Sariciftci NS, (eds), CRC Press, Boca Raton, Florida (2005), p. 453.
50. Karg S, Riess W, Dyakonov V, Schwoerer M, *Synth. Metals*, **54**, 427 (1993).
51. Marks R, Halls J, Bradley D, Friend R, Holmes A, *J. Phys.: Cond. Matter*, **6**, 1379 (1994).
52. Antoniadis H, Hsieh B, Abkowitz M, Jenekhe S, Stolka M, *Synth. Metals*, **62**, 265 (1994).
53. Frankevich E, *et al*., *Phys. Rev*. *B* **46**, 9320 (1992).
54. Rothberg L, *et al., Synth. Metals*, **80**, 41 (1996).
55. Hauch JA, Schilinsky P, Choulis SA, Childers R, Biele M, Brabec CJ, *Solar Energy Matl and Solar Cell* **92**, 727–731 (2008).
56. Hoth C, Schilinsky P, Choulis S, Brabec C *Nano-letters* **8**, 2806–2813 (2008).
57. Tipnis R, Bernkopf J, Jia S, Krieg J Li S, Storch M Laird D *Solar Energy Materials and Solar Cells* **93**, 442–446 (2009).
58. Dennler G, Lungenschmied, Neugebauer H, Sariciftic N, *J Material Research* **20**, 3224–3232 (2005).
59. Green BR, Parson WW, Kluwer Academic Publishers, Dordrecht 2003.
60. Blankenship RE, in Sun S, Sariciftci NS, (eds) *Organic Photovoltaics: Mechanism, Materials, and Devices*, CRC Press, Taylor Francis Group, Boca Raton 2005.
61. Forester T, *Discussion Faraday Transactions* **27**, 7 (1959).
62. Dexter DL, *Journal of Chemical Physics* **21**, 836 (1953).
63. van Amerongen H, *Photosynthetic excitons*, World Scientific, Singapore, 2000.
64. Blankenship RE, *Photosynthesis Research* **33**, 91–111 (1992).
65. Brettel K, *Biochimica et Biophysica Acta-Bioenergetics* **1318**, 322–373 (1997).
66. Lee I, Lee JW, Stubna A, Greenbaum E, *Journal of Physical Chemistry B* **104**, 2439–2443 (2000).
67. Zankel KL, Reed DW, Clayton RK, *Proceedings of the National Academy of Sciences of the United States of America* **61**, 1243 (1968).
68. Hiyama T, *Physiologie Vegetale* **23**, 605–610 (1985).
69. Blankenship RE, *Molecular Mechanisms of Photosynthesis*, Blackwell Science, Oxford, UK, 2002.
70. Armaroli N, Balzani V, *Angewandte Chemie-International Edition* **46**, 52–66 (2007).
71. Nakamura Y, Aratani N, Osuka A, *Chemical Society Reviews* **36**, 831–845 (2007).

72. Mcdermott G, Prince SM, Freer AA, Hawthornthwaitelawless AM, Papiz MZ, Cogdell RJ, Isaacs NW, *Nature* **374**, 517–521 (1995).
73. Koepke J, Hu XC, Muenke C, Schulten K, Michel H, *Structure* **4**, 581–597 (1996).
74. Trinkunas G, Herek JL, Polivka T, Sundstrom V, Pullerits T, *Physical Review Letters* **86**, 4167–4170 (2001).
75. Hess S, Feldchtein F, Babin A, Nurgaleev I, Pullerits T, Sergeev A, Sundstrom V, *Chemical Physics Letters* **216**, 247–257 (1993).
76. Burrell AK, Officer DL, Plieger PG and Reid DCW, *Chemical Reviews* **101**, 2751–2796 (2001).
77. Wurthner F., You CC, Saha-Moller CR, *Chemical Society Reviews* **33**, 133–146 (2004).
78. Wojaczynski J, Latos-Grazynski L, *Coordination Chemistry Reviews* **204**, 113–171 (2000).
79. Aratani N, Takagi A, Yanagawa Y, Matsumoto T, Kawai T, Yoon ZS, Kim D, Osuka A., *Chemistry–a European Journal* **11**, 3389–3404 (2005).
80. Nakamura Y, Hwang IW, Aratani N, Ahn TK, Ko DM, Takagi A, Kawai T, Matsumoto T, Kim D, Osuka A, *Journal of the American Chemical Society* **127**, 236–246 (2005).
81. Choi MS, Aida T, Yamazaki T, Yamazaki I, *Angewandte Chemie-International Edition* **40**, 3194 (2001).
82. Balzani V, Campagna S, Denti G, Juris A, Serroni S, Venturi M, *Accounts of Chemical Research* **31**, 26–34 (1998).
83. Cotlet M, Vosch T, Habuchi S, Weil T, Mullen K, Hofkens J, De Schryver F, *Journal of the American Chemical Society* **127**, 9760–9768 (2005).
84. Hahn U, Gorka M, Vogtle F, Vicinelli V, Ceroni P, Maestri M, Balzani V, *Angewandte Chemie-International Edition* **41**, 3595–3598 (2002).
85. Montano GA, Bowen BP, LaBelle JT, Woodbury NW, Pizziconi VB, Blankenship RE, *Biophysical Journal* **85**, 2560–2565 (2003).
86. Psencik J, Ikonen TP, Laurinmaki P, Merckel MC, Butcher SJ, Serimaa RE, Tuma R, *Biophysical Journal* **87**, 1165–1172 (2004).
87. Balaban TS, *Accounts of Chemical Research* **38**, 612–623 (2005).
88. Balaban TS, Linke-Schaetzel M, Bhise AD, Vanthuyne N, Roussel C, Anson CE, Buth G, Eichhofer A, Foster K, Garab G, Gliemann H, Goddard R, Javorfi T, Powell AK, Rosner H, Schimmel T, *Chemistry–a European Journal* **11**, 2268–2275 (2005).
89. Pierson BK, Olson JM, in Amesz, J, (ed.) *Photosynthesis*, Elsevier, Amsterdam 1987, pp. 21–42.
90. Deisenhofer J, Epp O, Miki K, Huber R, Michel H, *Nature* **318**, 618–624 (1985).
91. Allen JP, Feher G, Yeates TO, Komiya H, Rees DC, *Proceedings of the National Academy of Sciences of the United States of America* **84**, 5730–5734 (1987).
92. Debus RJ, Feher G, Okamura MY, *Biochemistry* **25**, 2276–2287 (1986).
93. Frank HA, in Deisenhofer J, Norris JR, (eds.) *The Photosynthetic Reaction Center*, Academic Press, San Deigo 1993, pp. 221–237.
94. Balzani V, Credi A, Venturi M, *Chemsuschem* **1**, 26–58 (2008).
95. Wasielewski MR, *Journal of Organic Chemistry* **71**, 5051–5066 (2006).
96. Gust D, Moore TA, Moore AL, *Journal of Photochemistry and Photobiology B–Biology* **58**, 63–71 (2000).
97. Das R, Kiley PJ, Segal M, Norville J, Yu AA, Wang LY, Trammell SA, Reddick LE, Kumar R, Stellacci F, Lebedev N, Schnur J, Bruce BD, Zhang SG, Baldo M, *Nanolett*. **4**, 1079–1083 (2004).
98. Faulkner CJ, Lees S, Ciesielski PN, Cliffel DE, Jennings GK, *Langmuir* **24**, 8409–8412 (2008).
99. Brettel K, Leibl W, *Biochim. et Biophys. Acta–Bioenergetics* **1507**, 100–114 (2001).
100. Carmeli I, Frolov L, Carmeli C, Richter S, *J Amer Chem Soc* **129**, 12352 (2007).
101. Lee I, Lee JW, Stubna A, Greenbaum E, *J. Phys Chem. B* **104**, 2439–2443 (2000).

102. Terasaki N, Yamamoto N, Hiraga T, Yamanoi Y, Yonezawa T, Nishihara H, Ohmori T, Sakai M, Fujii M, Tohri A, Iwai M, Inoue Y, Yoneyama S, Minakata M, Enami I, *Ang. Chem.Int Ed*. **48**, 1585–1587 (2009).
103. Ye SX, Discher BM, Strzalka J, Xu T, Wu SP, Noy D, Kuzmenko I, Gog T, Therien MJ, Dutton PL, Blasie JK, *Nano Letters* **5**, 1658–1667 (2005).
104. Miller RA, Presley AD and Francis MB, *Journal of the American Chemical Society* **129**, 3104–3109 (2007).
105. Liang Y, Guo P, Pingali SV, Pabit S, Thiyagarajan P, Berland KM, Lynn DG, *Chemical Communications* 6522–6524 (2008).

17

Transparent Conducting Oxides for Photovoltaics

Alan E. Delahoy[1,2] and Sheyu Guo[1,3]

[1]*EPV SOLAR, Inc., Robbinsville, NJ, USA,* [2]*New Millennium Solar Equipment Corp., Robbinsville, NJ, USA,* [3]*Yiri Solartech Co., Ltd., Suzhou, P. R. China*

17.1 INTRODUCTION

17.1.1 Transparent Conductors

Transparent conductors are materials, generally in the form of thin films, which simultaneously possess the properties of optical transmission and electrical conductivity. These properties are most commonly realized through the use of a heavily doped, wide-bandgap semiconductor (usually a metal oxide), although metal films (e.g. Ag), doped organic polymers, or metal nitrides (e.g. TiN) are occasionally used as a transparent conductor. The most common transparent conducting oxides (TCO) are based on tin oxide, indium oxide, zinc oxide, and cadmium oxide. Combinations of these materials (e.g. ZnO-SnO_2) have also been prepared, as have ternary compounds such as cadmium stannate (Cd_2SnO_4). Thin films of these materials are largely transparent in the visible portion of the spectrum since the photon energy E_{ph} (1.8–3.0 eV) is less than the bandgap E_g of the material (typically 3.2–3.8 eV) and hence the photon cannot be absorbed. Most oxides are highly insulating. However, the TCO is rendered conductive by free carriers resulting from the introduction of suitable substitutional impurity atoms, from deviations from stoichiometry, or occasionally from other impurities. The commonly used TCOs are *n*-type, meaning that the free carriers are electrons. The high free carrier concentration leads to a third property of a TCO, viz. high infrared reflectivity.

The most extensively used material is fluorine-doped tin oxide (SnO_2:F) which is used as a heat-reflecting coating on architectural glass (low emissivity glass for energy-efficient windows) and as a transparent electrode for thin film amorphous silicon (a-Si) and cadmium telluride (CdTe) based solar cells. The next most extensively used material is tin-doped indium oxide (ITO) which is used in flat-panel displays (FPD), high-definition TVs, touch screens, and certain types of solar

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

cells based on crystalline Si wafers, a-Si, or copper indium gallium diselenide (CIGS). Zinc oxide (ZnO) is also used in thin film Si and CIGS PV technologies. For several of these applications the transparent conducting oxide is deposited onto a rigid substrate such as glass, whereas for touch screens or thin film PV on foils the TCO is deposited onto a flexible polymer sheet (e.g. polyethylene terephthalate or PET) or a flexible coated metal foil. Other applications include heated freezer doors and cockpit windows. Existing applications for ZnO further include varistors, gas sensors, and surface acoustic wave devices. Emerging applications for ZnO include LEDs, lasers, OLEDs, OLED displays, transparent high-mobility TFTs, nanostructured devices, and spintronic devices. There is huge industrial use of TCOs. Consequently, enormous benefits stand to be gained from the development of lower cost and higher performance TCOs tailored to particular applications.

17.1.2 Transparent Conducting Oxides for Photovoltaics

In the photovoltaic arena, most types of thin film solar cells require a TCO as the current-collecting electrode on the sun-facing side of the cell. This is because the lateral conductivity of doped thin film semiconductors that are sufficiently thin to possess high optical transmission is too high for carrier collection over significant distances. The principal types of solar cells that use TCOs are shown schematically in Figure 17.1. Both a-Si and CdTe superstrate thin film PV technologies typically use SnO_2:F as the TCO, although next-generation a-Si/nc-Si cells tend to use ZnO doped

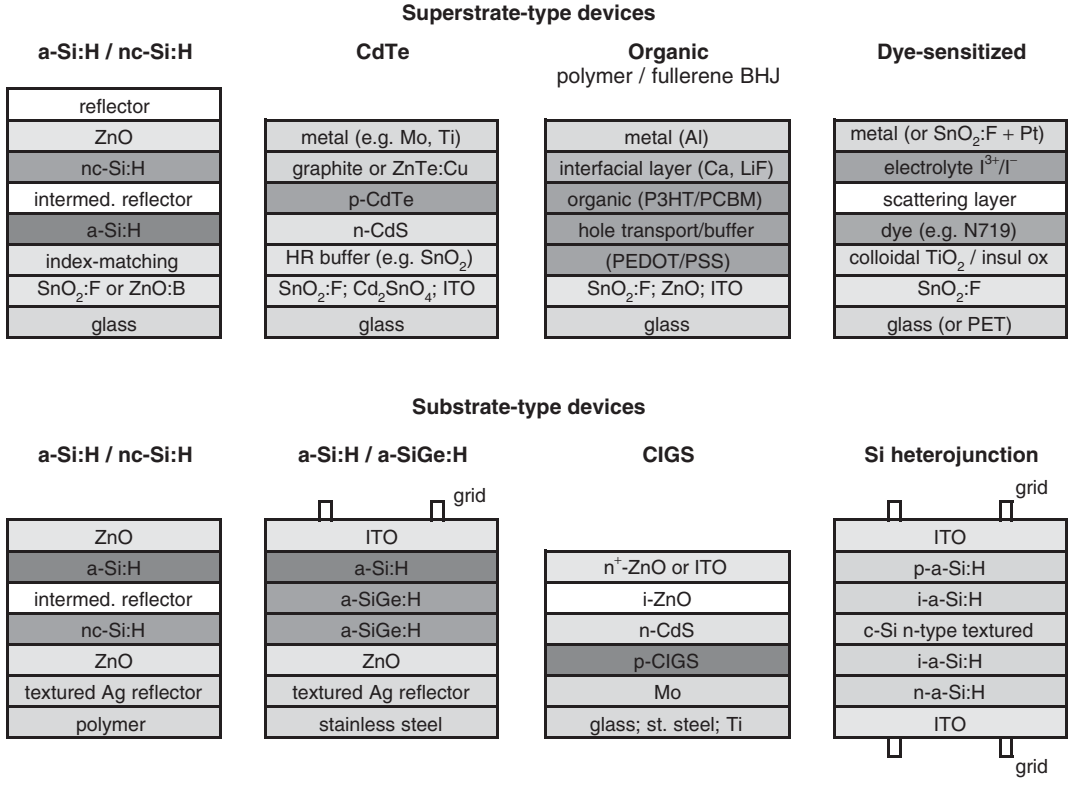


Figure 17.1 The principal types of solar cells that utilize one or more TCO layers in their construction. Top row: superstrate type devices; bottom row: substrate-type devices. See Plate 4 for the colour figure

with either Al (ZnO:Al) or B (ZnO:B). The dye-sensitized TiO_2 type cells also use SnO_2:F, while organic cells have been fabricated on most major types of TCO. Substrate-based technologies, e.g. a-Si on a polymer or steel foil, or CIGS on glass or metal foil, may use either ZnO:Al or ITO as the sun-facing TCO. Although crystalline Si wafer-based cells typically rely on a heavily doped emitter and metal grid to collect current, rather than a TCO, the Sanyo HIT (heterojunction with intrinsic thin layer) cell uses ITO and a metal grid on top of the *p*-type a-Si/*n*-type c-Si base for current collection. Yet another configuration, and one not shown in the figure below, is the mechanically stacked tandem cell. This type of cell requires a rear TCO electrode for the top cell. These applications will be discussed in greater detail in Section 17.5.

17.1.3 Properties, Selection and Trade-offs

The optical transmission as a function of wavelength for a doped TCO exhibits three distinct regions, namely, absorbing for sufficiently short wavelengths ($E_{ph} > E_g$), transmitting for visible wavelengths, and reflecting for $\lambda > \lambda_p$, where λ_p is the plasma wavelength (this typically lies in the IR, see Section 17.4.2). In the high transmission region the optical absorption can be small, but is usually not negligible. The intrinsic electrical properties are determined by the free carrier concentration n_e (/cm^3) and the carrier mobility μ (cm^2/Vs) which together determine the conductivity σ(S/cm) $= n_e e \mu$, where e is the electron charge. The resistivity $\rho(\Omega\,\text{cm}) = 1/\sigma$. The sheet resistance R_{sh}, expressed in ohms per square ($\Omega/\square$), of a thin film is then given by

$$R_{sh} = \frac{\rho}{t} \tag{17.1}$$

where t (cm) is the film thickness. The value of R_{sh} is independent of the size of the square.

Selection of the most appropriate TCO for a given application requires consideration of a large number of properties, not only related to optical and electrical performance, but also to surface morphology, stability during processing, environmental stability, toxicity, and production cost. Other yet more specialized properties, for example, flexibility, hardness, work function, deposition effect on underlying layers, or ease of patterning, may also be relevant. After material selection, the optimization of a TCO is a complex issue involving choice of deposition method and deposition parameters, choice of doping concentration, and choice of thickness. Increasing the dopant concentration will increase the carrier concentration, and hence conductivity, but will also increase optical absorption. Beyond a certain level of doping the carrier mobility will be reduced and the film conductivity may no longer be increased. For solar cell applications in which the TCO is required to carry current laterally, i.e. in the plane of the film, the finite TCO resistivity leads to power dissipation in the film via Joule heating (I^2R losses). While the sheet resistance of a film of given resistivity can be decreased by making the film thicker (hence reducing I^2R losses) this will lead to increased optical absorption in the film. The optimal film thickness required to minimize the total power loss (optical plus electrical) then depends on the integrated optical absorption across the wavelength range of interest and the distance over which current has to travel in the TCO. This is analyzed in detail in Section 17.4.3.

The following sections survey the different types of thin-film TCO, their preparation, their optoelectronic properties, and their application to diverse photovoltaic devices and modules. To support the discussion, the underlying theory of electron transport and optical behavior of these materials is introduced. Specific topics include TCO deposition by sputtering and CVD, the Drude theory for conductivity, the calculation of optical properties, the multiple roles of TCOs within modern PV devices, the growth of textured films, module optimization, film characterization and environmental stability. The concluding section outlines the prospects for future TCO development via concise reviews of selected topical areas such as commercial TCOs, the quest for high electron mobility, and alternative transparent conducting materials such as carbon nanotubes and amorphous TCOs.

17.2 SURVEY OF MATERIALS

In this section we give a quick and broad review of the materials used for transparent conducting oxides, with a focus on those used most commonly for PV applications. For convenience, several tables of material properties or film results are included. Many topics will be discussed in greater detail in the subsequent sections of this chapter.

17.2.1 Classification and Important Types

Transparent conducting oxide (TCO) thin films can be made from binary compounds, ternary compounds, as well as multicomponent oxides [1, 2]. The ternaries and quaternaries are formed through combinations of binary compounds. To date, the TCO films used in commercial PV applications are made from binary compounds as these appear easier to control than ternary or multicomponent materials. The relevant cations can be grouped as divalent (Cd^{2+} and Zn^{2+}), trivalent (Ga^{3+} and In^{3+}), and tetravalent (Sn^{4+}). Examples of these different types of TCOs are as follows:

Binaries: In_2O_3, SnO_2, ZnO, CdO, TiO_2
Ternaries (combination of binaries): Cd_2SnO_4, $CdSnO_3$, Zn_2SnO_4, $CdIn_2O_4$, $Zn_2In_2O_5$, $MgIn_2O_4$, $In_4Sn_3O_{12}$
Quaternaries (combination of ternaries): $Zn_2In_2O_5-MgIn_2O_4$, $ZnIn_2O_5-In_4Sn_3O_{12}$, $GaInO_3-In_4Sn_3O_{12}$, or $In_2O_3-Ga_2O_3-ZnO$

An overview of the more common or important *n*-type TCOs is given in Table 17.1. This table also lists the deposition method and doping elements commonly used for these films, the resistivity range for doped films, and the bandgap of the undoped film. The bandgap of the doped film is generally higher, e.g. the bandgaps of doped CdO and ZnO are 3.1 and 3.7 eV, respectively, compared with 2.4 and 3.3 eV for the undoped film. Despite its very low resistivity, CdO is not used because of toxicity concerns. Cadmium stannate (Cd_2SnO_4) can be prepared with low resistivity

Table 17.1 Overview of important TCOs

TCO	Common deposition methods	Doping element	Resistivity range ($10^{-4}\,\Omega\,cm$)	Bandgap (undoped) (eV)
SnO_2	APCVD, spray pyrolysis	F, Sb, Cl	3–8	3.6
ZnO	Sputtering, PLD, LPCVD, APCVD	Al, Ga, B, In, F	1–8	3.3
In_2O_3	Sputtering, PLD,	Sn, Mo, Ti, Nb, Zr	1–3	3.7
CdO	MOCVD	In, Sn	0.5–20	2.4
TiO_2	Sputtering, PLD	Nb, N	$9-10^6$	3.2
β-Ga_2O_3	Sputtering	Si, Sn	200^*-10^6	4.9
Cd_2SnO_4	Sputtering, sol-gel, spray pyrolysis	self-doped	1.2–10	3.1
Zn_2SnO_4	RF sputtering (annealed at 600 °C to form the spinel phase)	self-doped	100–500	3.4
a-Zn_2SnO_4	RF sputtering (T_s 375–430 °C) [5]	self-doped	30–60	
a-$ZnSnO_3$	RF sputtering (T_s RT – 300 °C)	self-doped	40–100	
$Zn_2In_2O_5$	DC or RF sputtering	self-doped	2.9	2.9
a-IZO	DC sputtering	self-doped	3.0–5.0	3.1

*Achieved for bulk material, not thin film.

($1.3 \times 10^{-4}\ \Omega$ cm) by room-temperature sputtering in O_2 of an oxide target followed by annealing at 580–700 °C to form a single phase spinel structure [3]. The resistivity of zinc stannate (Zn_2SnO_4), regardless of whether it is polycrystalline or amorphous (a-Zn_2SnO_4), is much higher because of a low carrier concentration [4, 5]. Polycrystalline zinc indate ($Zn_2In_2O_5$) has been reported to have a refractive index in the range 2.1–2.4, compared with the usual index of 1.8–2.0 for SnO_2, In_2O_3, and ZnO in the visible range [6].

Most high-quality TCO films (e.g. SnO_2:F, ZnO:Al, and high-temperature ITO) are polycrystalline, but interestingly, several amorphous TCO films have been prepared, some of which are shown near the foot of Table 17.1. Thus, ITO deposited at low temperatures, low temperature Zn_2SnO_4, and indium zinc oxide (IZO with Zn content of 10–42%) are all amorphous. Remarkably, a-IZO can have a resistivity as low as $3.0 \times 10^{-4}\ \Omega$ cm [2] and a mobility greater than 50 cm^2/V s [2, 7]. Although it might be expected that mobility correlates with a high degree of crystal perfection, this obviously cannot be the complete story, and perhaps correlation with lack of defects might be a promising way to view this result. The multicomponent TCOs are suitable for applications requiring reduced In content or better defined wet etch behavior for improved lithography. As will be seen, some of the multicomponent TCOs have important emerging applications in the PV field.

All the TCOs that have been mentioned so far are n-type. While Cu_2O is p-type, its bandgap (2.2 eV) is too small for it to be considered a TCO. Recently, a group of p-type TCOs having the delafossite structure was discovered [8]. These p-type TCOs, namely, $CuAlO_2$, $CuGaO_2$, $CuInO_2$, and $CuCrO_2$ exhibit positive Hall and Seebeck coefficients. Their optical transmission tends to be lower than that of the traditional TCOs. Typical properties of $CuAlO_2$ are: bandgap 3.5 eV, resistivity about 1 Ω cm, mobility 10 cm^2/V s. Another p-type TCO is Cu_2SrO_2, and many other p-type and transparent or partially transparent materials have since been discovered. There is also much effort to prepare p-type ZnO. While nitrogen on an oxygen site is an acceptor, it is accompanied by strong compensation, and p-type behavior so far does not seem to be unambiguously proven [9]. The discovery of p-type TCOs makes it possible to fabricate transparent transistors and facilitates fabrication of light-emitting diodes. However, no efficient PV device that makes use of a p-type TCO has yet been constructed.

17.2.2 Doping

In this section we will discuss doping using ZnO as an example. ZnO crystallizes with the wurzite crystal structure. It consists of a hexagonal close packed lattice with a diatomic base. The atoms are tetrahedrally bonded with a Zn atom bonded to four oxygen atoms and vice versa. The intrinsic defect states in ZnO include the donors $Zn_i^{\bullet\bullet}$, $Zn_i^{\bullet}$, Zn_i^{x} (Zn interstitials), $V_O^{\bullet\bullet}$, $V_O^{\bullet}$, V_O^{x} (oxygen vacancies) and the acceptors V_{Zn}'', V_{Zn}' (Zn vacancies). The Kröger–Vink notation used to denote these defects is described in Section 17.4.1.1.

It is well known that the electrical carrier concentration in ionic crystals can be manipulated by oxidation/reduction reactions that modify stoichiometry. Thus, heating in a partial pressure of oxygen reduces the conductivity and heating in hydrogen increases it. In the case of ZnO, a reaction balance occurs between the oxygen content of the crystal and interstitial zinc atoms or oxygen vacancies. In the former case, excess Zn ions and electrons migrate to join with oxygen to form ZnO:

$$2\ Zn_i^{\bullet\bullet} + 4e + O_2 \leftrightarrow 2\ ZnO \tag{17.2}$$

Nominally undoped ZnO can be prepared with low resistivity, but such films are unstable at relatively low temperatures. Extrinsic doping of TCOs (as in ZnO:Al) leads to much better high-temperature stability. In ZnO, extrinsic doping is accomplished by introducing group IIIa (group 13 in IUPAC Periodic Table) metal atoms substitutionally on Zn sites in the ZnO lattice.

These atoms can be B, Al, Ga, or In. Since their ionization energy (in ZnO) is low (53 meV for Al, 54.5 meV for Ga), the additional electron is donated at room temperature to the conduction band. Alternatively, a halogen atom (e.g. F, Cl) can be substituted for O at some of the anion sites. It has also been established that H acts as a shallow donor in ZnO with an ionization energy of 35 meV [10]. Hydrogen is always present in vacuum systems, and it is now suspected that H contributes to the near universal *n*-type behavior of ZnO. Finally, it has been reported that the oxide-forming metals Ti, Zr, Hf, and Y dope ZnO, suggesting oxygen vacancies as the donor. Most heavily doped TCOs are doped to degeneracy so that the Fermi level lies above the conduction band edge. Such TCOs exhibit metallic behavior, i.e. a positive temperature coefficient of resistivity.

17.2.3 Properties of TCOs Used in PV Applications

The following is an overview of the preparation and properties of the TCOs most commonly used in PV applications.

ZnO. Transparent conductive ZnO is generally prepared by DC or RF magnetron sputtering or by LPCVD. For high-quality ZnO doped with Al or Ga, the resistivity can be as low as $3 \times 10^{-4}\,\Omega\,\text{cm}$ and the average optical transmittance is over 85%. Although Al has been the traditional dopant, its high reactivity with oxygen suggests that Ga might offer advantages. ZnO:B films deposited from LPCVD can be grown with a textured surface morphology suitable for light trapping in thin film solar cells. The substrate temperature range for ZnO preparation is from room temperature to about 200 °C, which is lower than the temperature for SnO_2 and ITO deposition. ZnO is readily chemically attacked by acids or bases, and is reported to be hygroscopic. Commonly used etchants are diluted acid or ammonium chloride. ZnO is commonly reported to have a superior ability to resist a hydrogen plasma and is the preferred TCO for nc-Si solar cell fabrication.

SnO_2. Tin oxide coated glass for PV applications is commercially available and is made by the APCVD process. The substrate temperature is about 650 °C, and the films can be grown with a textured surface morphology. The doping element is fluorine (SnO_2:F). Film resistivities are in the $5–8 \times 10^{-4}\,\Omega$ cm range. The optical absorption coefficient in the visible for commercial SnO_2:F is often quite high ($\sim 400–600\,\text{cm}^{-1}$). Tin oxide has a high work function of 4.9 eV. It is much more chemically and thermally stable than ZnO, and is hard to etch. One method of etching SnO_2 is to use HCl plus zinc powder. Tin oxide films are readily reduced by a hydrogen plasma and their optical absorption is increased. The high process temperatures that are apparently needed for effectively doped SnO_2 films prevent their use on temperature-sensitve PV devices.

In_2O_3. As a TCO, In_2O_3 is not widely used in solar cell manufacturing, except in HIT cells and some a-Si based cells. One reason is that indium is a relatively rare and expensive element. However, compared with other TCOs, a higher carrier mobility can be obtained in doped In_2O_3. A high carrier mobility makes it possible for the TCO to achieve low optical absorption in the NIR region and still maintain high conductivity. Sputtering a pre-doped ceramic target (In_2O_3 with ~10% wt SnO_2) is a popular method to deposit In_2O_3. The substrate temperature is kept at 150–300 °C. The resistivity of the film is about $2 \times 10^{-4}\,\Omega$ cm. ITO can be prepared with the very smooth surface required for FPD manufacturing. Because of increasing demand, the price of indium has risen from about $100/kg in 2002 to about $600/kg today (2010).

TiO_2. A resistivity of less than $3 \times 10^{-4}\,\Omega\,\text{cm}$ has been achieved for Nb-doped TiO_2 deposited on a single-crystal substrate by the PLD process [11]. On a glass substrate, $9.5 \times 10^{-4}\,\Omega\,\text{cm}$ resistivity was achieved by annealing sputter-deposited TiO_2:Nb in a pure H_2 environment [12]. The refractive index of TiO_2 is around 2.5 and is therefore notably larger than that of the above TCOs. Recent progress in TiO_2 preparation is reviewed in greater detail in Section 17.9.4. TiO_2 is more resistant to a hydrogen plasma than is SnO_2.

Table 17.2 Record high carrier mobility (or in some cases low resistivity) of certain TCOs prepared by different methods

TCO	Abbreviation	μ (cm^2/V s)	ρ ($10^{-4}\,\Omega$ cm)	Deposition method	Reference
ZnO:Al	AZO	47.6	0.85	PLD	[15]
ZnO:Ga	GZO	30.9	0.81	PLD	[16]
ZnO:Al		44.2	3.8	RF sputtering	[17]
ZnO:Al		49.5	2.9	Hollow cathode sputtering	[18]
ZnO:Al		52	7.1	RF sputtering H_2/Ar treatment	[19]
ZnO:Al			2.0	RF sputtering (sub. perpendic.)	[14]
ZnO		120	4.6	RF sputtering	[13]
In_2O_3:Ti	ITiO	105	1.9	RF sputtering	[20]
In_2O_3:Ti		159	0.9	PLD	[21]
In_2O_3:Mo	IMO	130	1.7	Reactive evaporation	[22]
In_2O_3:Mo		70.2	1.6	Hollow cathode sputtering	[23]
In_2O_3:Sn	ITO	103	0.44	e-beam with zone-confinement	[24]
In_2O_3:Sn		42	0.77	PLD	[25]
In_2O_3: H	IO:H	140	2.9	RF sputtering w. post-anneal	[26]
SnO_2:F	FTO	70	7.0	APCVD w. H plasma treatment	[199]
TiO_2:Nb	TNO	22	2.4	PLD (epitaxial)	[11]
Cd_2SnO_4	CTO	68		RF sputtering	[28]
a-Zn_2SnO_4	ZTO	32	57	RF sputtering	[5]

Some notable record results for either high mobility or low resistivity for some of these TCOs prepared by different methods are given in Table 17.2. Since the literature now abounds with abbreviations for TCOs, we take the opportunity to list them as they occur in this Table. Most groups obtain a best resistivity of about $4–5 \times 10^{-4}\,\Omega$ cm for ZnO:Al deposited on unheated substrates by RF sputtering from a ceramic target (ZnO with 2% wt Al_2O_3). A typical set of parameters might be $n_e = 6.05 \times 10^{20}/cm^3$, $\mu = 25.7\,cm^2/V\,s$, $\rho = 4.0 \times 10^{-4}\,\Omega$ cm. Using an external solenoid [13] and substrates placed perpendicular to the target, a resistivity for ZnO:Al of $2 \times 10^{-4}\,\Omega$ cm was achieved [14]. High mobilities in the range $70–160\,cm^2/V$ s are obtainable for doped In_2O_3, and notably for In_2O_3:Mo and In_2O_3:Ti, compared with a mobility of $30\,cm^2/V$ s for commercially produced ITO ($n_e = 1.3 \times 10^{21}/cm^3$). The very low resistivity for ITO of $4.4 \times 10^{-5}\,\Omega$ cm reported in [24] appears to be an isolated result. Earlier efforts had achieved $1.2 \times 10^{-4}\,\Omega$ cm using plasma-assisted e-beam evaporation at a substrate temperature of 280 °C [27].

For reference purposes, the common properties of the principal TCOs SnO_2, In_2O_3, and ZnO are listed in Table 17.3. All three possess direct optical bandgaps (see Section 17.4.2 for additional information regarding In_2O_3). The progression in oxide formation energy ΔH_f from 3.6 to 6.0 eV per metal atom for ZnO, In_2O_3, and SnO_2 accounts for the increasing chemical stability of these TCOs. The data for this table was assembled from several sources, including references [29, 30].

The work function φ of a material is the energy required to remove an electron from the Fermi level at the surface to the vacuum level. The work function can be changed by doping, by a surface dipole, or by band bending at the surface. Although typical values of the work functions of the principal TCOs are given in Table 17.3, their variability should be borne in mind. The band alignment at interfaces is important in devices and depends not only on nominal work function, but also on growth order and the intial growth stages.

Again, for reference purposes, we mention both early and recent reviews of TCOs [31–34], sources for semiconductor physics and data [35, 36], and another recent book on ZnO [37].

Table 17.3 Properties of SnO_2, In_2O_3, and ZnO

Parameter	Unit	SnO_2	In_2O_3	ZnO
Mineral		Cassiterite		Zincite
Lattice		Tetragonal	Cubic	Hexagonal
Structure		Rutile	Bixbyite	Wurtzite
Space group		$P4_2/nmm$	$Ia3$	$P6_3mc$
a, c	nm	0.474, 0.319	1.0117	0.325, 0.5207
Density	$g\,cm^{-3}$	6.99	7.12	5.67
Band gap (E_g)	eV	3.5-3.6 (dir.)	3.6-3.75 (dir.) 2.75 (indir.)	3.3–3.4 (dir.)
Thermal cond. (κ)	$W\,m^{-1}\,K^{-1}$	$98_{\|}55\perp$		$69_{\|}60\perp$
Exp. coeff. (α)	$10^{-6}\,K^{-1}$	$3.7_{\|}4.0\perp$	6.7	$2.92_{\|}4.75\perp$
Hardness	Moh scale	6.5	5	4
$\varepsilon(0)$		$9.58_{\|}13.5\perp$	8.9	$8.75_{\|}7.8\perp$
$\varepsilon(\infty)$		$4.17_{\|}3.78\perp$	4.6	$3.75_{\|}3.70\perp$
Refractive index (n)			1.89 (633 nm)	$2.029_{\|}2.008\perp$
T_m	°C	1620	1910	1975
T_m (metal)	°C	232	157	420
T (vap. press. 10^{-3} Pa)	°C	882	670	208
ΔH_f/metal atom	eV	6.0	4.8	3.6
Other properties				piezoelectric
Exciton binding energy	meV			60
m_e^*/m_e		$0.23_{\|}0.3\perp$	0.35	0.28
Effective CB DOS (N_c)	cm^{-3}	3.7×10^{18}	4.1×10^{18}	3.7×10^{18}
μ_e (lattice)	$cm^2\,V^{-1}\,s^{-1}$	255	210	200
m_h^*/m_e				0.59
μ_h (lattice)	$cm^2\,V^{-1}\,s^{-1}$			5–50
Work function (φ)	eV	4.9	4.7	4.5

The symbols || and ⊥ denote directions or polarization directions parallel and perpendicular to the c-axis

17.3 DEPOSITION METHODS

Transparent conductive oxides can be prepared by various deposition techniques, and new techniques continue to be invented. The most suitable technique depends on the intended application. The driving forces in TCO development have been to address product needs and processing limitations, to improve TCO performance, to increase throughput and reduce manufacturing cost, and to accomplish robust process control. The optical, electrical, and surface properties of the TCO are related to the deposition method and deposition parameters. The dominant technologies for large-scale production are sputtering and CVD. However, PLD, thermal evaporation, sol–gel, spray pyrolysis and other processes are also widely studied. A useful review of many of these techniques is provided in the book by Gläser [38].

17.3.1 Sputtering

Sputter coating is performed in a vacuum system containing a cathode assembly on which is mounted the target material to be sputtered, a working gas (Ar) at low pressure, and the substrate to be coated. The sputtering process involves establishment of a glow discharge in the gas to create Ar^+

Table 17.4 Overview of sputtering configurations

Type of cathode	Excitation and target type		
	RF	DC	MF
Planar magnetron	C	M, C	M, C
Cylindrical magnetron		M, C	M, C
Hollow cathode		M	M
Dual planar		M, C	M, C
Dual cylindrical		M, C	M, C

Target: M denotes metallic, C denotes ceramic

ions, application of a negative bias to the target to attract the ions, with consequent ion bombardment of the target and ejection of atoms from the target. The sputter yield (sputtered atoms per ion) depends on the energy and angle of incidence of the ion, the mass of the ion and target atoms, and the binding energy of the surface atoms. The sputtered atoms land on the substrate with an average of about 5 eV of kinetic energy, although this figure can be influenced by Ar pressure or target-substrate distance [39]. In magnetron sputtering, a tunnel of magnetic field lines is formed over the target using permanent magnets in order to confine secondary electrons. This increases the plasma density, lowers the operating pressure and increases the sputtering rate. The target is generally a rectangular slab of material that is nonuniformly eroded in a closed loop called the racetrack. Improved magnetron design has increased target utilization from 25% to about 40%. The power source used for sputtering can be DC or pulsed DC, MF (i.e. mid-frequency AC), RF, or RF superimposed on DC.

Magnetron sputtering is eminently suitable and cost-effective for large-scale TCO production. This process is used to manufacture ITO for the flat-panel display industry and ZnO:Al for a-Si and CIGS modules. It is generally conducted using multiple-chamber in-line coaters with either horizontal or vertical substrate transport. For horizontal transport, the sputtering is usually downwards onto substrates carried on rollers or wheels. A static mode is sometimes used for FPD manufacturing. An overview of sputtering configurations is shown in Table 17.4, their classification being made according to the type of cathode, sputtering power supply, and target material.

Oxide films such as In_2O_3, ZnO, or TiO_2 are often deposited directly from ceramic sputtering targets which are made by pressing and sintering oxide powders pre-mixed with doping elements or oxides in a certain weight ratio. Different targets such as In_2O_3 doped with Sn, Mo, Ti, Si or Ge [20, 40–42], ZnO doped with Al or Ga [17, 43] and TiO_2 doped with Nb [44] have been studied. The sputtering gas is generally Ar with a small quantity of oxygen added. High-resistivity targets can only be sputtered using RF power at relatively low deposition rates. Conductive targets can be sputtered using DC or MF power. Ceramic targets are much more expensive than metallic ones, and the low utilization of planar targets can yield an unacceptably high manufacturing cost for thick ($\sim$1 μm) TCOs for PV applications. Another limitation is that the film doping from a given ceramic target cannot be varied because of the fixed composition of the target. Despite the target cost, the use of ceramic targets is less complicated and often has a higher yield than reactive sputtering (described below).

Cylindrical magnetron cathodes are now available. They use a rotating cylindrical target (essentially a metal tube) so that the entire surface of the target can be eroded uniformly. The advantages are increased target utilization, increased target volume, longer running time, and reduced problems of re-deposition and debris that can be encountered with planar targets. A ceramic tube target for ZnO:Al deposition can be made in segments. It consists of ZnO and $ZnAl_2O_4$ and can be made with >95% of theoretical density and with homogeneous microstructure.

Oxide films can also be made by reactive sputtering. The primary advantage of this process is the ability to use a low-cost metallic target. Oxygen is supplied uniformly across the substrate width in order to form the oxide film by reaction with the deposited metal. However, the method is complicated by target oxidation (target poisoning). Reactive sputtering of a metal target can be conducted in three target modes: metallic, oxidized, or transition (partially oxidized). These conditions are often characterized by a hysteresis curve resulting from a plot of cathode voltage versus reactive gas flow [45]. For Zn, the cathode voltage drops from about 650 to 400 V in switching from metallic to oxidized mode. Metallic mode operation usually leads to desorption of excess Zn from the growing film that ultimately contaminates the deposition system.

The best TCO films (neither absorbing nor over oxidized) are obtained in the unstable transition mode that requires active feedback to maintain the favorable intermediate target state. The process is controlled in transition mode by controlling the oxygen partial pressure. Using a signal derived from cathode voltage, oxygen partial pressure or optical emission, the partial pressure set point can be stabilized using a PID controller via control of discharge power or the oxygen gas flow. For example, the optical emission from the metal atoms, say the Zn line at 307 or 481 nm, can be ratioed to that obtained in the metallic mode or to the oxygen line at 777.4 nm, and this voltage is used as the control signal. Another complication of reactive sputtering is the disappearing anode caused by coating of internal surfaces with an insulating oxide. The use of mid-frequency pulsed power in which the cathode voltage is periodically reversed to discharge dielectric surfaces is helpful in reducing arcing. Another solution is to utilize twin magnetrons supplied with out-of-phase AC power so that each target is alternately the cathode and the anode and each is alternately cleaned during sputtering. This process effectively improves the target surface metallic property and allows a high deposition rate. The optimum substrate temperature for reactive sputtering is generally higher than for ceramic target sputtering. Detailed results in these areas are given in a number of papers [46–49]. Figure 17.2 shows a diagram and photo of the dual magnetron arrangement.

With PV module production capacity per factory steadily rising, the demand is for in-line magnetron sputtering machines of higher throughput. Some insights into throughput considerations can be gleaned by examining the evolution of sputtering systems designed for architectural glass coating. The trend has been towards larger compartment size, higher pumping speed, improved access for maintenance and improved software for glass transport control [50].

One aspect of planar magnetron sputtering is that the properties of the TCO film deposited on a static substrate vary as a function of distance perpendicular to the long axis of the cathode. Usually, the film resistivity increases in positions directly facing the target racetrack. The origin of this effect seems to be due to oxygen ions (O^-). The ions are accelerated in the cathode fall, undergo charge transfer during transit to the substrate, and finally bombard the growing film [51] potentially causing lattice defects. The effect is generally stronger for reactive sputtering from a metal target than for sputtering from a ceramic target [45]. In a dynamic mode (with substrate translation) these different properties are layered into the resulting film, causing a degraded average resistivity. The increased resistivity has been attributed principally to degradation of carrier mobility. The effect can be ameliorated by superimposing RF power during DC magnetron sputtering of AZO ceramic targets, which has the effect of reducing the DC sputtering voltage [52]. The addition of about 0.5% H_2 to the sputtering gas was also found to improve the spatial distribution of resistivity. The deleterious effect of bombardment by oxygen ions can also be ameliorated through the use of stronger magnets in the magnetron to reduce the cathode voltage and hence the ion energy.

We briefly mention some other developments in sputter technology. One is called high target utilization sputtering (HiTUS) which uses a remote RF plasma together with launch and steering magnets to replace the magnetron [53]. In this way, racetrack formation is avoided and the target can in principle be uniformly eroded. The method has been used to deposit ITO on glass with a resistivity of $1.6 \times 10^{-4}\ \Omega\,\text{cm}$.

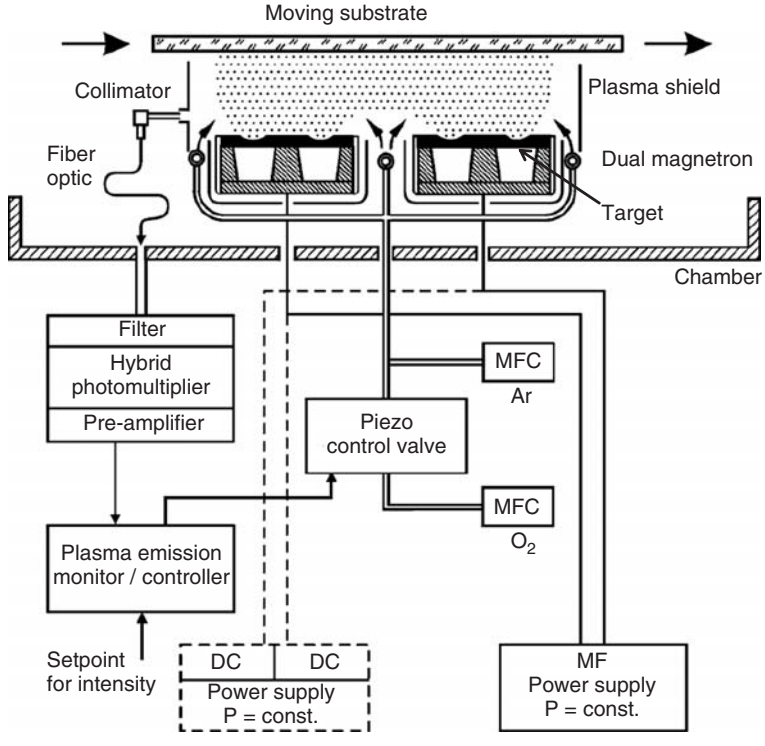


Figure 17.2 Dual magnetron reactive sputtering of ITO in in-line coater. Top: schematic of the deposition scheme with plasma-emission-based process control (adapted from [49]; bottom: arrangement of targets and cathode in opened system (reproduced from May C, Strumpfel J, Schulze D, *43rd SVC Tech. Conf*., 137 (2000) [49] by permission of the Society of Vacuum Coaters)

An alternative method of generating a high density plasma suitable for sputtering other than by magnetic confinement is to utilize a hollow cathode discharge [54]. In a hollow cathode, an intense plasma is confined geometrically by static electric fields. Electrons oscillate back and forth across the cathode cavity, being reflected after each pass by the electric field of the cathode fall and leading to efficient ionization of the argon working gas. In a practical arrangement, a channel is defined between two planar targets separated by 1–3 cm. The targets are sputtered in a large Ar flow that is constrained to pass through the channel. The sputtered atoms are carried by the Ar out of the cathode cavity and to the substrate. This method is called hollow cathode sputtering or gas flow sputtering. A reactive gas, e.g. O_2, is introduced outside of the channel [55], or more specifically, at the exit of the channel [56, 57] and is decomposed in the plasma afterglow. The targets are always sputtered in metallic mode, allowing an intrinsically stable and reproducible reactive process. In contrast to magnetron sputtering, where the operating pressure is 1–10 mTorr, leading to high-energy particles that can damage the lattice of the growing film, the high working pressure of hollow cathode sputtering (hundreds of mTorr) means that the sputtered atoms and other particles are thermalized before reaching the substrate thereby resulting in a soft deposition. Significant work has been conducted by two groups in developing linear hollow cathode sputtering sources suitable for large-area coating using a reactive process [55, 57].

17.3.1.1 Sputtering processes and results

In all magnetron sputtering processes the sputtering pressure is an important parameter. High-quality films are usually obtained in the pressure range 2–4 mTorr. Higher pressures result in increased thermalization of the sputtered atoms and a lower deposition rate while low pressures can allow re-sputtering of the film surface via negative ion bombardment. The flux and type of energetic particles can affect film properties such as stress, crystal size, roughness, texture, and point defects.

High-quality ZnO:Al films have long been prepared by RF magnetron sputtering from a ceramic ZAO target. On unheated, stationary substrates, resistivities of 4.0×10^{-4} and $6.25 \times 10^{-4}\,\Omega\,\text{cm}$ can be obtained with 2 and 1% (by weight) Al_2O_3 in the target, respectively. A resistivity as low as $1.4 \times 10^{-4}\,\Omega\,\text{cm}$ was achieved at $T_s = 230\,^{\circ}\text{C}$ [58]. With the goal of increasing the deposition rate, the pulsed DC sputtering of conductive ceramic targets was investigated. With the addition of 0.3% O_2 to improve film transmission, a resistivity of $2.2 \times 10^{-3}\,\Omega\,\text{cm}$ was obtained without substrate heating [45]. With substrate heating, a resistivity of $7 \times 10^{-4}\,\Omega\,\text{cm}$ was achieved at 200 °C, which remains considerably higher than that achievable with RF or reactive sputtering [59].

To avoid the need for expensive ceramic targets, reactive sputtering of metallic Zn:Al alloy targets has been pursued. Using a Zn:Al 1.5 wt % target and MF sputtering, with transition mode control using the ratio of the optical signals Zn 481 nm/O 777.4 nm, a resistivity of $2.9 \times 10^{-4}\,\Omega\,\text{cm}$ was achieved at 200 °C [46]. A very high dynamic deposition rate of 220 nm m/min was achieved using a dual DC magnetron with Zn:Al 2 wt % targets. The substrate temperature was limited to 170 °C in order to avoid junction damage to CIGS cells. Achieving good lateral sheet resistance uniformity with reactive sputtering was reportedly more difficult than with ceramic target sputtering. Reactive sputtering in the oxide and transition modes has also been conducted using a twin cathode driven at 40 kHz ac [48]. Several advantages relative to DC reactive sputtering were noted, including stable conditions during the entire target lifetime and reduced arcing. A resistivity of $4 \times 10^{-4}\,\Omega\,\text{cm}$ was achieved at 300 °C. Typical RMS roughness was 2 nm and grain size 30 nm. The reactive mid-frequency magnetron sputtering of dual Zn:Al targets in transition mode to prepare ZnO:Al has been successfully scaled up to coat 1.0×0.6 m substrates [60]. A resistivity of $3.4 \times 10^{-4}\,\Omega\,\text{cm}$ was obtained at T_s of 150–200 °C at a dynamic deposition rate of up to 80 nm m/min with negligible Zn desorption from the film. Processes for ZnO:Al intended for subsequent texturing are discussed in Section 17.6.

The relationship of film property to film structure has occupied a great deal of research effort. In the context of TCOs we may inquire about the effect of grain size and preferred orientation on electrical resistivity, and in particular, on electron mobility. We note that in ZnO the intra-grain mobility is anisotropic because of piezoelectric scattering for the *c*-axis direction. In one study, using reactively-sputtered ZnO:Al, it was found that, with increasing film thickness, the degree of *c*-axis orientation (texture index) increased and the in-plane resistivity generally decreased [61]. Thus, resistivity was found to correlate with texture index, and may also have correlated with grain size although this variable was not studied. It was later found that the growth rate (based on atomic areal density as determined by RBS) increased with deposition time and saturated at about 300 s [62]. The origin and evolution of texture is yet another area where a full understanding remains elusive. Self-textured films, often grown on amorphous substrates, can result from preferential nucleation or by evolutionary selection of faster-growing grains. It was demonstrated by high resolution TEM that reactively-sputtered ZnO films grown in metallic mode grew with (0001) planes parallel to the substrate surface forming a *c*-axis textured columnar structure starting at the substrate-film interface [63]. On the other hand, films grown in transition mode were reported to exhibit random orientation of initial nuclei. A review of texture development in ZnO concluded that underlying mechanisms could not yet be conclusively identified [64].

There have been occasional efforts to produce doped SnO_2 films by sputtering. For example, SnO_2:F was prepared by DC reactive sputtering in Ar–O_2–freon using a tin target [65]. However, the achievement of electrically active donors appears to be difficult.

A wide variety of oxides and nitrides have been very successfully prepared by hollow cathode sputtering, including ZnO, In_2O_3, SnO_2, Al_2O_3, $CuAlO_2$, AlN, InN:O, TiO_2:Nb and TiN [23, 56, 66]. In the case of ZnO:Al, the absence of any shift in the (002) X-ray peak was interpreted to mean that the films were largely stress free [67]. High-performance TCOs such as In_2O_3:Mo and In_2O_3:Ti having high mobilities of 80 $cm^2/V\,s$ have been produced [23, 68, 69]. The elements Zr and Nb were also demonstrated to be efficient and useful dopants in In_2O_3. The EPV SOLAR group has termed their process reactive-environment hollow cathode sputtering (RE-HCS). The properties of some TCO films produced by RE-HCS are shown in Table 17.5.

17.3.2 Chemical Vapor Deposition

In the CVD process, the precursors in vapor form are decomposed into atoms or molecules and chemically react with each other at the hot substrate surface to form a nonvolatile product coating. The precursor decomposition can be triggered by high temperature, plasma or light. The operating pressure of CVD can be in the range of a few torr to above atmospheric pressure. Based on the activation method, type of precursor, and working pressure, the processes are categorized as atmospheric pressure CVD (APCVD), low pressure CVD (LPCVD), plasma-enhanced CVD (PECVD), photo-CVD, and metalorganic CVD (MOCVD).

APCVD does not need vacuum, gives a high deposition rate and offers great cost benefits. This process has been adopted by the flat glass industry and others for F-doped SnO_2 film production [71, 72]. Thus, float glass, as it exits the tin bath and is still hot, can be directly coated with SnO_2:F during glass cooling. Such a plant can coat a 10–15 mile length of 130 inch-wide glass in a day. This on-line coating can be smooth or textured. Many different precursors can be used for the film deposition. For example, $SnCl_4$ (tin tetrachloride or TTC) [73], $(CH_3)_2SnCl_2$ (dimethyltin dichloride or DMTC) [74] and *n*-$C_4H_9SnCl_3$ (monobutyltin dichloride or MBTC) [75], mixed with O_2 or H_2O and a doping material such as HF or an organofluorine, are often used in industry. The most common process is

$$SnCl_4 + 2(H_2O) \rightarrow SnO_2 + 4(HCl) \tag{17.3}$$

Table 17.5 Properties of TCO films prepared by reactive-environment, hollow cathode sputtering (RE-HCS) on soda-lime glass

Material	T_s (°C)	t (μm)	R_{sh} (Ω/sq)	ρ (10^{-4} Ω cm)	N (10^{20}/cm^3)	μ (cm^2/V s)	References
ZnO:Al	250	0.36	21.1	7.59	2.32	35.5	[57]
ZnO:Al	340	1.03	2.8	2.86	4.42	49.5	[70]
ZnO:B	160	0.63	9.0	5.7			[57]
ZnO:B	110	0.44	17.8	7.8	2.1	37.8	[23]
ZnO:Ga	175	0.85	12.0	10.2	2.67	23.0	unpublished
In_2O_3	260	0.89	20.0	17.0	0.95	38.5	[68, 23]
In_2O_3:Mo	290	1.00	1.9	1.9	4.1	80.3	[56, 23]
In_2O_3:Zr	250	0.56	4.0	2.3	4.3	63.3	[68, 23]
In_2O_3:Nb	260	0.86	3.6	3.1			[23]
In_2O_3:Ti	300	0.54	3.3	1.8	4.3	80.6	[68, 23]
In_2O_3:Sn	280	0.34	5.8	2.0			[57]

The process temperature is in the range 590–650 °C. At this temperature, the diffusion of ions from the soda-lime glass into the coating will degrade the electrical properties of the SnO_2. A sodium barrier layer such as a thin SiO_2 or SiO_xC_y film is therefore deposited on the glass before the SnO_2 is deposited. In some cases two underlayers are used, resulting in the structure glass/SnO_2(~25 nm)/SiO_2(~20 nm)/SnO_2:F. Such underlayers, due to differences in refractive index, reduce the amplitude of interference fringes and therefore serve as color suppression layers to yield a more neutral appearance [71]. Since the thermal expansion coefficient of soda lime glass is greater than that of tin oxide (~8.6×10^{-6}/°C below 300 °C, versus 4.0×10^{-6}/°C) it would be expected that commercial tin oxide is under compressive stress, although based on XRD measurement of the lattice constant just the opposite (tensile stress) has been reported [76].

Tin oxide films deposited by APCVD can be prepared with rough surfaces which is a critical feature for a-Si/μc-Si solar cell efficiency improvement. By controlling the deposition conditions, the morphology of the films can be manipulated. For example, Asahi Glass Co. in Japan has produced so-called U-type and A-type SnO_2:F films having different surface texture [77]. Recent R&D work using TTC as the precursor has demonstrated SnO_2:F films having a surprisingly small visible absorption coefficient minimum of about 100 cm^{-1} compared with about 400–600 cm^{-1} typical of most commercial tin oxide [73]. The overcoating of SnO_2:F with TiO_2 by APCVD has been demonstrated using $Ti(OC_3H_7)_4$ at 500 °C [78]. Other TCO films deposited by APCVD include ITO [79], ZnO:F [80, 81], and ZnO:Al [82]. With fluorine contents of about 0.5 at.%, minimum ZnO:F film resistivities of 6×10^{-4} Ω cm using diethylzinc, $Zn(C_2H_5)_2$, and 5×10^{-4} Ω cm using a chelated diethylzinc precursor were obtained at growth temperatures of 400 °C and 480–500 °C, respectively. For ZnO:Al grown by APCVD a resistivity of 3×10^{-4} Ω cm was achieved. Plasma-enhanced APCVD of ZnO:Al using He, DEZ, CO_2 and trimethylaluminum (TMA) has also been reported [83]. A schematic diagram of an APCVD deposition system suitable for continuous production is shown in Figure 17.3. In general, such a machine consists of a belt transport system, multiple temperature zones, heated injector heads, and gas delivery systems including bubblers for liquid precursors.

Low-pressure CVD has become an important process to deposit boron-doped ZnO films for a-Si and a-Si/μc-Si solar cell application [84–87]. One feature of this approach is that ZnO:B films can be produced having a large-grain morphology eminently suitable for light trapping when used as a superstrate TCO. Another feature is that the process only needs a low substrate temperature in the range 140–160 °C. This makes it possible to additionally deposit, by LPCVD, textured

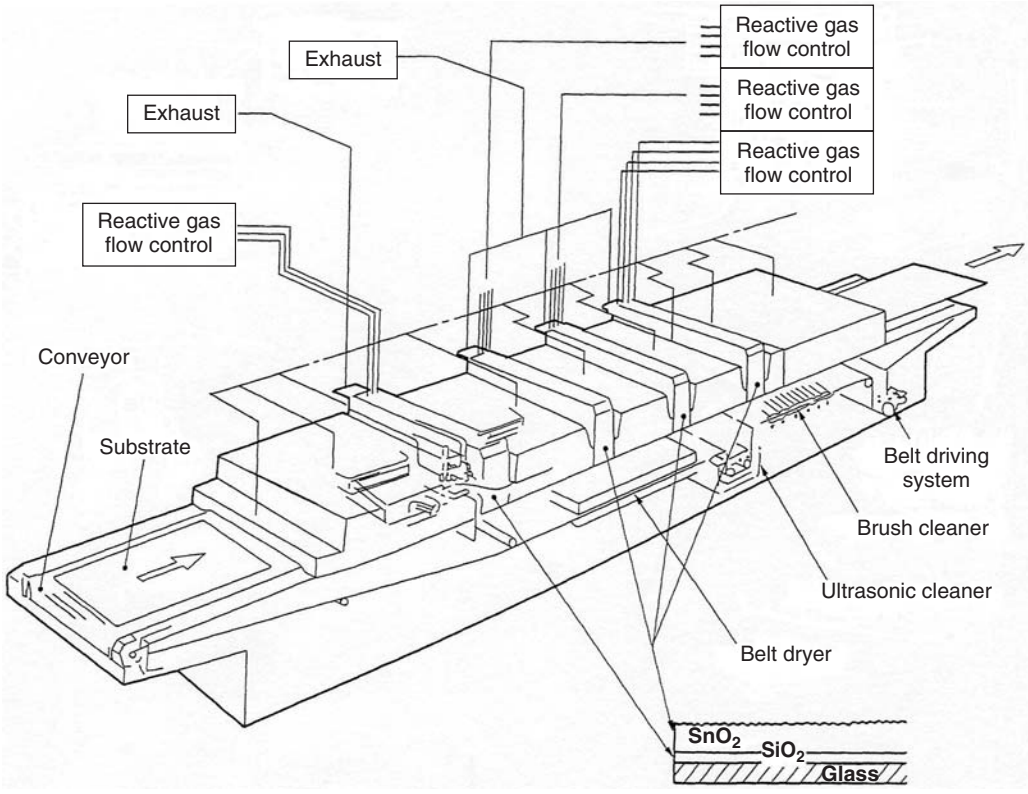


Figure 17.3 Schematic diagram of atmospheric pressure chemical vapor deposition (APCVD) system for coating SiO_2/SnO_2 on glass. Adapted from Sato K *et al*., *Reports Res. Lab. Asahi Glass Co., Ltd*, 42 2 (1992), with permission from AGC [77]

ZnO as a rear contact after the junction is formed. Undiluted diethylzinc (DEZ), and water (H_2O) vapors are used as precursors. The H_2O/DEZ ratio is kept slightly greater than unity. Diborane (B_2H_6) diluted in He is used as the doping gas. The pressure is from as low as 0.37 Torr to several Torr. In this regime, the reactions are kinetically limited, rather than being limited by diffusion of reactants to the substrate, and the deposition rate increases with increasing temperature, rising from 20A/s at 145 °C to 25A/s at 155 °C. For pressures > 10 Torr, premature reactions may occur in the gas phase. A strong morphological transition occurs for $T_s > 145$ °C, with large tetragonal grains emerging and XRD revealing a strong $11\bar{2}0$ orientation [84]. A small-grained nucleation layer remains at the substrate surface. By controlling the B_2H_6/DEZ ratio, pressure and substrate temperature, the properties of the ZnO:B (morphological, optical and electrical) can be optimized for solar cell application. These aspects are further discussed in Section 17.6. A minimum resistivity of about $1.0 \times 10^{-3}\ \Omega$ cm can be achieved, although higher values are more suitable for a TCO for a-Si/μc-Si solar modules.

The reaction intermediates for the deposition of ZnO by MOCVD from DEZ and H_2O have been studied via their capture on a cold substrate and subsequent analysis by simultaneous thermogravimetric and differential thermal analysis (TGA-DTA) [88].

As an example of the growth of more complex oxides, we mention that $Zn_2In_4Sn_3O_{14}$ films have been prepared by MOCVD with quite respectable properties ($\mu = 33.2\,cm^2/V\,s$ and

$n_e = 4.3 \times 10^{20}\,\mathrm{cm}^{-3}$) [89]. This film can be interpreted as consisting of the bixbyite phase of In_2O_3, with Zn^{2+} and Sn^{4+} co-substituted for two In^{3+}.

17.3.3 Pulsed Laser Deposition

Compared with sputtering and CVD, the PLD set-up is much simpler. A pulsed, high-power laser beam is focused on a target in a vacuum chamber and causes evaporation of the target material. The pulse energy is typically 300–600 mJ/pulse while the instantaneous power density is about $10^9\,\mathrm{W/cm}^2$. A luminous and directional plasma plume is created which also absorbs laser radiation. A heated substrate faces the target, and ablated material condenses on it to grow a film. The process results in approximately stoichiometric transfer of atoms from a multi-element target to the substrate. The development of PLD has benefited particularly from the drive to produce high-T_c superconducting materials such as $YBa_2Cu_3O_7$. Laser sources such as the excimer gas lasers KrF (248 nm) and ArF (193 nm) are used for PLD as well as solid state Nd:YAG lasers (256, 355 nm). The laser plume contains mostly single atoms and ions, as well as electrons, but also atom clusters, particulates and molten droplets. To reduce particle damage and droplet generation, the laser energy must be limited. For oxide film growth, an oxygen environment is needed to maintain the same stoichiometry in the film as in the target. A supersaturation appears during PLD pulse duration and creates a high nucleation density. Many high-quality TCO films have been made by the PLD technique [15, 16, 21, 25]. PLD also can used to produce ternary TCOs [90] and *p*-type TCOs [91]. It appears that PLD films often grow with very smooth surfaces and crystallize at low substrate temperatures because of the high kinetic energy of the plasma species. The latter feature could be useful for TCO growth on plastic substrates. On the other hand, it has been reported that ZnO films optimized for high mobility (at low carrier concentration, $\sim 3 \times 10^{16}\,\mathrm{cm}^{-3}$) show increased roughness [29]. Despite the simplicity of the PLD technique the mass transfer rate is low ($< 10^{-4}\,\mathrm{g\,s}^{-1}$) and scale-up appears difficult.

17.3.4 Other Deposition Techniques

Reactive thermal evaporation is frequently used to prepare ITO films in the laboratory. Thus, InSn (10 wt %) can be evaporated from a BN crucible in a W heater with a partial pressure of oxygen in the mid 10^{-4} Torr range and T_s of 175 °C to achieve a film resistivity of $2.4 \times 10^{-4}\,\Omega\,\mathrm{cm}$. RF plasma activation can be added, if desired. Alternatively, pellets of $In_2O_3 + 9$ mol% SnO_2 can be evaporated using an *e*-beam source, again in a partial pressure of oxygen, and preferably near $T_s = 300\,^\circ\mathrm{C}$ to obtain a low film absorptance [92].

A couple of other vacuum-based methods deserve mention. The first of these is atomic layer deposition (ALD). In this technique, a precursor and oxidizer are alternately dosed into the growth chamber and purged from it. Very thin precursor layers, in some cases just a monolayer, are thereby adsorbed and then oxidized. This four-step cycle is then repeated many times. As an example, ZnO:B is deposited by ALD by introducing DEZ ($+B_2H_6$) as precursor and H_2O as oxidizer into a chamber at about 0.1 Torr [93]. The ALD method is inherently slow, but yields highly conformal films. The method has also been used to grow Zr-doped In_2O_3 [94]. The second method of note involves the reaction of metal atoms with atomic oxygen. For example, zinc oxide has been prepared by the simultaneous supply of zinc and atomic oxygen to heated substrates in a vacuum chamber. The atomic oxygen was generated remotely by a confined, capacitively coupled RF discharge in oxygen arranged either as a point or line source [95]. A peak deposition rate of 40 A/s was achieved, with a resistivity of $2 \times 10^{-3}\,\Omega\,\mathrm{cm}$ for ZnO:In.

We now discuss some non-vacuum routes to TCO preparation. The spray pyrolysis technique is a low-cost method for deposition of various TCOs. The precursors are dissolved in a solvent

and the solution (including a metal dopant) is sprayed by high-pressure spray nozzles or ultrasonic nebulizer towards a heated substrate. The reactants pyrolyse on the substrate surface and the film is formed. The substrate temperature affects the crystallite orientation, morphology and other film properties. Unfortunately, many non-obvious parameters exist and process control is difficult. All of the binary TCOs (SnO_2, In_2O_3, ZnO, and CdO) can be achieved through this technique [96–98], as well as ternary oxides (e.g. $CuAlO_2$). Of note, a mobility of 43.8 $cm^2/V\,s$ and resistivity of $4.1 \times 10^{-4}\ \Omega$ cm were achieved for SnO_2:F films produced by ultrasonic spray pyrolysis in a belt furnace at 500–530 °C using DMTC with NH_4F and HF for doping [99].

Another simple way to obtain TCOs is by sol–gel methods. Metal alkoxides or metal–organic compounds that are dissolved in alcohol constitute the precursor solution. The thin films are coated on a substrate by spinning, dipping or drawing. The films may be subsequently annealed or sintered. *C*-axis-oriented ZnO films have been produced by such a method [100]. However, films prepared by the sol–gel route are often porous and of lesser quality than those achieved by vacuum or PLD processes. Finally, inkjet printing of TCOs has attracted considerable attention for possible low-cost printed electronics and solar cells [101, 102].

17.4 TCO THEORY AND MODELING: ELECTRICAL AND OPTICAL PROPERTIES AND THEIR IMPACT ON MODULE PERFORMANCE

In this theory and modeling section we discuss some important aspects of TCOs in greater detail. These aspects include TCO electrical properties (doping, carrier transport and scattering), optical properties (fundamental and free-carrier absorption), and monolithic module design taking into account TCO sheet resistance and absorption. In all cases we present useful equations that can be used both for both conceptual and quantitative understanding and for modeling purposes.

17.4.1 TCO Electrical Properties

17.4.1.1 Generation of free carriers

Transparent conducting oxide films are wide-bandgap semiconductors. At room temperature, a truly intrinsic film is very resistive. In order to achieve a conductive film, suitable impurities have to be introduced (extrinsic doping) or the films must deviate from stoichiometry. In Section 17.2 we discussed doping in ZnO. In this section we will discuss doping mechanisms in Sn-doped crystalline In_2O_3 (ITO). Indium oxide crystallizes in the cubic bixbyite structure. This is related to the fluorite structure which consists of a face-centered cubic array of cations with anions occupying all of the tetrahedral interstitial positions so that the cations sit in the center of a cube of anions. In the bixbyite structure, two oxygen atoms are missing from each cube, either along a face or a body diagonal, resulting in two types of In site [103]. There appear to be three mechanisms that can generate free carriers [89]. Using Kröger-Vink (K–V) notation, these models can be written as:

$$2In_{In}^{\,x} + 2SnO_2 \rightarrow 2Sn_{In}^{\bullet} + In_2O_3 + \tfrac{1}{2}O_2(g) + 2e' \tag{17.4}$$

$$2In_{In}^{\,x} + 2SnO_2 \rightarrow (2Sn_{In}^{\bullet}O_i'')^{x} + In_2O_3 \tag{17.5a}$$

$$(2Sn_{In}^{\bullet}O_i'')^{x} \rightarrow 2Sn_{In}^{\bullet} + \tfrac{1}{2}O_2(g) + 2e' \tag{17.5b}$$

$$O_O^{\,x} \rightarrow \tfrac{1}{2}O_2(g) + V_O^{\bullet\bullet} + 2e' \tag{17.6}$$

In the K–V notation [104], ionic species or defects are denoted by a capital letter, their site is denoted using a subscript, and their effective charge is denoted using a superscript (• for positive and

′ for negative). Thus, Equation (17.4) describes direct doping by Sn atoms in which a Sn atom substitutes for an In atom (In^{3+}) in the sublattice and the ionized Sn atom (Sn^{4+}) contributes an electron. In Equation (17.5a), a self-compensating reaction leads to the formation of neutral Sn:O_i associates involving O interstitials; in Equation (17.5b), reduction activates the donor species by removal of the O interstitials. This implies that Sn in In_2O_3 may or may not be activated. Under highly reducing conditions, O vacancies are formed and serve as electron donors, as described by Equation (17.6). Each vacancy provides two electrons. The doping efficiency of ITO (electrons per Sn atom) at low Sn concentrations ($C_{Sn} \sim 2$ at.%) has been variously reported to range from a little over unity to as low as 0.25. It is generally observed that the carrier density saturates around $C_{Sn} = 10$ at.%. In this range, the Sn^{4+}, which is normally sixfold coordinated with the O^{2-} anions, is neutralized by excess interstitial O ions [105]. Furthermore, the inactive Sn atoms produce neutral scattering centers that reduce the electron mobility. No matter what the growth process for ITO, we may say that the carrier concentration depends on the Sn concentration, the oxygen partial pressure, and the temperature.

In introducing cadmium stannate (Cd_2SnO_4) in Section 17.2 it was described as being self-doped, and to the best of the authors' knowledge no extrinsic dopants for this material are known. The origin of carriers has been variously attributed to oxygen vacancies, Cd interstitials, or, on the basis of first-principles calculations, Sn on Cd antisites [106].

17.4.1.2 Carrier transport and scattering

Electrons that have been freed as described above render the film electrically conductive. It is a standard result in semiconductor theory that such electrons respond to an external force according to classical mechanics, provided their real mass m_e is replaced by their effective mass m_e^* to account for the presence of the lattice. Thus, in the presence of an electric field E, the electrons feel a force $eE = m_e^* \, dv/dt$. As a result of collisions with imperfections in the crystal the electrons acquire an average velocity component in the direction of the force. This is the drift velocity v_d which is proportional to the applied field E and the average time τ between collisions, and inversely proportional to the electron effective mass

$$v_d = \frac{e\tau}{m_e^*} E = \mu_e E \tag{17.7}$$

The electron mobility μ_e is therefore $e\tau/m_e^*$ and is proportional to the time between collisions ($\sim 5 \times 10^{-15}$ to 1×10^{-14} s in a TCO). The current density J is given by $J = n_e e v_d = \sigma E$ where n_e is the free carrier (electron) concentration in the film and σ is the electrical conductivity of the film material. Thus,

$$\sigma = \frac{1}{\rho} = n_e e \mu_e = n_e e^2 \frac{\tau}{m_e^*} \tag{17.8}$$

This model is similar to the Drude theory of a classical gas of electrons in metals. The drift velocity is in general much smaller than the electron thermal velocity $(3kT/m_e^*)^{1/2}$. More correctly, if the electrons are regarded as particles obeying the Pauli exclusion principle (and hence Fermi–Dirac statistics), then the electron velocity V_F at the Fermi surface is $V_F = (3\pi^2)^{1/3}(\hbar/m_e^*)n_e^{1/3}$ [107] which is of the order of 5×10^5 m/s. The electron mean free path l is then $l = V_F \tau = V_F m_e^* \mu_e / e$; for $\mu_e = 50\,\mathrm{cm^2/V\,s}$ and $n_e = 5 \times 10^{20}\,\mathrm{cm^{-3}}$, l is approximately 8 nm. This is smaller than the usual grain size of TCOs.

From Equation (17.8), in order to achieve high conductivity, n_e must be raised by doping. However, heavy doping can result in several drawbacks: ionized dopant atoms are scattering centers that reduce the electron mobility; if the impurity concentration exceeds the solubility limit, phase separation can occur; a high electron concentration results in light absorption at long wavelengths

(we will show this shortly). To the extent that this is possible, a preferred strategy to increase conductivity is to increase the electron mobility. Consequently, we will now briefly identify the major mechanisms that control electron mobility; these are more fully reviewed elsewhere [108–110].

(a) *Ionized impurity scattering*. In a TCO film with high free carrier concentration, there exist a large number of ionized impurities which can be oxygen vacancies, dopants, or excess metal atoms. They serve as strong scattering centers to electrons. Dingle derived an expression for the electron mobility μ_i in degenerate semiconductors considering scattering at a screened Coulomb potential (Brooks–Herring–Dingle model) [111]. The screening of the electric field around the ionized impurity results from the free carrier gas around it. The result is

$$\mu_i = \frac{3(\varepsilon_0\varepsilon_r)^2 h^3}{Z^2 {m_e}^{*2} e^3} \frac{n_e}{n_i} \frac{1}{F_i(\xi)} \tag{17.9}$$

here $F_i(\xi) = \ln(1+\xi) - \dfrac{\xi}{1+\xi}$ and $\xi = (3\pi^2)^{1/3} \dfrac{\varepsilon_0 \varepsilon_r h^2 n_e^{1/3}}{m_e^* e^2}$ where ε_0 is the permittivity of free space, ε_r is relative static permittivity, h is Planck's constant; Z and n_i are the charge and concentration of the impurities, respectively. For an uncompensated fully ionized semiconductor, $n_e = n_i$. We note that μ_i depends on $1/Z^2$ and is temperature-independent.

(b) *Lattice vibration scattering*. Increasing temperature means stronger lattice vibrations, deformation of the lattice, and limitation of the electron mobility by acoustic phonon scattering. J. Bardeen and W. Shockley [112] calculated that the mobility μ_l arising from lattice scattering in nonpolar material is given by

$$\mu_l = \frac{2\sqrt{2\pi} e \hbar^4 C_l}{3 {m_e}^{*5/2} E_d^2 (kT)^{3/2}} \tag{17.10}$$

where C_l is the longitudinal elastic constant, E_d is the deformation potential constant in eV, $\hbar = h/2\pi$, k is Boltzmann's constant, and T is the absolute temperature of the semiconductor.

(c) *Grain boundary scattering*. The grain boundaries of polycrystalline semiconductor films contain interface states that can trap charges. These charges form potential barriers that hinder the passage of electrons between neighboring grains, especially if the grain size is comparable to the mean free path of the electrons. For electron traps, a depletion region forms on either side of the grain boundary, with current transport described by thermionic emission over the barrier [113]. The mobility μ_g resulting from grain boundary scattering has been described by the Petritz model [114], extended by Seto [115] and Baccarani [116]. The basic result is

$$\mu_g = \left(\frac{L^2 e^2}{2\pi m_e^* kT}\right)^{1/2} \exp\left(-\frac{\phi_b}{kT}\right) \tag{17.11}$$

where L is grain size and ϕ_b is the grain boundary potential (barrier height). The barrier height is $\varphi_b = e^2 N w^2 / 2\varepsilon\varepsilon_0$ (where w is the depletion width and N is the donor density). Since the interface states are below E_F they are filled with electrons; the interface charge is equal to the charge in the two depletion regions and so $eN_T = 2eNw$ where N_T is the surface trap density (typically falling in the range $0.3–3.0 \times 10^{13}$ cm^{-2}). We find $\varphi_b = e^2 N_T^2 / 8\varepsilon\varepsilon_0 N$. In a typical TCO, the free carrier concentration is high and the mean free path is only a few nanometers, i.e. less than the grain size, so the effect of grain boundary scattering would appear to be relatively small.

(d) *Neutral impurity scattering*. For scattering by neutral atoms the mobility μ_n is given by

$$\mu_n = \frac{m_e^* e^3}{20\varepsilon_r \varepsilon_0 \hbar^3 n_N} \tag{17.12}$$

where n_N is the density of neutral centers [117, 118].

Other scattering mechanisms such as electron–electron scattering only have a minor effect on the free carrier mobility in TCO films. The principal scattering processes discussed above may be combined (by adding collision probabilities) to determine the resulting free carrier mobility μ_e:

$$\frac{1}{\mu_e} = \frac{1}{\mu_i} + \frac{1}{\mu_l} + \frac{1}{\mu_g} + \frac{1}{\mu_n} \tag{17.13}$$

For most TCOs that are used in solar cells as a window electrode, n_e is in the range $10^{19}\,\mathrm{cm}^{-3}$ to $10^{21}\,\mathrm{cm}^{-3}$ and μ_e is less than $100\,\mathrm{cm}^2\,\mathrm{V}^{-1}\,\mathrm{s}^{-1}$ (generally about $30\,\mathrm{cm}^2\,\mathrm{V}^{-1}\,\mathrm{s}^{-1}$). For relatively high carrier concentrations ($>2 \times 10^{20}\,\mathrm{cm}^{-3}$) it is generally agreed that mobility is dominated by ionized impurity scattering. In the range 10^{19}–$10^{20}\,\mathrm{cm}^{-3}$ Minami concluded that, in ZnO, grain-boundary scattering is dominant [1] (although some others concluded that neutral impurity scattering dominates [119, 120]). Figure 17.4 shows the results of Minami's calculations, together with experimental data on impurity-doped and undoped (although still relatively conductive) ZnO films prepared by different methods. For comparison, three doped ZnO films prepared by hollow cathode sputtering are included. We might remark that the maximum room-temperature mobility in single crystals of ZnO is 180–$200\,\mathrm{cm}^2\,\mathrm{V}^{-1}\,\mathrm{s}^{-1}$.

It has been found experimentally that the electron optical effective mass for ITO increases with carrier concentration [121]. This suggests that, in this degenerate semiconductor, the $E(k)$ dispersion relation is non-parabolic [122]. A similar conclusion has been reached for ZnO:Al [123, 124]. For ITO, an empirical equation for effective mass m_e^* as a function of electron concentration n_e is given [109] as:

$$m_e{}^* = (0.066 n_e^{1/3} + 0.3) m_e \tag{17.14}$$

where m_e is the electron mass.

The scattering mechanism can also be investigated by measurement of the Seebeck coefficient S (the symbol α is also used), with units $\mu\mathrm{V\ K}^{-1}$ (see Section 17.7.1). In a degenerate

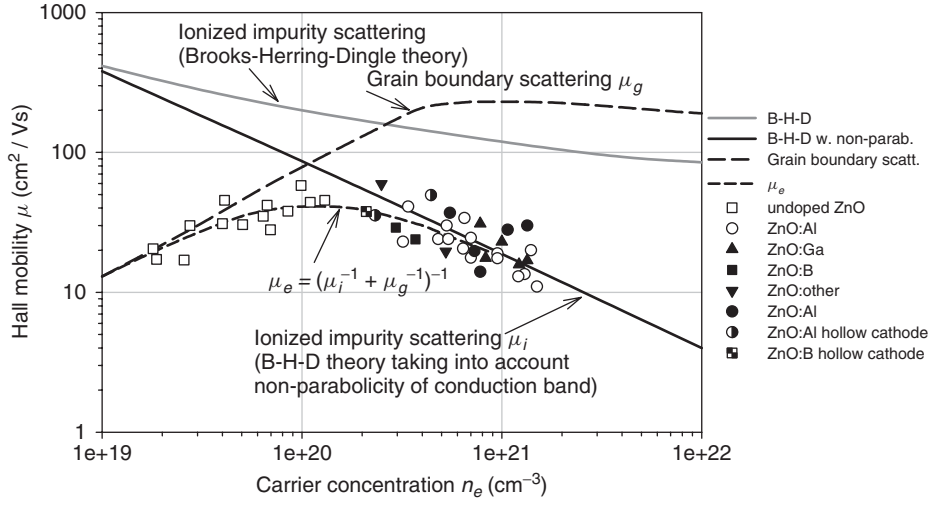


Figure 17.4 Measured Hall mobility versus carrier concentration of undoped and impurity-doped ZnO films. Two scattering mechanisms, ionized impurity scattering and grain boundary scattering, were used in the modeling. Majority of data from Minami T, *MRS Bulletin* **25**, 38 (2000) [1]

semiconductor, the dependence of S on temperature T and carrier concentration n is given by

$$S = -\frac{k}{e}\left(\frac{\pi^2}{3}\right)\left(r + \frac{3}{2}\right)\left(\frac{kT}{E_F}\right) \tag{17.15}$$

where $k/e = 86.2\,\mu\text{V/K}$, $E_F = (\hbar^2/2m^*)(3\pi^2 n)^{2/3}$ for a degenerate electron gas, and the scattering exponent r depends on the dominant scattering process [118, 119]. The magnitude of S varies as $n^{-2/3}$ and hence increases with decreasing carrier concentration.

17.4.2 TCO Optical Properties

Various optical absorption processes are possible in a semiconductor. Some examples are the fundamental absorption (valence band to conduction band transition), impurity absorption, free carrier absorption, exciton absorption, higher-energy transitions, intraband transitions, and donor–acceptor transitions [125]. The absorption coefficient α is defined by the fractional decrease in light intensity I in the direction of propagation according to

$$\alpha = -\left(\frac{1}{I}\frac{\mathrm{d}I}{\mathrm{d}x}\right).$$

In general, $\alpha(h\nu) = A\Sigma P_{if} n_i n_f$ where P_{if} is the probability for the initial state to final state transition, n_i the density of electrons in the initial state, and n_f the density of unoccupied final states, and the summation is over all transitions separated by the photon energy $h\nu$. Here we discuss the fundamental and free carrier absorption processes. The equations given below enable the optical transmission of a TCO to be calculated as a function of wavelength.

For a direct electron transition across the energy gap from the valence band to the conduction band, representing the strong, fundamental absorption of the semiconductor, the absorption coefficient for $E > E_g$ is:

$$\alpha(E) = \frac{A}{E}(E - E_g)^{1/2} \tag{17.16}$$

where E_g is the bandgap, E is the photon energy and A is a fitting constant [118, 126]. The conduction band is assumed to be parabolic and to be unoccupied. Equation (17.16) predicts that a plot of $(\alpha E)^2$ versus E for $E > E_g$ yields a straight line whose intercept on the E axis is the energy gap E_g. The term "direct" means that electron momentum is conserved. In an indirect transition, an electron in the valence band can access any empty state in the conduction band by conserving momentum through a phonon interaction. In indium oxide, for example, an indirect gap has been reported at about 2.6 eV (the direct gap being about 3.7 eV) [127]. Indeed, recent work has confirmed that direct optical transitions from the valence band maximum (VBM) to the conduction band minimum (CBM) are parity forbidden and that the onset of strong optical absorption occurs from valence bands 0.8 eV below the VBM [128].

In doped films, the conduction and valence band edges are perturbed by Coulomb interaction with ionized impurities, resulting in tails of states extending into the energy gap. The optical absorption coefficient for transitions between the band tails follows Urbach's rule

$$\alpha_U(E) = A_U \exp\left[-\frac{\xi}{kT}(E_g - E)\right] \tag{17.17}$$

where A_U is the fitting constant and ξ is a parameter describing the steepness of the sub-gap absorption edge obtainable by plotting $\ln(\alpha)$ versus E for $E < E_g$ (see, for example, [129]).

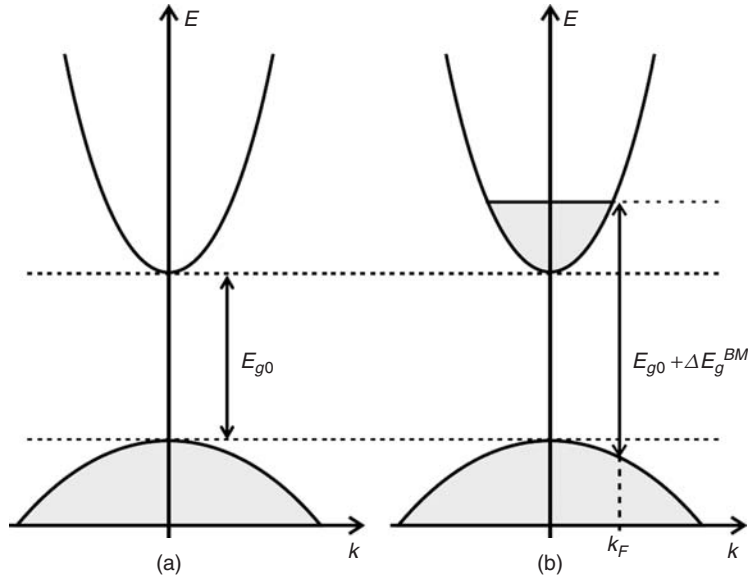


Figure 17.5 Schematic band structure of a TCO with parabolic bands separated by the fundamental gap E_{g0} and with vertical optical transitions. Sufficiently high doping fills the lowest states in the conduction band so that the optical gap is widened by the Burstein–Moss shift ΔE_g^{BM}. Redrawn after Sernelius B *et al.*, *Phys. Rev*. **B 37**, 10244 (1988) [130]

In heavily doped semiconductors, for example, TCOs having a high carrier concentration, the Fermi level lies in the conduction band (i.e. the semiconductor is degenerate) and the lowest states in the conduction band are filled. (These states start to fill when $n_e > N_c$ where N_c is the effective density of states in the conduction band; for ZnO, $N_c = 3.7 \times 10^{18}$ cm^{-3}.) Photons then need a larger energy than the fundamental band gap E_{g0} to excite a transition from the valence band to unoccupied states in the conduction band. This is shown in the simplified band structure diagram of Figure 17.5.

The increment of energy ΔE_g^{BM} is called the Burstein–Moss (B-M) shift [131].

$$\Delta E_g^{BM} = \frac{\hbar^2 k_F^2}{2m_{vc}^*} = \left(\frac{h^2}{2m_{vc}^*}\right)\left(\frac{3n_e}{8\pi}\right)^{2/3} \tag{17.18}$$

In this equation, k_F is the Fermi wavevector, $m_{vc}{}^*$ is the reduced effective mass $(m_{vc}^*)^{-1} = (m_v{}^*)^{-1} + (m_c{}^*)^{-1}$ and m_v^* and m_c^* are the valence and conduction band density-of-states effective masses. The optical bandgap E_g now becomes:

$$E_g = E_{g0} + \Delta E_g^{BM} \tag{17.19}$$

It may be noted that Equation (17.18) predicts relatively large values for ΔE_g^{BM} at large electron concentrations, e.g. $\Delta E_g^{BM} = 0.44$ eV at $n_e = 2 \times 10^{20}$ cm^{-3}, assuming a reduced effective mass of $0.26\, m_e$. However, once n_e exceeds a critical value of about 5×10^{19} cm^{-3}, a bandgap shrinkage mechanism sets in and partially offsets the increase due to Equation (17.18) [132]. We might also comment that the transparency of a TCO in the visible portion of the spectrum requires that low-energy transitions from the Fermi level into unoccupied states higher in the conduction band have a low probability. This is partly assured because TCO cations have filled *d*-shells:

$3d^{10}$ for Ga and Zn, $4d^{10}$ for Cd, In and Sn. Beyond this, the absorption of a photon requires the absorption or emission of a phonon to accommodate the change in electron momentum. To calculate the transition probability quantum mechanically, three particles (photon, electron, and phonon) are involved together with an intermediate state [118]. In what follows we will give a classical treatment of free-carrier absorption. The bandgap defined by Equation (17.19) determines the demarcation energy between strong TCO absorption and transparency at the high-frequency end of the spectrum.

At lower frequencies, the optical transmittance of the TCO film is determined by absorption and reflection of the free carrier plasma. To model these effects we again use the Drude model (see Section 17.4.1.2 above) which treats the electrons as a classical gas. The interaction between the light and the free carriers occurs via an AC electric field $E(t) = E_1 \exp(-i\omega t)$ so that the equation of motion for an electron becomes $m_e^*(\mathrm{d}^2x/\mathrm{d}t^2) + (m_e^*/\tau)\mathrm{d}x/\mathrm{d}t = eE(t)$ [133]; as introduced earlier, τ is the average time between collisions. This equation has the solution $x = (e/m_e^*)E_1 \exp(-i\omega t)/(-\omega^2 - i\omega/\tau)$. Knowing x, we may calculate $v = \mathrm{d}x/\mathrm{d}t$ and so find $\boldsymbol{\sigma}$ from $J = \boldsymbol{\sigma}E = nevE$ where $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\omega)$ is the complex, frequency-dependent conductivity of the TCO. Hence,

$$\sigma(\omega) = \frac{n_e e^2 \tau}{m_e^*} \frac{1}{(1 - i\omega\tau)} = \varepsilon_0 \varepsilon_\infty \omega_p^2 \tau \frac{1 + i\omega\tau}{1 + \omega^2\tau^2} \tag{17.20}$$

where ω_p is called the plasma frequency and is given by:

$$\omega_p = \left(\frac{n_e e^2}{\varepsilon_0 \varepsilon_\infty m_e^*} \right)^{\frac{1}{2}} \tag{17.21}$$

The plasma wavelength $\lambda_p = 2\pi c/\omega_p$ and as before, $n_e e^2 \tau / m_e^*$ is the DC conductivity σ. Note that the imaginary part of the conductivity peaks at a value $\sigma/2$ at a frequency given by $\omega\tau = 1$. At sufficiently high frequencies the material becomes a perfect dielectric with permittivity ε_∞.

From Maxwell's equations it can be deduced that plane waves can exist, traveling say in the x direction, having transverse field components E_y and H_z of the form $\exp\{i\omega(t - \boldsymbol{n}x/c)\}$, where $\boldsymbol{n}$ is the complex refractive index, and that $\boldsymbol{n}^2 = \mu(\varepsilon_\infty - i\boldsymbol{\sigma}/\omega\varepsilon_0) = \mu\varepsilon$ [134]. The expression $\varepsilon_\infty - i\boldsymbol{\sigma}/\omega\varepsilon_0$ is identified as the complex AC permittivity $\boldsymbol{\varepsilon}$:

$$\varepsilon(\omega) = \varepsilon_\infty - i\frac{\sigma(\omega)}{\varepsilon_0 \omega} = \varepsilon_r(\omega) + i\varepsilon_i(\omega) \tag{17.22}$$

The real and imaginary parts ε_r and ε_i of the permittivity, are, respectively:

$$\varepsilon_r = \varepsilon_\infty \left(1 - \frac{\omega_p^2 \tau^2}{1 + \omega^2 \tau^2} \right) \tag{17.23}$$

$$\varepsilon_i = \frac{\varepsilon_\infty \omega_p^2 \tau}{\omega(1 + \omega^2\tau^2)} \tag{17.24}$$

Thus, in the Drude model, the optical behavior is determined by the frequencies ω_p and $1/\tau$. For a TCO, $\omega_p \gg 1/\tau$. There are three regimes of behavior: (i) $\omega < 1/\tau$; (ii) $1/\tau < \omega < \omega_p$; (iii) $\omega > \omega_p$. In the theory of metals these are called the Hagen–Rubens regime, the relaxation regime, and the transparent regime, respectively. When $\omega = \omega_p$, the electrons resonate with the light field and absorb energy strongly. At this frequency, $\varepsilon_r = 0$. For $\omega < \omega_p$, ε_r is negative and the plasma becomes reflective. Solar cells normally operate in the regime where $\omega > \omega_p$. Assuming

$\mu = 1$, the complex refractive index is $\boldsymbol{n} = n + ik = \sqrt{(\varepsilon_r + i\varepsilon_i)}$, and so the refractive index and extinction coefficient are

$$n = \frac{\sqrt{2}}{2}\sqrt{(\varepsilon_r^2 + \varepsilon_i^2)^{1/2} + \varepsilon_r} \tag{17.25}$$

$$k = \frac{\sqrt{2}}{2}\sqrt{(\varepsilon_r^2 + \varepsilon_i^2)^{1/2} - \varepsilon_r} \tag{17.26}$$

Having derived the optical constants in terms of material parameters, we may proceed with optical modeling for a plane parallel film on a substrate. The theory of thin film optics is widely treated in the literature [135, 136] and is based upon the solutions of Maxwell's equations subject to the tangential components of the electric and magnetic vectors being continuous across the interface between the two media [134]. We will treat the case of normal incidence for which all polarizations of the incident light are equivalent. For a single thin film layer (e.g. a TCO) with optical indices n and k and thickness d on an infinitely thick, transparent substrate with refractive index n_s, we may calculate the reflectance (at the air–film side) and the transmittance into the substrate, at a given wavelength λ, using the following equations [135]:

$$R(\lambda) = \frac{(g_1^2 + h_1^2)e^{2a_1} + (g_2^2 + h_2^2)e^{-2a_1} + A\cos 2\gamma_1 + B\sin 2\gamma_1}{e^{2a_1} + (g_1^2 + h_1^2)(g_2^2 + h_2^2)e^{-2a_1} + C\cos 2\gamma_1 + D\sin 2\gamma_1} \tag{17.27}$$

$$T(\lambda) = \frac{n_s[(1+g_1)^2 + h_1^2][(1+g_2)^2 + h_2^2]}{e^{2a_1} + (g_1^2 + h_1^2)(g_2^2 + h_2^2)e^{-2a_1} + C\cos 2\gamma_1 + D\sin 2\gamma_1} \tag{17.28}$$

where

$$a_1 = \frac{2\pi kd}{\lambda}, \gamma_1 = \frac{2\pi nd}{\lambda} \text{ (in radians)}$$
$$A = 2(g_1g_2 + h_1h_2),\ B = 2(g_1h_2 - g_2h_1),\ C = 2(g_1g_2 - h_1h_2),\ D = 2(g_1h_2 + g_2h_1)$$

and

$$g_1 = \frac{1 - n^2 - k^2}{(1+n)^2 + k^2}, h_1 = \frac{2k}{(1+n)^2 + k^2}, g_2 = \frac{n^2 - n_s^2 + k^2}{(n+n_s)^2 + k^2}, h_2 = \frac{-2n_sk}{(n+n_s)^2 + k^2}$$

When the glass substrate is of finite thickness, the reflection at the rear surface perturbs the values of R and T given above. Substrates are usually sufficiently thick that the reflections are incoherent and we may simply add the intensities (not amplitudes) of waves successively reflected at the front and rear surfaces [136]. As a first approximation, the reflectance R' (again from the film side) is $R' = R + T^2R_s$, and the transmittance T' (air–film–glass–air) is $T' = T(1 - R_s)$ where $R_s = (n_s - 1)^2/(n_s + 1)^2$ is the reflectance ($\sim$0.04) of the substrate–air interface.

Using the equations presented above we have modeled the optical behavior of a 500 nm thin film TCO layer on a glass substrate having a refractive index $n_s = 1.52$. Material parameters for the TCO were chosen to be similar to those for In_2O_3:Ti, namely $m_e^* = 0.35m_e$, $\varepsilon_\infty = 4.48$ [68]. Three different films were modeled each having a sheet resistance of 6.24 Ω/□, but achieved with different combinations of carrier concentration and mobility while keeping their product constant (recall $\sigma = n_e e\mu$). Thus Film 1 had $n_e = 1 \times 10^{21}\,\text{cm}^{-3}$ and $\mu = 20\,\text{cm}^2\,\text{V}^{-1}\text{s}^{-1}$, Film 2 had $n_e = 5 \times 10^{20}\,\text{cm}^{-3}$ and $\mu = 40\,\text{cm}^2\,\text{V}^{-1}\,\text{s}^{-1}$, and Film 3 had $n_e = 2 \times 10^{20}\,\text{cm}^{-3}$ and $\mu = 100\,\text{cm}^2\,\text{V}^{-1}\,\text{s}^{-1}$. Figure 17.6 shows the calculated T and A as a function of wavelength. Film 1, with a high n_e and low μ, exhibits a strong absorption peak near its plasma wavelength λ_p of 1320 nm, and this absorption cuts into and reduces the film transmission in the visible spectrum. Although not shown in this figure, the reflectivity of this film rises strongly in the infrared for wavelengths $\lambda > \lambda_p$. Film 1 also exhibits a Burstein–Moss shift of its absorption edge to shorter wavelengths as a result of the large carrier concentration. By increasing the mobility to $100\,\text{cm}^2\,\text{V}^{-1}\text{s}^{-1}$ while

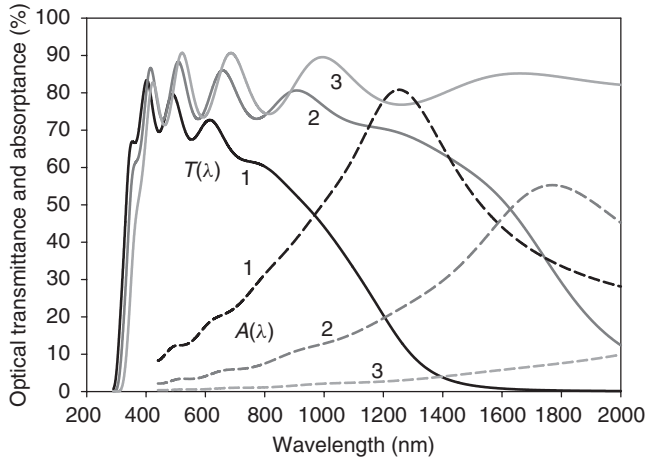


Figure 17.6 Calculated optical transmittance $T(\lambda)$ and absorptance $A(\lambda)$ of three TCO films (labeled 1, 2, and 3) having the same sheet resistance (6.24 Ω/□), but different carrier concentrations and mobilities as follows:

$$\text{Film 1: } n_e = 1 \times 10^{21}/\text{cm}^3;\ \mu = 20\,\text{cm}^2/\text{V s}$$
$$\text{Film 2: } n_e = 5 \times 10^{20}/\text{cm}^3;\ \mu = 40\,\text{cm}^2/\text{V s}$$
$$\text{Film 3: } n_e = 2 \times 10^{20}/\text{cm}^3;\ \mu = 100\,\text{cm}^2/\text{V s}$$

The film thickness is 500 nm and the substrate is glass. The other parameters used in the modeling were $m_e^* = 0.35m_e$ and $\varepsilon_\infty = 4.48$. The plasma wavelengths for these three films are $\lambda_p = 1320$ nm, 1870 nm, and 2960 nm, respectively.

reducing carrier concentration to $2 \times 10^{20}\,\text{cm}^{-3}$ the plasma absorption is reduced and pushed out to longer wavelengths until its effect on transmission for wavelengths utilized by solar cells becomes negligible. The shift in λ_p with n_e is evident from Equation (17.21). This example shows that to achieve a TCO of given sheet resistance it is desirable to increase the mobility and decrease the carrier concentration in order to reduce the optical absorption [137]. As a counter-example, the high carrier density in thin film titanium nitride (TiN) results in a resistivity as low as 30 μΩ cm, but also shifts the plasma wavelength to about 0.7 μm so that the film both reflects and absorbs red light.

We now discuss a figure of merit for TCO materials. It has been suggested that the quantity σ/α is an appropriate parameter. Certainly this parameter increases with increasing conductivity σ or with decreasing optical absorption coefficient α. Since the absorbance A of a film of material of thickness t is $A = 1 - \exp(-\alpha t)$ we have $\alpha = -(1/t)\ln(T + R)$ since $T + R + A = 1$. We also have $\sigma = 1/(R_{sh}t)$ and so $\sigma/\alpha = -[R_{sh}\ln(T + R)]^{-1}$. This expression has dimensions Ω^{-1} and does not depend on the thickness of the material. (Note that we have implicitly assumed that σ is a material constant independent of film thickness; in practice σ usually increases with thickness via an increase in grain size.) Values of the parameter σ/α pertaining to the visible region of the spectrum have been reported for various materials. Some of these are: SnO_2:Sb 0.4; ZnO:B 2.0, SnO_2:F 3.1; ZnO:Al 5.1; Cd_2SnO_4 7.0 [138, 139]. Recent optimization of some of these materials resulting in higher mobility suggests that the relevant figures of merit should be updated. The figure of merit is meant as a suggestive guide only. We do not recommend that it dictate use of a particular material during product design as many other important factors must be considered.

We may also derive an expression for σ/α based on the theory given in Section 17.4.1.2 and earlier in this section. From the expression for the field components of a plane wave

$\exp\{i\omega(t - \boldsymbol{n}x/c)\}$ given earlier it can be shown that the power flux decreases as the wave propagates in a medium with refractive index $\boldsymbol{n} = n + ik$ according to $\exp(-2\omega kx/c)$. The absorption coefficient α is therefore $\alpha = 2\omega k/c = 4\pi k/\lambda$ where λ is the wavelength in free space. We now need an expression for k. From Equation (17.26) we find that, in the visible region where $\varepsilon_i \ll \varepsilon_r$, $k = \varepsilon_i/(2\varepsilon_r^{0.5})$. Using Equations (17.23) and (17.24), and recalling that $\sigma = n_e e^2\tau/m_e{}^*$ (Equation 17.8) we find the following expression for σ/α:

$$\frac{\sigma}{\alpha} = 4\pi^2\varepsilon_0{\varepsilon_\infty}^{0.5}c^3\frac{\tau^2}{\lambda^2} \tag{17.29}$$

This interesting equation shows that, at a given wavelength, the fundamental material property that most strongly controls the TCO figure of merit σ/α is the average time between collisions τ for electrons. Moreover, the figure of merit is proportional to τ^2. Since $\tau = \mu m_e{}^*/e$ we may again think of the mobility as playing a major role in determining TCO quality. We note that the figure of merit also depends weakly on the high-frequency permittivity ε_∞ via its square root. For the record, the figures of merit σ/α calculated from Equation (17.29) (at $\lambda = 550\,\text{nm}$) for the three films modeled in Figure 17.6 are 1.0, 4.2, and 26 for films 1, 2, and 3, respectively.

It can also be shown that the absorption coefficient in the visible region, α_{vis}, arising from the tail of the free-carrier absorption, is approximately

$$\alpha_{vis} = \frac{Cn_e\lambda^2}{\varepsilon_\infty^{0.5}\mu m_e{}^{*2}} = \frac{377\sigma}{n(\omega\tau)^2} \tag{17.30}$$

where $C = e^3/(4\pi^2\varepsilon_0 c^3)$. The reader is reminded that Equation (17.30) has been derived from the classical Drude model. A quantum mechanical treatment yields a somewhat different dependence on wavelength [118]. In the second (compact) form of α_{vis}, the intrinsic impedance of free space $Z_0 = (\mu_0/\varepsilon_0)^{0.5} = 377\,\Omega$, n is the refractive index, σ is the DC conductivity and τ is the electron scattering time. Thus, if for ZnO:Al, $\sigma = 1250\,\text{S/cm}$, $n = 1.94$, $\mu = 35\,\text{cm}^2/\text{V s}$, then $\alpha_{vis}(550\,\text{nm}) = 430\,\text{cm}^{-1}$. The first form of Equation (17.30), which shows that $\alpha_{vis} \propto n_e/\mu$, confirms the need to reduce the carrier concentration n_e and to increase the carrier mobility μ to reduce the visible absorption coefficient.

We have shown that, at a given wavelength, the optical absorption coefficient for free carrier absorption is $\alpha = cn_e/\mu$ (Equation 17.30) where c is a constant. For $\alpha t \ll 1$, the optical absorptance A of a film of thickness t is approximately αt. Hence $A = cn_e t/\mu$. Since the sheet resistance $R_s = \rho/t$, we have $A = cn_e\rho/(\mu R_s) = cn_e/(\mu R_s n_e e\mu)$, i.e. $A \propto 1/(\mu^2 R_s)$. For a given R_s, then, the optical absorptance of a TCO film varies inversely with the square of the mobility. In solar cell applications we generally have to achieve a TCO sheet resistance of a certain value to keep I^2R losses low. Clearly, if two TCO materials are under consideration, then the material with the higher mobility should be chosen. This important point is further analyzed in the context of CdTe cells in Section 17.5.3.1 (see Figure 17.11).

As we have not yet shown graphically the behavior of spectral reflectance, we show in Figure 17.7 the spectral dependence of transmittance and reflectance for two TCO films on glass. We also take the opportunity to show the effect of the high-frequency dielectric constant ε_∞. In the visible region of the spectrum, ${\varepsilon_\infty}^{0.5}$ is approximately equal to the refractive index of the film, n. It is therefore not surprising that the amplitude of the oscillations in $T(\lambda)$ and $R(\lambda)$ should depend on ε_∞. The shift in plasma wavelength to larger values with increasing ε_∞ is again predictable from Equation (17.21) and a reduced absorption in the visible is a natural consequence. Although the average visible transmittance is lower for the film having the higher ε_∞ (due to a higher reflectance) it cannot be assumed that this conclusion would hold in an actual solar cell since the refractive index of the absorber layer is much higher than that of glass. By analogy with the observation that

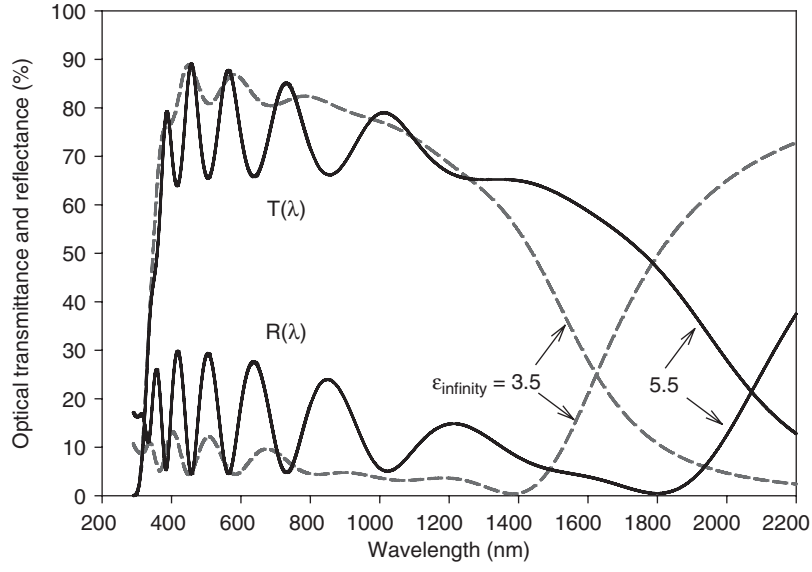


Figure 17.7 Calculated optical transmittance and reflectance of two TCO films with values of $\varepsilon_\infty = 3.5$ and 5.5. The fixed parameters were: film thickness 500 nm; carrier concentration $5 \times 10^{20}\,\mathrm{cm}^{-3}$; mobility $40\,\mathrm{cm}^2/\mathrm{V\,s}$; and effective mass $m_e^* = 0.35m_e$. The substrate is glass. The plasma wavelengths for these films are $\lambda_p = 1650$ nm and 2070 nm, respectively

SiO_2:HfO_2 can exhibit a significantly larger dielectric constant κ than SiO_2 with the introduction of a relatively small amount of high-κ HfO_2 [140], a similar strategy has been explored to increase the ε_∞ of a TCO (in this case In_2O_3 and ITO) via the addition of ZrO_2 [141].

Finally, we observe that the film refractive index $n \approx {\varepsilon_r}^{0.5} \approx {\varepsilon_\infty}^{0.5}(1 - {}^1\!/_2\omega_p^2/\omega^2)$ for $\omega_p^2/\omega^2 \ll 1$, i.e., in the visible and near-IR region. More conveniently, we may write this as $n \approx {\varepsilon_\infty}^{0.5}(1 - {}^1\!/_2\lambda^2/\lambda_p^2)$. From this we see that the refractive index of the material declines with increasing wavelength. Since ω_p^2 is proportional to the carrier concentration n_e we see that the rate of change of the refractive index n with increasing wavelength, $\mathrm{d}n/\mathrm{d}\lambda$, is proportional to $-n_e\lambda$ and is hence controlled by n_e. Since the optical performance of solar cells depends on the refractive index of the front TCO, this little-appreciated result is actually of some significance.

17.4.3 Influence of TCO Electrical and Optical Properties on Module Performance

The performance of monolithically integrated, thin film modules is directly affected by the electrical conductivity and optical transmission of the current-collecting window electrode (TCO). The conversion efficiency η of a unit cell may be written

$$\eta = \frac{I_{MPP}V_{MPP}}{P_{in}}(1 - f) \tag{17.31}$$

Here the various power loss factors under consideration are lumped into the overall loss factor f. I_{MPP} and V_{MPP} are the values of the cell current and voltage at the maximum power point (for zero losses) and P_{in} is the incident power falling on the entire unit cell.

In this section we will consider the optimization of module design (cell width) taking the TCO properties as given. We first calculate the total loss, taking into account the following three factors: the dead area resulting from the three patterning steps; the joule (I^2R) losses resulting from the sheet resistance R_{sh} ($\Omega/\square$) of the TCO; and the linear interconnect resistance R_c (Ω cm) resulting from contact resistance at the metallization-TCO interface in scribe #2 [142, 143]. Figure 17.9 appearing at the beginning of Section 17.5 may help with the visualization. (R_c is related to the more fundamental specific contact resistivity ρ_c (with units Ω cm^2) via $R_c = \rho_c/\Delta l$ where Δl is the width of scribe #2.) If the total cell width is $w_{tot} = w + w_d$ (where w is the active width and w_d is the dead width occupied by the interconnect region) the power loss due to the dead area is $J_{MPP}V_{MPP}w_dL$ and the cell power in the absence of this loss mechanism is $J_{MPP}V_{MPP}(w + w_d)L$ which we approximate as $J_{MPP}V_{MPP}wL$ since $w_d \ll w$. Hence the fractional power loss is simply w_d/w. To calculate joule losses in the TCO we note that the photocurrent builds up linearly in the TCO across the width of the cell before entering the interconnect at $x = w$ and so $I(x) = J_{MPP}Lx$, where L is the length of the cell (parallel to the scribe direction) and J_{MPP} is the cell current density. The power dissipated in the TCO is therefore

$$P = \int_0^w I^2(x)R_{sh}\,\mathrm{d}x/L = \frac{1}{3}J_{MPP}^2Lw^3R_{sh} \tag{17.32}$$

and the fractional power loss is $(1/3)J_{MPP}w^2R_{sh}/V_{MPP}$. This depends on the square of the cell width. Finally, it is easy to show that the fractional power loss due to interconnect resistance is $J_{MPP}wR_c/V_{MPP}$. Hence the total power loss fraction f is

$$f = \frac{w_d}{w} + \frac{1}{3}\frac{J_{MPP}}{V_{MPP}}w^2R_{sh} + \frac{J_{MPP}}{V_{MPP}}wR_c \tag{17.33}$$

For $R_c = 0$, the cell width w_{opt} that minimizes the loss fraction $f(w)$ can be shown to be

$$w_{opt} = \left(\frac{3}{2}w_d\frac{V_{MPP}}{J_{MPP}}\frac{1}{R_{sh}}\right)^{1/3} \tag{17.34}$$

and, it turns out, $f(w_{opt}) = (3/2)w_d/w_{opt}$ [41]. More generally, the variation of the three loss fractions considered in Equation (17.33) are depicted in Figure 17.8a as a function of active cell width w. The calculations were performed using parameters appropriate for a dual-junction (same bandgap), superstrate-type a-Si:H module deposited on SnO_2:F, although the curve shapes are similar for any type of integrated thin film module. The assumed parameters are shown in Table 17.6. Since the Joule loss in the thin-film metallization (e.g. Al) would be calculated in a manner identical to that used for the TCO, we added the values for R_{sh} and R_{met} to calculate the total losses for both materials. From Figure 17.8a it can be seen that, at 1 sun irradiance, the total power loss (relative to the maximum achievable) attains a minimum value of about 5.4% at a cell width of 0.90 cm. In practice, to reduce the number of scribes needed on the plate, a cell width of about 1.7 cm might be used in manufacturing, corresponding to a loss of 7.9%. At an irradiance of 0.5 suns, the losses are minimized at a larger cell width $w = 1.15$ cm. Clearly, the final choice of w for the module is dictated by the desired optimization criterion. A logical target would be to optimize the energy delivered by the module integrated over a full year of deployment. With this criterion, the temporal distribution of plane-of-array irradiance will result in a choice of cell width larger than w_{opt} evaluated at one sun.

This analysis can be extended so that the optimal thickness of the TCO can be determined. Since the optical absorption of commercial SnO_2:F is significant this means that the thickness dependence of the absorption must be included in the model. Taking as a starting point the 14 $\Omega/\square$ TCO used in the analysis to generate Figure 17.8, and assuming it has an absorbance of 0.05, we

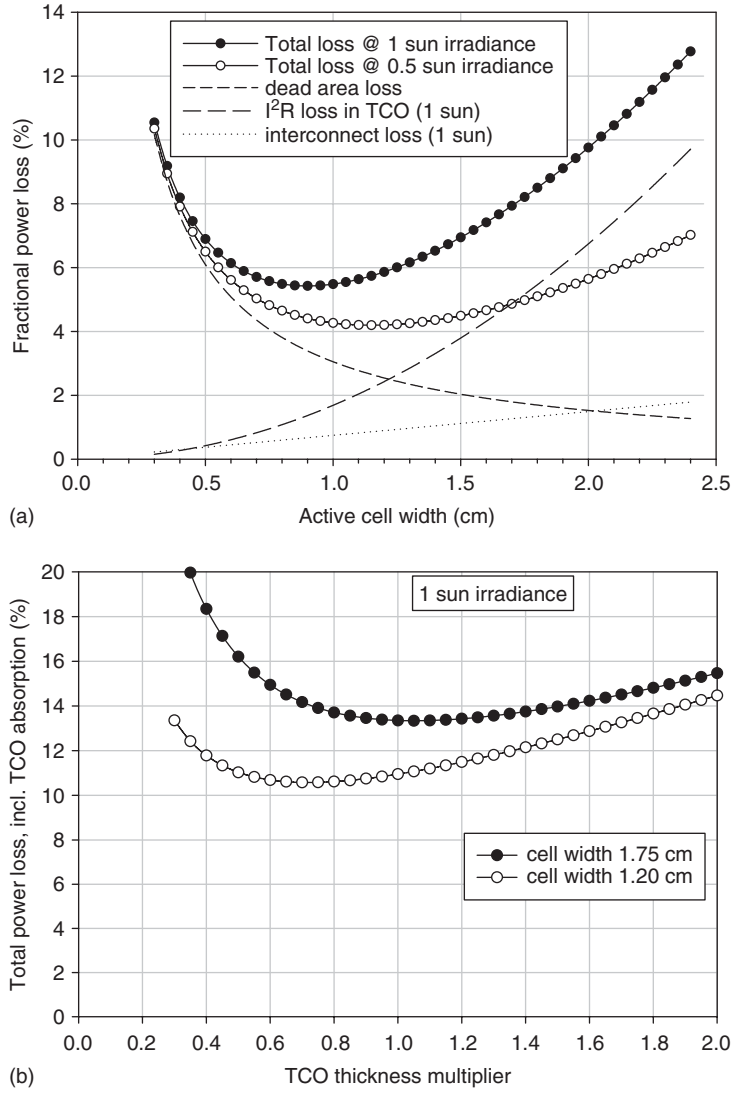


Figure 17.8 (a) Total and individual fractional power loss terms in thin film modules for the three mechanisms: (i) dead area; (ii) Joule loss in TCO, and (iii) Joule loss at interconnect, each as described by Equation (17.33) and plotted as a function of cell width w. The calculations were performed for dual-junction a-Si:H modules assuming 1 sun irradiance. The total loss at 0.5 suns is also plotted. (b) Total power loss including optical absorption in the TCO as a function of TCO thickness, for two different cell widths

may ask whether further gains in module performance might be secured by changing the thickness of the TCO. If the original TCO thickness is t_0 and the thickness is changed to mt_0, where m is a thickness multiplier, then in a simple linear model the sheet resistance will become $14/m\ \Omega/\square$

Table 17.6 Parameters assumed for hypothetical dual-junction a-Si:H module

Parameter	Value
J_{MPP} (1 sun)	5.0 mA cm^{-2}
V_{MPP}	1.443 V
w_d	0.0305 cm
R_{sh}	14.0 Ω/□
R_{met}	0.6 Ω/□
R_c	2.15 Ω cm

and the absorbance will become $0.05m$. With this additional loss term under consideration the expression for the total fractional power loss becomes

$$f(m) = A(m) + T\frac{w_d}{w} + \frac{T^2}{3}\frac{J^0_{MPP}}{V^0_{MPP}}w^2\frac{R_{sh}}{m} + T^2\frac{J^0_{MPP}}{V^0_{MPP}}wR_c \tag{17.35}$$

where $A(m) = 0.05m$, $T = T(m) = 1 - A(m)$, J^0_{MPP} is the value of J_{MPP} for zero loss (i.e. for $A = 0$ and $w_d = 0$) and as before we add R_{met} to R_{sh}/m. Figure 17.8b shows the total power loss f (including optical absorption in the TCO) as a function of the TCO thickness multiplier m for two different cell widths. At a cell width of 1.75 cm the overall power loss is minimized for $m = 1$, i.e. with no change to the original TCO thickness. However, at a cell width of 1.20 cm this thickness of TCO is no longer optimal and should be reduced by a multiplicative factor $m = 0.75$. This type of analysis may be further extended to include the dependence of haze on thickness. Yet another consideration for process optimization is the dependence of sputtering machine throughput and production cost on the thickness of the TCO layer.

17.5 PRINCIPAL MATERIALS AND ISSUES FOR THIN FILM AND WAFER-BASED PV

The properties of a transparent conductive oxide (TCO) used as a front electrode for thin film and other types of solar cells and modules play a significant role in determining the maximum attainable energy conversion efficiency. These properties may be categorized as electrical, optical, morphological, and chemical. The properties are manifested in the device in the following ways. The electrical parameters of the TCO, namely carrier concentration and mobility, determine its sheet resistance and therefore its I^2R power loss (with impact on device fill factor) and they also control optical transmission in the infra-red portion of the spectrum. Wavelength-dependent optical parameters such as transmission and absorption influence the amount of light that can be absorbed in the active layers of the device and hence its short-circuit current density. Morphological features on various length scales determine the degree of light scattering and hence the potential for light trapping. In addition, in superstrate devices, certain morphological features can result in cracks in the subsequently grown silicon layers thereby giving rise to shunting. Finally, the chemical nature of the TCO surface determines not only its electron affinity and work function that in turn influence the nature of the electrical contact between the TCO and the active semiconductor or doped layer, but also the resistance of the TCO to plasma attack and the types of elements that might diffuse into the semiconductor layers or participate in interfacial reactions.

17.5.1 TCOs for Superstrate-type Devices

In a superstrate device, the active semiconductor is deposited on a transparent substrate that has been pre-coated with a TCO. The substrate is usually glass, but polyimide is another possibility. In operation, the superstrate faces the sun and sunlight enters the device after passing through the superstrate and TCO layer. The TCO serves as a transparent front electrode (window layer) for the device. The main superstrate device types are based on amorphous silicon (a-Si:H), tandem amorphous silicon/nanocrystalline silicon (a-Si:H/nc-Si:H), cadmium telluride (CdTe), dye-sensitized titanium dioxide, and organic polymers. The general structure for monolithically interconnected, superstrate-type, thin film PV modules is shown in Figure 17.9. It can be seen that the back contact layer of one cell is arranged to deposit on the TCO electrode of the next cell in the region where the semiconductor has been removed by ablation by laser scribe #2.

17.5.1.1 TCOs for thin film Si devices

The vast majority of glass-based amorphous silicon (a-Si:H) modules sold today are prepared on a glass superstrate coated with fluorine-doped tin oxide, SnO_2:F, as the transparent conducting oxide (TCO). Indium tin oxide (ITO) is not suitable because of its susceptibility to reduction [144, 145] and lack of texture. The thickness and sheet resistance of the tin oxide coating are typically in the range of 550–900 nm and 8–16 Ω/□, respectively. Fortuitously, the refractive index of most TCOs (∼2.0) lies between that of glass (∼1.5) and that of a-Si:H (∼4.0 [146, 147]) thereby helping to couple light into the silicon. The tin oxide coating is deliberately grown to have a significant surface roughness of about 40 nm RMS. The texture imparts a hazy appearance to the TCO layer. A haze value of between 10–18% is typical (haze is defined in Sections 17.6 and 17.7). The surface texture is important for two principal reasons, namely, improving the adhesion of the a-Si:H layer, and

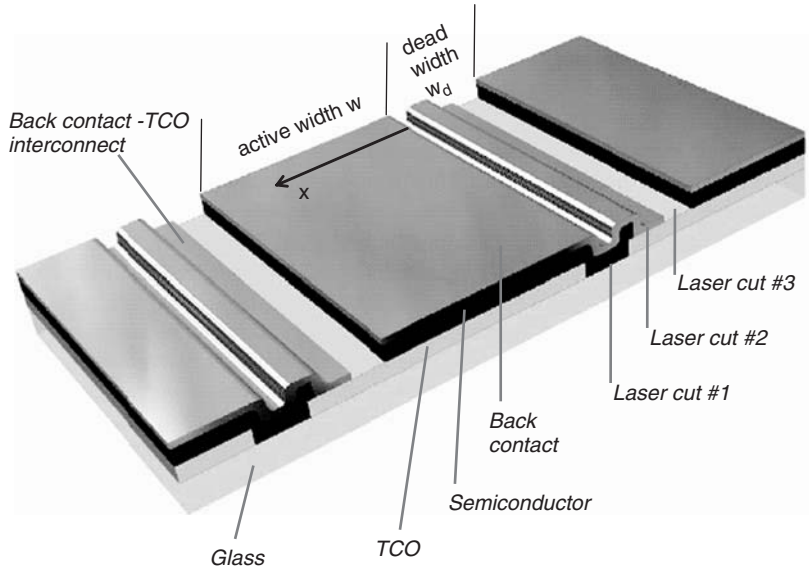


Figure 17.9 General structure of superstrate-type thin film module showing the thin film layers (TCO, semiconductor and back contact), and the interconnect region. Figure adapted from Oerlikon Solar presentation and used with permission

increasing the short-circuit current density J_{sc} of the device via improved light trapping. A third, although minor, effect of the nonplanar interface between the tin oxide and the a-Si:H is an effective grading of the refractive index that reduces the reflection loss at this interface. For superstrate a-Si:H or tandem a-Si:H/nc-Si:H cells, then, the TCO morphology is of paramount importance. Textured TCO glass (floatline SnO_2:F or PV TCO) is commercially available from several glass companies. The SnO_2:F coating is prepared pyrolytically by atmospheric pressure chemical vapor deposition (APCVD) at temperatures of almost 650 °C as described in Section 17.3. A few module manufacturers possess belt furnaces for producing their own tin oxide coated glass by APCVD.

The precise morphology of the TCO is of fundamental importance as it controls both light trapping and electrical shunting in the device. In order to reduce deposition time and hence production cost, the absorbers of solar cells should be made as thin as possible without compromising cell performance. For sufficiently thin absorbers, light trapping techniques must be employed to ensure that the incident sunlight can be effectively absorbed. Light trapping is achieved by using a physically textured TCO as the front window layer and a back contact scheme with high optical reflectivity (see Section 12.4) [148]. The textured TCO serves to scatter incident light and, more importantly, to partially replicate the texture after absorber growth at the absorber/reflector interface. Consequently, weakly-absorbed long-wavelength light that penetrates to the back of the cell is reflected back into the absorber over a range of angles because of the textured rear reflector. Unabsorbed light that reaches the front of the cell at an angle larger than the critical angle will undergo total internal reflection because the silicon has a higher refractive index than the front TCO. These internally reflected photons may bounce between the front and rear contacts many times N, thereby increasing the probability of their absorption in the active layer. The degree of light trapping is determined by the shape and size of the surface features of the TCO. It can be shown that the limiting value of effective absorption enhancement (relative to a single, normal pass) is $4n^2$, where n is the refractive index of the absorber [149].

It is of interest to understand where light in thin Si:H cells is ultimately absorbed. The fraction of light available for current generation is given by the effective transmission $T_{eff}(\lambda) = 1 - R_{cell}(\lambda) - A_{gl/TCO}(\lambda)$, where R_{cell} is the total reflectance of the cell and $A_{gl/TCO}$ is the true absorbance of the glass/TCO substrate measured using an index-matching fluid. The realized useful absorption is given by the quantum efficiency $QE(\lambda)$ (measured under reverse bias, if necessary). The difference $T_{eff}(\lambda) - QE(\lambda)$ represents the optical losses in the cell. These include absorption in the p-layer, and enhanced glass/TCO and back reflector losses due to the N light passes resulting from light trapping. The losses become considerable at longer wavelengths. For example, for a-Si:H on a high-haze TCO, $T_{eff}(750\,\text{nm}) = 0.6$ while $QE(750\,\text{nm}) = 0.2$. The contributing losses have been quantified in several studies [146, 150, 151].

For at least a decade it had appeared that commercial SnO_2:F suffered from two drawbacks that would prevent its use in next-generation thin film silicon modules utilizing nanocrystalline silicon (nc-Si:H). One drawback is that exposure of SnO_2:F to atomic hydrogen produced during PECVD can, to a greater or lesser extent, chemically reduce the film, thereby producing free Sn that causes optical absorption and film darkening (see also Section 17.5.7). Another drawback is that the standard TCO has an inappropriate morphology for effective light trapping for wavelengths ($\sim$1000 nm) near the band edge of the nc-Si:H, i.e. for wavelengths longer than those near the band edge of a-Si:H ($\sim$700 nm). For these reasons much attention was focused on the development of large area zinc oxide (ZnO) for use as a TCO in modules utilizing nc-Si:H. Compared with tin oxide, zinc oxide was reported to have superior stability in hydrogen plasmas. Zinc oxide also has lower optical absorption, and can be prepared with higher haze and a morphology more effective in providing light trapping in nc-Si:H. Because nc-Si:H has a smaller optical absorption coefficient, α, than a-Si:H (comparing their α values in their respective wavelength ranges of application), thicker layers must be used for effective absorption of sunlight. Light trapping then becomes particularly

important in helping to keep the cell thickness down. Textured TCOs are discussed in greater depth in Sections 17.6 and 17.9.

17.5.1.2 TCOs for CdTe devices

Thin film CdTe devices are heterojunctions with CdS as the *n*-type junction partner. The basic structure is glass/TCO/CdS/CdTe/back contact. Commercial CdTe modules generally use commercially-available SnO_2:F on soda-lime float glass as the front TCO. Occasionally, ITO is used, but In tends to diffuse into the subsequently deposited layers. The use of In_2O_3:F has also been reported [152]. Superior optical performance can be obtained with a TCO consisting of Cd_2SnO_4 [3, 153, 154]. This material not only has lower absorption than SnO_2:F, but also higher mobility (see Table 17.9). Annealed Cd_2SnO_4 apparently has a near defect-free microstructure [154, 155]. An obvious requirement for any TCO to be used for CdTe is that it withstand the high deposition temperature of CdTe (generally 500–600 °C) and the subsequent anneal in a halide vapor ($CdCl_2$). Under these conditions, ZnO:Al sputtered at room temperature is not suitable as a TCO. However, by keeping all processing temperatures < 390 °C, an all-sputtered ZnO:Al/CdTe/CdS cell having an efficiency of 14% has been prepared [156]. Because of the very high optical absorption coefficient of CdTe ($>10^4\,cm^{-1}$ for $\lambda < 0.73\,\mu m$) and a typical CdTe thickness of 6 μm, most light above the bandgap is absorbed within a depth of 1–2 μm. Light trapping is therefore not needed in conventional CdTe cells and TCO texture is not required. On the contrary, a smooth TCO (<3% haze) is necessary to minimize shunting. For SnO_2:F, the smoothness can be achieved by limiting the thickness and increasing the dopant concentration (see Table 17.9 in Section 17.6.4). The absorption loss in commercial tin oxide used for CdTe modules is about 6% compared with 3% for the tin oxide typically used for a-Si:H modules. In all mature CdTe cell designs, a thin, high-resistivity layer is overcoated on the TCO to improve cell performance. The high-resistivity layer can consist of intrinsic SnO_2, Ga_2O_3, or Zn_2SnO_4. The numerous reasons for its use are described in Section 17.5.3.1.

For use in high-performance, polycrystalline tandem cells targeting >20% efficiency, the top, higher-bandgap cell must be metalized with a TCO back electrode having a high NIR transmittance. One of the best top cells so far is a 13.9% CdTe cell having a thin, transparent Cu_xTe back contact backed by an ITO layer and an MgF_2 AR coating [157]. This cell achieves a very good NIR transmission of about 50%, although further reduction of NIR absorption is necessary (see Section 17.9.2).

17.5.1.3 TCOs for dye-sensitized solar cells (DSSC)

For glass-based DSSC, the photoelectrode traditionally consists of a 10-μm-thick nanoporous TiO_2 layer deposited on SnO_2:F [158]. A monolayer of dye adsorbed on the TiO_2 serves as a visible light absorber. The counter-electrode also consists of TCO-coated glass (SnO_2:F), but coated with a Pt catalyst [159]. To use ITO as the TCO, the ITO/TiO_2 is usually annealed in air or oxygen at 300–450 °C to sinter the TiO_2 nanoparticles, adhere the TiO_2 film to the ITO, and to remove residual organics. Unfortunately, this step increases the ITO resistivity by filling some of the oxygen vacancies. The ITO could be overcoated with SnO_2:F to prevent oxygen access to the ITO. However, the low resistivity of the ITO can be more simply recovered by annealing in a reducing (H_2) atmosphere at 350 °C for 2 h. Finally, the TiO_2 must be reoxidized which can be performed electrochemically in 0.1 *M* NaOH. (Without reoxidation, oxygen vacancies on the surface of the TiO_2 particles act as recombination centers shunting the TiO_2 to the electrolyte.) With these treatments, ITO-based photoelectrodes have yielded DSSC cell efficiencies comparable to those on SnO_2:F [160]. It was argued that the H_2 reduction step lowered the Schottky barrier at the ITO/TiO_2 interface. In the case of devices on polymer foil (e.g. PET), sputtered ITO can be used as the TCO.

17.5.1.4 TCOs for organic solar cells

One of the most common organic solar cell types is the bulk heterojunction [161, 162]. This consists of an interpenetrating network of a conjugated polymer (donor) and organic electron acceptor (soluble fullerene). The conjugated polymer serves to absorb photons and generate excitons. The excitons diffuse and dissociate at the interface with the acceptor. The general structure is glass/TCO (e.g. ITO)/interfacial layer (e.g. 30 nm PEDOT:PSS)/donor (e.g. P3HT):acceptor (e.g. PCBM)/hole-blocking layer (e.g. LiF)/back contact (e.g. Al). The ITO is generally 15–20 Ω/□. Before use, it is typically cleaned in an ultrasonic bath of acetone and isopropanol, rinsed in deionized water, and dried in an oven. It can be patterned using 33 wt % HCl at 70 °C. Finally, it is often treated with UV-ozone or oxygen plasma to remove residual organics and to increase the work function. The role of the PEDOT is to reduce the roughness of the ITO and to effect better band alignment with the HOMO level of the P3HT.

17.5.2 TCOs for Substrate-type Devices

In a substrate-type device, the active semiconductor either is deposited on a (usually) opaque substrate or it consists of a crystalline silicon wafer. The TCO is deposited on top. The main substrate-type devices are based on amorphous silicon (a-Si:H) and its double-junction or triple-junction variants, copper indium gallium diselenide or CIGS ($Cu(In,Ga)Se_2$), and the HIT cell (Heterojunction with Intrinsic Thin layer) on a c-Si wafer.

17.5.2.1 TCOs for a-Si:H alloy or a-Si:H/nc-Si:H cells

In the amorphous silicon case, the substrates are most commonly thin stainless steel or polyimide, the a-Si:H layers are deposited in the *n-i-p* sequence, and the top TCO is generally ITO in the highest efficiency R&D cells and either ITO, ZnO:Al or ZnO:B in manufactured products. In all cases a relatively gentle deposition process is required to avoid junction damage. Some of the recent work on flexible a-Si/nc-Si devices has employed a periodically textured polyethylene (PET) substrate (embossed in a roll-to-roll process) and a device structure PET/80 nm Ag/60 nm ZnO:Al/nc-Si:H/ZnO:B/a-Si:H/ZnO:B, as shown in Figure 17.10 [163–165]. This structure is notable for using transparent conducting ZnO in three locations and for different functions. The back reflector layers (Ag and ZnO:Al) are sputtered, whereas the intermediate reflector (see Section 17.5.4) and the top TCO consist of LPCVD ZnO:B (see Section 17.6.3).

17.5.2.2 TCOs for CIGS cells

In the case of CIGS, the most commonly used TCO is ZnO:Al. For production of CIGS cells in an R&D setting the TCO sheet resistance can be as high as 50 Ω/□ when used in conjunction with a grid having a small finger spacing. This sheet resistance can be achieved at a ZnO:Al thickness of about 200 nm. For CIGS modules, a sheet resistance of about 10 Ω/□ is desirable because of the high J_{sc} (~34 mA/cm^2) and a typical cell width of about 5 mm. To maintain transparency with these much thicker (~1 μm) layers, a high quality TCO with low optical absorption is necessary. A commonly encountered phenomenon is that the sheet resistance of ZnO:Al sputtered onto the CdS buffer layer is larger than that of the same film codeposited on glass, sometimes by as much as 50%. It is possible that this effect results from the filling of oxygen vacancies in the ZnO:Al with sulfur atoms, in which case the implication would be that there must be contributions to the conductivity from both Al dopant atoms and O vacancies, and that the latter portion of the conductivity is reduced by sulfur uptake.

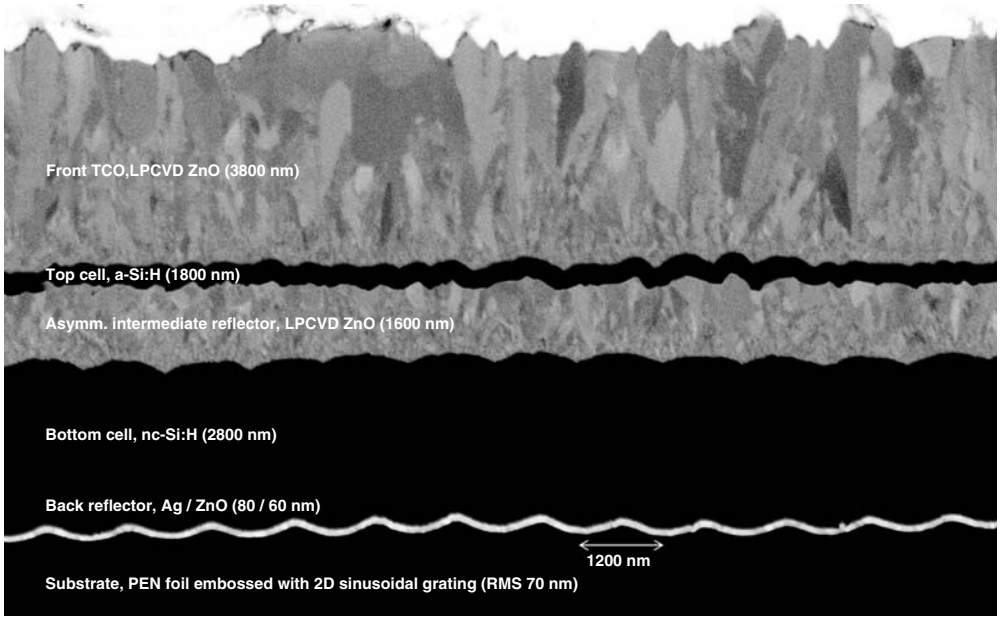


Figure 17.10 Cross-section image obtained by ion milling and SEM of a tandem a-Si/nc-Si cell on a periodically textured polymer substrate. The cell utilizes conducting and transparent ZnO layers for three functions: back reflector, intermediate reflector, and top TCO (bare SEM micrograph courtesy of Söderström T, IMT, EPFL, Switzerland)

In place of ZnO:Al, ITO can be used on CIGS cells. More recently, the use of amorphous indium zinc oxide has been explored. Other cell configurations may require the CIGS back contact to be transparent. For these applications (semi-transparent top cell in a tandem cell, or bifacial single-junction CIGS cell) it has been shown that ITO is a suitable contact to replace the Mo layer, provided the CIGS deposition temperature is kept below 520 °C [166].

17.5.2.3 TCOs for HIT cells

The HIT cell uses a doped a-Si:H layer deposited on crystalline Si to create a heterojunction. Avoiding epitaxial growth, an a-Si:H *i*-layer is inserted between the doped and crystalline layers. The cell is unique among commercially produced cells on c-Si in that it uses an ITO layer to effectively reduce the emitter sheet resistance. The structure used by Sanyo is: finger grid/80 nm ITO/15 nm p^+i a-Si:H/*n* CZ-Si (textured)/20 nm in^+ a-Si:H/80 nm ITO/back contact/finger grid. The role of the top ITO is twofold: to achieve high lateral conductivity and to function as an anti-reflection coating. The latter role dictates a well-controlled ITO thickness of 80–85 nm [167] and the former a grid spacing of 2 mm [168]. The role of the rear ITO is likewise twofold: to block metal diffusion and to balance stresses in the wafer to prevent warpage. This offers an effective route to the use of unusually thin wafers, with a 22.8% cell having been announced on a 98 µm wafer [169]. The rear ITO also allows the fabrication of bifacial modules that can utilize ground-scattered sunlight. The HIT cell features excellent surface passivation, unusually high V_{oc} (up to 743 mV), low temperature processing (<200 °C), and a low temperature coefficient. A gentle and low-temperature deposition process is required for the ITO to avoid junction damage. In R&D labs, the ITO is frequently prepared by reactive evaporation from a metallic source material plus oxygen at a substrate temperature of 200 °C or lower. In manufacturing, sputtering at a low power density

is most likely used. For an ITO resistivity of $1.7 \times 10^{-4}\,\Omega\,\text{cm}$ a sheet resistance of $20\,\Omega/\square$ can be obtained. Since minimization of absorption is important, electron mobility must be optimized and a higher resistivity is probably desirable. With the discovery of high-mobility TCOs such as In_2O_3:Ti [68] and In_2O_3:H [26] it seems likely that the ITO layers in the HIT cell will be replaced by one of these materials before long.

17.5.3 TCO/High-resistivity Layer and Other Bilayer Concepts

17.5.3.1 TCO/high-resistivity layer combinations

Most PV technologies that use a degenerately doped TCO as a window layer interpose a high-resistivity layer (HR layer) between the TCO and the semiconductor layers. The cell structure becomes TCO/HR layer/semiconductor junction/back contact. The presence of this high-resistivity layer reduces shunting effects and increases cell efficiency. The basic idea is that, without an HR layer, a localized pinhole or defect in the semiconductor layer will present a low-resistance shunt to the TCO layer. Because of the low lateral resistance of the TCO layer, this shunt is communicated across the entire cell, reducing FF and V_{oc}. If a layer having a relatively high resistivity is interposed, the voltage drop across the thickness of this layer due to the normal photocurrent (in the non-shunted areas of the device) remains small because of the thinness of the layer. However, a high current density can no longer flow through the defect as this would lead to a local voltage drop across the HR layer that exceeds the cell voltage. An HR layer resistivity in the range 10^0–$10^4\,\Omega\,\text{cm}$ is usual. Thus, a photocurrent of $20\,\text{mA/cm}^2$ passing through a 100 nm HR layer having a resistivity of $10^3\,\Omega\,\text{cm}$ would give rise to a voltage drop of 0.2 mV across the HR layer, representing a negligible power loss.

To calculate the impact of a pinhole we may consider the following model, namely, a large area cell operating at P_{max} in which a circular pinhole of area A and radius r_1 connects the metallic back contact to the TCO. The pinhole will sink photocurrent over a circular region up to a radius r_2 centered on the pinhole until the radial voltage drop in the TCO, $V(r_2) - V(r_1)$, is equal to the operating voltage of the cell V_{max}. It is then easy to show that the radius of influence r_2 of the pinhole is:

$$r_2 = k\sqrt{\frac{V_{\max}}{J_{\max} R_{sh}}} \tag{17.36}$$

For example, taking parameters appropriate for a CIGS cell, namely, $V_{max} = 0.52\,\text{V}$, $J_{max} = 30\,\text{mA/cm}^2$, TCO sheet resistance $R_{sh} = 15\,\Omega/\square$, then for a pinhole of area $A = 1 \times 10^{-6}\,\text{cm}^2$ the dimensionless quantity $k \approx 0.56$ and the radius of influence r_2 of the pinhole is about 0.6 cm. The current sunk into the pinhole is $J_{\max}\pi r_2^2 = 34\,\text{mA}$ and the effective shunt resistance presented by the pinhole is about $15\,\Omega$. On the other hand, if the pinhole is filled by the material of the HR layer with $\rho = 10^3\,\Omega\,\text{cm}$ and $t = 100\,\text{nm}$, then the resistance of this plug of material is $10\,\text{k}\Omega$. In other words, the effective shunt resistance associated with the pinhole has been increased by a factor of almost 700. Although we have imagined the shunt to be a zero-resistance short, a buffer layer also serves to protect against diode-like shunts in which weak diodes with low turn-on voltages exist in localized areas of the device. In this case a simple model might consist of a large-area cell having a small circular region in which the diode parameters are degraded. Once again, a plug of buffer material in series with the shunt (weak diode) serves to limit the radius of influence of the weak diode.

One of the most straightforward examples of the use of a high-resistivity layer is in CIGS cells and modules. The conventional structure is glass/Mo/CIGS/CdS/i-ZnO/n^+-ZnO, where the intrinsic ZnO (i-ZnO) fills the role of the high-resistivity layer. In this case the i-ZnO layer is about 500 Å (50 nm) in thickness and has a resistivity of about $50\,\Omega\,\text{cm}$ (sheet resistance of

$1 \times 10^7\ \Omega/\square$) [170]. The *i*-ZnO is preferably deposited from an undoped ZnO target by RF magnetron sputtering in an Ar/O_2 mixture. The partial pressure of oxygen controls the resistivity of the *i*-ZnO layer and is generally <1%. The *n*-layer is typically about 3500 Å in thickness (for small area devices) and can be deposited from a $ZnO:Al_2O_3$ (1–2 wt %) target by sputtering in pure Ar. For substrates on which multiple CIGS cells are prepared, the effect of the *i*-ZnO layer is to improve both the cell yield and efficiency (*FF* and V_{oc}) and to reduce the spread in efficiency. A parallel diode model similar to that discussed in the preceding paragraph has been analyzed to account for the observations [171]. If the *i*-ZnO is too thick (>70 nm) the *FF* declines apparently because part of it remains too resistive [172]. The protective effect of the *i*-ZnO layer has also been demonstrated via the deliberate introduction of defects and their observation using infrared thermography. The use of an *i*-ZnO buffer in CIGS devices can also be used with other highly conductive TCOs, e.g. *i*-ZnO/ITO. As we shall see in Section 17.8, this scheme appears to be more moisture-resistant.

High-resistivity buffer layers are also used in polycrystalline CdTe cells and modules [173, 174]. The structure of this type of device is generally glass/SnO_2:F/buffer layer/CdS/CdTe/back contact. If thick CdS layers are used the buffer layer generally confers no benefit. However, when thinner CdS layers are used, the buffer layer is definitely needed to avoid the deleterious effect of localized TCO/CdTe junctions [175, 176]. The most common buffer is undoped SnO_2, although Zn_2SnO_4 or $ZnSnO_x$ (ZTO), In_2O_3, ZnO, TiO_2 and Ga_2O_3 have also been used [155, 175, 177]. For thin (~300 Å) CdS films, a 30% gain in CdTe device efficiency can be obtained by inserting an 1800 Å *i*-SnO_2 layer deposited by LPCVD at 500 °C [174]. For CdTe applications, soda-lime float glass coated with SnO_2:F plus an undoped SnO_2 surface layer can be purchased commercially. In this TCO, the undoped SnO_2 is slightly reduced, and the presence of oxygen vacancies improves the adhesion of the subsequently deposited layers.

The performance of CdTe cells has been considerably advanced in an R&D setting through the use of: (i) a borosilicate glass substrate; (ii) a cadmium stannate (Cd_2SnO_4) TCO in conjunction with a ZTO buffer layer; and (iii) a nc-CdS:O window layer [154]. A cell efficiency of 16.5% was achieved (V_{oc} 845 mV, J_{sc} 25.9 mA/cm^2, *FF* 75.5%). The current gains resulting from the use of these materials are illustrated in Figure 17.11. The current gain from the borosilicate glass, relative to regular soda-lime glass that contains 0.1% iron oxide, is about 1.2 mA/cm^2; that from the improved TCO, Cd_2SnO_4, relative to the use of commercial SnO_2:F, is about 1.5 mA/cm^2; and that from the thinner and higher bandgap CdS, about 4.2 mA/cm^2. These current gains account for the increased J_{sc} of the NREL CdTe device (25.3 mA/cm^2) relative to that of the commercial CdTe device (18.4 mA/cm^2) analyzed in Figure 17.11. The parameters used for calculation of the free-carrier absorption (using the theory presented in Section 17.4.2) are given in the caption for this Figure. The current-equivalent of solar absorption in the TCO over the wavelength range 440–860 nm was found to be 0.7 mA/cm^2 for Cd_2SnO_4 and 2.5 mA/cm^2 for SnO_2:F. Assuming an average device QE of 0.84, a net 1.5 mA/cm^2 can be gained by switching to Cd_2SnO_4. The advantage of Cd_2SnO_4 relative to SnO_2:F stems from its higher electron mobility (see Section 17.4.2).

The zinc stannate buffer layer, Zn_2SnO_4, can be prepared by RF sputtering [178]. The film is amorphous if sputtered at room temperature, but becomes polycrystalline after annealing. The fundamental band gap was found to be 3.35 eV, with the optical gap rising to 3.89 eV with increasing carrier concentration. The ZTO buffer used in CdTe solar cells is reported as $ZnSnO_x$ [154]. The ZTO film is very resistive as grown, although its resistivity is reduced to 1–10 Ω cm after vacuum annealing at about 600 °C. It is found that interdiffusion of Zn into CdS and Cd into ZTO occurs during cell fabrication. This interdiffusion apparently confers several benefits, including improving TCO/CdS adhesion after the usual $CdCl_2$ treatment step, thinning of the CdS layer, and raising the bandgap of the CdS layer. The ZTO is also resistant to the nitric/phosphoric etchant used in back contact formation. This effect, together with the buffer layer effect, results in cells with high shunt resistances. Thus, in the case of CdTe cells, the buffer layer can play several distinct roles.

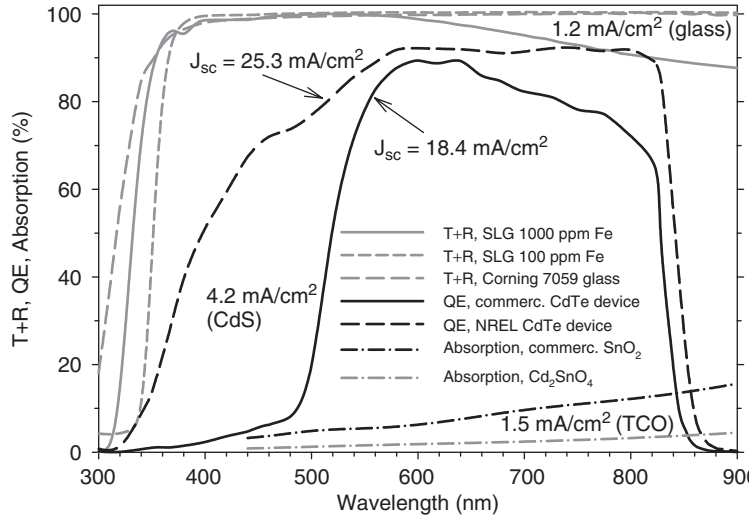


Figure 17.11 Reduction of current-loss factors in CdTe solar cells through use of improved TCO, buffer layer, and other materials. The modeling of free-carrier absorption in the TCOs assumed $\mu = 30\,\text{cm}^2/\text{V s}$, $n_e = 5 \times 10^{20}/\text{cm}^3$, $t = 400\,\text{nm}$ for the SnO_2:F, and $\mu = 50\,\text{cm}^2/\text{V s}$, $n_e = 8 \times 10^{20}/\text{cm}^3$, $t = 150\,\text{nm}$ for the Cd_2SnO_4. Both TCOs had a sheet resistance of 10.4 Ω/□ (Data for glass transmission and CdTe QE generated by T. Gessert and T. Coutts at NREL for the U.S. Department of Energy)

In a research environment, CdTe cells are sometimes deposited onto ITO as the TCO. The ITO is sputtered from a 90% In_2O_3 + 10% SnO_2 target to a thickness of about 400 nm. However, In diffuses into the CdTe layer. The In concentration in the CdTe can be significantly reduced through use of a 150 nm *i*-ZnO buffer layer. Alternatively, 50 nm of In_2O_3 can be used as a buffer layer [175]. It is unclear whether this does anything to limit In diffusion.

Although these examples of the use of buffer layers pertained to the amelioration of shunts and gain of other beneficial effects in polycrystalline devices (CIGS and CdTe), buffer layers are often used in a-Si:H devices. For example, a *p-i-n*-type device on tin oxide may have the structure glass/SnO_2:F/*p* a-SiC:H:B/*i* a-SiC:H (buffer)/*i* a-Si:H/*n* a-Si:H:P/Al. In this example, the function of the ∼10 nm a-SiC:H buffer layer is to increase the V_{oc} of the device. This is accomplished by a step in the conduction band edge preventing the back-diffusion of electrons from the *i*-layer to the *p*-layer, thereby reducing recombination. The role of buffer layers of various types, and their common occurrence in optimized PV devices, has been discussed in the literature [179, 180].

17.5.3.2 Bilayer TCOs

Multiple features not generally found in a single TCO layer may sometimes be obtained by combining two TCOs to form a bilayer TCO. Thus, the low cost and texture found in commercial SnO_2:F can be combined with the plasma resistance of ZnO by overcoating the tin oxide with a thin layer (∼10–20 nm) of ZnO. This is reported to increase the J_{sc} of a-Si:H cells by preventing reduction of the SnO_2 (see also Section 17.5.7). A bilayer TCO consisting of SnO_2:F/20 nm ZnO:Ga improves the efficiencies of nc-Si:H devices relative to those on bare tin oxide. Use of SnO_2:F overcoated with an ultra-thin (2 nm) layer of TiO_2 prepared by APCVD has been reported to increase the V_{oc} of a-Si:H solar cells by 30 mV [78]. To coat SnO_2 on wafer-based Si solar cells (to serve as both a

conductive layer and AR coating), interfacial layers of TiO_2 (or TiN) were developed as a barrier to SiO_2 formation [181].

The use of a SnO_2/TiO_2 bilayer offers the additional benefit of index-matching the SnO_2 to the thin Si layer [182–184] provided the thickness of the TiO_2 is optically significant (~50 nm). The refractive indices of the TCO, TiO_2, and Si layers are typically 1.9–2.0, 2.4–2.5, and 4.0, respectively. Ideally, the refractive index of the TiO_2 should be equal to the geometric mean of the indices for the TCO and Si layers, i.e. about 2.76, to minimize the reflectance. The TiO_2 can be prepared either by sputtering from an oxide target or by reactive sputtering from a metal target. A film resistivity of 10^4–$10^5\ \Omega$ cm is suitable, and can be obtained in a slightly sub-stoichiometric layer. In some cases, a trilayer structure consisting of $SnO_2/TiO_2/ZnO$ may offer the highest performance gain [182, 184]. Here the ZnO layer can be very thin (~10 nm); its role is to protect the TiO_2 against reduction and transmission loss.

Another concept is to combine a highly textured, undoped ZnO (*i*-ZnO) layer with a high mobility ($\mu \approx 80\ cm^2/V\ s$) doped TCO such as IMO (In_2O_3:Mo) or ITiO (In_2O_3:Ti) to form a bilayer such as *i*-ZnO/n^+-ITiO. In this structure, dubbed a TCLO (transparent conducting light trapping oxide), texture is provided by a first layer and high conductivity by a second layer, while both layers possess very low free-carrier absorption in the visible and near-IR region [185].

17.5.4 TCO as Intermediate Reflector

The basic structure of a superstrate thin film Si tandem cell is glass/front TCO/a-Si:H/nc-Si:H/back contact. In order to current match the top a-Si cell to the potentially high current of the bottom nc-Si cell, the thickness of the top *i*-layer has to be much larger than is desirable from the viewpoint of minimizing the Staebler–Wronski degradation of the top cell. By inserting an intermediate reflector (a ZnO:Al layer, for example) between the top and bottom cells some photons are reflected back into the top cell because of the difference in refractive index of the ZnO and Si layers. The cell structure is shown in Figure 17.1. The reflected light increases the J_{sc} of the top cell and decreases that of the bottom cell by roughly the same amount (~0.03 mA/cm^2/nm of reflector thickness for thicknesses <110 nm). The measured top-gain:bottom-loss current ratio is 4:3. The inclusion of the intermediate reflector allows the thickness of the top cell to be reduced to <300 nm, yielding a significant improvement in stability. A similar concept is applicable to substrate-type tandem cells. Glass-based, hybrid a-Si/nc-Si modules incorporating an intermediate reflector (or interlayer) have been produced prior to 2004 [186]. A ZnO:Al intermediate reflector can be deposited by RF magnetron sputtering. A suitable thickness range is 40–100 nm. With a surface-treated LPCVD ZnO front TCO, and thicknesses of top cell, bottom cell, and intermediate reflector of 290 nm, 3.0 μm, and 50 nm, respectively, a tandem cell with initial efficiency of 11.8% was achieved with $J_{sc,top} = 13.2$ mA/cm^2 and $J_{sc,bottom} = 12.8$ mA/cm^2 (sum of J_{sc} values 26.0 mA/cm^2) [187]. A further gain in performance has recently been achieved by substituting a phosphorus-doped a-Si,O:H intermediate reflector for the ZnO:Al reflector [188]. The SiO-based intermediate reflector has a crystalline volume fraction of 2–10% and has the advantage that it can be deposited *in situ* in the PECVD reactor. Yet another development has been the use of a so-called asymmetric intermediate reflector consisting of a relatively thick (1.6 μm) ZnO layer deposited by LPCVD in a-Si/nc-Si cells on a 2D-textured flexible PEN (polyethylene-naphthalate) foil (see Figure 17.10). In this case, the ZnO layer develops a texture during growth that considerably enhances the amount of light absorbed in a 200 nm top stack a-Si cell so that $J_{sc}^{top\ stack}$ increases from 9.5 mA/cm^2 (no reflector) to 12.5 mA/cm^2 (with LPCVD intermediate reflector) [163]. The feature size of the LPCVD intermediate reflector is about 300 nm so that green light is efficiently back-scattered into the top cell [165].

17.5.5 TCO Component of Back Reflectors

Light trapping in both superstrate and substrate thin film Si cells can be improved by an efficient back reflector. For example, with simple Al metallization as a back contact, the glass/front TCO/Si cell/back contact type of device suffers significant absorption of light in the Al layer. The absorption in the metal can be reduced by increasing the reflectivity at the rear Si interface through insertion of a TCO layer such as ITO [189] or, preferably, ZnO:Al [190]. The structure becomes glass/front TCO/Si cell/ZnO:Al/Al. The reflectivity is increased partly because of the downward step in refractive index at the a-Si:H/ZnO:Al (or nc-Si:H/ZnO:Al) interface and partly because the reflectivity at a dielectric/Al interface is higher for smaller values of the dielectric refractive index n, certainly for $n < 4$. Typical values for the reflectance at a-Si:H/reflector interfaces are 0.7 for Al, 0.82 for ZnO/Al, and 0.87 for ZnO/Ag [151]. The thickness of the ZnO:Al layer is not unduly critical and a suitable thickness is in the range 60–80 nm. The layer is generally prepared by magnetron sputtering. Since the refractive index n of ZnO:Al can be adjusted slightly through choice of deposition conditions it is clear that the smallest index should be sought. For example, a higher substrate temperature has been reported to lower the refractive index, with values $n(850\,\text{nm}, 25\,^{\circ}\text{C}) = 1.87$ and $n(850\,\text{nm}, 250\,^{\circ}\text{C}) = 1.79$ being obtained. Other approaches to increasing the reflectivity include alloying the ZnO with other materials (e.g. MgF_2) or the use of low-index/high-index multilayer stacks [191]. Alternatively, a patterned low-index dielectric can be used [146]. A higher reflectivity can be achieved with ZnO/Ag metallization, although care must be taken to limit Ag diffusion and to inhibit corrosion of the Ag layer. In back reflector applications the current flow is perpendicular to the plane of the ZnO film and since the transport distance is small, a high conductivity is not necessary. In fact, good cells have been made using i-ZnO rather than n-doped ZnO. However, low optical absorption in the ZnO remains important. The best optical performance can be achieved with a diffuse white reflector (DWR), the back contact scheme then being ZnO:Al/DWR. Here the ZnO:Al must be made much thicker (e.g. 500–1000 nm) in order to secure adequate lateral conductivity ($<20\,\Omega/\square$). A maximum in a-Si/nc-Si module power, representing the best compromise between sheet resistance and optical absorption, has been reported at a ZnO:Al thickness of 600 nm [192]. The diffuse white reflector is essentially a white paint consisting of TiO_2 pigments held in a binder. The reflection occurs via random Mie scattering from the TiO_2 particles. This type of reflector is effective not only in Si:H cells but also in thin crystalline Si cells [193].

For substrate-type thin Si cells on steel, the back reflector may consist of a textured Ag layer deposited at a substrate temperature of 350 °C overcoated with a sputtered ZnO:Al TCO that serves as a rear electrode, Ag diffusion barrier, and back reflector. The Ag texture is roughly 20 nm RMS. The optimal ZnO thickness depends on the texture. Modeling needs to take into account the Ag–ZnO surface plasmon absorption loss at 2.8 eV. The ZnO:Al can be further textured by chemical etching with 0.5% HCl to a RMS roughness of 60–80 nm. In the case of flexible plastic substrates (PET or PEN) the substrate can be textured by embossing.

17.5.6 Adjustment of TCO for Band Alignment

One conclusion of a theoretical analysis of CIGS solar cells using device simulation is that the highest efficiencies are achieved if the conduction band minimum of the window layer is about 0–0.4 eV higher than that of the CIGS layer [194]. This condition is fulfilled when a CdS buffer layer is used since its conduction band minimum (CBM) is 0.2–0.3 eV higher than that of CIGS. However, in a ZnO/CIGS junction, the CBM of ZnO is 0.2 eV lower than that of CIGS. This results in a barrier for electron injection in forward bias and leads to majority carrier (hole) recombination at defects at the ZnO/CIGS interface, lowering V_{oc} and FF [194, 195]. To make Cd-free devices, $Zn_{1-x}Mg_xO$ is often chosen as a replacement buffer layer for CdS to meet the conduction band offset requirement. The offset can be adjusted by changing the Mg content [196]. The introduction

of Mg into ZnO widens the bandgap principally by reducing the electron affinity (upward shift of CBM). For $Zn_{0.83}$ $Mg_{0.17}O$ the bandgap is 3.6 eV (widened from 3.24 eV) and the CBM is higher than that of CIGS ($E_g = 1.1$ eV) by 0.3 eV. For high-bandgap ($E_g \sim 1.3$ eV) CIGS cells with thin (<50 nm) CdS layers, the use of $Zn_{1-x}Mg_xO$ improves cell performance by reducing recombination at the CIGS/CdS interface [197]. The $Zn_{1-x}Mg_xO$ can be prepared by co-sputtering of ZnO and MgO targets or RF magnetron sputtering of a mixed target of fixed composition. The use of $Zn_{1-x}Mg_xO$ may also obviate the need for the usual *i*-ZnO layer.

17.5.7 Modification of TCO Properties

The bulk and surface properties of a TCO are not inviolate, but can be altered during processing or modified in a more deliberate manner by several types of treatment. To illustrate the various possibilities we will give some concrete examples. Of major interest to the a-Si:H and nc-Si:H community is the propensity of SnO_2:F to be reduced through exposure to atomic hydrogen at temperatures above a critical temperature (~200 °C) [198–202]. The reduction creates elemental tin at the surface, and if the exposure is sufficiently intense, small droplets of tin are formed. The deposition of a-Si:H by PECVD or by hot-wire CVD is accompanied by the production of atomic hydrogen, and when a-Si:H is deposited on SnO_2:F the freed oxygen reacts with Si atoms to form an SiO_x layer at the interface. Once formed, the SiO_x layer is a barrier to atomic hydrogen (which a-Si:H is not) and prevents further reduction of the SnO_2:F. The Sn constitutes a layer of up to 3 nm in thickness that results in optical absorption. It has also been reported that Sn can diffuse into the solar cell *p*-layer, thereby reducing its bandgap. Not all TCOs respond equivalently; while ITO undergoes a similar reduction in the presence of H, ZnO appears more resistant and does not change in transparency. However, a H-exposed ZnO surface does exhibit a downward band bending and consequently a lower work function. Conversely, exposure of ZnO to O increases its work function [203, and references therein]. Overcoating of SnO_2 with a 20 nm layer of ZnO (by sputtering at room temperature) is sufficient to protect the SnO_2 from damage by H [200]. A 7 nm SiO_2 layer formed by chemical transport also helps protect the SnO_2 surface [204].

The growth of a-Si:H *p-i-n* solar cells on ZnO (as opposed to SnO_2) has historically resulted in a lower device fill factor. There are several controversial issues surrounding this bare fact, including whether the ZnO is in fact reduced and whether a ZnO/*p*-layer contact problem exists. Despite suggestions to the contrary, the weight of evidence supports reduction of ZnO at the ZnO/a-SiC:H interface. Thus, the formation of SiO_2 is observed by XPS and the presence of elemental Zn is detected at the interface [205]. (In the absence of a covering layer, elemental Zn would presumably evaporate from a H-treated ZnO surface at comparatively low temperatures.) It is further known from real-time ellipsometry that H penetrates >200 Å into ZnO and increases the near-surface optical gap and conductivity [206]. It was then argued that the *p*-layer became depleted, resulting in series resistance that reduced the fill factor. A high fill factor was obtained using a TCO/*p*-layer structure: ZnO/20 nm (*p*) μc-Si/10 nm (*p*) a-SiC:H. Modeling of the TCO/*p-i-n* structure reveals a downward bending of the *p*-layer bands at the TCO interface, the magnitude depending on the work function of the TCO and the density of interface states, and representing a barrier to the extraction of holes [207]. However, it was later demonstrated by direct measurement of the TCO/*p*-layer contact resistance that the contact is ohmic and of low resistance (0.6–1.3 Ω cm^2) [208]. How then to explain the drop in fill factor? It had already been noted that the presence of ZnO and its surface condition has a major influence on the forward current and diode quality factor. Finally it was observed that overcoating ZnO:F with a-Ge:H in a 4-second deposition resulted in: (i) high fill factors; (ii) lowering of the diode quality factor from 2–3 (indicative of multiple forward current mechanisms) to 1.7; (iii) reduced tailing of both O and Zn into the a-Si:H layers [209]. It was further concluded that (without the a-Ge:H layer) ZnO is reduced to a greater extent than SnO_2. From this work we conclude that the fill factor problem is impurity related.

Table 17.7 Annealing of Asahi SnO_2:F at 300 °C for 30 minutes (after [202])

Annealing atmosphere	R_{sh} ($\Omega/\square$)	Mobility ($cm^2/V\,s$)
Initial	13	31
H_2/Ar	8	51
Ar	7	52
Air	9	40

It has been found that some types of commercially available SnO_2:F/glass substrates can be beneficially modified by thermal annealing treatments while others cannot. For those that exhibit the effect, the principal observation is that the Hall mobility can be substantially increased by annealing in a reducing or oxygen-free atmosphere at temperatures of 300–400 °C [202]. Table 17.7 shows some results for an Asahi substrate. The increase in mobility from 30 to 50 $cm^2/V\,s$ was found to be permanent. The annealing did not change the carrier concentration. Samples obtained at that time from AFG and LOF were unaffected by the treatment. In the case of the Asahi substrate, beneficial results were also obtained using a H_2 plasma treatment at 150 °C, i.e. below the threshold temperature for SnO_2 reduction. It was concluded that the mobility was limited by grain boundary scattering, and that the barrier height can be modulated by adsorption or desorption of O_2 (or an O-related species) at the grain boundaries [210, 202]. The mobility increase was found only for samples having a carrier concentration $<2 \times 10^{20}$ cm^{-3}. In samples with a higher carrier concentration the mobility is limited by ionized impurity scattering (see Section 17.4.2).

In earlier work by Asahi Glass Co., SnO_2:F samples having $n = 1.3 \times 10^{20}\,cm^{-3}$ and an initial mobility of 10 $cm^2/V\,s$ were overcoated with 100 Å ZnO and treated at 300 °C with an RF H_2 plasma. A peak mobility of 70 $cm^2/V\,s$ was attained after a treatment time of 60 s [199]. In the absence of a ZnO overcoat, a mobility of 45 $cm^2/V\,s$ was achieved by H_2 plasma treatment together with a decreased Si_xO_y interfacial layer after deposition of a-Si:H [211].

17.6 TEXTURED FILMS

The preceding discussion, both in the theory section (Section 17.4) and regarding several applications, has largely assumed that the TCO layer is smooth (specular). However, in many cases it is desirable that the TCO layer possess significant surface roughness. The literature variously refers to such layers as being granular, textured, or hazy. One of the most important features of a textured TCO is its ability to scatter light. One parameter to measure this ability is called the haze factor, or simply haze, $H_T(\lambda)$, which is the ratio of the diffuse transmittance $T_{diff}(\lambda)$ to the total transmittance $T_{tot}(\lambda)$:

$$H_T(\lambda) = \frac{T_{diff}(\lambda)}{T_{tot}(\lambda)} \tag{17.37}$$

where the total transmittance, $T_{tot}(\lambda) = T_{spec}(\lambda) + T_{diff}(\lambda)$, is the sum of the specular transmittance and the diffuse transmittance. Note that the haze factor is a function of wavelength. (An even simpler measure is a white-light haze; see Section 17.7.2.2 for experimental details.)

For thin film solar cells in which the optical absorption coefficient is low in the near-IR region, and especially for a-Si:H and nc-Si:H solar cells, the use of light trapping technology

becomes critical. A textured TCO not only reduces the reflection loss but, with its ability to scatter light, increases the optical path length and enhances the trapping of light in the solar cell. As a result, the overall light absorption in the active layers of the solar cell can be significantly increased.

17.6.1 Morphological Effects in a-Si:H Devices

The effect of TCO morphology on a-Si:H solar cell performance has been studied by several groups [212–214]. In an early US study, textured tin oxide films doped with fluorine, SnO_2:F, were grown from tetramethyltin, bromotrifluoromethane and oxygen by APCVD at 590 °C [212]. The film roughness and surface feature size, inferred from diffuse transmission at 633 nm, increased with increasing film thickness. Indeed, an increase in TCO roughness with thickness is generally observed, regardless of deposition method. A convenient measure of light trapping in a-Si:H *p-i-n* cells fabricated on various TCO layers is the value of the quantum efficiency at 700 nm (i.e. in the weakly absorbing region), provided the *i*-layer thickness and bandgap are held constant. For a 0.5 μm *p-i-n* cell (with Al metallization), $QE(700\,\text{nm})$ is about 0.08 for a smooth tin oxide film, but rises rapidly to about 0.28 for a diffuse transmission of 5–10% and plateaus at this value for larger diffuse transmissions up to 30% [212]. Thus, for a-Si:H, the larger roughness values are not necessary, and if achieved by increasing TCO thickness, are likely counterproductive because of increased optical absorption. This study also demonstrated a monotonic decline in open-circuit voltage, V_{oc}, with increasing diffuse transmission, from 880 mV for a smooth film to 780 mV at 30% diffuse transmission. A variety of film textures were achieved, including smooth, jagged, rod-like, and needle-like. The last two morphology types led to shunting and reduction of device fill factor.

17.6.2 Targeted Development of Textured SnO_2:F

Pioneering work in analyzing the relationship between TCO morphology and a-Si:H solar cell performance was also conducted by Fuji Electric [213–215]. Although successful in increasing J_{sc}, the so-called Type-A TCO (SnO_2:F) with a sharp-edged pyramidal structure led to a loss in V_{oc} [213]. This was attributed to defects in the a-Si:H resulting from steep V-shaped valleys in the TCO [214]. The V_{oc} loss can be severe for the thin top cell of an a-Si/a-Si tandem device [215]. Further development work by Asahi Glass Co. led to the creation of Type-U TCO with crystalline facets and with a shallower angle in the valleys [77]. The correlation between surface structure and solar cell performance parameters was studied via TEM images of cross-sections and SEM images of the surface. An SEM micrograph of a Type-U SnO_2:F TCO film (thickness 0.76 μm) is shown in Figure 17.12 (a). It was grown by APCVD using $SnCl_4$, methanol, and HF as reactants. This type of TCO has a lower fluorine concentration and hence lower carrier concentration. It is apparently grown slowly, resulting in higher crystalline perfection. These two factors presumably account for the lower optical absorption in the visible range of Type-U TCO (approximately 2.5%) compared with that of type-A TCO (approximately 5%) since the free carrier absorption tail extends into the visible range (see Section 17.4.3) and crystal defects or a suboxide may also contribute to absorption. The Type-U TCO results in minimal V_{oc} loss and is generally regarded as a high-quality standard against which other TCOs are judged. Indeed, a more recent comparison of a-Si:H cells on SnO_2:F grown by APCVD using TMT, MBTC, and TTC precursors (see also Section 17.3.2) revealed the reference Asahi Type-U TCO to yield the highest quality cells as judged by PV parameters, diode quality factor and reverse saturation current [216]. However, the cost of Type-U TCO glass has traditionally been too high for it to be used in module manufacturing.

Type-U TCO designates SnO_2:F with a certain shape of pyramidal surface structure, and can be prepared over a range of thicknesses. As film thickness is increased, so is the RMS roughness and consequently the haze. Figure 17.12 (b) shows a Type-U TCO having a thickness of 1.4 μm.

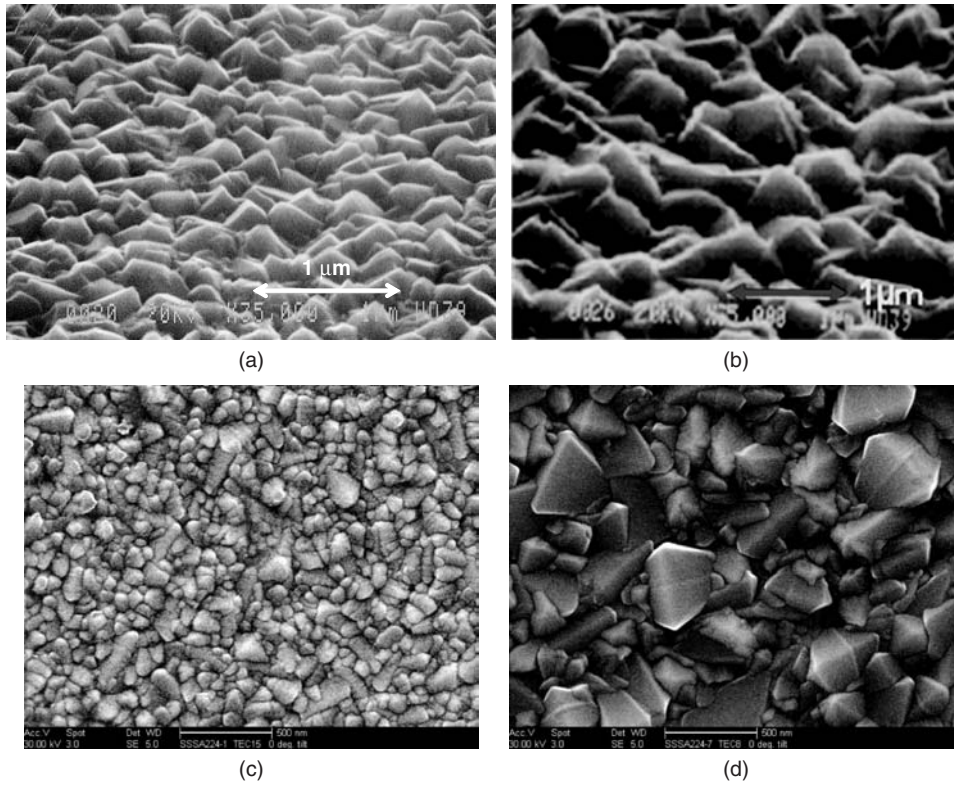


Figure 17.12 SEM micrographs of textured SnO_2:F: (a) Asahi Type-U, 0.76 μm, 16% haze, RMS 42 nm; (b) Asahi Type-U, 1.4 μm, 48% haze, RMS 62 nm; (c) Pilkington (now part of the NSG group) TEC15, 0.32 μm, haze 0.8%, RMS 12 nm; (d) Pilkington TEC8, 0.65 μm, haze 10-13%, RMS 35 nm. Micrographs (a) and (b) courtesy of Asahi Glass Company Ltd; (c) and (d) courtesy of Dr. David A. Strickler, NSG-Pilkington, and used with permission

Naturally, the optical absorption is also increased. Micrographs of the commonly used Pilkington TEC15 and TEC8 products are also shown in Figures 17.12c, d. The trends in product properties as a function of thickness are shown in Table 17.8 for Type-U TCO. As we have mentioned, the haze at a given wavelength is defined as the ratio of the diffuse (scattered) transmittance to the total transmittance. Figure 17.13 shows the haze as a function of wavelength for the three TCO films of different thickness described in Table 17.8. The haze of these films decreases rapidly with increasing wavelength. This represents a serious drawback for application of this type of TCO in cells whose response may extend to 1000 nm, e.g. nanocrystalline silicon cells.

17.6.3 Preparation and Properties of Textured ZnO

Textured ZnO has attracted increasing interest for thin Si:H cells, especially those involving nc-Si:H, because, compared with textured tin oxide, it possesses superior electrical and optical properties, excellent light scattering features and hydrogen plasma resistance. There have been sporadic reports of textured ZnO films being produced directly by magnetron sputtering. Of particular interest is the report of ZnO:B films with a resistivity as low as $4 \times 10^{-4}\ \Omega$ cm and a haze value of 28%

Table 17.8 Properties of Asahi Type-U TCO as a function of thickness

Sample	Thickness (μm)	*RMS* (nm)	R_{sh} (Ω/□)	*T* @ 550 nm (%)	Absorption @ 550/800 nm (%)	Haze @ 550 nm (%)
(a)	0.76	45	8	88	3.0/3.3	16
(b)	1.4	62	5	86		42
(c)	2.1	80	4	85	6.4/8.6	58

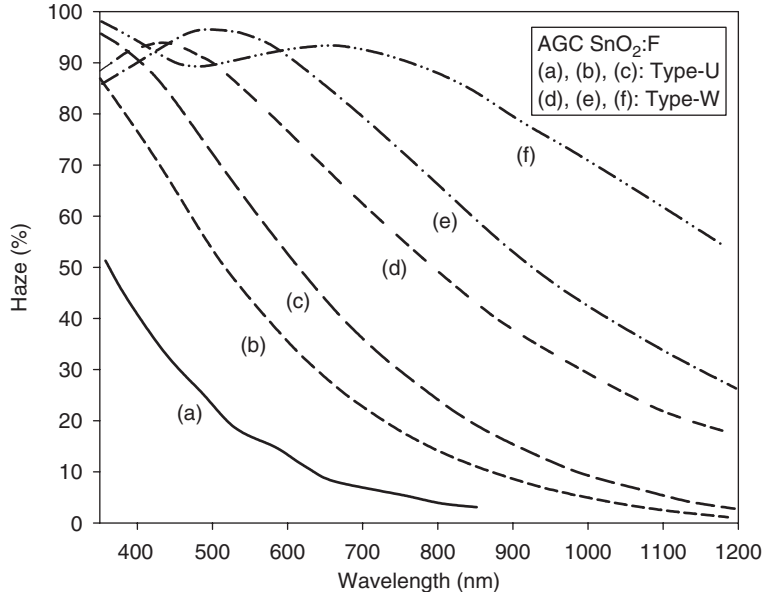


Figure 17.13 Haze of glass/SnO_2:F system as a function of wavelength. Curves labeled (a), (b), (c): Type-U TCO films of different thickness and with RMS roughness of 45, 62, and 80 nm as described in Table 17.8; curves labeled (d), (e), (f): Type-W TCO films with RMS roughness of 109, 122, and 150 nm as described in Section 17.9. Figure constructed from data in Taneda N, Oyama T, Sato K, *Tech. Digest PVSEC 17*, p. 309 (2007) [217]

at 550 nm having been prepared by DC magnetron sputtering of an undoped ceramic ZnO target in a 2.5% B_2H_6–Ar gas mixture [218]. Textured ZnO:F has been grown by APCVD [81] (see Section 17.3.2) and has been used for a-Si:H device fabrication with ultimately very good results (see Section 17.5.7). More recently, the global effort to develop textured transparent conductive ZnO films for PV applications has focused on two principal methods of preparation: (a) low-pressure chemical vapor deposition (LPCVD) by which textured films of ZnO:B can be grown directly; and (b) deposition of smooth ZnO:Al films by magnetron sputtering with subsequent chemical etching (e.g. in 0.5% HCl) to produce the desired texture. Significant progress has been made in scaling up these two techniques for high volume production. We will discuss them in turn.

It has long been recognized that ZnO thin films having a rough surface can be obtained by LPCVD [219, 220] with textured ZnO:B films for solar cells also having been systematically developed [221]. More recently, large-area textured ZnO:B films have been grown by LPCVD using

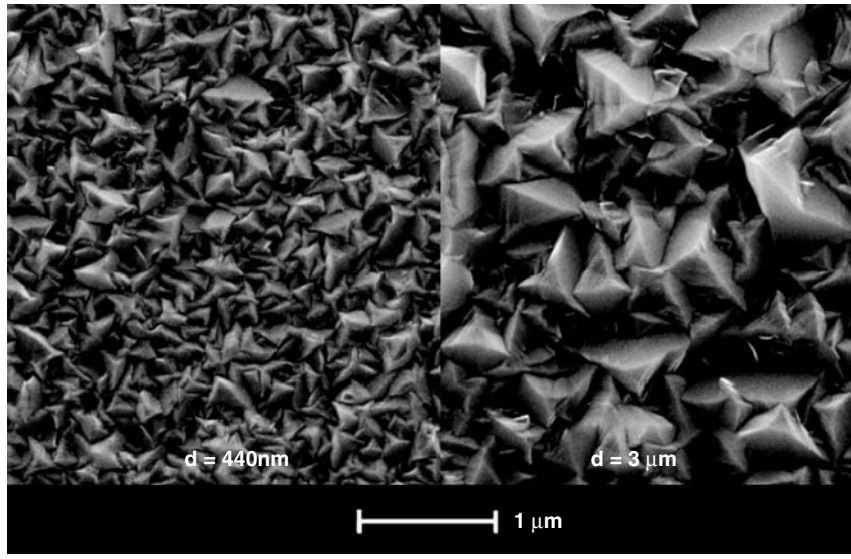


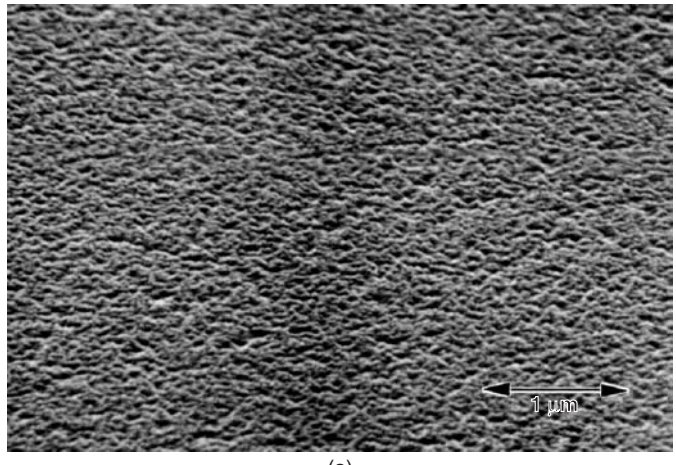
Figure 17.14 SEM micrographs of LPCVD ZnO:B films having thicknesses of 440 and 3000 nm. Reproduced from Faÿ S *et al*., *Thin Solid Films* **515**, 8558 (2007), with permission from Elsevier [86]

diethylzinc (DEZ) and water as precursors and diborane (B_2H_6) as the doping gas [85, 222]. The gases are introduced via a gas showerhead, with a less reactive oxygen source perhaps being useful to limit premature gas-phase reaction. The deposition process and suitable deposition parameters have already been discussed in Section 17.3.2. The morphology of LPCVD-deposited ZnO:B is shown in Figure 17.14 for two different film thicknesses. Since the surface feature size increases roughly linearly with thickness, the light scattering capability and haze also strongly increase with thickness, thereby increasing the photogenerated current in nc-Si:H cells [85].

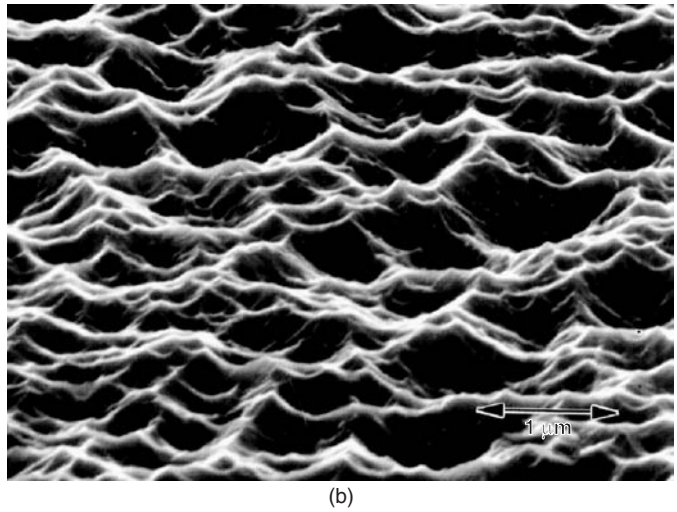
Textured ZnO:Al films can also be produced by post-deposition etching of sputtered ZnO:Al films using dilute hydrochloric acid (HCl) [223–226]. This process has been known since the late 1980s [227]. First, a ZnO:Al film is deposited by RF or MF magnetron sputtering of a ceramic target or by DC or MF reactive sputtering from a metal target. The ceramic target process is currently being transferred from planar to dual cylindrical magnetrons [228]. As deposited, the film is quite smooth. Then the ZnO:Al film is etched for less than a minute in 0.5% HCl. The electrical properties are determined only by deposition parameters and are almost unaffected by the etching process. The surface morphology after etching depends sensitively on deposition parameters and can be optimized for thin film solar cell application. Figure 17.15 shows SEM images of a film before and after etching. A uniform set of craters can be obtained that result in effective light trapping.

17.6.4 Other Methods to Prepare Textured TCO Film

Despite being highly developed, the LPCVD approach and the etching approach to producing textured films are not without drawbacks. The LPCVD process is very temperature sensitive, making large-area uniformity problematic, and LPCVD ZnO is not particularly stable. Furthermore, the precursor DEZ is expensive. And on the other hand, the etching of ZnO involves a cumbersome wet chemical process, the results of which are dependent on the precise nucleation and sputtering conditions of the ZnO. A new process to directly deposit a textured ZnO TCO has recently been



(a)



(b)

Figure 17.15 SEM images of RF-sputtered ZnO:Al: (a) as-deposited; (b) after etching in dilute HCl. Reproduced from Kluth O *et al*., *Thin Solid Films* **351**, 247 (1999), with permission from Elsevier [223]

developed [18, 229]. The process utilizes reactive-environment hollow cathode sputtering (RE-HCS). In this method, the deposition pressure is about 0.3 mbar (225 mTorr) which is much higher than that of traditional magnetron sputtering. The films also show competitive electrical and optical properties (see Section 17.3.1.1). One type of surface morphology obtainable using this technique is shown in Figure 17.16.

Other methods to deposit textured ZnO, such as expanding thermal plasma CVD [230], and a two-step, photo-MOCVD/ALD process [231], have also been reported. In the latter process, one of the functions of the ALD layer is to prevent the deterioration of the conductivity of the MOCVD layer [232].

We conclude this section by showing for comparison purposes the properties of various types of TCO used in PV device fabrication in the superstrate configuration. All of the TCOs are

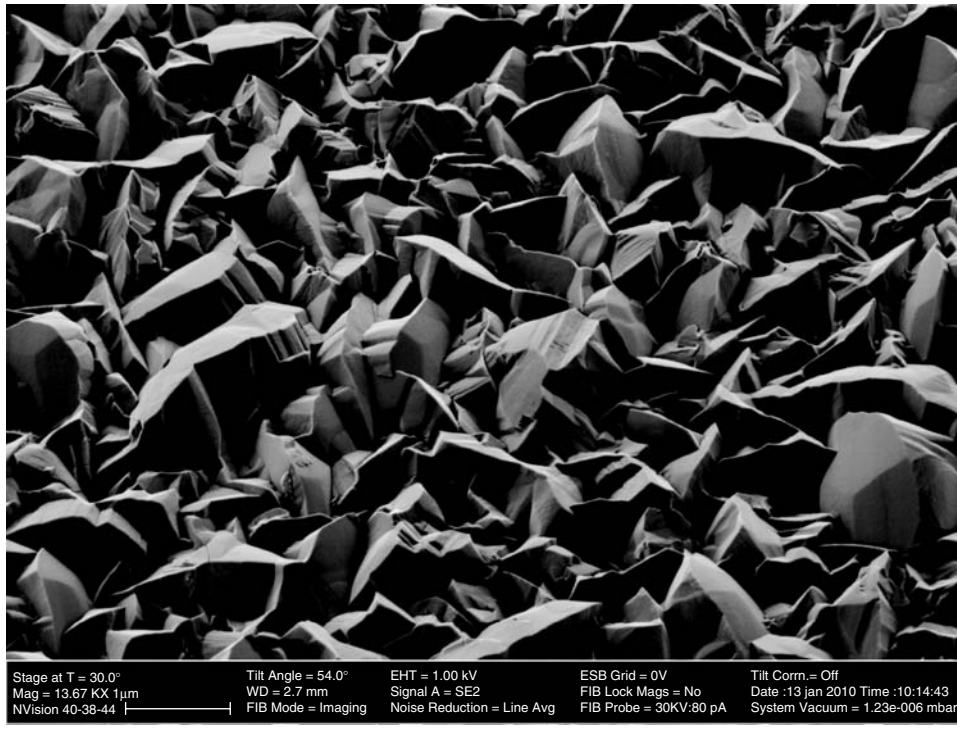


Figure 17.16 SEM image of textured ZnO:Al produced by RE-HCS (film prepared at EPV SOLAR, Inc., NJ; image courtesy of Stuart Boden, University of Southampton)

textured except for the two types used for CdTe. It should be pointed out that ongoing improvement of the extensively used SnO_2:F products, both floatline and off-line, is taking place. The previously described Type-U TCO is now being replaced by a Type-VU product with improved transmission and mobility. Thus, in Table 17.9 below we show the approximate properties of two types of floatline SnO_2:F, the higher grade Type-U and Type-VU TCOs that are produced off line, an example of LPCVD ZnO:B optimized for cell performance, an example of textured ZnO:Al produced by RE-HCS (see Section 17.3.1.1), and an example of sputtered (and annealed) Cd_2SnO_4. The properties given for the floatline SnO_2:F appropriate for dual-junction a-Si module production are approximately those of the AGC SOLAR (North America) AN14 TCO and those for CdTe module production are approximately those of the NSG-Pilkington TEC15 TCO.

17.6.5 Textured TCO Films: Description and Light Scattering

Additional characterization of TCO surfaces can be provided by surface height analysis and the angular dependence of optical scattering. It has been shown that both the vertical and the lateral dimensions of TCO surface features play a role in determining the effectiveness of the TCO in producing light trapping [233]. The root-mean-square (RMS) roughness δ_{rms} is defined as

$$\delta_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_i - z_{av})^2} \tag{17.38}$$

Table 17.9 Properties of TCOs used in superstrate PV devices in various areas of application

Property	Floatline SnO_2:F	Floatline SnO_2:F	Type-U SnO_2:F	Type-VU SnO_2:F	LPCVD ZnO:B	RE-HCS ZnO:Al	Sputtered Cd_2SnO_4
Application	2J a-Si	CdTe	R&D	a-Si/nc-Si	a-Si/nc-Si	R&D	CdTe (R&D)
Glass type	SL	SL	SL	SL, low Fe	SL, low Fe	SL	Corning 7059
Transmission* (%)	83–85	85	87	87.5		85.5	~90 (10 Ω/□)
Haze (%)	15–16	0.8	~ 16	25	50	33	
Film thickness (nm)	600	320	900	890	3000	1030	510
Sheet resistance (Ω/□)	14.3	13	8.7	8.3	10.0	2.8	2.6
Mobility ($cm^2/V\,s$)	30.6	30	37.0	56	28.0	49.5	54.5
Carrier conc. ($\times 10^{20}\,cm^{-3}$)	2.4	5.0	2.1	1.5	0.75	4.4	8.9
Resistivity ($\times 10^{-4}\,\Omega\,cm$)	8.6	4.2	7.8	7.4	30	2.9	1.3

*Total transmission, immersion method.

where N is the number of data points, z_i is the surface height at the ith data point and z_{av} is the average surface height. From scalar scattering theory, the haze in reflection or transmission can be written as a function of the RMS surface roughness δ_{rms}; in reflection $H_R(\lambda)$ is given by [234]

$$H_R(\lambda) = \frac{R_{diff}}{R_{tot}} = 1 - \exp\left[-\left(\frac{4\pi\,\delta_{rms} n \cos\theta_i}{\lambda}\right)^2\right] \tag{17.39}$$

where n is the refraction index of the scattering material and θ_i is the angle of incidence. If the correlation length $\sigma \gg \lambda$ (see below), the scattering is proportional to $(\delta_{rms}/\lambda)^2$. However, the assumptions under which Equation (17.39) was derived may not be valid for certain TCO surfaces of interest. The haze at 700 nm has been reported to correlate reasonably well with δ_{rms} [235].

While δ_{rms} is a measure of vertical variations, the correlation length σ is a measure of lateral variations. The correlation length is defined as the distance over which the autocovariance function falls to $1/e$ of its initial value. The autocovariance function can be calculated from

$$G(m) = \frac{1}{N}\sum_{i=1}^{N-m} z_i z_{i+m} \quad m = 0, 1, 2, \ldots (N-1) \tag{17.40}$$

Thus the data set is shifted through a distance (called the lag length, τ) equal to the product of the point spacing and m, after which the average of the height products of the two data sets is calculated. Often, the autocovariance function is a Gaussian $G(\tau) = {\delta_{rms}}^2 \exp(-\tau^2/\sigma^2)$. Another function of interest is the power spectral density, PSD, of the surface. This is essentially the squared Fourier transform of the surface; it therefore resolves the rough surface into components of various spatial wavelengths. The power spectral density and the autocovariance function form a Fourier transform pair [236]. For a textured TCO, spatial wavelengths in the range 0.1–5.0 μm are mostly of interest. It has been shown that the shape of the PSD function versus spatial wavelength is strongly different for unetched ZnO:Al, Asahi Type-U SnO_2:F, and texture-etched ZnO:Al, with larger features resulting in a higher PSD for wavelengths longer than 1 μm [237]. In general, the PSD rises with increasing spatial wavelength and saturates around the correlation length. The angular distribution of light scattering by a TCO is controlled by its PSD function.

Measurements of the diffuse transmittance as a function of angle (see Section 17.7.2.2) can also easily distinguish different TCOs. For example, the angle-resolved scattering data shown in Figure 17.17 shows a peak in the scattered power $S(\theta)$ at about $\theta = 40^\circ$ for commercial tin oxide and a much higher peak at about $\theta = 20^\circ$ for textured zinc oxide produced by hollow cathode

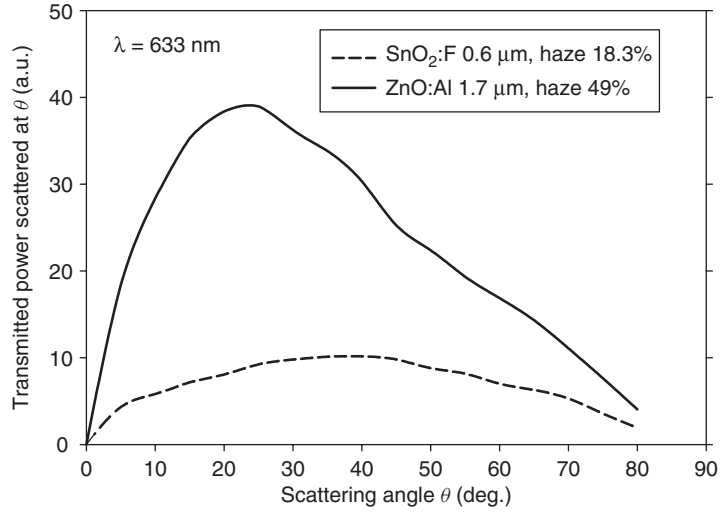


Figure 17.17 Angle-resolved scattering for two textured TCOs: APCVD SnO_2:F with a white-light haze of 18.3% and HC ZnO:Al with a white-light haze of 49%

sputtering [229]. While the haze is an indicator, but not a reliable predictor, of the short-circuit current density J_{sc} [233, 235], it appears that the J_{sc} of microcrystalline solar cells correlates well with the amount of scattering into large angles [237].

The quantification of surface morphology in other engineering disciplines, such as load-bearing analysis, has resulted in the definition of other distinct and useful parameters. For example, the features of a textured surface can be described by the so-called Birmingham 14 parameters [238]. These parameters can be divided into three categories: *amplitude parameters* (e.g. RMS roughness, surface skew, and kurtosis), which describe the distribution of surface height about a designated plane; *spatial parameters* (e.g. autocorrelation length, density of summits, and texture aspect ratio), which describe the spatial height distribution within the sample area, and *hybrid parameters* (e.g. mean summit curvature, interfacial area ratio), which describe the combined effects of amplitude and spatial attributes on topographical parameters. Surface skew S_{sk} is a measure of the asymmetry of the height distribution, with $S_{sk} < 0$ representing a surface with rounded peaks and sharp valleys, $S_{sk} = 0$ pertaining to a symmetrical Gaussian height distribution, and $S_{sk} > 0$ representing a surface with sharp peaks and rounded valleys. Values of S_{sk} have been reported for textured SnO_2:F and ZnO:Al [229]. The distribution of the local surface tilt angle has also been studied for various rough substrates [239]. More work is needed to identify the most relevant surface texture parameters for textured TCOs that perform well in solar cell applications, and to correlate these parameters with light scattering ability and solar cell performance.

17.6.6 Textured TCO Optimization

The effect of a textured TCO on a-Si or a-Si/μc-Si solar cell performance has been studied both experimentally and theoretically [151, 240–243]. Several inter-dependent factors need to be considered for TCO optimization, including optical, electrical, and morphological. A low free carrier absorption (FCA) is important, especially in the case of nc-Si:H solar cells. This goal can be realized by reducing the dopant concentration while maximizing the carrier mobility to secure adequate conductivity. Increasing TCO haze results in larger parasitic absorption in the TCO because of the

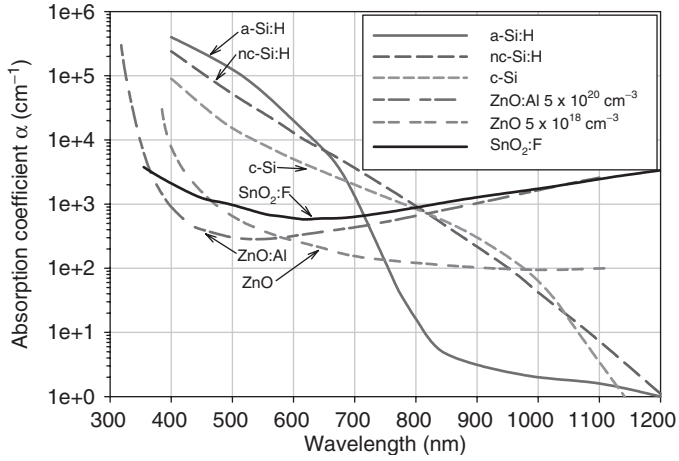


Figure 17.18 Spectral dependence of the optical absorption coefficient of amorphous silicon, nanocrystalline silicon, crystalline silicon, doped ZnO:Al, and undoped ZnO (data from reference [245]) and of commercial SnO_2:F (data from reference [73])

multiple light passes associated with light trapping. Figure 17.18 shows the optical absorption coefficients $\alpha(\lambda)$ for materials relevant to thin film silicon solar cells. The rise in $\alpha(\lambda)$ for $\lambda > 550$ nm for ZnO:Al relative to undoped ZnO is due to free-carrier absorption (see Section 17.4.2). This data surprisingly shows that, for $\lambda > 900$ nm, the optical absorption per light pass is probably greater in a lightly doped ZnO:Al TCO than in the nc-Si:H itself [244]. This confirms the need to limit FCA. An optimal carrier concentration in the range from $2.0 \times 10^{20}\,\text{cm}^{-3}$ to $2.5 \times 10^{20}\,\text{cm}^{-3}$ was calculated for a-Si/nc-Si solar modules taking into account parasitic absorption and joule losses (see Section 17.4.3) [244]. In the case of a-Si/nc-Si solar cells grown on LPCVD ZnO:B, the trend has been to reduce the carrier concentration in the ZnO:B to 1×10^{20} cm^{-3} or even lower (see Table 17.9). A useful consequence is an improvement in mobility. To achieve the desired sheet resistance of about $10\,\Omega/\square$, the film thickness is increased to $2-3\,\mu\text{m}$ [86].

Considering now the morphological aspects of a textured TCO, we first note that the ratio of the texture feature size to the light wavelength in the material is a critical factor influencing its light scattering ability. For example, LPCVD ZnO is optimized differently for nc-Si cells than for a-Si cells, since a larger feature size is required. The light scattering capacity of LPCVD ZnO directly depends on the pyramidal grain size and, as we saw in Section 17.6.3, the grain size depends on film thickness [85]. Furthermore, a smaller doping ratio (B_2H_6/DEZ) also leads to a higher haze [246]. However, as-grown large-grain LPCVD ZnO, although securing a high photocurrent, yields poor *FF* and V_{oc} values for nc-Si:H solar cells [247]. This effect has been traced to cracks and defects in the silicon emanating from the base of V-shaped valleys in the ZnO [248]. The cracks were modeled by a parallel diode that gives rise to an additional dark current. An effective solution to this problem is surface modification of LPCVD ZnO by plasma etching. A low-pressure RF discharge in gases such as Ar, O_2 or CO_2 can be used for this purpose (with Ar being preferred) and a treatment time of up to 80 minutes. The plasma treatment transforms the surface morphology from one having V-shaped valleys to U-shaped depressions that do not give rise to cracks in the nc-Si. The treatment also removes smaller-sized pyramids and asperities [249, 250].

In the case of a textured TCO obtained by post-deposition etching of sputtered ZnO:Al it has been shown that deposition pressure and substrate temperature T_s are important variables. For RF sputtering from a ceramic target three distinct outcomes from post etching are possible [251]. At low

Table 17.10 Summary of preparation conditions and film properties for an optimal textured ZnO:Al front TCO prepared by RF sputtering and post-deposition etching. Data compiled from Berginski M *et al*., *J. Appl. Phys*. **101**, 074903 (2007) [253]

Sputter parameters	
Ceramic target Al_2O_3 concentration	0.5–1.0 wt.%
Pressure	0.3 Pa
Substrate temperature	360–410 °C
Film (TCO) properties	
Initial thickness	800 nm (RMS 15 nm)
Carrier concentration	3×10^{20} cm^{-3}
Mobility	40 cm^2/V s
Resistivity	5.2×10^{-4} Ω cm
Thickness after etching	650 nm
Sheet resistance	8.0 Ω/□
Topography	Type 2; crater-like, uniform;
Crater size	lateral dimensions 1–2 μm; depth 200–400 nm
RMS roughness	>125 nm
Mean crater opening angle and FWHM	120–135°; 25–45°
Haze	30–40% at $\lambda = 1$ μm
Cell properties (μc-Si:H , 1.9 μm)	
Best J_{sc}	26.8 mA/cm^2
Peak $J_{ph}(\lambda)$ enhancement factor f	$f \approx 14$ at $\lambda = 900$–950 nm

pressure and high T_s an extremely compact film results that upon etching yields only a few scattered large craters. To obtain a uniform distribution of craters a somewhat higher pressure and lower T_s is required. At high pressure and low T_s the film is porous and etches uniformly without producing craters. The relationship of these results to the Thornton model of film growth and structure [252] has been discussed. A further important variable is the target alumina concentration, *TAC*, and in later work a 2D map of etching results versus *TAC* and T_s was presented [253]. As a potential low-cost, high-rate process, ZnO:Al films suitable for etching have also been produced by reactive sputtering using Zn:Al (Al 0.5%) metal targets with appropriate choice of oxygen partial pressure [254]. Table 17.10 presents a summary of the preparation and properties of the textured ZnO:Al that gives optimal nc-Si solar cell performance. Despite the intensive study of post-deposition etching, there does not yet seem to be a mechanistic explanation of the observed morphologies.

It has also been shown that, by post-etching vacuum annealing of ZnO:Al films at about 550 °C for one hour, their carrier concentration can be decreased, and hence the free-carrier absorption decreased, without loss of mobility. This process may be used, if necessary, to adjust the carrier concentration to its optimal range of $2–2.5 \times 10^{20}$ cm^{-3} [255]. A reduction in carrier concentration from 5×10^{20} cm^{-3} to 2×10^{20} cm^{-3} yields a ΔJ_{sc} of about +1.5 mA/cm^2 for a 0.9 μm single-junction nc-Si:H cell [244].

17.6.7 Application of Textured TCO to Solar Cells

The potential superiority of ZnO:Al as a TCO for a-Si:H devices has been demonstrated by codeposition of single-junction a-Si:H solar cells onto textured ZnO:Al and commercial SnO_2:F substrates [184]. The ZnO:Al was prepared by reactive-environment hollow cathode sputtering (RE-HCS, see Section 17.6.4). The short-circuit current density, J_{sc}, of solar cells on ZnO:Al was found

Table 17.11 Performance of glass/TCO/a-Si(*p-i-n*)/Al cells as a function of TCO type and haze

TCO	Haze (%)	V_{oc} (mV)	J_{sc} (mA/cm^2)	*FF* (%)	*Eff* (%)
SnO_2:F	16.4	843	12.14	69.3	7.09
ZnO:Al	11.1	849	11.38	70.0	6.76
ZnO:Al	16.0	844	12.78	70.7	7.63
ZnO:Al	55.0	782	12.31	66.0	6.35

to be up to 5.3% higher than on SnO_2:F. The higher J_{sc} results from the better optical transmission and light scattering of the ZnO:Al. The data is shown in Table 17.11. It may be remarked that some types of ZnO:Al (e.g. that produced by RF magnetron sputtering) need a special surface treatment in order to secure a suitable ZnO/p-layer interface and competitive fill factor; the RE-HCS ZnO:Al was untreated. The last three entries of Table 17.11 show the effect of haze on cell performance. Mirroring the results earlier described for tin oxide, optimal cell performance on ZnO:Al is obtained at intermediate haze values [229].

For single-junction nc-Si:H cells grown on LPCVD ZnO:B, the plasma surface treatment mentioned previously results in dramatic improvements in V_{oc} and *FF* so that the cell efficiency is more than doubled [247]. Furthermore, the carrier collection is improved so that reverse bias is no longer needed to achieve the full photocurrent. A best single-junction cell efficiency of 9.9% was achieved in this study.

The increase in the short-circuit current density, J_{sc}, of thin Si:H cells arising from the improved long-wavelength *QE* resulting from the use of substrates with textured TCOs is significant for a-Si:H and very large for nc-Si:H. Figure 17.19 (a) shows the *QE* data for single-junction a-Si:H, and Figure 17.19 (b) shows data for single-junction nc-Si:H solar cells on different smooth and textured sputtered ZnO:Al films. A smooth ZnO:Al front contact yields $J_{sc} = 15.6$ mA/cm^2, while the textured ZnO:Al films yield 23.0–26.8 mA/cm^2 [256]. The lower doping concentration further improves the QE at long wavelengths because of reduction of free carrier absorption. For a-Si:H/nc-Si:H tandem cells deposited on LPCVD ZnO:B, a total J_{sc} of 27.7 mA/cm^2 has been achieved [164]. The ZnO:B, with initial RMS roughness of 180 nm, was plasma treated to yield a U-shaped morphology having a final roughness of 120 nm. Preliminary experiments with triple-junction a-Si:H/nc-Si:H/nc-Si:H cells on sputtered and etched ZnO:Al yielded a cell efficiency of 12.1% with $J_{sc} = 9.3$ mA/cm^2 (total $J_{sc} = 27.5$ mA/cm^2) [244].

The optimization of a-Si:H/nc-SiH cells on textured TCOs is a complex process. Despite the advantages of a minimum absorbance of 2% for ZnO compared with 4% for SnO_2, and a higher haze for the ZnO, the cell recipes (and especially that of the p_1 layer) need to be tailored to the TCO, and in the final analysis the performance differential appears to be small, with 12.3 and 12.0% cells having been achieved by the same group on ZnO and SnO_2 substrates, respectively [257].

The application of textured TCOs to thin film solar cells made from polycrystalline CdTe and CIGS has also been studied, either for thickness reduction or performance improvement. In the case of very thin (0.6 μm) CdTe solar cells, their long-wavelength *QE* can be observed to depend on the haze of the TCO. A 5% increase in *QE* was found for deposition on textured SnO_2:F with 37% haze relative to that on smooth ITO [258]. For substrate-type CIGS solar cells, the ZnO:Al is deposited after the junction has been formed. Consequently, there is a process limitation regarding the preparation of the textured ZnO since its deposition temperature must be less than 200 °C to avoid damaging the junction. Nevertheless, an 8% increase in J_{sc} was achieved [229]. Figure 17.20 shows SEM images of the surface of CIGS solar cells with nontextured and (as-deposited) textured

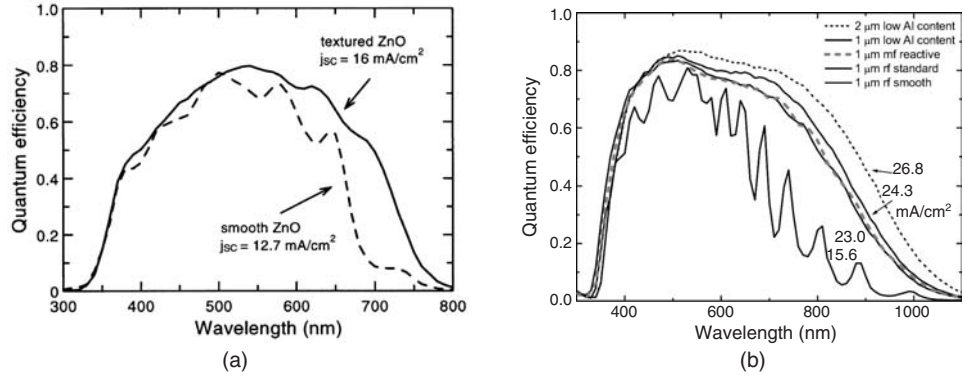


Figure 17.19 Quantum efficiency curves of (a) a-Si:H and (b) μc-Si:H *p-i-n* solar cells on smooth and texture-etched RF-sputtered ZnO:Al front contacts: (a) a-Si:H cell with 0.36 μm *i*-layer on ZnO:Al from 2 wt.% alumina ceramic ZnO target. Reproduced from Kluth O *et al*., *Thin Solid Films* **351**, 247 (1999), with permission from Elsevier [223]; (b) μc-Si:H cell with *i*-layer thickness of 1.0 μm on ZnO:Al from standard (1 wt.%) and low (0.5 wt.%) alumina concentrations in the ceramic ZnO target, also on a surface-textured, reactive m.f.-sputtered ZnO film, and a μc-Si:H cell with *i*-layer thickness of 2.0 μm on low Al concentration ZnO. Reproduced from Rech B *et al*., *Thin Solid Films* **511–512**, 548 (2006), with permission from Elsevier [256]

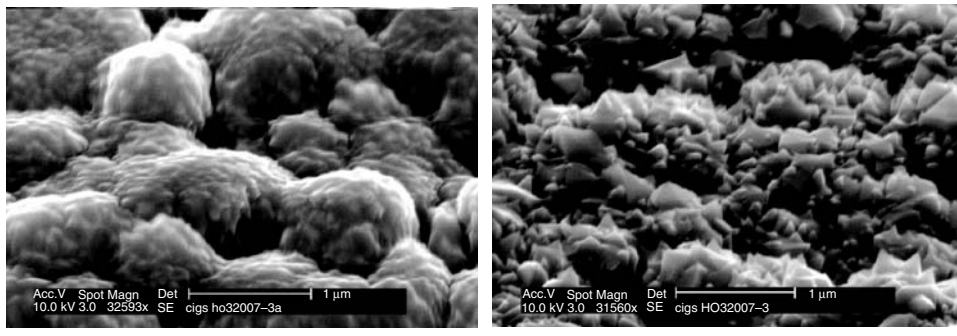


Figure 17.20 Surface morphology of CIGS solar cells coated with: (a) nontextured ZnO:Al deposited by RF sputtering from a ceramic target; and (b) textured ZnO:Al deposited by hollow cathode sputtering. The short-circuit photocurrent is increased by 8% through use of the textured ZnO:Al coating. Reproduced from Guo S *et al*., *Proc. SPIE* **6651**, 66510B, (2007), with permission from SPIE [229]

ZnO:Al. The large feature size in Figures 17.20 (a) and (b) is due to the CIGS, while the smaller feature size clearly visible in Figure 17.20 (b) is due to the textured ZnO:Al.

17.7 MEASUREMENTS AND CHARACTERIZATION METHODS

The methods used to characterize transparent conductive oxides are similar to the approaches used to study semiconductor properties. For PV application of TCO films, the electrical and

optical properties are of primary interest. Properties such as crystalline structure, morphology, composition, adhesion and other physical and chemical properties are also directly related to the performance of both the TCO and the device of which it is a part. We give a summary of these characterization methods.

17.7.1 Electrical Characterization

The sheet resistance of a thin film TCO can be quickly determined using a four-point probe [259]. The probe consists of four spring-loaded metal probes mounted in a straight line and equally spaced. The probe tips are often made of tungsten carbide and have a radius of about 50 μm. In operation, a constant-current source supplies a current I that is passed through the film via the two outer probes and the voltage V picked up by the two inner probes is measured. Provided the film thickness is much less than the spacing between the probe tips (as is the case for TCO films that are of the order of 1 μm in thickness) and the dimensions of the sample are much greater than the probe spacing, the sheet resistance is given by:

$$R_{sh} = \frac{\pi}{\ln(2)} \frac{V}{I} = 4.532 \frac{V}{I} \tag{17.41}$$

The units of sheet resistance are ohms per square (Ω/□). The use of separate current and voltage probes eliminates errors due to contact resistance effects. The resistivity ρ of the thin film may be calculated from the equation $R_{sh} = \rho/t$. The film thickness t can be determined using a stylus profilometer or optically (see Section 17.7.2). In a manufacturing environment, a contactless measurement of sheet resistance may be preferred, especially as the sheet resistance of a TCO could be checked after deposition of additional semiconductor layers. The measurement principle makes use of RF-induced eddy currents.

The electrical conductivity may be decomposed into the parameters carrier concentration, carrier mobility, and carrier type by making use of the Hall effect. In the usual arrangement, a magnetic field $\boldsymbol{B}$ is applied perpendicular to the plane of the film while a steady current flows in the film. The magnetic field exerts a force (the Lorentz force) on the carriers and displaces them in a direction perpendicular to both the field and the current. The result is an electric field $\boldsymbol{E}$. In equilibrium the net force $\boldsymbol{F} = e\boldsymbol{E} + e\boldsymbol{v} \times \boldsymbol{B}$ is zero [134]. Since $\boldsymbol{j} = ne\boldsymbol{v}$ we have $\boldsymbol{E} = -\boldsymbol{j} \times \boldsymbol{B}/(ne) = -R_H(\boldsymbol{j} \times \boldsymbol{B})$ where R_H is the Hall coefficient:

$$R_H = \frac{E_z}{j_x B_y} = \frac{1}{ne} \tag{17.42}$$

For a rectangular strip of film of thickness t and width w carrying a current $I_x = j_x wt$ and the magnetic field normal to the plane of the film, the transverse Hall voltage is $V_H = E_z w$ and so

$$V_H = \frac{I_x B_y}{net} \text{ or equivalently, } n_s = \frac{I_x B_y}{eV_H} \tag{17.43}$$

where $n_s = nt$ is the sheet density of carriers (m^{-2}). Note that the MKS units for the above equations are, for magnetic induction B the Tesla ($= 1\,\mathrm{V\,s/m^2}$ or 10^4 Gauss), for carrier density $n\,\mathrm{m}^{-3}$, for $R_H\,\mathrm{m}^3/\mathrm{C}$. Thus, from Hall effect measurements we may learn the sign of the carrier charge e (from R_H) and the surface or sheet density of the carriers n_s from the Hall voltage V_H.

A suitable sample configuration for Hall effect measurements is the classic cloverleaf geometry or simply a square of the thin film TCO with four very small ohmic contacts placed on the periphery at the corners of the square (van der Pauw [260]). For definiteness we denote these contacts 1, 2, 3, 4 in sequence around the sample (see Figure 17.21).

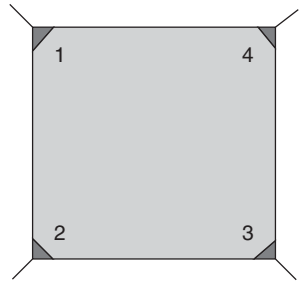


Figure 17.21 Square sample of TCO film on glass prepared with four contacts at the corners for Hall effect measurements

A current I_{13} is applied to opposing contacts 1 and 3 and the voltage $V_{24}(B)$ is measured between contacts 2 and 4 using a high-impedance voltmeter. Misalignment of the contacts will give rise to a voltage $V_{24}(0)$ even in the absence of a magnetic field and this must be subtracted from the measured value of $V_{24}(B)$. Several other spurious voltages can arise from thermoelectric or thermomagnetic effects. The recommended protocol involving the systematic reversal of the current and magnetic field and the interchange of contacts should be followed to ensure reliable data [259, 261]. From the current and Hall voltage parameters and the value of the magnetic field (which can itself be measured using a commercial Hall effect magnetometer), n_s can be calculated using Equation (17.43). In addition, the sheet resistance R_{sh} of the same TCO sample can be determined using the van der Pauw geometry by applying a current I_{12} to adjacent contacts 1 and 2 and measuring the voltage V_{43}. The sheet resistance is given by:

$$R_{sh} = \frac{\pi}{2\ln(2)}\left(\frac{V_{43}}{I_{12}} + \frac{V_{14}}{I_{23}}\right) f \tag{17.44}$$

where f is a factor that deviates slightly from unity as the resistance ratio $R_{43,12}/R_{14,23}$ deviates from unity [118]. Finally, the Hall mobility μ_H can be calculated from:

$$\mu_H = \frac{1}{en_s R_{sh}} \tag{17.45}$$

Determination of n_s, R_{sh} and μ_H can be completed without knowing the film thickness. After measuring the film thickness t, the sample resistivity $\rho = R_{sh}t$ and the carrier concentration $n = n_s/t$ can be calculated. For n-type TCO films, the carrier concentration and mobility are typically in the $1-10 \times 10^{20}$ cm^{-3} and $10-100\,\text{cm}^2\,\text{V}^{-1}\,\text{s}^{-1}$ ranges.

Sheet resistance measurements on very thin films can present some difficulties of interpretation. First, the growth of very thin layers may be sensitive to nucleation effects and therefore reproducibility should be checked. It should also be checked whether atmospheric exposure plays a role (e.g. in donor deactivation) and whether measurements should be made under nitrogen. Moreover, the results may be affected by the existence of space-charge regions near the interfaces. Even in the normal range of thicknesses, it should be borne in mind that both the mobility and carrier concentration generally increase away from the substrate interface. This effect can persist until a thickness of 0.5–2.0 μm is reached. In contrast, the transport properties of ITO tend to be preserved even in relatively thin films (down to 100 nm).

The temperature dependence of the dc electrical conductivity of TCOs may be studied using a vacuum dewar or cryostat to prevent moisture condensation below room temperature. A suitable sample may be prepared in gap cell configuration by sputtering ohmic contacts (e.g. Mo) on top of the TCO. A four-wire measurement configuration is desirable. The resistivity of ZnO:Al

deposited by RF sputtering from a ceramic target (1.3 wt% Al_2O_3) has been measured from 100 to 380 K and a temperature coefficient of resistance (TCR) of $+2.9 \times 10^{-4}$ K^{-1} around room temperature was determined [45]. The TCR is positive (metallic behavior) consistent with the degenerate doping of the sample.

The carrier type in TCOs can be determined by the sign of the thermopower. The Seebeck coefficient S, with units $\mu V\ K^{-1}$, can be measured using a sample in gap cell configuration and establishing a temperature difference in the sample from one contact to the other. In steady state the diffusion of carriers from the hot side to the cold side (resulting in an electric field and hence an emf) is balanced by an opposite drift current. The absolute measurement of the thermopower ideally requires correction for the (small) thermoelectric effect due to the contacting metals [119]. Measurement of S can shed light on the carrier scattering mechanism (see Section 17.4.1.2).

17.7.2 Optical Characterization

17.7.2.1 Specular films

The optical properties of TCO films can be determined by measurement of the transmittance $T(\lambda)$ and reflectance $R(\lambda)$ using a UV-Vis-NIR spectrophotometer. A common wavelength range is 250–2500 nm using deuterium and tungsten–halogen light sources. The absorption can be calculated from $A = 1 - T - R$. Expressions for $T(\lambda)$ and $R(\lambda)$ for a planar film on a transparent substrate have been given in Section 17.4.2. Methods for determining the thickness t and the optical constants $n(\lambda)$ and $\alpha(\lambda)$ $(= 4\pi k(\lambda)/\lambda)$ of the film from the measured $T(\lambda)$ are given in [262]. Similar information may also be obtained from spectral reflectometry [263]. In this technique, light can be delivered by a central optical fiber and collected after normal reflection by surrounding fibers. Suitable models for $n(\lambda)$ and $k(\lambda)$ must be chosen, and software is then used to find the best fit values of t, n, and k taking into account the interconnection of n and k imposed by the Kramers–Kronig relations that connect dispersion and absorption [133]. Modern spectral reflectometry instruments and software can also handle multilayer films. Reflectance data may be obtained at even longer wavelengths (e.g. up to 25 μm) using an FTIR spectrometer.

A more powerful and robust tool for the optical characterization of all types of thin films and multilayer structures is ellipsometry [264, 265]. Ellipsometry involves measurement of the change in the polarization state of incident light after non-normal reflection. The fundamental equation of ellipsometry is

$$\rho = \frac{r_p}{r_s} = \tan(\psi)\exp(i\Delta) \tag{17.46}$$

where r_p and r_s are the complex reflection coefficients for p- and s-polarized light, $\tan(\psi)$ is the amplitude ratio of the s and p coefficients and Δ is their differential phase change. For certain samples, e.g. ZnO where the c-axis does not coincide with the sample normal, more generalized equations are required. In spectroscopic ellipsometry (SE), the measurements of ψ and Δ are performed over a range of wavelengths. Again, a layered optical model must be entered in order to extract the material properties of interest [266]. For example, the dielectric functions $\varepsilon_r(E) + i\varepsilon_i(E)$ of ZnO:Ga and In_2O_3:Sn samples having a range of carrier concentrations were recently determined by ellipsometry for the structure TCO/SiO_2/c-Si [267]. The TCO was modeled as having a rough surface layer (thickness t_s) on a bulk layer (thickness t_b) with a dielectric function $\varepsilon(E) = \varepsilon_{TL}(E) + \varepsilon_D(E)$ based on the Tauc–Lorentz and Drude models. The surface layer was modeled using the Bruggeman effective-medium approximation. The calculated n and k spectra exhibited dispersion increasing with carrier concentration. A similar investigation, although different in detail, has been presented for ZnO:Al [268].

Optical absorptance spectra can also be measured using photothermal deflection spectroscopy (PDS) [269]. This sensitive technique is especially useful in the case of weak absorption.

17.7.2.2 Textured TCO film characterization

The total and diffuse transmittance and reflectance of textured films can be measured as a function of wavelength using a spectrophotometer fitted with an integrating sphere. The spectral haze $H(\lambda)$ is defined as the ratio $T_{diff}(\lambda)/T_{tot}(\lambda)$. For textured films, absorption figures uncritically calculated from $A = 1 - T - R$ are too large because light backscattered from the TCO/air interface is lost by light trapping in the glass substrate, the light being ultimately absorbed or lost through the sample edges. The measured transmittance also underestimates the amount of light that would be transmitted to a solar cell absorber because of the difference between a TCO/air interface and a TCO/solar cell absorber interface. In order to perform more realistic measurements, an index-matching fluid, for example, dichloromethane (CH_2Cl_2, index 1.74 at 586 nm), is generally used [270]. One method is to attach a thin (nonabsorbing) glass slip to the TCO surface with a very thin intervening layer of the index-matching fluid. This greatly reduces the scattering at the TCO surface.

A simple measurement of texture can be made using a haze meter. In this instrument, a white light source is used and a parallel beam is formed. The integrating sphere has an inlet port, an exit port, an off-axis detector, and is coated with a white diffuse reflector (e.g. barium sulphate). With a glass/TCO sample at the inlet port and a white reflector covering the exit port, the detector current I_{tot} is recorded (corresponding to the scattered plus specular light, i.e. total light, emanating from the sample). With the exit port open, the detector current I_{diff} is due only to the scattered (diffuse) light. The haze is $I_{diff}/I_{tot} \times 100\%$, or equivalently, $T_{diff}/T_{tot} \times 100\%$.

The spectral haze factor defined above gives the total scattered power at a given wavelength, but no information about the angular distribution of the scattered light. Angle-resolved light scattering measurements (ARS) can be performed to obtain the scattered power dependence on scattering angle [229, 237, 271]. The ARS analysis can be conducted in air using a laser beam incident on the glass/TCO sample and a detector (Si photodiode) mounted on a rotatable arm to measure the scattered light intensity (in transmission). The beam is chopped and the photocurrent is measured using a current preamplifier and lock-in amplifier. A typical set up is shown in Figure 17.22. The method is also described in reference [271]. Care must be taken to eliminate pick up of trapped light emanating from the edge of the substrate (e.g. by taping the edges). The

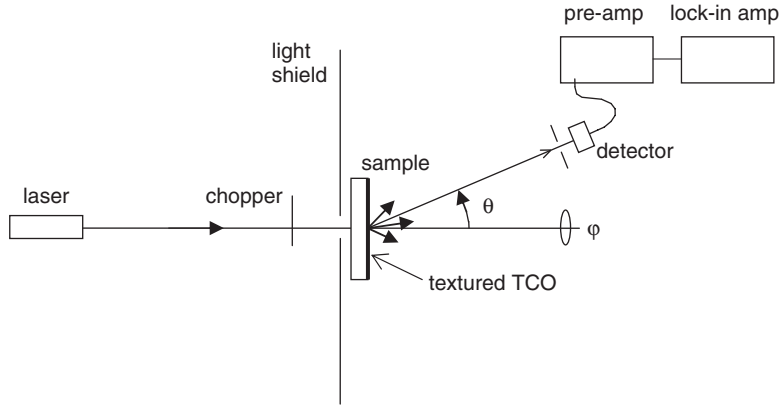


Figure 17.22 Typical set-up for measurement of angle-resolved light scattering

total power $S(\theta)$ scattered at a given scattering angle θ is obtained by integrating the measured intensity over the azimuth angle φ. This is equivalent to multiplying the photocurrent by $\sin(\theta)$. For measurement of scattered light in a reflective mode, the sample can be coated with 80–100 nm Ag. Wavelength-dependent data can be acquired using a white-light supercontinuum laser source and spectrometer. A full analysis of ARS data sets in terms of the statistical topographic properties of the TCO surfaces apparently remains to be completed.

17.7.3 Physical and Structural Characterization

The thickness of thin film TCO layers deposited on a substrate can be measured using stylus profilometry. The profilometer includes a stage that can be translated in a plane without vertical excursions. The substrate is set on the stage, a diamond-tipped stylus is lowered onto the film, and the substrate is translated under the stylus so that the stylus rides over the surface of the film. The stylus is coupled to a slug that moves between the two secondary coils of a linear variable differential transformer (LVDT). The coils are excited by high-frequency AC and their differential voltage is proportional to the deviation of the slug position from the null position, thereby enabling measurement of the height of the stylus. As the stylus rides over a step in the film (created by masking, etching, ablation, or scribing) the change in stylus height indicates the film thickness. It is desirable that the step be abrupt. If the step is formed by film removal it should be verified that the surface of the substrate remains intact. For textured films, the stylus will record the overall height of the film and not the average thickness since the radius of curvature of the stylus (perhaps 12.5 μm) is usually large compared with the surface feature size of the film. This can affect the calculation of carrier density and resistivity.

For TCO applications involving thin Si wafers or plastic substrates, film stress is an important issue. The total internal stress σ_f in a TCO film resulting from a particular deposition process can be determined by depositing a film onto very thin glass cover slips or 50–150 μm micro-sheet glass. The stress can be calculated from the resulting substrate curvature K (m^{-1}) using the Stoney equation

$$\sigma_f = -\frac{EKt_s^2}{6(1-\nu)t_f} \tag{17.47}$$

where E and υ are the Young's modulus and Poisson's ratio of the substrate, and t_s and t_f are the thickness of the substrate and film, respectively [272]. For ITO deposited by facing target sputtering a compressive stress of $0.2–1.0 \times 10^9\,\mathrm{N\,m^{-2}}$ can be achieved while $3 \times 10^9\,\mathrm{N\,m^{-2}}$ is typical for magnetron sputtering [273].

The surface topography of TCO films can be examined with high depth of field by scanning electron microscopy (SEM). Modern field-emission electron sources now give this technique adequate resolution for such investigations. The reader is reminded that the incident beam penetrates a short distance into the sample (depending on beam voltage) and results in the emission of low energy secondary electrons that are attracted by a detector; the SEM image therefore appears as though illuminated from the detector and viewed from the beam direction [274]. Information about film nucleation and growth can be inferred by examination of fracture cross-sections. A columnar structure is often apparent. Alternatively, a dual FIB-SEM instrument can be employed to generate a cross-section suitable for SEM imaging by focused ion beam (FIB) milling using, for example, a Ga^+ ion beam. The recently introduced He ion microscope is similar to a SEM, but uses a He ion beam instead of an electron beam to achieve sub-nanometer-scale resolution. Transmission electron microscopy (TEM) of cross-sectional samples can also be used to detect nucleation layers and estimate column growth angles. High-resolution TEM (HRTEM) can be used to image grain boundaries. With the beam focused on an individual grain, a selected area diffraction (SED)

pattern can be obtained. This can be used to confirm the crystal structure, e.g. tetragonal for SnO_2. Electron diffraction has also been used to detect other phases at the interface and grain boundaries of ZnO:Al [275]. More quantitative topographic information can be acquired by atomic force microscopy (AFM). Built-in software packages can also provide results for RMS roughness, height distribution, correlation length, and 2D isotropic PSD (power spectral density, see Section 17.6.5).

The crystallinity, preferred orientation and crystallite size of a TCO layer on a noncrystalline substrate can be determined by X-ray diffraction (XRD). In this technique, X-ray interference peaks occur when the Bragg condition $2d\sin(\theta) = \lambda$ is satisfied for reflection from a set of crystal planes, where d is the interplanar spacing, θ is the glancing angle of incidence and reflection of the beam and λ is the wavelength of the X-rays. Measurements using a diffractometer are conducted in the usual θ–2θ configuration [276]. The interplanar spacing and Miller indices hkl are related, for the cubic lattice, by $1/d^2 = (h^2 + k^2 + l^2)/a^2$ where a is the lattice constant or $1/d^2 = 4(h^2 + hk + k^2)/(3a^2) + l^2/c^2$ for the hexagonal lattice. A shift in the lattice parameter is indicative of a uniform strain in the film. A quick indication can be obtained that crystallites are not distributed with random orientation if the ratios of the integrated intensities for different (hkl) planes for a thick film do not agree with the tabulated values [277] for a powder specimen. For example, the value of $I(400)/I(222)$ for magnetron-sputtered ITO was found to be 2.0, indicating a <100> preferred orientation, rather than the value of 0.33 pertaining to a random orientation [27]. Phase identification is particularly useful for multi-component TCOs, e.g. for the ZnO-In_2O_3 system or for ternary TCOs such as $CuInO_2$. While the presence of impurity phases can usually be detected by XRD, metallic inclusions that cause film darkening are usually not detectable. For a homogeneous sample, the crystallite size can be inferred from the broadening of the X-ray peaks given by the Scherrer formula:

$$d_{hkl} = \frac{0.9\lambda}{FWHM\cos(\theta)} \tag{17.48}$$

Here, the crystallite diameter d, strictly in a direction normal to the reflecting planes (hkl), can be calculated from the full width at half maximum (FWHM) of the X-ray peak (corrected for instrumental broadening) assuming that the broadening is due to a crystallite size effect and not stress. It should be appreciated that the penetration depth of X-rays is considerable and that a layer of small crystallites at the base of the film can contribute to peak broadening. Other XRD analysis techniques (pole figures, rocking curves) may be used to study crystallite orientation and quality in greater detail.

Raman spectroscopy, which measures the energy shift between incoming and outgoing photons due to interaction with phonon modes, can sometimes be used to identify crystalline phases in thin film materials. As an example, the Raman lines for anatase TiO_2 are at 144, 198, 399, 516, and 640 cm^{-1}, while those for rutile TiO_2 are at 143, 240, 447 and 612 cm^{-1} [278].

The crystallization temperature of an amorphous TCO film may be determined by differential thermal analysis (DTA). An exothermic peak can be found upon crystallization.

17.7.4 Chemical and Surface Characterization

Several analytical techniques are available for investigating the composition of TCO films. These include X-ray photoelectron spectroscopy (XPS), Auger electron spectroscopy (AES), and energy-dispersive X-ray analysis (EDX). Precise determination of stoichiometry is, however, difficult. Of particular relevance to TCO characterization is measurement of the doping element concentration and binding state. The dopant/host atomic concentration ratio can be conveniently determined (in a non-vacuum instrument) by inductively coupled plasma atomic emission spectroscopy (ICP-AES). In this method the film is dissolved in nitric acid which is nebulized and injected into an Ar plasma

where the atoms are excited and subsequently emit photons; a triple-grating monochromator is used to identify the wavelengths and hence the elements that are present. A standard solution is used to calibrate the instrument. Both the film composition and dopant valence state can be identified by XPS. As examples, the F/Sn ratio can be determined for SnO_2:F and the Al^{3+} state can be identified in ZnO:Al. The concentration of metallic elements, e.g. In and Sn in ITO, can also be determined by X-ray fluorescence (XRF). Impurities in the TCO can be measured by secondary ion mass spectrometry (SIMS), XPS, or Fourier transform infrared spectroscopy (FTIR). Thus impurities such as C or H in ZnO:Al can be determined using SIMS. While hydrogen can be detected by SIMS, it is hard to quantify. A better technique for this purpose is hydrogen forward scattering spectrometry (HFS). Hydrogen content can also be inferred from thermal desorption spectroscopy (TDS). The concentration of Al in ZnO:Al can also be determined by nuclear reaction analysis (NRA) using the $^{27}Al(p,\gamma)^{28}Si$ resonance reaction. The areal density N_A (atoms/cm^2) of heavy elements can be determined by Rutherford backscattering spectroscopy (RBS). Typically, the incident beam consists of 2.0 MeV He^+ ions while the backscattered ions are detected by a Si surface barrier detector. RBS spectra can be simulated by a standard program called RUMP in order to fit the sample spectrum. A highly accurate measurement of N_A can be obtained via RBS analysis by increasing the count statistics, and the result is independent of surface roughness or void fraction.

In several types of devices that utilize TCOs (including some solar cells and organic electroluminescent devices) the work function of the TCO can modify the interfacial band diagram. The work function of the film, usually defined as the energy required to remove an electron from the Fermi level to infinity ($E_{vac} - E_F$), can be measured by ultraviolet photoemission spectroscopy (UPS). The work function depends on the position of E_F relative to the band edges (and hence on doping concentration) and a possible surface dipole [279]. The surface dipole can be modified by oxygen or cation termination or by electropositive or electronegative adsorbates. For ITO, it has been reported that the work function is 4.3 eV for an Ar^+ sputter-cleaned surface (resulting in lowered surface oxygen content), 4.5 eV after cleaning with organic solvents (leaving up to 20 at.% carbon on the surface), and 4.75 eV after UV-ozone treatment (removes carbon contamination) [280].

A specialized, although important, laboratory technique has been devised to rapidly evaluate the susceptibility of a TCO layer to delaminate from its glass substrate [281]. The technique applies an electrochemical stress to the TCO/glass interface. It was devised principally to evaluate the durability of SnO_2:F on soda-lime glass. To perform the test, an indium pad is first soldered to the glass side of a TCO/glass sample and a copper wire is soldered to the TCO at the edge of the sample. The sample is placed glass side down on a pre-heated metal hot plate at about 185 °C (see Figure 17.23). As the sample heats, the indium melts and electrically connects the underside of the glass to the metal plate. A voltage of 100 V DC is applied for 15 minutes between the metal plate (+ ve) and the copper wire (−ve), after which the sample is removed from the power supply and

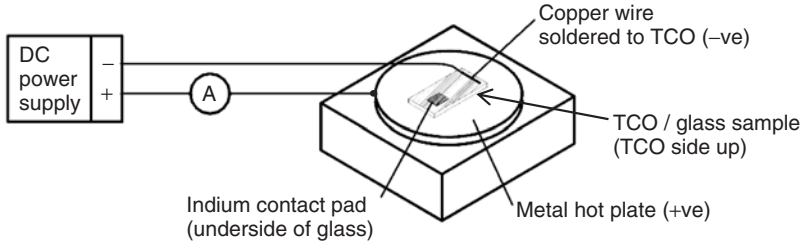


Figure 17.23 Set-up for TCO/glass delamination test showing indium contact to the glass and connections to the electrically-isolated DC power supply. Adapted from Jansen K, Delahoy A, *Thin Solid Films* **423/2**, 153 (2003) [281]

hot plate and allowed to cool in room air (typically 40–60% relative humidity). The function of the bias is to establish an electric field to drive Na^+ ions from the glass towards the portion of the glass/TCO interface that is above the indium pad. To a greater or lesser degree the TCO is darkened as the Na^+ ions are reduced to Na metal at the cathode. In the presence of moisture, subsequent chemical reactions at the interface will cause the TCO to crack and delaminate if it is prone to do so. If the TCO does not visibly delaminate after 10 minutes or so it may be scratched with a razor blade (or exposed to greater humidity) to further stress the sample. A TCO that delaminates under the conditions of this test is not suitable for use in PV products.

17.8 TCO STABILITY

The durability of photovoltaic modules is of prime importance. In this section we focus on the intrinsic stability of the thin film TCO component of a module. Other aspects of module reliability including absorber and junction degradation, and module encapsulation are dealt with in individual chapters for each cell technology.

The pyrolytic tin oxide coating (SnO_2:F) used in TCO glass is widely acknowledged to be extremely durable and chemically resistant. One highly specific weak point of tin oxide, when used as a superstrate for a-Si:H or nc-Si:H module manufacturing, is its susceptibility to reduction by atomic hydrogen [198]. Another specific area of concern involves delamination of the coating. In general, pyrolytic tin oxide had proven over the years to be extremely adherent, and it was a surprise when delamination was documented to occur in some fielded modules and, at one point in time, after a-Si:H deposition if the plates were not promptly processed into modules. It was only after development of a convenient laboratory technique for quickly inducing such delamination [281] that these specific failures were definitively linked to an electrochemical corrosion mechanism. This mechanism involves field-driven drift of sodium ions towards the glass/TCO interface and their subsequent accumulation and chemical reaction in the presence of moisture [281–284]. (A related, but different, electrochemical corrosion mechanism had been observed in earlier work [285].) The rash of product failures that prompted development of this delamination technique resulted from a process drift in the chemistry of the silicon oxycarbide barrier and color-suppression layer that is deposited on the glass prior to the tin oxide coating. These changes in the undercoat layer rendered associated batches of TCO susceptible to the delamination mechanism. The "hotplate test" for delamination has been universally adopted by TCO manufacturers as part of their quality control, and this type of delamination problem does not appear to have recurred. The test is described in detail in Section 17.7.4. Regardless of the TCO adhesive strength, it is recommended that module strings be grounded at the negative pole to suppress positive ion migration from the glass to the TCO. The normal durability of thin film modules fabricated on tin oxide is reflected by data extending over seven years accumulated by NREL that shows a degradation rate of about 1%/year [286].

High-temperature processing during the fabrication of certain types of solar cell can degrade TCO properties. For example, the exposure of indium tin oxide (ITO) to temperatures above 300 °C in air can increase its resistivity by a factor of three. In the fabrication of dye-sensitized solar cells (DSSC), a paste of TiO_2 is sintered on the TCO at 400–600 °C. While SnO_2:F (FTO) is stable in air at these temperatures, the use of ITO in DSSC is generally precluded. By overcoating a 700 nm ITO film with 100 nm FTO a bilayer FTO/ITO TCO was formed that was found to be thermally stable at 600 °C [287]. The films were prepared by spray pyrolysis. The FTO/ITO bilayer possessed a sheet resistance of 1–2 Ω/□ and a resistivity of 1.4×10^{-4} Ω cm (i.e. properties considerably superior to those of FTO alone) and yielded DSSC efficiencies higher than those obtained with FTO.

Zinc oxide (ZnO) is commonly used as the top TCO for CIGS modules and also as the front and rear contacts for some thin Si modules. Doped ZnO appears relatively stable when heated in Ar to temperatures up to 400 °C, although a direct comparison of sputtered ZnO:Al and APCVD

ZnO:F would be of interest. However, the susceptibility of ZnO to degradation via damp heat conditions presents both a barrier and uncertainty regarding the wider introduction of ZnO into commercial thin film modules. There is also concern that acetic acid, which can be produced by hydrolysis of the vinyl acetate units in EVA encapsulant [288], would rapidly attack ZnO in thin film Si modules should it gain access to the ZnO. Indeed, corrosion of Al in contact with EVA has been observed in damp heat testing of glass/EVA/Al-glass structures [288, 289].

CIGS modules have in the past encountered difficulties in passing the IEC 1215 (now replaced by the IEC 61646) damp heat test. This standard required modules to degrade less than 5% after 1000 hours in an environment characterized by 85 °C and 85% relative humidity ("85/85"). (Such a requirement is not contained in the latest (2008) version of the standard.) Accelerated tests with unencapsulated CIGS modules subjected to 500 hours of damp heat (85/85) have shown, as well as cell- and Mo-related effects, two TCO-related effects [290]. First, the conductivity of the ZnO:Al declined steadily, so that after 500 hours the resistivity had increased by about 50%. Secondly, the contact resistance at the ZnO:Al/Mo interconnect rose rapidly from approximately $1.5 \times 10^{-3}\ \Omega\ \text{cm}^2$ and saturated at about $5 \times 10^{-2}\ \Omega\ \text{cm}^2$ within 100 hours. The latter effect seemed to be the dominant factor in accounting for the observed drop in module fill factor. For unknown reasons, widely different stabilities of ZnO/Mo interconnects have been observed [291]. The use of gridded modules has been suggested as a way of securing a metal–metal interconnect that might be more stable than a metal–TCO interconnect [292]. In contrast to the contributions of ZnO to module degradation, experiments with bare CIGS cells have shown that the presence of the high-resistivity *i*-ZnO layer tends to protect the underlying device against damp heat degradation [293].

The increase in ZnO resistivity has been attributed to the reduction in carrier concentration due to the reaction of H_2O with oxygen vacancies in the film. It is easy to speculate that the detailed grain structure of ZnO could affect the diffusion of water through the film and hence the spatial distribution of resistivity. One reported approach to forming moisture-resistant ZnO is heavy doping of ZnO with Ga [294]. The ZnO was produced by off-axis RF magnetron sputtering. An increase in Ga concentration (from the usual 4.3 wt % to 23.1 wt %) appears to result in the *c*-axis of crystallites being more randomly oriented rather than being normal to the substrate. The usual columnar structure is replaced by randomly oriented discrete grains, thus eliminating boundaries that run from the bottom of the film to the top. At the same time, the film becomes smoother. Upon damp heat soaking, more stable values of carrier concentration n_e and mobility μ are found, especially around 10–12 wt% Ga, albeit at lower values than those obtained at 4.3 wt% Ga. It is also conjectured that some of the excess Ga can become activated to compensate for the usual carrier loss.

It has been shown that substrate roughness has a large effect on the degradation of sputtered ZnO:Al [295]. In this study the films were deposited by RF magnetron sputtering from a ceramic $ZnO:Al_2O_3$ (1 wt %) target. Thus, on smooth quartz, the lateral resistivity of ZnO:Al increased by only a factor of two after 1000 h of damp heat exposure, whereas on rough (bead-blasted) quartz the resistivity increased by two orders of magnitude. Despite this dramatic effect, optical analyses revealed only a minor decline in n_e and μ within the grains. Thus, extended grain boundaries may serve both to facilitate water penetration and to impede current transport after degradation. It also appeared that thicker films (e.g. 0.5 μm) are more stable than thinner films (e.g. 0.1 μm). The increased damp heat sensitivity of rough ZnO:Al layers is consistent with earlier observations on CIGS modules prepared by different techniques [291].

In a study aimed at screening TCOs for use in CIGS solar cells, ZnO:Al, ZnO:B, amorphous indium zinc oxide (IZO) and amorphous ITO were subjected to 85/85 damp heat testing [296]. The polycrystalline films exhibited strong degradation in conductivity (and changes in XRD patterns) while the amorphous films were much more stable.

The use of multi-layer coatings consisting of a stack of paired Al_2O_3/polymer layers has also been investigated as a potential encapsulant for CIGS and CIGSS cells [297]. The stack provides

a tortuous path for moisture diffusion in the polymer layers after passage through defects in the Al_2O_3 layers. It was demonstrated that such coatings greatly retarded degradation of mini-modules and that moisture finally penetrates from the edges of the mini-modules. This suggests that an enhanced edge seal is required. Bare cells exhibited an increase in the sheet resistance of the ZnO, followed by an increase in the diode reverse saturation current density J_0. It was further observed that coated devices having a top TCO consisting of ITO (rather than ZnO) and a smooth surface degraded much more slowly than coated devices having a ZnO top TCO and a rough surface. For the former type of device, with a multi-layer Al_2O_3/polymer coating, no degradation was found at 1500 h in an 85/85 environment, with 25–30% degradation at 3200 h.

In a study focusing on LPCVD ZnO:B films grown at 155 °C, exposure to 100% RH at 40 °C for 800 h resulted in a huge drop in Hall mobility from 33 to 2 $cm^2/V\,s$ [298]. From the relative constancy of the optically deduced mobility it was inferred that the degradation mechanism occurred at the grain boundaries. For a similar exposure regime, the resistivity-increase factors for three films with different initial carrier concentrations, namely LPCVD ZnO:B with $8 \times 10^{19}/cm^3$, LPCVD ZnO:B with $2 \times 10^{20}/cm^3$, and sputtered ZnO:Al with $4 \times 10^{20}/cm^3$, were 20, 5, and 1.2, respectively. It was concluded that heavily doped ZnO films are less sensitive to damp heat exposure.

Several similar results were obtained in a study aimed at evaluating ZnO:Al (AZO) and ZnO:Ga (GZO) for potential use in liquid crystal displays [299]. In these damp heat tests (60 °C and 90% RH), thin films of undoped indium oxide and ITO were found to be stable while AZO and GZO always degraded in conductivity. Decreases in both n_e and μ were observed with the decrease in μ being stronger. The decrease in mobility was attributed to grain boundary scattering. For doped films, the best stability was found for [Al] in the range 5–8 at% and [Ga] around 5 at%. Increasing instability was found for films of decreasing thickness, while for films of equal thickness AZO was more stable than GZO. It was noted that a higher Hall mobility and a larger crystallite size both tended to confer improved stability. Interestingly, it was found that the original conductivity could always be regained by annealing DH-treated films in a reducing atmosphere at 300 °C. Annealing in vacuum had elsewhere been observed to largely recover the conductivity [300].

It has further been demonstrated using deuterated water (D_2O damp heat) that gaseous penetration of the entire film depth occurs within 24 h [301]. Subsequent exposure to H_2O reduced the deuterium concentration implying a dynamic D–H exchange, consistent with weak bonding and the film's ability to be annealed. This study also showed that texture-etched ZnO:Al films degraded more rapidly than unetched films, perhaps because of etching at grain boundaries. The dependence of the gaseous permeability of TCO films on film microstructure has been studied via the out-diffusion of implanted rare gas atoms [302]. Despite the degradation observed in bare ZnO layers, unencapsulated glass-based microcrystalline silicon thin film modules have been prepared having both a ZnO front TCO and a ZnO/Ag rear metallization that are stable under 1000 h 85/85 damp heat exposure [302].

Indium tin oxide (ITO) deposited at about 300 °C is a durable and stable material. For ITO films deposited at room temperature by RF magnetron sputtering from an oxide target it was found that thin layers (160 nm) were amorphous, while thicker layers (350 nm) were polycrystalline [303]. Thicker layers with (400) preferred orientation were found to be stable when subjected to 85/85 damp heat exposure for 1000 hours. Improved flexibility has also been reported for amorphous ITO [304]. We note that, as prepared by certain manufacturers, multijunction a-Si:H modules deposited on stainless steel with ITO as the top TCO layer have passed the damp heat qualification test. However, ITO is not suitable for all TCO applications, not only because of the cost of indium, but sometimes on account of indium diffusion, e.g. into a-Si:H or into the organic layers of an LED. In summary, for the three common TCOs ZnO:Al, ITO, and SnO_2:F, the susceptibility of bare TCO layers to degradation under DH conditions is greatest for ZnO:Al and least for SnO_2:F. There is evidence that indium zinc oxide (IZO) is considerably more resistant to DH than ZnO.

Thus, approaches to improving the stability of modules incorporating TCOs include: (a) the use or development of more damp heat resistant TCOs; (b) control of TCO microstructure; (c) effective thin-film capping of a damp heat sensitive TCO; and (d) use of a dessicant-containing edge sealant for glass–glass modules.

17.9 RECENT DEVELOPMENTS AND PROSPECTS

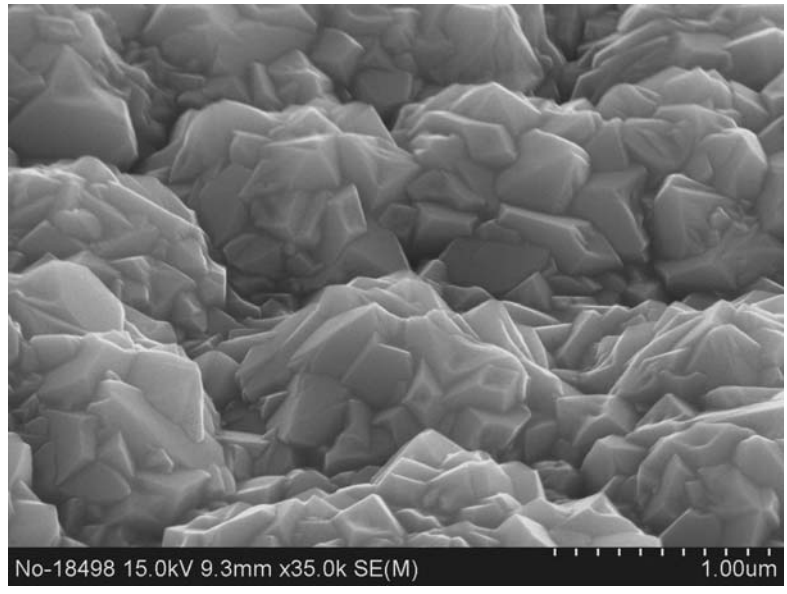
In this section we review selected current activities in the area of transparent conductors for PV applications. With the advent of large-scale production technology for CdTe and a-Si/nc-Si tandem modules, the improvement of optical transmission in the NIR region, and for the Si cells the improvement of light trapping capability, are attracting huge attention. A higher-performance TCO than ITO is being sought for HIT cells. As will be shown, advances in TCO materials are taking many exciting twists and turns. Furthermore, new PV materials, photonic effects and device architectures are leading to multi-functional application of TCOs. The relentless increase in the cost of indium is also spurring research into alternative or reduced In-content TCO materials. The need to withstand the rigors of outdoor deployment is inspiring new approaches to improving TCO stability or flexibility. Finally, although outside the scope of this review, the field is being increasingly supported by understanding and predictions resulting from theoretical modeling based on molecular dynamics and the *ab initio* density-functional approach for electronic structure calculations for both crystalline and amorphous TCOs.

17.9.1 Evolution of Commercial TCO-coated Glass

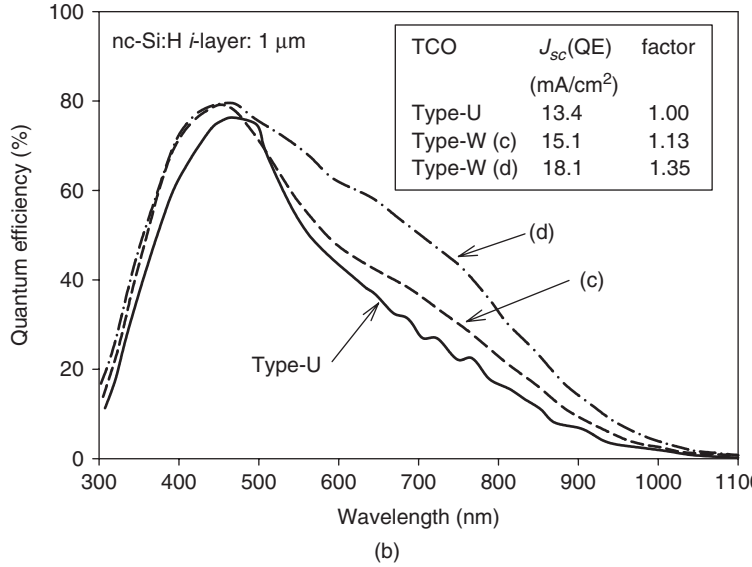
The high volume manufacturers of SnO_2:F coated glass (glass companies) are responding to customer requests for higher-performance products for a-Si:H and a-Si:H/nc-Si:H applications. Known to be in the pipeline are FTO products on reduced iron content glass, with antireflection coatings, and with improved grain structure and orientation. All these features are designed to increase light coupling to the PV cell [305]. As mentioned in Section 17.6.4, carrier mobility in low carrier concentration Type-VU TCO can be increased to 55 $cm^2/V\,s$. The Type-VU product is also heat-strengthened.

A new type of double-textured SnO_2:F film (designated W-texture TCO) has recently been developed at Asahi Glass Co. The double-textured TCO films show high light scattering ability over the entire light spectrum to which a-Si:H and nc-Si:H solar cells are sensitive [306]. The film, with thickness $>2\,\mu m$, is made by APCVD and has a surface with two distinct types of morphology. The first results from large, micrometer-scale bumps of tin oxide, and the second results from an overcoat with a sub-micrometer-scale textured tin oxide. This is clearly shown in Figure 17.24(a). The net result is a morphology having two characteristic length scales and an RMS roughness up to 150 nm. This type of TCO exhibits a remarkably high haze throughout the wavelength region 400–1200 nm, and can achieve a haze value of 88% at 800 nm (see curve (f) in Figure 17.13). The use of W-type SnO_2:F provides very effective light diffusion and light trapping in nc-Si:H solar cells. Consequently, a larger spectral response at long wavelengths is obtained relative to that on Type-U SnO_2:F. As shown in Figure 17.24b, QE(700 nm) can be raised from 30% to 50%. (On the other hand, W-texture TCO does not improve the QE of a-Si:H cells.) This type of TCO is expected to be useful for a-Si:H/nc:H-Si tandem module fabrication [307]. However, the cost of large-scale production is still an open question.

The production of TCO-coated glass for PV applications seems likely to diversify and undergo lively competition in the future. Several companies are able to offer APCVD machines for off-line tin oxide coating (as opposed to on-line coating at a float glass plant). Such machines



(a)



(b)

Figure 17.24 (a) FE-SEM image of double-textured TCO SnO_2:F having RMS roughness 150 nm. (b) The external quantum efficiency of nc-Si:H solar cells on two double-textured TCO samples with RMS roughness of 122 nm (curve (c)) and 150 nm (curve (d)), and on one Type-U sample. (The spectral haze for the Type-W TCOs is shown in Figure 17.13.) Figure (a) courtesy AGC; Figure (b) data re-plotted from Oyama T *et al*., *Mater. Res. Soc. Symp. Proc*. **1101**, KK02 (2008) [306]

could be operated either by those interested in selling the product or by end-users of the product. It is known that increased optical absorption in SnO_2:F can result from impurities (e.g. Cl, or C if organometallic precursors are used), interstitial F atoms (non-doping F), or film thickness (if thick films have to be used to achieve the necessary haze value). It remains to be seen how well these effects can be controlled at the different production scales. Zinc oxide produced either by LPCVD or APCVD may also enter the commercial fray. Several types of TCO produced by sputtering-based technologies could also be offered commercially using, for example, large in-line sputtering systems with rotating cathode magnetrons or other types of cathodes.

17.9.2 Quest for High Carrier Mobility

The importance of securing a high carrier mobility to reduce TCO optical absorption has been stressed both in this chapter and in the literature. So far, the highest carrier mobility was achieved in Sn-doped CdO grown epitaxially on single-crystal MgO (111) substrates by pulsed laser deposition (PLD) at 750 °C [308]. A peak carrier mobility of 609 $cm^2/V\,s$ was obtained at a Sn doping concentration of 2.5% (carrier concentration $4.74 \times 10^{20}\,cm^{-3}$) corresponding to a resistivity of $2.2 \times 10^{-5}\,\Omega\,cm$. Also using PLD, a carrier mobility of 159 $cm^2/V\,s$ (with $n_e = 4.3 \times 10^{20}\,cm^{-3}$) was obtained in Ti-doped In_2O_3 deposited on a sapphire substrate at 600 °C and in 0.56 mTorr O_2 [21]. Returning to the CdO work, a mobility of only 27 $cm^2/V\,s$ was obtained for CdO deposited on glass. The X-ray FWHM figures for the CdO films on MgO (111) and glass substrates were 0.19° and 4.2°, respectively, indicating a strong deterioration in crystal quality. A mobility of about 70 $cm^2/V\,s$ was, however, achieved for In-doped CdO deposited on glass by MOCVD at 410 °C [309].

For PV applications, it will be challenging to achieve a high mobility TCO on glass combined with compatibility for large-scale production. To accomplish this, a better understanding of doping, carrier transport and optical transparency are needed. Despite advances in the theoretical understanding of carrier generation and band structure [310, 311], some phenomena still can not be adequately explained. For example, the recent demonstration of high mobility in In_2O_3 films doped with Mo, Ti, or Zr [22, 23, 319] is still not fully understood. A new theory involving magnetic interactions has been advanced to explain the experimental results [312]. Thus, as a result of exchange-splitting of the Mo *d* states, carriers of one spin are scattered only by Mo *d* states of the same spin and not by the remaining half with the opposite spin, i.e. the effective concentration of Mo scattering centers is reduced by 50%. As another example, an unusually high mobility of 140 $cm^2/V\,s$ ($n_e = 1.5 \times 10^{20}\,cm^{-3}$) has been obtained in hydrogen-doped In_2O_3 (IO:H) deposited on glass by RF magnetron sputtering [26, 313]. The sputtering was conducted without substrate heating in Ar and O_2 gases and a partial pressure of H_2O of about 7.5×10^{-7} Torr. The resulting amorphous In_2O_3:H film was then annealed in vacuum at 200 °C, causing it to crystallize. Remarkably large grains were obtained (>90 nm) with a hydrogen content of 2–4%. At lower partial pressures of H_2O the mobility after annealing was only 20–30 $cm^2/V\,s$. Also, by adding hydrogen to the sputtering gas (H_2/Ar 0.2–0.3%), a mobility of over 50 $cm^2/V\,s$ was obtained in ZnO:Al films sputtered from ZnO:Al_2O_3 (0.1 or 0.2 wt %) targets [19]. The high mobility was maintained over a broad range of substrate temperature (from room temperature to 300 °C). The In_2O_3:H results are particularly striking, with the H apparently both facilitating grain growth and acting as a shallow donor, yet in both this case and the ZnO:Al case the mechanisms by which hydrogen enhances the mobility remain to be clarified.

Films of ZnO:Al are generally degraded when heated in air at temperatures above 300 °C due to a strong increase in resistivity [314]. However, when capped with a film of Si or a-Si:H, ZnO:Al can be heated to over 600 °C without degradation and the resistivity is in fact reduced [315, 316]. In particular, with a slow ramp heating (to 650 °C) and cooling of the structure borosilicate glass/SiN diffusion barrier/ZnO:Al (RF-sputtered, 300 °C)/*n*-type a-Si:H (PECVD, 50 nm) in a

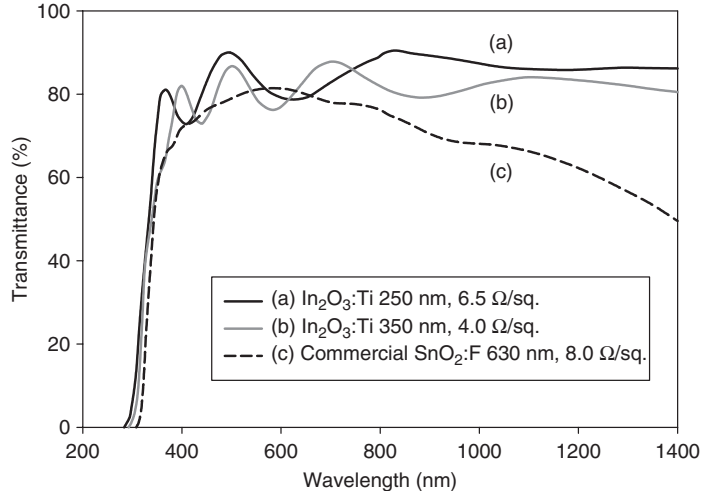


Figure 17.25 Transmission spectra of: (a) ITiO (250 nm, mobility 105 cm^2/Vs); (b) ITiO (350 nm); and (c) FTO (630 nm, mobility 25.2 cm^2/Vs). Reproduced from Bowers J *et al*., *Prog. Photovolt.: Res. Appl*. **17**, 265 (2009), with permission from Elsevier [320]

tube furnace in N_2, the resistivity of the ZnO:Al decreased in the best case from 2.7×10^{-4} to 1.4×10^{-4} Ω cm resulting from a mobility increase from 42 to 67 cm^2/V s. This mobility in ZnO:Al films had previously only been achieved by epitaxial growth on sapphire at 750 °C (70 cm^2/V s at 1×10^{20} cm^{-3}) [317]. The capping concept may allow fabrication of thin poly-Si solar cells in superstrate configuration using either solid-phase crystallization or Al-induced crystallization.

Both the mobility and carrier concentration of In_2O_3:Ti have been increased by vacuum annealing. Thus, a film deposited at 500 °C by RF magnetron sputtering on soda-lime glass was annealed at 530 °C and a mobility of 105 cm^2/V s was achieved [318]. The transmission spectra of ITiO films are compared with that of a commercial SnO_2:F (FTO) sample in Figure 17.25. The superior transmission of the ITiO is evident. This is partly due to titanium being an efficient dopant in In_2O_3 (yielding one carrier per Ti atom [319]) thus minimizing scattering and allowing a high mobility.

The use of high-mobility TCO films in the fabrication of solar cells with various absorber types has been reported in recent publications. Of particular interest is the use of a textured, high conductivity, low-absorption i-ZnO/n^+-ITiO bi-layer (see Section 17.5.3.2) in nc-Si:H cells to increase J_{sc} [321]. With a view to their ultimate utilization in tandem- or triple-junction cells or bifacial cells, CdTe solar cells were made on Ti-doped In_2O_3 (ITiO) films with mobility over 100 cm^2/V s [20]. Compared with SnO_2:F and ITO of comparable conductivity, it was shown that ITiO gives much better optical transmission through the cell for $\lambda > 900$ nm. High-mobility ITiO films have also been used as a front TCO in CIGS solar cells [69], as a back contact for bifacial CIGS solar cells [322], and in dye-sensitized TiO_2 solar cells for long-wavelength light transmission improvement [320, 323]. In the bifacial CIGS cell, the rear/front efficiency ratio was improved from 0.59 for ITO to 0.79 for ITiO. In a-Si:H/c-Si heterojunction (HIT) solar cells, the use of high-mobility, H-doped In_2O_3 (deposited at room temperature and annealed at 140 °C) having a low carrier concentration (1.5×10^{20} cm^{-3}) helps to reduce not only the absorption loss, but also the reflection loss at the TCO/a-Si:H interface by maintaining a high refractive index for the TCO out to 1000 nm (see Section 17.4.2). Thus, the refractive indices of IO:H and ITO at 600 and 1000 nm are: IO:H 2.0 and 1.88; and ITO 1.80 and 1.20 [26].

The foregoing examples show that for the future production of high-mobility TCOs, the choice of material, deposition process, doping element and post-processing treatment are still very much open questions.

17.9.3 Enhancement of Scattering and Useful Absorption

In thin film Si:H cells, the reflection and scattering of light at the back contact enhances light trapping in the cell. We saw in Section 17.5.5 that insertion of a low-index layer (e.g. a TCO) between the Si and the metal (Al or Ag) increases the reflectivity. To increase the light scattering ability of the contact, a new approach is to prepare Ag nanoparticles on the front side of the TCO. By annealing a 20 nm Ag layer at 180 °C for several hours, discrete nanoparticles of Ag with an average diameter of 300 nm can be formed. For an *n-i-p* nc-Si:H cell formed on a glass/Ag/TCO/Ag nanoparticle substrate it was found that the quantum efficiency at long wavelengths is improved by the presence of the nanoparticles, despite competing absorption losses [324]. Unfortunately, both FF and V_{oc} were reduced. The scattering was found to be wavelength-dependent.

There has been much excitement about the potential use in the PV arena of increased field intensities associated with surface plasmons. On the one hand it is known that surface plasmons can be responsible for parasitic absorption, e.g. in the Ag layer of the common textured Ag/ZnO reflector [325]. However, in an effort to harness the effect and enhance useful absorption, plasmonically active Ag nanoparticles were incorporated into a thinner-than-usual organic bulk heterojunction cell resulting in the structure: glass/ITO/Ag/PEDOT:PSS/P3HT:PCBM/Ba/Al. The 10 nm nanoparticles were formed by evaporating a 2 nm (by mass) island layer of Ag onto the ITO. The average J_{sc} of the devices was thereby increased from 4.6 to 7.3 mA/cm^2 [326]. Also, photocurrent enhancement was demonstrated in optically-thin GaAs solar cells with circular pillars of Ag deposited on the AlGaAs window layer [327]. While surface plasmons may improve an ultra-thin device, it remains unclear whether this effect is suitable to improve optimized, high-efficiency PV devices.

The use of vertical structures such as ZnO, SnO_2, or TiO_2 nanorods or platelets, onto which active PV layers are then grown, offers exciting possibilities to increase the probability of carrier collection while also increasing the optical absorption in devices such as thin Si:H, dye-sensitized, or organic solar cells (see, for example [328]). Related concepts have long been appreciated, with the vertical multijunction cell having been analyzed as early as 1972, and the black appearance of highly structured surfaces also being well known. More recent work has included analysis of the radial junction nanorod solar cell [329], the fabrication of ITO/*n*-a-Si:H/*p*-Si (VLS method)/Ta_2N Si nanowire solar cells [330], and demonstration of strong optical absorption enhancement in a-Si:H nanocone arrays formed by reactive ion etching [331]. It can be anticipated that, by using a nanostructured TCO, thin Si:H devices having a high stabilized efficiency will soon be realized. The nanostructured TCO can be prepared by forming TCO nanorods, either by direct nanorod growth onto a planar TCO or by etching of the TCO. (Alternatively, columnar holes could be etched into a TCO layer.) To complete superstrate-type devices, the nanostructured TCO is overcoated with the Si:H layers, followed by conformal metallization. These possibilities have been outlined in reference [332].

17.9.4 Doped TiO_2 and Other Wide-gap Oxides

Recent advances have made doped TiO_2 films having the anatase structure a promising candidate for a viable TCO material. Anatase TiO_2 has a bandgap of 3.2 eV and an electron effective mass of $1.0m_e$ as opposed to $8-20m_e$ for the rutile phase of TiO_2. These basic properties indicate that TiO_2 might be capable of being made highly conductive with excellent optical transmission. However, early studies found it difficult to achieve these two features simultaneously. The breakthrough came via the epitaxial growth of Nb-doped TiO_2 films onto $SrTiO_3$ (100) substrates by PLD at

about 550 °C in which a film resistivity as low as $2.3 \times 10^{-4}\,\Omega\,\text{cm}$ was achieved at 6% Nb doping together with an internal optical transmittance of more than 97% in the visible region [11, 333, 334]. With 3% Nb, $n_e = 1.2 \times 10^{21}\,\text{cm}^{-3}$ and $\mu = 22\,\text{cm}^2/\text{V s}$. The material behaved as a typical degenerate semiconductor. On alumina substrates the films grew with the rutile structure leading to a much higher resistivity of $0.12\,\Omega\,\text{cm}$. Epitaxial TiO_2:Nb (with 15% Nb) deposited by RF magnetron sputtering onto (100) $SrTiO_3$ at 375 °C attained a resistivity of $3.3 \times 10^{-4}\,\Omega\,\text{cm}$, with $n_e = 2.4 \times 10^{21}\,\text{cm}^{-3}$, $\mu = 7.9\,\text{cm}^2/\text{V s}$ and a transparency of about 80% [335].

The next step in development was the achievement of a resistivity as low as $4.6 \times 10^{-4}\,\Omega\,\text{cm}$ on alkali-free glass. This was accomplished by depositing an amorphous $Ti_{0.94}Nb_{0.06}O_2$ film onto unheated substrates by PLD at 1×10^{-4} Torr O_2 and then rapidly annealing the film in 1 atmosphere hydrogen to 500 °C. The optical absorbance was reduced by the annealing step to less than 10% in the visible region [336]. However, for TiO_2 to become a suitable TCO for solar applications it is desirable that a high-quality film can be deposited on glass by sputtering. This goal is being approached with the report that a resistivity of $9.5 \times 10^{-4}\,\Omega\,\text{cm}$, mobility of $3.9\,\text{cm}^2/\text{V s}$ and an absorbance of less than 10% were achieved by reactively sputtering a metallic $Ti_{0.94}Nb_{0.06}$ target at room temperature onto alkali-free glass followed by annealing of the amorphous TiO_2:Nb film to 600 °C in H_2 [12]. Large-area deposition of TiO_2:Nb has been accomplished by pulsed DC sputtering of a ceramic $Ti_{0.94}Nb_{0.06}O_2$ target in dynamic mode. By sputtering at somewhat higher pressures (6 mTorr) the anatase phase of TiO_2 was achieved (after annealing) and a film resistivity of $1.8 \times 10^{-3}\,\Omega\,\text{cm}$ was obtained by vacuum annealing at 360 °C [337].

Other wide-gap oxides include the classic insulators Al_2O_3, BaO, CaO, MgO and SiO_2. These materials have so far failed to be developed as transparent conductiong oxides, partly because electrons resulting from oxygen vacancies appear to be strongly localized near the vacancy (see the literature on F-centers). Nevertheless, there may exist a possibility for certain multi-component oxides, such as $InAlZnO_4$, to be developed as a TCO.

17.9.5 Other Types of Transparent Conductor

An interesting material that can be prepared in thin film form to perform the TCO function is the carbon nanotube (CNT). The electrical conductivity of an individual carbon nanotube can be as high as that of a bulk metal [338]. By forming a network of single-wall carbon nanotubes (SWCNTs), a thin layer can be largely transparent while yielding good lateral conductivity. Carbon nanotubes can be prepared by an arc discharge method or by water-assisted CVD using ethylene. The latter technique yields stunning, dense SWNT forests up to 2.5 mm in height, the CNT diameter being about 2 nm [339]. Carbon nanotubes have great chemical, thermal and mechanical stability. Because of the one-dimensional nature of the tubes, and because SWCNTs usually consist of a mixture of metallic and (doped) semiconducting tubes, the details of electrical transport in networks are complex. While Hall effect measurements on SWCNT networks do not yield meaningful results, the Seebeck coefficient is positive, implying hole conduction [340]. A $50\,\Omega/\square$ layer can be obtained at a thickness of about 40 nm. Optical transmission is very high in the NIR (normal free-carrier absorption is not seen) but falls with increasing photon energy in the visible, e.g. $T = 80\%$ at 1 eV and 60% at 3 eV at $50\,\Omega/\square$. Initial exploration of the use of SWCNT electrodes in different types of solar cells is already under way, as briefly described below.

In the organic photovoltaic area, a SWCNT electrode was employed to replace the standard hole-collecting ITO electrode to form a solar cell having the structure quartz/SWCNT/PEDOT:PSS/P3HT-PCBM/Ga:In [341]. Both in this paper and in subsequent work in which a multi-wall CNT layer was prepared on ITO-coated glass, it was pointed out that the polymer layers infiltrate the voids in the CNT layer leading to 3D hole collection. Thus, in the ITO/MWCNT device, hole collection occurs both in the MWCNT network *and* at the exposed portions of the

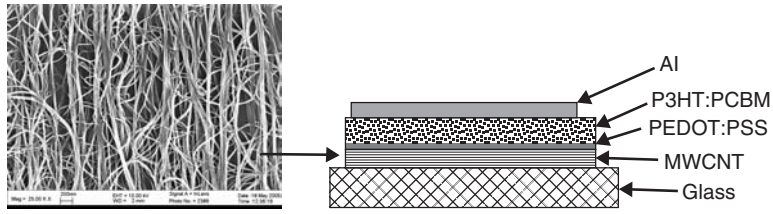


Figure 17.26 Left: SEM image of the liquid-densified multiwall CNT sheet having a thickness of 50–100 nm and $R_{sh} = 600\,\Omega/\square$; the sheet retains its high porosity after deposition of a thin layer of (electron-blocking) PEDOT:PSS. Right: Schematic of the organic solar cell ($J_{sc}^{\mathrm{CNT}} = 5.5\,\mathrm{mA/cm^2}$) with CNT anode as a transparent conducting electrode. In a cell with a hybrid ITO/CNT anode, the pores between the nanotubes increase the area for hole collection so that $J_{sc} = J_{sc}^{\mathrm{ITO}} + J_{sc}^{\mathrm{CNT}} = 11\,\mathrm{mA/cm^2}$. Reproduced from Ulbricht R *et al*., *Sol. Energy Mater. Sol. Cells* **91**, 416 (2007), with permission from Elsevier [342]

ITO electrode, leading to an almost doubled J_{sc} of 11 mA/cm^2 in this hybrid anode device [342]. The structure of the organic solar cell is shown in Figure 17.26.

Transparent electrodes consisting of SWCNTs have also been applied to CIGS solar cells to replace the conventional TCO layer. The best result (8.2%) was obtained with the structure 100 Ω CNT/100 nm Parylene/1000 Ω CNT/CdS/CIGS/Mo in which the 1000 Ω/□ CNT layer coated with a 100 nm Parylene-N layer is used in place of the usual *i*-ZnO buffer layer [343]. In the CdTe area we have already discussed the use of ITO as part of a transparent rear electrode for CdTe cells designed as the top cell of a tandem device (see Section 17.5.1.2). The absence of free-carrier absorption in SWCNT networks suggests that replacement of the ITO layer by a SWCNT layer could be beneficial for this type of cell. A 12.4% transparent CdTe cell with a Cu_xTe/100 Ω/□ SWCNT network back contact has been reported [344].

Another wave of interesting devices is likely to include the use of graphene, a two-dimensional conductor consisting of a single layer of carbon atoms. Unsupported graphene films have an ultra-high electron mobility larger than that of any other semiconductor. Large-area films can be produced by acid oxidation of flake graphite, chemical exfoliation of the oxide, deposition on quartz, and finally reduction. Consisting of multiple graphene sheets, a film of 10 nm thickness has a sheet resistance of 1.8 kΩ/□ ($\sigma = 550$ S/cm) and a transmittance of 71% at 1000 nm at the present stage of development. Such a graphene film has been used in a dye-sensitized TiO_2 solar cell to replace the SnO_2:F. A reasonable V_{oc} of 0.7 V was obtained [345]. Graphene films have also been used in organic solar cells [346]. A new hybrid graphene-carbon nanotube material is also thought to be promising as a transparent conductor [347]. Yet another fascinating development is a solution-grown Ag nanowire mesh with $T_{Solar} = 80\%$, $R_{sh} = 16\,\Omega/\square$ and with good light scattering capability [348].

17.9.6 Amorphous TCOs

It is not unreasonable to predict widespread application of amorphous TCOs based on already discovered outstanding properties. At the time of writing, transparent thin film transistors (TFTs) based on *a*-$InGaZnO_4$ (*a*-IGZO) are being used in prototype front-drive, full-color e-paper and flexible AM-OLEDs [349]. Amorphous InZnO (*a*-IZO, with In content of 90%) is being used in OLEDs and LCDs [2]. In the PV arena, *a*-IZO is being applied to CIGS cells in an R&D setting as a potential replacement for ZnO:Al [350]. Other amorphous TCOs of interest are *a*-ZTO and *a*-ZIO. Perhaps the most surprising feature of these materials is their electron mobility. In *a*-IGZO

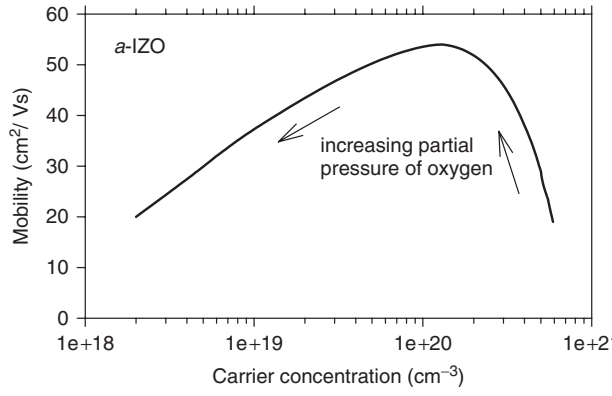


Figure 17.27 General relationship between Hall mobility and carrier concentration for amorphous indium zinc oxide films sputtered at room temperature using oxide targets having In metal contents ranging from 60 to 84%. Figure constructed with data from Leenheer A *et al.*, *Phys. Rev*. B **77**, 115215 (2008) [353]

TFTs the field-effect mobility μ_{FE} is $>10\,\mathrm{cm^2/V\,s}$ (compared with $1\,\mathrm{cm^2/V\,s}$ in a-Si:H) [351], in *a*-ZTO μ_{FE} is typically $20-50\,\mathrm{cm^2/V\,s}$ [352], while in *a*-IZO Hall mobilities in the range of $52-60\,\mathrm{cm^2/V\,s}$ have been demonstrated [7, 90, 353].

Whereas the electrical conductivity of conventional polycrystalline TCOs can be controlled by substitutional doping, the conductivity of *a*-IZO prepared by DC magnetron sputtering of IZO targets can be adjusted by the quantity of oxygen added to the sputtering gas. While the conductivity and carrier concentration both decline monotonically with increasing oxygen, the mobility passes through a peak at a partial pressure of oxygen of about 2% [353]. This work employed oxide targets with atomic In concentrations, In/(In + Zn), of 0.84, 0.80, 0.70, and 0.60, and the sputtering was conducted at ambient temperatures. A more or less universal curve of Hall mobility versus carrier concentration n_e was obtained, with oxygen being the controlling parameter, regardless of the metal ratio In/Zn in the source target. This relationship is shown in Figure 17.27. The optical bandgap E_g of the film is about 3.1 eV at 4% oxygen and increases with decreasing oxygen to about 3.9 eV at zero added oxygen, consistent with a Burstein–Moss shift. The slope of E_g versus $n_e^{2/3}$ implies $m_{vc}^* = 0.56m_e$ (see Section 17.4.2). For $n_e < 1 \times 10^{19}\,\mathrm{cm^{-3}}$ the carriers may need to navigate a distribution of potential barriers in the conduction band so that the conductivity is controlled by a percolation process. A minimum resistivity of about $5.2 \times 10^{-4}\,\Omega\,\mathrm{cm}$ (at $n_e = 6 \times 10^{20}\,\mathrm{cm^{-3}}$) was observed in this work, while $2.7 \times 10^{-4}\,\Omega\,\mathrm{cm}$ has been obtained both for *a*-IZO films grown at 110 °C by PLD [90] and for *a*-IZO films deposited at room temperature by RF magnetron sputtering [354]. Ultra-flat amorphous IZO has also been prepared on large area substrates ($>1\,\mathrm{m^2}$) by DC sputtering with a minimum resistivity of $3.25 \times 10^{-4}\,\Omega\,\mathrm{cm}$ being obtained at $T_s = 200\,^\circ\mathrm{C}$ [355]. These highly conductive films behave like traditional degenerate TCOs with dominant ionized impurity scattering.

Early studies of amorphous indium oxide (*a*-IO) had examined the effect of Sn addition and had shown that the resistivity of *a*-ITO is higher than that of *a*-IO (for equal n_e) [356]. It was later confirmed that neither Zn nor Sn in *a*-IO contribute carriers [7], and neither does Al_2O_3 in *a*-IZO [357]. The source of the free carriers in *a*-IO and *a*-IZO would therefore appear to be oxygen vacancies. A similar conclusion was reached for *a*-$Zn_{1.2}In_{1.9}Sn_{0.1}O_{4.25-\delta}$ [358]. The Seebeck coefficient is negative implying that the carriers are electrons. In contrast to *a*-IO, which crystallizes at 180 °C, *a*-IZO does not crystallize until >500 °C [359]. Another important difference is that the formation process for *a*-IZO is much more robust with the deposited film remaining amorphous

even at $T_s = 300\,^{\circ}\mathrm{C}$. There are also some significant differences in transport behavior between these oxide materials and a-Si:H. In a-Si:H, transport is via hopping between localized tail states. The covalent bonding in Si involves directional hybrid sp^3 or p orbitals, while in the ionic oxide semiconductors it is commonly suggested that the conduction band minimum consists of s-orbitals with spherical symmetry and large radius that are unaffected by a disordered structure [349]. Also, the Hall sign anomaly of a-Si:H is absent in *a*-IZO. Two of the most attractive features of amorphous materials, as opposed to polycrystalline materials, are their natural large-scale uniformity and absence of boundaries running throughout the film thickness. The latter feature may impart superior environmental stability (see Section 17.8). Combined with competitive electronic properties, it seems safe to say that amorphous TCOs have a bright future.

As we have seen, transparent conducting oxides play a vital, and sometimes complex, role in many of the PV technologies that are themselves described in greater detail in the other chapters of this handbook. From just Section 17.9 alone, it should be obvious that the TCO field is burgeoning with new materials, nanostructures, deposition methods, and research directions. It is hoped that the contents of this chapter will prove to be informative, inspiring, and useful to practitioners and students alike.

REFERENCES

1. Minami T, *MRS Bulletin* **25**, 38 (2000).
2. Minami T, *Thin Solid Films* **516**, 1314 (2008).
3. Wu X, Mulligan W, Coutts T, *Thin Solid Films* **286**, 274 (1996).
4. Wu X, Coutts T, Mulligan W, *J. Vac. Sci. Technol*. **A15** (3), 1057 (1997).
5. Kluth O, Agashe C, Hüpkes J, Müller J, Rech B, *3rd WCPEC-3*, **2**, 1800 (2003).
6. Minami T, *Semicond. Sci. Technol*. **20**, S35 (2005).
7. Yaglioglu B, Huang Y-J, Yeom H-Y, Paine D, *Thin Solid Films* **496**, 89 (2006).
8. Kawazoe H, Yanagi H, Ueda K, Hosono H, *MRS Bulletin* **25**, 28 (2000).
9. Meyer B, Sann J, Hofmann D, Neumann C, Zeuner A, *Semicond. Sci. Technol*. **20**, S62 (2005).
10. Hofmann D *et al*., *Phys Rev. Lett*. **88**, 045504 (2002).
11. Furubayashi Y *et al*., *Appl. Phys. Lett*. **86**, 252101 (2005).
12. Yamada N, *Thin Solid Films* **516**, 5754 (2008).
13. Minami T, Nanto H, Takata S, *Appl. Phys. Lett*. **41**, 958 (1982).
14. Minami T, Nanto H, Takata S, *Jpn. J. Appl. Phys*. **23** L280 (1984).
15. Agura H, Suzuki A, Matsushita T, Aoki T, Masahiro, Okuda M, *Thin Solid Films* **445**, 263 (2003).
16. Park S-M, Ikegami T, Ebihara K, *Thin Solid Films* **513**, 90 (2006).
17. Agashe C, Kluth O, Hüpkes J, Zastrow U, Rech B, Wuttig M, *J. Appl. Phys*. **95**, 1911 (2004).
18. Guo S *et al*., *Proc. 23*rd *European Photovoltaic Solar Energy Conference*, 2482 (2008).
19. Duenow J, Gessert T, Wood D, Young D, Coutts T, *J. Non-Cryst. Solids* **354**, 2787 (2008).
20. Calnan S *et al*., *Thin Solid Films* **517**, 2340 (2009).
21. Gupta P, Ghosh K, Mishra S, Kahol P, *Mater. Lett*. **62**, 1033 (2008).
22. Meng Y *et al*., *Thin Solid Films* **394**, 219 (2001).
23. Delahoy A, Guo S, *J. Vac. Sci. Technol*. **A 23**, 1215 (2005).
24. Rauf I, *Mater. Lett*. **18**, 123 (1993).
25. Ohta H, Orita M, Hirano M, Tanji H, Kawazoe H, Hosono H, *Appl. Phys. Lett*. **76**, 2740 (2000).
26. Koida T, Fujiwara H, Kondo M, *Sol. Energy Mater. Sol. Cells* **93**, 851 (2008).
27. Shigesato Y, Takaki S, Haranoh T, *J. Appl. Phys*. **71**, 3356 (1992).
28. Coutts T, Young D, Li X, Mulligan W, Wu X, *J. Vac. Sci. Technol*. **A18**, 2646 (2000).

29. Ellmer K, Klein A, Rech B, *Transparent Conductive Zinc Oxide: Basics and Applications in Thin Film Solar Cells*, Springer-Verlag, Berlin (2008).
30. Norton D *et al*., *Materials Today*, 34 (June 2004).
31. Haacke G, *Ann. Rev. Mater. Sci*. **7**, 73 (1977).
32. Chopra K, Major S, Pandya D, *Thin Solid Films* **102**, 1 (1983).
33. Granqvist C, *Sol. Energy Mater. Sol. Cells* **91**, 1529 (2007).
34. Exarhos G, Zhou X-D, *Thin Solid Films* **515**, 7025 (2007).
35. Grundmann M, *The Physics of Semiconductors: An Introduction Including Devices and Nanophysics*, Springer-Verlag, Berlin (2006).
36. Madelung O (Ed.) *Semiconductors – Basic Data* (2nd edn) Springer-Verlag, Berlin (1996).
37. Morkoç H, Özgür Ü, eds. *Zinc Oxide: Fundamentals, Materials and Device Technology*, WILEY-VCH Verlag GmbH & Co. Weinheim (2009).
38. Gläser H, *Large Area Glass Coating*, von Ardenne Anlagentechnik GmbH, Dresden (2000).
39. Westwood W, *Sputter Deposition*, AVS Education Committee Book Series, Vol. 2, AVS, New York (2003).
40. Nanto H, Minami T, Orito S, Takata S, *J. Appl. Phys*. **63**, 2711 (1988).
41. Yoshida Y, Gessert T, Perkins C, Coutts T, *J. Vac Sci. Technol*. **A21**, 1092 (2003).
42. Maruyama T, Tago T, *Appl. Phys. Lett*. **64**, 1395 (1994).
43. Fortunato E *et al*., *Thin Solid Films* **451-452**, 443 (2004).
44. Das C *et al*., *Sol. Energy Mater. Sol. Cells* **93**, 973 (2009).
45. Delahoy A, Cherny M, *Mater. Res. Soc. Symp. Proc*. **426**, 467 (1996).
46. Malkomes N, Vergöhl M, Szyszka B, *J. Vac. Sci. Technol*. **A19**, 414 (2001).
47. May C, Menner R, Strümpfel J, Oertel M, Sprecher B, *Surf. Coat. Technol*. **169-170**, 512 (2003).
48. Jäger S, Szyszka B, Szczyrbowski J, Bräuer G, *Surf. Coat. Technol*. **98**, 1304 (1998).
49. May C, Strümpfel J, Schulze D, *Proc. 43rd SVC Ann. Tech. Conf*., p. 137 (2000).
50. Schulze D, *Proc. 46th SVC Ann. Tech. Conf*., p. 233 (2003).
51. Tominaga K, Kume M, Yuasa T, Tada O, *Jpn. J. Appl. Phys*. **24**, 35 (1985).
52. Minami T, Miyata T, Ohtani Y, Mochizuki Y, *Jpn. J. Appl. Phys*. **45**, L409 (2006).
53. Calnan S, Upadhyaya H, Thwaites M, Tiwari A, *Thin Solid Films* **515**, 6045 (2007).
54. Ishii K, *J. Vac. Sci. Technol*. **A7**, 256 (1989).
55. Jung Th, Kälber T, Heide V, *Surf. Coat. Technol*. **86-87**, 218 (1996).
56. Delahoy A, Guo S, Paduraru C, Belkind A, *J. Vac. Sci. Technol*. **A22**, 1697 (2004).
57. Delahoy A *et al*., *4th World PVSC*, 327 (2006).
58. Igasaki Y, Saito H, *J. Appl. Phys*. **70**, 3613 (1991).
59. Ruske F, Sittinger V, Werner W, Szyszka B, *Proc. 48th SVC Ann. Tech. Conf*., p. 302 (2005).
60. Szyszka B *et al*., *Proc. 44*th *SVC Ann. Tech. Conf*. 272 (2001).
61. Birkholz M, Selle B, Fenske F, Fuhs W, *Phys. Rev*. **B 68**, 205414 (2003).
62. Fenske F, Selle B, Birkholz M, *Jpn. J. Appl. Phys*. **44**, L662 (2005).
63. Jiang X, Jia C, Szyszka B, *Appl. Phys. Lett*. **80**, 3090 (2002).
64. Kajikawa Y, *J. Crystal Growth* **289**, 387 (2006).
65. Martel A *et al*., *Surf. Coat. Technol*. **122**, 136 (1999).
66. Guo S, Shafarman W, Delahoy A, *J. Vac. Sci. Technol*. **A24**, 1524 (2006).
67. Takeda H, Sato Y, Iwabuchi Y, Yoshikawa M, Shigesato Y, *Thin Solid Films* **517**, 3048 (2009).
68. Delahoy A *et al*., *19th European Photovoltaic Solar Energy Conference*, p. 1686 (2004).
69. Delahoy A, Chen L, Akhtar M, Sang B, Guo S, *Solar Energy* **77**, 785 (2004).
70. Delahoy A *et al*., *23rd European Photovoltaic Solar Energy Conference*, 2069 (2008).
71. Gerhardinger P, McCurdy R, *Mater. Res. Soc. Symp. Proc*. **426**, 399 (1996).
72. van Mol A, Chae Y, McDaniel A, Allendorf M, *Thin Solid Films* **502**, 72 (2006).
73. Sheel D *et al*., *Thin Solid Films* **517**, 3061 (2009).

74. Giunta C, D. Strickler A, Gordon R, *J. Phys. Chem*. **97**, 2275 (1993).
75. Buchanan J, McKown C, *J. Non-Cryst. Solids* **218**, 179 (1997).
76. Li X, Pankow J, To B, Gessert T, *Proc. 33rd IEEE Photovoltaic Specialists Conference*, p. 76 (2008).
77. Sato K, Gotoh Y, Wakayama Y, Hayashi Y, Adachi K, Nishimura H, *Reports Res. Lab. Asahi Glass Co. Ltd*, **42**, 129 (1992).
78. Kambe M, Fukawa M, Taneda N, Sato K, *Sol. Energy Mater. Sol. Cells* **90**, 3014 (2006).
79. Maruyama T, Tabata K, *Jpn. J. App. Phys*. **29**, L355 (1990).
80. Hu J, Gordon R, *Solar Cells* **30**, 437 (1991).
81. Liang H, Gordon R, *J. Mater. Sci*. **42**, 6388 (2007).
82. Hu J, Gordon R, *J. Appl. Phys*. **71**, 880 (1992).
83. Barankin M, Gonzalez II E, Ladwig A, Hicks R, *Sol. Energy Mater. Sol. Cells* **91**, 924 (2007).
84. Faÿ S, Kroll U, Bucher C, Vallat-Sauvain E, Shah A, *Sol. Energy Mater. Sol. Cells* **86**, 385 (2005).
85. Faÿ S, Feitknecht L, Schlüchter R, Kroll U, Vallat-Sauvain E, Shah A, *Sol. Energy Mater. Sol. Cells* **90**, 2960 (2006).
86. Faÿ S, Steinhauser J, Oliveira N, Vallat-Sauvain E, Ballif C, *Thin Solid Films* **515**, 8558 (2007).
87. Addonizio M, Diletto C, *Sol. Energy Mater. Sol. Cells* **92**, 1488 (2008).
88. Velasco A, Oguchi T, Kim H, *J. Cryst. Growth* **311**, 2731 (2009).
89. Freeman A, Poeppelmeier K, Mason T, Chang R, Marks T, *MRS Bulletin* **25/8**, 45 (2000).
90. Mikawa M *et al*., *Mater. Res. Bull*. **40**, 1052 (2005).
91. Ginley D *et al*., *Thin Solid Films* **445**, 193 (2003).
92. Hamberg I, Granqvist C, *J. Appl. Phys*. **60**, R123 (1986).
93. Yamada A, Sang B, Konagai M, *Appl. Surf. Sci*. **112**, 216 (1997).
94. Asikainen T, Ritala M, Leskelä M, *Thin Solid Films* **440**, 152 (2003).
95. Delahoy A, Ruppert A, *Proc. 2nd World Conference on Photovoltaic Solar Energy Conversion*, (European Commission, Ispra) 668 (1988).
96. Aukkaravittayapun S *et al*., *Thin Solid Films* **496**, 117 (2006).
97. Wienke J, Booij A, *Thin Solid Films* **516**, 4508 (2008).
98. Sawada Y, Kobayashi C, Seki S, Funakubo H, *Thin Solid Films* **409**, 46 (2002).
99. Veluchamy P *et al*., *Sol. Energy Mater. Sol. Cells* **67**, 179 (2001).
100. Bao D, Gu H, Kuang A, *Thin Solid Films* **312**, 37 (1998).
101. Noh Y, Cheng X, Sirringhaus H, Sohn J, Welland M, Kang D, *Appl. Phys. Lett*. **91**, 043109 (2007).
102. Cranton W *et al*., *Thin Solid Films*, **515**, 8534 (2007).
103. Hwang J-H, Edwards D, Kammler D, Mason T, *Solid St. Ionics* **129**, 135 (2000).
104. Kröger F, *The Chemistry of Imperfect Crystals*, North-Holland Publishing Company, Amsterdam (1964).
105. Yamada N, Yasui I, Shigesato Y, Li H, Ujihira Y, Nomura K, *Jpn. J. Appl. Phys*. **38** 2856 (1999).
106. Zhang S, Wei S, *Appl. Phys. Lett*. **80**, 1376 (2002).
107. Kittel C, *Introduction to Solid State Physics*, John Wiley & Sons, Inc., 4th edn (1971).
108. Ellmer K, *J. Phys. D: Appl. Phys.* **34/21**, 3097 (2001).
109. Kulkarni A, Knickerbocker S, *J. Vac. Sci. Techol*. **A 14**, 1709 (1996).
110. Zhang D, Ma H, *Appl. Phys*. **A 62**, 487 (1996).
111. Zawadzki W, in Moss T, Ed, *Handbook on Semiconductors*, North-Holland, Amsterdam (1982).
112. Bardeen J, Shockley W, *Phys. Rev*. **80**, 72 (1950).
113. Orton J, *The Story of Semiconductors*, Oxford University Press (2004).
114. Petritz R, *Phys. Rev*. **104**, 1508 (1956).

115. Seto J, *J. Appl. Phys*. **46**, 5247 (1975).
116. Baccarani G, Ricco B, Spadini G, *J. Appl. Phys*. **49**, 5565 (1978).
117. Erginsoy C, *Phys. Rev*. **79**, 1013 (1950).
118. Seeger K, *Semiconductor Physics*, Springer-Verlag Berlin, 3rd edn (1985).
119. Young D, Coutts T, Kaydanov V, Gilmore A, Mulligan W, *J. Vac Sci. Technol*. **A 18**, 2978 (2000).
120. Shigesato Y, Paine D, *Appl. Phys. Lett*. **62**, 1268 (1993).
121. Ohhata Y, Shinoki F, Yoshida S, *Thin Solid Films* **59**, 255 (1979).
122. Pisarkiewicz T, Zakrzewska K, Leja E, *Thin Solid Films* **174**, 217 (1989).
123. Minami T, Sato H, Ohashi K, Tomofuji T, Takata S, *Journal of Crystal Growth* **117**, 370 (1992).
124. Brehme S *et al*., *Thin Solid Films* **342**, 167 (1999).
125. Pankove J, *Optical Processes in Semiconductors*, Prentice-Hall, Englewood Cliffs, NJ (1971).
126. Sapoval B, Hermann C, *Physics of Semiconductors*, Springer-Verlag New York (1995).
127. Weiher R, Ley R, *J. Appl. Phys*. **37**, 299 (1966).
128. Walsh A *et al*., *Phys. Rev. Lett*., **100**, 167402 (2008).
129. Keil T, *Phys. Rev*. **144**, 582 (1966).
130. Sernelius B, Berggren K-F, Jin Z-C, Hamberg I, Granqvist C, *Phys. Rev*., B **37**, 10244 (1988).
131. Burstein E, *Phys. Rev*. **93**, 632 (1954).
132. Roth A, Webb J, Williams D, *Solid State Comm*. **39**, 1269 (1981).
133. Dressel M, Grüner G, *Electrodynamics of Solids*, Cambridge University Press (2002).
134. Bleaney B, Bleaney B, *Electricity and Magnetism*, Oxford University Press (1965).
135. Heavens O, *Optical Properties of Thin Solid Films*, Dover Publications, Inc. New York (1991).
136. Macleod H, *Thin-Film Optical Filters*, Macmillan Publishing Co., New York, 2nd edn (1986).
137. Coutts T, Young D, Li X, *MRS Bulletin* **25/8**, 58 (2000).
138. Gordon R, *Mater. Res. Soc. Symp. Proc*. **426**, 419 (1996).
139. Gordon R, *MRS Bulletin* **25/8**, 52 (2000).
140. Wilk G, Wallace R, Anthony J, *J. Appl. Phys*. **89**, 5243 (2001).
141. Gessert T, Yoshida Y, Fesenmaier C, Coutts T, *J. Appl. Phys*. **105**, 083547 (2009).
142. Gupta Y, Liers H, Woods S, Young S, DeBlasio R, Mrig L, *Proc. 16th IEEE Photovoltaic Specialists Conference*, p. 1092, (1982).
143. van den Berg R *et al*., *Sol. Energy Mater. Sol. Cells* **31**, 253 (1993).
144. Carlson D, *IEEE Trans. ED-* **24**, 449 (1977).
145. Raniero L *et al*., *Thin Solid Films* **511-512**, 295 (2006).
146. Ellis Jr. F, Delahoy A, *Sol. Energy Mater*. **13**, 109 (1986).
147. Janki S, Baumgartner F, Ellert C, Feitknecht L, *Proc. 20th European Photovoltaic Solar Energy Conference*, p. 1620 (2005).
148. Schade H, Smith Z, *J. Appl. Phys*. **57**, 568 (1985).
149. Yablonovitch E, Cody G, *IEEE Trans. Electron. Devices* **ED-29**, 300 (1982).
150. Lechner P, Geyer R, Schade H, Rech B, Müller J, *Proc. 28th IEEE Photovoltaic Specialists Conference*, p. 861 (2000).
151. Hegedus S, Kaplan R, *Prog. Photovolt.: Res. Appl*. **10**, 257 (2002).
152. Romeo N, Bosio A, Canevari, Podestà A, *Solar Energy* **77**, 795 (2004).
153. Oehlstrom K, Sittinger V, Friedmann S, Abken A, Reineke-Koch R, Parisi J, *Proc. 17th European Photovoltaic Solar Energy Conference*, p. 1172 (2001).
154. Wu X, *Solar Energy* **77**, 803 (2004).
155. Wu X *et al*., *Proc. 28th IEEE Photovoltaic Specialists Conference*, p. 470 (2000).
156. Gupta A, Compaan A, *Appl. Phys. Lett*. **85**, 684 (2004).
157. Wu X *et al*., *Prog. Photovolt.: Res. Appl*. **14**, 471 (2006).
158. Grätzel M, *J. Photochem. & Photobiol. C* **4**, 145 (2003).
159. Kroon J *et al*., *Prog. Photovolt.: Res. Appl*. **15**, 1 (2007).

160. Lee S *et al*., *J. Phys. Chem. C* **113**, 7443 (2009).
161. Yu G, Gao J, Hummelen J, Wudl F, Heeger A, *Science* **270**, 1789 (1995).
162. Li G *et al*., *Nature Mater*. **4**, 864 (2005).
163. Haug, F.-J, Söderström T, Dominé D, Ballif C, *Mater. Res. Soc. Symp. Proc*. **1153**, 1153–A13 (2009).
164. Despeisse M *et al*., *Proc. SPIE* **7409**, 7409B–1 (2009).
165. Söderström T, Haug F.-J, Terrazzoni-Daudrix V, Ballif C, *J. Appl. Phys*. **107**, 014507 (2010).
166. Nakada T, *Thin Solid Films* **480-481**, 419 (2005).
167. Plá J, Tamasi M, Centurioni E, Rizzoli R, Summonte C, Durán J, *Proc. 17th European Photovoltaic Solar Energy Conference*, p. 3027 (2001).
168. Taguchi M *et al*., *Proc. 31st IEEE Photovoltaic Specialists Conference*, p. 866 (2005).
169. Taguchi M *et al*., *Proc. 24th European Photovoltaic Solar Energy Conference*, p. 1690 (2009).
170. Contreras M *et al*., *1st WCPEC*, p. 68 (1994).
171. Rau U, Schmidt M, *Thin Solid Films* **387**, 141 (2001).
172. Ishizuka S *et al*., *Sol. Energy Mater. Sol. Cells* **87**, 541 (2005).
173. Wu X *et al*., *NCPV Photovoltaics Program Review*, AIP Conf. Proc. **462**, 37 (1999).
174. Li X *et al*., *NCPV Photovoltaics Program Review*, AIP Conf. Proc. **462**, 230 (1999).
175. McCandless B, Dobson K, *Solar Energy* **77**, 839 (2004).
176. Feldman S *et al*., *Proc. 31st IEEE Photovoltaic Specialists Conference*, p. 271 (2005).
177. Mamazza R *et al*., *Proc. 31st IEEE Photovoltaic Specialists Conference*, p. 283 (2005).
178. Young D, Moutinho H, Yan Y, Coutts T, *J. Appl. Phys*. **92**, 310 (2002).
179. von Roedern B, Bauer G, *Mater. Res. Soc. Symp. Proc*. **557**, 761 (1999).
180. von Roedern B, *Photovoltaic Materials, Physics of* in *Encyclopedia of Energy*, Vol. **5**, Elsevier (2004).
181. Kurtz S, Gordon R, *Sol. Energy Mater*. **15**, 229 (1987).
182. Matsui T, Fujibayashi T, Sato A, Sonobe H, Kondo M, *Proc. 20th European Photovoltaic Solar Energy Conference*, p. 1493 (2005).
183. Berginski M, Das C, Doumit A, Hüpkes J, Rech B, Wuttig M, *Proc. 22nd European Photovoltaic Solar Energy Conference*, p. 2079 (2007).
184. Delahoy A *et al*, Photovoltaic Cell and Module Technologies II, edited by Bolko von Roedern, Alan E. Delahoy, Proc. of SPIE Vol. **7045**, 704506, (2008).
185. Anna Selvan J, Delahoy A, Guo S, Li Y-M, *Sol. Energy Mater. Sol. Cells* **90**, 3371 (2006).
186. Yamamoto K *et al*., *Solar Energy* **77**, 939 (2004).
187. Dominé D, Bailat J, Steinhauser J, Shah A, Ballif C, *Proc. IEEE 4th World Conference on Photovoltaic Energy Conversion*, p. 1465 (2006).
188. Buehlmann P *et al*., *Appl. Phys. Lett*. **91**, 143505 (2007).
189. Deckman H, Wronski C, Witzke H, Yablonovitch E, *Appl. Phys. Lett*. **42**, 968 (1983).
190. Kothandaraman C, Tonon T, Huang C, Delahoy A, *Mater. Res. Soc. Symp. Proc*. **219**, 475 (1991).
191. Jones S, Tsu D, Liu T, Steele J, Capangpangan R, Izu M, *Mater. Res. Soc. Symp. Proc*. **808**, 599 (2004).
192. Hedler A *et al*., *Proc. 34th IEEE Photovoltaic Specialists Conference*, p. 1102 (2009).
193. Berger O, Inns D, Aberle A, *Sol. Energy Mater. Sol. Cells* **91**, 1215 (2007).
194. Minemoto T *et al*., *Sol. Energy Mater. Sol. Cells* **67**, 83 (2001).
195. Schmid D, Ruckh M, Schock H, *Sol. Energy Mater. Sol. Cells* **41/42**, 281 (1996).
196. Minemoto T, Hashimoto Y, Satoh T, Negami T, Takakura H, Hamakawa Y, *J. Appl. Phys*. **89**, 8327 (2001).
197. Minemoto T *et al*., *Sol. Energy Mater. Sol. Cells* **75**, 121 (2003).
198. Schade H, Smith Z, Thomas III J, Catalano A, *Thin Solid Films* **117**, 149 (1984).
199. Sato K, Matsui Y, Adachi K, Gotoh Y, Hayashi Y, Nishimura H, *Proc. 23rd IEEE Photovoltaic Specialists Conference*, p. 855 (1993).

200. Wanka H, Schubert M, Lotter E, *Sol. Energy Mater. Sol. Cells* **41/42**, 519 (1996).
201. Wallinga J, Arnoldbik W, Vredenberg A, Schropp R, van der Weg W, *J. Phys Chem. B* **102**, 6219 (1998).
202. Hegedus S, *J. Appl. Phys*. **92**, 620 (2002).
203. Nuruddin A, Abelson J, *Thin Solid Films* **394**, 49 (2001).
204. Wanka H, Bilger G, Schubert M, *Appl. Surf. Sci*. **93**, 339 (1996).
205. Böhmer E, Siebke, Rech B, Beneking C, Wagner H, *Mater. Res. Soc. Symp. Proc*. **426**, 519 (1996).
206. An I, Lu Y, Wronski C, Collins R, *Appl. Phys. Lett*. **64**, 3317 (1994).
207. Smole F, Topič M, Furlan J, *J. Non-Cryst. Solids* **194**, 312 (1996).
208. Hegedus S, Kaplan R, Ganguly G, Wood G, *Proc. 28th IEEE Photovoltaic Specialists Conference*, p. 728 (2000).
209. Ganguly G *et al*., *Appl. Phys. Lett*., **85**, 479 (2004).
210. Shanthi E, Banerjee A, Dutta V, Chopra K, *Thin Solid Films* **71**, 237 (1980).
211. Sato K, Gotoh Y, Hayashi Y, Nishimura H, *Proc. 21st IEEE Photovoltaic Specialists Conference*, p. 1584 (1990).
212. Gordon R, Proscia J, Ellis F Jr., Delahoy A, *Sol. Energy Mater*. **18**, 263 (1989).
213. Sakai H *et al*., *J. Non-Cryst. Solids* **115**, 198 (1989).
214. Sakai H, Yoshida T, Hama T, Ichikawa Y, *Jpn. J. Appl. Phys*. **29**, 630 (1990).
215. Ichikawa Y, Fujikake S, Yoshida T, Hama T, Sakai H, *Proc. 21st IEEE Photovoltaic Specialists Conference*, p. 1475 (1990).
216. Löffler J, van Mol A, Grob F, Rath J, Schropp R, *Proc. 19th European Photovoltaic Solar Energy Conference*, p. 1493–1496 (2004).
217. Taneda N, Oyama T, Sato K, *Tech. Digest PVSEC-17*, 309 (2007).
218. Nakada T, Ohkubo Y, Murakami N, Kunioka A, *Jpn. J. Appl. Phys*. **34**, 3623 (1995).
219. Souletie P, Wessels B, *J. Mater. Res*. **3**, 740 (1988).
220. Major S, Chopra K, *Sol. Energy Mater. Sol. Cells* **17**, 319 (1988).
221. Yamada A, Wenas W, Yoshino M, Konagai M, Takahashi K, *Proc. 22nd IEEE Photovoltaic Specialists Conference*, p. 1236, (1991).
222. Benagli S *et al*., *Proc. 22nd European Photovoltaic Solar Energy Conference*, p. 2177 (2007).
223. Kluth O *et al*., *Thin Solid Films* **351**, 247 (1999).
224. Müller J *et al*., *Thin Solid Films* **392**, 327 (2001).
225. Hüpkes J *et al*., *Sol. Energy Mater. Sol. Cells* **90**, 3054 (2006).
226. Müller J *et al*., *Thin Solid Films* **442**, 158 (2003).
227. Delahoy A, Ellis F, Kothandaraman C, Schade H, Tonon T, Weakliem H, Semi-Annual Technical Progress Report, Subcontract No ZB-7-06003-1 (1989).
228. Zhu H, Bunte E, Hüpkes J, Siekmann H, Huang S, *Thin Solid Films* **517**, 3161 (2009).
229. Guo S, Sahoo L, Sosale G, Delahoy A, in Photovoltaic Cell and Module Technologies, von Roedern B, Delahoy A, (eds), *Proc. SPIE* **6651**, 66510B (2007).
230. Groenen R *et al*., *Thin Solid Films* **392**, 226 (2001).
231. Sang B, Yamada A, Konagai M, *Jpn. J. Appl. Phys*. **37**, L206 (1998).
232. Sang B, Dairiki K, Yamada A, Konagai M, *Jpn. J. Appl. Phys*. **38**, 4983 (1999).
233. Dekker T, Metselaar J, Schlatmann R, Stannowski B, van Swaaij R, Zeman M, *Proc. 20th European Photovoltaic Solar Energy Conference*, p. 1517, (2005).
234. Bennett J, Mattsson L, *Introduction to surface roughness and scattering*, Optical Society of America, Washington D.C. (1989).
235. Lechner P, Geyer R, Schade H, Rech B, Kluth O, Stiebig H, *Proc. 19th European Photovoltaic Solar Energy Conference*, p. 1591, (2004).
236. Elson J, Rahn J, Bennett J, *Appl. Opt*. **22**, 3207 (1983).
237. Kluth O, Zahren C, Stiebig H, Rech B, Schade H, *Proc. 19th European Photovoltaic Solar Energy Conference*, p. 1587 (2004).

238. Sosale G, M. Eng. Dissertation, McGill University (2007).
239. Hüpkes J *et al*., *Proc. 24th European Photovoltaic Solar Energy Conference*, p. 2766, (2009).
240. Hegedus S, Deng X, *Proc. 25th IEEE Photovoltaic Specialists Conference*, p. 1061 (1996).
241. Poruba A *et al*., J Appl. Phys. **88**, 148 (2000).
242. Springer J, Rech B, Reetz W, Muller J, Vanecek M, *Sol. Energy Mater. Sol. Cells* **85**, 1 (2005).
243. Leblanc F, Perrin J, Schmitt J, *J. Appl. Phys*. **75** 1074 (1994).
244. Berginski M *et al*., *Proc. 21st European Photovoltaic Solar Energy Conference*, p. 1539 (2006).
245. Müller J, Rech B, Springer J, Vanecek M, *Solar Energy* **77**, 917 (2004).
246. Steinhauser J, *et al*., *Proc. 20th European Photovoltaic Solar Energy Conference*, p. 1608 (2005).
247. Bailat J *et al*., *IEEE 4th World Conference Photovoltaic Energy Conversion*, p. 1533, (2006).
248. Python M *et al*., *J. Non-Cryst. Solids* **354**, 2258 (2008).
249. Ballif C *et al*., *Proc. 21st European Photovoltaic Solar Energy Conference*, p. 1552 (2006).
250. Addonizio M, Manoj R, Usatii I, *Proc. 22nd European Photovoltaic Solar Energy Conference*, p. 2129 (2007).
251. Kluth O, Schöpe G, Hüpkes J, Agashe C, Muller J, Rech B, *Thin Solid Films* **442**, 80, (2003).
252. Thornton J, *Ann. Rev. Mater. Sci*. **7**, Annual Reviews, Inc., Palo Alto, CA 239 (1977).
253. Berginski M *et al*., *J. Appl. Phys*. **101**, 074903 (2007).
254. Ruske F, Jacobs C, Sittinger V, Szyszka B, Werner W, *Thin Solid Films* **515**, 8695 (2007).
255. Berginski M, Hűpkes J, Reetz W, Rech B, Wuttig M, *Thin Solid Films* **516**, 5836 (2008).
256. Rech B *et al*., *Thin Solid Films* **511–512**, 548 (2006).
257. Sheng S *et al*., *Proc. 24th European Photovoltaic Solar Energy Conference*, p. 2850 (2009).
258. Amin N, Isaka T, Yamada A, Konagai M, *Sol. Energy Mater. Sol. Cells* **67**, 195 (2001).
259. Runyan W, *Semiconductor Measurements and Instrumentation*, McGraw-Hill Book Company, New York (1975).
260. van der Pauw L, *Philips Research Reports* **13**, 1 (1958).
261. http://www.eeel.nist.gov/812/hall.html.
262. Swanepoel R, *J. Phys. E* **16**, 1214 (1983).
263. Ohring M, *Materials Science of Thin Films*, Academic Press, San Diego, 2nd edn (2002).
264. Theeten J, Aspnes D, Ellipsometry in Thin Film Analysis, in Huggins R, Bube R, Vermilyea D, Eds, *Annual Review of Materials Science* **11**, 97 Annual Reviews Inc., Palo Alto (1981).
265. Jellison Jr. G, *Thin Solid Films* **234**, 416 (1993).
266. Synowicki R, *Thin Solid Films* **313–314**, 394 (1998).
267. Fujiwara H, Kondo M, *Phys. Rev*. **B71**, 075109 (2005).
268. Pflug A, Sittinger V, Ruske F, Szyszka B, Dittmar G, *Thin Solid Films* **455–456**, 201 (2004).
269. Jackson W, Amer N, Boccara A, Fournier D, *Appl. Opt*. **20**, 1333 (1981).
270. Mizuhashi M, Gotoh Y, Adachi K, *Jpn. J. Appl. Phys*. **27**, 2053 (1988).
271. Stover J, *Optical Scattering: measurement and analysis*, SPIE, Bellingham, Washington, USA, 2nd edn, (1995).
272. Smith D, *Thin-Film Deposition: principles and practice*, McGraw-Hill, New York (1995).
273. Hoshi Y, Kato H, Funatsu K, *Thin Solid Films* **445**, 245 (2003).
274. Wells O, *Scanning Electron Microscopy*, McGraw-Hill, New York (1974).
275. Sieber I *et al*., *Thin Solid Films* **330**, 108 (1998).
276. Cullity B, *Elements of X-ray Diffraction*, Addison-Wesley, Reading MA, 2nd edn, (1978).
277. X-ray powder diffraction data is available from the International Centre for Diffraction Data (ICDD, formerly JCPDS), Newtown Square, PA, USA.
278. Mwabora J *et al*., *Thin Solid Films* **516**, 3841 (2008).
279. Klein A *et al*., *Thin Solid Films* **518**, 1197 (2009).
280. Sugiyama K, Ishii H, Ouchi Y, *J. Appl. Phys*. **87**, 295 (2000).

281. Jansen K, Delahoy A, *Thin Solid Films* **423/2**, 153 (2003).
282. Carlson D *et al*., *Prog. Photovolt.: Res. Appl*. **11**, 377 (2003).
283. Osterwald C, McMahon T, del Cueto J, *Sol. Energy Mater. Sol. Cells* **79**, 21 (2003).
284. McMahon T, *Prog. Photovolt.: Res. Appl*. **12** (2,3), 235 (2004).
285. Mon G, *20th IEEE Photovoltaic Specialists Conference*, 108 (1988).
286. Osterwald C, Adelstein J, del Cueto J, Kroposki B, Trudell D, Moriaty T, 32nd IEEE Photovoltaic Specialists Conference, p. 2085 (2006).
287. Goto K, Kawashima T, Tanabe N, *Sol. Energy Mater. Sol. Cells* **90**, 3251 (2006).
288. Kempe M, Jorgensen G, Terwilliger K, McMahon T, Kennedy C, Borek T, *IEEE 4th World Conference on Photovoltaic Energy Conversion (WCPEC-4)*, p. 2160 (2006).
289. Jorgensen G *et al*., *Sol. Energy Mater. Sol. Cells* **90**, 2739 (2006).
290. Wennerberg J, Kessler J, Stolt L, *16th European Photovoltaic Solar Energy Conference*, p. 309 (2000).
291. Powalla M *et al*., *Thin Solid Films* **431–432**, 523 (2003).
292. Wennerberg J, Kessler J, Stolt L, *Sol. Energy Mater. Sol. Cells* **75**, 47 (2003).
293. Kessler J, Norling J, Lundberg O, Wennerberg J, Stolt L, *16th European Photovoltaic Solar Energy Conference*, p. 775 (2000).
294. Nakagawara O, Kishimoto Y, Seto H, Koshido Y, Yoshino Y, Makino T, *Appl. Phys. Lett*. **89**, 091904 (2006).
295. Greiner D, Papthanasiou N, Pflug A, Ruske F, Klenk R, *Thin Solid Films* **517**, 2291 (2009).
296. Sundaramoorthy R *et al*., *Proc. SPIE, Reliability of Photovoltaic, Cells, Modules, Components, and Systems II*. Edited by Dhere N, Wohlgemuth J, Ton D, **7412**, 74120J (2009).
297. Olsen L, Gross M, Kundu S, *33rd IEEE Photovoltaic Specialists Conference*, p. 166 (2008).
298. Steinhauser J *et al*., *phys. stat. sol. (a)* **205**, 1983 (2008).
299. Minami T, Kuboi T, Miyata T, Ohtani Y, *phys. stat. sol.(a)* **205**, 255 (2008).
300. Tohsophon T *et al*., *Thin Solid Films* **511–512**, 673 (2006).
301. Owen J, Hüpkes J, Nießen L, Zastrow U, Beyer W, *24th European Photovoltaic Solar Energy Conference*, p. 2274 (2009).
302. Beyer W, Hüpkes J, Stiebig H, *Thin Solid Films* **516**, 147 (2007).
303. Guillén C, Herreo J, *Surf. Coat. Technol*. **201**, 309 (2006).
304. Matsumoto F, Tani M, Enomoto T, *US Patent 5,105,291* (1992).
305. Cording C, *Proc. SPIE, Photovoltaic Cell and Module Technologies II*, **7045**, 704507–1 (2008).
306. Oyama T, Kambe M, Taneda N, Masumo K, *Mater. Res. Soc. Symp. Proc*. **1101**, KK02 (2008).
307. Kambe M *et al*., *Proc. 24th European Photovoltaic Solar Energy Conference*, p. 2290 (2009).
308. Yan M, Lane M, Kannewurf C, Chang R, *Appl. Phys. Lett*. **78**, 2342 (2001).
309. Jin S *et al*., *Chem. Mater*. **20**, 220 (2008).
310. Kiliç Ç, Zunger A, *Phys. Rev. Lett*. **88**, 095501 (2002).
311. Narushima S, Orita M, Hirano M, Hosono H, *Phys. Rev*. **B 66**, 035203 (2002).
312. Medvedeva J, *Phys. Rev. Lett*. **97**, 086401 (2006).
313. Koida T, Fujiwara H, Kondo M, *Jpn. J. Appl. Phys*. **46**, L685 (2007).
314. Minami T, Oohashi K, Takata S, *Thin Solid Films* **193/194**, 721 (1990).
315. Lee K *et al*., *Appl. Phys. Lett*. **91**, 241911 (2007).
316. Ruske F, Roczen M, Hüpkes J, Gall S, Rech B, *Proc. 24th European Photovoltaic Solar Energy Conference*, p. 2353 (2009).
317. Lorenz M *et al*., *Solid-State Electronics* **47**, 2205 (2003).
318. Hashimoto R, Abe Y, Nakada T, *Appl. Phys. Expr*. **1**, 015002 (2008).
319. van Hest M, Dabney M, Perkins J, Ginley D, Taylor M, *Appl. Phys. Lett*. **87**, 032111 (2005).
320. Bowers J, Upadhyaya H, Calnan S, Hashimoto R, Nakada T, Tiwari A, *Prog. Photovolt.: Res. Appl*. **17**, 265 (2009).

321. Anna Selvan J, Li Y-M, Guo S, Delahoy A, *Proc. 19th European Photovoltaic Solar Energy Conference*, (2004).
322. Nakada T, Miyano T, Hashimoto R, Kanda Y, Mise T, *Proc. 22nd European Photovoltaic Solar Energy Conference*, p. 1870 (2007).
323. Bowers J, Upadhyaya H, Nakada T, Tiwari A, *Solar Energy Materials and Solar Cells* (2009). *In press*
324. Moulin E, Sukmanowski J, Schulte M, Gordijn A, Royer F, Stiebig H, *Thin Solid Films* **516**, 6813 (2008).
325. Haug F-J, Söderström T, Cubero O, Terrazzoni-Daudrix V, Ballif C, *J. Appl. Phys*. **104**, 064509 (2008).
326. Morfa A, Rowlen K, Reilly III T, Romero M, van de Lagermaat J, *Appl. Phys. Lett*. **92**, 013504 (2008).
327. Tanabe K, Nakayama K, Atwater H, *Proc. 33rd IEEE Photovoltaic Specialists Conference*, p. 129 (2008).
328. Fortunato E, Ginley D, Hosono H, Paine D, *MRS Bulletin* **32**, 242 (2007).
329. Kayes B, Atwater H, Lewis N, *J. Appl. Phys*. **97**, 114302 (2005).
330. Tsakalakos L, Balch J, Fronheiser J, Korevaar B, Sulima O, Rand J, *Appl. Phys. Lett*. **91**, 233117 (2007).
331. Zhu J *et al*., *Nano Lett*. **9**, 279 (2009).
332. Vanecek M *et al*., *Proc. 24th European Photovoltaic Solar Energy Conference*, p. 2286 (2009).
333. Furubayashi Y, Hitosugi T, Hasegawa T, *Appl. Phys. Lett*. **88**, 226103 (2006).
334. Furubayashi Y *et al*., *Thin Solid Films* **496**, 157 (2006).
335. Gillispie M, van Hest M, Dabney M, Perkins J, Ginley D, *J. Appl. Phys*. **101**, 033125 (2007).
336. Hitosugi T *et al*., *Appl. Phys. Lett*. **90**, 212106 (2007).
337. Junghähnel M, Heimke B, Hartung U, Kopte T, *Proc. 24th European Photovoltaic Solar Energy Conference*, p. 2641 (2009).
338. Ebbesen T, Lezec H, Hiura H, Bennett J, Ghaemi H, Thio T, *Nature* **382**, 54 (1996).
339. Hata K, Futaba D, Mizuno K, Namai T, Yumura M, Iijima S, *Science* **306**, 1362 (2004).
340. Barnes T *et al*., *Phys. Rev*. B **75**, 235410 (2007).
341. Du Pasquier A, Unalan H, Kanwai A, Miller S, Chhowalla M, *Appl. Phys. Lett*. **87**, 203511 (2005).
342. Ulbricht R *et al*., *Sol. Energy Mater. Sol. Cells* **91**, 416 (2007).
343. Contreras M *et al*., *Proc. IEEE 4th World Conference*, **1**, 428 (2006).
344. Barnes T *et al*., *Appl. Phys. Lett*. **90**, 243503 (2007).
345. Wang X, Zhi L, Müllen K, *Nano Lett*. **8**, 323 (2008).
346. Wu J, Becerril H, Bao Z, Liu Z, Chen Y, Peumans P, *Appl. Phys. Lett*. **92**, 263302 (2008).
347. Tung V *et al*., *Nano Lett*. **9**, 1949 (2009).
348. Lee J-Y, Connor S, Cui Y, Peumans P, *Nano Lett*. **8**, 689 (2008).
349. Kamiya T, Hosono H, *NPG Asia Mater*. **2**, 15 (2010).
350. Sundaramoorthy R *et al*., *Proc. 34th IEEE Photovoltaic Specialists Conference*, 001576 (2009).
351. Nomura K, Ohta H, Takagi A, Kamiya T, Hirano M, Hosono H, *Nature* **432**, 488 (2004).
352. Chiang H, Wager J, Hoffman R, Jeong J, Keszler D, *Appl. Phys. Lett*. **86**, 013503 (2005).
353. Leenheer A, Perkins J, van Hest M, Berry J, O'Hayre R, Ginley D, *Phys. Rev*. B **77**, 115215 (2008).
354. Martins R *et al*., *J. Non-Cryst. Solids* **352**, 1471 (2006).
355. Betz U, Marthy J, Atamny F, *Proc. 46th SVC Ann. Tech. Conf*., p. 175 (2003).
356. Bellingham J, Phillips W, Adkins C, *J. Phys.: Condens. Matter* **2**, 6207 (1990).
357. Tominaga K *et al*., *J. Vac. Sci. Technol*. A **23**, 401 (2005).
358. Phillips J *et al*., *Appl. Phys. Lett*., **67**, 2246 (1995).
359. Jung Y, Seo J, Lee D, Jeon D, *Thin Solid Films* **445**, 63 (2003).

18

Measurement and Characterization of Solar Cells and Modules

Keith Emery

US National Renewable Energy Laboratory

18.1 INTRODUCTION

Methods of assessing the performance of photovoltaic (PV) cells and modules are described in this chapter. The performance of cells and modules can be described by their current versus voltage ($I-V$) and spectral responsivity versus wavelength ($S(\lambda)$) characteristics. Measurement equipment, procedures, and artifacts are discussed for $I-V$ and $S(\lambda)$. The most common performance indicator is the PV efficiency under standard reporting conditions (temperature, spectral irradiance, total irradiance). The efficiency is the maximum electrical power divided by the total irradiance. Procedures for accurately determining the efficiency or the maximum power with respect to reference conditions are described. Alternatives to the standard peak watt rating and how they compare with actual field performance are discussed. Because photovoltaics must operate for 20–30 years, with a degradation of less than 1% per year, procedures for assessing the durability of PV modules are also discussed.

18.2 RATING PV PERFORMANCE

A variety of performance indicators have been employed by the photovoltaic community to rate the performance of PV cells and modules [1–4]. Domestic and international consensus standards have been adopted to rate the performance of PV cells and modules in terms of the output power, or equivalently, their efficiency with respect to standard reporting conditions defined by a temperature, spectral irradiance, and total irradiance [5–15]. Modules and systems are rated by their peak power under standard reporting conditions because manufacturers sell and customers purchase PV modules and systems according to price per watt of power produced. Modules and systems are sometimes evaluated with respect to their energy produced with respect to a typical meteorological year, reference day, or at a given site. For convenience and ease of comparison among technologies, the

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

energy is sometimes divided by the peak watt rating to give a performance ratio. Other performance indicators may be more appropriate for niche markets, such as aesthetics for building-integrated PV, liters per day for water pumping, or low light-level operation for consumer electronics [4].

The actual output of a PV module or system in the field is a function of orientation, total irradiance, spectral irradiance, wind speed, air temperature, soiling, and various system-related losses. Various module and system rating methods attempt to ensure that the actual performance is comparable to the rated performance to keep the resulting level of customer satisfaction high.

18.2.1 Standard Reporting Conditions

The PV performance in terms of standard reporting conditions (SRC) or standard test conditions (STC) is commonly expressed in terms of a peak watt rating or an efficiency. At the research level, an internationally accepted set of standard reporting conditions is essential to prevent the researcher from adjusting the reporting conditions to maximize the efficiency. The procedures for measuring the performance with respect to SRC must be quick, easy, reproducible, and accurate for the research cell fresh out of the deposition system or for the module on a factory floor with production goals. The PV conversion efficiency (η) is calculated from the measured maximum or peak PV power (P_{max}), device area (A), and total incident irradiance (E_{tot}):

$$\eta = \frac{P_{\max}}{E_{tot} A} 100. \tag{18.1}$$

Parameters which directly influence the measurement or calculation of the efficiency, and therefore must be well controlled and well-defined, are the area of the device, the spectrum and intensity of incident radiation which determine E_{tot}, and the temperature of the device. The term "reporting," rather than "reference" or "test," is used because in practice a test is performed at other than SRC and then corrected to be equivalent to being measured at SRC. ASTM standards use SRC while the international IEC standards use STC. The currently accepted SRC for rating the performance of cells and modules are summarized in Table 18.1 [15–17]. The direct and global air

Table 18.1 Standard reference conditions (SRC) for rating photovoltaic cells, modules and systems. The irradiance listed is the reference irradiance and the reference spectrum may not integrate to this value

Application	Irradiance ($W\,m^{-2}$)	Reference spectrum	Temperature (°C)
Terrestrial nonconcentrating			
Cells	1000	Global [14, 15]	25 cell [5, 6, 11, 12]
Modules, systems	1000	Global [14, 15]	25 cell [7] or NOCT [7, 12]
Modules, systems	1000‡	prevailing	20 ambient
Terrestrial concentrating			
Cell†	>1000	Direct [10, 14]	25 cell [5]
Module	850 direct	prevailing	20 ambient [13]
Extraterrestrial	1366 [8], 1367 [16, 17]	AM0 [8, 16, 17]	25 [17], 28 cell [27]

†linear regression of power to project test conditions, $850\,W\,m^{-2}$ with a 5° field of view for concentrator systems.
‡At present, no consensus standards exist, *de facto* conditions are listed.

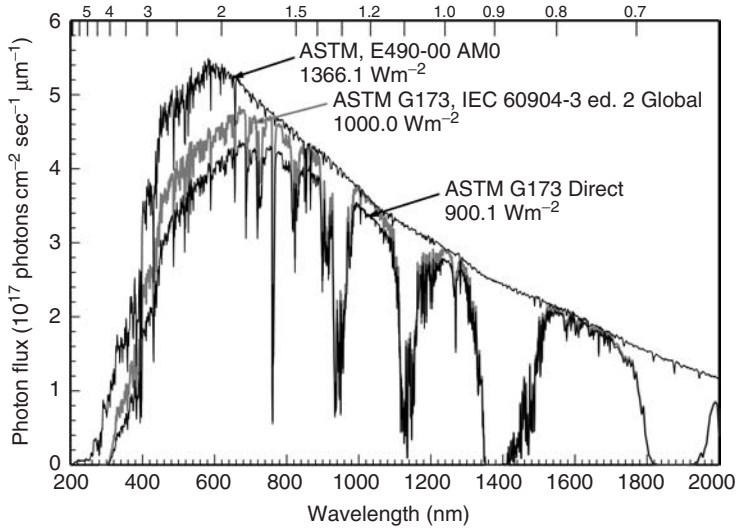


Figure 18.1 Global, Direct, and AM0 reference spectra [8, 14, 15]. The integrated irradiance values are not the reference irradiance values

mass 1.5 (AM1.5), and air-mass 0 (AM0) reference spectra are shown graphically in Figure 18.1 and in tabular form on the Web site http://rredc.nrel.gov/solar/spectra/am1.5/.

It should be noted that neither the direct reference spectrum nor the global reference spectrum actually integrates to the 1-sun reference total irradiance of $1000\,W\,m^{-2}$ [10, 14, 15, 18]. In 2008, the international terrestrial standards community revised the reference spectrum from a slightly modified version of ASTM G159 to a slightly modified version of ASTM G173 [10, 14, 15]. The differences are in the digits of precision versus significant digits and how the data past 4000 nm are treated [10, 14, 15]. The ASTM standards committee attempted to have the global spectrum ASTM G173 integrate to $1000\,W\,m^{-2}$ using the open-source spectral irradiance model SMARTS 2 developed by Gueymard [19, 20]. The global reference spectrum integrates to about $1000.4\,W\,m^{-2}$, and the direct reference spectrum integrates to about $900.1\,W\,m^{-2}$. The structure in the spectral irradiance is due to atmospheric absorption and scattering. The full width at half maximum (FWHM) bandwidth in the spectral irradiance at any given wavelength is approximately the difference in wavelength between adjacent points and is a function of the bandwidth of the measurement system.

The PV community has arbitrarily taken the term "one sun" to mean a total irradiance of $1000\,W\,m^{-2}$ [18]. In fact, the spectral irradiance of the ASTM G173 direct reference spectrum normalized to a "1-sun" value of $1000\,W\,m^{-2}$ exceeds the AM0 spectral irradiance in the infrared (IR), which is not physically possible without concentration. The term "global" refers to the spectral irradiance distribution on a 37°-tilted south-facing surface with a solar zenith angle of 48.19° (AM1.5). The term "direct" refers to the direct-normal component (5° field of view) of the global spectral irradiance distribution [14, 19]. The term AM1 or AM1.5 is often used to refer to standard spectra, but the relative optical air mass is a geometrical quantity and can be obtained by taking the secant of the zenith angle or the sine of the solar elevation (see Chapter 22). For AM1, the zenith angle is 0°. The relative optical air mass can be pressure corrected to an absolute air mass by multiplying by the barometric pressure and dividing by the sea-level pressure. In outer space, the pressure is zero so the absolute air mass is always zero. The fact that the reference spectrum

only approximates the "real-world" spectra at solar noon is unimportant as long as the differences between the photocurrents are the same for various PV technologies.

The technical basis for the direct spectra has recently been reexamined and found to have a diffuse component that is substantially greater than concentrators would normally encounter [21, 22]. Examination of the US solar radiation database has found that when the global-normal irradiance is near $1000\,W\,m^{-2}$, the direct-normal component is near $850\,W\,m^{-2}$ and not the $767\,W\,m^{-2}$ to which the direct standard spectrum integrates [21]. This difference has been attributed to a high turbidity [22]. This has not been a problem for concentrators in the past because of their relative insensitivity to the specific direct spectra [23, 24]. Recent high-efficiency structures such as the GaInP/GaAs/Ge triple-junction solar cell exhibit a significant difference in the efficiency between the global and direct reference spectra ($>10\%$). It has been suggested that the global reference spectrum may be a better spectrum than the direct reference spectrum to optimize concentrator cells for use in sunny climates [22, 25]. The previous ASTM G159 direct-beam reference spectrum may be more appropriate for regions with high aerosol content that have direct-beam resources above $5\,kW\,hm^{-2}d^{-1}$ such as Saudi Arabia [26]. The IEC TC82 Working Group 7 Standards Committee is considering what additional reference spectra should be considered for evaluating concentrator cells and modules. At present there is no consensus among the calibration labs around the world on alternatives to ASTM G173 direct and there is no IEC standard being drafted to address applications where concentrators are deployed in high aerosol climate. It is the author's recommendation that G159 direct be used as the reference spectrum for concentrator applications in sunny but high aerosol regions of the world.

The extraterrestrial spectral irradiance distribution at one astronomical unit distance from the sun is commonly referred to as the AM0 spectrum. International consensus standards for AM0 measurements have been developed [17]. Measurements of the total AM0 irradiance used by the aerospace community have varied from 1353 to $1372\,W\,m^{-2}$ [8, 16, 27–30]. Many groups still rely on the less accurate value of $1353\,W\,m^{-2}$ total AM0 irradiance [27, 28]. Recently, a new ASTM AM0 standard has been adopted that uses more accurate spectral irradiance measurements given in Figure 18.1 [8]. The best estimate for the solar "constant" is $1367\,W\,m^{-2}$ recommended by the World Radiation Center [16], or 1366.1 recommended by ASTM [8]. Both of these values were obtained from long-term monitoring of the solar irradiance with an active-cavity radiometer on the Solar Max and Nimbus 7 and other satellites [30]. Fortunately, the $1353\,W\,m^{-2}$ total AM0 irradiance, used by many groups for efficiency measurements and reporting purposes, does not enter into the spacecraft PV power measurements. This is because primary balloon or space-based AM0 reference cells are calibrated at whatever irradiance exists at the time of calibration, corrected for 1 astronomical unit distance from the sun. However ISO standard 15387 allows terrestrial based calibrations with respect to this synthetic AM0 spectral irradiance, as discussed in Section 18.3.3 and Chapter 22 [17].

A variety of definitions for cells and modules have been proposed [1, 5, 32, 33]. A module consists of several encapsulated, environmentally protected, electrically interconnected cells. The area of a cell is taken to be the total area of the space-charge region including grids and contacts. The standard definitions of cell area replace the term "space-charge region" with "frontal area," but this term does not adequately account for devices with multiple cells on a single substrate or superstrate. The area of a concentrator cell is based on the area that is designed to be illuminated [5]. This area is taken to be the area of the space-charge region minus the area of any peripheral bus bars or contacts. A submodule or minimodule is an unencapsulated module.

The PV efficiency (η) is inversely proportional to the area definition used (Equation 18.1). In fact, differences in the area definition often account for the greatest differences in reported efficiencies between various groups and values published in the literature [33, 34]. The largest differences occur when the so-called active area (total device area minus all area that is shaded or not active)

is used. The use of an active area in the efficiency neglects the trade-off between lower resistance losses and increased shading. Several thin film PV device structures do not have any shading losses, so the active and total areas are the same. To prevent an artificial increase in the efficiency, care must be taken to ensure that light outside the defined area cannot be collected by multiple internal reflections or incomplete electrical isolation. Incomplete electrical isolation is always a possibility when the device area is defined by the contact area and the junction area is larger than the device area. This effect increases as the cell size decreases. Larger perimeter-to-area ratios increase the potential effect of current being collected outside the defined area. This phenomenon is why a 1-cm^2 minimum area is required for inclusion in the *Progress in Photovoltaics* efficiency tables [35]. To be sure the region enclosed by the total area is the only active region, an aperture should be used [35].

At the module level, the total area including the frame is used. For prototype modules, where the frame design is less important than the encapsulation and cell interconnections, an aperture-area definition is often used. The aperture-area definition is the total area minus the frame area. This aperture area may be defined by opaque tape if there is no frame to eliminate the possibility of the module collecting current outside the defined aperture area by multiple internal reflections or light piping. Plastic tape may or may not be opaque enough in the IR depending on the tape and on the PV materials.

The most common method of performance rating for modules is the PV power conversion efficiency under SRC (Table 18.1). The power or peak watt rating on the module's nameplate is usually given with respect to SRC, as shown in Table 18.1 using a 25 °C module temperature. Unfortunately, prevailing conditions under natural sunlight do not commonly match nameplate conditions. The nameplate rating that the manufacturer assigns to a given module model number is often higher but rarely lower than the measured power output in the field [36–38]. While the nameplate rating is determined with the module temperature controlled at 25 °C, the actual power produced is often less than this because the module will typically run around 35 °C above air temperature on a sunny day. The temperature coefficient for the peak power is negative. The nameplate rating also does not include long-term degradation or system losses. System losses include the power-conditioning unit's efficiency, ability of the power conditioner to operate at the maximum PV power point, orientation, shading, resistance losses in the wiring, and mismatch in the power of different modules.

The nominal operating cell temperature (NOCT) is a rating designed to give information about the thermal qualities of a module and a more realistic estimate of the power in the field on a sunny day at solar noon. The NOCT of a module is a fixed temperature that the module would operate at if it is exposed to the nominal thermal environment (20 °C air temperature, 800 $W\,m^{-2}$ total irradiance, and a wind speed of 1 $m\,s^{-1}$) [7, 39]. Typical NOCT values found on the module nameplate range from 35 to 45 °C. The term "standard operating conditions" (SOC) is sometimes used for flat-plate or concentrator terrestrial modules operating at NOCT. The actual determination of the NOCT of a module with an uncertainty of less than ±2 °C has proved difficult because of difficulties in measuring the temperature of cells in an encapsulated module, uncertainties in the total irradiance, and secondary environmental effects such as wind direction, ground reflections, mounting, and electrical loading [39, 40]. The installed NOCT is up to 15 °C warmer for roof-mounted applications than a free-standing module depending on the stand-off distance between the module and roof [39, 40]. The module temperature can be calculated from the NOCT or installed NOCT and air temperature using

$$T = T_{air} + (\mathrm{NOCT} - 20\,^{\circ}\mathrm{C})E_{tot}/800\,\mathrm{W\,m^{-2}}. \tag{18.2}$$

A wind-speed correction can also be applied to Equation (18.2) [7, 39].

For a fair and meaningful comparison of efficiencies between technologies, the measurements should be performed after any initial degradation or transient behavior has stabilized. Commercial silicon modules have shown small changes in performance after the first hours of

operation [41, 42]. At the present time, all amorphous silicon PV technologies degrade when exposed to sunlight. Fortunately, this degradation stabilizes at a level of 80–90% of the initial value. Partial recovery occurs in the field during the summer when the higher module temperature leads to partial annealing or when amorphous silicon modules are annealed in the laboratory at 60–70 °C [43, 44]. The efficiency continues to decrease after 500 hours of light exposure at lower temperatures even if the light level is reduced [43–45]. For a fair and meaningful comparison of improvements in amorphous silicon module development, the performance at SRC is now reported after illumination of about $1000\,\mathrm{W\,m^{-2}}$, at a module back-surface temperature of nominally 50 °C, for at least 1000 hours, with a resistive load near $P_{\max}$, and low humidity [33, 34]. These conditions were chosen to approximate one year of outdoor exposure without the humidity or temperature cycling. Thin-film module stabilization procedures allow for shorter times based upon the power changing less than 2% over two consecutive periods of $43\,\mathrm{kWhm^{-2}}$. Other thin-film module technologies may undergo reversible and irreversible changes during the first few hours of light exposure [34, 46, 47].

18.2.2 Alternative Peak Power Ratings

A variety of groups have suggested and adopted alternative rating schemes to compare module and system performance between the various PV technologies. These schemes are based on measurements of a module's performance in the field and performing a regression analysis on the data. The site-specific power production is more relevant for bulk power generation than the power with respect to a particular theoretical reference spectrum and module operating temperature.

One popular method was adopted by Pacific Gas and Electric Company and the Photovoltaics for Utility-Scale Applications (PVUSA) project in California, USA, to rate and purchase PV systems. They perform a linear regression analysis on the actual measured system or module power produced (P), air temperature (T_a), wind speed (S), and total plane-of-array irradiance (E_{tot}) as measured with a pyranometer or radiometer:

$$P = P_{\max}(E_{\mathrm{tot}}, T_{\mathrm{a}}, S) = E_{\mathrm{tot}}(C_1 + C_2 E_{\mathrm{tot}} + C_3 T_{\mathrm{a}} + C_4 S), \tag{18.3}$$

where C_1, C_2, C_3, and C_4 are the regression coefficients [13, 37, 48]. The goal of performing a multiple-regression analysis on the measured power to a fixed set of environmental conditions is to accurately represent the average power output under clear-sky conditions near midday, or the energy on an hourly to yearly basis at a given site using typical meteorological year data sets. The power can be measured at the maximum direct-current (dc) power point, or on the dc side of the inverter, or at the alternating-current (ac) power out of the inverter. The last two power-measurement locations will include most system losses. This site-specific rating scheme takes into account the different thermal characteristics of modules and spectral sensitivities since it is not referenced to a standard spectrum or module temperature. The power rating is evaluated using Equation (18.3) at $T_a = 20\,^{\circ}\mathrm{C}$, $S = 1\,\mathrm{m\,s^{-1}}$, and $E_{tot} = 1000\,\mathrm{W\,m^{-2}}$ for flat-plate collectors. For concentrators, the direct-normal incidence sunlight within a 5° or 5.7° field of view of $850\,\mathrm{W\,m^{-2}}$ is used for E_{tot}. ASTM has recommended a wind speed of $4\,\mathrm{m\,s^{-1}}$ based on resource data. The difference between the fields of view is because an absolute-cavity radiometer has a 5° field of view and some less accurate but less expensive normal-incidence pyrheliometers have a 5.7° field of view.

The primary advantage of basing the reference temperature on the air temperature is that the different thermal characteristics of the module, array, and system are included in the rating, and the power rating is closer to what is actually observed. The different spectral conditions at the different sites are also accounted for by not referencing the performance to a fixed spectrum, but rather, referencing the power to the actual spectrum that was incident on the module. If a PV reference cell whose spectral response matches the cells in the module is used to measure E_{tot},

then the power would be with respect to a reference spectrum at all light levels. Spectral mismatch issues associated with E_{tot} measured with a thermal or spectrally matched detector are discussed further in Section 18.3.1.

18.2.3 Energy-based Performance Rating Methods

The SRC-based peak power rating gives the output under a unique set of rarely-occurring conditions. Despite its widespread acceptance, this peak power rating (i.e. maximum instantaneous watts) does not capture the differences among the plethora of flat-plate and concentrator module designs with different total irradiance, diffuse irradiance, spectral irradiance, and temperature sensitivities. Energy-based ratings (i.e., integrated power over time in kW h) capture the module performance in the "real" world. It is easy to integrate the measured PV power produced over a time interval to obtain the total energy produced compared with the incident energy. A variety of rating criteria exist besides the standard reference conditions listed in Table 18.1, depending on the application in Table 18.2.

The AM/PM method, proposed by ARCO/Siemens Solar Industries, attempts to rate a module in terms of the PV energy produced during a standard solar day with a given reference temperature and total irradiance distribution [49]. The AM/PM method is appealing because it is an energy-rating method that is not site-specific. A variation on the AM/PM energy-rating method was developed in which a regression analysis of the measured power and irradiance data to a nonlinear response function was summed over a standard day defined by a fourth-order polynomial [50].

A rating scheme based on the PV energy delivered over a standard day has been proposed for a small set of standard days [34–47]. These five days were obtained from the typical meteorological year database corresponding to a hot-sunny, cold-sunny, hot-cloudy, cold-cloudy, and a nice day [54, 55]. The meteorological data for the standard days include latitude, longitude, date, air temperature, wind speed, relative humidity, and direct, diffuse-horizontal, and global-normal irradiances. The direct-beam and plane-of-array spectral irradiances are then computed for hourly intervals throughout the day using a spectral model [56]. The model developed by Nann requires only the meteorological parameters listed in the standard days [4, 56]. Figure 18.2 shows the meteorological characteristics of the hot-sunny standard day [52–54]. The hot-sunny day was taken from the meteorological data for Phoenix Arizona, USA, on June 24, 1976 [54, 55].

Other schemes for energy rating based on site-specific conditions instead of standard days have also been developed. In 1990, a rating based on realistic reporting conditions (RRC) was proposed. This method measured the performance of PV modules under different irradiances and temperatures and predicted the module's output under various operating conditions [3, 4, 56–60]. This method has been used to compare commercial modules, highlighting the different

Table 18.2 Photovoltaic rating criterion for PV applications

Application	Relevant PV parameter
Grid-connected, hydrogen production	Annual energy delivered
Power for peak utility demand	Power near solar noon
Remote system for cooling	Temperature coefficient and NOCT
Remote system with storage	Energy during cloudy day
Pump system for agriculture	Energy during growing season
Small power consumer products	Efficiency under room light
High value (Space)	High Efficiency, radiation and thermal stability, mass

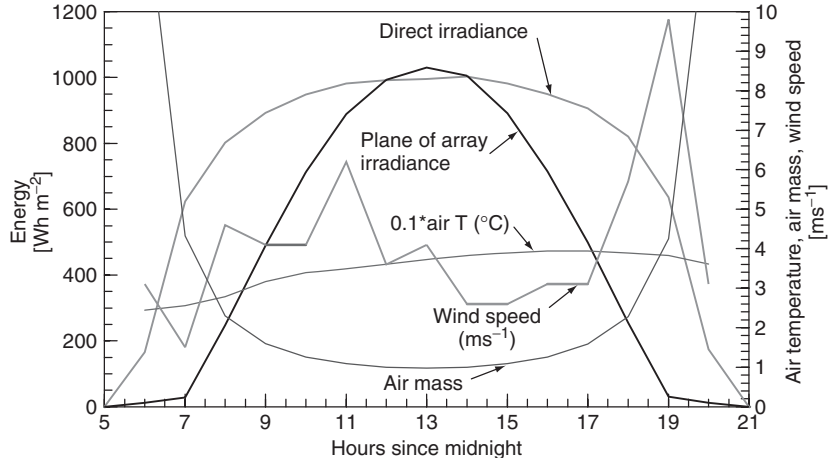


Figure 18.2 Meteorological conditions for the hot-sunny reference day [51–54]

dependencies on light level and temperature. The National Renewable Energy Laboratory (NREL), in conjunction with Sandia National Laboratory and in partnership with the US Department of Energy (DOE) Solar Energy Technologies Program (SETP), developed the Solar Advisor Model (SAM) to determine the levelized cost of energy for a variety of technologies. This model uses the typical meteorological year data base and translation equations to predict the energy for a given location in the US https://www.nrel.gov/analysis/sam/ [61, 62]. Table 18.3 illustrates the comparative ratings from the SAM model for selected module technologies for various locations in the United States representative of Hot/Dry, Cool/Dry, Hot/Humid, and Cool/Humid climate zones.

Table 18.3 The yield (annual kW h output of a PV module under STC normalized to the predicted array rated power) for four different cell technologies at four different climates. The expected yield is assuming the module operated at STC conditions for the number of 'sun hours' at each location. The percentage decreases, due to primarily to real operating temperatures greater than 25 °C and to intensity dependent device parameters, were calculated using the SAM (https://www.nrel.gov/analysis/sam/) [61, 62]

Phoenix, Arizona Hot, Dry (33° N, 112° W)		Miami, Florida Hot, Humid (25° N, 80° W)	
Expected output: 2372 kW h/kW		Expected Output: 1898 kW h/kW	
Hi Eff. C-Si	−23%	Hi Eff. C-Si	−21%
Multi-Si	−26%	Multi-Si	−23%
CdTe	−18%	CdTe	−15%
3J a-Si	−21%	3J a-Si	−16%
Billings, Montana Cool, Dry (45° N, 108° W)		**Boston, Massachusetts Cool, Humid (42° N, 71° W)**	
Expected Output: 1483 kW h/kW		Expected Output: 1377 kW h/kW	
Hi Eff. C-Si	−19%	Hi Eff. C-Si	−18%
Multi-Si	−19%	Multi-Si	−18%
CdTe	−14%	CdTe	−13%
3J a-Si	−17%	3J a-Si	−17%

The inputs to the SAM model were coefficients measured at Sandia for various technologies using King method described in the next section in Equations (18.6–18.15). The modules are installed at latitude tilt [50, 55]. The efficiency (%) and maximum power temperature coefficients (%/°C) of power for the modules was as follows: high efficiency monocrystalline Si: 19.3, −0.38; standard multicrystalline Si: 13.3, −0.50; CdTe thin film: 7.6, −0.22; and triple junction a-Si/a-SiGe/a-SiGe thin film: 5.7, −0.21. It should be noted that these efficiency values are below the average value of some commercially available products. As expected the predicted performance based upon the peak watt rating is significantly less than actually occurs because modules operate above 25 °C and have different temperature and irradiance coefficients. In many studies, covering many climate conditions over a wide range of latitudes on three continents, a-Si reportedly out produces c-Si by 10–20% in terms of kW h/kW [63–69]. These differences are more pronounced the lower the c-Si module efficiency or shunt resistance.

18.2.4 Translation Equations to Reference Conditions

Translation equations for the maximum power as a function of temperature and irradiance are useful for translating from measured conditions to reference conditions or for predicting the energy for a reference day or year. The most basic translation and often most accurate equations for a solar cell are based on the diode model with series and shunt resistances discussed in Chapter 3. Assumptions about the independence of the resistance, dark current, and diode quality factor with light level can be relaxed if the solar cell parameters are extracted from accurate 1-sun data. This model has been extended to modules by combining them in series and parallel combinations [70].

To a first order, short-circuit current (I_{sc}), open-circuit voltage (V_{oc}), P_{max}, and fill factor (FF) are linear with temperature, and I_{sc} is also linear with E_{tot} [54, 71–76]. The assumption that P_{max} is linear with E_{tot} and temperature along with the module's NOCT allows the performance under standard reference conditions to be translated as a function of air temperature and irradiance available in resource databases for energy-based rating methods [7, 39]. This has been implemented in the Web-based software package PVWatts, available at http://www.nrel.gov/rredc/. Numerous other publicly available and commercial software packages are available for system sizing. A technical report summarizing PV, hybrid system, and battery storage models was recently published [77]. Typical temperature coefficients for maximum power, or efficiency, of various PV technologies are summarized in Figure 18.3.

A set of translation equations for current and voltage based on the work of Sandstrom has been implemented in consensus standards [76, 78]. These equations translate the entire current–voltage curve for temperature and irradiance. Using the notation of the international standard in reference [78], the following equations allow one to translate the current I_1 and voltage V_1 measured from temperature T_1 to T_2 and irradiance E_1 to E_2:

$$I_2 = I_1 + I_{sc1}\left(\frac{E_2}{E_1} - 1\right) + \alpha(T_2 - T_1) \tag{18.4}$$

$$V_2 = V_1 - R_s(I_2 - I_1) - I_2 K(T_2 - T_1) + \beta(T_2 - T_1), \tag{18.5}$$

where α and β are the temperature coefficients, R_s is the series resistance, and K is a curve-shape correction factor. A typical value for the voltage temperature coefficient β is 2 mV/°C. A typical value for the K is 0.001 Ω/°C. Applying Equations (18.4) and (18.5) at a fixed irradiance ($E_2 = E_1$) and assuming no series resistance ($R_s = 0$), the value of K is determined that best translates the $I-V$ characteristics for temperature.

Translation equations for I_{sc}, V_{oc}, V_{max}, and I_{max} as a function of E_{tot}, T_c, absolute air mass (AM_a), and angle of incidence (AOI) based on multiple-regression analysis of field data have been

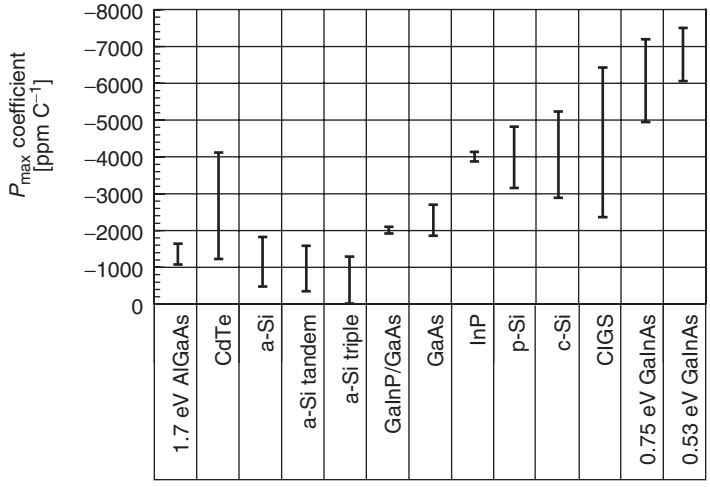


Figure 18.3 Typical temperature coefficients of various PV technologies [72]

proposed by King and implemented in SAM https://www.nrel.gov/analysis/sam/.:76,77, [79]

$$I_{\rm sc}(E, T_{\rm c}, {\rm AM_a}, {\rm AOI}) = (E/E_0) f_1({\rm AM_a}) f_2({\rm AOI})[I_{\rm sc0} + \alpha_{\rm Isc}(T_{\rm c} - T_{\rm o})] \quad (18.6)$$

$$E_{\rm e} = I_{sc}(E, T_{\rm c} = T_0, {\rm AM_a}, {\rm AOI})/I_{\rm sc0} \quad (18.7)$$

$$I_{\rm max}(E_{\rm e}, T_{\rm c}) = C_0 + E_{\rm e}[C_1 + \alpha_{\rm Imax}(T_{\rm c} - T_0)] \quad (18.8)$$

$$V_{oc}(E_e, {\rm i}) = V_{oc0} + C_2 \ln(E_e) + \beta_{Voc}(T_c - T_0) \quad (18.9)$$

$$V_{max}(E_e, T_c) = V_{max0} + C_3 \ln(E_e) + C_4 [\ln(E_e)]^2 + \beta_{Vmax}(T_c - T_0) \quad (18.10)$$

$$I_{sc0} = I_{sc}(E = E_0, T_c = T_0, AM_a = 1.5, AOI = 0^\circ) \quad (18.11)$$

$$V_{oc0} = V_{oc}(E_e = 1, T_c = T_0) \quad (18.12)$$

$$V_{max0} = V_{max}(E_e = 1, T_c = T_0) \quad (18.13)$$

$$f_2 = \frac{\frac{E_0}{I_{sco}} I_{sc}(AM_a = 1.5, T_c = T_0) - E_{diff}}{E_{dir}\cos(\theta)}, \quad (18.14)$$

where E is the plane-of-array solar irradiance in W/m^2, E_e is the effective irradiance in units of suns (0–1 for non-concentrated light), E_0 is the one-sun irradiance of 1000 W/m^2, E_{diff} is the diffuse irradiance in the plane of the module, E_{dir} is the direct-normal irradiance, and AOI is the solar angle-of-incidence on the module; T_c is the temperature of the cells inside the module, T_0 is the module rating reference temperature, and α_{Isc}, α_{Imax}, β_{Voc}, and β_{Vmax} are the temperature coefficients of I_{sc}, I_{max}, V_{oc}, and V_{max}, respectively. The pressure-corrected relative optical air-mass AM_a can be written as [80]:

$$AM_a = \frac{P}{P_0}\left[\cos(\theta) + 0.50572(96.07995^\circ - \theta)^{-1.6364}\right]^{-1}, \quad (18.15)$$

where P is the barometric pressure, P_0 is the pressure at sea level, and θ is the angle between the sun and zenith in degrees. The function $f_1(AM_a)$ is empirically obtained from the temperature- and irradiance-corrected I_{sc} versus air mass and assumes that the only spectral dependence is the zenith angle. Data are collected over a range of irradiances, incident angles, air masses, and temperatures and a multiple-regression analysis is applied. These translation equations have been compared with simple linear translation equations derived from simulator-based measurements

for several modules using outdoor data [54]. These translation equations give similar results to Equation (18.3) when temperature, maximum power tracking, and spectral issues are considered [53]. Recently, the bilinear method has gained attention because of its ability to translate the $I-V$ curve for temperature and irradiance based on a minimum number of other $I-V$ curves at other temperatures and irradiances [81, 82]. The bilinear method has recently been included in IEC standards for $I-V$ translation equations [78]. Other translation equations for current and voltage developed for the space program in the 1970s are possible [27].

18.3 CURRENT–VOLTAGE MEASUREMENTS

A solar cell can be modeled as a diode in parallel with a current generator, whereas a module is a series-parallel network of solar cells, as discussed in Chapter 7. Measurements of cell or module $I-V$ behavior allow the diode characteristics to be determined, along with other important parameters including the maximum power point, P_{max}. A typical $I-V$ measurement system is composed of a simulated or natural light source, test bed to mount the device under test, temperature control and sensors, and a data acquisition system to measure the current and voltage as the voltage across the device or current through the device is varied with an external load or power supply.

18.3.1 Measurement of Irradiance

The measurement of the irradiance, E_{tot}, incident on the PV device in Equation (18.1) is typically performed with a thermal detector (pyranometer, cavity radiometer) for outdoor measurements and reference cells for simulator-based measurements. If the goal is to determine the PV efficiency or power with respect to standardized or different reference conditions, then a spectral error will exist due to differences in spectral sensitivity between the cell or module and the reference device. Outdoor measurements of PV systems or modules are often made with respect to the total irradiance incident on the module. If a broadband thermal detector with a constant spectral responsivity is used, then the spectral error is zero. If a silicon-based pyranometer is used to measure the performance based on the total irradiance, then there will be a spectral error unless its spectral responsivity, SR matches the cell or module under test. For PV measurements with respect to a reference spectrum, the spectral error in the measured short-circuit current I_{sc} of a PV device can be written in general as [2]:

$$I'_{sc} = \frac{I_{sc}}{M} = I_{sc}\frac{\int_{\lambda_1}^{\lambda_2}\int_{\theta_1}^{\theta_2}\int_{\phi_1}^{\phi_2} E_{Ref}(\lambda,\theta,\phi)S_T(\lambda,\theta,\phi)\,d\lambda\,d\theta\,d\phi \int_{\lambda_1}^{\lambda_2}\int_{\theta_1}^{\theta_2}\int_{\phi_1}^{\phi_2} E_S(\lambda,\theta,\phi)S_R(\lambda,\theta,\phi)\,d\lambda\,d\theta\,d\phi}{\int_{\lambda_1}^{\lambda_2}\int_{\theta_1}^{\theta_2}\int_{\phi_1}^{\phi_2} E_{Ref}(\lambda,\theta,\phi)S_R(\lambda,\theta,\phi)\,d\lambda\,d\theta\,d\phi \int_{\lambda_1}^{\lambda_2}\int_{\theta_1}^{\theta_2}\int_{\phi_1}^{\phi_2} E_S(\lambda,\theta,\phi)S_T(\lambda,\theta,\phi)\,d\lambda\,d\theta\,d\phi}, \tag{18.16}$$

where the spectral responsivity of the device under test (S_T), spectral responsivity of the reference detector (S_R), reference spectral irradiance (E_{Ref}), and source spectral irradiance (E_S) are a function of wavelength (λ), and incident azimuth (ϕ) and zenith (θ) angles. Spectral responsivity is measured in units of A/W but can be converted to units of quantum affiance in units of electron/photon. Spectral irradiance is measured in units of W/m^2/unit wavelength. If quantum efficiency is desired for the spectral correction equation then the spectral irradiance must be converted to photon flux in units of photon/cm^2/s/unit wavelength. This general form allows the reference detector to be noncoplanar with the device under test and the source angular distribution of the source spectrum to be nearly arbitrary. In practice, the test device and reference detector used to

measure the total irradiance are usually coplanar to minimize errors associated with measuring the orientation. The direct and AM0 reference spectra have no angular dependence, and measurements are normally performed at normal incidence so the angular dependence in Equation (18.16) drops out. If the angular dependence of the reference detector used to measure E_{tot} and the global reference spectra follow an ideal cosine response, then Equation (18.16) simplifies to [83–85]:

$$M = \frac{\int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda) S_R(\lambda)\, d\lambda \int_{\lambda_1}^{\lambda_2} E_S(\lambda) S_T(\lambda)\, d\lambda}{\int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda) S_T(\lambda)\, d\lambda \int_{\lambda_1}^{\lambda_2} E_S(\lambda) S_R(\lambda)\, d\lambda}. \tag{18.17}$$

If the reference detector is a thermal detector, S_R is independent of wavelength and drops out. The source spectral irradiance (E_S) and spectral responsivity of S_R and S_T need only be relative in Equation (18.17), because any multiplicative error sources will drop out. Ideally, the limits of integration λ_1 and λ_2 in Equation (18.17) should encompass the range of the reference detector and reference spectrum, or else an error can arise [86]. If the spectral irradiance of the light source is the same as the reference spectrum or if the relative spectral responsivity of the reference detector matches the relative spectral responsivity of the test device, then M is unity. Manufacturers of PV cells and modules for multi-million-dollar satellites require the lowest possible measurement uncertainty, so they require primary balloon or space-calibrated reference cells of the same manufacturing lot, and purchase solar simulators with the closest spectral match to AM0 that is technically possible so they can assume M is unity.

The short-circuit current of the reference cell under the source spectrum ($I^{R,S}$) is used to determine the effective irradiance using the following equation:

$$E_{tot} = \frac{I^{R,S} M}{CV}, \tag{18.18}$$

where CV is the calibration value of the instrument used to measure the incident irradiance in units of A W^{-1} m^2. If a thermal detector is used, then CV has the units of V W^{-1} m^2, and $S_R(\lambda)$ is constant.

18.3.2 Simulator-based $I-V$ Measurements: Theory

The short-circuit current of a test device ($I^{T,R}$) at the reference total irradiance (E_{Ref}) can be written as: [2, 4, 84, 85]

$$I^{T,R} = \frac{I^{T,S} E_{ref} CV}{I^{R,S} M} = I^{T,S} \frac{E_{ref}}{E_{tot}}, \tag{18.19}$$

where $I^{T,S}$ is the short-circuit current of the test device measured under the source spectrum, M is from Equation (18.17), and $I^{R,S}$ is the short-circuit current of the reference cell under the source spectrum. This is the standard simulator-based calibration procedure. Many groups assume M is unity because of the difficulty in obtaining a spectral irradiance of the source spectrum and knowledge of the spectral responsivities of the test and reference device. Typically, the simulator is adjusted so that E_{tot} is equal to E_{Ref} from Equation (18.18) or

$$1 = \frac{E_{Ref} CV}{I^{R,S} M} = \frac{I^{R,R}}{I^{R,S} M}. \tag{18.20}$$

Typically, a reference cell is made of the same material and technology as the devices that it will be used to test, causing M to be closer to unity because $S_R \approx S_T$. Because of stability questions for research cells it is often desirable to use a color glass filter over a Si cell to simulate the cells responsivity. The responsivity of organic PV, dye-sensitized PV and amorphous Si can be best simulated with a Schott color glass filter such as model KG5 being placed over a mono- or multicrystal Si cell. Some small-area filtered Si cells such as those fabricated by Hamamatsu part numbers S1133 or S1787-4 are suitable for use as 1-sun reference cells. At 1-sun light levels most commercial detectors saturate causing the photocurrent to not be equal to the short-circuit current because the detector is series resistance limited. Larger-area filtered Si reference cells are also available commercially from PV reference cell companies and some calibration labs. Ideally, the angular response of the reference package should be similar to the device under test. This is essential for outdoor measurements [2, 87–89]. Consensus standards have been developed, giving guidance for reference cells [90–93]. If the detector package has a window and an air gap between the window and cell, then the package should be completely illuminated and used only with simulators to prevent reflection-related artifacts [34, 94]. Recently, the terrestrial community has proposed a standard package design for the World Photovoltaic Scale [92, 93]. This package was designed by international terrestrial PV calibration laboratories to accommodate their various PV calibration equipment, while having standardized connectors to facilitate international intercomparisons.

18.3.3 Primary Reference Cell Calibration Methods

Perhaps the most straightforward method of determining the short-circuit current with respect to a set of reference conditions is to measure the absolute external spectral responsivity at the reference temperature, $S_T(\lambda)$, and to integrate it with the reference spectrum, $E_{Ref}(\lambda)$, at the reference total irradiance, E_{tot}, using the following equation:

$$I_{sc} = \frac{E_{tot} A \int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda) S_T(\lambda)\, d\lambda}{\int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda)\, d\lambda}. \tag{18.21}$$

For I_{sc} to be in units of amps, wavelength (λ) must be in units of μm, PV area (A) in m^2, $E_{Ref}(\lambda)$ in $W\,m^{-2}\,\mu m^{-1}$, E_{tot} in $W\,m^{-2}$, and $S_T(\lambda)$ in $A\,W^{-1}$. The limits of integration should encompass the range of $E_{Ref}(\lambda)$. If $E_{Ref}(\lambda)$ is normalized to integrate to E_{tot}, then the limits of integration should encompass the response range of the device. The relationship between the spectral responsivity and the quantum yield is discussed in Section 18.4. Equation (18.21) assumes that $S_T(\lambda)$ is uniform over the PV device and that $S_T(\lambda)$ is independent of voltage bias, E_{tot}, and $E_{Ref}(\lambda)$. These assumptions can be relaxed for single-junction devices by applying an external bias light operating at E_{tot}. Several groups have gone to great lengths to minimize the various errors associated with Equation (18.21) [95, 96]. Several intercomparisons show that differences in the absolute spectral response of more than 10% are possible from well-known PV calibration laboratories [92, 93–99]. Accurate spectral response measurements require measuring the total power incident on a small area of the PV device (or the power density over the entire PV device) and the current produced at that wavelength. The major sources of uncertainty are the detector calibration used to measure the power (typically μW) and errors in measuring a small ac current (typically μA or less) produced by the chopped monochromatic light, in the presence of a large dc offset (mA to A) from the broadband bias light. Because a reference cell calibrated using Equation (18.21) will be fully illuminated when used to set a solar simulator or measure E_{tot} in natural sunlight, variations in the responsivity over the sample surface may result in an error in I_{sc}. Some groups employ a laser to more accurately measure the absolute response of a reference cell at one

wavelength because the light power is in the mW range – instead of μW, associated with diffraction-grating or filter-based systems – and reference detectors used to measure the monochromatic light power can be more accurately calibrated for laser lines.

A primary reference cell can be calibrated under natural sunlight using Equation (18.19) and a thermal detector [1, 2, 9, 83, 85, 100–103];

$$CV = \frac{I^{T,S}}{E_{tot}} \frac{\int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda) S_T(\lambda)\, \mathrm{d}\lambda}{\int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda)\, \mathrm{d}\lambda} \frac{\int_{\lambda_1}^{\lambda_2} E_S(\lambda)\, \mathrm{d}\lambda}{\int_{\lambda_1}^{\lambda_2} E_S(\lambda) S_T(\lambda)\, \mathrm{d}\lambda}, \tag{18.22}$$

where the cell short-circuit current ($I^{T,S}$), solar spectra ($E_S(\lambda)$), and total irradiance (E_{tot}) are measured at the same time. The spectral responsivity of the reference detector is constant, so $S_R(\lambda)$ is constant and drops out of Equation (18.22). The incident irradiance E_{tot} is measured with a thermal detector that is traceable to the world radiometric reference scale, such as an absolute-cavity radiometer or pyranometer. The solar constant and solar power density in units of $\mathrm{W\,m^{-2}}$ for solar applications are based on the world radiometric reference scale, which is derived from a family of primary absolute-cavity radiometers maintained at the World Radiation Research Center in Davos, Switzerland [102]. The field of view for the cell and spectroradiometer must be matched. Some investigators prefer to use a pyranometer mounted coplanar to the spectroradiometer and solar cell(s) on a horizontal surface [9, 101]. Investigators at the National Renewable Energy Laboratory (NREL) use an absolute-cavity radiometer because it is the primary instrument used to calibrate pyranometers and has a field of view of 5°, minimizing field-of-view-related error sources [1, 100, 103, 104]. The pyranometer-based calibration method requires that the spectral irradiance over a wavelength range of 300–2500 nm be known for the spectral correction factor, outside of which the contribution is negligible [9, 105]. If a cavity radiometer is used to measure the direct beam irradiance and a pyranometer is used to measure the diffuse irradiance then limiting the measured spectral irradiance to a 300–2500 nm range assumes that the total irradiance outside this range is identical to the reference spectrum because the cavity radiometer responds from the UV to far IR. Typically, CV is an average of many short-circuit current measurements taken over several days. A current-to-voltage converter was developed in 1983 by the PV Cell and Module Performance Characterization Team at NREL because of the long lead lengths (required to reach from the indoor test station to the outdoor test facility) and the desire to measure the short-circuit current with an uncertainty of less than 0.02% while maintaining the bias at the cell to within several mV of zero volts. The converter is given in Figure 18.4. This method only requires that the relative spectral response and spectral irradiance be known, thereby eliminating all error sources that are multiplicative and wavelength independent.

If the absolute spectral irradiance $E_S(\lambda)$ of the light source in the test plane is known, as is the case for a standard lamp, blackbody, or absolute spectral irradiance measurement, then Equation (18.18) reduces to [83, 105–107]:

$$CV = I^{T,S} \frac{\int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda) S_T(\lambda)\, \mathrm{d}\lambda}{\int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda)\, \mathrm{d}\lambda \int_{\lambda_1}^{\lambda_2} E_S(\lambda) S_T(\lambda)\, \mathrm{d}\lambda}. \tag{18.23}$$

This method is appealing because the reference spectrum may correspond to the blackbody spectrum, and it can be readily performed in the laboratory. This method only requires the relative spectral responsivity measurements. If a standard lamp is used and the reference spectrum is the terrestrial reference spectra in Table 18.1, then the spectral correction factor in Equation (18.23) is typically 12 [107]. This method is sensitive to errors in $S_T(\lambda)$, $E_S(\lambda)$, positioning, and stray

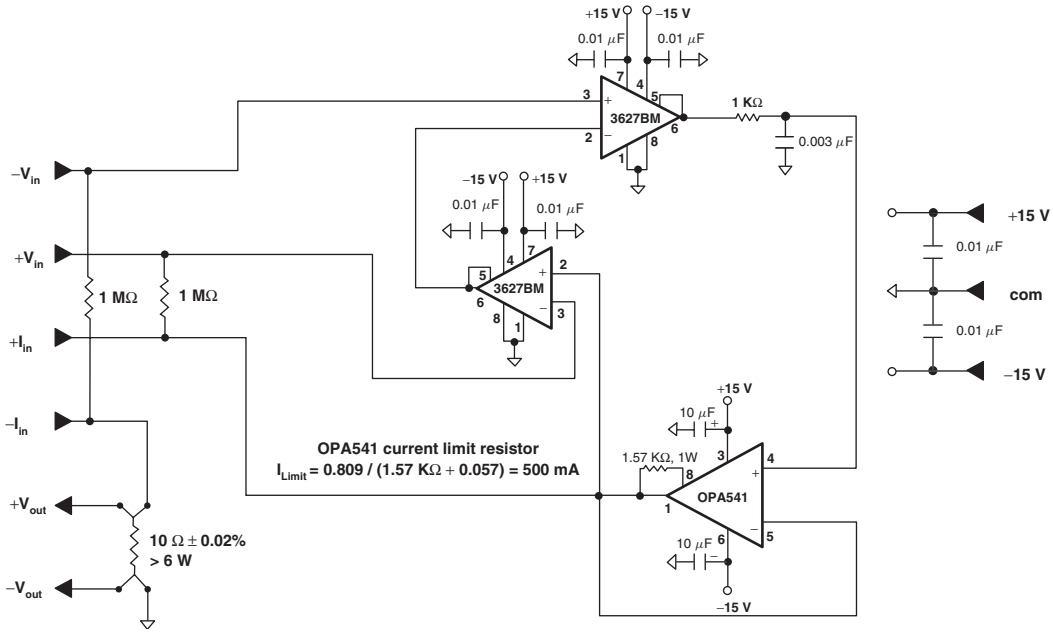


Figure 18.4 Current to voltage converter, the accuracy of which is limited by the accuracy and temperature dependence of the current sense resistor, designed by Carl Osterwald and used by the PV Cell and Module Performance Characterization Team in the National Center for Photovoltaics at NREL since 1983

light [86, 107]. The sensitivity to positioning is because the light source (i.e. standard lamp or blackbody) is not collimated, and changes in total irradiance of 1% per millimeter distance from the source are not uncommon [107]. If a solar simulator is used as the light source in Equation (18.23), then the spectral correction factor is greatly reduced, minimizing the sensitivity to errors in $S_T(\lambda)$, $E_S(\lambda)$, and positioning errors are reduced to less than 0.1% per mm of distance from the source [105].

Once the short-circuit current is known under a given $E_{old}(\lambda)$, it can be translated to any other $E_{new}(\lambda)$ with the following equation:

$$I_{new} = \frac{I_{old} E_{tot}^{new}}{E_{tot}^{old}} \frac{\int_{\lambda_1}^{\lambda_2} E_{old}(\lambda)\,\mathrm{d}\lambda \int_{\lambda_1}^{\lambda_2} E_{new}(\lambda) S_T(\lambda)\,\mathrm{d}\lambda}{\int_{\lambda_1}^{\lambda_2} E_{new}(\lambda)\,\mathrm{d}\lambda \int_{\lambda_1}^{\lambda_2} E_{old}(\lambda) S_T(\lambda)\,\mathrm{d}\lambda}. \tag{18.24}$$

Equation (18.24) assumes that the current is linear with intensity. This method is especially useful for translating the calibration of primary AM0 reference cells to terrestrial reference spectra or vice versa.

The AM0 reference spectrum is, by definition, the extraterrestrial solar spectrum at 1 astronomical unit distance from the sun. The AM0 community uses this fact to calibrate reference detectors by measuring their response in space or very high altitudes. In this case the true value is determined by the actual solar spectrum at the time of calibration and not a tabular reference spectrum, as in the case of terrestrial-based calibration methods. This means that a small random

error will exist because the solar spectrum varies slightly with solar activity. By definition, there is no spectral error.

Extraterrestrial calibration procedures include using spacecraft, balloons, and high-altitude aircraft and the calibration procedures defined by Equations (18.21–18.23) [17, 108–114]. There are no spectral corrections for balloon and spacecraft calibrations because the data are taken above the atmosphere. The high-altitude aircraft calibration procedure involves a Langley plot of the logarithm of I_{sc} versus absolute or pressure-corrected air mass over a typical range 0.25–0.5 [110, 111, 113, 114]. The data are collected above the tropopause, thereby eliminating water vapor and most scattering, with the dominant spectral feature arising from ozone absorption [110, 114]. The Jet Propulsion Laboratory (JPL) balloon calibration program requires a custom package design for standardized mounting, data acquisition system, and thermal considerations [108, 109]. ISO 15387 allows for ground based calibrations with respect to a synthetic AM0 spectrum (Figure 18.1) using Equations 18.21, 18.22, or 18.23 [17].

18.3.4 Uncertainty Estimates in Reference Cell Calibration Procedures

All measurements have an uncertainty between the measured or derived value and the true value. In the case of terrestrial PV reference cell calibrations, the true value is the value under reference conditions given in Table 18.1. For extraterrestrial calibrations, the true value is determined by the actual solar spectrum at the time of calibration and not the tabular reference spectrum, as in the case of terrestrial calibrations. Any variation in the primary AM0 calibration because of the varying solar constant is not considered an error. The spectral responsivity of PV cells change as a function of radiation damage as discussed in Chapter 9. For an accurate assessment of the performance as a function of radiation damage using procedures that assume the spectral correction factor is unity at least three matched reference cells are required (beginning of life, mid life, and end of life) to minimize the spectral errors. A formal uncertainty analysis is required for ISO 17025 accredited calibration laboratories and facilities that certify that a module has passed the IEC design qualification and type approval requirements [115]. An uncertainty analysis can serve as a guide to reducing the dominant error sources in a measurement.

The uncertainty that is expected to include 95% of the results (U_{95}) in a calibration can be expressed in terms of the Type A and Type B error sources [116]. Type A error sources are statistical in nature such as the standard deviation, whereas Type B error sources are all others that can be written as:

$$U = k\sqrt{\sum\left(\frac{c_i u_i}{p_i}\right)^2}. \tag{18.25}$$

For 95% confidence, the coverage factor k is 2. The sensitivity factor c_i involves partial derivatives and becomes unity if the uncertainty can be expressed as a product of uncorrelated terms that are in units of percentage of value. The probability distribution p_i is $\sqrt{2}$ for rectangular probability and 2 for Gaussian probability distribution. If the distribution is unknown, then it is conservative to assume that it is rectangular. An example of an uncertainty analysis of Equation (18.25) using the NREL direct-beam calibration method required for NREL's ISO 17025 accreditation by A2LA as a calibration laboratory is given in Table 18.4 [116–120]. To perform the uncertainty analysis, Equation (18.22) is expanded to include all the relevant correction factors:

$$\langle CV\rangle = \langle CV_U kT_c\rangle = \left\langle \frac{I^{T,S}}{E_{tot}} kT_c \right\rangle. \tag{18.26}$$

Table 18.4 Uncertainty analysis of the NREL PV Cell and Module Performance Characterization Team's implementation of tabular calibration method equation following the ISO Guide to the expression of uncertainty following Equation (18.22) taken from the required uncertainty analysis for ISO 17025 accreditation [98, 116, 119, 120]

Uncertainty component	Source of uncertainty	Value (%)	Coverage
Type A error sources			
$U_{\mathrm{CVu(25\,^{\circ}C)}}$	Corrected calibration value (all data sets, 3 days)	0.27	$n = 35$
U_{CVu}	Uncorrected calibration value (one data set)	0.083	$n' = 85$
Type B error sources			
U_{Etot}	Measured total irradiance (cavity radiometer)	0.34	Gaussian
U_{Tc}	Temperature correction to I_{sc} to each data set	0.14	rectangular
U_{k}	Spectral correction to each data set	0.80	Gaussian
U_{Isc}	Measured I_{sc} for one point in one data set	0.034	rectangular
U_{meter}	Reference cell DMM ($\mathrm{I_{sc}} = 123\,\mathrm{mA}$ typical value)	0.023	Gaussian
	1-year HP34401, 1 V, of reading	0.004	Gaussian
	1-year HP34401, 1 V, of range	0.0007	Gaussian
	1-year HP34401, 1 V, 1 line cycle	0.001	Gaussian
	HP34401, 1 V, temperature $23 \pm 10\,^{\circ}\mathrm{C}$	0.005	Gaussian
U_{resistor}	1-year current sense resistor uncertainty	0.02	Gaussian
$U_{\mathrm{stability}}$	1 Ω current-sense resistor 1-year stability	0.002	Gaussian

The spectral correction factor, k, is the inverse of M in Equation (18.18). Its uncertainty, estimated using Monte Carlo methods, is 10–20% of the magnitude of the spectral correction factor [86, 118]. Uncertainties related to the temperature correction factor T_c involve an additive and multiplicative component, but can be transformed into a single multiplicative temperature correction factor to the current, voltage, or power whose uncertainty can be estimated. A detailed uncertainty analysis of the error in determining the maximum power of a module measured outdoors using the ISO methodology was published by Whitfield and Osterwald [121].

The uncertainty analysis in Table 18.4 is a simplified version of a more comprehensive analysis in reference [120]. The analysis is for equipment that minimizes all known error sources in Equation (18.26). For example, the current sense resistor introduces an error of less than 0.02% because a stable low temperature coefficient resistor was chosen instead of a less expensive typical 0.25% precision with a temperature coefficient above 100 ppm/°C. The test bed minimized meter related uncertainties by choosing a 6 1/2 voltmeter. The dominant error sources in Equation (18.26) are the absolute cavity radiometer used to measure E_{tot} and the spectral correction factor. The other principle error source is the random day to day error after all error sources have been corrected for of 0.3%. Since the error is random collecting data on more than 3 days will reduce this contribution to the total error. The direct method was chosen for outdoor primary reference cell calibration because the global method requires a pyranometer to measure E_{tot} which has significantly higher error sources related to measuring the diffuse component incident on the cell being calibrated. When using instruments outdoors it should be noted that additional temperature related may not be include in the calibration labs uncertainty estimate. An example of this is a pyranometer calibration where some instruments are known to have a significant temperature dependence, but the calibration lab only reports the temperature that the data was taken at with no estimate of the error for using the unit at any other temperature. The temperature coefficient of pyranometers used to measure the broadband solar irradiance are usually not stated by the manufacturer and can be a significant error source [88]. Voltmeters in outdoor conditions often exceed the temperature range that the

manufactures uncertainty estimate was based upon. A rigorous uncertainty estimate should include this additional error source [120]. When a detector is calibrated the calibration temperature is given but the temperature coefficient is usually not reported so use of the detector at any other temperature will result in an additional error with the error being approaching 5% /°C near the energy gap.

18.3.5 Intercomparison of Reference Cell Calibration Procedures

Typically, groups claiming to be able to calibrate primary reference cells with respect to reference conditions claim an uncertainty of ±1% [92, 99, 85, 112, 122]. Intercomparisons among the various calibration methods are the best way to determine if the uncertainty estimates are valid. Formal intercomparison of terrestrial calibration procedures sponsored by the Photovoltaic Energy Project (PEP) were conducted in 1985, 1987, and 1993 [92, 99, 85]. The PEP '85 intercomparison involved PV calibration laboratories from the Commission of European Communities, France, Germany, Italy, Japan, United Kingdom, and the United States. The differences in I_{sc} with respect to the global reference spectrum between the laboratories in the PEP '85 intercomparison were almost 8% for single-crystal and multicrystal silicon and 20% for amorphous silicon. However, in the PEP '85 intercomparison, six out of the eight agencies agreed within 3% for the crystalline cells and 6% for the amorphous cells [99]. In the PEP '87 and PEP '93 intercomparisons, the participants provided uncertainty estimates. The level of agreement between the laboratories for the PEP '87 intercomparison was 4% for single- and multicrystal silicon cells and 14% for amorphous silicon [98]. This level of agreement was 2–10 times the laboratories' estimated uncertainties, which ranged from ±0.7 to ±5%, indicating that the uncertainty estimates for some of the participants were overly optimistic [98]. Several of the participants based their estimated uncertainties on the standard deviation of repeated calibrations, thereby neglecting nonrandom error sources. The PEP '93 intercomparison also showed a rather large spread of 12% for the single- and multicrystal silicon cells, even though the estimated U_{95} uncertainties ranged from ±1.0 to ±2.7% – again indicating that several laboratories underestimated their errors [92]. After excluding laboratories whose calibrations were based on reference cells and laboratories that had more than 50% of their calibrations exceed ±2% of the mean, the resultant average deviation was 1.1% [92]. Because the terrestrial PV calibration laboratories around the world could not agree on a calibration procedure that had a proven U_{95} uncertainty of less than ±2%, it was decided to establish a set of reference standards called the World Photovoltaic Scale (WPVS) [92, 93]. The four laboratories from the PEP '93 intercomparison that performed primary calibrations and were within ±2% of the mean were the National Renewable Energy Laboratory in the United States, Japan Quality Assurance Organization/Electrotechnical Laboratory in Japan, Tianjin Institute of Power Sources in the Peoples Republic of China, and Physikalisch-Technische Bundesanstalt (PTB) in Germany. A formal mechanism was established to include other laboratories in the future, provided their calibrations agreed with the four established WPVS calibration laboratories. Recently, the European Solar Test Installation – Joint Research Center in Ispra, Italy, was included by performing blind intercomparisons on their global calibrations with three of the other four laboratories [122]. The WPVS reference cells reside with the laboratory that provided them, providing each participating laboratory with reference cells traceable to the WPVS. A set of technical drawings was developed for future WPVS reference cells to prevent the problem with the existing set of twenty WPVS reference cells having incompatible cables, mounting holes, and temperature sensors [92, 93]. Sixteen of the twenty WPVS reference cells were recently recalibrated, along with six new candidate WPVS reference cells [123]. The new WPVS calibration values changed by 0.4% at most, with an average decrease of 0.2% [123]. The results of the 1993 PEP terrestrial intercomparison were much closer than earlier intercomparisons, where deviations of ±3% were common and deviations of 5% from the mean were observed [125]. The year-to-year repeatability appeared to be no better than ±3% [125]. Recent intercomparisons at the module level hosted by NREL and the European

union show little reduction in differences since the PEP'93 intercomparison [126–130]. The NREL hosted intercomparison among certified calibration or qualification laboratories where the participants were only given commercial modules with no further information or "matched" reference cells indicated a ±3.5% deviation in P_{max} from the average for Si and larger for thin films [126]. The Fraunhofer hosted intercomparison showed a deviation in P_{max} for Si commercial Si modules of ±2% [128].

Surprisingly, the spread in I_{sc} temperature coefficients for the laboratories that performed them was greater than 50%, even though the temperature of the cells could be controlled and they had temperature sensors permanently attached to them [92]. This variation in temperature coefficient can be partly understood by noting that the temperature coefficient is a function of the light source by which the cells are illuminated [72]. For cells having a narrow response range, or for multijunction cells, the sign of the coefficient can even change depending on the light source [72, 131, 132].

The AM0 community has conducted intercomparisons between various groups over the years [108, 112, 124, 133]. In general, the agreement of primary AM0 calibrations on spacecraft, balloon, and high-altitude aircraft calibrations is better than ±1% over many years [108–110]. AM0 calibrations based on terrestrial measurements sometimes agree with primary high-altitude calibrations within 1% [109] and at other times the agreement is poor (>10%) [124].

18.3.6 Multijunction Cell Measurement Procedures

Spectrum-splitting series-connected multijunction devices are critical to achieving high-efficiency III–V-based (Chapter 8) and a-Si-based (Chapter 12) solar cells. The procedures to measure the $I-V$ characteristics of a multijunction device with respect to reference conditions are the same as those for a single-junction device, but with the added constraint that the simulator should be set so each junction operates at the proper photocurrent. This is accomplished by satisfying the following j equations for the j junctions in a multijunction PV device [134–137]:

$$I_j^{R,R} = M_j I_j^{R,S} \tag{18.27}$$

$$M_j = \frac{\int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda) S_{R,j}(\lambda)\,\mathrm{d}\lambda \int_{\lambda_1}^{\lambda_2} E_S(\lambda) S_{T,j}(\lambda)\,\mathrm{d}\lambda}{\int_{\lambda_1}^{\lambda_2} E_{Ref}(\lambda) S_{T,j}(\lambda)\,\mathrm{d}\lambda \int_{\lambda_1}^{\lambda_2} E_S(\lambda) S_{R,j}(\lambda)\,\mathrm{d}\lambda}. \tag{18.28}$$

This procedure involves adjusting the simulator in Equation (18.27), measuring its spectrum, calculating M_j, and readjusting the simulator. If reference cells are used whose relative spectral responsivity efficiency matches each individual junction, then the spectral correction M_j is unity, and only Equation (18.27) needs to be satisfied. This is possible for the high-efficiency crystalline material systems, where the other junctions can be shorted out without significantly affecting the relative spectral responsivity [134]. The adjustment of the solar simulator to satisfy Equation (18.27) can be problematic. Figure 18.5 illustrates the various approaches that researchers have taken in the past. The first approach, illustrated in Figure 18.5a, involves combining an ultraviolet/visible (UV/VIS) light source L_1 and infrared (IR) light source L_2 with a dichroic filter assembly [134, 135, 138, 139]. This approach is particularly useful for two- and three-junction devices where the top-cell energy gap is around 600–700 nm and the middle-cell energy gap is also around 600–700 nm. This is because of the relatively limited choices in the transition wavelength for standard dichroic filters. Figure 18.5b works for any material system because it uses a simulator whose spectral match is close to the reference spectrum, but brings in extra light that can be filtered at will for each junction [140]. The primary disadvantage of this approach is that the supplemental light sources

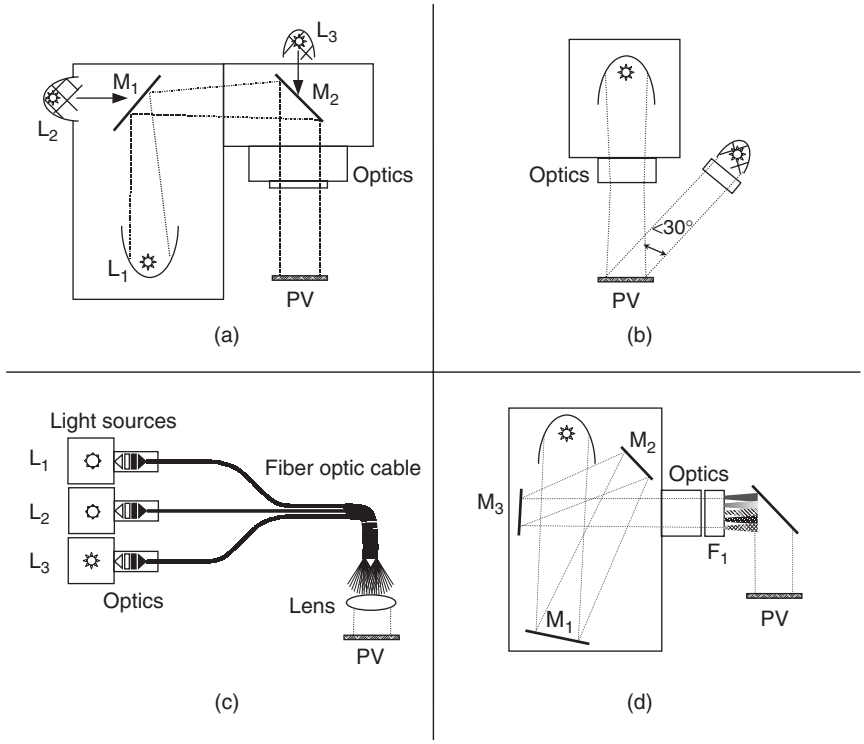


Figure 18.5 Methods of adjusting the spectral content of solar simulators. The light sources are L_1 and L_2 and L_3 and the mirrors are M_1, M_2, and M_3

are not collinear with the broadband light, giving the possibility of large variations in the spectral irradiance in the test plane. A fiber-optic solar simulator, shown schematically in Figure 18.5c, is useful because a wide variety of laser and incoherent light sources can be combined into one fiber bundle, which then illuminates the test plane [136, 141, 142]. This approach has the disadvantage of being restricted to small illumination areas, typically less than 2 cm in diameter at 1-sun, although commercial vendors such as SphereOptics/LabSphere offer to build systems with 10 cm or larger diameter at 1-sun with multiple spectral bands for any conceivable high-efficiency multijunction structure. Another approach, in use at NREL since 1988, shown in Figure 18.5d, is to place filters and apertures close to the integration optics of a large-area solar simulator [136, 137, 143]. This method is particularly useful for large-area ($>100\,cm^2$) samples. Its primary drawback is that the light sources are not separately adjustable for each junction. This concept can also be applied to pulsed simulators, where the distance between the flash lamp(s) and the test plane is usually large and a wide range of intensities is possible. This method works for any multijunction technology because standard high-pass, low-pass, and band-pass filters are available to cover any combination of bandgaps. The limitation of the method is that it requires multiple iterations where the spectrum of the light source is adjusted, the spectrum remeasured, and the spectral error recalculated until no further change in the simulator spectrum to achieve the desired spectral matching is required.

An entirely new procedure has recently been developed where the intensity of each light source is adjusted only once [137, 144, 145]. The approach is based on a simulator setup with multiple light sources arranged so that the intensity of each source may be adjusted without causing a change in the relative spectral irradiance. Each light source, $E_{S,i}(\lambda)$ for each junction, j, can

therefore be characterized by its relative irradiance. The adjustment of the intensity of the ith light source to an absolute level is mathematically described by the scaling factor C_i. Under these premises, the condition $I_j^{T,S} = I_j^{T,R}$ yields a system of j linear equations,

$$\sum_i C_i \int E_{S,i}(\lambda) S_{t,j}(\lambda)\, d\lambda = \int E_{ref}(\lambda) S_{t,j}(\lambda)\, d\lambda, \tag{18.29}$$

which can easily be solved for the unknown scaling factors C_i. The intensity of the ith light source $E_{S,i}(\lambda)$ for the jth junction of the test device $S_{t,j}(\lambda)$ is adjusted until the measured short-circuit current of reference cell $I_j^{T,R}$ with an absolute spectral response $S_{R,j}(\lambda)$ is obtained using

$$I_j^{T,R} = C_i A \int E_{S,i}(\lambda) S_{R,j}(\lambda)\, d\lambda. \tag{18.30}$$

The system of equations requires only relative spectral responses and relative irradiances and is therefore equivalent to the calculation of the mismatch factor in Equation (18.28). The equations allow for the same reference cell to be used to set the simulator sources [137, 144, 145]. This procedure only requires one adjustment of each light source, and that the relative spectral irradiance be measured only once for each light source, unlike the procedure described by Equations (18.27) and (18.28), which requires a spectral irradiance measurement after each simulator adjustment. This allows the performance of a multijunction PV device under varying reference spectra or current matching conditions to be rapidly determined.

18.3.7 Cell and Module $I-V$ Systems

A wide variety of $I-V$ measurement systems have been developed to measure the performance of PV devices – from 0.01-cm^2 area cells to multi-kilowatt arrays [2, 146]. A generic $I-V$ system is shown in Figure 18.6. The voltage across the PV device (from a cell to an array) is biased with a variable load, with the current being sensed by a precision four-terminal shunt resistor or magnetic transducer. A Kelvin connection to the device under test (DUT) is required to eliminate resistance losses up to the location where the voltage is sensed. For cells the voltage sense would be adjacent to the current contact. For modules the voltage sense is at the junction box where the current lead is connected. Domestic and international standards have been developed for the minimum characteristics of typical $I-V$ measurement systems [5–7, 9, 11, 15]. The critical parameters on

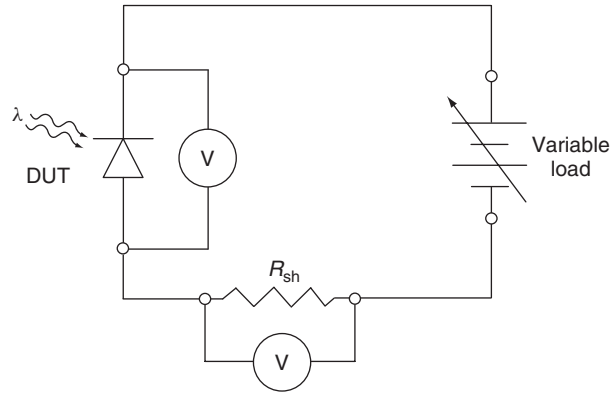


Figure 18.6 Typical current–voltage measurement system

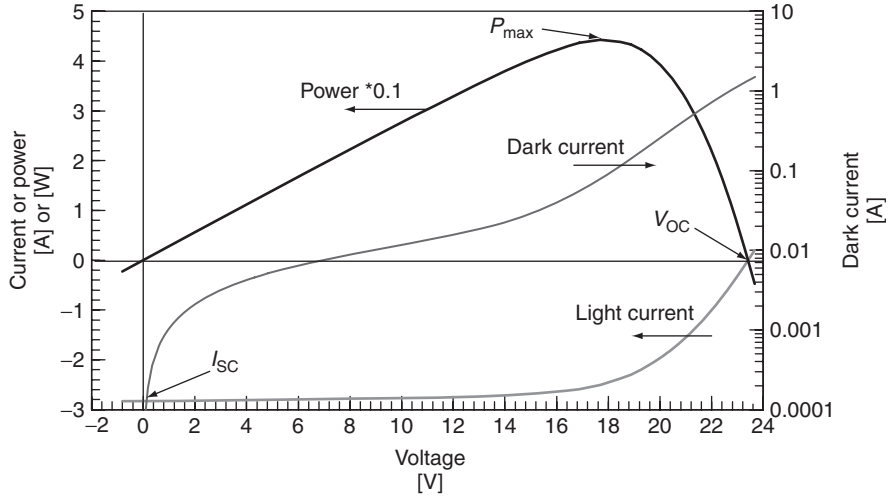


Figure 18.7 Typical light and dark current versus voltage curve of a commercial 50 W PV module

the $I-V$ curve are the open-circuit voltage (V_{oc}), short-circuit current (I_{sc}), and maximum power point (P_{max}). Figure 18.7 shows a typical $I-V$ curve for a 50-W module in the light at SRC and in the dark. The fill factor (FF) is a normalized parameter indicating how ideal the diode properties are, and it is calculated by the following expression:

$$FF = \frac{P_{max}}{V_{oc} I_{sc}}. \tag{18.31}$$

The fill factor is often expressed as a percentage by multiplying Equation (18.31) by 100.

The open-circuit voltage can be determined from a linear fit to the $I-V$ curve around the zero current point or by measuring the voltage with the load disconnected. The value of V_{oc} is often obtained by linear interpolation of the two $I-V$ points closest to zero current. Performing a linear regression using more than two points can reduce the uncertainty in V_{oc}; however, care must be taken not to include points resulting from blocking diodes in series with a module or fitting in nonlinear regions. One manufacturer of commercial equipment includes the twenty $I-V$ points in the power quadrant closest to zero current. Another approach that works for all types of cells and modules is to include in the linear regression fit all points that satisfy the constraint that: (1) the absolute value of the voltage be within 10% of the voltage interpolated to zero current; and (2) the absolute value of the current be less than 20% of the current interpolated to zero voltage.

The value of I_{sc} is usually determined by linear interpolation of the two points closest to zero voltage. Performing a linear curve fit using more than two points can reduce the uncertainty in I_{sc}; however, care must be taken to not include points resulting from bypass diodes in parallel with a module or fitting in nonlinear regions. One manufacturer includes the twenty $I-V$ points in the power quadrant closest to zero voltage. Another approach that works for a wide variety of cells and modules is to include all $I-V$ points in the linear fit that satisfy the constraint that: (1) the current is within 4% of the current interpolated to zero voltage; and (2) the absolute value of the voltage be less than 20% of the voltage interpolated to zero current. These constraints make no assumptions about the spacing between points (fit to a fixed number of $I-V$ points) or the shape of the curve (avoiding the inclusion of nonlinear regions) while including as many $I-V$ points in the linear regression as possible.

P_{max} is often taken to be the largest measured power. A more accurate method is to perform a fourth-order or higher polynomial curve fit to the measured power versus voltage data points within 80–85% of P_{max}. [6, 7, 146]. To prevent erroneous results on low fill factor devices, the power–voltage data selected for curve fitting must be restricted to voltages greater than 80% of the voltage at the measured maximum power (V_{max}). This algorithm can be improved by selecting the order of polynomial that gives the best fit to the data up to a fifth order. An approach recommended by ASTM is to perform a fourth-order polynomial fit to the data where the measured current is greater than $0.75I_{max}$ and less than $1.15I_{max}$ and the measured voltage is greater than $0.75V_{max}$ and less than $1.15V_{max}$ [6, 7, 146]. The preceding constraints seem to work for cells and modules measured by the PV Performance team at NREL with fill factors between 40 and 95% [146].

It is best to measure the incident irradiance at each current–voltage point and to use separate meters for the same sampling interval. A less accurate method is to measure the current and voltage sequentially with the same meter and only measure the intensity once. This method assumes that there are no temporal fluctuations in the intensity of the simulated or natural light during the measurement period and that the voltage is constant. This may not be the case if the bias rate or capacitance is large or if there are transients in the device [34, 146]. One commercial $I-V$ system averages forty voltage readings and then forty current readings to obtain a single $I-V$ point. A better approach would be to take forty pairs of current and voltage points and then average the voltage and current.

Useful additions that are rarely found in commercial $I-V$ equipment include checking for valid contacts and limiting the maximum current through the device. To prevent damage to PV cells, the resistance between the current and voltage contacts should be measured with manual or automatic measurements prior to the $I-V$ measurement. If the sample is too small for Kelvin contacts, then the current at zero voltage should be monitored while making contact to the cell. Most commercial systems are polarity dependent because bipolar power supplies or loads are much more expensive. The current near zero volts is obtained by a low-voltage load of opposite polarity. This feature can damage cells and modules with a low reverse-bias breakdown, such as amorphous silicon or GaAs, if a diode is not placed in parallel with the sample to prevent a reverse-bias voltage from being applied. The polarity of the PV with respect to the polarity of the data acquisition system can easily be determined from the sign on the voltage near zero current or the sign on the current near zero voltage. Commercial systems rarely allow the I-V curve to be swept in both directions which is critical to understand and minimize bias rate artifacts. This allows the operator to not be concerned about which connection is positive and prevents excess bias voltages from accidentally being applied by automatically choosing a safe range for the maximum forward and reverse bias.

Commercial $I-V$ systems developed for the semiconductor industry by Hewlett Packard/Agilent Technologies Inc., Keithley Instruments Inc., and others are readily available. Some are based on source-measurement unit (SMU) which combines the functions of sourcing or sinking V or I while measuring I or V. Unlike most power supplies or electronic loads commercial SMUs typically have a high input impedance, allowing the voltage to be remotely sensed without loading down the sample. SMUs also have a distinct advantage in measuring low currents because they automatically compensate for the variable setting time with current range. The major limitation of commercial SMU's is that they do not have a high enough current range to be able to measure most commercial Si cells. Electronic loads and SMU's can also oscillate because of the high capacitance of some PV materials. These units can be operated manually or with a computer with a wide range of features and capabilities including bipolar operation. The primary problem with units designed for transistor and diode analysis is the cost and limited maximum current. This limitation is only a problem for groups that need to perform $I-V$ measurements at biases greater than $\sim$100 V and $\sim$5 A. Programmable electronic loads are manufactured by a variety of companies including American Reliance, Kepco and Agilent. These transistor loads cover a broad range of current and voltage for a given power rating making

them ideal as a low-cost module $I-V$ tester. These generic $I-V$ systems require software to download and save the data and to calculate relevant PV parameters. A variety of companies also manufacture PV test equipment for $I-V$ measurements, including Berger Lichttechnik GmbH, Daystar Inc., Pasan Beval S.A., Spectrolab Inc., Spire Inc., Wacom Electric Co. Ltd., and numerous other small companies around the world. Commercial $I-V$ measurement software typically is designed for industrial applications and lacks the capability to detect bias-rate artifacts by changing the bias direction, or variable bias, or load slew rate. The current and voltage ranges are also adequate for production testing, but not for research and development. Commercial software also has a fixed format for saving the data and plotting the results that may be difficult or impossible for the user to modify. Most commercial $I-V$ software also measures only the light $I-V$ characteristics in the power quadrant. This is not sufficient for analyzing effects of nonohmic contacts or voltage-dependent photocurrent collection.

Many groups have developed custom data acquisition systems with commercial components. These systems are reviewed in reference [2] and consist of electronics to measure the current and voltage and a power supply [1], operational amplifier [147, 148], capacitor [149], or transistors as the load [148–152]. Ideally, the current sense resistor should have a low 1-year calibration drift, low temperature coefficient and have a power rating at least six times the maximum expected load power to prevent errors from changing ambient temperature or resistor heating. With these constraints the a 1% accurate resistor can be calibrated to the 1-year stability specification. It should be noted that $I-V$ systems built around custom circuits and software can be difficult and time-consuming to maintain when the developer is gone. Isolating the high and low from earth ground avoids unintentional errors from ground loops. This is rarely done in commercial systems. In determining the most appropriate $I-V$ data acquisition system for devices that may range from single research-level cells to arrays, the present and all possible future applications should be considered. Any group that is considering building its first $I-V$ system, upgrading an old reliable $I-V$ system, or expanding existing capabilities should consider the following factors:

- Desired outputs from the system – tabular and graphical display and hard copy of data, database or update of a simple directory text file, control over format and content of saved data, meteorological parameters
- Minimum and maximum current and voltage range
- Cost in time and money available for design and development
- Cost in time and money available for maintenance (repairs, enhancements, and expansion)
- Compatibility with existing hardware, software, and databases
- Flexibility to detect and compensate for artifacts – flexibility in bias direction and bias range, manual control, control over pre-measurement illumination and bias state.

Assuming that the PV device is actually at SRC, the error in the fill factor is primarily affected by the connections to the data acquisition system. First and foremost, four connections should always be used to make contact with the device. These four connections consist of two two-terminal connections (known as Kelvin connections), providing separate voltage and current connections to each side of the device. If only one connection, serving as both the voltage and current contact to one side of the device is used, then any wire or contact resistance between the device and the measurement instrument will appear as a series resistance, reducing FF and hence P_{max} and η. At the component level of modules and above, wire resistance losses are included. For cells without attached wires, the goal is to simulate as accurately as possible the module contacting scheme. This generally means placing a current and preferably a voltage probe contact at each wire bond pad.

For production PV testers, the ribbons attached to grids on Si cells are simulated by using linear arrays of probes (possibly spring-loaded to a specific force). At the cell level, many PV

materials will be damaged if the probe penetrates the relatively thin contact pad or misses the pad and touches the semiconductor. When the contact geometries are small, a micromanipulator and microscope eyepiece may be required, and contacting problems tend to be more difficult. The typical probe contact procedure in the PV performance characterization laboratory at NREL is to choose the appropriately sized Kelvin probe mounted on a three-axis manipulator and make contact to the device, while monitoring the resistance between the voltage and current Kelvin contact. Kelvin probes are used with an attached coaxial cable manufactured by Accuprobe Inc., with CuBe tip diameters between 12.7 and 381 μm. The Accuprobe Z-probe series comes with inline or side-by-side Kelvin probe needles.

A temperature-controlled vacuum plate is used for compatible structures where at least one of the positive or negative contacts is on the front surface (mono- and multicrystalline wafers or thin film devices deposited on a substrate) to provide a large-area, low-resistance contact. This low, but finite, contact resistance will appear as a power loss in P_{max} unless a separate voltage contact is used. This voltage contact may be a miniature, spring-loaded, blunt-tipped, gold-plated probe; a patterned, metallized Kapton or ceramic; a printed-circuit board placed in a narrow slot in the vacuum plate; or some other method. The surface area of this voltage contact will introduce an error in the voltage when light is incident on the cell because of nonuniform heat transfer to the temperature-controlled plate.

For large-area cells with currents above several amps that are designed to have ribbons attached, differences in fill factors can occur of more than 50% between groups [2, 153]. The source of these differences can be attributed to contacting and spatial nonuniformities in the light source. Differences at the 2% level or less can occur because of shading and contacting-related differences when a linear array of spring-loaded probes is used to simulate the ribbons. The magnitude of the differences is related to the ribbon or busbar width, probe assembly height and width, reflections off the probe assembly, and light source divergence. These differences can appear as nonrandom differences in I_{sc} between calibration labs depending on their procedures.

Other sources of errors can be variations in shadowing and reflections, difficulty controlling cell temperature, and distributed resistance losses. Differences in I_{sc} of 2% between NREL and other groups have been attributed to light reflected off of metal probe(s), nearby operators in white lab coats, and fixturing. If the distance between the test plane and the nearest optical surface is short, the possibility may exist for reflection-related artifacts [34]. These reflection-related artifacts occur because the field of view of the reference cell and test device are not identical and can occur from reflections off simulator optics, reference cell packages, probe fixtures, the test-station enclosure, and the region underneath the test device. Light reflecting from the region under the test device is especially important for bifacial cells and superstrate structures. Custom fixturing or optics to direct the light upward is often required when both contacts are on the side not being illuminated, as is the case for point-contact, interdigitated all-back contact, or wrap-around Si cells, or for superstrate structures such as amorphous silicon (a-Si) or cadmium telluride (CdTe) on transparent conducting oxide-coated glass. Achieving temperature control and Kelvin contacting for these structures is problematic. Groups have used patterned circuit boards or Kapton to achieve contacts and temperature control. The contact area is often the junction area for small-area thin film devices on insulating superstrates, allowing a metal spool with a vacuum hold-down to make thermal and electrical contact to the metallized cell and a probe or wire to make contact to the transparent conducting oxide layer. Many of the best thin film devices have a thick layer of indium metal bonded to the transparent conducting oxide around the cell border to reduce resistance losses. High-efficiency research cells on insulating substrates, such as Cu(Ga,In)(S,Se), often also have an In border around the cell area to reduce resistance losses.

18.3.8 Concentrator Measurement Issues

Evaluating the performance of concentrator cells poses several challenges. As of 2010, there are no consensus standards to evaluate concentrator cells. In 2008, the International Electrotechnical Commission Technical Committee TC-82, Working Group 7 for PV concentrators, adopted the de-facto standards given in Table 18.1. Issues of temperature measurement and control are aggravated by the large heat load. Typically, concentrator cells are evaluated under flash systems with a 1-ms pulse duration or under continuous illumination. If a continuous light source is used, then the temperature of the space-charge region cannot be measured directly – any temperature sensor will affect the temperature because of the small thermal mass and large light level on the sample. There can be large differences in the temperature of the space-charge region, which is the cell temperature, and the temperature of the vacuum plate or measured front-surface temperature. One approach is to first set the temperature-controlled vacuum plate to a given temperature with the cell in the dark or with minimal heat load, and then measure V_{oc} as a function of time as a high-speed shutter is opened, exposing the sample to the full light level [154]. The V_{oc} will rise as the shutter is opened and the cell is exposed to the concentrated light; it will go through a maximum as the cell heats up, corresponding to the V_{oc} at the known temperature measured with no heat load. The plate temperature can then be reduced until this V_{oc} is obtained. A spectrally adjustable concentrator simulator is required to ensure that the photocurrent for each junction is at the same level as if illuminated by the reference spectrum, or equally the photocurrent ratios for the various junctions should be the same under the simulator spectrum as under the referenced spectrum. This has been implemented for two junctions with the lamp voltage used to adjust the spectrum and apertures used to adjust the light level [155]. Recently Spectrolab has developed a commercial multi-source solar simulator to tune the spectrum to properly evaluated high-efficiency multijunction cells.

The other primary problem with concentrators is the measurement of E_{tot}. The simplest approach is to determine the one-sun I_{sc} and assume that this value is linear with light level. Another approach is to use calibrated neutral-density filters [2, 156, 157]. Neutral-density filters can be calibrated to better than 1% at a given wavelength using lasers. Typically, however, they have a $\pm 5\%$ variation in transmissivity for wavelengths between 400 and 1100 nm, limiting their usefulness to single-junction devices, such as Si and GaAs, that are insensitive to spectral errors. The linearity can also be inferred by a series of measurements with aperture or changing flash-lamp voltage [137]. Other approaches to determining the linearity involve exposing the cell to low-level periodic sunlight and concentrated sunlight [158, 159]. Ideally, a calibrated linear reference cell should be used, but spatial nonuniformity of the concentrated beam can lead to a larger error than if linearity were assumed. Again, ideally, the spectral responsivity should be measured as a function of bias light level to address the issue that, for nonlinear devices, the spectral error M will change with total irradiance. Groups have developed spectral response systems capable of measuring the responsivity as a function of bias light level to about 200 suns [160, 161].

Concentrator modules cannot normally be measured in solar simulators because the optics are not a point source; the bulb(s) or integration optics will be imaged on the cell, resulting in a much larger spatial variation in intensity than the module would encounter under natural sunlight. Recently, a low-cost, large-area parabolic mirror has been developed to collimate the light from flash simulators for concentrator modules. The beam quality is sufficient for manufactures to verify the performance of the modules prior to shipment [162]. For this reason, concentrator modules are typically evaluated outdoors under natural sunlight over some period of time. The PVUSA method in Equation (18.2) has been adopted as an ASTM standard to evaluate the dc or ac power or energy rating of concentrator modules systems [13, 37, 48]. Concentrator modules and arrays evaluated at Sandia National Laboratories have consisted of the performance (P_{max}, or $I-V$ characteristics) as a function of direct-beam irradiance and heat-sink temperature [156, 157].

18.3.9 Solar Simulators

Solar simulators are used to simulate natural sunlight for repeatable and accurate indoor testing of the $I-V$ characteristics of PV cells or modules. The ideal solar simulator should: (1) have less than $\pm 1\%$ variation in the light level during the $I-V$ measurement period; (2) have less than a $\pm 1\%$ spatial variation in irradiance in and several centimeters above the test plane; and (3) introduce less than a 1% spectral mismatch error between the test and reference cell. These constraints are essential to ensure an uncertainty in the efficiency of less than $\pm 2\%$. Solar simulators are classified according to the spatial nonuniformity of the total irradiance, temporal instability of irradiance, total irradiance within a given field of view, and spectral match to the reference spectrum [163, 164]. Groups that evaluate multijunction cells and modules have found that single-source simulators can introduce large errors in I_{sc}, and P_{max}. Multisource solar simulators suitable for evaluating reference cells have been discussed in Section 18.3.6 [134, 135].

One can correct for the temporal variation in light level during the $I-V$ measurement if the intensity and device current are measured at the same time for each $I-V$ data point. Most commercial and custom $I-V$ systems for continuous light sources do not correct for this temporal variation in the light level, although most groups have procedures in place to correct for long-term drift in the simulator total irradiance over a period of hours or longer. The spatial uniformity varying with time for arc lamps cannot be readily corrected for, although placing the intensity monitor as close to the test device as possible can minimize these effects. A spatially nonuniform light source presents a measurement challenge of determining the average illumination level for a cell or the illumination level of the current-limiting cell in a module [153, 165]. The efficiency will always be reduced for nonuniform illumination compared to uniform illumination at the cell [165, 166] or module [168] level. The standard for classification of spatial nonuniformity assumes a uniform distribution without small area local nonuniformities. This is normally true but may not be the case where the distance to the simulator bulb is less than the size of the bulb. A Class A module solar simulator for spatial nonuniformity according to the standards was recently found to be Class C if the pixel size were reduced from the maximum allowed 20×20 cm to a size of 2×2 cm [163, 164].

There are three types of illumination sources typically used for solar simulators: continuous arc, pulsed arc, and filament lamps. The merits and problems of these different simulators have been compared [33, 169–171].

Commercial continuous Xe arc-lamp solar simulators have an excellent spectral match to the AM0 or terrestrial spectrum, and their point source (small arc volume) with integrating optics achieves a ± 1 to $\pm 3\%$ variation in spatial uniformity. The spectrum of these lamps shifts slightly from the blue to the red during the bulb life, with most of the spectral shift occurring in the first 100 hours of operation [169]. The intensity of continuous arc lamps is controlled by changing the distance from the lamp to the test plane or by changing the current. Pulsed simulators are especially useful for characterizing concentrator cells and for large-area modules. The intensity of pulsed light sources is adjusted by changing the distance from the lamp to the test plane, adjusting an aperture near the flash lamp, or by changing the voltage at which the lamp flashes. The spectrum of pulsed lamps shifts from the blue to the red the greater the number of flashes on the lamp, and is difficult to quantify because of the difficulty in measuring the spectral irradiance of pulsed light sources. This effect is caused by metal from the electrodes being deposited on the lamp envelope with each flash. The spectral match of arc-lamps in the UV and visible spectra is excellent, but is poor in the red (>700 nm) because of the numerous Xe emission lines. Custom filters reduce the magnitude of these lines to manageable levels. These emission lines are reduced for pulsed Xe lamps.

The least expensive small-area light source is a tungsten–halogen lamp with a dichroic filter. The lamp spectrum depends strongly on the operating voltage or current [169]. A shift in the spectrum with bulb age has been attributed to tungsten–halogen lamps, but a careful study revealed

that although the intensity drops with bulb age at a constant current, the shift with bulb age is less than the variation from bulb to bulb out of the same case [169]. To minimize spectral shifts of filament light sources with bulb age, the sources should be run at the same current throughout their life. The distance between the bulb and test plane should be varied to maintain the proper light level. There are a wide variety of tungsten–halogen bulbs with different wattage, lifetime, and voltage ratings. The choice of the most appropriate bulb is a compromise; for example, the ELH bulb has one of the highest wattages, but operates at 120 V and has a short lifetime of 35 hours. Low-voltage bulbs such as the HLX, ELC, or HMM operate below 40 V, eliminating the safety hazard of higher-voltage bulbs, and they have a longer life, but a lower wattage. As with continuous arc-lamps, the lamp lifetime is reduced by frequently turning the lamp on and off. At least one module manufacturer uses an array of tungsten–halogen bulbs for production testing of modules in their multi-megawatt plant. They operate the bulbs at a low "simmer" voltage between $I-V$ measurements to greatly increase the lifetime of the bulbs. They use "matched" reference cells to minimize their sensitivity to spectral errors. The spectral irradiance of filament lamps is characterized by a blackbody spectrum of 3200–3450 K. These lamps are deficient in the blue region of the solar spectrum because the AM0 spectrum can be approximated as a 5900 K blackbody. Sulfur lamps or other light sources with a large area tempered Schott KG3 or KG5 filter have proven to be useful in testing organic and amorphous silicon cells because their spectral output matches the response range of the cells, minimizing extra heating, and the lamps have a long life.

18.4 SPECTRAL RESPONSIVITY MEASUREMENTS

The spectral responsivity ($S(\lambda)$) or quantum efficiency (QE(λ)) is essential for understanding current generation, recombination, and collection mechanisms in photovoltaic devices. PV cell and module calibrations often require a spectral correction factor that uses the spectral responsivity (i.e. Equations (18.16–18.22)). The spectral responsivity is measured either in units of current produced per unit power and can be converted to quantum yield, or in electron–hole pairs collected per incident photon through the equation:

$$QE(\lambda) = \frac{qS(\lambda)}{\lambda hc}. \tag{18.32}$$

The factor hc/q equals 0.80655 for the wavelength in units of μm and the spectral responsivity in units of A/W. The quantum yield, in units of electron per photon, is often multiplied by 100, giving the quantum efficiency.

Typically, the spectral responsivity is measured at short-circuit current because it is easy to define and usually is the same as the photocurrent except for cells exhibiting voltage-dependent current collection such as a-Si devices. PV devices normally operate near their maximum power point. The shape of the spectral responsivity is assumed to be the same at the maximum-power and short-circuit points. Voltage-dependent spectral responsivities have been reported for a-Si [172, 173], CdTe [174], and Cu(Ga,In)(S,Se) [175] material systems. This is an error in the spectral correction factor only if the change with voltage is wavelength dependent.

The PV community has designed a variety of spectral response measurement systems, including ones based on interference filters, grating monochrometers, and interferometers [92, 95, 98, 99, 160, 161, 176–183]. For a single-junction solar cell, $S(\lambda)$ is determined by illuminating the cell with periodic (i.e. "chopped") monochromatic light and continuous broad-band bias light of much greater intensity. The ac photocurrent from the device due to periodic monochromatic light is converted to an ac voltage and measured with a lock-in amplifier. An ac voltmeter may be used instead of a lock-in amplifier if the ac signal is large compared with the ac noise. The measured

a.c. photocurrent due to the monochromatic chopped light is often in the μA to mA range. The broadband dc bias light is to duplicate the generation rate and carrier density in the device when it is at or near its intended operating point, e.g. 1-sun. Significant differences in SR between no bias and ~1 sun bias conditions suggest a large degree of space charge trapping, photoconductivity, or light dependent field or lifetimes.

For a two-terminal multijunction device, measuring the spectral responsivity of the individual junctions requires that the junction to be measured is determining the photocurrent through the device (e.g. is the current-limiting junction). Limiting the current is normally achieved by illuminating the other junctions *not* being measured with a dc bias light whose spectral irradiance covers their response range [179, 183]. To measure $S(\lambda)$ for the *top* cell in a two-junction device, the bottom cell must be illuminated with "red" light that is absorbed mostly in the bottom cell. To measure $S(\lambda)$ for the *bottom* cell in a two-junction device, the top cell must be illuminated with "blue" light that is absorbed mostly in the top cell. In practice, the intensity of the bias light is increased until $S(\lambda)$ for the junction being measured is a maximum and $S(\lambda)$ values for the other junctions are minima. If the multijunction device terminals are at zero volts, then the cell being measured is at some reverse-bias voltage because the other junction is forward biased due to the bias light [179, 183]. Because $S(\lambda)$ can depend on voltage, the cell being measured should be at zero volts [175], which is accomplished by forward-biasing the multijunction cell. If each junction of a two-terminal device has about the same V_{oc}, then the cell should be forward biased to half the V_{oc} of the tandem cell. In practice, the V_{oc} of the individual junctions is not well known; therefore, the forward-bias voltage must be adjusted to maximize the $S(\lambda)$ of the cell being measured and to minimize the $S(\lambda)$ of the other junctions. In practice, for an unknown multijunction device, the procedure is iterative – increasing the bias light intensity and adjusting the bias voltage to maximize $S(\lambda)$ of the cell being measured while minimizing $S(\lambda)$ of the other junctions.

18.4.1 Filter-based Systems

A filter-based spectral responsivity $S(\lambda)$ measurement system is created by shining broadband light source through interference filters and directing the light using a mirror to the device under test, as shown in Figure 18.8 [176]. The filter wheel can be rotated with stepping solenoids controlled by digital logic or stepper motors. The shutter shown in Figure 18.8 is essential when using an ac voltmeter to measure the signal and no monochromatic light is incident on the sample; it is less important, however, when using a lock-in amplifier to measure the periodic monochromatic signal. The monochromatic beam power is measured with a pyroelectric radiometer or calibrated Si detector. The reference detector can measure either the power versus wavelength in real-time, or the power versus wavelength data can be stored in a file. The advantage of real-time calibrations is that intensity fluctuations in the monochromatic beam can be corrected for. The advantage of a stored calibration file is that the measured power is much higher, thus minimizing sensitivity to background light, the system is less complex, and polarization effects associated with a beam splitter are not present.

It is often desirable to measure $S(\lambda)$ of modules consisting of multiple cells in series. The simplest approach would be to illuminate the whole module with ac monochromatic and dc broadband light with the module at 0 V, just as in the case of cells. Because of their high monochromatic light power density and large beam area, filter-based $S(\lambda)$ systems are capable of fully illuminating any commercial module. The problem with this method is that different cells may be current limiting at various wavelengths, and the bias point of the current-limiting cell whose $S(\lambda)$ is being measured is not at 0 V. This problem is solved by voltage biasing, similar to the multijunction $S(\lambda)$ measurements [76, 179, 180, 183]. Figure 18.9 illustrates the geometry for measuring the spectral responsivity of an individual cell in a packaged module where the individual

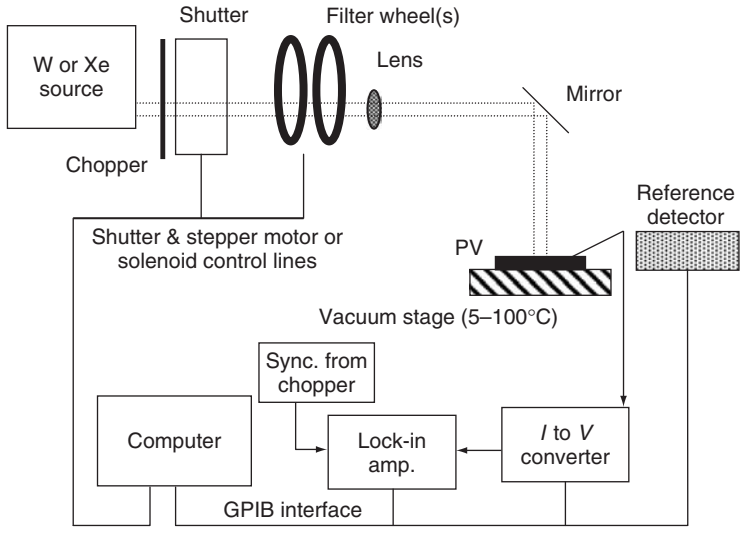


Figure 18.8 Typical narrow-band interference filter based spectral responsivity measurement system. The light is projected from the light source through the filters (monochrometer) on to the mirror which directs the beam to the test plane

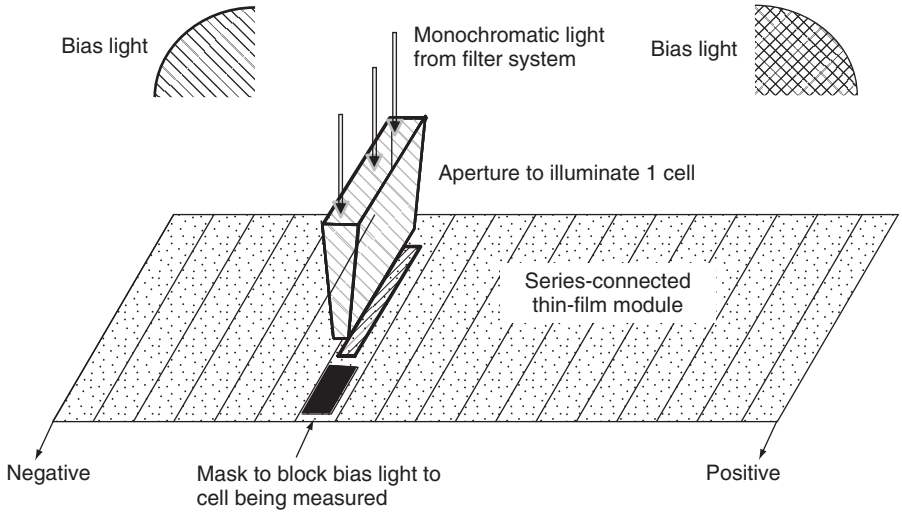


Figure 18.9 Apparatus for measuring the spectral responsivity of a single cell in a multi-cell module

cells are inaccessible. The following sequence of steps is the solution to measuring $S(\lambda)$ of a single cell in a module [12, 176]:

1. Bias the module with light to simulate "1-sun."
2. Forward bias the module to the measured module V_{oc} under the bias light in the previous step multiplied by $(n-1)/n$, where n is the number of cells in series. Another procedure is to apply

monochromatic light at a wavelength that the cell responds to and then reduce the forward-bias voltage from the measured V_{oc} toward 0 V until the ac signal is a maximum.
3. Shine the monochromatic light on only one cell.
4. Reduce the bias light on the cell that sees the monochromatic light in regions where there is no monochromatic light to ensure that this cell is current limiting.
5. Finally, measure the $S(\lambda)$ of the chosen cell in the module.

If the cell's $S(\lambda)$ in the module is not a function of bias light intensity, then the region where the monochromatic light illuminates the cell does not need to be illuminated by a dc bias light. To measure each junction in a multijunction module, the spectral content of the bias light must be adjusted and the voltage bias and intensity of the bias light must be iterated. These procedures have been shown to produce the same relative spectral responsivity for an electrically isolated cell in a module as when all of the cells in the module were series-connected [176, 180].

18.4.2 Grating-based Systems

The grating system shown in Figure 18.10 was developed to measure the responsivity of thermophotovoltaic cells from 400 to 3200 nm. Grating-monochrometer-based systems are especially useful for their broad wavelength range and high spectral resolution. If a double-grating monochrometer is used, then the stray light in the UV can be eliminated, which is important for UV or high-bias-light measurements [95, 98, 161, 177]. Long-wave-pass, order-sorting filters are commonly used to suppress modes (e.g. $\frac{1}{2}\,\lambda/m$ where $m = 2$, 3, etc.) at shorter wavelengths. For example, a Schott WG360 color glass filter is commonly used for $S(\lambda)$ measurements in the 400–700 nm region, and a Schott RG630 filter is used as an order-sorting filter for measurements over a 700–150 nm wavelength range. If a single-grating monochrometer is used with a tungsten light source, a band-pass filter may be needed for measurements in the 300–600 nm wavelength region to suppress modes at longer and shorter wavelengths (e.g., $\frac{1}{2}\,\lambda$ and 2λ). The light power of the grating-monochrometer-based system can be focused to a rectangular spot of about 1 × 3 mm by imaging the monochrometer exit slits onto the test plane with a magnification of less than 1. Chromatic

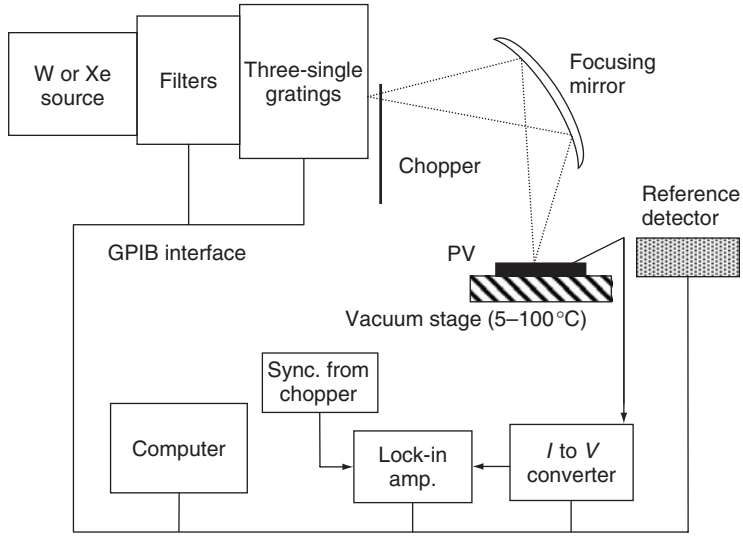


Figure 18.10 Typical grating-monochrometer-based spectral responsivity measurement system

aberrations in the len(s) cause the beam size to change with wavelength. This effect can be eliminated by eliminating all lens and using a spherical – or better yet, a parabolic – mirror. Typically, grating-based systems have lower optical throughput (lower intensity), but higher spectral resolution than filter-based systems. The FWHM can be adjusted by varying the monochrometer slit width or grating resolution.

18.4.3 Spectral Responsivity Measurement Uncertainty

Spectral responsivity measurements involve the measurement of the photocurrent produced by light of a given wavelength and power. The spectral responsivity is typically measured with bias light simulating reference conditions, because the device may be nonlinear [92, 95, 98, 99, 176–180, 183]. Typically, the spectral correction factor for efficiency measurements is calculated based on $S(\lambda)$ measurements near 0 V and is assumed to be the same as at the maximum power point. This assumption is valid for most PV systems and results in a negligible error for amorphous silicon, which has a voltage-dependent spectral responsivity [172–175], assuming a reasonably well-matched reference cell is used, such as a Schott KG5 filtered mono-Si cell.

The photocurrent is measured with a current-to-voltage converter. Many groups use a high-current, low-noise operational amplifier with a gain of 10–10 000 as the current-to-voltage converter because commercial current amplifiers saturate around 10 mA. A current-to-voltage converter developed by Carl Osterwald and used by the PV Cell and Module Performance Characterization Team at NREL since 1983 is based on an Apex power-operational amplifier (±40 V, 8 A) with computer-controlled gain resistors (50–10 000 units). This converter is useful for wide bias ranges and signal levels (Figure 18.11) [176]. A simple current-sense resistor may be adequate for systems that measure the same type of PV device all the time. The major limitation is that resistor and thermal noise at the microvolt level limits the measured currents to microamperes. Commercial current preamplifiers typically have a maximum current rating of 1–10 mA, limiting their usefulness for measuring the spectral responsivity with bias light (i.e. a 1-cm^2 device with short-circuit current density $J_{sc} = 30$ mA/cm^2 produces 30 mA of dc bias current with a 1-sun bias light). An operational amplifier configured as a current-to-voltage converter allows the insertion of a power supply in series with the PV device, giving a wide range of bias voltages. This feature is critical when measuring modules, multijunction devices, or devices with a voltage-dependent spectral responsivity [132, 172–176, 179, 180, 183]. Most groups use a lock-in amplifier to detect the periodic ac signal, but this is not required because the intense monochromatic light afforded with interference filters allows the ac signal to be amplified to the range where ac voltmeters are quite accurate [88]. Modern digital lock-in amplifiers have rapid auto-ranging capabilities and will outperform an ac voltmeter for noisy signals because of their large dynamic range. Error sources related to the measurement of the photocurrent are summarized in Table 18.5. If semiconductor-based calibrations are employed with the same electronics used to measure the test and reference device, then all multiplicative errors drop out. For pyroelectric-radiometer-based calibrations, the absolute photocurrent must be measured for absolute $S(\lambda)$ measurements.

For absolute current measurements, the measured lock-in signal must be multiplied by a waveform correction factor that relates the measured root-mean-square (RMS) signal with the peak signal. This factor is $\sqrt{2}/2$ for a sine wave, $2\sqrt{2}$ for a square wave, and $2\sqrt{2} \cdot a \cdot \sin(\pi/a)/\pi^2$ for a trapezoid, with the constant π/a being the radian angle at the top of the rising edge of the trapezoidal waveform [184].

The response time of PV devices to chopped light can be a problem for electrochemical cells or those with many deep-level recombination centers. Similar to results reported elsewhere, chopping frequencies below 4 Hz are required to keep the ac photo-response independent of frequency [185]. This effect is more pronounced at low light levels and in the infrared. However,

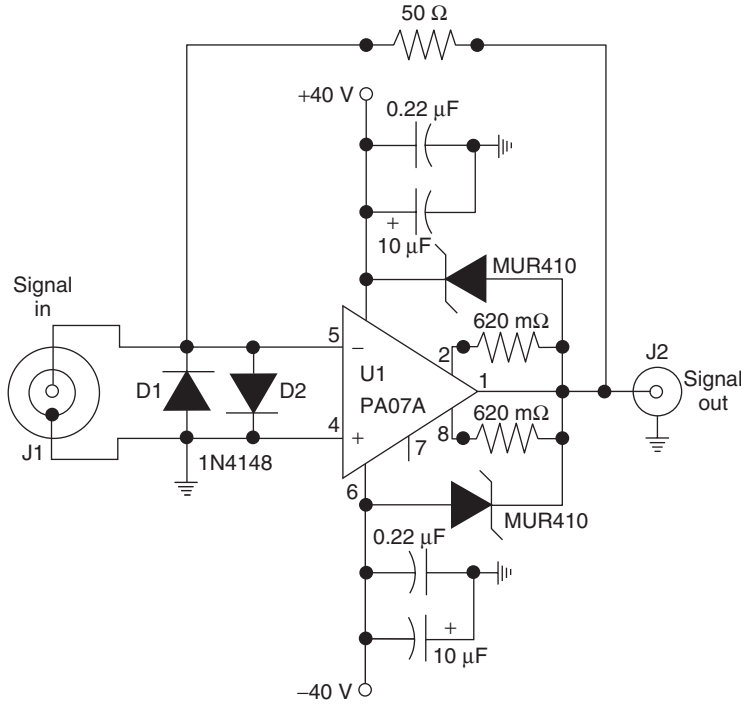


Figure 18.11 Current to voltage converter used for spectral responsivity measurements by the PV Cell and Module Performance Characterization Team in The National Center for Photovoltaics at NREL since 1983 [120, 176]

Table 18.5 Spectral responsivity error sources for measurement of the photocurrent

I. Electrical Instrumentation
  A. current-to-voltage (I to V) converter
    1. commercial current amplifier, lock-in amplifier or custom amplifier gain, linearity, noise, offset, shunt resistor, calibration, drift, thermovoltages
  B. signal from I to V converter measured with
    1. lock-in amplifier
      calibration, resolution, accuracy, waveform to sine wave correction factor, overloading, noise, dynamic range, time-constant, procedures for using lockin amplifier
    2. an ac voltmeter
      gain, offset from noise level, linearity, time-constant

II. PV cell or module
  A. temperature, response-time to periodic light, linearity of PV device, white-light bias spatial uniformity, monochromatic light spatial uniformity, voltage bias of cell being measured, spectral content of bias light, device sensitivity to polarization of light

III. Mechanical
  A. mechanical movement of optics, mechanical vibration, monochromatic beam wandering with wavelength, chopped stray monochromatic light

most inorganic solar cells can be adequately measured between 50–100 Hz, a comfortable range for most mechanical choppers and lock-in amplifiers. It is important that light from the bias light source not be allowed to go through the light chopper. A simple procedure to determine if this artifact is present is to turn off the monochromatic light source and measure the test device's ac response as a function of bias light intensity (it should be zero).

Semiconductor-based calibrations are useful where the photocurrent is known within a multiplicative constant. Their primary limitation is their temperature sensitivity near the energy gap (wavelengths greater than 900 nm for Si) and their limited response range. Calibrated semiconductors are not commercially available for wavelength ranges beyond 300–1700 nm (Si-Ge hybrid detector). Accurate calibrations for semiconductor-based detectors are difficult to obtain for wavelengths greater than 1800 nm. If the same amplifier is used to measure the reference and unknown PV devices, then uncertainties in the gain drop out. For semiconductor calibrations using a dual channel lock-in amplifier, the chopper phase is irrelevant, whereas the electrically calibrated pyroelectric radiometer typically requires that the chopper be manually adjusted until the phase is correct. Semiconductor-based calibrations allow the test and reference signals to be filtered independently to maximize the signal-to-noise ratio. Various error sources associated with measuring the monochromatic light power are listed in Table 18.6. The measurement of the monochromatic light power can be performed with radiometric detectors or semiconductor detectors. When a quartz slide is used as the beam splitter, errors in the power can then arise because of polarization effects. The light off the monochrometer is polarized, and the polarization angle can change with a grating change. The bandgap, photoluminescence, and absorption coefficient for PV devices can be sensitive to the polarization angle. The light reflected from a glass surface will have a different polarization from the light reaching the test plane and will be of much lower intensity. These effects are minimized if a calibration is performed with the detector in the test plane and the file is stored to disk. This

Table 18.6 Spectral responsivity error sources for measurement of the light power

I. Filament or Xe arc light source
   A. Intensity fluctuations, change in spectrum with age and current
II. Real-time calibration
   A. Source-light polarization with a glass beam splitter, signal to noise, detector characteristics, calibration drift with time of monitor detector
III. Stored calibration file
   A. Monochromatic source calibration drift with time
IV. Stray light
   A. Detector sees light that cell does not see, area of detector different from device area, different field of views
   B. Monochrometer – incomplete attenuation of higher and lower grating orders, single versus double grating
   C. Narrow-bandwidth filters – pinholes in the filter, degradation of blocking filter, insufficient blocking of light ($>10^{-4}$)
V. Detectors and associated electronics in general
   A. Calibration, resolution, accuracy, gain, phase, offset, linearity, spatial uniformity of detector element
   B. Drift in temperature of room, change in the detector's field of view, degradation of detector
   C. Spectral response of detector
VI. Pyroelectric detector
   A. Time constant of detector, microphonics, signal to noise, phase-angle adjustment, waveform factor (square wave assumed)

Table 18.7 Spectral responsivity error sources related to the monochromatic light

I.	Bandwidth, filter defects, polarization variation with wavelength
II.	Wavelength offset, wavelength error, wavelength variation with room temperature
III.	Beam wanders with wavelength
IV.	Beam larger than the test device
	A. Detector area versus PV area, position of detector and PV different, spatial uniformity of beam
V.	Beam smaller than detector and device area
	A. Partially shaded regions, spatial variation in responsivity of PV

procedure is required at least once for real-time calibrations. Real-time calibrations account for the change in spectrum with lamp age, current, and time.

If the beam is larger than the sample, then the spatial uniformity of the monochromatic beam is important. For the NREL filter monochrometer system, spatial nonuniformities of $\pm 10\%$ are typical; more importantly, these errors can change with wavelength because of variations in the transmission of the filter and spatial variation in the output of the Xe arc-lamp. Electrically calibrated pyroelectric detectors are spectrally flat from the UV to far IR and have a low broadband error of less than $\pm 2\%$; but they are sensitive to microphonics, temperature changes within the detector's field of view, and have noise at the $0.01\,\mu W\,cm^{-2}$ level. Semiconductor-based detectors are not sensitive to light outside their relatively narrow response range and can be measured with the same electronics used to measure the test device, thus eliminating any wavelength-independent multiplicative error sources. Semiconductor-based detectors can drift with age [186] and have temperature coefficients exceeding $1\%/^{\circ}C$ near their bandgap. Table 18.7 summarizes the error in the quantum efficiency that can occur because of the monochromatic light. The bandwidth of the monochromatic light can contribute to the error near the bandgap or when the light transmitted through a bandpass filter is highly asymmetric [184]. These errors have a small effect on the spectral correction factor when the FWHM bandwidth of the interference filters are less than 10 nm and the response range is broad, as with Si [184].

18.5 MODULE QUALIFICATION AND CERTIFICATION

PV modules for grid-tied power generation applications are designed to last 20 years or more in the field with no maintenance. Modules are designed to withstand daily thermal cycling, hail, wind, sand, and storms, along with prolonged exposure to UV light. For safety reasons, PV cells and wiring must be isolated from the frame, edges, or module surface to prevent a hazardous electrical potential from forming. To verify that a particular product is reliable, domestic and international consensus standards have been developed for the design qualification and type approval of terrestrial crystalline silicon and thin-film modules [187–190]. A survey of the history behind the development of these standards was recently published [191].

Ideally, modules should be subjected to long-term exposure under natural conditions. Flat-plate silicon PV modules had poor reliability in the first years of development. An extensive program was conducted in the late 1970s and early 1980s by the Jet Propulsion Laboratory for the US government in an effort to improve the reliability of PV modules [192]. In the 1980s, the flat-plate solar array program at JPL identified the following 13 principal Si module flat-plate module degradation mechanisms: open-circuited cell regions; shorted cells; open-circuited interconnects; gradual cell power degradation; optical degradation of the module package; front-surface soiling; glass breakage; open circuits in the module wiring; hot-spot failures of cells in a module; shorted

bypass diodes; shorts to the frame or ground; delamination of the module encapsulant; and life-limiting wear-out [192].

Accelerated testing procedures have been developed to simulate 20 or more years of exposure in the field and to identify the major failure mechanisms. Unfortunately, passing the tests does not guarantee a 20-year life because not all of the failures mechanisms and modes have been identified and not all stress factors, such as UV exposure, simulate extended exposure. Accelerated tests have been developed by various governmental organizations such as the American Society of Testing and Materials (ASTM) [189] or the Commission of European Communities [190, 192–194]. The tests described are from the International Electrotechnical Commission (IEC), the international standards organization relevant for photovoltaics [187, 188]. These tests include safety tests, mechanical integrity tests, and thermal cycling.

The thermal-cycle test determines the ability of a module to withstand thermal expansion coefficient mismatch, fatigue, and other stresses caused by repeated changes in temperature, and it includes 200 thermal cycles from -40 to $+85\,^{\circ}\mathrm{C}$. The damp-heat test, designed to determine the module's ability to withstand the effects of long-term exposure to moisture in high-humidity environments, consists of 1000 hours at 85% relative humidity at $85\,^{\circ}\mathrm{C}$. The humidity-freeze test determines the ability of the module to withstand the effects of high temperature and humidity followed by sub-zero temperatures, and it consists of 10 cycles from 85% relative humidity and 85 to $-40\,^{\circ}\mathrm{C}$. This is not a thermal-shock test because the maximum change in temperature with time above freezing temperatures is $100\,^{\circ}\mathrm{C\ h^{-1}}$. This humidity-freeze test reveals the detrimental accumulation of liquid water inside the module under high-humidity conditions. Because modules can have a substantially inferior performance under low light levels when they have a low shunt resistance, modules are tested at an irradiance of $200\,\mathrm{W\,m^{-2}}$ ($\sim$only 20% of 1-sun). Modules exposed to 60 kW h $\mathrm{m^{-2}}$ of total solar irradiation are required to sustain a loss in power of less than 5% to reveal any synergistic degradation effects that may not be detected by other tests.

The hot-spot test is designed to ensure that modules will not fail when hot-spots occur because of mismatched cells, partial shading, or interconnect failures. A hot-spot is a localized region in a module that is operating at a significantly higher temperature ($\sim$5 to $40\,^{\circ}\mathrm{C}$) than the rest of the module. The hot-spot test consists of five 1-hour exposures at $1000\,\mathrm{W\,m^{-2}}$ irradiance under the worst-case hot-spot condition. The modules must also survive a hail test of 25-mm-diameter ice balls directed at 11 impact locations with a velocity of $23\,\mathrm{m\,s^{-1}}$ [187, 188]. Modules are also subjected to a twist test to detect module defects that might arise when the module is mounted on a nonplanar surface. The twist test consists of supporting the module on three of its corners and deflecting the fourth corner 1.2° with respect to the plane defined by the other three corners [187, 188]. Modules are also expected to withstand wind, snow, or ice loads without mechanical or electrical failure. The static-load test simulates a wind load of $130\,\mathrm{km\ h^{-1}}$, and consists of mounting the module in the manner proscribed by the manufacture and applying a force of 2400 Pa uniformly to the front surface for 1 hour and then to the rear surface for 1 hour. The insulation test is a safety test designed to determine if the current-carrying parts of the module are sufficiently well isolated from the module edges or frame and requires a resistance of greater than $50\,\mathrm{M\,\Omega}$ at 500 V dc resistance to ground, and less than $50\,\mu\mathrm{A}$ current to ground at an applied voltage of 1000 V plus twice the maximum system voltage [187, 188]. For frameless modules, ground is obtained by attaching a metal conductor to the outside perimeter.

The wet-leakage-current test is a stringent safety test to ensure the current-carrying parts are well isolated, preventing a ground-fault condition from occurring, even when the module is wet. The wet-leakage test is also designed to verify that moisture does not enter the active part of the module, where it might cause delamination or corrosion. The module is placed in a tank of water with resistivity of $3500\,\Omega$ cm or less at $22 \pm 3\,^{\circ}\mathrm{C}$ and a surface tension less than $3\,\mathrm{N\,m^{-2}}$, and the

leakage current at 500 V is measured. The maximum allowed leakage current for the wet-leakage test (sometimes called the wet hi-pot test) is 10 μA plus 5 μA times the surface area in m^2. The wet-leakage test is performed before and after the various stress tests. The number of modules, sequence of events, and the pass/fail criteria are specified in the standard documents [187, 188]. Thin film module testing procedures also require the $I-V$ characteristics under SRC to be measured periodically during light exposure (between 800 and 1000 $W\,m^{-2}$, between 40 and 60 °C) until the module power changes by less than 2% over two consecutive light soaking periods of at least 43 $kWhm^{-2}$ to determine the stabilized performance before and after stress [188].

Recently, the IEC standards committee (IEC TC-82 Working Group 7) developed qualification standards for concentrator modules [196]. The standard is similar to the flat-plate standard but requires on-sun testing. Separate segments can be qualified because the whole module or segment may be too large for environmental chambers. The standard had to take into account the wide range of maximum operating temperature from ambient to over 40 °C above ambient temperature. For example, the dish in a dish-based concentrator, or the front lens in a lens-based one are at ambient temperature while the cell and heat sink are above ambient temperature. The committee is currently developing a tracker standard and a performance rating standard for concentrator modules.

The standards community is considering the development of test-to-failure (TTF) protocols to uncover failure mechanisms and modes that are not evident from normal qualification testing [197]. The TTF protocol is not a true reliability test or service-life prediction test that is designed to determine the mean-time-between failure because it does not uncover every failure mode or mechanism. By extending the qualification test's number of thermal cycles and hours of damp heat at 85 °C, 85% relative humidity and adding voltage bias a quantitative assessment of the relative reliability of different modules can be obtained.

18.6 SUMMARY

The state-of-the-art in rating the performance of cells and modules has been summarized. The uncertainty in the peak watt rating for modules is larger than the community desires. The limitation of ±1% variation among the most accurate recognized primary terrestrial and extraterrestrial calibration groups around the world is a significant contribution to this error [92, 108]. Additional module error sources related to spatial nonuniformity of the light source, bias rate, spectral error and other factors contribute to the ±2% variation in the maximum power rating for Si modules and significantly larger for thin-films [98, 99, 126–130]. Standards documents have been written giving general guidance on how to test PV. ASTM standards can be obtained from their web site http://astm.org/. The relevant photovoltaic standards are in volume 12.02 and are managed by committee E44.09. The ASTM standards mentioned in this chapter are given in references [6–8, 10, 12–14, 28, 32, 90, 103, 163, 177, and 189]. Internationally, IEC standards are most commonly used and can be purchased at http://www.iec.ch/. The relevant IEC PV standards committee is TC-82. There are 10 parts to IEC 60904 containing PV standards related to testing. The other major IEC PV module standards are 60891, 61215, 61646 and 62108 [78, 187, 188, 196]. The IEC standards mentioned in this chapter are given in references [11, 15, 78, 91, 164, 187, 188, and 196]. International standards for the space PV business are covered in ISO at http://www.iso.org. The most significant ISO PV performance standard is given in reference [17] which also documents the primary AM0 reference cell calibration methods used around the world.

The internationally accepted reference spectra can be purchased as standards documents from ASTM or IEC or downloaded from the NREL web site http://rredc.nrel.gov/solar/spectra/am1.5/ which submitted these standards to the ASTM and IEC committees for approval.

ACKNOWLEDGEMENTS

This work was supported under contract DE-AC36-08GO28308. Table 18.3 was created using SAM by Chris Thompson and Steve Hegedus of the Institute of Energy Conversion, University of Delaware. The contributions of the NREL PV Performance team (A. Anderberg, P. Ciszek, C. Mack, T. Moriarty, C. Osterwald, L. Ottoson, S. Rummel and R. Williams) are acknowledged.

REFERENCES

1. Emery K, Osterwald C, *Solar Cells* **17**, 253–274 (1986).
2. Emery K, Osterwald C, *Current Topics in Photovoltaics* **3**, Chap. 4 (1988).
3. Heidler K, Raicu A, Wilson H, *Proc. 21st IEEE PVSC*, 1017–1022 (1990).
4. Nann S, Emery K, *Solar Energy Matls*. **27**, 189–216 (1992).
5. Terrestrial Photovoltaic measurement Procedures, NASA Tech. Report TM 73702, June 1977.
6. Standard ASTM E948, *Standard Test Method for Electrical Performance of Non-Concentrator Photovoltaic Cells Using Reference Cells*, Amer. Society for Testing Matls., West Conshocken PA, USA.
7. Standard ASTM E1036, *Standard Test Methods for Electrical Performance of Nonconcentrator Terrestrial Photovoltaic Modules and Arrays Using Reference Cells*, Amer. Society for Testing Matls., West Conshocken PA, USA.
8. Standard ASTM E490-00, *Standard for Solar Constant and Air Mass Zero Solar Spectral Irradiance Tables*, Amer. Society for Testing Matls., West Conshocken PA, USA.
9. Standard Commission of the European Community, CEC 101, Issue 2, EUR-7078 EN, *Standard Procedures for Terrestrial Photovoltaic Measurements* (1981).
10. Standard ASTM G159-98, *Standard Tables for References Solar Spectral Irradiance at Air Mass 1.5: Direct Normal and Hemispherical for a 37° Tilted Surface*, Amer. Society for Testing Matls., West Conshocken PA, USA.
11. Standard IEC 60904-1, *Photovoltaic devices Part 1: Measurement of photovoltaic current-voltage characteristics*, International Electrotechnical Commission, Geneva, Switzerland.
12. *STANDARD ASTM E2236*, *Standard Test Methods for Measurement of Electrical Performance and Spectral Response of Nonconcentrator Multijunction Photovoltaic Cells and Modules*, Amer. Society for Testing Matls., West Conshocken PA, USA.
13. Standard ASTM E2527, Standard Test Method for Rating Electrical Performance of Concentrator Terrestrial Photovoltaic Modules and Systems Under Natural Sunlight, Amer. Society for Testing Matls., West Conshocken PA, USA.
14. Standard ASTM G173-03, *Standard Tables for References Solar Spectral Irradiance at Air Mass 1.5: Direct Normal and Hemispherical for a 37° Tilted Surface*, Amer. Society for Testing Matls., West Conshocken PA, USA.
15. Standard IEC 60904-3, *Measurement Principles for Terrestrial PV Solar Devices with Reference Spectral Irradiance Data editions 1 and 2*, International Electrotechnical Commission, Geneva, Switzerland.
16. Wehrli C, *Extraterrestrial Solar Spectrum*, Physikalish-Meterologisches Observatorium and World Radiation Center, Tech. Report 615, Davos-Dorf, Switzerland, July 1985.
17. ISO 15387:2005, Space systems – Single-junction space solar cells – Measurement and calibration procedures, International Organization for Standardization, Geneva, Switzerland.
18. Bird R, Hulstrom R, Riordan C, *Solar Cells* **14**, 193–195 (1985).
19. Gueymard, C., *Solar Energy* **71**, 325–346 (2001).
20. Gueymard, CA, Myers D, Emery K, *Solar Energy*, **73**. 443–467 (2002).
21. Kurtz S, Myers D, Townsend T, Whitaker C, *et al*., *Solar Energy Materials and Solar Cells* **62**, 379–391 (2000).

22. Myers D, Kurtz S, Emery K, Whitaker C, *et al*., *Proc. 28th IEEE PVSC*, pp 1202–1205 (2000).
23. King D, Siegel R, *Proc. 17th IEEE PVSC*, pp 944–951 (1984).
24. Osterwald C, *Proc. 18th IEEE PVSC*, pp 951–956 (1985).
25. Faine P, Kurtz S, Riordan C, Olson J, *Solar Cells* **312**, 259–278 (1991).
26. Aerosol Robotic Network (AERONET), Goddard Space Flight Center, http://aeronet.gsfc.nasa.gov.
27. *Solar Cell Array Design Handbook*, Jet Propulsion Laboratory Tech. Report SP 43-38, volume 1, October 1976.
28. Standard ASTM Standard E490-73, *Standard for Solar Constant and Air Mass Zero Solar Spectral Irradiance Tables*, Amer. Society for Testing Matls., West Conshocken PA, USA.
29. Makarova Y, Kharitonov A, Distribution of Energy in the Solar Spectrum, translated from Russian in NASA Tech. Report TT F-803, June 1974.
30. Frölich C, Lean J, Total Solar Irradiance Variations: the construction of a composite and its Comparison with Models, *Int. Astronomical Union Symp. 185: New Eyes to See inside the Sun and Stars*, Kluwer Academic Publ., Dortrecht The Netherlands, pp 89–102 (1998).
31. Lee R, Barkstrom B, Luther M *Proc. 6th conf. On Atmospheric Radiation*, American Meteorological Society, Boston, MA (1986).
32. Standard ASTM E1328, *Standard Terminology Relating to Photovoltaic Solar Energy Conversion*, Amer. Society for Testing Matls., West Conshocken PA, USA.
33. Emery K, *Solar Cells* **18**, 251–260 (1986).
34. Emery K, Field H, *Proc. 24th IEEE PVSC*, pp 1833–1838 (1994).
35. Green M, Emery K, Hishikawa Y, Warta W, Solar Cell Efficiency Tables (version 36), *Prog in Photovoltaics* **18**, 346–352 (2010).
36. Firor K, *Proc. 18th IEEE PVSC*, pp 1443–1448 (1985).
37. Jennings C, *Proc. 19th IEEE PVSC*, pp 1257–1260 (1987).
38. Taylor R, *Solar Cells* **18**, 335–344 (1986).
39. Fuentes M, A Simplified Thermal Model for Flat-Plate Photovoltaic Arrays, Sandia Tech. Report Sand 85-0330, May 1987.
40. Koltay P, Wenk J, Bücher K, *Proc. 2nd World Conf. PVSEC*, pp 2334–2337 (1998).
41. Fisher H, Pschunder W, *Proc. 10th IEEE PVSC*, pp 404–411 (1973).
42. DeWolf S, Choulat P, Szlufcik, Périchaud, *et al*., *Proc. 28th IEEE PVSC*, pp 53–56 (2000).
43. Luft W, von Roedern B, Stafford B, Waddington D, *et al*., *Proc. 22nd IEEE PVSC*, pp 1393–1398 (1991).
44. Luft W, von Roedern B, *Proc. 24th IEEE PVSC*, pp 850–853 (1994).
45. DelCueto J, von Roedern B, *Prog in Photovoltaics* **7**, 101–112 (1999).
46. Meyer T, Schmidt M, Harney R, Engelhardt F, *et al*., *Proc. 26th IEEE PVSC*, pp 371–374 (1997).
47. Ruberto M, Rothwarf A, *J. Appl Phys*. **61**, 4662–4669 (1987).
48. Hester S, Townsend T, Clements W, Stolte W, *Proc. 21st IEEE PVSC*, pp 937–943 (1990).
49. Gay C, Rumburg J, Wilson J, *Proc. 16th IEEE PVSC*, pp 1041–1046 (1982).
50. Gianoli-Rossi E, Krebs K, *Proc. 8th EC PVSEC*, pp 509–514 (1988).
51. Kroposki B, Myers D, Emery K, Mrig L, *et al*., *Proc. 25th IEEE PVSC*, pp 1311–1314 (1996).
52. Kroposki B, Marion W, King D, Boyson W, Kratochvil J, *Proc. 28th IEEE PVSC*, pp 1407–1411 (2000).
53. Whitaker C, Townsend T, Newmiller J, King D, *et al*., *Proc. 26th IEEE PVSC*, pp 1253–1256 (1997).
54. Marion B, Kroposki B, Emery K, del Cueto J, *et al*., Validation of a Photovoltaic Module Energy Ratings Procedure at NREL, NREL Tech. Report NREL/TP-520-26909, Aug. 1999.

55. Marion W, Urban K, (1995), User's Manual for TMY2s (Typical Meteorological Years) – Derived from the 1961-1990 National Solar Radiation Data Base, NTIS/GPO Number: DE95004064, NREL Tech. Report TP-463-7688, 1995.
56. Nann S, Riordan C, *Journal of Appl. Meteorology* **30**, 447–462 (1991).
57. Raicu A, Heidler K, Kleiss G, Bücher K, *Proc. 11th EC PVSEC*, pp 1323–1326 (1992).
58. Raicu A, Heidler K, Kleiß G, Bücher K, *Proc. 22nd IEEE PVSC*, pp 744–749 (1991).
59. Bücher K, *Proc. 23rd IEEE PVSC*, pp 1056–1062 (1993).
60. Bücher K, Kleiss G, Bätzner D, *Proc. 26th IEEE PVSC*, pp 1187–1191 (1997).
61. Gilman, P; Blair, N; Mehos, M; Christensen, .; Janzou, S; Cameron, C (2008). Solar Advisor Model User Guide for Version 2.0, NREL Report No. TP-670-43704, 2008.
62. King D, Boyson W, Kratochvil J, Photovoltaic Array Performance Model, Sandia Report No. SAND2004-3535, 2004.
63. Jardine C, Conibeer G, Lane K, *Proc. 17th European PVSEC*, pp 724–727 (2001).
64. Gregg A, Parker T, Swenson R, *Proc. 31st IEEE PVSC*, pp 1587–1592 (2005).
65. van Cleef M, Lippens P, Call J, Proc. 17th European PVSEC, pp 565–568 (2001).
66. Akmahmad K, Kitamura A, Yamamoto F, Okamoto H, Hamakawa Y, *Solar Energy Materials and Solar Cells* **46**, 209–218 (1997).
67. Bergman J, T*ech. Digest PVSEC-14*, pp 1071–1074 (2004).
68. Holland S M, *Photon International*, 10–11 (2000).
69. Itoh M, Takahashi H, Fuji T, Takakura H, Hamakawa Y, and Matsumoto Y, *Solar Energy Materials and Solar Cells* **67**, 435–440 (2001).
70. King D, Dudley J, Boyson W, *Proc. 26th IEEE PVSC*, pp 1295–1297 (1997).
71. Green M, *Solar Cells Operating Principles, Technology and System Application*, University of New South Wales, Kensington, Australia (1992).
72. Emery K, Burdick J, Caiyem Y, Dunlavy D, *et al*., *Proc. 25th IEEE PVSC*, pp 741–744 (1996).
73. Osterwald C, Glatfelter T, and Burdick J, *Proc. 19th IEEE PVSC*, pp 188–193 (1987).
74. Kameda M, Sakai S, Isomura M, *et al*., *Proc. 25th IEEE PVSC*, pp 1049–1052 (1996).
75. Hisikawa Y, Okamoto S, *Solar Energy Matr. and Solar Cells*, **33**, 157–168 (1994).
76. Sandstrom J, A Method for Predicting Solar cell Current-Versus Voltage Characteristics as a Function of Solar Intensity and Cell Temperature, JPL Tech. Report TR 32-1142, Jet Propulsion Laboratory, Pasadena, CA, USA, July 1967.
77. Klise G and Stein J, Models Used to Assess the Performance of Photovoltaic Systems, Sandia Report No. SAND2009-8258, 2009.
78. Standard IEC 60891, *Procedures for temperature and Irradiance corrections to measured I-V characteristics of crystalline silicon photovoltaic devices*, International Electrotechnical Commission, Geneva, Switzerland.
79. King D, Kratochvil J, Boyson W, Bower W, *Proc. 2nd World Conf. PVSEC*, 1947–1952 (1998).
80. Kasten F, Young A, *Appl. Optics* **28**, 4735–4738 (1989).
81. Hishikawa Y, Tsuno1 Y, Kurokawa K, *Proc. 21st EEC PVSEC*, pp 2093–2096 (2006).
82. Marion B, Rummel S, Anderberg A, *Prog in Photovoltaics* **12**, 593–607 (2004).
83. Emery K, Osterwald C, Cannon T, Myers D, *et al*., *Proc. 18th IEEE PVSC*, pp 623–628 (1985).
84. Emery K, Osterwald C, *Solar Cells* **27**, 445–453 (1989).
85. Osterwald C, *Solar Cells* **18**, 269–279 (1986).
86. Field H, Emery K, *Proc. 23rd IEEE PVSC*, pp 1180–1187, (1993).
87. Shimokawa R, Miyake Y, Nakanishi Y, Kuwano Y, *et al*., *Solar Cells*, **19**, 59–72 (1987).
88. Emery K, Waddington D, Rummel S, Myers D, Stoffel, T, *et al*., SERI Results from the PEP 1987 Summit Round Robin and a Comparison of Photovoltaic Calibration Methods, NREL Tech. Report TR-213-3472. March 1989.

89. Myers D, Emery K, Stoffel T, *Solar Cells* **27**, 455–464 (1989).
90. Standard ASTM E1040, *Standard Specification for Physical Characteristics of Nonconcentrator Terrestrial Photovoltaic Reference Cells*, Amer. Society for Testing Matls., West Conshocken PA, USA.
91. Standard IEC 60904-2, *Requirements for Reference Cells*, International Electrotechnical Commission, Geneva, Switzerland.
92. Osterwald C, Anevsky S, Barua A, Dubard J, *et al*., *Proc. 26th IEEE PVSC*, pp 1209–1212, (1997) also Results of the PEP'93 Intercomparison of Reference Cell Calibrations and Newer Technology Performance Measurements: Final Report, NREL Tech. Report NREL/TP-520-23477, March 1998.
93. Osterwald C, Anevsky S, Bücher K, Barua A, *et al*., *Prog in Photovoltaics* **7**, 287–297 (1999).
94. Shimokawa R, Nagamine F, Nakata M, Fujisawa K, *et al*., *Jpn. J. Appl. Phys*. **28**, L845–L848 (1989).
95. Metzdorf J, *Appl. Optics* **26**, 1701–1708 (1987).
96. Bruce S, *Optical Metrology and More*, NIST Tech. Report NISTIR 5429, U.S. Department of Commerce, National Institute of Standards and Technology, Gaithersburg, MD, USA (1994)
97. Allison J, Crab R, *Proc. 12th IEEE PVSC*, pp 554–559 (1976).
98. Metzdorf J, Wittchen T, Heidler K, Dehne K, *et al*., *Proc. 21st IEEE PVSC*, pp 952–959 (1990), also The Results of the PEP '87 Round-Robin Calibration of Reference Cells and Modules – Final Report, PTB Tech. Report PTB-Opt-31, Braunschweig, Germany, ISBN 3-89429-06706, November 1990.
99. Ossenbrink H, Van Steenwinkel R, Krebs K, The Results of the 1984/1985 Round-Robin Calibration of Reference Solar Cells for the Summit Working Group on Technology, Growth and Employment, Tech. Report EUR 10613 EN, Joint Research Center, ISPRA Establishment, ISPRA Italy, April 1986.
100. Osterwald C, Emery K, Myers D, Hart R, *Proc. 21st IEEE PVSC*, pp 1062–1067 (1990).
101. Gomez, T, Garcia L, Martinez G, *Proc. 28th IEEE PVSC*, pp 1332–1335, (2000).
102. Romero J, Fox N, Fröhlich C, *Metrologia* **32**, 523–524 (1995/1996).
103. Standard ASTM E1125, *Standard Test Method for Calibration of Primary Non-Concentrator Terrestrial Photovoltaic Reference Cells Using a Tabular Spectrum*, Amer. Society for Testing Matls., West Conshocken PA, USA.
104. Osterwald C, Emery K, *Journal of Atmospheric and Oceanic Technology* **17**, 1171–1188 (2000).
105. Shimokawa R, Nagamine F, Miyake Y, Fujisawa K, *et al*., *Jap. J. of Appl. Physics* **26**, 86–91 (1987).
106. King D, Hansen B, Jackson J, (1993), *Proc. 23rd IEEE PVSC*, pp 1095–1101 (1993).
107. Bücher K, Stiening R, Heidler K, Emery K, *et al*., *Proc. 23rd IEEE PVSC*, pp 1188–1193 (1993).
108. Anspaugh B, Downing R, Sidwell L, Solar Cell Calibration Facility Validation of Balloon Flight Data: a Comparison of Shuttle and Balloon Flight Results, JPL Tech. Report 85-78, Jet Propulsion Laboratory, Pasadena, CA USA (1985).
109. Anspaugh B, *Proc. 19th IEEE PVSC*, pp 542–547 (1987).
110. Brandhorst H, *Appl. Optics* **4**, 716–718 (1968).
111. Brandhorst H, *Proc. International Colloquium Organized by ECOSEC*, 565-574 in *Solar Cells*, (1971).
112. Bucher K, *Prog in Photovoltaics* **5**, 91–107 (1997).
113. Emery K, Osterwald C, Kazmerski L, Hart R, *Proc. 8th PVSEC*, pp 64–68 (1988).
114. Jenkins P, Brinker D, Scheiman D, *Proc. 26th IEEE PVSC*, pp 857–860 (1997).
115. General Requirements for the Competence of Testing and Calibration Laboratories, Geneva, Switzerland, ISO/IEC standard 17025 (2005).

116. ISO Guide to the Expression of Uncertainty in Measurement, International Organization for Standardization, Geneva, Switzerland, ISBN 92-67-10188-9 (1995).
117. Emery K, Osterwald C, *Solar Cells* **27**, 445–453, (1989).
118. Emery K, Osterwald C, Wells C, *Proc. 19th IEEE PVSC*, pp 153–159 (1987).
119. American Association for Laboratory Accreditation, www.A2LA.org, certification number 2236.01.
120. Emery K, Uncertainty Analysis of Certified Photovoltaic Measurements at NREL, NREL tech. Rep. NREL/TP-520-45299, August 2009.
121. Whitfield K, and Osterwald C, *Prog in Photovoltaics* **9**, 87–102 (2001).
122. Zaaiman W, Mullejans H, Dunlop E, Ossenbrink H, *Proc. 21st PVSEC*, pp 64–68 (2006).
123. Emery K, The Results of the First World Photovoltaic Scale Recalibration, NREL Tech. Report NREL/TP-520-27942, March 2000.
124. Matsuda S, Flood D, Gomez T, Yiqiang Y, *Proc. 2nd World Conf. PVSEC*, pp 3572–3577 (1988).
125. Treble F Krebs K, *Proc. 15th IEEE PVSC*, pp 205–210 (1981).
126. Rummel S, Anderberg A, Emery K, King D, *et al*., *Proc. 4th World Conf. PVSEC*, pp 2034–2037 (2006).
127. Betts T, Gottschalg R, Infield D, Kolodenny W, *et al*., *Proc. 21st PVSEC*, pp 2447–2450 (2006).
128. Kiefer K, Warta W, Hohl-Ebinger J, Herrmann W, *et al*., *Proc. 21st PVSEC*, pp 2493–2496 (2006).
129. Herman W, *et al*., *Proc. 22nd EU PVSEC*, p. 2506 (2007).
130. Herrmann W, Zamini S, Fabero F, Betts T, *et al*., *Proc.23rd EU PVSEC*, pp 2719–2723 (2008),
131. Friedman D, *Proc. 25th IEEE PVSC*, pp 89–92 (1996).
132. Virshup G, Chung B, Ristow M, Kuryla M, Brinker D, *Proc. 21st IEEE PVSC*, pp 336–338 (1990).
133. Bogus K, Larue J, Masson J, Robben A, *Proc. 6th EC PVSEC*, pp 348–354 (1985).
134. Glatfelter T, Burdick J, Czubatyj W, *Proc. 2nd PVSEC*, pp 106–109 (1986).
135. Glatfelter T, Burdick J, *Proc. 19th IEEE PVSC*, pp 1187–1193 (1987).
136. Virshup G, *Proc. 21st IEEE PVSC*, pp 1249–1255 (1990).
137. Emery K, Meusel M, Beckert R, Dimroth F, *et al*., *Proc. 28th IEEE PVSC*, pp 1126–1130 (2000).
138. Bickler D, *Solar Energy* **6**, 64–68 (1962).
139. Bennett M, Podlesny R, *Proc. 21st IEEE PVSC*, pp 1438–1442 (1990).
140. Shimokawa R, Nagamine F, Hamyashi Y, *Jpn. J. of Appl. Phys*. **25**, L165–L167 (1986).
141. Sopori B, Marshall C, Emery K, *Proc. 21st IEEE PVSC*, pp 1116–1121 (1990).
142. Sopori B, Marshall C, *Proc. 22nd IEEE PVSC*, (1991).
143. Krut D, Lovelady J, Cavicchi T, *Proc. 2nd World Conf. PVSEC*, pp 3671–3674 (1998).
144. Adelhelm R, Bücher K, *Sol. Ener. Mater. Sol. Cells* **50**, 185–195 (1998).
145. Muesel M. Adelehlm R, Dimroth F, Bett A, *et al*., *Prog in Photovoltaics* **10**, 243–255 (2002).
146. Emery K, Osterwald C, *Proc. 21st IEEE PVSC*, pp 1068–1073 (1990).
147. Schultz P, Meilus A, Hu S, Goradia C, *IEEE trans. On Inst. and Meas*. **IM26**, 295–299 (1977).
148. Briskman R, Livingstone P, *Sol. Energy Matr. and Sol. Cells* **46**, 187–199 (1997).
149. Cox C, Warner T, *Proc. 16th IEEE PVSC*, pp 277–1283 (1982).
150. Skolnik H, *Solar Energy* **14**, 43–54 (1972).
151. Kern R, Wagemann W, *Proc. 7th EC PVSEC*, pp 314–318 (1986).
152. Mantingh E, Zaaiman W, Ossenbrink H, *Proc. 24th IEEE PVSC*, pp 871–873, (1994).
153. Heidler K, Fischer H, Kunzelmann S, *Proc. 9th EC PVSEC*, pp 791–794 (1989).
154. Moriarty T, Emery K, *Proc. 4th NREL TPV Conf*., AIP proceedings **460**, 301–311 (1998).

155. Kiehl J, Emery K, Andreas A, *Proc. 19th PVSEC*, paper 5BV.2.11 (2004).
156. Gee J, Hansen R, *Solar Cells* **18**, 281–286 (1986).
157. Nasby R, Sanderson R, *Solar Cells* **6**, 39–47 (1982).
158. Dondero R, Zirkle T, Backus, C, *Proc. 18th IEEE PVSC*, pp 1754–1755 (1985).
159. Martin S, Backus C, *Proc. 6th EC PVSEC*, pp 290–294 (1985).
160. Chambers B, C. Backus C, *Proc. 3rd EC PVSEC*, pp 418–422 (1980).
161. Stryi-Hipp G, Schoenecker A, Schitterer K, Bucher K, *et al*., *Proc. 23rd IEEE PVSC*, pp 303–308 (1993).
162. Domínguez C, Antón I, Sala G, Martínez M. *Proc. 4th International Conference on solar concentrators*, 4B–3 (2007).
163. Standard ASTM E927, *Standard Specification for Solar Simulation for Terrestrial Photovoltaic Testing*, Amer. Society for Testing Matls., West Conshocken PA, USA.
164. Standard International Electrotechnical Commission IEC 60904-9, *Solar Simulator Performance Requirements*, Geneva, Switzerland.
165. Cuevas A, Lopez-Romero S, *Solar Cells* **11**, 163–173 (1984).
166. Dhariwal S, Mathur R, Gadre R, *J. Phys. D: Appl. Phys*. **14**, 1325–1329 (1981).
167. Schönecker A, Bücher K, *Proc. 22nd IEEE PVSC*, pp 203–208 (1991).
168. King D, Dudley J, Boyson W, *Proc. 25th IEEE PVSC*, pp 1295–1297, (1996).
169. Emery K, Myers D, Rummel S, *Proc. 20th IEEE PVSC*, pp 1087–1091 (1989).
170. Seaman C, Anspaugh B, Downing R, Esty R, *Proc. 14th IEEE PVSC*, pp 494–499 (1980).
171. Matson R, Emery K, Bird R, *Solar Cells* **11**, 105–145 (1984).
172. Wronski C, Abeles B, Cody G, Morel D, *et al*., *Proc. 14th IEEE PVSC*, pp 1057–1061 (1980).
173. Shafarman W N, Klenk R, McCandless B, *J. Appl. Phys*. **79**, 7324–7328 (1996).
174. Fardig D A, Phillips J E, *Proc. 22nd IEEE PVSC*, pp 1146–1150 (1991).
175. Hegedus S, *Prog in Photovoltaics* **5**, 151–168 (1997).
176. Emery K, Dunlavy D, Field H, Moriarty T, *Proc. 2nd World Conf. PVSEC*, pp 2298–2301 (1998).
177. Standard ASTM E1021, *Standard test methods for measuring spectral response of photovoltaic cells*, Amer. Society for Testing Matls., West Conshocken PA, USA.
178. Bücher K, Schönecker A, *Proc. 10th EC PVSEC*, pp 107–110 (1991).
179. Muesel M, Baur C, Letay G, Bett A, *et al*., *Prog in Photovoltaics* **11**, 499–415 (2003).
180. Tsuno Y, Hishikawa Y, Kurokawa K, *Proc. 23 EU PVSEC*, pp 2723–2727 (2008).
181. Jing-Gui C, Xiong-Jun G, Pei-Neo Y, Yu-Xue W, *Proc. 3rd PVSEC*, pp 743–748 (1987).
182. C Schill, K Bücher, A Zastrow, *Proc. 14th EC PVSEC*, pp 309–312 (1997).
183. Burdick J, Glatfelter T, *Solar Cells* **18**, 301–314 (1986).
184. Field H, *Proc. 26th IEEE PVSC*, pp 471–474 (1997).
185. Sommeling P, Riefe H, Kroon J, Van Roosmalen J, *et al*., *Proc. 14th EC PVSEC*, pp 1816–1819 (1997).
186. Stock K, Heine R, *Optik* **71**, 137–142 (1985).
187. International Electrotechnical Standard Commission IEC 61215, *Crystalline silicon terrestrial photovoltaic (PV) modules – Design qualification and type approval*, Geneva, Switzerland.
188. International Electrotechnical Commission Standard IEC 61646, *Thin-film terrestrial photovoltaic (PV) modules – Design qualification and type approval*, Geneva, Switzerland.
189. Osterwald C, ASTM Standards Development Status, *Proc. 18th IEEE PVSC*, pp 749–53 (1985).
190. Commission of the European Communities, Joint Research Center, ISPRA Establishment, specification 502, *Qualification Test Procedures for Photovoltaic Modules*, May 1984.
191. Osterwald C, Mcmahon T, *Prog in Photovoltaics* **17**, 11–33 (2009).
192. Ross R, *Proc. 17th IEEE PVSC*, pp 464–472 (1984).
193. Krebs K, Ossenbrink H, Rossi E, Frigo A, *et al*., *Proc. 5th EC PVSEC*, pp 597–603 (1983).

194. Dunlop E, Bishop J, Ossenbrink H, *Proc. 2nd World Conf. PVSEC*, pp 459–462 (1998).
195. Bishop J, Ossenbrink H, *Proc. 25th IEEE PVSC*, pp 1191–1196, (1996).
196. International Electrotechnical Commission Standard IEC 62108, *Concentrator Photovoltaic (CPV) Modules and Assemblies – Design qualification and type approval*, Geneva, Switzerland.
197. Osterwald C, Terrestrial Photovoltaic Module Accelerated Test-to-Failure Protocol, NREL tech. Rep. NREL/TP-520-42893, March 2008.

19

PV Systems

Charles M. Whitaker[1,*], Timothy U. Townsend[1], Anat Razon[1], Raymond M. Hudson[1] and Xavier Vallvé[2]

[1]*BEW Engineering, San Ramon, CA, USA*
[2]*Trama TecnoAmbiental, Barcelona, Spain*

19.1 INTRODUCTION: THERE *IS* GOLD AT THE END OF THE RAINBOW

19.1.1 Historical Context

Over its first 50 years, the PV industry's notion of a system has grown from milliwatts to megawatts, from low voltage dc to high voltage ac, from satellites to power plants. This chapter is not a timeline of significant technical and commercial milestones – Becquerel's 1839 identification of the photovoltaic effect, Bell Laboratories' 1954 announcement of a 6% efficient silicon cell, and Bell's 1962 PV-equipped Telstar satellite have been well-documented elsewhere [1, 2]. Rather, this section provides a contemporary look at system-level design and performance issues associated with this rapidly growing sector of the energy economy.

Of particular significance is the fact that PV texts have typically regarded PV as a "tomorrow" technology. PV has been termed an intriguing technical solution today, but only in niche markets, with broad world markets remaining an elusive 5–10 years distant. This text marks the beginning of a new commercial era, as many of the promises associated with massive scale PV are now being realized.

* Charles M. Whitaker (Chuck), a photovoltaic veteran with over 25 years' experience in design, project management, and most significantly, in international codes and standards development, passed away on May 3, 2009 while leading the writing effort on this chapter. His co-authors have worked to complete it in a manner consistent with Chuck's thinking. Chuck was with us long enough to see the PV industry blossom. He was gratified to see his tireless efforts laying much of the standards groundwork that a long-promising industry would eventually rely on had finally borne fruit in the form of both worldwide PV growth and in the success of his own engineering consulting firm.

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

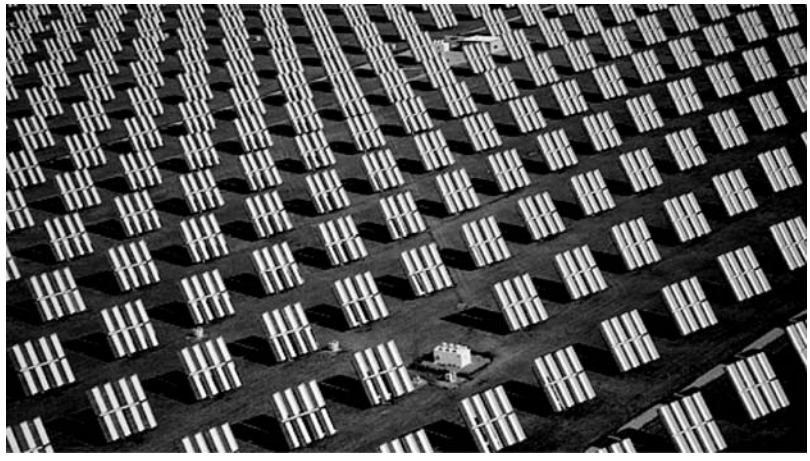


Figure 19.1 Carrizo Plain, 1984: Arco Solar 6.5 MW on 800 2-axis tracker mounts, each containing 128 modules of nominal 40 W_P each. Most trackers, like those shown here, had side reflectors to increase the irradiance on the modules. (Courtesy of George Lepp)

19.1.2 Contemporary Situation

The publication of this book's 2nd edition coincides with the recent 25th anniversary of the world's first MW-scale PV system at the Carrizo Plain in central California (see Figure 19.1). While it is now common to see a dozen or more multi-MW deployments per year, this rapid increase in both scale and number of systems has really only flourished over the past half-dozen years. The emergence of multi-MW PV systems has been concentrated in Europe, most notably in Germany, Spain, and Portugal, with Korea and the U.S. adding several since 2003 as well [3] (see Figure 19.2).

Figure 19.2 Olmedilla, 2008: NOBESOL 60 MW_P (Courtesy of Nobesol) See Plate 5 for the colour figure

The total installed capacity worldwide has grown from roughly 2 GW at the time of this book's first edition to a projected 25 GW through 2010 [4]. Demand over the past decade has increased at a remarkable 50% annual growth rate, with 4500 MW, nearly one-third of the cumulative world total, installed in 2008 alone [5]. While clear signs of softening in module price and demand forecasts started cropping up in 2009 in response to a global economic downturn, the PV industry is nevertheless on a strong growth curve [6].

The commercial importance of PV has been driven upwards by a combination of technology advancement, heightened environmental concern, and rising conventional electricity costs. However, the most significant factor has been strong government incentives to encourage investment in grid-connected applications, discussed in more detail in Chapter 2. The rapid worldwide growth exhibited this past decade has been marked by a distinct shift in the proportion of grid-connected end uses for PV. In 2000, roughly 75% of PV was used off-grid. By 2009, that ratio had essentially flipped, with 75% of PV now destined for on-grid applications. Off-grid applications, whether for space applications or numerous terrestrial uses, are still growing steadily in absolute numbers, but have been surpassed in relative market share by the large-scale subsidized systems. Off-grid systems are often economic without subsidies, as the cost of providing new utility grid connections in rural areas is often prohibitively high. This is especially true in remote locations and in developing countries where the utility grid infrastructure is relatively sparse.

19.2 SYSTEM TYPES

There are myriad end uses for PV, with a broad variety of system complexity. A range of applications is shown in the chart in Figure 19.3. On-grid versus off-grid applications share certain attributes but the PV systems satisfy distinctly different needs. For example, both on-grid and

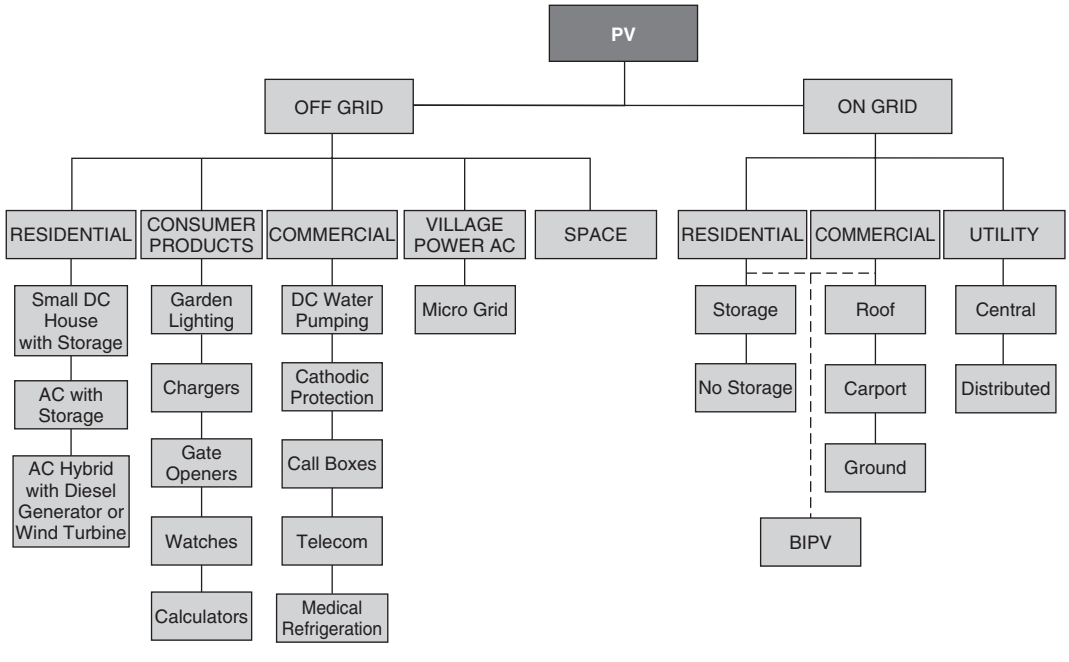


Figure 19.3 PV system taxonomy chart [7]

off-grid PV systems may use the same module technology, be mounted in the same manner, be deployed in the same climate, and deliver the same amount of AC energy to a hypothetical customer. The on-grid system will almost certainly be less expensive per kW to install and maintain and will operate more efficiently than its off-grid counterpart.

However, if no grid exists, it is usually prohibitively expensive to extend grid service into a remote area. In such a case, despite their relatively high cost and lower efficiency, off-grid PV systems are often the best solutions compared with the traditional options of fossil-fueled generators, regular battery swap-outs, or foregoing electric power altogether. Similarly, all portable applications of PV, regardless of cost, achieve something no grid can ever accomplish by providing power to non-stationary end uses.

The following section describes several common PV system configurations.

19.2.1 Small Off-grid DC System

DC systems include applications that directly use the DC energy produced by PV modules to supply power to DC loads. These include space-based systems, portable solar devices and small consumer products, very small residential systems, water pumping, and other small applications that are typically less than 1 kW in size and for which all of the loads only require DC electricity to function. A simplified block diagram of a DC system is given in Figure 19.4. Figure 19.5 is a photograph showing this type of system.

19.2.2 Off-grid AC System

Off-grid systems are those in which the PV energy is converted to AC power, but there is no utility grid available. The loads in this type of system run from AC power. The inverter in this type of system acts to regulate the AC voltage to all of the loads. Energy storage (i.e. usually batteries) is typically included in off-grid systems to allow for the required power balance between the intermittent PV energy source and the load requirements (i.e. lighting at night). A common example of this is a remote home with a system composed of solar modules on the roof or on a pole nearby, an off-grid inverter, and a battery bank for storage. A simplified block diagram of an off-grid system is given in Figure 19.6. Figure 19.7 shows this type of system.

19.2.3 On-grid Systems

Utility grid-connected systems are those in which the energy produced by PV modules is converted to AC electricity and either used on-site or injected into the utility grid. To accomplish this, the DC

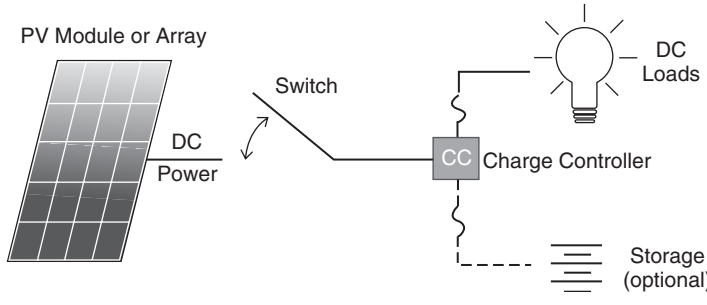


Figure 19.4 Simplified diagram of DC PV system (source TTA)

Figure 19.5 Small DC off-grid system Jurte family, Mongolia (Courtesy of Peter Adelmann, 2007)

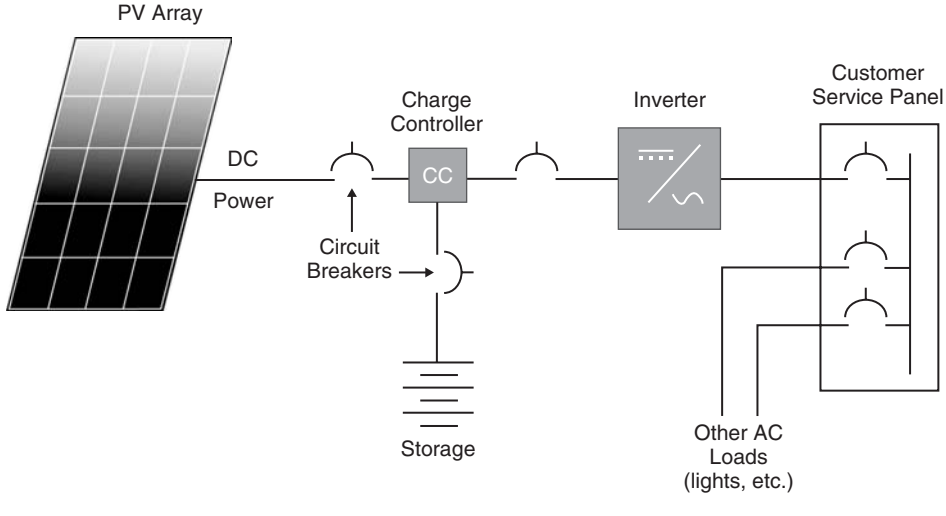


Figure 19.6 Off-grid system simplified block diagram

PV output must be converted to AC current by an inverter (as described further in this chapter and in great detail in Chapter 21). Like any power plant, this AC current must be synchronized with the utility grid that it is being interconnected with. This includes the AC voltage and frequency. As the excess PV generated energy is passed into the utility grid which is also being energized by other power generating sources, it is distributed to all of the loads connected to the grid and not to

Figure 19.7 Off-grid AC system, Ecuador (Courtesy of TTA)

specific equipment. A simplified block diagram of a grid connected system is given in Figure 19.8. Figures 19.9–19.13 show examples of residential, commercial, and utility-scale systems.

19.2.4 Hybrid PV Systems

The term hybrid can be broadly used to denote a PV system that is used in conjunction with one or more auxiliary sources of power. Traditionally this has meant a second source such as wind or hydroelectric turbines, however, many modern PV systems employ auxiliary dispatchable

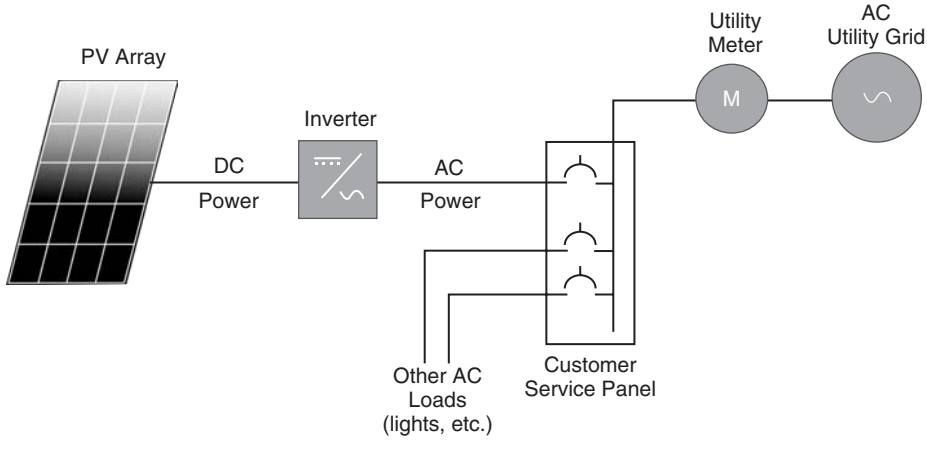


Figure 19.8 Grid-connected system simplified block diagram

Figure 19.9 Residential on-grid system (Courtesy of Watson family solar house, Lexington, MA, USA)

Figure 19.10 Ota City, Japan distributed 550 on-grid residential systems, 2.2 MW (Courtesy of Hironobu Igarashi) See Plate 6 for the colour figure

(on-demand) sources such as on-site fossil-fueled generators or the utility grid. The term multi-mode is often reserved to describe the special case of a hybrid system which operates either in parallel with an external AC source (on-grid) or as a stand-alone AC source (off-grid) when grid power is unavailable. A block diagram of a typical hybrid system is shown in Figure 19.14.

Energy can be made available on a continuous basis during the night or during unfavourable weather conditions by the battery. Hybrid systems are usually sized so that a high fraction of the energy is of solar origin and under favourable weather conditions the consumer's total energy demand is met by the PV generator. Surplus energy is stored in the battery. A back-up generator can simultaneously provide electricity to both the loads and to charge the battery, thus increasing the reliability as well as eliminating the need for an oversized PV array. Hybrid systems capture the best attributes of each energy resource and provide "grid-equivalent" electricity. The utility meter must allow for two way energy flow and net energy metering.

Figure 19.11 Denver, Colorado 300 kW_P commercial grid-tied rooftop ballasted system (Courtesy of Brad Eccles, BEW Engineering)

Figure 19.12 Currently one of the world's largest utility scale system (fixed mounting), 54 MW_P incorporating 225 000 Si modules and covering 1.3 km^2, in Strasskirchen, Germany (Courtesy of Raymond Hudson, BEW Engineering)

All electricity generating components are connected via separate charge controllers and rectifiers to a DC bus bar, where the battery is connected. The battery supplies power to the loads in response to the demand. In order to protect the battery, the PV charge controller disconnects all loads before the battery becomes deeply discharged. At the same time, when the battery status is low or if the loads need more power than the inverter can deliver, the back-up generator can be automatically started; it then supplies AC power directly and charges the battery via a rectifier or a bidirectional inverter.

Figure 19.13 Utility central station PV, Springerville, AZ 0.18 km^2 and 4.5 MW [7] (Courtesy of Tucson Electric Power)

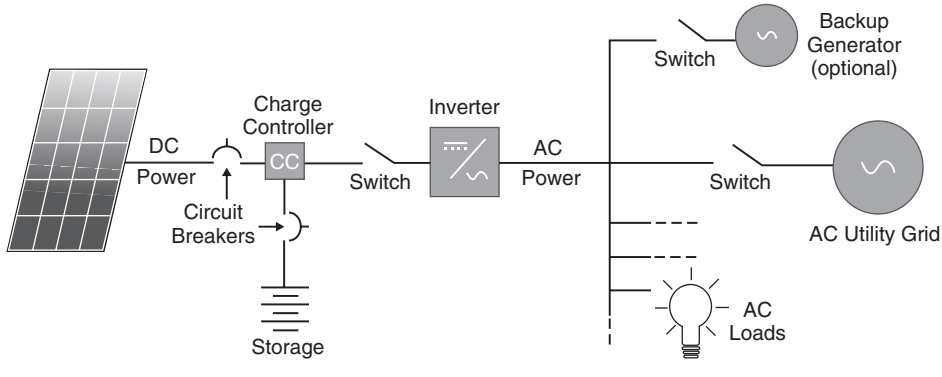


Figure 19.14 Simplified block diagram of a hybrid system

19.2.5 Micro-grids

Off-grid PV systems are well established for supplying single rural homes and separate small loads. When several remote houses are clustered forming a village, an emerging option is to design a larger plant to supply each household with a standard AC single-phase service (see Figure 19.15). With a central PV plant some economies of scale are realized. Installation and operation costs are lower than the collective cost for multiple single-house PV systems. Typically, micro-grids have a capacity up to 100 kW. One hurdle that is unique to micro-grids is developing the metering technology and a collective sense of responsibility in managing, and paying for, a jointly used, limited energy resource.

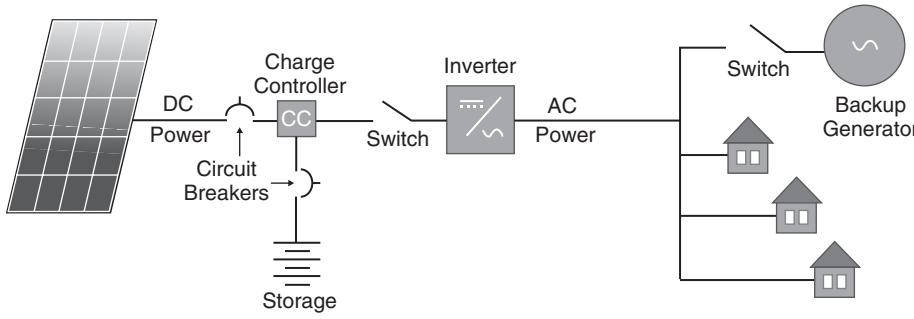


Figure 19.15 Simplified block diagram of micro-grid

19.2.6 Smart Grid

Another contemporary topic is the "smart grid." The definition of exactly what makes up a smart grid is not formalized at this time but generally consists of:

- A combination of central and distributed energy sources such as PV
- Switching equipment for connecting/disconnecting loads and generators
- Enhanced communications including generation and load monitoring
- An overall system monitoring system that optimally dispatches generations and loads
- Often incorporate dispatchable energy storage

The overall goal of the smart grid approach is to improve the reliability and efficiency of electricity delivery while effectively incorporating distributed and variable generation sources. PV systems are thought to integrate well with the smart grid concept because of their environmental attributes and their fundamentally distributed nature.

19.3 EXEMPLARY PV SYSTEMS

The best PV systems will always demonstrate good:

1. Performance
2. Safety
3. Reliability

The *performance* of a system can be measured both by the amount of energy produced and/or by the system meeting its financial expectations. For off-grid systems, the measure of success is simpler and almost indistinguishable from reliability: does it supply the power demanded of it at all times, regardless of weather or time of year? For grid-tied systems, delivering the maximum possible energy to the local utility-tied load is often the goal. Proper design, installation, and maintenance are all key to achieving high system performance.

Safety is also an essential characteristic of successful PV systems. It is critical that the system is compliant with the local electrical code, employs components certified to internationally recognized safety standards, is properly grounded to prevent shock hazards, and does not create a fire hazard. As the PV industry expands, safety considerations must evolve, too, with at least three characteristic hurdles to overcome:

1. the "newness" or inexperience factor
2. the "hurry-up" or shortcut factor
3. the competitive "field machismo" factor

For a workforce accustomed to relatively safe low voltage off-grid dc systems, there is often a large gap in understanding the deadly hazards of working with utility power. Safety is further compromised when workers are uncomfortably subjected to extremes of heat, cold, wind, rain, and sometimes snow while trying to maintain their concentration during long hours, often spent making simple and repetitious electrical and mechanical connections.

Finally, *reliability* forms the third leg of support common to all good PV systems. Contemporary PV systems are typically intended to last for a nominal minimum of 25+ years and are expected to operate without fail at nearly their original levels throughout that period. As more grid-tied PV systems are deployed, utilities will rely more heavily on these systems performing reliably as they become significant portions of the electrical grid. In the case of off-grid systems, reliability may be even more crucial, since the grid does not act as a backup source of electricity. Lack of reliability adversely impacts performance by reducing the amount of electricity produced and increasing the cost of maintaining the system. Unlike most electronic components, PV is exposed to harsh outdoor conditions where extremes of temperature and humidity increase failure rates and hasten component degradation.

19.4 RATINGS

There are several ways to "rate" a PV system. It is very common for ratings to be stated in terms of the nominal (also known as "peak") DC or AC power. These are often denoted as W_P, kW_P, or MW_P as shorthand notations to represent peak DC power, or as W_{AC}, kW_{AC}, or MW_{AC} as shorthand notations for nameplate AC power. None of these terms require any field measurements. They are simply based on manufacturers' nameplate values and over-predict the actual performance of an array by 20–30% because they do not account for real system losses.

IEC Standard 61 724 [8] defines several terms that are helpful for establishing ratings. These include the reference yield Y_r, perhaps better understood as the number of peak sun hours per year on a given array plane; the array yield Y_a, or DC kWh per kW per year, the final yield (Y_f, or annual kWh produced per installed kW_P), and the performance ratio (PR, or ratio of final yield divided by reference yield).

The kW_P rating is simply the product of the number of PV modules in the system and the nameplate (W_P) ratings of the modules, without accounting for real and unavoidable losses. It should be noted that the nameplate rating of PV modules is usually based on Standard Test Conditions (STC = 1,000 W/m^2, 25 °C cell temperature, and air mass (AM) 1.5 spectrum) [9], a combination that is convenient to establish indoors but exceptionally rare in the field. Module rating methods are also discussed extensively in Chapter 18.

Just as a system's DC rating is usually defined as the sum of the nameplate ratings of the modules, the AC rating is often defined as the sum of the nameplate ratings of its inverters. With such a definition, there is an implicit assumption that there are enough PV modules (i.e., DC capacity) in the system to support that peak AC power level. Both DC and AC peak power ratings are imprecise terms, best used for gathering a general understanding of the size of the system. AC nameplate ratings are lower than DC nameplate ratings because there are several unavoidable losses that occur during the conversion from DC to AC power. Many PV systems will have DC nameplate ratings 10–30% larger than their AC nameplate ratings. This is an appropriate level of over-sizing, given the fact that parasitic losses from wire resistance, temperature, dust, inverter efficiency, and

several other effects are unavoidable. This significant difference between DC and AC nameplate ratings has caused confusion and misinterpretation when costs and yields are discussed without clearly stating whether the rating corresponds to a DC or AC basis.

There have been efforts to define PV system AC power ratings in other ways. Some methods are *analytical*, some rely instead on *measured* data (see Footnote[1] and Reference [10]). Analytical AC ratings tend to start with nameplate DC capacity, then step through a series of loss factors, as noted above to arrive at a diminished, but more realistic, AC power rating. Often, the chosen rating conditions are STC, but not always. One alternate is Standard Reporting Conditions (SRC) [11], which substitute a more realistic module temperature of 45 °C, but are otherwise like STC. Depending on PV technology type ratings made under SRC are 5–10% lower than ratings made under STC.

An *analytical* AC rating example is that used by the California Energy Commission (CEC) under the California Solar Initiative (CSI). The CEC assigns an AC rating to each system seeking a financial rebate under the CSI [12]. The CEC-AC rating is the product of the system nameplate DC capacity, a temperature-based DC efficiency adjustment, and a representative inverter DC-to-AC efficiency adjustment. The CEC-AC rating is typically 83%–87% of the nameplate DC rating. Note the CEC-AC rating is somewhat generous, since it ignores ohmic wire loss, mismatch, dust, shade and other small losses that usually reduce power by an additional 5–10%.

$$kW_{CEC-AC} = \#\,\text{modules} \cdot (kW_p/\text{mod.}) \cdot \text{temperature.adj.} \cdot \text{inv.adj} \tag{19.1}$$

The CEC's calculation is based on PVUSA Test Conditions (PTC)[2], [13], a more common and realistic combination of sunlight and temperature than STC. Module and inverter manufacturers must supply several certifications and performance specifications in order to obtain a CEC listing. The CEC calculates the module efficiency adjustment based on the product's STC efficiency, its Nominal Operating Cell Temperature (NOCT) [14], and its temperature coefficient of maximum power. The inverter adjustment is based on test data from an independent laboratory that measures and weights inverter efficiency over a range of power levels.

An example of a *measured*, or field-determined, rating is the PVUSA test method. The PVUSA rating, not surprisingly, is based on PTC. It relies on a regression analysis of actual system data fitted to a linear power equation. The power equation is:

$$kW_{AC} = C_1 \cdot Irr + C_2 \cdot Irr^2 + C_3 \cdot Irr \cdot T_{AIR} + C_4 \cdot Irr \cdot WS \tag{19.2}$$

where Irr = Irradiance in the plane of the array, WS = wind speed, m/s, T_{AIR} = air temperature (not module temperature), and C_1 through C_4 are regression-fitted coefficients.

The first term is intended to describe the obvious and direct linear relationship between irradiance and power. Though it may not be apparent at a glance, the other three terms are associated with the thermal behavior of an in-place system. The fitted coefficients are unique to each system and reflect all types of losses, whether caused by dust, wire loss, or any other effect. Once fitted, the custom equation can be used to predict a given system's power at PTC or any other combination of conditions. The equation's accuracy is limited by both the prevailing weather conditions used to fit the model and by the degree of extrapolation needed to predict power under the desired rating conditions. At least several dozen data points are usually needed to obtain a satisfactorily

[1] No ASTM standard currently exists for rating flat-plate PV systems, though a proposed standard was submitted in the fall of 2009 by A. Kimber in the paper *Improved Test Method to Verify the Power Rating of a Photovoltaic (PV) Project*.

[2] PVUSA was a CEC-sponsored PV research and demonstration project headquartered in Davis, CA that ran from 1988–2000. PVUSA Test Conditions are defined as 1000 W/m^2 plane of array irradiance (850 W/m^2 direct normal irradiance for concentrators), 20 °C air temperature, and 1 m/s wind speed at 10 meters above grade.

accurate fitted equation. The accuracy of the equation is dependent on how similar the desired rating conditions and the prevailing conditions during the data gathering period are. Ideally, the prevailing conditions are richly distributed around the desired rating condition, such that the equation is only interpolating within the range of test data. High-quality PTC ratings can be obtained in the spring in many sunny Northern Hemisphere climates. At that time, there tends to be a good mix of rainy days to clean the array, sunny days to provide irradiance points at and above 1 kW/m^2, daytime temperatures around 20 °C, and mild winds. The PVUSA equation was found to work well 8–9 months per year, its chief limitation being a distinct tendency to over-predict the true PTC rating by 5–10% if mid-winter conditions are used to fit the model.

For example, a system with a nameplate DC rating of 20 kW$_P$ might equivalently be rated (with no field measurements) at about 17 kW$_{CEC\text{-}AC}$. After collecting several dozen data points in April, its PTC rating via the PVUSA power Equation (19.2) might result in a rating of about 16 kW$_{PTC}$ after plugging in the PTC conditions into the following fitted Equation (19.3), with results shown as Equation (19.4):

$$kW_{AC} = 19 \cdot Irr - 1.1 \cdot Irr^2 - 0.1 \cdot Irr \cdot T_{AIR} + 0.1 \cdot Irr \cdot WS \tag{19.3}$$

$$16\ kW_{AC} = 19 \cdot (1\ kW/m^2) - 1.1 \cdot (1\ kW/m^2)^2 - 0.1 \cdot (1\ kW/m^2) \cdot (20\,^{o}\mathrm{C}) + 0.1 \cdot (1\ kW/m^2) \cdot (1\ \mathrm{m/s}) \tag{19.4}$$

Energy yield ratings of PV systems are valuable for understanding the amount of energy that a given system will be able to generate in a given period of time (typically annually). The energy rating is developed through calculations or a simulation of the performance of all of the components configured as they are in the actual system with the input being determined from the solar resource and weather conditions at the installation site. This results in a typical first year energy yield in kWh. This can be adjusted annually to include system degradation. This type of rating is especially valuable for grid-tied systems that have payments and incentives based on the amount of energy produced as is becoming increasingly common.

Measured final yields for several systems ranging in size from 1 kW to 19 kW in Davis[3], California are shown in Figure 19.16. Many designers might project an expected annual yield of 1,500 kWh/kW$_P$ for this climate, which receives an average of 5 kWh/m^2/day global horizontal insolation (GHI) at an average air temperature of 16 °C (equivalent to 5 'sun hours', i.e. 5 hours per day of 1 kW/m^2 intensity).

The initial projections of annual yield change quickly and can be wildly in error relative to the true long-term value. A system such as System R, which began operating in mid-summer, quickly declines as the poorer fall and winter months begin to have an influence. The opposite is true for systems such as System T, which initially seem to perform poorly because they were started in mid-winter. The lifetime annual yield trends for each system clearly dampen out after about three years, at which point the starting month becomes unimportant. It is visually helpful, but not really necessary to keep continuous monthly records, as long as intermittent monthly energy readings are taken in the later years to demonstrate the trend. After 5–7 years, a slight amount of degradation is apparent, and it is clear that lifetime yields range a lot even in the same climate, depending on orientation, degree of shading, system reliability, and the degree to which actual module power matched its nameplate rating. None of the systems has produced 1,500 kWh/kW$_P$, and even the best-designed, oriented, and maintained system (system B with only 1% shading, S orientation and optimally tilted at 30°) has fallen over 5% short of the 1,500 kWh/kW$_P$ benchmark. The GHI over these years has matched the long-term average so lower-than-average irradiance cannot be blamed. Instead, the combined impacts of overstated manufacturers' module ratings, significant shading,

[3] Davis's climate matches that of Sacramento, 30 km east.

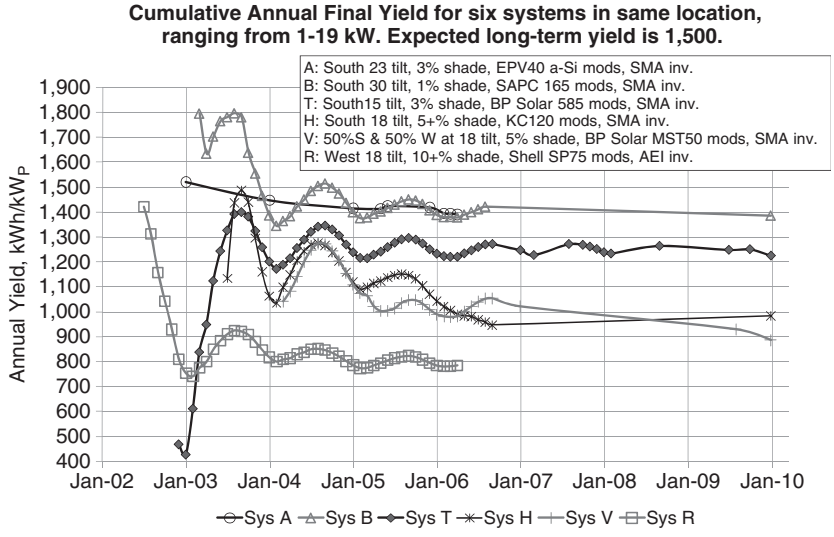


Figure 19.16 Lifetime yields for six sample systems in an identical climate. This figure shows the classic time-dependent damping that approaches the long-term average yield in kWh/kWp/yr. Note seasonal swings in insolation overwhelm lesser effects such as temperature; also, consecutive monthly data points were not available throughout the study. Legend fields are as follows: orientation, tilt angle, shade loss %, module type, and inverter brand.

orientation, maintenance practices, and inverter reliability have caused these systems to fall short of expectations. In some cases, poor west-facing orientation is chiefly to blame, in others, heavy shading and inverter failures have had large effects on lifetime yield.

19.5 KEY SYSTEM COMPONENTS

PV systems can be exceptionally simple, such as direct-coupled arrangements with few components other than modules, wire, and connections to an intermittently operated load. Most, however, consist of a related set of components designed to extract the maximum energy from the PV modules and reliably and safely deliver that energy to the intended load. While the specific components vary from application to application, some components are common to most applications, especially among grid-tied systems. The following sections describe components typically employed in a wide mix of PV applications.

19.5.1 Modules

While deep technical details about PV cells and modules are covered in other chapters of this text, this section is intended to provide an overview of the characteristics that are important for practical system applications.

A small number of PV module types now dominate the market. Well over 99% are flat-plate types, with concentrator PV making up the rest. Wafer-based silicon cells are the major constituent of the PV modules, with encapsulant, cell tabbing, front and rear cover sheets, framing, and connecting wire being the other main components. Wafer-based modules have the highest efficiencies

commercially available and are typically fabricated from single-crystal (c-Si) or multicrystalline silicon (multi-Si) ingots. Together, these two types account for approximately 85% of all PV now being used in systems throughout the world, in roughly equal quantities, with various thin films of amorphous silicon or CdS/CdTe (cadmium sulfide/cadmium telluride) making up most of the other 15% of the commercial market. [15] Ribbon silicon is a die-drawn variant of multi-Si that is manufactured by a very small number of companies, without the requirement to saw wafers from ingots. (See Table 1.5 in Chapter 1 for a detailed breakdown of percentages).

Significant increases in production of c-Si and multi-Si based modules have occurred over the past few years. This includes a major influx from China and other low-cost production areas of the world. The result has been a considerable reduction in cost of these modules which greatly reduces the overall system cost and benefits the overall performance goals.

Most other commercially available PV modules are generally described as thin film PV. A number of materials in addition to silicon are used. Thin film, or amorphous silicon, was introduced in the mid-1980s. It is generally less than half the efficiency of the more prominent wafer type Si modules. Other materials used for making PV modules include cadmium telluride (CdTe), copper indium gallium selenide (CIGS), copper indium selenide (CIS), and gallium arsenide (GaAs). The beginning chapters in this book detail the various properties and commercial aspects of each of the main types of materials used in PV modules. It should be noted that at the time of publication the amount of thin film modules being installed in systems is growing more rapidly than the industry as a whole, with an emphasis on large ground-mount systems. This is due to the relatively low efficiency of the modules which results in requiring a relatively large amount of area to support a given amount of power capacity. The complementary fact is also true that thin film modules are rarely used in systems where the amount of space is constrained such as roof mount systems. The cost of thin film PV modules has dropped significantly with a corresponding benefit in financial appeal.

The most important technical characteristics for identifying and distinguishing PV modules for system applications are their electrical and physical specifications. The key electrical characteristics include: nameplate power rating, V_{OC}, I_{SC}, V_{MP}, I_{MP}, fill factor, temperature coefficients, and efficiency. The key physical characteristics are: dimensions, weight, cover material, packaging, mounting requirements, and grounding method. The impacts that these characteristics have on the system performance and design are discussed in the following sections of this chapter. Visual appearance, shape, and physical flexibility are important in building integrated designs, as discussed in Chapter 23.

19.5.2 Inverters

After modules and sometimes, structures, inverters are the next most significant equipment cost element of a PV system. For example, the inverter typically represents 5–10% of the total system cost for a commercial scale on-grid system and 15–25% of the total system cost for a residential off-grid system. The topic of power conversion is described in Chapter 21 of this book; this section is intended to give an overview of the key requirements for the power electronics from a system perspective.

19.5.3 On-grid Inverters

The basic function of the inverter is to convert the DC power produced by the PV modules to AC power for system electrical loads. This is accomplished through the use of transistor-based power electronics circuitry. The power transistors are switched on and off at a high frequency in a manner that draws power from the PV modules at their maximum power point and passes the power to

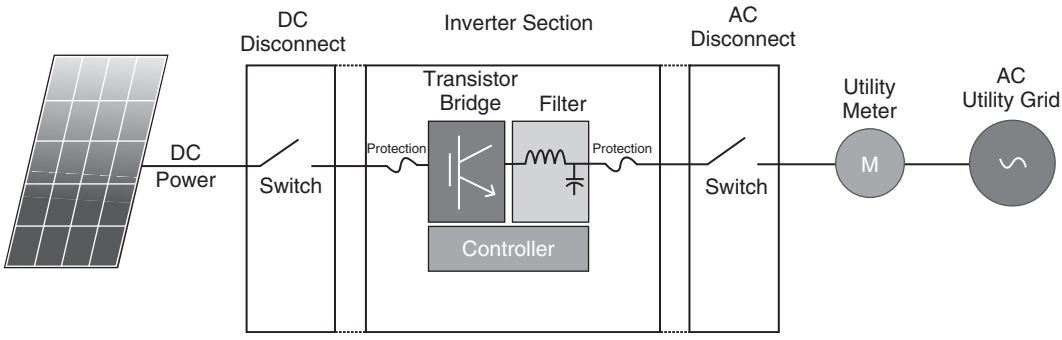


Figure 19.17 Inverter block diagram

the AC grid (for grid-tied systems) or to the local loads (for off-grid systems). The block diagram (Figure 19.17) gives an overview of the key components of a grid-tied inverter.

The PV array is connected to the inverter through a DC disconnect switch. Similarly, an AC disconnect switch is between the AC outputs of the inverter and the AC grid. These switches allow for maintenance and repairs to take place safely with the inverter components de-energized. The disconnect switches are often integrated into the inverter enclosure. There are also over-current protection devices on both the AC and DC sides. These are shown as fuses but this function may also be accomplished with circuit breakers. Additional protection is typically included on the AC and DC sides for transient surge suppression such as lightning induced voltages. The transistor bridge is made up of power semiconductor switches, typically mosfets, bipolar junction transistors (BJTs), or insulated gate bipolar transistors (IGBTs), which are operated in either a conducting or blocking state. This state is switched on and off at a high frequency (2–20 kHz) such that the resulting AC waveform is in the shape of a sinusoid. A filter is used to remove switching components of the current that are not at the fundamental AC frequency (50 or 60 Hz). A key component of the inverter is the controller. It implements the algorithms that control the transistors, provides maximum power point tracking, implements grid interface requirements, provides metering of the system input and output, and communicates to the user and overall system monitoring/control systems. For off-grid systems, the controller implements the battery the battery charging function.

For grid-tied systems, the inverter must regulate the AC current that is flowing into the grid and meet the utility requirements for power quality and interaction functionality. These standards are changing quickly and are a topic of significant discussion in many countries. In the USA, the grid interactivity requirements are generally defined by the IEEE-1547 "Standard for Interconnecting Distributed Energy Resources with Electric Power Systems" and UL-1741 "Standard for Safety Inverters, Converters Controllers and Interconnection System Equipment for Use with Distributed Energy Resources" standards. These standards include a requirement for "anti-islanding." This requirement ensures that the inverter disconnects from the utility grid if there is a loss of regulation of the AC voltage by the utility. The IEEE-1547 and UL-1741 standards also presently do not allow the inverter to operate at a variable power factor angle or attempt to regulate the AC voltage.

In Europe, the situation is somewhat different in that there are different standards from country to country and sometimes from utility to utility within one country. All inverters deployed in Europe must bear a "CE" mark. Additionally, a new IEC standard that is attempting to provide some common requirements is IEC 62 109 "Safety of power converters for use in photovoltaic power systems." Unlike in the USA, where standards are geared towards quickly disconnecting from a faulty grid, the prevailing European practice is to have PV inverters remain online during

minor faults in order to support the utility grid. This includes riding through voltage dips and operating at variable reactive power levels to maintain normal grid voltage.

Another important system consideration concerning inverters is the environmental rating of their enclosures. Some inverters are rated for only indoor environments such as NEMA 1, 2, or 12 (from NEMA 250–2003) or IEC IP20 (from IEC 60529). For outdoor environments enclosure ratings such as NEMA 3R/4 or IP 65 are used. It is important that the inverter is installed in the system in a manner consistent with its environmental protection rating. This is also true for the ambient temperature range specified for the inverter. These factors impact where and how the inverter is physically mounted. Best practices include placing the inverter in a well-ventilated, shaded area such as the north side of a building. Inverters located indoors are protected from the elements but do add a small and generally unwanted source of waste heat.

An important characteristic of a PV inverter is its efficiency. Inverters are typically less efficient at low power levels, reach a peak in the middle of their power range, and then drop off slightly at high power levels. Inverter efficiency also varies with the DC PV voltage. A number of standards have been developed to provide a weighted-average, or blended, efficiency for various climates. These are selected to reflect the frequency of typical solar resource conditions. Two examples of efficiency weighting are those established by the California Energy Commission (CEC) and those applied by the IEC, more commonly termed "European" weightings. As these methods weight the inverter efficiency at various power levels in different ways, the same piece of equipment may have multiple weighted-average efficiency ratings. The Figure 19.18 below shows a typical inverter efficiency curve. Table 19.1 lists the weighting factors for both of the blending methods.

19.5.4 Off-grid Inverters

Off-grid inverters are required to regulate the AC voltage that is being supplied to the local loads. This requires a fast dynamic response to keep the system operating in a stable manner. The electricity storage element (i.e. batteries) of the off-grid system provides a buffered source of power to withstand transients in the system. Often the control of the power flowing to and from the storage device is provided by a device known as a charge controller. This function may be integrated into the inverter or provided by a separate device. The sizing of the inverter for off-grid systems is

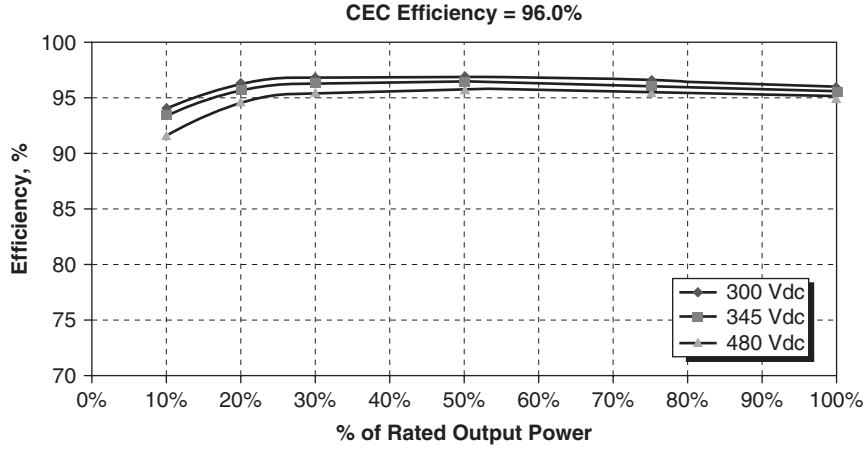


Figure 19.18 Example of inverter efficiency curve for a 250 kW grid-tied inverter [16]

Table 19.1 CEC and European efficiency weighting table

Inverter power level (%)	Weighting factor California (CEC)	Weighting factor Europe
5	0.00	0.03
10	0.04	0.06
20	0.05	0.13
30	0.12	0.10
50	0.21	0.48
75	0.53	0.00
100	0.05	0.20

critical for reliable system performance. The off-grid inverter must be capable of providing for any surges that the load may require, such as motor inrush (i.e. inductive kick) currents and at the same time, have a high efficiency at marginal power which is the most common steady state operating condition.

A special case of DC power conversion does not require an inverter, but does require a component that is functionally similar to an inverter. DC–DC converters, as used in PV applications, only change one DC voltage to another DC voltage. DC–DC converters are used to maintain an array at its maximum power point while delivering DC power to a load such as a battery or a motor at its proper voltage. Recently, direct DC converters to universal serial bus (USB) have come into existence for charging phones, radios, and digital music players (e.g. mp3).

19.5.5 Electrical Balance of System (BOS) and Switchgear

This includes all the wiring, fuses, combiners, fittings, grounding connections, switchgear, and metering. This material must be robust enough to withstand many years outdoors in wet conditions, extreme temperatures, corrosion and exposure to UV radiation. All BOS equipment needs to be rated for the voltage and current class it will be subjected to as well as the indoor or outdoor environment it will be operating in. PV wire comes in several grades of temperature and voltage class, as well as special designations for fire retardant properties, UV resistance, and wet conditions.

19.5.6 Storage

Energy storage is considered essential in almost all off-grid applications. Usually this means storing solar-generated electricity in a battery, although pumped water storage is the preferred energy storage medium in some direct-coupled applications. While a variety of battery technologies are finding increasing use (Pb-Sb, Li-ion, NiMH, NiCd, Zn-air, and others), lead–acid types are the long-standing commercial standard form of storage. Lead–acid batteries are available as flooded or sealed, with absorbed glass mat (AGM) or gel cell types popular forms of sealed lead–acid batteries. Most solar applications use a deep-discharge version of flooded lead–acid technology, which feature thicker lead plates for increased durability and cycle life. It is a common design practice to size batteries such that they are not routinely discharged more than about 50%. It is also a common operational practice to periodically charge flooded lead–acid batteries to their full capacity, as long as the floating set points are properly adjusted to avoid excessive gassing, loss

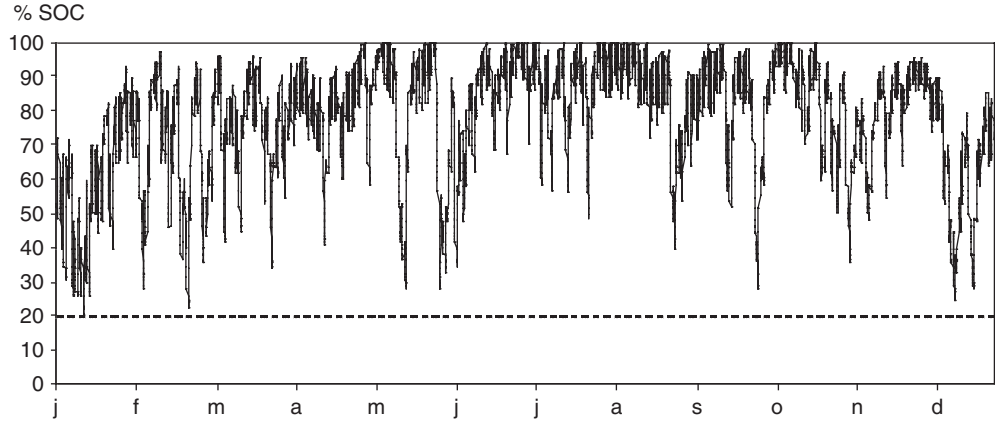


Figure 19.19 Battery state of charge (SOC) pattern for one year in a PV-supplied remote household (Source TTA)

of electrolyte, and reduced efficiency. Properly designed, sized, and maintained batteries can attain service lives of 10 years or more.

Energy storage must make up for the periods with load and without sufficient solar irradiation, such as during the night or during cloudy weather. Both the daily and seasonal patterns of charge and discharge cycles drive storage sizing and design. In a typical daily cycle the battery is charged during the day and is discharged during the night. There are also seasonal cycles that are a direct consequence of irradiance levels and seasonal demand variations (see Figure 19.19). These cycles are the main parameters that influence the lifetime of a properly maintained battery, although other parameters such as battery temperature, current, and voltage are significant as well. Figure 19.19 clearly shows the marginal ability to maintain adequate charge during the winter and the comparative ease of maintaining a high SOC in the summer. The seasonal extremes range from a low of 20% in the winter to 100% in the summer with an annual average of about 70%. In the winter the charge controller's low-voltage disconnect is occasionally shedding loads to prevent battery damage. Typically on a monthly basis, the battery should become fully charged at least 30% of the days and preferably more to assure long life and adequate equalization. Depending on how carefully the user manages the PV plant, the battery may frequently suffer long periods of intermediate state of charge without reaching full charge, even under average irradiation conditions, which presents an additional stress which shortens battery lifetime.

19.5.7 Charge Controllers

Typically, designing off-grid systems for high reliability exclusively with photovoltaic power leads to an oversized PV system because it must be sized to provide enough power for the maximum load expected during the cloudiest periods. Consequently, it produces much more energy than needed many other times of the year. Without a grid to receive the excess power, the PV system has to be occasionally idled to avoid overcharging the batteries. The purpose of a charge controller is essentially to direct traffic: it steers current into the batteries when needed, it permits current to flow to the inverter as loads dictate, maintains the array at its peak power point (not all charge controllers have this capability), and most importantly, provides disconnect capability to prevent over-discharge of the battery by either disconnecting or disabling certain components.

19.5.8 Structures

Several types of structures are in common use for mounting modules in PV systems. With some overlap, structural options differ depending on the application, such as ground-mounted, carport, rooftop, or façade mounting.

For residential roof top systems there are several solutions ranging from building-integrated PV (BIPV, see Chapter 23), whether as PV roofing tiles or as laminates that adhere to an underlying roof membrane such as a standing seam metal type, to rack-mounted systems that are offset from the roof, and tilted racks. BIPV may save money by being incorporated into the original design, removing the need for some roofing material, although it can be harder to service and operates warmer than rack-mounted PV. BIPV can take the form of roofing, shingles, windows, awning, or skylights. Solar tiles are interlocking tiles placed on top of an existing roof, and allow for air circulation between the PV and roof. Flexible modules can be also attached directly to the roof by either adhesion (so-called peel-and-stick) or fasteners. However, a separate racking system can be a good solution because convective cooling is important in PV systems. Module efficiency is higher at lower temperature, with the secondary benefit of reducing unwanted heating of the roof membrane and the building interior. Additionally, the roof may last longer because it receives less sun (UV) exposure and stays cooler. Fasteners require roof penetrations that require care to waterproof properly; they can be sealed with caulking, flashing, or rubber boots.

Commercial roofs – modern warehouse, office park, and retail buildings – are usually nearly flat, and not normally designed to withstand heavy dead loads. Therefore, if ballasted structures are used, the tilt angle must be kept to very low angles in order to reduce wind uplift and ballasting requirements. Ballasting is often done using gravel or paving blocks. Attached systems are more complex, but enable the use of steeper tilt angles. Carports are often a good structural option for commercial applications. The elevated structures cost more than a standard ground mounting and are a bit more difficult to service, but provide a welcome secondary benefit by providing shade for cars and do not require additional land.

For ground-mounted systems there are several ways to create a foundation, ranging from simple ballast blocks that lie on the ground, to poured concrete piers, to driven piles or helical screws. These systems can either be fixed or tracking. Tracking structures can move about one or two axes. Tracking systems can boost annual output by up to 50% relative to a fixed horizontal surface. This increase comes at a cost; the structures are more expensive to buy, install, maintain, and operate. Tracking systems require more area to avoid row to row shadowing. Most tracking systems operate electrically, but some types are passively driven using refrigerants. (For more information on tracking see Chapter 22).

As with flat-plate systems, concentrator systems can be deployed using several types of structures. While there are similarities in the sense that either fixed, single-axis, or dual-axis tracking structures can be used, CPV remain distinctly different because of the unique geometries and materials used. For example, lenses and mirrors are essential components unique to CPV. Securing and positioning these components accurately and reliably introduces a much different set of design challenges than the flat-plate designer faces. CPV systems separate the functions of collection and conversion of solar radiation to electricity, as discussed in Chapter 10. These systems require accurate tracking and an ability to cool down the PV material. [17]

In the off-grid micro-grid case, many times situating the central PV array on a building erected especially for the micro-grid serves the dual purpose of providing a structure for the modules, inverter, and battery as well as creating a public building such as a community center, school, or clinic.

19.5.9 Standards

PV system design is governed by both local and national/international standards and codes. Applicable codes often cross beyond electrical design practices into disciplines such as structural-related building codes and civil engineering standards. The following list, while not intended to be comprehensive, is nevertheless representative of the range of codes and standards that are typically considered in the course of designing a modern commercial system. Of course, references to National Electrical Code (NEC) apply only to USA projects.

General standards referring to the solar radiation measurement process and instrument calibrations are defined by ISO 9845-1, DIN 5043-2, and IEC 61 7725.

International standards and certifications [18] referring to solar cells and modules include: EN 50380 (data sheet and nameplate information), IEC 60891 (procedures for temperature and irradiance corrections for C-Si), IEC 60904 (PV devices), IEC 61277 (terrestrial PV guide), IEC/PAS 62011 (renewable energy in rural decentralized electrification), IEC 61215 C-Si (terrestrial PV modules), IEC 61345 (UV testing for PV modules), IEC 61646 (thin film terrestrial PV modules), IEC61701 (salt mist corrosion), IEC 61721 (susceptibility to accidental impact damage), JRC-ISPRA 503 (qualification test procedures for C-Si PV modules), IEC 61829 (C-Si PV array), IEEE 929 (recommended practice for utility interface of residential and intermediate PV systems), IEEE 1262 (recommended practice for qualification of PV modules), and IEEE 1513 (recommended practice for qualification of concentrator photovoltaic modules). Organizations such as TUV Rheinland, CSA, and ETL have all offered PV certification services, often in accordance with companion IEC or UL listings.

Standards also exist for PV system advising planning, implementation, and safety. They include: IEC 60364-7-712 (electrical installations of buildings), IEC 61194 (off-grid characteristics), IEC 61702 (dc PV pumping systems), IEC 61724 (performance monitoring), IEC 61727 (PV utility interface), IEC 61683 (Power conditioners – Procedure for measuring efficiency), IEC/TR2 61836 (terms and symbols), IEC 62124 (off-grid PV design qualification and approval), IEEE 928 (recommended criteria for terrestrial PV power systems), IEEE 1373 (recommended practice for field test methods and procedures for grid-connected PV systems), and IEEE 1374 (guide for terrestrial PV power system safety).

Common codes used for U.S. systems include the National Fire Protection Association's (NFPA) National Electric Code (NEC), especially section 690 pertaining to PV, as well as design, installation, and operation practices stipulated under various Occupational Safety and Health Administration (OSHA) guidelines and the Uniform Building Code (UBC).

Finally, standards and certification for modules or other parts/components of photovoltaic systems, including battery-backed systems, include: IEC 61173 (overvoltage protection), IEC 61683 (Power conditioners – procedure for measuring efficiency), IEC 61427 (secondary cells and batteries), IEEE 937 (recommended practice for installation and maintenance of lead-acid batteries for PV systems), IEEE 1144 (sizing of industrial nickel-cadmium batteries for PV systems), IEEE 1145 (recommended practice for installation and maintenance of nickel-cadmium batteries for PV systems), and IEEE 1361 (recommended practice for determining performance characteristics and suitability of batteries in PV systems).

19.6 SYSTEM DESIGN CONSIDERATIONS

PV systems evolve over a series of design steps. Beginning with a site analysis, the subsequent system design involves selecting, sizing, and integrating components. This is followed by installation,

an economic lifetime analysis of system operation and maintenance, and finally, a salvage or redeployment stage. This section reviews design steps and shows how each contributes to a successful PV system in terms of performance, safety, and reliability.

19.6.1 Site Analysis

Surveying the site prior to designing a PV system is fundamental to having a PV system that meets the needs of the end user and is optimal for the site in terms of load, intended use, and climate.

19.6.2 Location

The latitude and longitude of the site defines the sun paths, angles and the pattern of daylight hours. Of equal importance is the local climate, nominally expressed in terms of equivalent peak sun hours per day, as kWh/m^2/day. Combined, the geographical coordinates and local climate determine the potential amount of solar resource available at a specific site. [See Figure 1.3 in Chapter 1 for the global solar resource map.] It may be difficult to obtain historical meteorological data from the exact proposed location. Picking the most appropriate weather source is an arduous task and at times rather subjective. This is partly due to microclimate or small gradients between monitoring stations and the site, and is also due to uncertainty in the method of measuring solar radiation. There are several different kinds of weather resource data sources available.

- Ground-based weather stations
- Satellite-based data
- Computer-generated interpolations, some using ground and satellite hybrid data sources.

In addition to long-term average data, multiple-year data sets enable designers to characterize year to year variability (See Chapter 22). After solar radiation, the ambient temperature range experienced at the specific site is the next most important climate variable. Extremes of low and high temperature as well as average daytime temperatures are all important for PV design. When designing a system it is prudent to define the maximum temperature as the average of the record high temperature and the warmest month's daily average maximum temperature. Similarly, defining the minimum system temperature as the average of the record low temperature and the average minimum temperature coldest month's daily. These temperature ranges are used for sizing the number of modules in each series string so that the PV DC voltage range is compatible with the inverter and other balance of system components.

19.6.3 Orientation and Tilt

The orientation and tilt of a system impacts how much of the available irradiance the system can collect. There are general rules of thumb to follow when orienting a fixed tilt system. Theoretically, the optimal orientation, or surface azimuth, is true south (not magnetic south) and the optimal tilt is equal to the latitude (Chapter 22). However, empirically, it is generally preferable to have the system facing the equator and tilted at approximately 10–15° less than the local latitude. This is principally a consequence of poor weather being concentrated in the winter months. Other factors that influence the optimal orientation and tilt are: (1) convenience (an existing slope is often less expensive to install upon); (2) local obstructions (shading due to trees and surrounding buildings); (3) asymmetrical microclimates (consistent morning fog or afternoon showers); and (4) sensitivity to time-of-delivery generation. California's incentive programs provide the largest rebate to systems optimized for May to October energy. In this case the system should be at a slightly lower tilt angle

than otherwise would be chosen to maximize annual generation. For tracking systems orientation issues are less significant since the system tracks the sun throughout the day.

19.6.4 Shading

Shading has a surprisingly disproportionate impact on PV output. Shade on as little as 5–10% of an array can predictably reduce its output by over 80%. To minimize this loss a site shading survey should always be performed. Two types of shading can be defined for design purposes: (1) near-field and (2) global (also referred to as horizon).

At any moment, near-field shading affects only a portion of an array, while horizon type shading can be viewed as affecting either all or none of an array. Examples of near-field shading would include local obstructions such as trees, walls, rooftop equipment, or neighboring rows of panels. Near-field shading is electrically equivalent to mismatch. If one module in a string is shaded it may have the same effect as if the entire string were shaded, as the entire string can only carry the same amount of current as its weakest link.

Examples of horizon shading would include distant hills or even nearby objects if they are very large relative to the size of the array such as adjacent rooftops or buildings. Horizon shading prevents any beam radiation from striking the array, and its effects are uniform and simultaneous. Thus, a horizon-shaded array will only be exposed to diffuse radiation.

Near–field shading is difficult to model but fairly easy to avoid once a site survey is done. Conversely, horizon shading is easy to model but difficult or impossible to mitigate. There are several tools that can be used to evaluate the extent of shading at a given spot. Some examples include: the Solar Pathfinder [19], the Sun Selector [20], the Wiley Electronics ASSET [21], and the Solmetric SunEye [22].

For simulation purposes, some newer software is sophisticated enough to model both near-field and horizon effects. Older software (>10 yr) tends to ignore shading or treats it as a constant loss factor. Simulation software is discussed in a subsequent section of this chapter.

For multi-row commercial systems, row to row shading is inevitable, but designers have the ability to choose array geometry to satisfactorily minimize this type of shading loss. Latitude and climate are the chief factors that influence what array geometry will be needed to achieve a design shading loss target, but for a given location, the principal design variables include: row spacing, array tilt and azimuth, and module orientation (portrait or landscape). Many designers try to limit annual shading losses to 2–4%. In practice, this means spacing rows such that the setback ratio is at least 2:1 in sunny, lower latitude regions and at least 3:1 for cloudier mid-latitude regions. The setback ratio (SBR) is defined as the horizontal distance, or gap, between rows, divided by the vertical distance between the high and low sides of adjoining rows. According to Figure 19.20, the SBR, the critical shade angle (α), the ground cover ratio (GCR), and the array tilt (β) are related.

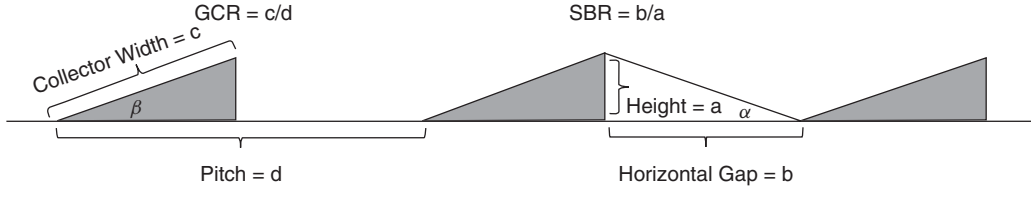


Figure 19.20 Row spacing geometry with module facing to the left at tilt angle β

The GCR is defined as the array area divided by the ground area, or, for an array of unit depth, as the row width, c, divided by the row spacing, d.

$$\alpha = \tan^{-1}(1/\mathrm{SBR}) = \tan^{-1}(\mathrm{a/b})$$
$$\mathrm{GCR} = \mathrm{c/d} = (\cos(\beta) + \mathrm{SBR}^{*}\sin(\beta))^{-1}$$

Figure 19.21 illustrates how shading affects design geometry (tilt and spacing) for contemporary multiple row arrays[4]. This example is situated in Geneva, Switzerland. The array has a setback ratio of 3:1, which corresponds to a noon time critical shade angle, or constant limit angle, of 18.4° for a south-facing surface. The upper line represents the output as a function of tilt angle for a single row and shows an optimal tilt angle of 36° (which happens to be 10° less than the 46° latitude at Geneva). The lower line represents the output for a realistic array with a constant setback ratio of 3:1 for any choice of tilt angle. The perhaps surprising and fortuitous conclusion is that the optimal tilt angle for this (and many) commercial array(s) is several degrees less than might otherwise be selected in the absence of row to row shading. In this case a 30° array tilt would not only be the best choice for maximizing energy production, it would also reduce wind loading and, potentially, structural cost since the structural element would not need to be as strong or as large.

The choice of portrait versus landscape module orientation becomes important whenever significant row-to-row shading occurs regularly. As most arrays are aligned south-facing, E–W running shade lines will creep along the lower edge of neighboring rows. Because of the physical

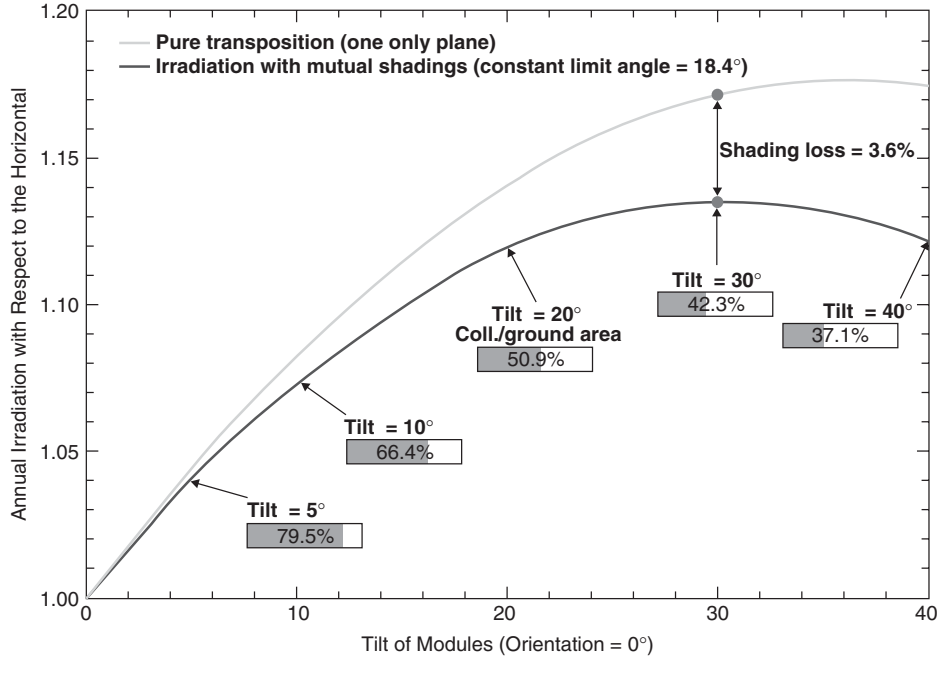


Figure 19.21 Geneva, Switzerland, optimal tilt angles for one unshaded row and for multi-row shading, for a constant setback ratio of 3:1

[4] PVSYST v4.37, University of Geneva, Switzerland.

layout of bypass diodes within most modules, it is advantageous to mount them in a landscape direction so that a shaded "stripe" can be bypassed without sacrificing the entire module's output. For N–S shade lines originating from objects east or west of an array, a portrait module alignment would be advantageous. Finally, most thin-film modules are less susceptible to shading. Some include bypass diodes anyway, but their "softer" $I-V$ curves (that is, lower fill factors) and "sliver" shaped cells make them more tolerant to irregular illumination. The end result is that thin film modules tend to exhibit shade-related reductions in output that are more nearly linearly proportional to the shaded fraction of total array area.

For modern tracking systems, a feature called backtracking is normally used. Backtracking is the purposeful retreat of a tracker away from its optimal rotation angle to avoid row-to-row shading. The principle behind backtracking is that the energy sacrificed due to non-optimal higher incidence angles is less than the energy that would be lost due to shading. In essence, backtracking permits a little higher packing density relative to the non-backtracked case.

The ground cover ratio (GCR), or the active array planar area relative to the ground footprint area, is more significant than array layout (e.g. rectangular or "soldier" layout versus hexagonal layout) in terms of shade impact. Most fixed-tilt arrays at tilt angles optimized for maximum annual energy will be installed with GCRs in the 0.4–0.5 range, with shallow tilt arrays often packed in much more densely. Horizontal one-axis trackers can achieve shade losses comparable to fixed-tilt arrays with only slightly less dense GCRs, typically in the 0.35–0.5 range. Tilted one-axis and two-axis tracking arrays generally need to be spaced at GCRs in the 0.20 range or less in order to achieve comparably low annual shading losses of about 2%. [23, 24]

19.6.5 Dust and Soiling

Dust, soiling, leaves, bird droppings, soot, snow, and frost will all reduce the amount of electricity that can be produced by a system. These effects tend to be seasonally dependent and can also vary significantly among climate types. Dust accumulation is influenced by local weather patterns, local soils, air and automobile traffic, and agricultural activities. In places where soiling causes a significant loss in electricity production timed array washings should be considered. This is, of course, a labor-intensive process that requires significant quantities of water, so the choice of washing depends on the relative financial benefits associated with the expected increase in generation. Soiling can cause monthly losses of up to 25% and annual losses of 7% if not mitigated properly. Studies have found a single washing in the middle of the dry season can reduce annual dust losses in half, from a typical 6% loss to just 3% [25]. Snow losses have not been well characterized. Until recently, few commercial PV systems have been deployed in areas that regularly receive significant and lasting quantities of snow. Small off-grid systems have been successfully deployed in snowy climates for many years, but the prevailing design wisdom has been to simply tilt the modules at high angles of 60° or more to ensure rapid shedding of snow. Modern commercial and utility scale systems on low-slope roofs and large ground-mounted structures are far more prone to snow losses. Annual losses of 2% or more are likely to occur in snowy climates, with one rooftop study in southern Germany finding a six-year range of 0.3–2.7% per year loss attributable to snow. [26]

19.6.6 Roof and Ground Considerations

The choice of mounting a PV array on the roof or on the ground is governed by some simple considerations. In addition to a roof's orientation and area, its suitability for array mounting is influenced by characteristics such as:

- Membrane type: shingles, tile, metal, built up (tar and gravel), roll roofing, thermoplastic (TPO), PVC, concrete, or rubberized (EPDM)
- Age and basic condition
- Strength; dead load, live load, wind load, and seismic load limits
- Accessibility: equipment staging, fire and service access

For ground-mounted arrays some typical considerations include:

- Soil type
- Drainage
- Trenching feasibility
- Vegetation control
- Habitat needs
- Security

Carport/shade structures represent a special case of ground-mounted arrays. They offer the dual benefit of providing shade for vehicles as well as producing electricity (see Figure 19.22). While the elevated structures are more costly to install and maintain than standard ground mount structures, they allow a second use for land that has already been compromised by development. This may have added public relations benefit because some people like to associate the solar-generated electricity with the charging of electric vehicles parked below.

19.6.7 Interconnection Equipment

A key part of site evaluation includes an assessment of the available electrical infrastructure. PV systems represent a second source of power to supplement the utility feed. The common practice

Figure 19.22 Carport system, Diablo Valley College, Pleasant Hill, California (Courtesy of SolarWorld)

is to interconnect the PV source with the grid at either a tap point on the utility side of a facility's main service disconnect (but still on the customer's side of the utility revenue meter), or as a branch within an existing service panel. Some considerations include:

- Main service panel ampacity, voltage, and phasing.
- PV system ampacity.
- Main service panel age, condition, and available space for additional branch.
- Distance and routing from PV switchgear to facility switchgear (to avoid excessive wire loss and cost).
- Local utility metering requirements (service access rules differ among utilities, where some require full isolation capability for meter access).

19.6.8 Load Data

The magnitude and shape of a facility's daily load pattern is instrumental in helping determine the optimal PV system size. Load information is not relevant for PV systems operating under feed-in tariffs, but it is for the majority of PV systems, whether on-grid or off-grid.

Many commercial utility rates are structured with both energy and peak demand charges. Energy charges are customarily differentiated by price according to time of day and season. Demand charges are customarily based on the highest 15-minute power consumption per month. Residential rates, though exempt from demand charges, are often tiered in progressively more expensive energy blocks to encourage conservation. Residential, commercial, industrial, and agricultural customer classes typically exhibit distinctly different daily and seasonal load patterns, which must be accounted for and mapped against PV output profiles. Thus, a system designed to maximize the ability of the PV to offset the most expensive portion of the load might be different from one whose goal is to maximize the annual kWh generated.

For off-grid systems, load characteristics dictate storage needs, as well as PV system size and backup power capacity, if needed. For example, an off-grid home occupied year-round will typically include enough storage capacity to satisfy energy requirements during extended periods of poor weather, also known as days of autonomy, where 3–10 days is a common design range. Some designers use the term loss of load probability (LOLP) instead of days of autonomy. Indirectly, the size of the battery bank influences the PV array size, since the array must be large enough to recharge the battery bank, while supplying the loads, in a reasonably short time under average weather conditions.

19.6.9 Maintenance Access

While PV requires very little maintenance, it is still important to be able to access the system to replace failed components and to perform scheduled routine inspections and maintenance. Oftentimes, rooftop PV systems have ground-mounted inverters, switchgear, and data acquisition cabinets. Access to ground equipment is often controlled with fencing or by being located indoors and therefore requires appropriate levels of coordination in order to service equipment. Important maintenance access items include:

- Security clearances to allow for access to the system.
- Storage for spare parts and tools.

- Water for array washing.
- Assess service areas for existing rooftop equipment and fire protection.

19.6.9.1 Sizing

Almost invariably, sizing is modeling; often, sizing is iterative; sometimes, sizing is the crux of the classic optimal system design problem. It is the solution to converting a power-based commodity (kW_P) into an energy-based commodity (kWh). The economic "sweet spot" for optimal sizing is fairly forgiving for grid-tied systems, but fairly critical for stand-alone systems. The reason is that utilities will seamlessly and automatically deliver the power not supplied by a grid-tied PV system, but the off-grid PV system owner must assure that sufficient power is available at all times.

19.6.9.2 Grid-tied

There are several good criteria for sizing on-grid PV systems. Only one can truly be considered an optimal design, that is, sizing for best economic value. However, the other methods listed below are in wide use and are generally easier to follow:

- Sized to meet 100% or some other target fraction of annual kWh load.
- Sized to offset 100% or some other target fraction of peak power kW demand.
- Sized to fit available area (ground, roof, carport, BIPV).
- Sized to obtain maximum incentive (i.e. rebates).
- Sized for optimal economic value (this requires a comprehensive evaluation of all costs and benefits).

19.6.9.3 Off-grid

Off-grid system sizing is usually done by sizing the PV system, or the PV system in conjunction with supplementary power source(s), to match the load requirements. The load requirements however, are more complex. Each of the following conditions must be satisfied:

- Sufficient inverter capacity to meet peak power demands (for DC systems substitute sufficient charge controller, breaker, and wire capacity).
- Array sized to meet energy needs after a minimum number of days of autonomy under seasonally worst weather patterns (or more generally, the worst combination of weather and load).
- Array sized to meet maximum allowable days of backup generator use.

19.7 SYSTEM DESIGN

The essence of system design is to make use of the results of the site analysis, then select and integrate components that will meet the performance targets. PV designs are often iterative, that is, system size or other parameters are sometimes honed to optimize economic value. The first design and decision fork, of course, is to determine whether the system will be on-grid or off-grid. Almost all off-grid systems will use low voltage DC (up to $\approx$50 V). For smaller on-grid systems up to $\approx$2 kW, designers can opt between providing low voltage DC ($\approx$50 V) or higher voltage DC (up to 600 V) to the inverter. The main considerations between these two voltage classes are the relative

safety and shade tolerance of sub-50 V systems versus the higher efficiency, lower wire losses, and reduced system cost for 600 V systems.

19.7.1 Component Selection Considerations

19.7.1.1 Modules

At 50% or more of the cost of a typical PV system, modules are the most significant single cost element. Despite their importance to overall system cost, modules are often priced similarly enough such that non-price considerations end up governing module selection.

Some of the many non-price considerations include:

- *Efficiency*. Often, area constraints dictate whether low-efficiency modules (5–8% DC-STC module efficiency) are suitable. High-efficiency modules (>13%) have the additional advantage of reducing all area-related costs for items such as structure, wiring, and land.
- *Ease of installation*. Framed modules with pre-drilled mounting holes tend to be favored over frameless modules, though some frameless modules such as peel-and-stick laminates are even easier to install than framed ones. Modern plug connectors have largely displaced traditional junction boxes with terminals. Some manufacturers offer better grounding methods than others, and feedback from installers suggests this can be a strong consideration.
- *Availability*. World demand for PV modules has at times outpaced supply, causing delays or last minute expensive product substitutions.
- *Weight*. Lighter weight panels are cheaper to ship, easier to handle, and present less structural demands.
- *Warranty*. Most warranties carry similar terms of 1–3 year workmanship and 20–25 year performance level.
- *Physical size/aspect ratio*.
- *Degradation*. Historically, thinfilm PV has tended to degrade more rapidly than wafer-based PV.
- *Current, voltage, and power parameters*. Each of these must be compatible with the DC and AC characteristics of the inverter and the limitations of the AC interconnection.
- *Aesthetics*. Some products such as BIPV (see Chapter 23) are specifically designed and marketed to visually blend in well with their surrounding architecture, appearing as uniform featureless black rectangles (no grids, no frames, no white background between cells).
- *Recognized certifications or listings*. Testing organizations such as IEC, UL, CE, CSA, TUV Rheinland, and ETL/Intertek all provide certifications that increase consumer confidence and in some cases, eligibility for incentives.
- *Manufacturer history*.

19.7.1.2 Module nameplate binning practices

In the context of system design, performance targets will be affected by manufacturers' module binning practices. Variations in power from module to module are inevitable, with ranges of ±10% common at the time of the first edition of this book. As manufacturing techniques have improved, production tolerances have tightened to the ±5% range. Until recently, many manufacturers practiced unbalanced binning, where modules exceeding the nameplate power rating would be rebranded as a new product with a higher nameplate power rating. The pool of modules destined for the consumer under the original nameplate rating would be, on average, somewhat short of the

advertised value. Typical shortfalls of 2–3% were common. The contemporary emphasis on actual performance, as evidenced by European style feed-in tariffs, has encouraged manufacturers to meet their advertised nameplate power ratings. It is now common for manufacturers to cite *balanced* bin ranges as narrow as ±2%, with the secondary benefit of reduced mismatch.

19.7.1.3 Degradation

Most PV systems lose power over time in some combination of two of the following three ways:

- Staebler–Wronski (S–W) degradation.
- Light-induced degradation (LID).
- Long-term degradation.

S-W degradation occurs in thin film silicon modules and generally causes a reduction of 15–25% in power over the first 1000 h of exposure. While this type of degradation is partially reversible by annealing for prolonged periods at slightly elevated temperatures (50–150°), the S–W process will recur upon renewed exposure to sunlight (see Chapter 12).

LID is an irreversible loss of 1–3% in power that occurs in wafer type silicon modules. It is thought that oxygen impurities that get trapped during the ingot formation process are the main cause of this small but common phenomenon [27].

Long-term degradation applies broadly and probably universally among PV technologies [28]. The literature varies considerably on what the magnitude and mechanisms are behind this phenomenon. Annual power declines of 0–2% or more have been reported, with most hovering in the 0.3–1% range [29–31]. Degradation in this chapter's context is meant to include both module-level and any additional system-level effects.[5] [32, 33] Most of the lower 0.3–0.5% data are associated solely with modules; reports in the 0.5–1% range tend to be system-level results [34]. Module-level degradation has been associated with encapsulant deterioration, internal delaminations, and increased series resistance due to solder bond thermal fatigue [35]. Increasing series resistance progressively reduces the maximum power point voltage and therefore, power. Although manufacturers endeavor to minimize series resistance losses, daily thermal cycling of modules deployed outdoors results in a gradual increase in series resistance [36]. The causes of system-level degradation are less well understood. Corrosion at wire terminations, degradation of inverter circuitry, and in many cases, unaddressed module failures within a larger field of panels have all been cited as potential system-level degradation mechanisms.

19.7.1.4 Inverters

The primary consideration in the selection of an inverter for PV systems is that it is compatible with the array output DC voltage characteristics and that it has the proper interface capabilities for the grid or loads to which its output is connected.

[5] A survey of long-term system degradation literature shows repeated references to the 0.5–2% range. PowerLight (now SunPower) reports – 0.5% based on field measurements. The PVUSA project observed 1% degradation as typical for flat-plate crystalline silicon PV systems. Others, such as NREL, the Southwest Technology Development Institute and Ben-Gurion University, Israel, have reported 1% degradation rates at past NREL performance and reliability workshops. David King of Sandia reports module degradation of 0.5–2% for 1991 to 1998 modules. Virtually all manufacturers' warrantees allow for 1% degradation, in most cases starting from the low end of their nameplate power tolerance. Kristopher Davis and H. Moaveni conducted a 4-year study of two systems and found a 1.2%/year degradation for the c-Si system and a 2.1%/year degradation for the a-Si system.

The DC voltage range from the output of the PV array must be within the allowed input DC voltage range (see string sizing section below). This requires that the maximum allowed inverter voltage be higher than the maximum open-circuit voltage (V_{OC}) in the worst case conditions (cold cells at high irradiance). The minimum voltage that the array will typically ever produce significant power at (hot cells at high irradiance) should also be above the inverter minimum operating voltage limit. Naturally, the array's maximum power point voltage (V_{MP}) range should also fit squarely within the inverter's operating range. Examples are given in the later sections of this chapter. A common range of PV DC voltages for utility scale inverters in the US is 300 V_{DC} to 600 V_{DC}. It is common for utility scale inverters in Europe to operate at a higher level such as 450 V_{DC} to 1000 V_{DC} due to differences in the codes and standards that allow a higher maximum DC level in equipment such as this. Also of note is that DC arrays are typically grounded on either the negative or positive DC side of the array in the US. In Europe it is common for the array to operate in an ungrounded configuration. This is also due to differences between electrical codes and standards and the inverter must be designed to operate safely with the grounding system it is connected to (see section on grounding below).

For successful system operation, the PV inverter must operate at the same voltage and frequency as the utility grid for grid-tied systems. This is typically 220 V_{AC} at 50 Hz in Europe and many parts of the world and 120/240 V_{AC} at 60 Hz in North America for residential single-phase systems. The voltages are higher for three-phase commercial and utility-scale systems. Examples are 480 V_{AC} in the USA and 690 V_{AC} in Europe. Some large utility-scale system inverters step the output voltage up to medium voltages (commonly 12–70 kV) through the use of a transformer. For off-grid systems, the AC voltage and frequency that the inverter produces must match those required by the loads that are powered by the system.

The final consideration is the power rating of a PV inverter. Residential scale inverters from a given manufacturer will be available in discrete power ratings, such as 3, 4 and 6 kW. If the power available from the PV array is above the power rating of the inverter then the inverter will "clip" the power and limit its output. As the power available from the array is almost always less than the nameplate power of the modules alone, it is common, but not essential, that the array power rating be somewhat larger than the inverter power rating. A common factor of 1.2 times or more is used in system design. This would mean that an array with a set of modules that could produce 12 kW_{DC} at STC may be connected to a 10 kW_{AC} inverter. There may be some times that the inverter must forego power by operating off of the array's maximum power point, but in a properly designed system only a small fraction of a percent of potential generation will be lost [37].

The above descriptions are for systems with "central" or "string" inverters in which multiple PV modules are connected to a single inverter. Another option is to use a "microinverter" which is a dedicated inverter connected to each PV module. These are relatively small and fit within the space of the module frame. Microinverters are typically rated at 100 to 300 W and they are designed to match the PV module voltage characteristics. They have at least three advantages. (1) Microinverters can maximum power point track each module accurately to maximize the amount of AC power output. This helps in situations where shading may be an issue so that any one module being shaded will not impact the output of an entire string. (2) They are intrinsically modular and a system can be relatively easily expanded by adding additional modules with their associated microinverters. (3) They are individually addressable via software, thus allowing higher-resolution (module-level) monitoring to diagnose performance issues as opposed to inverter level monitoring. However, the disadvantages are that they are presently expensive on a $ per Watt (or € per Watt) basis compared with central inverters, the high number of total system components, and reliability is of a great concern as the environment underneath a PV module can be very harsh for electronic equipment. They will be exposed to the same temperature and humidity as the back surface of the PV module. These disadvantages may well decrease over time as manufacturers gain more experience and feedback from the field.

19.7.1.5 Electrical BOS and switchgear

Auxiliary components such as wire, conduit, switches, combiner boxes, fuses, and all small fittings need to be sized to support the current, voltage, and power characteristics of the proposed system and need to carry the appropriate specifications for the intended service environment. Of course, as with the PV system's major components, these components need to comply with local electric codes, which in some cases mean carrying listings from recognized laboratories (most US jurisdictions require UL-listing for each of the above-mentioned parts).

19.7.1.6 Storage

Many different qualities of batteries are available. To choose the best option for a particular application the more relevant data to analyze are installation labor costs, purchase price, maximum recommended depth of discharge, projected operational lifetime and cycle lifetime from the data sheet.

For a long-term application, where access to the PV system is difficult and the heavy batteries are expensive to transport, there is a strong incentive to purchase the highest-quality batteries with the longest lifetimes (8–12 years). It may even be advisable to over-size the PV array to assure frequent full charge status of the battery. On the other hand, in a region where batteries are readily available and easily replaced, it may be preferable to select a more economical option that will need to be replaced more frequently (3–5 years). Responsible battery disposal or recycling should be considered as an upfront design issue, especially when lower grade batteries are selected.

Excessive discharge can damage the battery and must be avoided by limiting the routine daily discharge to a minimum SOC that, depending on the type of battery, ranges between 20 and 70%. If this lower set point is reached, the loads have to be disconnected. The *practical or useful* capacity is the fraction of the nameplate rated capacity that can be extracted before reaching the minimum state of charge as recommended by the manufacturer. For the same required *practical* capacity, a battery with a lower allowed SOC will have a smaller nameplate rated capacity. All batteries experience a slow self-discharge even when not connected to a load. Prolonged periods of neglect, especially when not immediately followed by a full recharge, will age and compromise the *practical* capacity. Temperature also strains battery life. Temperatures higher than 20 °C increase the rate of self discharge and temperatures below freezing retard the chemical reaction and limit the available amps that may be drawn from the battery. It is essential that batteries be shielded from temperature extremes and that they be kept in a dry, well-ventilated location to avoid build-up of explosive gases (Figure 19.23).

The battery of an off-grid PV system is sized so that its *practical* capacity endure 3–10 days of autonomy, sometimes more in critical applications, to supply the average daily energy requirements. For example, if there is a backup generator, or if part of the average demand are nonessential loads, then a shorter autonomy is sufficient. For a winter daily average load of 2 kWh, and a 3-day autonomy target, a practical capacity of 6 kWh is needed. For a depth of discharge of 50% a 12 kWh battery is required. For a typical nominal battery bank voltage of 48 V, 250 Ah of battery storage would be required.

For small PV plants that supply daily loads in the range of 100–500 Wh, 12 V is the most common system voltage, often because certain DC appliances can be directly connected to the DC bus. For larger system sizes that typically supply all of their loads through an inverter, 24 and 48 V inverters and battery banks are common.

Figure 19.23 Battery bank in off-grid PV household. The cabinet is located on the shaded part of the building to avoid overheating, is easily accessible, well ventilated, with safety goggles, densimeter, and funnel at hand (source TTA)

19.7.1.7 Structures

The principal design consideration for structures is that they secure the modules to the underlying material, whether rooftop or ground-mounted, over a lifetime of potentially harsh outdoor conditions. Typically the modules are securely fastened to the structure, but the structure itself can either be securely fastened (bolted), or merely weighted down on the surface underneath (ballasted). Sometimes local climates preclude the use of certain structural materials; for example, some coastal jurisdictions prohibit the use of aluminum because of its susceptibility to salt corrosion. Structures must be designed to withstand the weight of the array components, the upward, downward, and lateral wind forces under local design extremes, seismic forces in accordance with local codes, as well as live loads during installation and servicing. Tracking structures have the additional design challenge of maintaining accuracy between service intervals, which in practice means ensuring that snow, debris, or electrical transients do not disrupt operations.

For low-slope commercial rooftops, Table 19.2, adapted from a 2009 article in SolarPro magazine, points out some differences between attached and ballasted structural options [38].

19.7.1.8 Grounding

Grounding is done to reduce both fire and personnel shock hazards. The grounding methods favored in North America differ from those used elsewhere in one fundamentally major way. While safety, or equipment, grounding is done on virtually all systems above 25 V, whether DC or AC, North American systems differ from others in that one of the *current-carrying* conductors, also defined

Table 19.2 Commercial low-slope roof PV structure considerations (adopted from Ryan Mayfield, *Solar Pro* (February/March 2009)

Design issue	Structurally attached	Fully ballasted
Cost for racking components	Lower	Higher
Installation labor cost	Higher	Lower
Coordination between trades	High: penetrations need flashing and sealing	Low: pre- and post-installation may suffice
Typical roof loading	Typically lower than ~3 psf	5–10 psf or higher
Roof maintenance and replacement	Easier	More difficult
Typical tilt angle range (°)	5–45	0–20
Accommodates low profile obstruction on roof	Yes	Less accommodating
Accommodates roof pitch changes	Yes	Less accommodating
Impact on roof drainage	Minimal	Design around potential problem areas required
Meets seismic code requirements	Yes	May need adhesives or structural attachments
Dust susceptibility	Less prone	More prone

as the *neutral conductor*, is intentionally grounded. The principle is that if the neutral conductor is grounded at all times, a flaw developing elsewhere will complete the circuit through the ground. The faulted circuit will effectively keep equipment that is not intended to carry current de-energized, making it inherently safer to service a system that is known to have a problem. Unfortunately, this type of system is extremely hazardous to touch under normal operating conditions. With ungrounded systems, the opposite is true: a normally operating circuit can be contacted without hazard, but a faulted circuit can energize non-current-carrying equipment, making it more hazardous to service.

19.7.2 Economics and Design

Maximizing annual energy is normally, but not always, the optimal system design choice. The reasons for this are different for off-grid versus on-grid systems. For off-grid systems, it is very common to design systems that are maximized for winter output at the expense of annual output. There is usually a surplus of energy produced in the summer and it often more important to minimize either outage time or generator run time.

For on-grid systems, the structure of system incentives, i.e. subsidies, can have a strong influence on system design, as can the structure of utility rates. Feed-in-tariffs promote designs that maximize yield. One-time upfront rebates and tax credits have an increased emphasis on system capacity. Many utilities offer rates that vary by time of use, a trend being emphasized in the emerging smart-grid concept. Commercial rates for larger customers usually include a demand charge component, too. While this component is usually much smaller than the energy-based component of the bill, for certain customers it may be significant enough to influence the orientation and size of the system. Finally, residential utility customers often pay increasing rates for progressive blocks, or tiers, or energy use. Each of these considerations will influence system design.

Several measures of economic value are in wide use. Perhaps the simplest is the cost of energy, normally expressed as a long-term discounted measure called the levelized cost of energy (LCOE) [39]. This term factors in the capital cost or loan terms of the solar investment plus operating costs, less upfront and non-performance-based incentives, and levelizes the annual and often irregular payments via a selected discount factor.

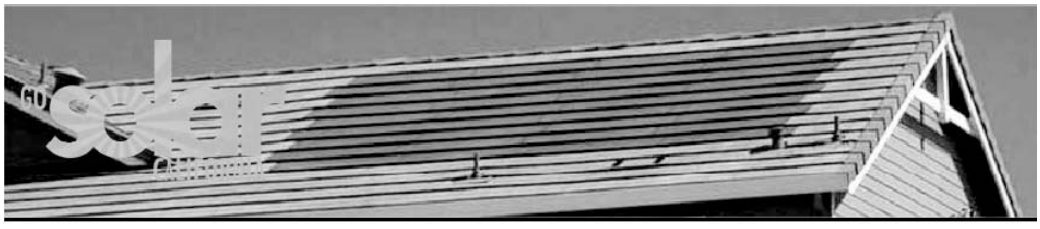
When performance-based benefits such as feed-in rates, power purchase energy sales, avoided energy costs, and renewable energy credits (i.e. green tags) are factored in, more comprehensive and traditional figures of economic merit such as net present value (NPV), benefit/cost (B/C) ratio, payback, and internal rate of return (IRR) are often used.

One of the strongest influences on economics and design is the type of incentive. These are broadly termed as either capacity-based, as in upfront rebates, or energy-based, as in feed-in tariffs. Much of California's PV development over the past decade has been subsidized by a one-time state rebate, which has decreased from a high of US$ 3.70/W_P ($4.50/$W_{CEC-AC}$) in 2002 to US$ 1.10/W_P for most customer classes in 2010. When the California state rebate is considered in conjunction with US Federal incentives of a 30% tax credit and a 5-year accelerated depreciation allowance, the combination skews the total of a system's lifetime benefits to be 80% non-performance-based. With such a small portion, 20%, of lifetime system benefits at risk due to actual performance, there tended to be many systems installed with severe shading, inferior orientation, and overly dense row spacing. None of these performance drawbacks would have a negative influence on system economics. In fact, there was a distinct incentive to place as much capacity as a given area could hold.

That situation has shifted, with all California systems greater than 30 kW now compensated via a 5-year performance-based incentive that rewards actual energy delivery instead of capacity. While this is still a lot different than the European style feed-in tariff administered over a typical 20 year period, it nevertheless shifts the lifetime ratio of non-performance-related benefits to performance-related benefits from an 80/20 mix to about a 60/40 mix and thereby encourages more effective designs.

The factors that influence economic value are numerous and complicated. To provide some practical context to the difficulty of evaluating economics and design, the following example case is presented to roll up each of the relevant system design and economics terms into a concise set of results. The example uses a popular free web-based US software program called Clean Power Estimator (CPE) [40]. The system is rated at 1 MW_{CEC-AC}/1.2 MW_P. The installation is in the San Francisco area at a facility that currently pays $30,000 per month for electricity. The proposed system faces southwest at a 10° tilt. It is assumed to have an installed cost of $6/$W_P$ before incentives, with a 5% loss assumed for dust and shading. Grid-supplied electricity is selected to escalate at 3.5% per year, $1000 per year is specified for maintenance, and the system is paid for with a 7%, 10-year loan.

Figure 19.24 summarizes the various inputs and key results. The proposed system is expected to produce 1,618 MWh per year, or 79% of the customer's energy needs. The corresponding yield is 1,383 kWh/kW_P. Figure 19.25 shows the cumulative discounted cash flow, which, because of the loan structure, never goes negative and finishes the 30-year period with a net present value (in 2010 dollars) of $2.7 million on the basic $7 million investment. The cumulative discounted cash flow gives a truer expression of payback than the 8.6 year (simple) payback term listed in the system summary table within Figure 19.24. The payback listed in the summary table is normally labeled as simple payback, as it only relates first year energy value to initial net system cost. The incentives in this example are threefold: one is a 30% Federal tax credit, another is the Federal and State accelerated depreciation allowance, and a third is California's 5-year performance-based incentive payment at $0.10/kWh.



System Summary

Net system cost after all incentives:	$4,268,980
PV system electricity production:	1,618,001 kWh/year
Electricity production supplied by system:	79%
Carbon dioxide emission reduction:	1,791,412 lbs per year
Internal rate of return:	17%
Net present value:	$2,669,411
Years to payback:	8.6

General Assumptions

City, State	San Ramon CA
Utility	Pacific Gas and Electric Company (PG&E)
Rate	Net Metered Small General Time-of-Use Service (Schedule A-6)
Technologies	PV only
Recalculation	After each input

PV System

Size	1.0 MW-ac PV System
Cost	$7.000 per kW-ac
Tilt	10º
Orientation	SW
PV Output Adjustment	-5%
O and M Cost	$1,000 per year

Financial Assumptions

Incentive Option	EPBB
Tax Status	Taxable
Taxable Income	$10,000,000 per year
Electric Bill	$30,000 per month
Escalation	3.5% per year
Payment	Loan
Loan Life	10 years
Loan Rate	7.00%

Place cursor over the ? symbols for help

Figure 19.24 Clean Power Estimator inputs and results for 1.2 MW_P fixed-tilt system in northern California

19.7.3 System Integration

System integration is the process of matching the specifications of the various system components to the operating conditions presented by a given location. A one-size-fits-all approach is not effective for PV. In short, not all modules are compatible with all inverters, and certainly not in all environments. The system integration process must begin with an understanding of the temperature extremes likely to occur at a given location, as these parameters are not within the designer's control.

As noted above in the inverter selection discussion, it is both the lowest array voltage on the hottest day and the highest array voltage on the coldest day that determine the combination of module type, series string length, and inverter type that will work well. Record highs and lows are not useful for system integration, nor are simple average temperatures. Fortunately, values in between the extreme and the average temperatures are very useful. Practices vary among designers,

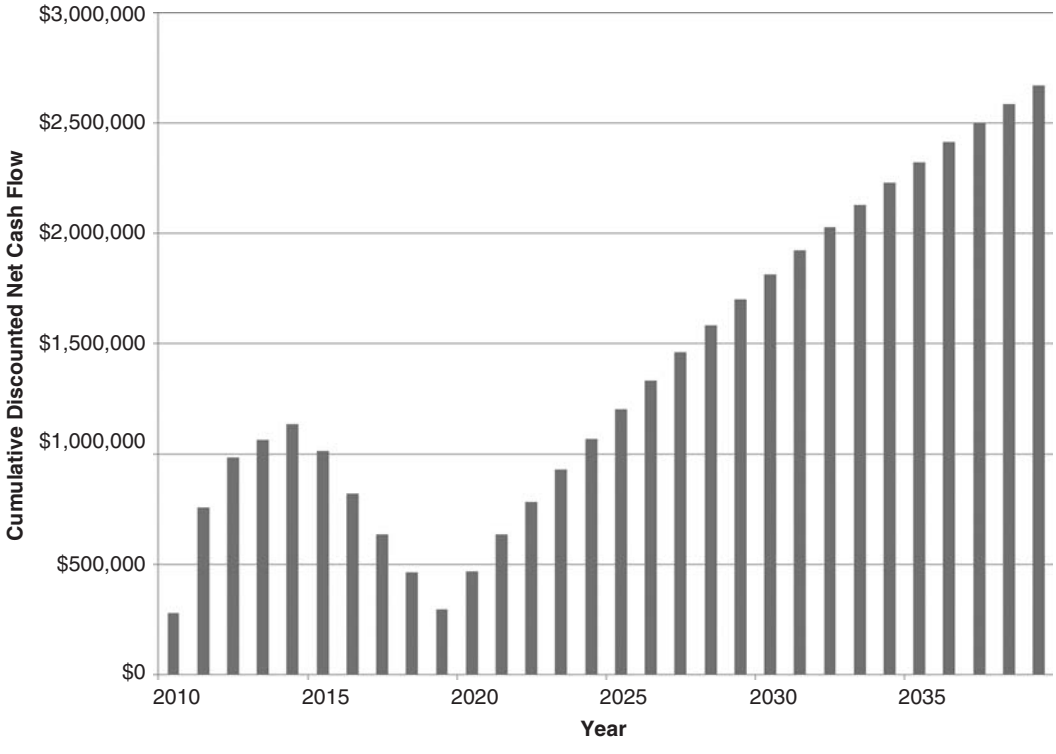


Figure 19.25 Clean Power Estimator cumulative discounted cash flow for 1.2 MW_P fixed tilt system in northern California

but it is common to select a low design temperature that has a statistically low chance of being exceeded in the winter months. Similarly, a high design temperature is usually chosen such that it has a small chance of being exceeded during the summer.[6]

In practice, choosing a value halfway between the record temperature and the average temperature for either the coldest or warmest months produces a result in good agreement with the ASHRAE seasonal 2% exceedance number. For example, ASHRAE's 2% (warmest summer month) value for Fresno is 40 °C, meaning the air temperature will exceed 40 °C, but only for 14 hours per month. While this special temperature is relatively obscure, Fresno's long-term data are readily available at several free internet sites such as www.weather.com. There, both its record high of 46 °C and its typical daily high for the warmest month of 36 °C are posted. The average of those two values is 41 °C, just 1 °C different from the ASHRAE value. A similar calculation scheme can be used to determine coldest design temperatures.

One additional step is needed to determine a design high temperature for system integration. That step involves adding a solar gain term to the ambient high temperature determined above. Depending on how well ventilated the array is, peak module temperatures will be at least 30 °C

[6] The American Society of Heating, Refrigeration, and Air Conditioning Engineers (ASHRAE) periodically publishes its Handbook of Fundamentals, with statistically based guidelines for design high and low temperatures for hundreds of US locations. This document is for sale, but a free resource on the internet is www.weather.com, which lists average and record high and low temperatures worldwide.

above ambient and can reach 50 °C above ambient. In the US, the NEC applies adders for sun-exposed wiring that range from as little as 14 °C for arrays raised 0.3 – 1.0 m above a roof surface, to as much as 33 °C for arrays closer than 1.3 cm from the roof.

With the operating extremes known, if an inverter and module are already selected, then the only step left in integrating the major components is to determine series string length. The string length, or number of modules in series, must be large enough to assure unlimited operation on the warmest days, yet small enough to avoid exceeding the equipment's voltage rating on the coldest days (typically 600 V_{DC} in the US and Canada and 1000 V_{DC} elsewhere). Sometimes the integration process requires changing the module type or inverter type if a suitable string length cannot be achieved.

With the major equipment selected, the remaining system integration tasks are comparatively simple: choosing BOS components that are rated for the intended power, current, and voltage. This includes components such as properly sized DC and AC disconnects, wiring, combiner boxes, fuses, and fittings.

String sizing determines the number of modules to be connected in series. This means that the voltages are summed and the current remains the same. This defines the maximum voltage of a system. As part of the system integration it becomes clear that a module is not merely selected by its peak power rating. The module's voltage impacts the system's maximum voltage. Output parameters: V_{OC}, I_{SC}, I_{mp}, P_{max} of each module also have temperature coefficients that correspond to their various properties, as explained in Chapter 18. The most significant, and typically the largest in magnitude, is the temperature coefficient of open-circuit voltage V_{OC}, which is always negative for PV modules. During the summer months, high temperatures inversely impact V_{OC} and may cause up to a 25% decrease in electrical output relative to nameplate power. This is the reason it is important for modules to be mounted in a manner where they can readily dissipate heat.

19.7.4 Intermittency

Solar energy does not always correspond well with time of use of electricity. This intermittency of resource occurs on a number of time scales:

1. Dynamically, including rapidly moving clouds
2. Daily, including sunrise to sunset and general weather variations
3. Seasonally, as the amount of daylight changes throughout the year

For a grid-connected system, the item of concern is the dynamic variations. Examples of this type of variation are rapidly moving clouds and crawling shadows on to the PV modules. This can have an impact on the electricity production on a time scale of minutes or seconds. These variations in resource can cause the power being delivered by the system to vary and cause an impact on the utility distribution AC voltage. For large grid-tied systems, a dynamic electrical system model of the PV plant including the inverter characteristics and the resource input is often made to simulate the impact of the PV system to the utility grid at the point it is connected to. These simulations are performed using software simulation tools such as: Positive Sequence Load Flow (PSLF) [41] and Power System Simulator for Electrical Transmission (PSS®E) [42] that are used in many parts of the world for grid interconnection studies.

For off-grid systems, the longer term daily and seasonal intermittency aspect of the solar resource impact the size of the storage and backup generation. These components need to be sized in a manner that includes the time when solar resource may not be available or is limited. This is illustrated in the off-grid example at the end of this chapter.

19.7.5 Material Failure

To minimize material failure the static and dynamic forces as well as the environmental factors that a PV system will be subjected to over the course of its life must be taken into consideration. The structure, trackers, foundations, and other components will undergo many thermal cycles, exposure to ultra violet irradiation, extreme temperatures, wet conditions, strong winds, snow, and maybe even seismic events. Excess stress and fatigue failure can cause irreparable damage. Therefore the system should be designed to prevent corrosion and withstand the maximum design loads [43].

19.7.6 Modeling

PV modeling is usually thought of as a means of forecasting annual energy, though modeling of cash flows and other quantities such as temperature, solar radiation, and current and voltage are often desired, too. Sophisticated software is now widely available to perform detailed hourly simulations of energy and numerous other quantities. A review of modeling principles and a sample listing of current software follows below. First, however, consider that simple – and equally accurate – annual energy modeling can be achieved without detailed software by multiplying just three previously defined terms:

$$\text{Annual kWh}_{AC} = (\text{System kW}_P) \times (\text{Reference yield, } Y_r) \times (\text{Performance ratio, PR})$$

Of course, these terms must each be well characterized in order to achieve accurate annual results, but in many cases, they are. The system size in kW_P is a straightforward product of the number of modules and the modules' nameplate rating. The reference yield (in units of hours per year of peak irradiation at $1\,kW/m^2$) varies widely around the globe, and is strongly dependent on array orientation, but databases of solar resource for various orientations and locations are widely available. The performance ratio (PR) tends to hover in the range of 0.65–0.75 range, [44] with a strong dependence on system reliability, overall component quality, and temperature. It is the end result of numerous DC to AC loss mechanisms and is normally treated as a dimensionless quantity, though it is actually a ratio of average electric output per unit of solar radiation input, in units of kWh_{AC}/kW_P.

In an earlier example in this chapter, a northern California climate with a reference yield of 2,000 peak sun hours per year and a PR of 0.70 was found to have an annual output of $1{,}400\,kWh_{AC}$ for a $1\,kW_P$ system, hence a final yield of $1{,}400\,kWh_{AC}/kW_P$.

The best modeling techniques will cover four common elements in detail: solar resource; system mechanical and electrical characteristics; load characteristics; and economic parameters. Not all commercial software covers each element, and many users and design situations don't require the ability to model all four elements. A truly comprehensive program would go beyond the basic four elements and account for items such as year-to-year weather variability, long-term degradation, component replacements, general O&M expenses, outages, and finally, salvage or disposal costs. No such program is thought to exist yet. The four major modeling elements and some examples of commercial software follow.

19.7.6.1 Key modeling element 1: solar resources and weather

Characterizing a location's climate typically requires at least 10 years of data, ideally in hourly format, but often available only as monthly totals or averages. The core items include (see Chapter 22 for more detailed modeling and discussion of some of these parameters):

- Global horizontal irradiation.
- Air temperature.

Several other parameters are often used in simulation programs, but are not as widely available:

- Diffuse horizontal irradiation.
- Direct beam irradiation.
- Wind speed.

The following additional terms are sometimes helpful, but are not often included in weather databases:

- Wind direction.
- Dust/soiling/snow patterns.
- Humidity.
- Turbidity.
- Optical depth.
- Cloud cover.

19.7.6.2 Key modeling element 2: PV system characteristics

Many mechanical and electrical properties are often needed to run sophisticated models, though at a minimum, the nominal DC capacity and orientation need to be specified. The most basic models characterize the module output as linearly increasing with intensity, and linearly decreasing with temperature according to a single negative TC. More sophisticated models represent the module with several device parameters including their individual temperature and intensity dependence. More generally, the following terms are commonly used:

- Module type, number, and module specifications of current, voltage, temperature dependencies, physical dimensions, and often, more detailed terms such as series and shunt resistance.
- Inverter type, number, and efficiency and voltage minimum and maximum limits.
- Wire resistance (as conveyed by wire type and length).
- Tilt angle or tracking axis characteristics (rotation limits, backtracking).
- Orientation.
- Ground cover ratio or similar array geometry specification.
- Horizon shade profile.
- Thermal dissipation properties (such as nominal operating cell temperature, NOCT, or equivalent terms describing fixed and wind-related heat transfer).
- Expected lifetime of system (strongly affects LCOE).

19.7.6.3 Key modeling element 3: load characteristics

Whether residential or commercial, facility loads are often highly variable from month to month and year to year, and accurate data on an hourly or quarter-hourly basis are seldom available. Nevertheless, sizing and economics are strongly sensitive to load magnitude and pattern, so modelers must often make coarse estimates based on customer class and a sampling of monthly utility bills.

- Hourly power demand profile.
- Monthly utility bill.

19.7.6.4 Key modeling element 4: economic factors

Economic terms can range from as little as the cost of utility energy to as many as two dozen terms that influence the lifecycle costs and benefits of the PV system. Such terms may typically include:

- System installed cost.
- Scheduled replacement cost (inverter every 10 years, batteries every 6 years).
- Financing terms: loan duration, down payment %, discount rate.
- Electricity cost (by time of use if applicable) and escalation rate.
- Taxes on PV system.
- Income tax structure for buyer.
- Incentives: rebates, tax credits, depreciation, renewable energy credits, and any similar subsidies.
- Maintenance cost.

19.7.6.5 Example commercial simulation programs

Table 19.3 is a sample list of software programs are in common use as of 2010. While a detailed comparison of the features of these programs is outside the scope of this publication, a basic

Table 19.3 Sample PV modeling software

Name	Type: simulation, estimation, or both Execution: web or local	Licensing	Treatment of loss mechanisms
PVSYST (University of Geneva, Switzerland)	Type: both Execution: local	1000 CHF single-user	Highly detailed, time-varying
PV*SOL (Valentin Software, Germany)	Type: both Execution: local	468 single-user	Less detailed, but time-varying
PV Design-Pro (Maui Solar Soft.)	Type: simulation Execution: local	$259 US single	Less detailed, but time-varying
PVWATTS (NREL)	Type: Simulation Execution: Web	Free	All as constants
RetScreen (Canada Department of Natural Resources)	Type: estimation Execution: local	Free	All as constants
Clean Power Estimator (Clean Power Research)	Type: estimation Execution: web	Free	All as constants
SAM: Solar Advisor Model (NREL)	Type: simulation Execution: local	Free	Less detailed, some as time-varying
PV f-Chart	Type: Estimation Execution: Local	$500 US single-user	Most as constants, some as time-varying

list of characteristics is shown. The list is not intended to be comprehensive, partly because of the impracticality of assembling such a list and partly because many are only used for specialty purposes, not for full energy simulation.[7] One broad distinction among solar modeling programs is whether they are simulation-based or estimation-based with respect to inputting weather data; another is whether the program is web-based or executed on an individual computer. Simulation programs rely on individual timestep records, typically hourly, to calculate accurate energy (and other parameter) profiles for each day. Hourly weather data are not widely available worldwide, so simulation programs have limited applicability. Estimation programs typically begin with widely available monthly-average data and use correlations to synthesize daily generation profiles. Some programs are flexible enough to use either type of weather data as input. Another key distinguishing characteristic is the manner in which loss mechanisms are accounted for. Some programs simply treat factors such as inverter efficiency, shading, dust, and wire losses as constants while others treat such factors more realistically as time-varying quantities. Some are free, some require licenses.

19.8 INSTALLATION

Proper installation is crucial to the success of a PV system. The system should be designed for ease of installation and good accessibility. Over time, improvements in PV components have helped to streamline the installation process. Two examples include the introduction of structural attachments that permit "top-down" module mounting from above the module and simple plug type module wiring connectors. Each installation crew member must be skilled at performing their required tasks. Wires should be installed so that they are not subject to scuffing or cutting. Safety must be maintained during the installation to prevent injuries to the crew and damage to the structure and equipment. If necessary the design documentation should be updated to reflect the as-built system. Once the system is functional, commissioning must be done to test the equipment and installation before commercial operation commences [45].

19.9 OPERATION AND MAINTENANCE/MONITORING

PV systems are designed and installed with the intention that they be fully functional for a minimum of 20–30 years. The following is a list of common scheduled tasks:

- Module cleaning to remove soiling.
- Inspection of wires and electrical connections.
- Verification of proper inverter operation.
- Inspection of mechanical mounting system.
- Replacement of broken or damaged modules.
- Vegetation control.

The recommended practices in the component manuals should be followed. Typically, residential scale systems do not schedule professional O&M visits, as the annual revenues are too small. The above list of O&M responsibilities are usually shouldered by the owner. Larger commercial and utility scale systems usually generate enough revenue to warrant having scheduled O&M performed by PV professionals.

O&M costs for 4.5–5.0 MW systems installed by two utilities in the hot and dry southwestern U.S. were monitored over a seven-year period. It was found that annual O&M cost, on a

[7] The PV Resources website at http://www.pvresources.com/en/software.php has a lengthy list of both comprehensive and specialty software.

dollar per kW_P per year basis, ranged from 0.1 to 0.5% of initial system cost, depending mostly on inverter repair/replacement cost [46, 47]. At a typical system cost of \$8000 per kW_P, and by choosing a middle value of the O&M range at 0.25%, the resulting annual O&M cost is about \$20 per kW_P/yr. Another way to express O&M is on an energy basis rather than on an installed capacity basis. A common energy basis rule of thumb for larger grid-tied PV systems is that the O&M costs run in the range of 1–3 cents per kWh of energy generated. This cost increases as the system ages and increased maintenance is required to ensure good performance. Off-grid systems are typically more expensive than grid-tied systems to operate and maintain on a relative basis due to the additional components such as batteries.

Monitoring the performance of a PV system is important both for recording the power output of the system and for determining the system health. Since problems can occur between scheduled maintenance intervals, it is important to minimize the period of time that the system is either down or not performing at its optimal level. Monitoring a system can be a valuable part of the O&M regimen. Monitoring is a two-fold process. The system's output and the local environmental conditions need to be recorded. This allows detection of system malfunctions as well as system under-performance. In some large systems it may make financial sense to have a regular on-site crew member to monitor the system and be able to fix problems as they arise [48].

Many residential and commercial installers now offer remote monitoring services via internet-based tools. These monitoring systems often interface with the inverter and provide key system condition measurements and operating parameters. An example of a monitoring system display is given in Figure 19.26.

By monitoring the performance of a large number of systems in a given geographic area, systems whose relative output has fallen below the average of the others can be detected, independent of daily sunlight and weather conditions.

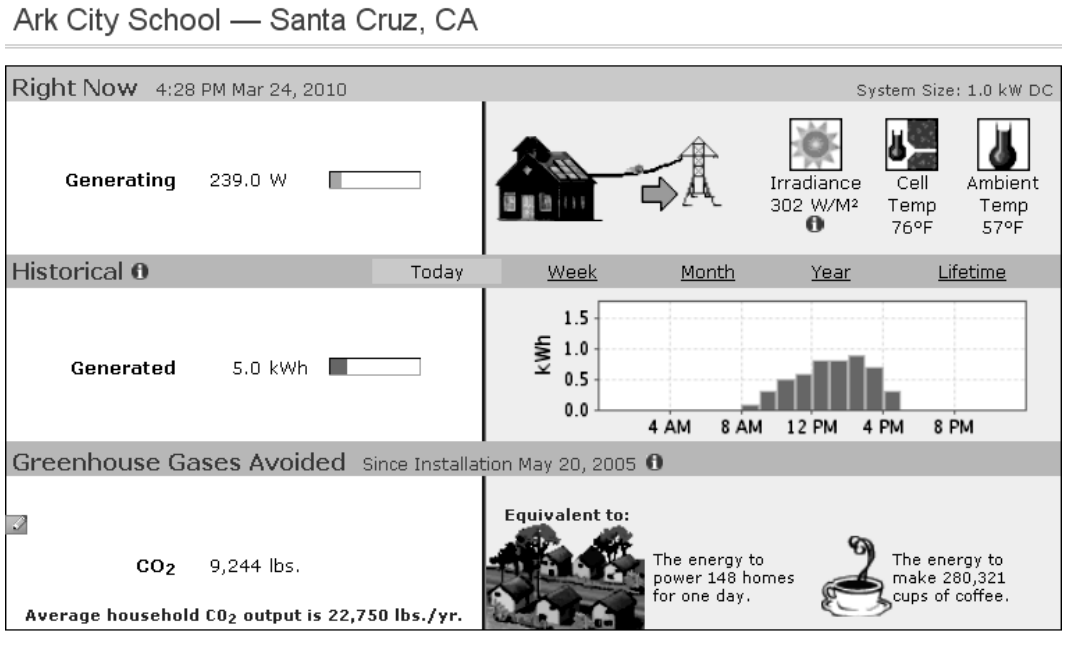


Figure 19.26 Monitoring display for 1 kW_P system [49]

19.10 REMOVAL, RECYCLING AND REMEDIATION

As part of looking at PV systems from cradle to grave, the removal of a PV system at the end of its usable life is something that must be taken into consideration both financially and environmentally. Part of PV's *raison d'etre* is motivated by the desire to provide electricity without harmful emissions. The ability to responsibly retire a system without causing any environmental harm should be incorporated into the system plan. With today's 20–30 year PV system operating lifetimes, the need to remove and recycle will fall on a future generation. There is an increasing awareness of the economic benefits of recycling and reusing the most valuable materials in the module (i.e. Al, Si wafers, In, Te, Ag, glass) as well as a responsibility to prevent heavy metals (Cd, Pb) from entering the environment. Groups in the US and Europe have been established to encourage life-cycle costing and materials recycling [50, 51]. Many manufacturers guarantee to collect their modules and much of the material can be salvaged or recycled. It is customary to expect that through recycling an economic value in the range of 10% of the initial system price could be recovered.

Chapter 1 provides information about the short energy payback times of 1–3 years and very high CO_2 reduction potential of PV, as well as the commitment to safe manufacturing and recycling within the PV community. An additional environmental benefit of PV systems is that they do not require significant amounts of water for operation, which is important considering that large scale systems are often installed in arid regions. This is an advantage over concentrating solar thermal plants which require substantial quantities of water.

19.11 EXAMPLES

19.11.1 Example Off-grid House/Cabin AC/DC/diesel/batteries

We will look at a rural household site located in the Pyrenees Mountains in Spain that receives, on average, 3.85 peak sun hours [kWh/m^2] per day at the horizontal and 3.2 peak sun hours [kWh/m^2] per day on the worst month with a corresponding optimum tilt angle of 55°. The average minimum temperature in the coldest month is 0 °C and the record low is – 9 °C, suggesting a design low temperature between these two extremes would be appropriate, or – 4 °C. The average high temperature in the warmest month is 30 °C and the record high is 40 °C, suggesting a design high temperature between these two extremes of 35 °C. We assume there are no significant losses due to shading or snow and that the PV modules will be periodically cleaned by the owner so there are no soiling losses. The needs of the house are 3,300 Wh/day on average all year around. Higher loads associated with more lighting needs during the winter months are offset by lower food refrigeration loads, giving a fair amount of daily load balance throughout the year. The quality requirements are that the service is to be provided at standard 230V–50 Hz AC electricity to all loads with an average energy daily availability of 3,300 Wh from the PV system for the worst month average climatic conditions. The rated maximum power to the loads is 3,446 W. The battery capacity must be sufficient for at least 3 days of storage at the average daily energy use. We say "at least" because the battery sizes are fixed, and we choose the next largest readily available battery size. The loads are served via a maximum 15 A circuit. The battery will be sized to provide three days of autonomy based on average daily usage. For electrical safety, a DC system voltage of 48 V will be used. Ample area on the roof and surrounding grounds are assumed to be available. *Estimate of the load*. The table of loads (Table 19.4) has been prepared based on the needs of the household and the following considerations: the main appliances are all high efficiency and the laundry machine is not heating water (cold wash or external heating only). All loads have an average daily utilization, except the washing machine that is used three times weekly. The deferrable loads, to be used only when there is surplus PV generation, are not considered for the daily energy calculation, but have

Table 19.4 Loads of the household

Main loads	Unit power (W)	Quantity	Total power (W)	Probable unit peak power (W)	Average use (hr/day)	Daily energy (W h/day)
Lamp	11	6	66	20	3	198
Lamp	20	4	80	30	2	160
Low-consumption TV	90	2	180	250	3	540
DVD	50	1	50	50	1	50
Portable Hi-Fi	35	1	35	35	2	70
Efficient refrigerator	120	1	120	500	Consumption based on data sheet	650
Efficient Deep Freezer	120	1	120	500	Consumption based on data sheet	650
Bi-thermal laundry washer	250		250	1200	400 Wh/wash, 3X/week	170
Kitchen appliances	80	3	240	160	0.5	120
Personal computer	205	1	205	300	2	410
Idle load of inverter	10	1			24	240
Deferrable loads						
Vacuum cleaner	1800	1	1,800	2500		
Water pump	300	1	300	600		
Total			3446	Maximum 2500		3,258

to be taken into account for the inverter power. From the table we have that the rounded off average demand is 3,300 Wh/day. The inverter has to be able to supply the total installed power (3,491 W) and respond to a surge of all the installed power plus the start up of the largest single appliance ($3{,}446 + 2{,}500 = 5{,}946$ W) and the base load is 100 W (an idealized load profile can be seen in Figure 19.27). The required inverter specifications are: rated power (30 min) $\geq$ 3,500 VA; rated surge power (5 s) $\geq$ 6,000 VA, and a desired efficiency at an operating power of 100 W $>$80%.

The battery useful capacity has to supply at least three days of demand. We chose a battery type that can handle a maximum depth of discharge of 80% at typical discharge rates of 120 hours and a voltage of 48 V_{DC}. This means that the nominal storage capacity has to be higher than 12,375 Wh since:

$$C_{120} = (3,300\,\text{Wh/day}) \times 3/0.8 \geq 12,375\,\text{Wh}$$

At a rated voltage of 48 V this means that the rated capacity is

$$C_{120} \geq 12,375\,\text{kWh}/48\,\text{V} = 258\,\text{Ah}$$

For this type of load profile, system orientation, and climate, prior experience suggests that approximately 60% of the required household energy will flow through the battery. Over a 10-year

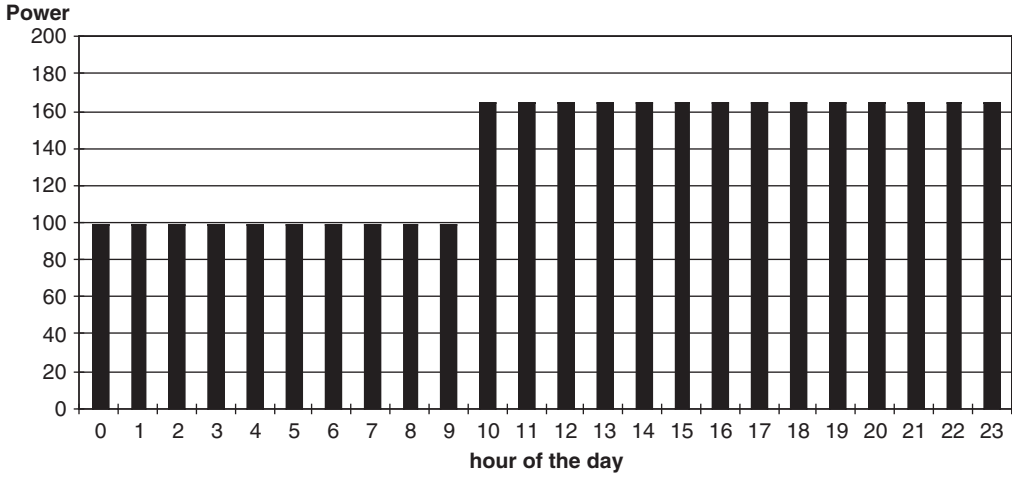


Figure 19.27 Idealized load profile (source TTA)

period the estimated battery throughput is

$$3,300\,\text{Wh/day} \times 0.6 \times 365 \times 10 = 7,227\,\text{kWh}$$

The selected batteries should have an expected lifetime useful capacity of at least 7,227 kWh. The useful energy per discharge cycle is

$$12,375\,\text{Wh} \times 0.8 = 9,900\,\text{Wh}$$

Therefore, the number of cycles over a ten-year period is approximately

$$72,27,000\,\text{Wh}/9,900\,\text{Wh/cycle} \approx 730\,\text{capacity cycles.}$$

Consider a battery whose specification sheet lists the following characteristics:

- Product: Varta 4 OPzS 200.
- 24 cells of 2 V each.
- C_{120} at 20 °C=300 Ah.
- 1500 cycles at 25 °C at 80% depth of discharge (DOD) [52].

The PV plant has to produce 3,300 Wh/day during the worst month; we will assume a performance ratio (PR) of 60%. The minimum nameplate PV capacity must be

$$P_{\text{pv}} = (3,300\,\text{Wh/day})/(0.6 \times 3.2\,\text{h/day}) \geq 1.7\,\text{kW}_\text{P}$$

An 85 W_P crystalline module with 36 cells and a V_{oc} of 21.5 V and an I_{sc} of 5.2 A will be used. The array will require series strings of 4 modules each, or 340 W_P per string. A minimum of 5 parallel strings will be needed to satisfy the minimum power requirement of 1.7 kW_P. The charge controller must be rated for a maximum voltage of

$$\begin{aligned} &\text{V}_{OC} \times \text{number of modules in series} \times \text{cold weather temperature factor} \\ &\quad = 21.5\,\text{V} \times 4 \times 1.2 = 103\,\text{V} \end{aligned}$$

Battery charge controllers (BCC) can be oversized in anticipation of future increases in array capacity or load. We will assume a 30% future expansion factor when sizing the charge controller. The minimum rated current for the BCC is

$$I_{BCC} > I_{SC} \times \text{number of strings in parallel} \times \text{future expansion factor} = 5.2 \times 5 \times 1.3 = 33.8\,\text{A}$$

The charge controller's voltage set points for equalization and floating will be adjusted to a customary 2.50 and 2.28 V per cell, respectively, and will be temperature compensated by the charge controller's built-in state of charge (SOC) algorithm. These values are typical thresholds for 2 V lead–acid cells. With 24 cells forming a 48 V battery, the set points will be $2.5 \times 24 = 60$ V for equalization and $2.28 \times 24 = 55$ V at 25 °C. The BCC will be set to disconnect the battery when its SOC reaches 20%, which corresponds to the maximum DOD allowed.

19.11.2 On-grid Example Systems

Residential and commercial grid-tied rooftop systems are now the most commercially significant types of PV systems. However, not all grid-tied systems, especially commercial ones, are designed for optimal economics. Numerous prominent public systems have been placed as environmental showcases or for good public relations. Unlike standalone systems, meeting the electric load of a grid-tied customer is not the governing design parameter. This does not mean that the fraction of building load supplied by PV is unimportant, it is simply not critical. Generally the limiting factors for utility-connected systems will either be the available area or budget.

19.11.3 On-grid House

Consider a system located near Sacramento, California (shown in Figure 19.28) that receives, on average, 5.5 peak sun hours [kWh/m^2] of plane of array (POA) radiation (south-facing 30° tilt) per day (also referred to in IEC 61724 as reference yield, Y_r)[8]. The average minimum temperature in the coldest month is 2 °C and the record low is −16 °C, suggesting a design low temperature in between these two extremes would be appropriate, or −7 °C. The average high temperature in the warmest month is 34 °C and the record high is 47 °C, suggesting a design high temperature in between these two extremes of 40.5 °C. Soiling in Sacramento will reduce un-cleaned system output by about 6% in a typical year. We assume there are no significant losses due to shading or snow. The available roof area is 24 m^2 of composition shingle roof in good condition. The system is mounted on the south-facing side at a 30° tilt. The modules are mounted parallel to the roof with a 9 cm offset using standard aluminum solar racking with top-down clamps.

The annual kWh usage for this household is equal to the California average of 10,000 kWh per year. In this climate, a well-designed PV system will generate about 1,400 kWh/kW$_P$ per year[9], so a 7 kW$_P$ system would be needed to fully offset this home's energy use. The designers found that a visually appealing layout of 14 Sharp 165 W modules totaling 2.3 kW$_P$ would fill 75% of the available rooftop area, leaving satisfactory room for service and emergency access,

[8] The National Renewable Energy Laboratory (NREL) publishes the National Solar Radiation Data Base where annual radiation for several orientations is available for 239 US locations.

[9] This yield is based on the product of the reference yield Y_r of 5.5 sun hours/day $\times$ 365 days/yr $\times$ 0.7 PR $\approx 1,400$ kWh/kW$_P$. The typical 0.7 performance ratio, PR, means 30% of the potential sunlight energy is lost in the conversion to ac electricity. About one-third of the 30% loss is attributable to temperature, one-third to inverter losses, and the remaining third due to the combined effects of wire resistance, mismatch, soiling, and similar small losses.

and integrate well with an SMA SWR1800U (1,800 W) inverter using two strings of seven modules in series. The proposed system can be expected to offset about one-third of the household's electricity needs.

The chosen inverter has a nominal power rating of 1,800 W_{AC} with a DC voltage range of 156–400 V_{DC}. The array-to-inverter loading ratio is 1.3, consistent with the manufacturer's recommendation and perfectly suitable in Europe, but slightly above what would typically be recommended in a sunny US climate. In sunny locations, an inverter loaded with this much DC capacity will probably sacrifice 1% or more of its potential generation due to high power clipping. Loading ratios of 1.2 or lower are more common in this climate since they virtually eliminate clipping losses.

For this array, the maximum DC voltage cannot exceed 400 V. This term is equal to the product of the number of modules in series (7), the module open-circuit voltage (44 V), and a cold weather voltage adjustment factor, which itself is equal to the temperature coefficient of open-circuit voltage (–0.4%/°C) multiplied by the temperature difference between the design minimum and the STC temperature of 25 °C. This string size satisfies the 400 V limit.

Under the warmest conditions this array will operate at a temperature 17 °C above the design maximum ambient temperature, which was estimated to be 40.5 °C. At 57.5 °C, the string voltage will be calculated in a manner similar to the open-circuit voltage equation above, with three substitutions. This string length satisfies the 156 V minimum voltage for the inverter.

With just two strings, no DC fusing is required for this system, as is usually the case for so-called string inverters in residential applications. The short-circuit current (I_{SC}) of the module, however, is needed for wire sizing. The module I_{SC} is 5.3 A. US designers following the NEC must satisfy two wire sizing criteria. The first requires applying a 1.56 multiplier to obtain the minimum basic ampacity rating at 30 °C. Here, the minimum is 8.3 A. The second criteria is that the selected wire must also be capable of carrying 125% of I_{SC} (6.6 A) under the conditions of use, which typically include ampacity de-ratings for temperature and sometimes for multiple (>3) conductors in one conduit. At 57.5 °C the temperature-related ampacity adjustment is 0.71, such that this second criteria will require a wire with a minimum basic rating at 30 °C of 6.6/0.71, or 9.3 A. Routing the DC wire to the inverter can be done with sunlight-resistant wire in free air, but is more commonly done by quickly transitioning the module "pigtail" wires in a junction box to a less expensive "building" wire in conduit.

A typical wire size and type that might be selected for this application would be #10 AWG THWN-2 (13.6 mm^2). This is a common nylon jacketed building wire, with the "-2" suffix indicating the ampacity is valid under wet or dry conditions up to 90 °C. This wire has a basic 30 °C rating of 40 A, de-rated by 0.71 at 57.5 °C for this exposed rooftop location. We apply an additional condition of use de-rating of 20% for having four current-carrying conductors in one conduit (2 strings × 2 conductors per string)[10]. The de-rated ampacity of the #10AWG wire is 40 × 0.71 × 0.8 = 23 A, well above the required 6.6 A target. A code compliant conduit with four current-carrying conductors plus a same-sized ground cannot fill more than 40% of the conduit's internal area. This wire has an area of 0.0211 in^2 (13.6 mm^2) including its nylon jacket. The minimum internal area, then, is 5 wires × 0.0211/0.4 = 0.26 in^2 (170 mm^2). A common US trade size conduit is ½ inch (1.25 cm) EMT (electrical metallic tubing), with an internal area of 0.3 in^2. This size would be suitable, since the internal area is slightly greater than the minimum 0.26 in [2].

Most residential systems will have similarly oversized DC wire, so the resistance losses tend to be very small. #10AWG wire has a resistance of 0.004 Ω/m, so the round-trip voltage drop

[10] According to NEC Table 310.15(B)(2)(a)

Figure 19.28 2.3 kW_P Berman family residential grid-tied system, near Sacramento, California

for a typical 25 m one-way run, with a nominal string voltage of 252 V_{DC}, and a nominal string current of 4.6 A, would only be:

$$0.004\,\Omega/\text{m} \cdot 25\,\text{m} \cdot 2 \cdot 4.6\,\text{A}/252\,V_{DC} = 0.4\%$$

The DC conduit run terminates at a DC switch next to the inverter. The smallest standard switch rated for this voltage class (<600 V_{DC}) is rated for 18 A_{DC}, well above the 6.6 A_{DC} design current.

The inverter's output voltage is 120 V_{AC} and it has a15 A_{AC} rating. This requires AC switchgear, wiring, and over-current protection at 125% of 15 A, or 19 A. The next largest standard sizes are a 30 A unfused AC disconnect and a 20 A circuit breaker mounted within the existing electrical supply panel as the point of common coupling (POCC). As before, a common #10AWG building wire is used to complete the AC design [53, 54]. As with the DC cabling, a suitably sized AC conduit would also be ½ inch (1.25 cm) EMT, with even more unused fill area than the DC run, since the AC conduit run would only contain three wires (one hot and one neutral and one ground), instead of the five conductors entering the DC disconnect switch. The AC run from the inverter to the supply panel is typically shorter than 10 meters, with a minor to negligible wire loss.

19.11.4 Commercial Roof

Commercial rooftops are usually of the low-slope variety, and tend to have relatively low dead load limits, although this is not universally true. Because of this, ballasted PV systems also tend to be placed at shallow tilts to reduce the amount of weight required. Optimally tilted systems tend to be attached through the roof membrane to the underlying building structure. One other general characteristic of commercial roofs is that PV system ground cover ratios tend to be smaller than

on residential or on ground-mounted systems because of the numerous obstructions presented by elevator shafts, parapet walls, skylights, HVAC equipment, and service corridors.

Consider a system located atop a large retail store in Manahawkin, NJ. A similar array is shown in Figure 19.29. Using the same method as described above, the minimal expected temperature for the area is −14 °C while the maximum ambient temperature is 35.5 °C. It is anticipated the system will lose about 2% of its potential annual generation due to summer dust and winter snow. The long-term average GHI for nearby Atlantic City is 1,486 kWh/m^2, with an average air temperature of 12 °C.

The array consists of 390 Kyocera modules, with a nameplate rating of 205 W each, for a nominal DC capacity of 80 kW$_p$. A single 75 kW Satcon inverter with a DC operating range of 315–480 V and a maximum DC input voltage of 600 V is mounted at street level. The inverter is lightly loaded with a DC to AC capacity ratio of 1.07 (80 kW$_P$/75 kW). For this climate, module type, and inverter, an optimal series string length of 15 modules has been determined.

The modules are ballast-mounted at a common fixed 10° tilt, facing southeast to conform to the orientation of the building. Rows are spaced to achieve a 0.67 ground cover ratio, with a corresponding setback ratio of 2.9. The modules are in landscape orientation to make best use of their bypass diodes and thereby minimize row-to-row shading, which for this geometry was modeled to have an annual impact of a 1% reduction in generation.

The strings are collected and distributed among four 8-pole DC combiner boxes with 15 A string fuses. The 84 A (I_{sc} × 1.25 × 8 poles) PV output circuits from the combiners terminate at two 100 A 3-pole unfused 600 V DC disconnect switches. Wires exposed to sunlight will have lower ampacities because of their increased temperature. The NEC stipulates a 17 °C temperature rise above design ambient for most sun-exposed rooftop wiring situations. The design wire temperature is therefore adjusted to 52.5 °C. For #2AWG THWN-2 wire (≈9.5 mm diameter), the NEC stipulates

Figure 19.29 Commercial ballasted rooftop system, New Jersey. SunLink's low-profile racking relies on a combination of module weight, racking weight, and concealed ballast blocks to anchor the array (Courtesy BEW Engineering)

an ampacity de-rate of 0.76 for temperatures between 5 and 55 °C. This type and size of wire has a basic ampacity rating of 130 A at 30 °C, reduced to 99 A at 53 °C. This is comfortably above the minimum 84 A necessary under the conditions of use.[11] The smallest allowable grounding conductor for these segments would need to have a basic rating at 30 °C of at least 84 A, which may be satisfied with a #4AWG wire. The smallest allowable EMT conduit, with each PV output circuit in its own conduit, is 1 inch trade size (equivalent to metric size 27).

On the AC side the system is connected to a customer-owned revenue-grade energy meter. The POCC is a line tap on the customer side of the facility's 400 A main supply panel. The 75 kW inverter has a rated current of 90 A, requiring a 200 A AC disconnect with 125 A fuses. A code-compliant design consisting of three #3AWG THWN-2 phase wires, a #3AWG neutral wire, and a #3AWG ground wire are run in a 2.5 inch EMT conduit (metric designator 63).[12] The system also contains a remote data-logging measurement system with weather instrumentation, inverter and meter communication, and a wireless communication data transmission system in order to monitor the performance of the system.

The expected final yield Y_f is 1,193 kWh/kW$_P$, the PR is expected to be 0.76, and the total annual generation is initially expected to be 95 MWh. In the long term, annual generation is expected to decline at a rate of about 0.5–1.0% [55, 56].

19.11.5 Utility-scale Ground-mounted Tracking

Large-scale PV systems, some utility-owned and most privately-owned, now dominate the PV industry, with over 20 countries hosting systems greater than 1 MW through 2009 [57]. Most of these are ground-mounted, and many employ tracking mechanisms to increase annual generation and broaden the daily power output curve. A good example is the 14 MW$_P$ ground-mounted, single-axis tracking system located at the Nellis Air Force Base near Las Vegas, Nevada shown in Figure 19.30. Installed in 2007, the Nellis system consists of over 70,000 modules from four different manufacturers. Most of the modules are mounted on N-S single-axis trackers tilted at 20°, but about 14% of the array capacity is mounted on horizontal N–S single-axis trackers. The 5,000 tilted-tracker segments are ballasted to the ground with a trio of large concrete bases; the remaining horizontal trackers are about 50 m long and employ a series of poured concrete foundations at 5 m intervals. There are 54 inverters distributed throughout the 0.6 km^2 site, 52 of which are 250 kW, with two more 100 kW units making up the total.

The long-term average GHI for Las Vegas is 2,078 kWh/m^2, with an average air temperature of 20 °C. The design cold and warm temperatures, using the same calculation as the previous two examples, are – 6 and 44 °C, respectively. While all of the DC source circuit wiring is exposed to sunlight and subject to a significant temperature de-rating, all of the DC wiring from the combiner boxes to the inverters is routed underground, as is all of the AC wiring from the inverters all the way up to the 12 kV overhead utility distribution lines, so each of the underground segments is not subject to temperature-related ampacity de-ratings.

The inverters are loaded lightly in terms of DC capacity. While each inverter is not identically loaded, the loadings are similar and average about 263 kW$_P$/250 kW, or a loading ratio of 1.05. Most of the field is populated at a GCR of 0.22, with the horizontal tracker section employing a GCR of 0.35. Backtracking is used throughout, with a rotation range of ±40°. While this geometry results in about a 9% loss relative to the theoretical radiation on an ideal tracker with no range limits or shading interference, the losses would be worse for a tracker that did not backtrack.

[11] According to 2008 NEC Table 310.16.
[12] According to 2008 NEC section 250.122 for grounding and Table 9 for conduit sizing.

Figure 19.30 Nellis Air Force Base 14 MW_P (0.6 km^2 [2]) PV system dedication, December 2007 (Courtesy Tim Townsend, BEW Engineering)

Considering these tracker and field geometry properties, the blended average plane of array (POA) radiation is 2,900 kWh/m^2, a boost of nearly 40% relative to a fixed horizontal surface. This is also about 25% more output than would be expected annually from an optimally-tilted south-facing fixed array [58]. The desert climate in Las Vegas receives less than 100 mm of rain per year, so soiling losses were estimated to be about 5% per year prior to construction; since then, anecdotal observations have suggested the dust losses are perhaps 50% less than anticipated.

The expected final yield Y_f is 2,240 kWh/kW_P, the PR is expected to be 0.77, and the total annual generation is initially expected to be 31 GWh_{AC}. In the long term, annual generation is expected to decline at a rate of about 0.5–1.0%. Annual maintenance costs averaged over the project's life are estimated to be about 1.5 cents/kWh or equivalently, \$32/kW_P/yr. In terms of operation and maintenance planning, this plant is large enough that it sets a US precedent regarding staffing; this is the first PV plant that justifies having a full-time employee on site.

Detailed analyses of performance, cost, reliability and maintenance of utility scale tracking power plants have recently been published for systems in the US, [59, 60] Spain, [61] and Germany [62].

REFERENCES

1. California Energy Commission. 2009. *Go Solar California: A Short History of Solar Energy and Solar Energy in California*. [Online] Available at: http://www.gosolarcalifornia.org/solar101/history.html [Accessed 11 February 2010].
2. Massey, D., The Porticus Center. 2006. *Bell Labs: The Solar Battery (Photovoltaics)*. [Online] Available at: http://www.porticus.org/bell/belllabs_photovoltaics.html [Accessed 3 March 2010].

3. Lenardic D, PV Resources. 2010. *Large-scale Photovoltaic Power Plants*. [Online] (Updated 22 March 2010) Available at: http://www.pvresources.com/en/top50pv.php [Accessed 27 February 2010].
4. Roth W, 2007. *German-Turkish TU9-Workshop on Sustainable Energy: Photovoltaics - Current Situation and Prospects (Market Overview)*. Gebze, Turkey November 2008. Fraunhofer-Institut für Solare Energiesysteme ISE: Freiburg, Germany.
5. Bradford T, Prometheus Institute. 2009. *State of the Solar Industry- Monthly Webinar*. [Online] http://www.solarelectricpower.org/events/webinars.aspx [Accessed 2 June 2009].
6. Englander D, 2009. 2009 Global PV Demand Analysis and Forecast. *Greentech Media*, February.
7. Moore L, Post H, Hansen T, Mysak T, 2006. Five Years of Operating Experience at the Springerville PV Generating Plant. *Sandia National Laboratories*, Albuquerque, NM.
8. International Electrotechnical Commission (IEC), Photovoltaic System Performance Monitoring-Guidelines for Measurement, Data Exchange, and Analysis, IEC Standard 61724, Geneva, Switzerland, 1998.
9. ASTM Std E1036.
10. Kimber A, 2009. No ASTM standard currently exists for rating flat-plate PV systems, though a proposed standard was submitted in the fall of 2009. This effort stems from a paper Improved Test Method to Verify the Power Rating of a Photovoltaic (PV) Project., Presented at the *34th IEEE PV Conference*, Philadelphia, PA, June 2009. In it, Kimber *et al*. suggest options for field data collection and modeling to determine AC ratings. IEC 61724 covers fielded system data measurement and analysis techniques.
11. Kevin L, 2001. Test Method for Photovoltaic Module Ratings. *Florida Solar Energy Center*. [Online] Available at: http://fsec.ucf.edu/ [Accessed 5 April 2010].
12. California Energy Commission. 2009. *Go Solar California: CSI guidelines*. [Online] Available at www.GoSolarCalifornia.ca.gov [Accessed 17 February 2010].
13. Hester SL, 1988. PVUSA: Lessons Learned from Startup to Early Operation. *IEEE*, [Online], pp 937–43, Available at: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00111757. [Accessed 22 February 2010].
14. Stultz JW, Wen LC, 1977. Thermal Performance Testing and Analysis of Photovoltaic Modules in Natural Sunlight. *Jet Propulsion Laboratory*, pp 5101–31.
15. See endnote 4
16. California Energy Commission. 2009. *Go Solar California: Inverter Efficiency Curve*. [Online] Available at www.gosolarcalifornia.org/equipment/inverter.php [Accessed 7 March 2010].
17. Dunlop JP, 2010. *Photovoltaic Systems*. 2nd edn, Illinois: American Technical Publishers, Inc.
18. Solar Buzz. 2009. *Solar Buss: Solar Industry Codes & Certifications*. [Online] Available at http://www.solarbuzz.com/ProductCertifications.htm [Accessed 22 February 2010].
19. Solar Pathfinder. 2009. *Solar Pathfinder*. [Online] Available at http://www.solarpathfinder.com/ [Accessed 3 February 2010].
20. The Solar Design Company. 2010. *Solar Site Selector: A sun-path indicator for solar engineers*. [Online] Available at http://www.solardesign.co.uk/pw-solar.php [Accessed 3 February 2010].
21. Wiley Electronics, LLC. 2010. *Acme Solar Site Evaluation Tool (ASSET)*. [Online] Available at http://www.we-llc.com/ASSET.html [Accessed 3 February 2010].
22. Solmetric. 2010. *Solmetric Suneye*. [Online] Available at http://www.solmetric.com/ [Accessed 3 February 2010].
23. Gordon JM, Wenger HJ, 1991. Central-station solar photovoltaic systems: Field layout, tracker, and array geometry sensitivity studies. *Solar Energy*, **4**, 211–17.
24. National Renewable Energy Laboratory. 2010. *PVWATTS simulation software*. [Online] Available at http://www.nrel.gov/rredc/pvwatts/changing_parameters.html [Accessed 6 April 2010]. http://www.nrel.gov/rredc/pvwatts/changing_parameters.html.

25. Townsend TU, Hutchinson PA, 2000. Soiling Analyses at PVUSA. *Solar 2000: Proceedings of the 2000 Annual Conference of the American Solar Energy Society*. Madison, WI, 2000.
26. Becker G *et al*., 2007. *An Approach to the Impact of Snow on the Yield of Grid Connected PV Systems*. Bavarian Association for the Promotion of Solar Energy.
27. Damiani B *et al*., 2003. Light Induced Degradation in Promising Multi-crystalline Silicon Materials for Solar Cell Fabrication, *3rd World Conference on Photovoltaic Energy Conversion*, Plenary, pp 927–30. Osaka: Japan.
28. Kaushikaa ND, Raib AK, 2007. An investigation of mismatch losses in solar photovoltaic cell networks. *Energy*, **32**, 755–59.
29. Dhere NG, Pethe SA, Kaul A, 2009. Outdoor monitoring and high voltage bias testing of PV modules as necessary test for assuring long term reliability. *Reliability of Photovoltaic Cells, Modules, Components, and Systems II, Proc. SPIE* Vol. 7412. Florida Solar Energy Center, University of Central Florida, Florida.
30. Kurtz S, Granatab J, Quintanab M, 2009. Photovoltaic-Reliability R&D toward a Solar-Powered World. *Reliability of Photovoltaic Cells, Modules, Components, and Systems II, Proc. SPIE* Vol. 7412. National Renewable Energy Laboratory, Golden, Colorado & Sandia National Laboratories, Albuquerque, New Mexico.
31. Osterwald C, Adelstein J, del Cueto J, Kroposki B, Moriarity T, 2006. *Comparison of degradation rates of individual modules held at Max Power*, *Proc. 4th WCPEC*, pp 2085–8.
32. BEW Engineering, 2007. *Photovoltaic System Installation Engineering Review*. San Ramon, California.
33. Davis K, Moaveni H, 2009. Effects of module performance and long-term degradation on economics and energy payback: Case study of two different photovoltaic technologies. *Reliability of Photovoltaic Cells, Modules, Components, and Systems II, Proc. SPIE* Vol. 7412. Florida Solar Energy Center, University of Central Florida, Florida.
34. Parrettaa A, Bombace M, Graditia G, Schioppo R, 2004. Optical degradation of long-term, field-aged c-Si Photovoltaic modules. *Solar Energy Materials and Solar Cells*, **86**, 349–354.
35. King DL, Quintana MA, Kratochvil JA, Ellibee DE, Hansen BR, 2000. Photovoltaic module performance and durability following long-term field exposure. *Progress in Photovoltaics: Research Applications*, **8**, 241–256.
36. Meyer EL, van, Dyk E, 2004. Assessing the Reliability and Degradation of Photovoltaic Module Performance Parameters. *IEEE Transactions on Reliability*, **53** (1), 83–92.
37. Macêdo WN, Zilles R, 2007. Operational Results of Grid-Connected Photovoltaic System with Different Inverter's Sizing Factors (ISF). *Progress in Photovoltaics: Research and Applications*, **15**, 337–352.
38. Mayfield R, 2009. Flat Roof Mounting Systems. *Solar Pro magazine*, February/March 2009, pp 46–61.
39. National Renewable Energy Laboratory. 2010. *Solar Advisor Model (SAM)*. [Online] Available at: https://www.nrel.gov/analysis/sam/background.html [Accessed 15 March 2010].
40. Go Solar California. 2010. *Clean Power Estimator*. [Online] Available at: http://gosolarcalifornia.cleanpowerestimator.com/gosolarcalifornia.htm [Accessed 22 February 2010].
41. GE Energy. 2010. *PSLF Software*. [Online] Available at: http://www.gepower.com/prod_serv/products/utility_software/en/ge_pslf/index.htm [Accessed 10 March 2010].
42. Siemens. 2010. *PSS®E: Transmission System Analysis and Planning*. [Online] Available at: http://www.energy.siemens.com/hq/en/services/power-transmission-distribution/power-technologies-international/software-solutions/pss-e.htm [Accessed 10 March 2010].
43. Messenger RA, Ventre J, 2004. *Photovoltaic Systems Engineering*. 2nd edn, Boca Raton, Florida: CRC Press.
44. International Energy Agency, 2007. *Photovoltaic System Programme: Cost and Performance Trends in Grid-Connected Photovoltaic Systems and Case Studies, IEA-PVPS T2-06:2007*, Erlenbach, Switzerland: International Energy Agency.

45. See endnote 43
46. Moore L. Post H, Hayden H, Canada S, Narang D, 2005. Photovoltaic Power Plant Experience at Arizona Public Service: A 5-year Assessment. *Progress in Photovoltaics: Research and Applications*, **13**, 353–363.
47. Moore L, Post H, 2008. Five Years of Operating Experience at a Large, Utility-scale Photovoltaic Generating Plant *Progress in Photovoltaics: Research and Applications*, **16**, 249–259.
48. See endnote 17
49. Fat Spaniel Technologies. 2010. *Compare and Contrast Historical Performance Data of Schools*. [Online] Available at: http://view2.fatspaniel.net/SSH/MainView.jsp?school = urban [Accessed 24 March 2010].
50. Colombia University. 2010. *Center for Life Cycle Analysis*. [Online] Available at: www.clca.columbia.edu [Accessed 22 March 2010].
51. PV Cycle. 2010. *PV cycles Association*. [Online] Available at: www.pvcycle.org [Accesses 22 March 2010].
52. VHB Industrial Batteries Ltd. 2001. *VARTA Specifications Sheet: OPzS range 4 OPzS 200 . . . 24 OPzS 3000*. Ontario: VHB Industrial Batteries Ltd.
53. See endnote 17
54. Masters GM, 2004. *Renewable and Efficient Electric Power Systems*. 4th edn, John Wiley and Sons, Inc
55. See endnote 54
56. Wiles J, 2007. *Photovoltiac Power Systems and the 2005 National Electrical Code: Suggested Practices*. 2nd edn, Las Cruces, New Mexico: Southwest Technology Development Institute, New Mexico State University.
57. See endnote 3
58. See endnote 23
59. Moore L, Post H, 2008. *Progress in Photovoltaics*, **16**, 249–259.
60. Moore L, Post H, Hayden H, Canada S, Narang D, 2005. *Progress in Photovoltaics*, **13**, 353–363.
61. Garcia M, Vera J, Marrayo L, Lorenzo E, Perez M, 2009. Solar-tracking PV Plants in Navarra: A 10 MW Assessment. *Progress in Photovoltaics: Research and Applications*, **17**, 337–346.
62. Rindelhardt, U., Bodach, M., 2007. Operational Experiences With Megawatt PV Plants in Central Germany. *Proc 22nd European PVSEC*, pp 2952–2955.

20

Electrochemical Storage for Photovoltaics

Dirk Uwe Sauer

Fraunhofer Institute for Solar Energy Systems ISE, Freiburg, Germany

20.1 INTRODUCTION

The availability of solar energy is not only different with respect to the annual yield, but varies within the seasons of the year, during day and night and from day to day due to the weather conditions. So do the electrical loads. To balance the differing time patterns of loads and solar energy production, energy storage must be included in almost all autonomous power supply systems. Thirty percent or even more of the lifetime costs of autonomous power supply systems based on solar energy may be attributed to the storage. Only very few autonomous photovoltaic (PV) power supply systems have no battery storage system. These are PV or wind-pumping systems where the difference between water demand and energy supply (solar radiation) is levelled out by means of a water storage tank.

PV-powered autonomous power supplies power the complete range of appliances from very small appliances (watches, calculators) with a few milliwatt power requirements to large village power supply systems with about 10 kW power requirements. This chapter is dedicated to applications of PV generators in the range of approximately 10 W to 10 kW. This excludes very small appliances where different storage concepts are used.

There are a number of technical solutions to the problem of energy storage. Storing energy in the electric field of a capacitor is a solution for storage times ranging from microseconds up to 10 s. Storing energy in the magnetic field of a coil is a technology that has been under development for many years (supraconducting coils). It has, however, until today, not resulted in a product that has entered the commercial market.

Medium- and high-speed flywheels are being operated in small numbers in grid support, uninterruptible power supplies and in underground railways or buses to overtake energy from regenerative breaking and to support acceleration. However, storage times are in the range of seconds. Compressed air storage in the MW h range, which after decades of development has not

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

reached the market, and pumped water storage are the other technologies that store the energy mechanically. Pumped water storage shows a pronounced economy of scale, which means that larger systems have lower specific investment costs and better storage efficiencies. This is the reason pumped storage has its application in small grids that supply power in megawatt hours per day rather than in the domain of energy systems based on photovoltaics, which are and will be designed to deliver some 10 kilowatt hours per day.

The most promising storage technologies today for the application of which we are discussing here, are electrochemical systems [1]. Therefore, this chapter is dedicated to electrochemical storage systems only.

Although a variety of storage technologies are under development, for the systems focused in this chapter the lead–acid battery still is, and will be for some years to come, the workhorse for autonomous power supply systems. The ageing effects that limit battery lifetimes are treated in detail. Optimised control strategies allow increasing the battery lifetime and a considerable reduction of the overall system costs.

For storing large energy quantities with low power requirements, electrochemical storage systems with separate storage and power conversion units are under development. These are namely hydrogen storage systems with an electrolyser and a fuel cell as converters and redox-battery systems. The latter use charged ions from metal salts dissolved in liquids as the storage medium and a converter unit quite similar to a fuel cell. These systems are getting more and more interesting as seasonal storage systems and for balancing power generation and power demand in grids with a high penetration of renewable power sources (mainly wind and sun).

There are numerous requirements for storage systems in autonomous power supply systems. Their importance in different applications is different; some of them are in contradiction to each other and therefore they cannot be fulfilled at the same time. Table 20.1 gives an overview of the most important requirements of batteries in autonomous power supply systems.

System designs for autonomous power supply systems should track the properties and the requirements of the storage system from the very beginning. Planning of a system and later on just adding the storage will neglect the numerous interactions between the storage, its peripherals and its operation strategy and the overall system design and control. Therefore, only an integrated

Table 20.1 Requirements to electrical storage systems in autonomous power supply systems (the order of appearance does not imply any weighting of their importance)

- High energy efficiency
- Long lifetime (years)
- Long lifetime in terms of capacity throughput
- Low costs
- Good charge efficiency even at very low currents
- Low self-discharge rate
- Low maintenance requirements
- High availability worldwide
- High power availability
- Easy estimation of state of charge and state of health
- Low exposure
- Easy recycling
- Low toxicity of materials
- Fail-safe behaviour at overcharging or deep discharge
- Easy extendibility in voltage and capacity through series and parallel connection
- Low voltage gap between charging and discharging (allows direct connection of loads to the battery)
- Fast charging ability
- No memory effect
- Low explosive potential
- High reliability in operation: high MTBF

Note: MTBF: mean time between failure

planning of the system allows to use all synergies and to design systems that can be operated during their lifetime at minimum costs.

20.2 GENERAL CONCEPT OF ELECTROCHEMICAL BATTERIES

20.2.1 Fundamentals of Electrochemical Cells[1]

20.2.1.1 The equilibrium potential

The basic element of each battery is the electrochemical cell. A positive and a negative electrode are immersed in an electrolyte. The reactive substances (the active materials) are stored in the electrodes.[2] Chemical and electro chemical reactions, the electrode reactions, occur at both electrodes, which release or absorb electrons according to

$$S(N)_{\text{red}} \rightarrow S(N)_{\text{ox}} + n \cdot \text{e}^- \quad \text{or} \quad S(P)_{\text{ox}} + n \cdot \text{e}^- \rightarrow S(P)_{\text{red}}$$

where N and P indicate negative and positive electrodes, S_{red} and S_{ox} indicate the reduced and oxidised states of the chemical compounds that react and n is the number of electrons involved in the process. The possibility of splitting up the cell reaction into two separate electrode reactions is a decisive prerequisite for the realisation of any electrochemical cell. Only then can the electron exchange connected with the electrode reactions be collected as a current that flows through the consumer (or the charging device), and the energy input or output connected with the chemical reaction be converted into electrical energy. Otherwise, the reaction would occur merely as a chemical reaction. The electrical charge would be exchanged directly between the reacting substances and the released energy would be converted predominantly into heat and to some extent into volume energy.

The electrochemical storage system is based on the conversion of chemical energy into electrical energy and vice versa. The amount of energy that can be stored in a cell is determined by the different energy content of chemical substances that represent the charged and discharged states. Consequently, the characteristic parameters of the system are determined by a number of electrochemical reactions and the energetic changes connected with these reactions. In total, these reactions result in the cell reactions that characterise the battery system.

The laws of thermodynamics generally apply to the equilibrium, which means that all reactions are balanced. In the electrochemical cell, these data can only be measured when no current flows through the cell or the electrodes. On account of this balance, the thermodynamic parameters do not depend on the reaction path; they depend only on the difference between the final and initial components of the electrochemical reaction.

Because of the equilibrium conditions, the laws of thermodynamics describe the possible upper limit of the performance data. As soon as the current flows through the cell, energy losses occur due to kinetic restrictions and ohmic resistance.

---

[1] This section is based on Chapters 2.1, 2.2 and 2.3 from the book of D. Berndt *Maintenance-free Bbatteries* [3] which can be highly recommended for a deeper insight into applied battery technology.

[2] Note: the wording used herein is characteristic for classical electrochemical secondary accumulators with solid active masses and a liquid electrolyte. In fact, batteries with solid electrolytes and liquid active masses exist as well. Examples are redox-flow batteries (see Section 20.5.1) or the NaS batteries (liquid Na and S as active masses, solid oxide ceramic as electrolyte).

The energy exchange, connected with electrochemical reactions, is described by the following thermodynamic parameters. As far as these parameters concern chemical or electrochemical reactions, they actually describe the difference between the parameters before the reaction started and after it was completed. Therefore, they are expressed as the difference between the initial and the final state of the reaction.

- *Enthalpy of the reaction* ΔH, which describes the amount of energy released or absorbed. It is derived from the energy content of the chemical compounds H.
- *Free enthalpy of the reaction* ΔG (also called the Gibbs free energy), which represents the (maximum) amount of chemical energy that can be converted into electrical energy, and vice versa.
- *Entropy* ΔS, which characterises the energy loss or gain connected with the chemical or electrochemical process. The product $T\Delta S$ represents the heat exchange with the surroundings when the process occurs reversibly. This is synonymous with minimal heat loss or gain of the system, which is true only when no current flows through the battery.

T is the absolute temperature. The most important relation among these parameters is given by the following formulae:

$$\Delta G = \Delta H - T \cdot \Delta S$$

As ΔG describes the amount of energy that can be converted into electrical energy, a simple relation between the Gibbs free energy and the equilibrium voltage[3] E_0 of the cell can be derived.

$$\Delta G = -n \cdot F \cdot E_0$$

where n is the number of exchanges of electronic charges, F is the Faraday constant (96 485 As) and nFE_0 describes the generated electrical energy.

Thermodynamic quantities such as ΔH and ΔG depend on the concentration of the reacting components as far as these are dissolved according to the relation

$$\Delta G = \Delta G_S + R \cdot T \cdot \sum_i \ln[(a_i)^{j_i}]$$

with

a_i = activity[4] of the reacting components i (closely related to the concentration),

j_i = number of equivalents of this component that take part in the reaction,

R = molar gas constant for an ideal gas ($R = 8.3413$ J/K/mole),

ΔG_S = standard value when all activities are unity.

From these equations, the Nernst equation describing the concentration dependence of the equilibrium voltage can be derived, where $E_{0,S}$ is the equilibrium potential in standard conditions.

$$E_0 = E_{0,S} + \frac{R \cdot T}{n \cdot F} \cdot \sum \ln[(a_i)^{j_i}]$$

[3] Occasionally, the equilibrium voltage is called the open-circuit voltage. However, strictly speaking this term only means a voltage without external current flow, and may concern a mixed potential as well. On account of secondary reactions, the rest potential in batteries are usually mixed potentials, but this is not strictly observed in practical languages.

[4] Activity described the effective concentration. Thermodynamic rules are derived for dilute solutions. The activity is equivalent to the concentration in very dilute solutions, but the activity can be different at higher concentrations as interactions among the ions in the solution need to be taken into account.

The temperature coefficient of the equilibrium cell potential can be derived from the following thermodynamic relation.

$$\frac{\mathrm{d}E_0}{\mathrm{d}T} = -\frac{\Delta S}{n \cdot F}$$

Thermodynamic calculations are always based on a complete cell, and the derived voltage refers to the potential difference between two electrodes. The potential difference between the electrode and the electrolyte, the "absolute potential", cannot be determined. The name electrode potential always refers to a potential difference measured with the aid of a reference electrode. To get a basis for the electrode-potential scale, the zero point was arbitrarily equated with the potential of the standard hydrogen electrode (SHE[5]).

20.2.1.2 Electrode kinetics at current flow

When current flows through a cell, the reaction must take place at a corresponding rate. For each delivered ampere second, a corresponding number of electron exchanges must have occurred. This means that at the electrodes the elementary processes

$$S_{\mathrm{red}} \rightleftarrows S_{\mathrm{ox}} + n \cdot \mathrm{e}^- \quad \text{or} \quad S_{\mathrm{ox}} + n \cdot \mathrm{e}^- \rightleftarrows S_{\mathrm{red}}$$

must take place $6.42 \times 10^{18}/n$ times (reciprocal of one elemental charge). To achieve this current flow, additional forces are required, which intensify the electron and ion flow in the required direction. These additional forces find their expression in deviations from the equilibrium data, which means irreversible energy loss.

Usually, the reaction path consists of a number of reaction steps that precede or follow the actual charge-transfer step, and the rate of each these reaction steps is determined by its kinetic parameters, such as exchange current density, diffusion coefficient or transport numbers. The slowest partial step of this chain is decisive for the rate of the overall reaction. As a consequence, limitations of the reaction rate are often not caused by the electron-transfer step itself, but by preceding or following steps such as the transport rates of the reacting ions to and from the electrode surface.

Transport processes play an important role, because electrochemical conversion can only take place when reaction partners and electrons are available at the same time. Frequently, the reaction substances must be brought to these places or transported away, for example, when the reactions include substances in the dissolved state.

Electrochemical equilibrium is always composed of two reactions, the actual reaction and its reversal. The resulting current/voltage relation is called the 'current–voltage' curve or the current–potential curve. It is composed of two curves, one for the forward and the other for the reverse reaction. At equilibrium, both reactions are in balance.

The forward and reverse reactions of an arbitrary electrode are shown in Figure 20.1. The horizontal axis represents the electrode potential related to the equilibrium potential E_0. The vertical axis represents the current density, which is synonymous with the reaction rate.

[5] The standard hydrogen electrode means a hydrogen electrode immersed in acidic solution with H^+ ion activity of 1 mole/dm^3 and H_2 pressure of 1 atm. The specification of the electrolyte concentration is required because the potential of the hydrogen electrode depends on the H^+ ion concentration and is shifted by -0.0592 V when the H^+ concentration is reduced by one decade. The potential of the standard hydrogen electrode at 25 °C is synonymous with the zero point of the potential scale. The temperature coefficient of the standard hydrogen electrode is $+0.871$ mV/K.

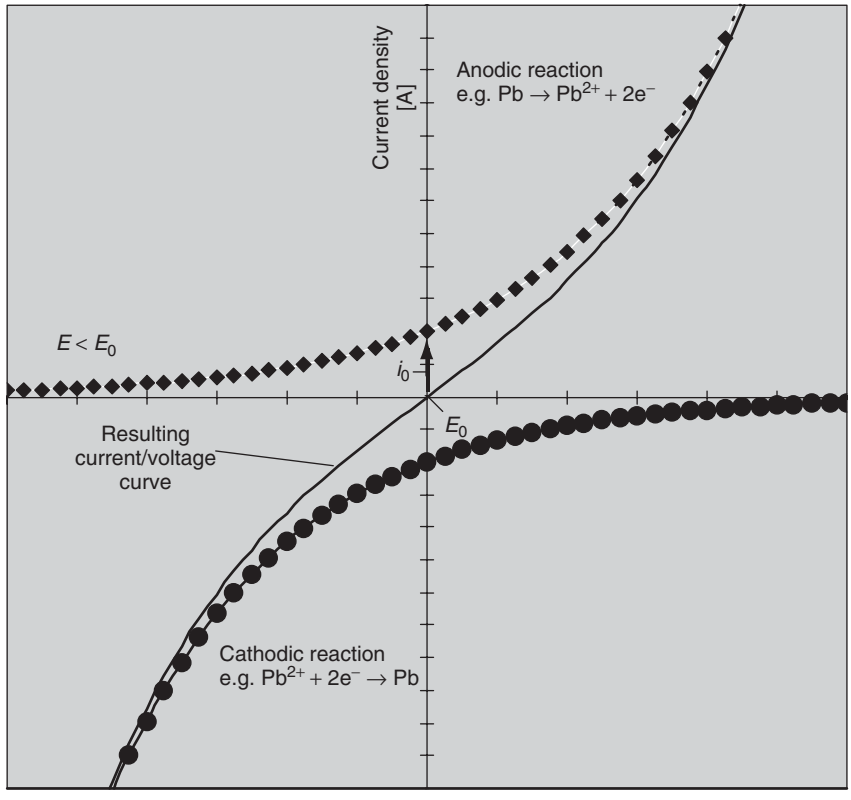


Figure 20.1 Current/voltage curve based on the Butler–Volmer equation (example taken from the lead electrode of a lead–acid battery)

The exponential relation between voltage and current is based on the fact that the charge/discharge reaction, in which the electrons are released or absorbed (the so-called transfer reaction) can be approximately described by an exponential law, called the Butler–Volmer equation.

$$i = i_0 \cdot \left\{ \exp\left[\frac{\alpha \cdot F}{R \cdot T} \cdot (E - E_0) \right] - \exp\left[-\frac{(1-\alpha) \cdot F}{R \cdot T} \cdot (E - E_0) \right] \right\}$$

where i is the current density, i_0 the exchange current density, E the actual potential, E_0 the open-circuit electrode potential and α the transfer factor describing the efficiency of the overvoltage on forward and backward reactions. The difference $E - E_0$ is called overvoltage or polarisation. The difference expresses the additional energy voltage required to force the current through the surface. The exponential relation between current and voltage means that the increase in current might be enormous when the overvoltage exceeds certain values.

The equilibrium voltage E_0 is determined by the point at which the forward and reverse reactions are equally fast. In lead–acid batteries, this is the point where metal dissolution and deposition balance each other, which means that the current densities of the forward and reverse reaction equal each other. This equilibrium potential represents a dynamic equilibrium: current flow occurs in both directions, but does not appear externally.

The current density for the forward and reverse reactions at the open-circuit potential is called the exchange current density i_0, which describes the rate at which this equilibrium is

adjusted. The exchange current density represents an important kinetic parameter. High exchange current density means that the equilibrium potential is rather stable, while a low exchange current density indicates that the electrode potential will be polarised, even when very small current densities flow through the electrode. On the other hand, it is important that unwanted side reactions have rather small exchange current densities. In the lead electrode, the exchange current density for the charge/discharge reaction is of the order of 10^{-5} A/cm^2 while it is only of the order of 10^{-13} A/cm^2 for the hydrogen production.[6] Therefore, the hydrogen evolution at open-circuit voltage is rather small.

In the literature, simplified versions of the Butler–Volmer equation can be found. For high overvoltages caused by the electrochemical charge-transfer (trans) process, the so-called Tafel equation is a proper approximation.

$$(E - E_0)_{\text{trans}} = \frac{R \cdot T}{\alpha \cdot F} \cdot \ln\left(\left|\frac{i}{i_0}\right|\right)$$

Using a semi-logarithmic plot results in straight lines, called the Tafel lines. For mathematical reasons, α is signed positive for positive currents and negative for negative currents.

For small overvoltage, the first-order approximation of the exponential terms in the Butler–Volmer equation results in the following equation:

$$(E - E_0)_{\text{trans}} = \frac{R \cdot T}{F} \cdot \frac{i}{i_0}$$

Electrochemical reactions, chemical reactions as well as transport processes that precede or follow the charge/discharge step, lead to changes in the concentration of the reacting substances at the electrode surface and thereby may change the current/voltage curves. Each of these steps can cause an overvoltage. If the diffusion of one of the reacting partners to the electrode surface is the slowest partial step, then the concentration of this substance is reduced more and more with increasing overvoltage. A limit is reached when the concentration of the reaction partner is reduced to zero at the electrode surface. From this point, further increase in overvoltage no longer increases the current. In fact, with rising overvoltage typically a side reaction becomes dominant and the current goes into this reaction. This is the case with the hydrogen evolution at the lead electrode. If the electrode is totally charged and the overvoltage is increasing, the current going into the hydrogen-evolution reaction takes over the complete current through the electrodes. This happens even though the current-exchange density for the hydrogen evolution is several decades smaller than that of the lead charging/discharging reaction.

Diffusion processes can be characterised by a limiting current i_{lim}, which describes the maximum flow of charge carriers that can be transferred through diffusion to the reaction site. The overvoltage of this diffusion process (diff) can be described by the following equation.

$$(E - E_0)_{\text{diff}} = \frac{R \cdot T}{n \cdot F} \cdot \ln\left(1 - \frac{i}{i_{\text{lim}}}\right)$$

An effect not often described explicitly is the ‘production’ of the charge carriers from a chemical process. Typically, this is included in the diffusion overvoltage, but for a deeper understanding of the battery processes and the effects of ageing (diffusion itself is not affected directly by ageing) it is worthwhile to separate these effects.

[6] It is worth noting that the current density caused by the current flow through the electrode during a charge or discharge of a lead–acid battery (approximately 10 h discharge or charge) is of the order 10^{-5}–10^{-6} A/cm^2 for the Pb electrode (assumptions: capacity of the lead electrode 3.865 g/A h, inner surface of Pb active material 0.5 m^2/g and discharge current 0.1 A/A h). This gives a feeling for the very high activity in equilibrium conditions.

To explain the effect, the lead–acid battery is taken as an example. Figure 20.10 in Section 20.4.7.1 will describe the process in more detail. From the electrochemical process described by the Butler–Volmer equation or the Tafel equation charged ions are released into the electrolyte during the discharge process. This increases the concentration c of charged ions in the electrolyte above the equilibrium concentration c_0 (defined by the solubility of the ions in the electrolyte) resulting in a concentration (conc) overvoltage. The following equation gives the mathematical formulation of this overvoltage.

$$(E - E_0)_{\text{conc}} = \frac{R \cdot T}{n \cdot F} \cdot \ln\left(\frac{c}{c_0}\right)$$

As soon as the concentration of any species in a solution deviates from its equilibrium concentration, chemical processes driven by concentration gradients occur. In the case of the lead electrode, dissolved Pb^{2+} ions form the lead sulphate crystals ($PbSO_4$) with $SO_4{}^{2-}$. The rate of formation of sulphate crystals (and the dissolution of the crystals during charging) determines the concentration of the charged ions in the electrolyte and therefore the concentration overvoltage. The rate of forming and dissolving of sulphate crystals depends strongly on the crystal size, shape and number. These parameters depend on the operating conditions of the battery and on the ageing of the active material. Therefore, battery ageing and active material structure is reflected by the values of the concentration overvoltage and the charge-transfer overvoltage.

The diffusion overvoltage describes the transport of ions that are available in sufficient volume (in the lead–acid battery these are the $SO_4{}^{2-}$ ions), thus a classical transport phenomenon. The concentration overvoltage describes the generation with respect to the absorption of ions from a chemical process.

Chemical processes are always driven by concentration gradients. Electrochemical processes are driven by the external currents. Therefore, the electrochemical process must take place exactly at the rate given by the external current. The rate of the chemical process depends only on ion concentration. With respect to the charge/discharge process in a battery, this means that the electrochemical process follows without delay any changes in the external current flow. Chemical processes have time constants as the rate of the process depends on the concentration in the electrolyte. In steady-state charge and discharge conditions, the rates of the electrochemical process and the chemical process need to be equivalent. This means that the ratio of the ion concentration and the equilibrium concentration in the electrolyte is constant.

All processes strongly depend on the temperature. The temperature dependence of the reaction rate k of a chemical reaction is described by the Arrhenius equation (C is a constant, E_A the activation energy).

$$k = C \cdot \exp\left(-\frac{E_A}{R \cdot T}\right)$$

The activation energy for many processes is of the order of 50 kJ/mol. From this experience, the rule of thumb '*increase of temperature by 10 K increases the reaction rate by a factor of 2*' is derived.

20.2.2 Batteries with Internal and External Storage

Electrochemical accumulators convert electrical energy into chemical energy. The energy is stored in a chemical compound. In secondary electrochemical batteries, this process is reversible. During discharging, the chemical energy is converted back into electrical energy. Thus, the converters determine the charging and the discharging power and the storage determines the energy capacity of the systems. This principle concept is described in Figure 20.2.

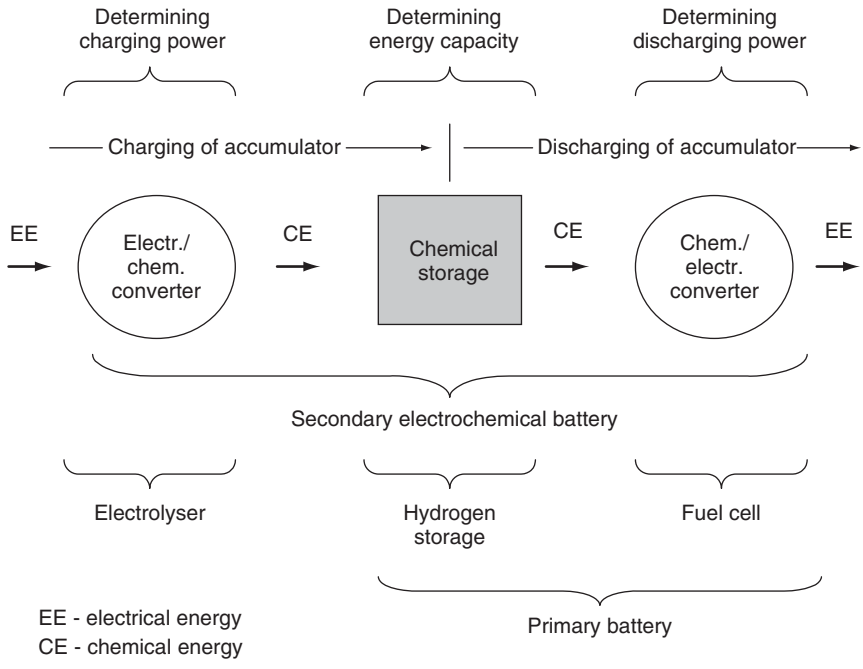


Figure 20.2 Schematics of electrochemical storage systems

In secondary electrochemical batteries with internal storage, the converter and the storage cannot be separated. The interface of the active material to the electrolyte is equivalent to the converter; the transformed active mass is the storage. In conventional secondary electrochemical batteries, power and capacity depend on each other and cannot be designed independent of each other. In practice, there is a small margin for the design. For high-power and low-capacity requirements, very thin electrodes having a high surface-to-capacity ratio are used. Section 20.4 will go into the details of this type of batteries.

Nevertheless, the margin is limited and this is a drawback for autonomous power supply systems where high energy capacity is required and the power requirement is moderate. Therefore, electrochemical storage systems with separated converters and storage are of interest. The electrolyser/hydrogen storage/fuel cell system is a well-known option for the problem, even though it is not yet a common commercially available solution. Details of this system will be discussed in Section 20.5.2.

A second class of storage systems with separated converter and external storage units are electrochemical Redox systems, where the reaction partners such as iron and chromium salts or vanadium salts are dissolved in liquids and stored separately in tanks. The converter functions quite similar to a fuel cell. During charging or discharging, the electrolyte is pumped into the converter. These systems are not available in large quantities in the market yet, but there are several R&D activities on these systems. Section 20.5.1 will discuss this technology in more detail.

For a clear understanding of this chapter and to get familiar with the wording used in the ‘storage community’, the following two sections explain some basics. They are oriented very much on the application of storage systems in autonomous power supply systems. They are by no means complete, but should help to understand this chapter without additional literature. For more information, References [2, 3] are highly recommended.

20.2.3 Commonly Used Technical Terms and Definitions

A battery is made from two or more electrochemical *cells* connected in series. *Primary* and *secondary* electrochemical cells can be distinguished. Secondary batteries – also called accumulators – have reversible reactions and are rechargeable. This chapter is centred around them.

An electrochemical cell consists of two *electrodes*. Commonly, one is called the 'positive' electrode and the other, the 'negative' electrode. The positive electrode has a more positive potential than the negative electrode with respect to the standard hydrogen electrode.[7] Each combination of charged and discharged active material has a specific electrochemical potential. The potential difference between the positive and the negative electrode is called the *cell potential* or *cell voltage*. The *equilibrium voltage* of a cell is a function of the electrolyte concentration and the temperature. The *open-circuit voltage* can be measured if no external current flows through the battery. It is identical to the equilibrium voltage if all the internal overvoltages mainly caused by diffusion processes have levelled out. The time until this stadium has been reached depends on the battery technology and the operation condition. It is in the range of some seconds to many hours.

The *capacity* of a cell is measured typically in ampere hours (A h). The capacity is determined by a constant current discharge down to a defined *end-of-discharge voltage*. The capacity depends significantly on the discharge current and the temperature. Battery manufacturers can define the discharge current and the end-of-discharge voltage on their own. Therefore, it is very important to check the reference conditions defined by the manufacturer while comparing the capacity of different products.

Typically, nominal cell voltages are in the range between 1.2 and 3.6 V. Therefore, several cells are usually connected in series to build a *string* of higher nominal voltage. The *nominal voltage* of a battery is therefore defined by the number of cells connected in series times the nominal cell voltage of a single cell. Batteries are often sold in so-called blocks or modules. Therein, several cells have been integrated and connected in series with only one set of terminals. A well-known example is the starting, lighting, ignition (SLI) battery for cars where six cells are connected in series but are sold as one 12 V block. To increase the capacity of a cell, often several sets of positive and negative electrodes are connected in parallel within a single cell. To increase the capacity even more, two or more strings can be connected in parallel. The *nominal energy* content (W h or kW h) of a battery is defined by the nominal battery voltage times the nominal battery ampere hour capacity.

The *state of charge* (*SOC*) gives the capacity that can be discharged from a battery at a certain moment. Hundred percent state of charge means a fully charged battery, 0% SOC means that the nominal capacity is discharged. State of charge is defined in more detail in Section 20.2.4. Often, instead of SOC the *depth of discharge* (*DOD*) is used in the literature or in data sheets. DOD is defined as 0% when the battery is fully charged and as 100% after the nominal capacity is discharged from the battery (DOD = 100% – SOC).

When looking up literature related to autonomous power supply systems, typically a *positive battery current* is defined to increase the SOC of the battery while a negative current decreases the SOC. However, please be aware that some authors use the opposite definition.

A *cycle* refers to a discharge followed by a recharge. Cycles used in data sheets always start from a fully charged battery up to a certain DOD. A nominal *full cycle* is a discharge down to 100% DOD. The *cycle lifetime* for a battery is given by the number of cycles as a function of the DOD. Nevertheless, in autonomous power supply systems cycles as defined above do not occur as can been seen from Figure 20.6. Many *partial cycles* within a *macrocycle* (time between two full

[7] The standard hydrogen electrode is a platinum electrode rinsed with hydrogen gas in 1 N electrolyte. Its potential is defined as 0 V.

charging states) occur, where a partial cycle is defined as the charge transfer within the time of the change of the direction of the battery current. Overall, charge transfer of batteries in autonomous power supply systems can be defined by the *capacity throughput*. It is given by the accumulated ampere hour discharged from the battery divided by the nominal capacity. The resulting number is formally equivalent to the number of 100% DOD cycles delivered during the battery life. This normalised number will be referred herein as the capacity throughput.

The *ampere-hour efficiency* η_{Ah} is defined as the ratio of the ampere hours discharged from the battery divided by the ampere hours charged to the battery within a certain period (typically one month or one year or within a period between two full charging processes). Often the *charge factor* is used instead of the ampere hour efficiency. It is defined as $1/\eta_{\mathrm{Ah}}$. For a sustainable battery operation, charge factors greater than one are necessary.

The *energy efficiency* η_{Wh} is the ratio of the energy discharged from the battery divided by the energy charged to the battery within a certain period (defined as above).

The size of a battery is given by its nominal energy content in the fully charged condition. To express the relative size of a battery concerning the load in autonomous power supply systems, often the term *days of autonomy* is used. The 'days of autonomy' is defined by the ratio of the nominal energy content of the battery (kW h) (sometimes the practical capacity according to Figure 20.3 instead of the nominal capacity is also used) to the average energy consumption per day (kW h/day). Therefore, the unit is days, and expresses how long a system can be supplied only from the fully charged battery.

Battery currents are usually given relative to the battery size. The reason is that the strains and the current-dependent electrical properties are related to the specific current loads to the electrodes with respect to the active materials. For larger capacities just formed from parallel-connected electrodes or cells or from larger electrodes, the normalisation of the current to the capacity is an appropriate measure. Therefore, battery currents are expressed as multiples of the ampere hour capacity or as multiples of the capacity-defining discharge current. For a battery with a capacity of $C = 100\,\mathrm{A\,h}$, a current of 10 A is defined as $0.1 \times C$. In the example, 100 A is called the *C-rate*. I_{10} is the current that discharges a fully charged battery within 10 h down to the defined end-of-discharge voltage. The typical nomenclature for the capacity is C_x where x is the time in which the battery is discharged. For example: $C_{10} = 10\,\mathrm{h} \times I_{10}$, or $C_{10} = 100\,\mathrm{A\,h}$, $I_{10} = 10\mathrm{A} = 0.1 \times C_{10}$.

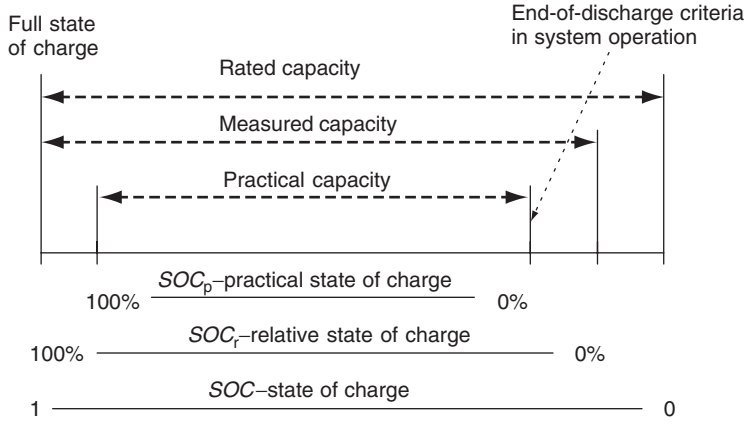


Figure 20.3 A comparison of the different definitions of battery capacity and the corresponding definitions of state of charge

Note that $1 \times I_{10}$ is not equivalent to $10 \times I_{100}$ as the C_{100} capacity is typically larger than the C_{10} capacity. For a more detailed example, see Section 20.4.7.3.

The *end-of-charge voltage* defines an upper voltage limit. Charging of the battery usually is not stopped on reaching the end-of-charge voltage (other than the end-of-discharge voltage), but the charge current is reduced accordingly to maintain the end-of-charge voltage over time.

The *lifetime* of a battery depends very much on the operating conditions and the control strategy. Manufacturers usually define two types of lifetime: the *float lifetime* (calendar lifetime) gives the lifetime under constant charging conditions without cycling (typical applications are uninterruptible power supplies), and for continuous cycling (*cycle lifetime*, typical applications are fork-lift trucks). Sometimes, the *shelf lifetime* is given. It defines the time for which a battery can be stored before usage.

Self-discharge describes the (reversible) loss of capacity on open-circuit conditions. It depends very much on the temperature.

The *state of health* is defined as the ratio of the actual measured capacity and the rated or nominal capacity. The state of health indicates to which extent the battery is still able to fulfil the requirements. According to the norms, lead–acid batteries are at the end of their lifetime if the state of health is under 80%. However batteries can be operated significantly longer, but the days of autonomy are reduced accordingly and a system might not fulfill the energy requirements any more in a proper way. Batteries operating at a state of health of 50% are found frequently especially in hybrid systems. As a consequence, the share of the motor generator is increasing.

20.2.4 Definitions of Capacity and State of Charge

For operation and energy management in autonomous power supply systems, the battery capacity and the actual state of charge of the battery are the most important parameters. State-of-charge determination is difficult in autonomous energy supply systems with renewable energies because full charging of the battery as it is done frequently with conventional battery chargers is very unusual.

If state of charge is displayed, the question that arises is, what is the meaning of the specific values. Figure 20.3 shows different definitions of the battery capacity and the corresponding definitions of state of charge. The measured capacity of a battery might be smaller or even higher than the rated capacity given by the manufacturer. During the lifetime, the measured capacity decreases more and more due to ageing effects. The practical capacity is less than the measured capacity. Owing to the special conditions of renewable energy sources, batteries are almost never completely recharged (number of charging hours is limited) [4]. The maximum state of charge that is reached during normal system operation is called a 'solar-full state of charge'. Further on, the system defines an end-of-discharge criterion to avoid deep discharging of the battery and therefore accelerated ageing, which usually differs from the end-of-discharge criteria used for capacity tests. Therefore, the practical battery capacity is lower than the measured capacity. In the literature and other publications, no common definition of the state of charge is used. Therefore, any data and results must be handled with care.

Within this chapter the following definitions according to Figure 20.3 and reference [5] will be used.

The rated or nominal capacity is defined as the 10-h discharge capacity C_{10}. This is the basis for the SOC determination. The rated or nominal capacity does not change during the life of a battery whereas the measured capacity changes with time. The state of charge with respect to the measured capacity is called relative state of charge (SOC_r). The practical capacity C_p is always

less than the measured capacity. The state of charge definition related to the practical capacity is the practical state of charge (SOC_p). SOC_p is 100% if a solar-full state of charge is obtained.

In reference [5], a complete review of the different definitions for the capacity, state of charge and full state of charge is given. Further on, definitions of open-circuit voltage and state of health are included, because some state-of-charge meters and algorithms use these definitions as well.

20.3 TYPICAL OPERATION CONDITIONS OF BATTERIES IN PV APPLICATIONS

To understand the requirements of storage systems for autonomous power supply systems, an analysis of the typical operating conditions is necessary. The operating conditions vary very much according to the location and application of the system, the load patterns, the installed power generators and the operation strategy.

The most important parameters for the classification of the operating conditions are the charge and discharge currents, the state-of-charge profile and the temperature.

20.3.1 An Example of an Energy Flow Analysis

Photovoltaic stand-alone systems may be roughly separated into two groups. There are systems consisting of the PV modules that charge the batteries. A charge controller prevents overcharging or deep discharging. The appliances are supplied from the battery either directly or via a DC/AC converter. The typical representative of this first system group is the Solar Home System (SHS), which is operating in hundreds of thousands of rural households. It will supply lights and TV sets and delivers, in a standard version, approximately 0.25 kW h of electricity per day. Larger systems may deliver up to 5 kW h per day.

The second group of systems combine a photovoltaic generator with a diesel gen-set and possibly with additional wind turbines or hydroelectric generators. Including a diesel gen-set as a controllable generator gives an additional degree of freedom to the system sizing.[8] It allows reducing the battery capacity, especially if the solar radiation undergoes strong seasonal variations. These systems are called hybrid systems. They are designed to deliver from 1 kW h per day to typically between 10 and 100 kW h per day. They may be used to supply power to telecommunication equipment, mountain lodges, hospitals or hotels in non-electrified rural areas and small villages.

A typical hybrid system with a 4.5-kW_p PV generator, a diesel generator and 32 kW h of lead–acid battery storage is in operation since 1992 in the Black Forest in Germany. It has to deliver 10 kW h per day to the hikers inn ‘Unterkrummenhof’. An analysis done in model calculations based on measured data shows the flow of energy in the system and the effect of internal losses (Figure 20.4). Two-thirds of the power delivered ($E_{consumer}$) is drawn from the battery storage. The numbers given in the diagram are normalised to the nominal energy production of the PV generator under the radiation conditions at the site. A detailed discussion of the energy-flow diagram is given in reference [6].

Figure 20.4 underlines the key role that battery storage has to fulfil in hybrid systems. More than 80% of the energy used goes via the battery storage. This is a typical value for all hybrid systems and is even higher in many pure PV battery systems.

[8] Diesel gen-sets are currently the most common solution for an additional controllable generator. Other solutions like thermoelectric, thermophotovoltaic or fuel cell generators have been developed in many places and might be alternatives in the near future.

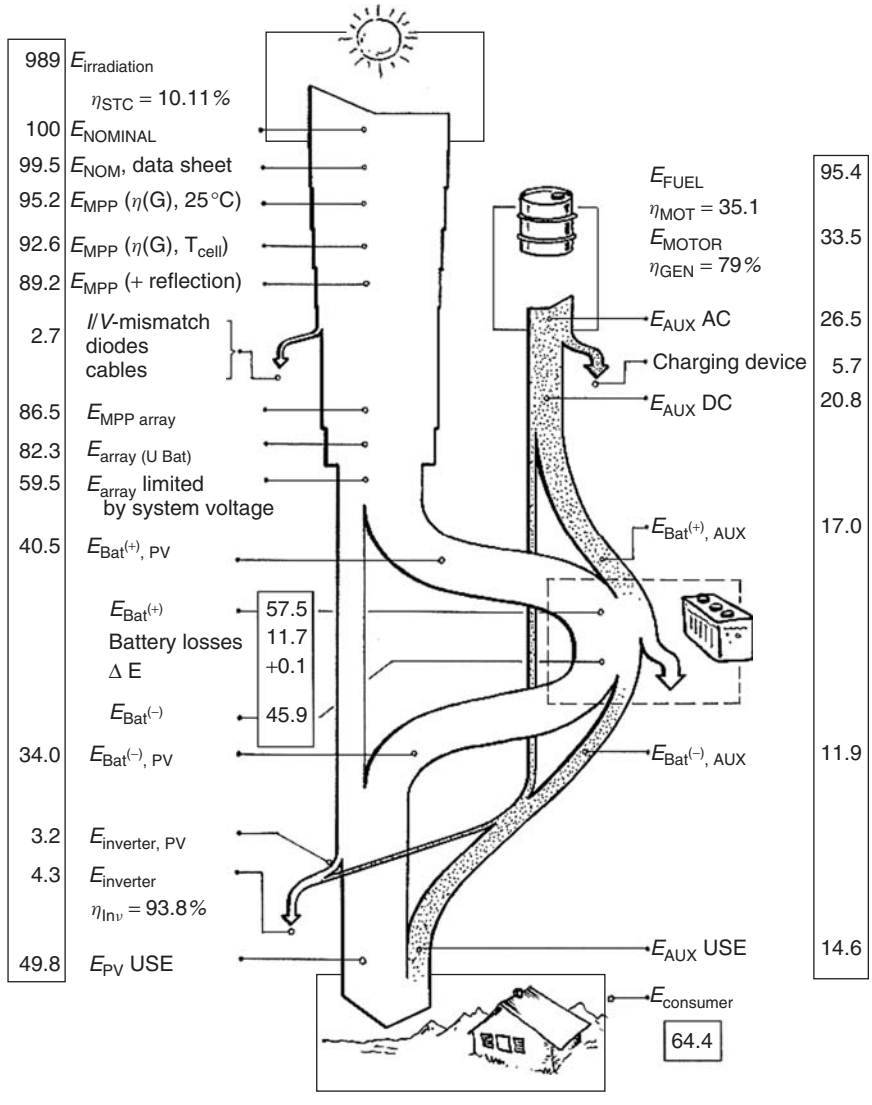


Figure 20.4 Energy-flow diagram on a one-year basis of the autonomous power supply system Unterkrummenhof (PV diesel battery) near Freiburg/Germany [6]

20.3.2 Classification of Battery Operating Conditions in PV Systems

An intensive study of the operational data of close to 30 batteries in stand-alone PV systems with and without a diesel generator was made. All systems were operated under European radiation conditions [7]. The study resulted in a classification of the battery-operating conditions into four classes. Figures 20.5 and 20.6 show measured data on the annual operating conditions of four systems selected to represent the four different classes. Figure 20.5 shows scatter graphs of the battery current versus battery voltage and Figure 20.6 shows the time series of the state of charge within a complete year.

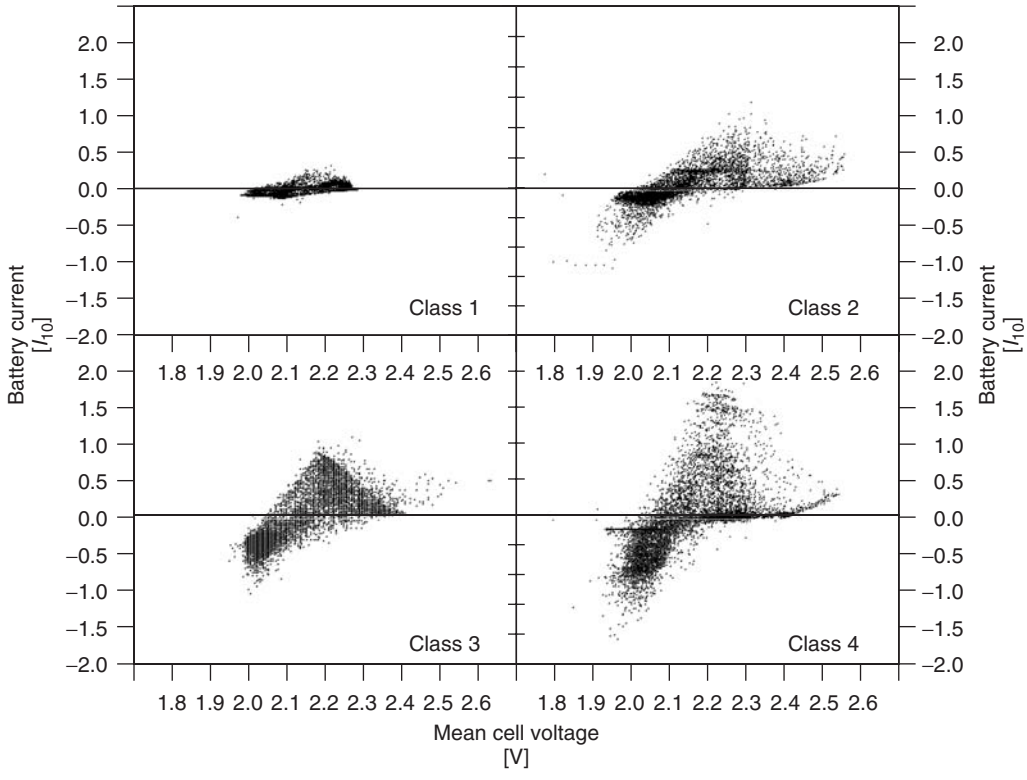


Figure 20.5 Battery current in units of I_{10} versus cell voltage for all data points. For each class, data of a complete year were taken from a representative system

The hybrid system 'Unterkrummenhof', characterised in Figure 20.4 is a system from Class 2. Class 3 and Class 4 represent systems with an increasing role of the diesel gen-set and relatively smaller PV generators and batteries. The Class 1 system is a system without a back-up generator, designed to operate with high reliability in Europe.

Solar home systems (SHSs), which were not included in the survey, operating under the favourable conditions of low latitudes would typically be equipped with a three- to five-'day' battery and would not show the pronounced long-lasting period of deep discharge during the winter months. Classical SHSs are very much like Class 2 or Class 3 systems and village power supply systems like Class 4.

An extension of the classification including the Southern Europe climatic conditions is necessary, but is not available yet.

The time series of a battery's state of charge (Figure 20.6) and the current–voltage representation (Figure 20.5) demonstrate that batteries in stand-alone systems have to operate under very specific conditions such as the following:

- The charging and discharging currents are small compared with the standard 10-h discharging current I_{10} (at least for system Classes 1 and 2).
- For long periods, sometimes weeks or even months, the batteries do not reach a fully charged state (SOC = 100%).

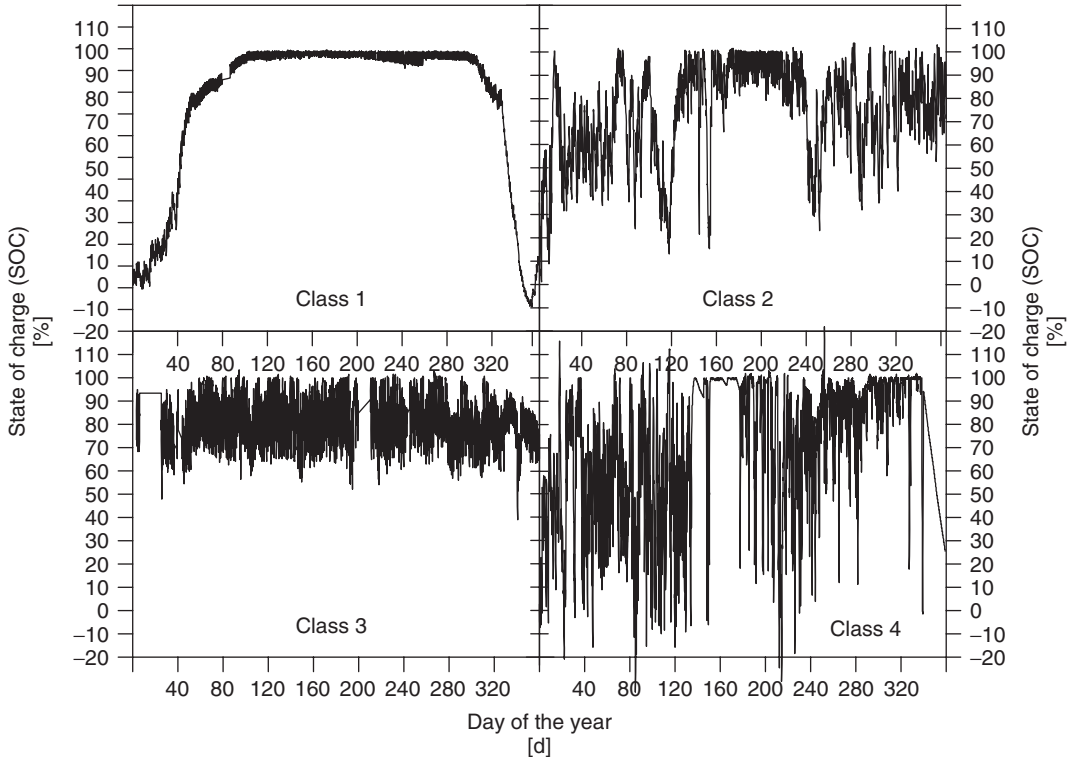


Figure 20.6 Time series of state of charge calculated from current, voltage and temperature using an ampere-hour balance with a voltage and temperature-dependent loss current [5, 8]

These features are distinct from other battery applications in, for example, uninterruptible power supplies, where the batteries are kept at a full state of charge for the longest time of the year or in vehicle traction applications, for example, fork-lift trucks, where the batteries are fully recharged regularly with high charging current. Figure 20.6 also demonstrates that in systems of Classes 1 and 2 – the same is true for the solar home system application – the daily discharged capacity is between 5 and 30% of the nominal battery capacity. This is equivalent to the annual full cycles between 20 and 100.

Table 20.2 gives a comprehensive view of the requirements of the battery in the different system classes. In Table 20.3, suggestions for the selection of adequate lead–acid battery types out of the variety of products available in the market on the basis of this classification are made.

On the basis of this classification, an evaluation of the properties of a battery according to the requirements of the systems is possible. Solar fraction and storage size in days of autonomy are the output of all commercial system design and simulation-software packages. Therefore, the classification allows the system designer to ask the battery manufacturers for an appropriate battery type by showing him the typical operating conditions. The differences in the operating conditions and the requirements listed in Table 20.2 show clearly that one 'solar battery' cannot exist. The range of operating conditions in autonomous power supply systems is very large and requires appropriate individual solutions.

An additional parameter for differentiating the operating conditions of batteries is the amount of capacity throughput caused by an AC ripple. Loads and generators are connected to the battery

Table 20.2 Identification of classes by typical system indicators (solar fraction, storage size) for the different classes of operating conditions and importance of battery features for the different classes. The storage size is given in units of battery capacity divided by the mean daily load (days of autonomy). The solar fraction is the amount of energy produced by the PV generator divided by the energy produced by all the energy sources within the system (including the diesel generator in hybrid systems) [9]

System indicators	Class 1	Class 2	Class 3	Class 4
Solar fraction/%	100	70–90	About 50	<50
Storage size/days of autonomy	3→10	3–5	1–3	About 1
Capacity throughput*	10–25	30–80	100–150	150–200
Necessary battery features				
Number of lifetime cycles*	Low (<300)			High (>1200)
Capability to withstand long periods in deep discharged states	Important			Less important
Low self-discharge rate	Important (<1% per month)			Less important (5% per month)
Measures against acid stratification	Important	Very important		Important
Resistance against corrosion	Important			Less important

*The capacity throughput is defined as the number of ampere-hours discharged from the battery divided by the nominal capacity of the battery. The given numbers are typical of the applications in the defined classes of operating conditions. A full cycle with regard to lifetime cycles is defined by a one capacity throughput. This is equivalent to a complete discharge (100% DOD) of a fully charged battery. In data sheets, often the cycle number is given for cycles with a depth of cycle other than 100% (e.g. 80%). This has to be taken into account while comparing the design capacity throughput of different products (see also Figure 20.24).

Table 20.3 Continuation of Table 20.2. Different product groups of lead–acid batteries are classified with respect to the different classes as defined in Section 20.3.2 (● optimum, ○ acceptable) table from reference [9]

Type of battery	Class 1	Class 2	Class 3	Class 4
SLI	–	–	–	–
Stationary	●	●	○	–
Traction	–	○	●	●
Electric vehicle	–	○	●	●
'Solar battery' (from SLI)	●	○	–	–
VRLA	●	●	●	○
Flooded	●	●	●	●

at the same time. This results, in many cases, in so-called microcycles in which the battery current changes from charging to discharging and vice versa with a frequency between 1 and 300 Hz. Measurements and calculation have shown that this can cause an additional capacity throughput of up to 30%. This in fact shortens the lifetime of the batteries. Taking into account the optimum operating strategies, battery ageing is related strongly to the capacity throughput. The amount of this additional capacity throughput depends very much on the system's sizing and the electrical properties of the loads, the inverters, the converters and the generators [10].

20.4 SECONDARY ELECTROCHEMICAL ACCUMULATORS WITH INTERNAL STORAGE

20.4.1 Overview

There are several secondary electrochemical accumulators available on the market. They differ in parameters concerning the materials of the electrodes and the electrolyte. This results in different electrical properties such as energy and power density, efficiency, lifetime, cycle life, operation temperature, inner resistance and self-discharge and last but not the least economic properties like battery costs and maintenance requirements.

Products such as lead–acid batteries, $ZnBr_2$, NiCd, NiFe, NiZn, nickel–metal hydride (NiMH), Zn–air, Li–ion, Li–polymer, Li–metal and rechargeable alkali mangan (RAM) are available. They operate at room temperature, but also high-temperature batteries such as NaS and $NaNiCl_2$ (ZEBRA) operating at 300–350 °C are possible. In addition, there are capacitors storing the energy in an electrostatic field instead of chemical bonds out of which the double-layer capacitors are the most interesting for autonomous power supply systems.

The specific energy densities of batteries are an important parameter to characterise the different battery types. For logistic reasons on supplying the batteries to the systems, gravimetric and volumetric energy densities are a relevant cost factor also for autonomous power supply systems. Figure 20.7 shows an overview of the energy densities of different secondary battery technologies based on an analysis of commercially available products. The figure does not show the theoretical limits of the different technologies.

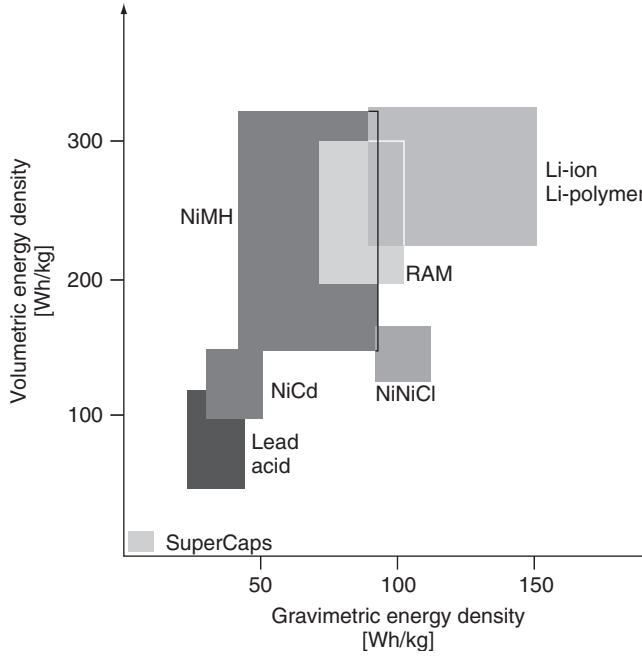


Figure 20.7 Practical specific volumetric energy density (W h/l) versus specific gravimetric energy density (W h/kg) for various secondary battery technologies (data from Table 20.4)

Table 20.4 gives an overview of the most important properties of several secondary electrochemical accumulators. A more detailed description is given on the NiCd (Section 20.4.2), the nickel-metal hydride (NiMH) (Section 20.4.3), the RAM (Section 20.4.4) and the Lithium batteries (Section 20.4.5) as well as on the double-layer capacitors (Section 20.4.6). Lead–acid batteries are treated in greater detail in Section 20.4.7. All other battery types have no relevance in the field of autonomous power supply systems. Details on all primary and secondary batteries can be found in reference [2].

20.4.2 NiCd Batteries

NiCd batteries have been available as a commercial product for many decades and are well-proven products. They have very good properties concerning lifetime and lifetime cycles. They are widely used in heavy-duty applications and in very cold climates. Standard NiCd-battery designs can be operated easily at temperatures of $-20\,^{\circ}\mathrm{C}$ and specially designed cells can be operated even up to $-50\,^{\circ}\mathrm{C}$. Nevertheless, NiCd batteries have a bad image due to the cadmium content, which is, known to be environmentally incompatible.

Several different types of NiCd batteries are in the market with differences in the plate technologies and the handling of evolved gases.

The basic reaction is the same for all construction types of NiCd batteries.

$$2\mathrm{NiOOH} + 2\mathrm{H_2O} + \mathrm{Cd} \rightleftarrows 2\mathrm{Ni(OH)_2} + \mathrm{Cd(OH)_2} \tag{20.1}$$

During discharge, trivalent NiOOH is reduced under consumption of water to divalent $\mathrm{Ni(OH)_2}$. Metallic cadmium is oxidised to $\mathrm{Cd(OH)_2}$. The reversible backward reactions proceed during charging.

The potassium hydroxide electrolyte (KOH) does not undergo a significant change in its concentration or density during charging or discharging. Only water, which is present in high concentrations, participates in the reaction. The electrolyte density is about $1.2\,\mathrm{g/cm^3}$.

NiCd batteries are available with liquid electrolyte and as sealed, maintenance-free types [3].

The rated voltage of NiCd cells is 1.2 V. Although the discharge rate and the temperature significantly affect the discharging behaviour of all electrochemical cells, the effect is noticeably less pronounced in NiCd batteries than in lead–acid batteries. As a result, NiCd batteries can be discharged at higher rates, without the accessible capacity falling much below the rated capacity. Even for discharge rates of $5 \times C_5$ a high-performance NiCd battery can supply 60–80% of the rated capacity. Also, the influence of the temperature on the capacity is comparatively small which is due to the fact that diffusion processes have less impact on the reaction kinetics compared with lead–acid batteries.

High temperatures in the range of $40\,^{\circ}\mathrm{C}$ and more should be avoided as the charging efficiency is getting very low and the self-discharge rate is increasing significantly. Self-discharge rates at $20\,^{\circ}\mathrm{C}$ are in the range of 20%/month. The energy efficiency is in the range of 60–70%, which is significantly less compared to lead–acid batteries.

A NiCd battery can withstand occasional deep discharge, inverse charging and also freezing of the cells without direct damage.

NiCd cells have a low internal resistance. Typical values for the DC resistance are between 0.4 and $2\,\mathrm{m\Omega}$ for a fully charged 100 A h cell. The internal resistance is largely inversely proportional to the cell size for all cell types. Falling temperatures and a decrease in the SOC increase

Table 20.4 Overview of the technical data of different secondary batteries based on actual available products. All numbers are typical data based on the data sheets of existing products. The data are not the theoretical limits for the different technologies. Products for special applications may have technical parameters outside the ranges displayed in the table

Battery technology	Electrolyte	Energy density [W h/kg]	Energy density [W h/l]	Efficiency η_{Wh} [%]	Life-time [years]	Typical. cycle lifetime [cycles]	Temperature for operation		Typical applications (examples)
							Standard charging [°C]	Discharging [°C]	
Lead–acid	H_2SO_4	20–40	50–120	80–90	3–20	250–500	−10 to +40	−15 to +50	Stationary application (UPS, autonomous power supplies), traction, SLI
NiCd	KOH	30–50	100–150	60–70	3–25	300–700	−20 to +50	−45 to +50	Power tools, hobby toys, consumer products, traction, applications as for lead–acid batteries with higher power requirements or lower ambient temperature, electrical cars
NiMH	KOH	40–90	150–320	80–90	2–5	300–600	0 to +45	−20 to +60	Laptop, mobile phones, camcorder, electric vehicles, hybrid cars, hobby toys
Li–ion, Li–polymer	Organic, polymers	90–150	230–330	90–95	–	500–1000	0 to +40	−20 to +60	Laptop, mobile phones, Camcorder, smart cards
RAM		70–100	200–300	75–90	–	20–50	−10 to +60	−20 to +50	Consumer products, hobby toys
SuperCaps		1–10	2–15	90–95	~10	500 000	−25 to +75	−25 to +75	For applications with typical cycle periods of less than 10 s at very high power requirements
NaNiCl	β-AlO_2	~100	~150	80–90	–	~1000	+270 to +300	+270 to +300	Hybrid vehicles, electric vehicles (prototypes available)

the internal resistance, but the internal resistance remains essentially constant up to a DOD of 60–80%, and only increases significantly at higher DODs. Thus, the internal resistance is not a suitable indicator to determine the state of charge.

Under normal operating conditions, a NiCd battery can reach up to 2000 100% DOD cycles, even under severe operating conditions. Depending on the application and the operating conditions, the lifetime can be between 8 and 25 years. Starter batteries for diesel generators reach lifetimes of about 15 years, batteries for train lighting achieve 10–15 years and stationary batteries have lifetimes of 15–25 years. Good charging (charge factors of approximately 1.2), frequent overcharging and frequent full discharges are necessary to achieve long lifetimes.

A number of factors are responsible for the high reliability and the very long lifetime of NiCd batteries: the design is mechanically very robust, the cells are not easily damaged by incorrect technical handling, such as inverse charging, overcharging and long idling periods at medium or low states of charge and the reactants involved are not very corrosive with regard to the electrodes and other components in the cell.

One drawback is the so-called memory effect, which occurs under some operating conditions. This term is used to describe the tendency of the battery to adapt its electrical properties to the cycling conditions in which it has been operated over a long period of time. This means that a battery that is cycled over prolonged periods up to a certain DOD, tends to limit discharging to this DOD, even if a higher discharge at high discharge current is planned. This effect can be resolved by discharging the battery several times with a low current. In modern NiCd batteries, this effect is not very pronounced any more.

Despite the good electrical properties, the market share of NiCd batteries in autonomous power supply systems is not very high due to the high costs. Investment costs for NiCd batteries are round about a factor of 3 higher than lead–acid batteries.

20.4.3 Nickel–Metal Hydride (NiMH) Batteries

The active material of the positive electrodes of a nickel–metal hydride (NiMeH or NiMH) battery in its charged state is NiOOH, the same material as in a NiCd battery. The negative active material in the charged state is hydrogen, a component of a metal hydride. The metal alloy is subjected to a reversible absorption/desorption process during charge and discharge of the cell. The reaction for the reversible charge/discharge process is indicated below [11].

$$\mathrm{NiOOH} + \mathrm{MH} \rightleftarrows \mathrm{Ni(OH)_2} + \mathrm{M} \tag{20.2}$$

An aqueous solution of potassium hydroxide is the main component of the electrolyte. Only a small amount of electrolyte is used in sealed nickel–metal hydride cells, most of which is absorbed in the separator and the electrodes. In the cell, oxygen can be transported from the positive to the negative electrode and can recombine there with hydrogen to form water. Thus, the cells can be used like dry cells and can also be installed in any desired position.

The discharge characteristics of sealed nickel-metal hydride cells are very similar to those of sealed NiCd cells. The open-circuit voltage is between 1.25 and 1.35 V/cell, and the nominal voltage is also 1.2 V.

The electrical characteristics are quite similar to NiCd batteries, even though their energy efficiency is about 80–90% and the maximum power available is less than that in NiCd batteries. The latter is of little relevance in autonomous power supply systems. Memory effects are less pronounced than in NiCd batteries. Self-discharge at 25 °C is also in the range of 20%/month, but at 45 °C it is as high as 60%/month.

Nickel–metal hydride batteries are not as robust against polarity reversal as NiCd batteries. If the positive electrode gains a negative potential, hydrogen is generated at the positive electrode. Some of the gas can be absorbed by the metal of the negative electrode, but the rest remains as gas in the cell. If the discharge is continued, oxygen is formed at the negative electrode, which further increases the gas pressure and leads to oxidation of the metal electrode. When the overpressure is large enough, the safety vent opens and the gas pressure falls again. To avoid this, appropriate measures must be taken especially if long series-connected strings are used. References [12, 13], and Chapter 19 of this book discuss a number of possible solutions for this problem.

Nickel–metal hydride batteries have replaced NiCd batteries in the market for portable appliances (e.g. mobile phones, camcorders and power tools) to a large extent due to their better environmental compatibility and their higher gravimetric energy density (Figure 20.7). However, NiMeH batteries are not commercially available in larger capacities as necessary for autonomous power supply systems. The reasons for this are mainly the costs that are actually approximately five times higher than the lead–acid batteries. Currently, there is no evidence that NiMeH batteries will play a major role in storage for autonomous power supply systems except for small technical appliances with some $10\,W_p$. It is more likely that lithium batteries may enter this market, but NiMeH batteries are seen as an interim technology between NiCd and Lithium batteries.

20.4.4 Rechargeable Alkali Mangan (RAM) Batteries

Alkali mangan cells are well known as primary batteries for several decades. Over the past few years this technology has entered the market as a secondary battery. In the beginning, the primary batteries were used and recharged. Meanwhile, rechargeable alkali manganese (RAM) cells are in the market specially designed as a secondary battery. RAM batteries are gas tight. The nominal voltage is 1.5 V/cell and is therefore 25% higher than that of NiCd or NiMeH batteries. Currently, only small batteries with capacities of up to 5 A h are in the market. They are significantly less expensive than the NiCd batteries. RAM cells have higher inner resistance than all other batteries discussed here. RAM cells are much more environmentally compatible than NiCd batteries as they contain no heavy metals.

The major drawback is the low deep-cycling lifetime of the RAM cells. Up to now only approximately 20–50 full cycles (100% DOD) are available. However, if only very shallow cycles are required (1–5% DOD), several thousands of cycles can be achieved. Even though RAM cells are currently not suited to larger autonomous power supply systems, they are interesting storage systems for small appliances with limited lifetime or usually very shallow cycling as, for example, some kinds of toys. Emergency lighting systems may be another field of application where normally only the surveillance electronics need any power (recharged by a small PV generator) and only in case of an emergency the full capacity is needed.

20.4.5 Lithium–Ion and Lithium–Polymer Batteries

Lithium batteries have been the most emerging battery technology over the last few years. Primary lithium batteries were already well known due to their very high energy density and shelf lifetimes of up to 10 years without any major self-discharge. Nowadays, lithium–ion and lithium–polymer batteries have captured the market for portable applications like camcorders, mobile and cordless phones and organisers. Even though they are not used in larger autonomous power supply systems today, it is worth having a closer look at this technology. Their electrical properties concerning efficiency and charge/discharge characteristics are very well suited to these applications. At the moment, lithium batteries are by far too expensive for applications where the high gravimetric

energy density is of little benefit. However, as this is an emerging technology and cost reductions in the manufacturing process are expected, they might play a role in some autonomous power supply applications in the next few years.

The lithium–ion rechargeable battery's operation is based simply on lithium ions migration between the cathode and anode. Lithium-ion rechargeable batteries are therefore fundamentally different from non-rechargeable lithium and, for example, secondary lead–acid or NiCd batteries in that the basic form of the cathode and anode materials does not change.

When the battery is charged, the lithium ions in the cathode material (lithium compound) migrate via a separator into the layer structure of the carbon material that forms the anode, and a charging current flows. During discharging, the lithium ions in the carbon material migrate backwards to the cathode material. This is known as the 'rocking-chair' principle. Even though a large number of different material combinations are known under the name of lithium ion batteries, the most important materials for the commercial products are of the LiCo and the LiMn type.

The reactions for the reversible charge/discharge process are indicated below.

$$\mathrm{Li}_{1-x}\mathrm{CoO}_2 + \mathrm{C}_n\mathrm{Li}_x \rightleftarrows \mathrm{LiCoO}_2 + \mathrm{C}_n \text{ (cobalt type)} \tag{20.3}$$

$$\mathrm{Li}_{1-x}\mathrm{Mn}_2\mathrm{O}_4 + \mathrm{C}_n\mathrm{Li}_x \rightleftarrows \mathrm{LiMn}_2\mathrm{O}_4 + \mathrm{C}_n \text{ (manganese type)} \tag{20.4}$$

Li–ion batteries of the modern types as marketed today have a nominal voltage of 3.6 V. As this is far above the water-electrolysis voltage of 1.23 V, no aqueous electrolytes can be used anymore. The electrolyte here is an organic solvent with dissolved lithium salts. The cathode material is lithium cobaltite (LiCoO_2) or lithium manganese oxide spinel ($\mathrm{LiMn}_2\mathrm{O}_4$). The anode material is graphite of coke (graphitised carbon).

Lithium–ion rechargeable batteries have a three-layer structure consisting of a porous separator sandwiched between sheet-like cathode and anode materials, which, in the case of a prismatic cell, are wrapped around in an elliptical form. These materials are impregnated in an electrolyte and sealed in a metal case. This metal case includes a safety vent to protect the battery by releasing gas externally if the pressure inside the cell builds up to extreme levels.

Lithium batteries are potentially risky due to their very high energy density and the reactivity of metallic lithium. Incorrect handling of a lithium rechargeable battery may cause heat, explosion or fire. Therefore, it is even more important with this battery type to assure overcharge protection, over-discharge protection, over-current protection, short-circuit protection and operation at too high temperatures. Today, lithium batteries are only supplied with an integrated control electronic as a protection device. It works independent of all external chargers or monitoring devices and is therefore fully controlled by the battery manufacturer.

The main differences between the lithium–ion and the lithium–polymer batteries can be described as follows. Lithium–ion batteries have a fluid organic electrolyte while the negative electrode is made from a lithium/carbon intercalation electrode. The electrolyte has a high conductivity. The non-metal electrode increases the safety in comparison with a Li metal electrode. What is sold today as a lithium–polymer battery is in fact a combination of a polymer electrolyte and a lithium/carbon intercalation electrode. The use of the polymer simplifies the manufacturing. Strictly speaking the so-called lithium–polymer batteries are polymer–lithium–ion batteries.

The lithium–polymer cell is just entering the market. In the long run, it is expected that lithium-polymer batteries can be manufactured at lower costs than lithium-ion batteries. Further, they allow very flexible battery designs. This makes lithium–polymer batteries an interesting solution for chip integration or smart cards, but also larger capacities for power applications are available for field demonstrations now.

Compared with NiCd or nickel–metal hydride batteries, a disadvantage of lithium batteries is that they are less tolerant to operations with high currents, which makes discharge at high currents noticeably more difficult. Also, they currently do not achieve the same cycle life as NiCd or nickel–metal hydride batteries. However, both points are subject to R&D and especially concerning the power rating, significant steps forward have been achieved.

Lithium batteries require constant current/constant voltage charging (Figure 20.23a). The recharge behaviour is very good. Full charging of the battery is not as important as with lead–acid batteries to achieve adequate lifetimes. However, the voltage limit must be observed very accurately. The end-of-charge voltage is limited to 4.1 V and must not be extravagated by more than 50 mV. High voltage causes the formation of metallic lithium. In series-connected cells, it must be assured that the voltage limits are kept within the acceptable limits for each individual cell.

The discharge of lithium batteries must be restricted to the material-specific end-of-discharge voltage. Again, over-discharge leads to the formation of metallic lithium. For the cobalt type, the end-of-discharge voltage is 2.3 V/cell and for the manganese type 2.7 V/cell. Figure 20.8 shows the discharge curves of a lithium–ion battery at different discharge currents. The battery capacity only slightly depends on the discharge current. In addition, Figure 20.9 shows the temperature dependence of the discharge curves. As the ion migration depends strongly on the temperature, the low-temperature performance is not too good.

20.4.6 Double-layer Capacitors

Conventional capacitors have a dielectric between the electrodes. Their capacity is determined by the dielectric number and the area of the electrodes. The so-called double-layer capacitors have an ion-conducting electrolyte between the electrodes. Therefore, an agglomeration of charge carriers at the interface between the electron-conducting and the ion-conducting interface is possible. The interface is called the electrochemical double layer. In contrast to secondary batteries, no chemical reaction and no charge transfer from the electrode to the electrolyte happened. Therefore, no changes in the material structure occur resulting in cycle lifetimes of several hundred thousands. The storing of energy only depends on the electrostatic effect. However, in contrast to classical capacitors in which in the dielectric only electrons are moved, in double-layer capacitors a movement of ions

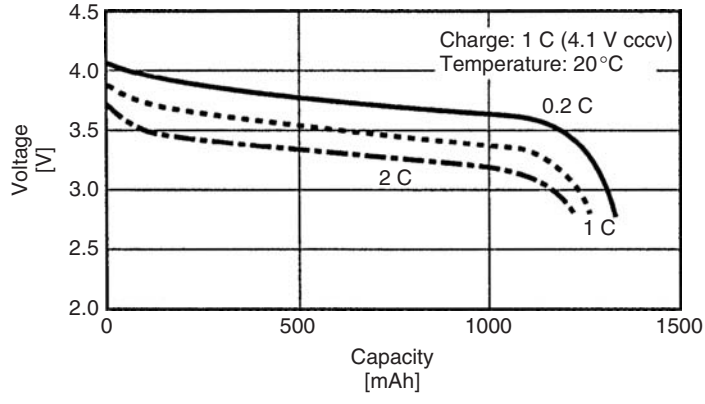


Figure 20.8 Voltage during discharge of a Li–ion battery with $C_5 = 1350\,\text{mA h}$ capacity as a function of the discharged capacity at different discharge currents and a temperature of 20 °C. The charging is done with the constant current–constant voltage (cccv or *IU*, see Section 20.4.7.6.1) regime with a charge current of 1 C and an end-of-charge voltage of 4.1 V [14]

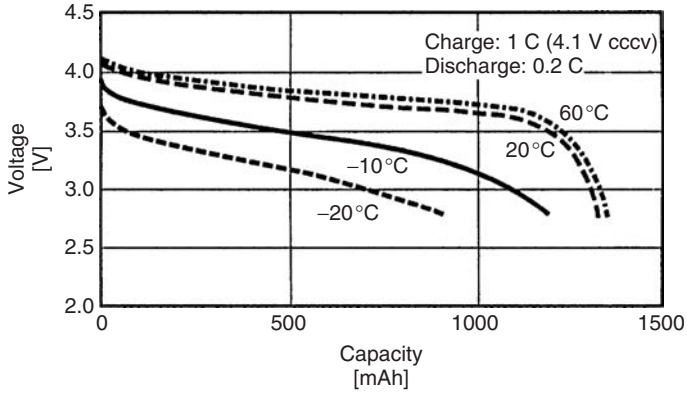


Figure 20.9 Voltage during discharge of a Li–ion battery with $C_5 = 1350\,\text{mA h}$ capacity as a function of the discharged capacity at different battery temperatures at a discharge current of 0.2 A. The charging is done as in Figure 20.8 [14]

and therefore a significant mass movement occurs. This results in diffusion-time constants during charging and discharging in the double-layer capacitor.

Depending on the electrode material, the capacity is approximately 20V–40 μF/cm^2. The electrode materials are typically made of carbon with very high surface areas of approximately 2000 m^2/g. The number of charge carriers in the double layer is limited because with increasing charge-carrier density the potential increases. If the potential is too high, the charge carriers are forced to penetrate the electrode/electrolyte interface, resulting in electrochemical reactions like in secondary batteries. However, in this case this is an irreversible effect and destroys the double capacitor. An additional problem is that in many double-layer capacitors aqueous electrolytes are used and the gassing must be avoided as well (start of water electrolysis at 1.23 V). Therefore, the maximum voltage needs to be limited to approximately 1.5–2.0 V. To avoid the electrolysis problem, organic electrolytes are used which allow maximum voltages of 3–4 V, but they have significantly lower conductivity than aqueous electrolytes. Therefore, for applications with very high power requirements capacitors with aqueous electrolytes are used; if higher energy density and lower power is required organic electrolytes can be used. Because overcharging of the double-layer capacitors will destroy them, a careful single-cell control is necessary when they are operated in long strings of series-connected cells.

Double-layer capacitors are often known by their brand names such as SuperCaps or GoldCaps. They all are based on the above-described technology.

The self-discharge of double-layer capacitors is in the range of 5%/day at 20 °C. Especially at higher temperatures, the self-discharge rate (approximately doubling of the self-discharge rate with a 10-K temperature increase as in all electrochemical systems) is hardly acceptable for autonomous power supply systems.

The electrical characteristics are dominated on one hand by the low inner resistance (resulting in high power) and on the other hand by the linear decrease in voltage with the state of charge. On one hand, this allows easy estimation of the state of charge, but on the other hand the voltage drop is very high and increases the requirements of the electronics or limits the usable energy from the double-layer capacitor (e.g. operation only between 1.7 and 2 V).

Today, double-layer capacitors are available in units of up to some thousand farads. Their gravimetric and volumetric energy density is very low (Figure 20.7), but they may have power

densities up to 5000 W/kg. Therefore, double-layer capacitors are most suited to applications with very high power requirements and low energy demand. As double-layer capacitors are a new and emerging technology, it is difficult to give definite cost figures. For orientation purposes, a cost of approximately 50 000 €/kW h can be estimated today. However, to supply a current of 200 A at 2 V for 2 s the cost is approximately €10 for the storage.

For autonomous power supply systems, double capacitors are an interesting technology in applications with peak power demand or for smoothing of power flow. These are, for example, pumping systems where pumps have a very high power demand to overcome the initial inertia. Another application might be grid-connected PV inverters with power quality control functionality. They are more efficient with a millisecond storage system. As a rule of thumb, it can be assumed that double-layer capacitors can find their place in applications with discharge times of less than 10 s per cycle (for 'power storage') or in combinations with conventional batteries. The big advantage of the capacitors is their almost unlimited number of cycles until the end of their lifetime (several hundred thousands).

20.4.7 The Lead–Acid Battery

Lead–acid batteries have been commercially used for more than 100 years for storing electrical energy. It has been the most widely used storage system for electrical energy for decades and still remains so. Lead–acid batteries cover a wide range of applications from SLI batteries in cars and trucks to uninterruptible power supplies, from load-levelling batteries for grid stabilisation to traction batteries (fork-lift trucks and others) and last but not the least autonomous power supply systems. Different battery designs have been developed for different applications to cover the various requirements.

Lead–acid is by far the cheapest battery type in comparison with all other readily available storage systems with appropriate characteristics according to the list given in Table 20.1. A major drawback of the lead–acid battery is the low specific gravimetric energy content due to the high atomic weight of lead. However, this is not a parameter of major importance in autonomous power supply systems as the battery is stationary.

It is expected that the lead–acid battery will remain as the working horse for autonomous power supply systems for many more years, probably decades. Therefore, this chapter gives a deeper insight into the lead–acid battery. The lead–acid battery chemistry, a detailed description of the battery design, ageing effects and recommendations for the operating strategy are the main topic of the following sections.

20.4.7.1 Lead–acid battery chemistry

Lead–acid batteries in the charged state consist of a positive electrode with lead dioxide (PbO_2) and a negative electrode with lead (Pb) as the active materials. Both electrodes contain a support grid, which is made from hard lead alloys. Sulphuric acid (H_2SO_4) diluted to 4M or 5M is used as the electrolyte. The following reaction equations describe the main reaction:

Positive electrode	$PbO_2 + 3H^+ + HSO_4^- + 2e^- \rightleftarrows PbSO_4 + 2H_2O$	(20.5)
Negative electrode	$Pb + HSO_4^- \rightleftarrows PbSO_4 + H^+ + 2e^-$	(20.6)
Cell reaction	$Pb + PbO_2 + 2H^+ + 2HSO_4^- \rightleftarrows 2PbSO_4 + 2H_2O$	(20.7)

PbO_2 and Pb are both converted to lead sulphate $PbSO_4$ during discharging (double sulphate theory). Sulphuric acid as the electrolyte is used up during the discharging of the battery. Therefore,

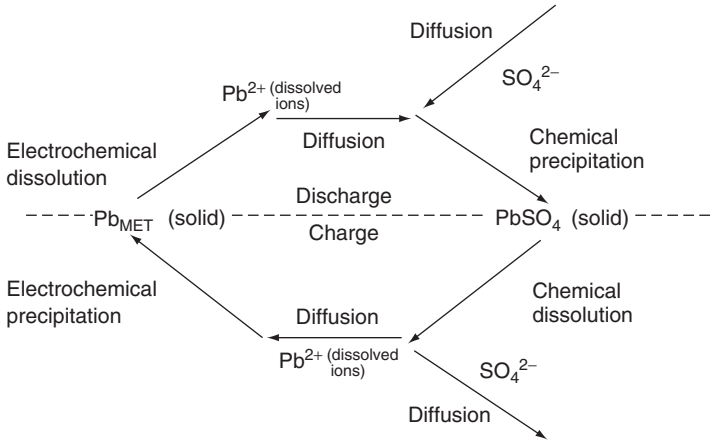


Figure 20.10 Schematic of the charge/discharge process in the lead electrode of a lead–acid battery

the concentration of the sulphuric acid decreases linearly with the state of charge. This is an important difference with respect to almost all other battery types, where the electrolyte has only the function of an ion conductor. In lead–acid batteries, it is in addition the source for the ions to counterbalance the charge dissolved in the electrolyte from the electrochemical process. Therefore, the electrolyte is subject to 'structural' changes, like the electrode materials themselves. This is an important reason for several battery characteristics and ageing effects as will be discussed later.

In Section 20.2.1.2, it was described that during charging and discharging not only the electrochemical process described by the Bulter–Volmer equation occurs, but a chemical process takes place as well. Figure 20.10 shows a schematic of this overall process for the lead electrode which is described by Equation (20.6).

The charged electrode consists of lead (Pb) in the solid state. When a discharge current occurs, two electrons are withdrawn from the metallic lead and dissolution of Pb^{2+} ions into the electrolyte occurs. Through diffusion, the charged ions are transported away from the reaction surface. As the charged ions unbalance the number of positive and negative ions in the electrolyte, negatively charged ions are necessary to counterbalance the positive surplus. They are provided as SO_4^{2-} ions from the sulphuric acid electrolyte. The SO_4^{2-} ions are transported by diffusion from the free electrolyte volume to the reaction site of the electrochemical reaction. There, the Pb^{2+} and the SO_4^{2-} ions meet and form $PbSO_4$ by a chemical precipitation process. This finally results in the formation of $PbSO_4$ crystals.

During charging, the reverse process takes place. Pb^{2+} ions are taken from the electrolyte to form solid Pb during the electrochemical precipitation process. These ions are transported by diffusion processes to the reaction site. To stabilise the Pb^{2+} ion concentration in the electrolyte, a chemical dissolution process of the $PbSO_4$ crystals takes place. Because the positive ions are removed from the electrolyte through the electrochemical precipitation process, the SO_4^{2-} ions need to be transported away from the reaction site to assure electrical neutrality.

All these processes cause overvoltages.

1. Electrochemical dissolution with respect to precipitation described by the Butler–Volmer equation.
2. Transport of Pb^{2+} ions described by the diffusion law resulting in diffusion overvoltages.

3. Transport of SO_4^{2-} ions described by the diffusion law, law of migration of charged ions in an electrical field and fluid dynamics caused by the change in the pore volume during charging and discharging resulting in diffusion overvoltages.
4. Chemical precipitation or dissolution of the $PbSO_4$ crystals forced by deviations of the ion concentration in the electrolyte from the equilibrium concentration resulting in concentration overvoltages.

All processes depend on the temperature. Further, the processes depend on the electrolyte concentration. The concentration influences the equilibrium current density of the Process 1, the diffusion rate of ions in Processes 2 and 3 and it has a strong impact on the equilibrium concentration of the Pb^{2+} ions and hence on the Process 4. Ageing of the battery and the operating conditions (high currents, small currents and pulse currents) have a significant impact on the overvoltages caused by Processes 1 and 4. This is mainly caused by changes in the inner active surface on the charged active material side (Pb) as well as on the $PbSO_4$ side.

The nominal voltage of a lead–acid battery is 2.0 V; the open-circuit voltage of a charged battery is about 2.1 V, depending on the concentration of the electrolyte.

The open-circuit potential of the positive electrode in a fully charged battery is approximately +1.75 V against the standard hydrogen electrode. The negative electrode potential is approximately −0.35 V against the standard hydrogen electrode. The relationship between the electrolyte concentration and the electrode potential with respect to the cell potential can be seen from Figure 20.14. The dependence of the open-circuit potential on the temperature is as small as 0.2 mV/K. Therefore, it can be neglected for practical reasons.

In addition, there is the main side reaction – the water electrolysis. As the electrolyte is aqueous and the cell voltage is approximately 2 V and can be as high as 2.5 V, water electrolysis takes place continuously. Hydrogen and oxygen are produced at the negative and the positive electrodes, respectively. Hydrogen production starts at electrode potentials more negative than 0 V against the standard hydrogen electrode. Oxygen evolution starts at electrode potentials more positive than 1.23 V. Fortunately, the so-called overvoltages at lead electrodes for the gas production are very high and therefore the gas production is inhibited to a high extent. This allows the lead–acid battery to be stable even at the high cell potential of 2 V. The self-discharge rate caused by the gassing is approximately 2–5% per month.

Positive electrode	$2H_2O \rightarrow O_2 + 4H^+ + 4e^-$	(20.8)
Negative electrode	$4H^+ + 4e^- \rightarrow 2H_2$	(20.9)
Cell reaction	$2H_2O \rightarrow 2H_2 + O_2$	(20.10)

Very comprehensive handbooks on lead–acid batteries have been written by Bode [15] on the fundamentals and by Berndt [3] on valve-regulated batteries.

20.4.7.2 Lead–acid batteries: technology, fundamentals, concepts and applications

All the different types of lead–acid batteries discussed in the following text are based on the reaction equation presented above. Whereas Figure 20.11 shows a complete battery system in an autonomous power supply system, Figure 20.12 shows the schematic construction of the electrochemical $Pb/H_2SO_4/PbO_2$ cell.

Solid lead grids, rods or plates serve to conduct the current (grid) and to mechanically stabilise the porous active mass in both electrodes. Depending on the battery type, different lead

Figure 20.11 A battery in a photovoltaic system with a rated capacity of $C_{10} = 37.5\,\mathrm{A\,h}$ and a rated voltage of 168 V (flat-plate technology, 28 blocks of 6 V connected in series, Picture courtesy Fraunhofer ISE)

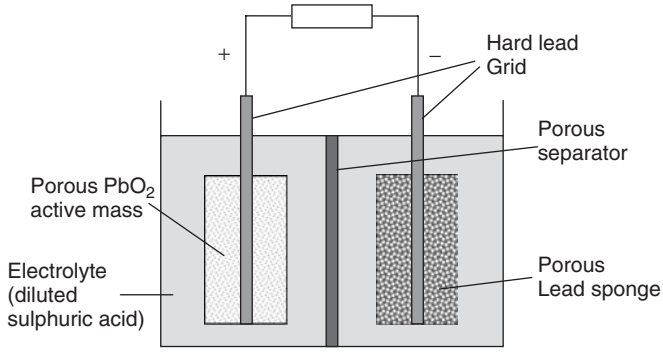


Figure 20.12 Schematic construction of a $Pb/H_2SO_4/PbO_2$ cell

alloys for the grid are used to increase the stability, improve the tooling properties and to reduce corrosion. The porous active material is attached to the grid PbO_2 (lead dioxide) for the positive electrode and to the Pb sponge for the negative electrode. Figure 20.13 shows a more detailed view of the structure of the active material. The active material has an internal surface area of approximately $0.5 - 5\,\mathrm{m^2/g}$ for the negative and the positive electrode in the fully charged state. The reaction speed and thus the charging and discharging properties are determined by the internal surface. Both electrodes are completely immersed in diluted sulphuric acid (H_2SO_4). As described earlier, the sulphuric acid plays a double role as the ion conductor between the electrodes and a reagent in the charge and discharge reactions. While the battery discharges, the PbO_2 and Pb are converted to $PbSO_4$ (lead sulphate). The sulphate ions (SO_4^{2-}) are drawn from the electrolyte, causing the electrolyte concentration to fall.

It is worth mentioning that the ratio of the specific volumes of $PbSO_4$ and Pb is 2.4 and the ratio between $PbSO_4$ and PbO_2 is 1.96. This means that the solid-phase volume of the active materials increases during discharge. This reduces the free electrolyte volume in the pores and causes mechanical stress on the active material.

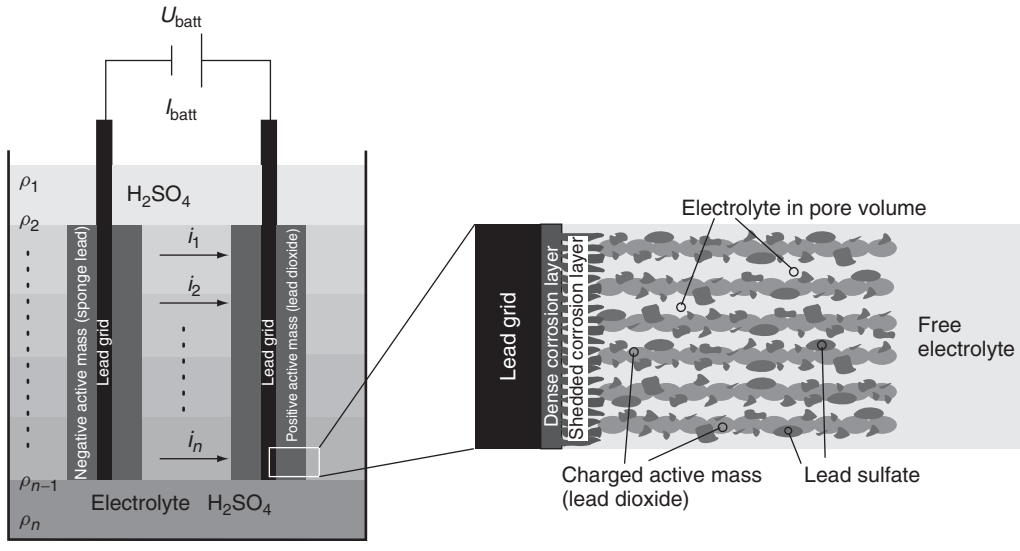


Figure 20.13 A more detailed schematic drawing of the lead–acid battery. The left-hand side shows a macroscopic view of the cell including effects like acid stratification represented by the different electrolyte densities in different horizontal heights of the battery followed by inhomogeneous vertical-current distribution within the electrodes. The right-hand side shows a 'microscopic' view of the active material in a partial state of charge

A separator is located between the electrodes, intended to prevent short circuits between the electrodes.

The above-described water electrolysis increases significantly as a function of voltage and temperature. As a rule of thumb, an increase in the so-called gassing rate by a factor of two is caused by an increase of 10 K in the temperature and by a factor of 3 by an increase in the cell voltage by 100 mV.

Regarding the hydrogen and the oxygen created as a result of the electrolysis reaction, two different technologies can be distinguished. In so-called flooded batteries, the electrolyte is in the liquid phase. To allow the gases to emerge from the battery, batteries with liquid electrolyte are not sealed gas tight. However, this results in a decrease in the water content of the battery and therefore the electrolyte level decreases and the concentration of the sulphuric acid increases. The water loss needs to be compensated during the maintenance that should take place once or twice a year. Deionised water must be used for refilling and not sulphuric acid or tap water.

The so-called valve-regulated lead–acid (VRLA) batteries are sealed gas tight with a valve. The valve allows the release of gas only in the case of overpressure in the battery. In normal operation, the gas is recombined to water within the battery. This effect is achieved by an immobilisation of the electrolyte. Two different techniques are state of the art: the electrolyte is transferred into a viscous gel by adding SiO_2 to the electrolyte or the electrolyte is absorbed within a highly porous glass matt (absorbed glass matt type, AGM). In both cases, the oxygen can pass through the electrolyte to the negative electrode. The recombination of oxygen and hydrogen occurs at the negative electrode. In VRLA batteries, the electrolyte is not in the liquid phase and therefore no spillage of electrolyte occurs in case of any break of the case or other accidents.

However, if the battery is incorrectly overcharged, more gas is generated than can recombine, so the gas must be able to leave the cell through the valve if an overpressure builds up. At the

same time, the valve must prevent ambient air from entering the battery. As these batteries cannot be refilled with water, blowing off the gas from the cell must be reduced to a minimum to prevent the cell from drying out. As a rule of thumb, after more than 10% water loss the battery is at the end of its lifetime. The water loss can be estimated by the weighting of the battery.

To achieve low gassing rates in VRLA batteries, normally lead–calcium alloys are used for the grids. Flooded batteries use mainly lead–antimony alloys with less than 2.5% antimony (Sb). This is a good compromise among the beneficial effects of antimony grids (good grid conditions for casting and good contact of the active material to the grid result in low contact resistance) and the harmful effect of the reduction of the hydrogen overvoltage caused by the antimony. However, as gassing needs to be minimised in VRLA batteries, antimony grids are not an appropriate choice. Especially, in the early days of the VRLA batteries, the antimony-free grids caused a significant reduction in battery lifetime through the so-called antimony-free effect. This effect is described in the literature as premature capacity loss (PCL) [16].

While the rated capacity of a cell depends on the geometry and the number of parallel-connected electrodes, the rated voltage of a cell is 2.0 V. The open-circuit voltage U_0 of the cell

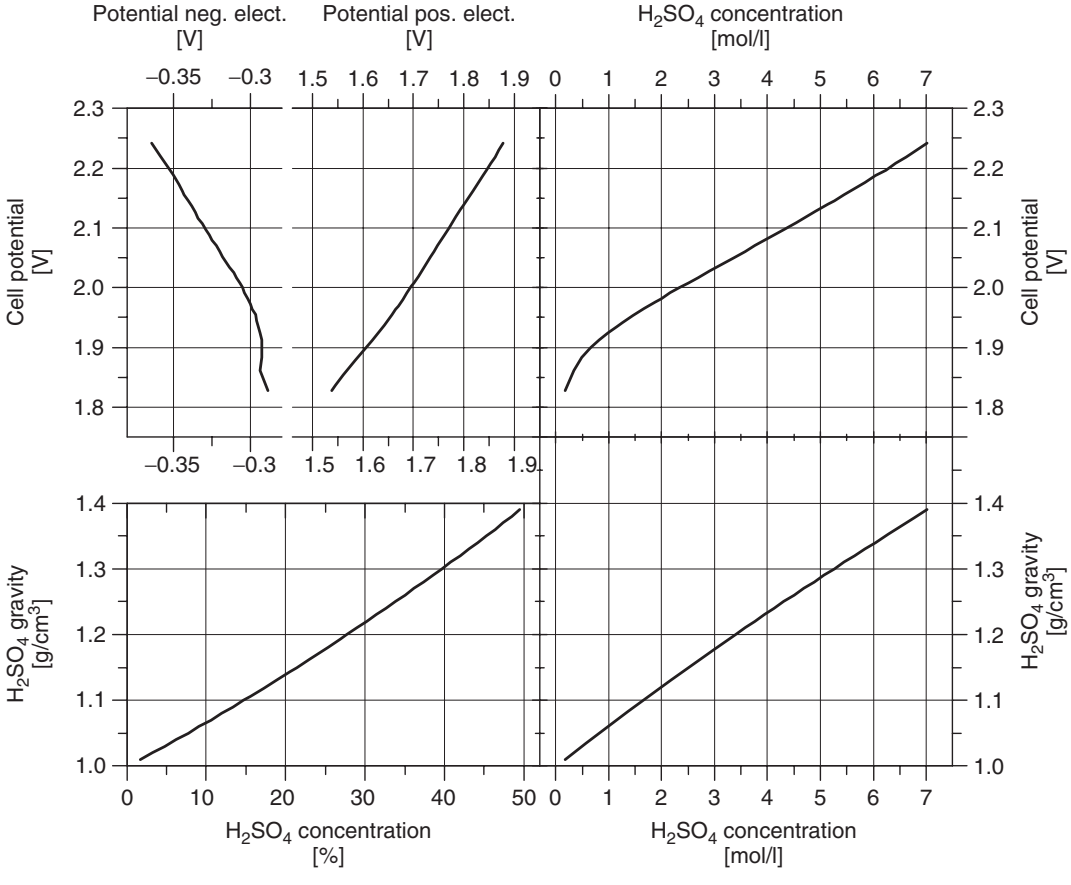


Figure 20.14 Correlation between the acid concentration in percent, mol/l or density and the equilibrium potential of the negative and positive electrodes and the cell potential. The graphs allow converting each value to the other

depends on the electrolyte concentration as shown in Figure 20.14, but for practical purposes the open-circuit voltage can be determined by the following rule of thumb:

$$\frac{U_0}{\mathrm{V}} = \frac{\rho}{\mathrm{g/cm^3}} + 0.84 \ldots 0.86 \tag{20.11}$$

where ρ is the density of the electrolyte. Electrolyte concentration and electrolyte density have an almost linear relation. As the electrolyte density can be easily measured, the electrolyte density is often used to express the electrolyte concentration. At 25 °C, 30% H_2SO_4 in H_2O has a density of about 1.22 g/cm^3 and 40% H_2SO_4 in H_2O has a density of 1.30 g/cm^3. Typical electrolyte densities in fully charged batteries are between 1.22 and 1.32 g/cm^3, depending on the application, the technological type and the climatic conditions. The acid density in the discharged state is between 1.18 and 1.05 g/cm^3. According to Equation (20.11), the open-circuit voltage also varies with the density. It is not a constant by any means.

Figure 20.14 shows the correlation between the electrode and cell potentials and the acid concentration. The acid concentration can be measured by means of the concentration in mol/l, the density in g/cm^3 and the percentage of acid in the solution $\%_{\mathrm{weight}}$. This allows the translation of all values to one another.

According to Equation (20.7), the electrolyte concentration decreases during discharge. According to Equation (20.11) the open-circuit voltage decreases in a manner directly proportional to the acid concentration. VRLA (sealed) batteries have less electrolyte per ampere hour capacity than flooded batteries. Therefore, the open-circuit voltage decreases more rapidly in sealed batteries during discharge than in flooded batteries. This must be taken into account, if the voltage is used as an SOC indicator.

Today, two different plate technologies are commonly used. The most common type of plate is the flat plate (Fauré type). The porous active mass is applied to the hard lead grid as a paste. The flat plate is simple and cheap to produce. The so-called tubular plates[9] are also widespread. A central lead rod, surrounded by active material, is inserted into a protective tube that is permeable to the electrolyte. A plate then consists of a row of adjacent tubes. While flat-plate electrodes have lower internal resistance and therefore higher specific power than tubular plates, the latter show more cycles in their lifetime. Manufacturing of tubular plates is more expensive than the manufacturing of flat plates. Figure 20.15 shows schematics of a flat-plate and a tubular-plate electrode. Currently, wounded lead cells come into the market consisting of a thin lead foil pasted with active material and with a very thin glass-mat separator between the electrodes. These batteries are primarily used in high-power applications like the ignition of motors. The first results show good cycle-life behaviour.

Lead–acid batteries are used today in many different applications, therefore a large variety of application-specific batteries are in the market. They can be distinguished by their specific power, their cycle life and their float lifetime. Float lifetime is the relevant parameter for batteries in uninterruptible power supplies where the batteries are subjected to only very few cycles in case of failure of the mains, but they should have long operating lifetimes while always being 100% charged. The main battery types for different applications and their typical operating conditions are as follows:

SLI (starting, lighting, ignition) batteries. Used for starting of engines; very high power capabilities even at low temperatures; traditionally only very low capacity throughput; subject to high temperature fluctuations; the largest global market for lead–acid batteries with respect to the capacity; production companies in almost all countries throughout the world; the operation profile is changing in modern cars as cars have a very high power demand beside starting owing to

[9] Tubular-plate electrodes are not very common in North America. Traditionally, tubular-plate batteries are more popular in Europe for cycling applications like, for example, fork-lift trucks.

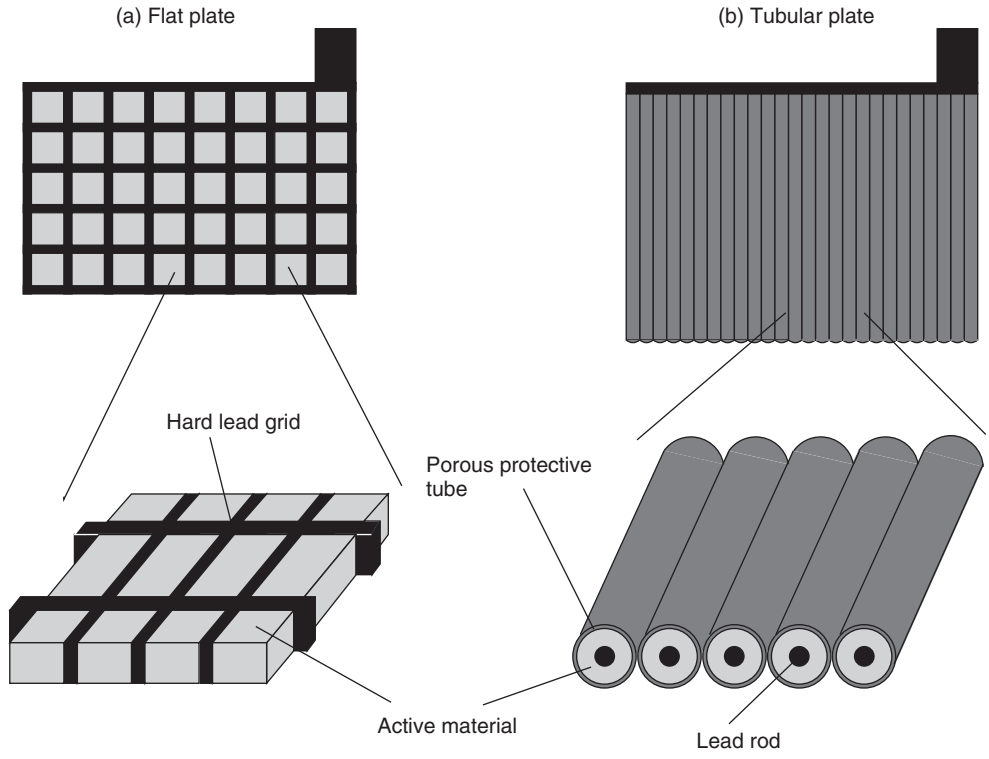


Figure 20.15 Schematic diagram of (a) a flat-plate electrode and; (b) a tubular-plate electrode

numerous electric applications in cars such as seat heating, electric window lift or HiFi systems. The batteries are made from very thin flat-plate electrodes to achieve high power. SLI batteries are a mass product, highly automated and are therefore very cheap.

UPS (uninterruptible power supply)/stationary batteries. Long idle periods at full SOC; rapid discharge when required (discharge time in the range of 10 min to 1 h, in some applications even longer); designs for lifetimes of up to 20 years available and market grows strongly in connection with the expansion of telecommunications and computer systems. The electrodes are thicker than for SLI applications to withstand corrosion for long periods.

Traction batteries. Application in fork-lift trucks, traction engines, underground mining vehicles and so on; designed for daily complete cycling with moderate currents and regular and controlled complete charging and cycle lifetimes of 1000–2000 cycles with 80% DOD can be achieved. The most common electrode technology in these applications is tubular-plate. Flooded batteries show longer lifetimes than VRLA and are widely used.

Electric-vehicle batteries. Widely fluctuating current profile; partial recharging phases (regenerative braking); inadequate lifetimes to date; expanding market and strong competition from other types of battery technology (see Table 20.3). Low gravimetric energy density is a major drawback in this application. Lead–acid batteries based on thin pasted lead foils and wounded design are currently under development and are already available in the market from some manufacturers to serve hybrid vehicles[10] that are seen at the moment as a more realistic option than purely

[10] Hybrid vehicles have a conventional motor, but with less power than in traditional cars. Acceleration is supported by electric motors powered by the batteries. The batteries are charged during regenerative breaking and from the

battery-powered electric vehicles. Wounded cells have very high power capability and therefore can serve electric motors for accelerating and regenerative breaking.

Batteries for photovoltaic systems. Operating conditions corresponding to the load profiles illustrated in Section 20.3; complete charging very seldom and many partial cycles. Two classes of so-called solar batteries are in the market. One class is the modified SLI battery with typically thicker grids than those used in SLI batteries, quite cheap (often from local production in developing countries [17]) but with limited lifetime. The other class of 'solar batteries' are modifications from high-quality batteries originally used for cycling or standby applications. In general, flooded batteries show better lifetimes in autonomous power supply systems than VRLA batteries. On the other hand, VRLA batteries have significant advantages concerning electrolyte spillage, maintenance and transport and very little release of corrosive and explosive gases. This reduces the requirements on the battery housing significantly. Therefore, a final choice must be made according to the specific application and the boundary conditions.

Operational experience, however, reveals that the lifetime of batteries in stand-alone applications based on solar energy is in general unsatisfactory compared with battery lifetimes in traditional applications. Batteries in solar home systems normally have to be exchanged after 2–3 years and batteries in hybrid systems after 3–8 years. Lifetime extensions to 5 years in solar home systems and 10 years in hybrid systems are achievable with advanced batteries designed for the purposes of autonomous power supply systems and appropriate system designs and operation strategies.

20.4.7.3 Discharge capacity

The capacity that can be withdrawn from lead–acid batteries depends strongly on the discharge conditions.

For stationary batteries usually the C_{10} or C_8 capacity, for starter batteries usually the C_{20} capacity and for traction batteries usually the C_5 capacity is specified. Solar batteries often are rated as C_{100} and C_{120} at 100 or 120-h discharge current, respectively. Typical end-of-discharge voltage is 1.8 V/cell or 1.85 V/cell for C_{10}, C_{20} and C_{100}. For C_5, 1.7 V/cell is commonly used. All other ratings and voltage limits can be found as well.

The measured and the practical capacity increase when the discharge current decreases. If a battery is discharged with a lower current than the rated current, a higher capacity than the rated capacity can be withdrawn. If the state of charge is specified with respect to the rated capacity (a reasonable convention), negative values for the state of charge can arise. This is the reason Figure 20.6 displays negative states of charge.

Figure 20.16 shows the voltage during discharge as a function of the discharged capacity at different discharge currents. The lead–acid batteries' capacity depends very much on the discharge current.

As a rule of thumb, it can be assumed that a battery with a nominal capacity of 100 A h at C_{10} has approximately 50 A h at C_1 and approximately 130 A h at C_{100}. Please note the fact that the corresponding currents I_1 and I_{100} are not equivalent to $10 \times I_{10}$ and $0.1 \times I_{10}$, respectively. In this example, I_1 is 50 A and I_{100} is 1.3 A.

With respect to the electrical properties, the temperature influences the inner resistance (increasing conductivity of the electrolyte with increasing temperature), the diffusion processes and the reaction at the electrochemical double layer. Therefore, the capacity of lead–acid batteries depends

main motor. This concept allows the conventional motor to run with small variations in power and therefore at higher efficiencies. Fuel consumption has been reduced with this concept in prototypes down to 2 l/100 km.

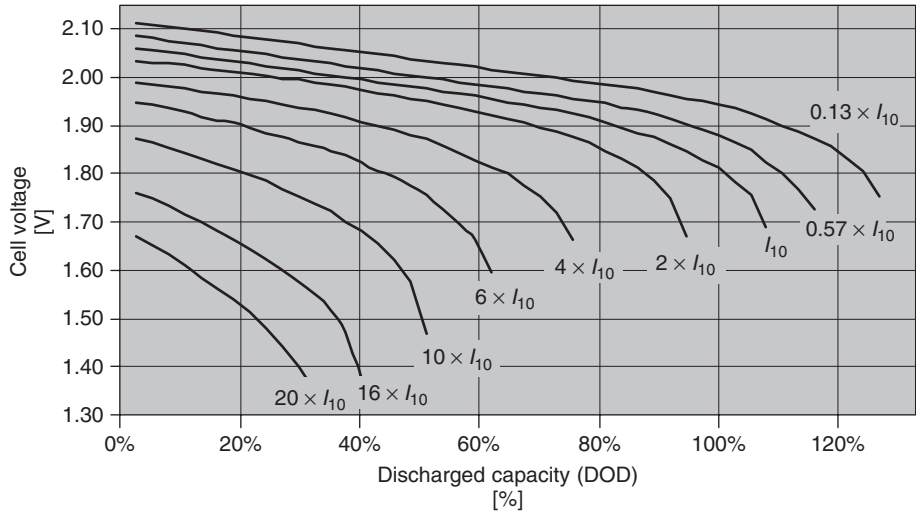


Figure 20.16 Voltage during constant discharge as a function of the discharged capacity at different discharge currents (tubular-plate lead–acid battery, C_{10} capacity defined at I_{10} and 1.8 V), data from Berndt D, *Blei-Akkumulatoren (Varta)*, VDI-Verlag, 11. Auflage, Düsseldorf (1986) [18]

strongly on the temperature. The capacity increases almost linearly by approximately 0.6%/K with an increase in temperature. The capacity decreases in the same way with decreasing temperatures. Depending on the battery technology, the temperature coefficient can be as high as 1.0%/K.

20.4.7.4 Ageing processes and their influence on the battery properties

A comparison with other battery systems, for example, NiCd batteries, reveals that the relatively short lifetime of lead–acid batteries is a significant disadvantage. However, the operating behaviour of the batteries is already affected during the lifetime by the processes responsible for ageing. Thus, it is helpful to be aware of the most important ageing processes and to avoid the conditions that accelerate them.

20.4.7.4.1 Acid stratification

Acid stratification is not an ageing process but affects the operating behaviour of the battery. On one hand it reduces the available capacity and changes the current/voltage characteristics. On the other hand it leads to inhomogeneous current distribution at the electrodes. The latter effect accelerates sulphation (Section 20.4.7.4.2), which is a major ageing effect. Therefore, acid stratification is a reason for ageing and not an ageing effect by itself. Acid stratification in flooded batteries can be removed immediately by stirring the electrolyte.

Because the electrolyte functions as an active component of the electrode reaction, local variations in the density can arise, with the result that the acid density decreases in the upper part of the cell and increases in the lower part. The potential difference associated with the concentration difference leads to discharging of the lower section of the electrodes, which can result in irreversible ageing effects (e.g. sulphation).

In batteries with liquid electrolytes, the acid stratification can be eliminated by deliberate overcharging, associated with gas production. The same mixing effect can be achieved by active

circulation of the electrolyte. Such electrolyte stirring systems are made from an air bubbling system (Figure 20.22).

Electrolyte stratification can also occur in batteries with immobilised electrolytes. In a gel battery, the effect is very small and therefore is of little relevance. In AGM batteries, the strength of acid stratification depends very much on the quality of the glass mat. Large battery cells from AGM technology are mounted vertically to avoid any acid stratification. While purchasing the batteries, it is important to check the manufacturer's specifications regarding vertical installation of the battery. The problem with VRLA batteries is that an existing acid stratification cannot be removed.

From theoretical consideration [19, 20], it is obvious that small currents in conjunction with acid stratification lead to a significant undercharging of the lower part of the electrodes. This effect is getting more and more pronounced with smaller battery currents. When an acid stratification occurs, the upper part of the electrodes is charge preferential and the lower part is discharge preferential. This results in differences in the local state of charge between the upper and the lower part of up to 30%. During the limited charging times, the upper part can reach a very high state of charge while the lower part is by far not completely charged. This means that the lower part of the electrode is cycled in lower states of charge than from the average state of charge of the electrode. Further, the lower part is cycled without a full charge for extended periods.

These findings are confirmed twofold by experimental results. On one hand, it was possible to show experimentally the effect of inhomogeneous current distribution and the state of charge within the electrode as a function of the charge/discharge currents [20]. On the other hand, almost all physicochemical analysis of lead–acid batteries from PV systems at the end of their lifetime shows a high degree of sulphation in the lower part of the electrodes [21, 22].

Another problem related to acid stratification in batteries with liquid electrolytes is that a measurement of the acid density, made at the only accessible position above the electrodes, does not give any direct information on the battery's state of charge.

As an example, Figure 20.17 shows the correlation between the acid density above the electrodes in a flooded battery and the battery's state of charge for a battery from a PV system, with a large number of partial cycles.[11] Without acid stratification, a measured acid density of, for example, 1.18 g/cm^3 corresponds to a real state of charge of approximately 30%, whereas it can range between approximately 30 and 75% if the acid is stratified. This measurement only allows a lower limit to be estimated. A measurement of the acid density above the electrodes can lead to appreciable errors in determining the state of charge and thus in associated operation-management measures.

20.4.7.4.2 Sulphation

When the electrodes are discharged, the active masses, PbO_2 and Pb, are transformed into $PbSO_4$. The size of the sulphate formed depends on the strength of the discharge current – high discharge currents result in small sulphate crystals. If a battery is not recharged soon after its discharge, the sulphate crystals grow as a result of recrystallisation processes. The rate of recrystallisation is linearly correlated with the solubility of sulphate ions. Unfortunately, the solubility of sulphate ions increases with decreasing acid concentration [15]. Therefore, periods of low states of charge (and hence low acid concentrations and high sulphate solubility) harm the battery by accelerating the growth of large sulphate crystals. During subsequent charging, large sulphate crystals with their relatively smaller active surface are redissolved more slowly than smaller ones, so that sulphate

[11] Figure 20.17 is based on a detailed battery model including modelling of the vertical acid-density distribution. The model was verified by measurements in a battery. The model and verification are described in [19]. Therefore, the state of charge and the acid density above the electrode displayed in Figure 20.18 are calculated by the model. The calculations are based on detailed measurements of the battery current, voltage and temperature in the system.

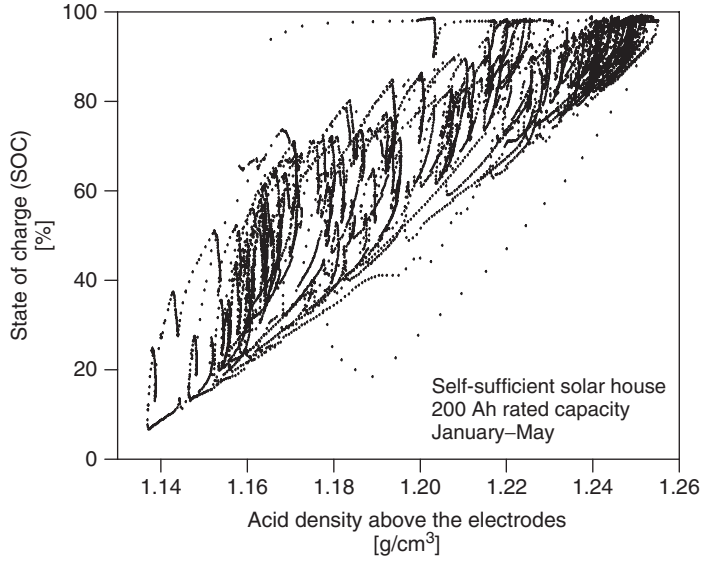


Figure 20.17 Acid density above the electrodes versus the actual state of charge, measured over five months for a battery from a photovoltaic system (200 A h cells, simulated acid densities based on measured initial data 10-minute average values)

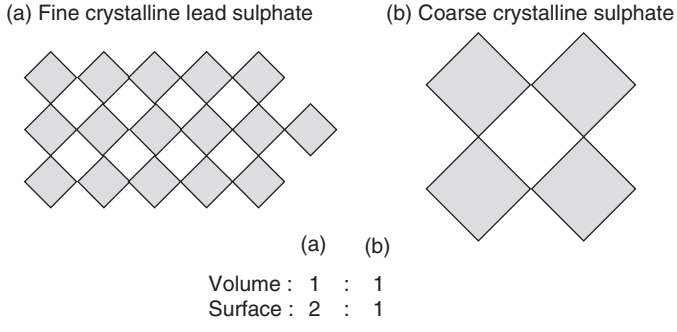


Figure 20.18 An example illustrating the effect of the crystal size on the active surface area of the electrodes. Mass ratio $a{:}b = 1{:}1$, surface area ratio $a{:}b = 2{:}1$

crystals are still present when charging is nearly finished. Figure 20.18 illustrates that for the same volume, small crystals have a larger surface area than large ones (two-dimensional representation of the three-dimensional effect).

During the course of the operation, these remaining sulphate crystals can accumulate, reducing the active mass and thus the accessible capacity [23]. Sulphation can be reduced to a minimum if each discharging process is rapidly followed by sufficiently complete charging. The effect of acid stratification is that complete charging is seldom achieved for the lower part of the electrode, so that strong sulphation occurs there. This sulphation effect can be clearly seen in the cross-sections of Figure 20.19.

As a result of sulphation, the amount of active material available for normal charging and discharging operations decreases. This reduces the capacity, and the voltage during discharge is

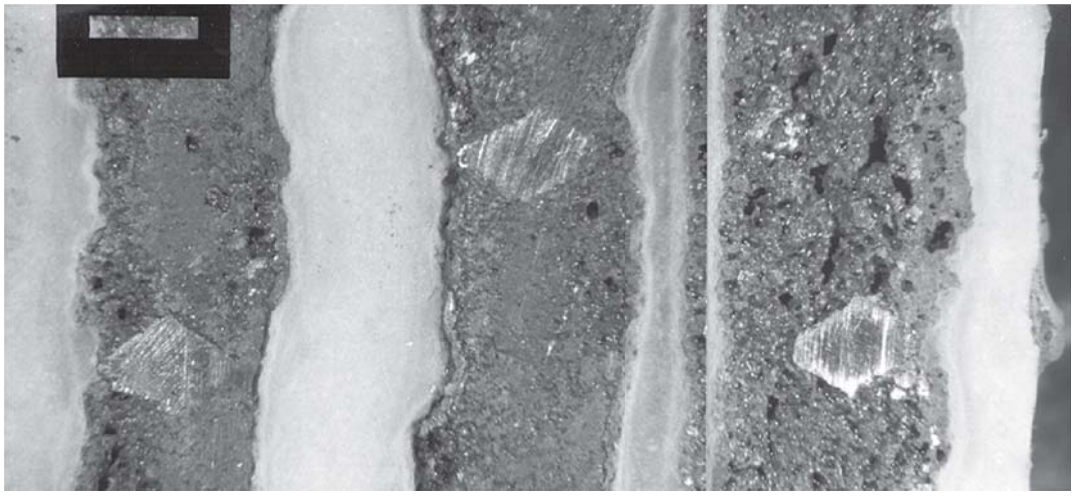


Figure 20.19 Cross-section of a negative flat-plate electrode after 3.5 years of operation in a photovoltaic system. The sections were taken from the upper, central and lower parts of the electrode (from left to right, respectively). Sulphation is clearly indicated in the lower section by the very coarse pores between large crystals, and by the noticeable broadening of the electrode, due to the difference in the specific volumes of Pb and $PbSO_4$ (Photo source: ZSW)

also shifted to lower values. If sulphation is too pronounced (as in the lower section shown in Figure 20.19), larger areas of the electrodes can become completely inactive.

20.4.7.4.3 Corrosion

The high positive potential at the positive electrode results in the corrosion of the lead grid [24]. On one hand, this causes the cross-section of the grid to decrease, so that the grid resistance increases. On the other hand, a layer consisting of lead dioxide, lead oxide and lead sulphate forms between the grid and the active material, which also raises the contact resistance. This becomes evident during charging and discharging as an increased ohmic voltage drop. Figure 20.20 shows the cross-section of a tubular electrode from a battery after 3.5 years of operation. The lead core (grid rod in the centre of the tube) has almost completely disappeared due to corrosion.

The corrosion rate depends on the acid density, the electrode potential, the temperature, the grid alloy, the active material coverage [25] and a most important factor, the manufacturing quality of the grid. Corrosion is particularly pronounced for cell voltages below 2.0 V and above 2.4 V [24]. Corrosion is minimal for cell voltages around 2.23 V. Corrosion is an irreversible ageing effect and increases the internal resistance of the battery. An indirect consequence is that the current distribution becomes more inhomogeneous in the vertical direction, so that sulphation is accelerated in the lower parts of the electrodes. In batteries for PV systems, thicker grids are used to reduce the effect of corrosion and thus extend the lifetime.

20.4.7.4.4 Erosion

Both electrodes are subjected to strong mechanical loads during cycle operation. The reason is that up to 50% of the active material is converted to lead sulphate during discharge. Lead sulphate has a volume per mole, which is 1.94 times larger than lead dioxide and 2.4 times larger than lead.

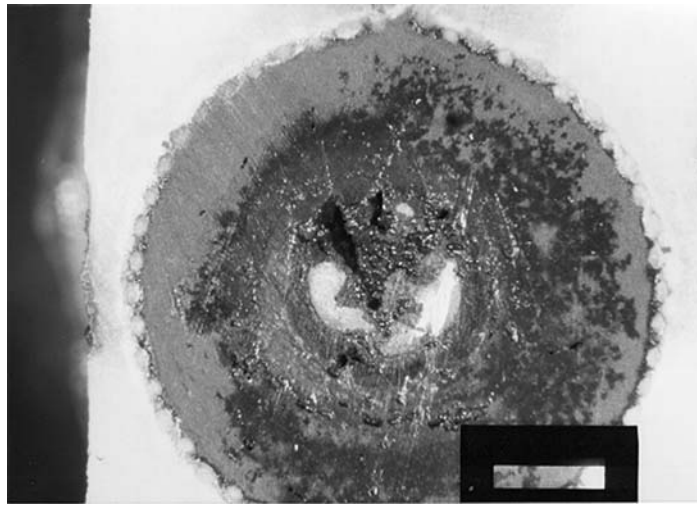


Figure 20.20 Cross-section through a tubular electrode of a battery after 3.5 years of operation. The lead rod (light-coloured area in the centre) has almost completely disintegrated as a result of corrosion. The diameter of the electrode is 8 mm (Photo source: ZSW)

These large changes in volume act to loosen the active material. This effect increases with increasing depth of discharge. Thus, deep discharge with low currents has an additional negative effect on sulphation. Once the active material has become loose, it can be separated from the electrode, for example, by gas movement, and gathers as sludge at the base of the battery. If the volume below the electrodes that contains the sludge is not large enough, there is a danger of short circuits between the electrodes.

The erosion effect is much less pronounced in sealed batteries, as the electrodes there can be mounted under pressure. The pressure compensates the forces arising from the volume change and increases the stability of the active masses.

The available active mass at the electrodes is reduced by the loss due to erosion. This corresponds directly to a reduction in the accessible capacity. Accordingly, the battery will be discharged earlier.

20.4.7.4.5 Short-circuits

In addition to the danger of short circuits in the sludge volume of batteries with liquid electrolytes, there are two further risks for short circuits.

The plate connectors from the positive electrode above the active material are also subject to corrosion. This results in detachment of large corrosion flakes, which can fall onto the electrodes and cause short circuits. This risk can be eliminated by including separators that extend upward well over the electrodes.

Further, there is a risk for all battery types that dendrites (microscopic short-circuits) may grow from the positive to the negative electrode through the separators. These dendrites are so fine that they are usually not visible even when the battery is investigated in the laboratory. Their growth is accelerated by long periods at a low state of charge and thus low acid concentrations. As described in Section 20.4.7.4.2, the solubility of $PbSO_4$ increases at low acid concentrations. For example, the solubility of $PbSO_4$ is 2 mg/l at a sulphuric acid density of 1.28 g/cm^3, and is already 35 mg/l

at a density of 1.02 g/cm^3. The higher the solubility the higher is the rate of recrystallisation and hence the rate of formation of large sulphate crystals and dendrites. The danger of dendrite growth can be counteracted by thicker separators and by operating the battery at high states of charge.

In general, a short-circuit is a defect that causes a sudden and complete breakdown in the battery. Short-circuits can occur from dendrites between the electrodes and the sludge in flooded batteries from active material accumulated as a result of active-material shedding under the electrodes. If the amount of material exceeds the free-electrolyte height under the electrodes, short circuits occur through the "mud". Microscopic short-circuits through the separator affect particularly the self-discharging properties of a battery.

20.4.7.4.6 Reverse charging

If a battery is subjected to a discharging current, even after the battery has been fully discharged, the potential changes its sign.

Reverse charging of an individual cell can occur in strings of series-connected cells. The battery voltage and therefore the depth of discharge protection are controlled on the basis of overall string voltage. A single cell within the string may have lower capacity due to manufacturing deviations or due to accelerated ageing. In consequence, the low-capacity cell can be over-discharged resulting in reverse charging.

In case of reverse charging, PbO_2 is formed at the original Pb electrode and vice versa. Although this type of reverse charging can sometimes increase the capacity for a short time, it is certainly detrimental to the cell lifetime in the long term. The main cause of damage is the oxidation of additives in the lead sponge of the negative electrode, which are included to maintain the high porosity of the electrode. If these additives are destroyed, large lead crystals form in the negative electrode resulting in a loss of internal surface area and therefore in an irreversible loss of capacity.

The danger of reverse charging can be reduced if the voltage of the individual cells is monitored and used as the criterion to end discharging (rather than the total voltage of the battery). Alternatively, it can be prevented by allowing charge equalisation between the cells of a series connection [12].

20.4.7.4.7 Ice formation

Figure 20.21 shows the dependence of the freezing point of diluted sulphuric acid on the electrolyte concentration. Ice must be prevented from forming in a battery under all circumstances, as it is then practically impossible to operate the battery (in particular, a frozen battery can hardly be charged) and it is possible that the cell housing may burst (battery breakdown and contamination of the surroundings with sulphuric acid).

For operating modes with very low discharge currents, or if the deep-discharge protection is inactive or non-existent, it is possible to reach extremely deep discharge, as the voltage does not break down until very late in the process. Batteries that are subjected to temperatures below the freezing point should be dimensioned such that after withdrawing $1.3 \times C_{100}$ or $1.7 \times C_{10}$, the acid density is still so high that freezing is not expected, according to Figure 20.21.

20.4.7.5 Battery peripherals

For proper battery operation, several battery peripherals must be used. The following gives a brief description of the most important devices.

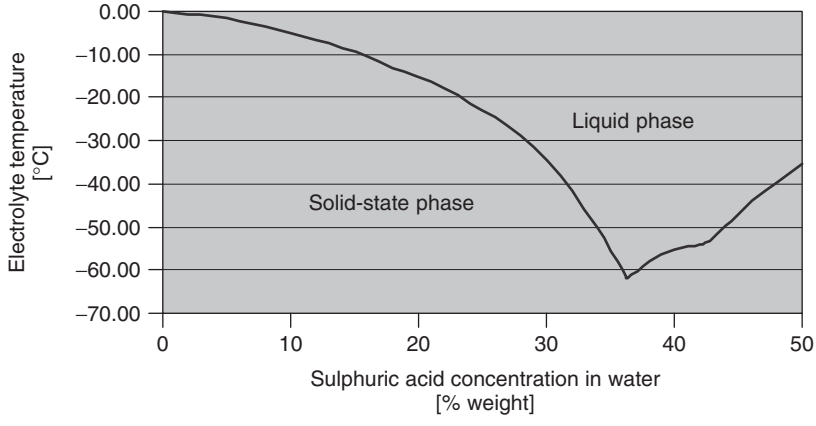


Figure 20.21 Freezing point of diluted sulphuric acid as a function of the acid concentration

Charge controller. Charge controllers are responsible for the charging strategies and the deep-discharge protection. They limit the power from the PV generator if necessary. More details on the operation strategies are given in Section 20.4.7.6 and on the hardware in Chapter 19.

Chargers. Chargers are AC/DC converters that use the power from motor generators to recharge the battery. They need a charge control as well to avoid overcharging of the battery. The charging regime should be the same as for charge controllers.

Charge equaliser. In long strings of series-connected cells, problems with individual cells like overcharging and reverse charging can occur owing to differences in the ageing processes or tolerances in the production. Charge equalisers avoid the detrimental effects by individual treatment of the cells. More details are given in [13] and in Chapter 19.

Monitoring. To get actual information of the state of the battery, monitoring systems can be used. A wide range of commercial products is available. They range from simple voltage monitoring of the complete battery to complete monitoring of temperature, current and voltage of individual blocks and cells as well as impedance of the battery.

State-of-charge meters. For proper battery operation (Section 20.4.7.6) and for the orientation of the user, it is helpful to have proper information on the actual state of charge of the battery. Several devices and algorithms are available, but only very few are really suited to autonomous power supply systems [26].

Electrolyte-agitation systems. To avoid the detrimental effects of acid stratification in flooded batteries, electrolyte-agitation systems are an effective solution. Figure 20.22 shows a battery with an agitation system. Most systems are based on compressed air that is pumped to the bottom of the cells. The ascending air bubbles cause an electrolyte circulation and mixing.

Recombinators. To reduce the loss of water from flooded batteries, recombinators are used. They consist of a catalyser that recombines the hydrogen and oxygen gas evolving from flooded batteries. Figure 20.22 shows a battery with recombinators.

20.4.7.6 Operation strategies

The operation strategy and the charging strategy have an important impact on the battery lifetime. Therefore, in the following paragraphs some ideas on appropriate strategies are discussed.

In most PV systems, the system and the battery control are realised through the charge controllers; in some cases energy management systems take over this job. The battery requires

Figure 20.22 Flooded, tubular-plate battery (2 × 240 A h in parallel connection, 12 V) with gas recombinators for reduction of water loss through gassing and an electrolyte-stirring system to avoid acid stratification (only connected to the right hand unit), the membrane air pump is mounted on the wall, operation of the pump: twice a day for 15 min (Picture courtesy Fraunhofer ISE)

proper handling of frequent full charges, gassing for electrolyte-stirring in flooded batteries, control of end-of-charge voltage and deep-discharge protection.

20.4.7.6.1 Charging

As a result of the specific operating conditions, full charging occurs very rarely. Charge and discharge cycles follow each other very frequently. Long charging times at constant power supply as available in grid-connected systems do not occur. Nevertheless, field experience showed that full charging is necessary to achieve long battery lifetimes.

The most common charging strategy is the constant current/constant voltage mode (*IU* or cccv, Figure 20.23(a)). In autonomous power supply systems, this means that the battery is charged with the fully available power until the battery voltage reaches the defined end-of-charge voltage. From this moment, the charging power to the battery is limited in a way that this voltage limit is not exceeded (constant voltage mode). The voltage drops at the moment when the battery charging is not high enough (due to decreasing power generation or increasing load) to maintain the battery voltage at the given limit. Most charge regulators and battery chargers use this charging procedure.

A more advanced charging method is shown in Figure 20.23(b). The charging starts with a constant current/constant voltage charging, but the maximum voltage is reduced after a certain

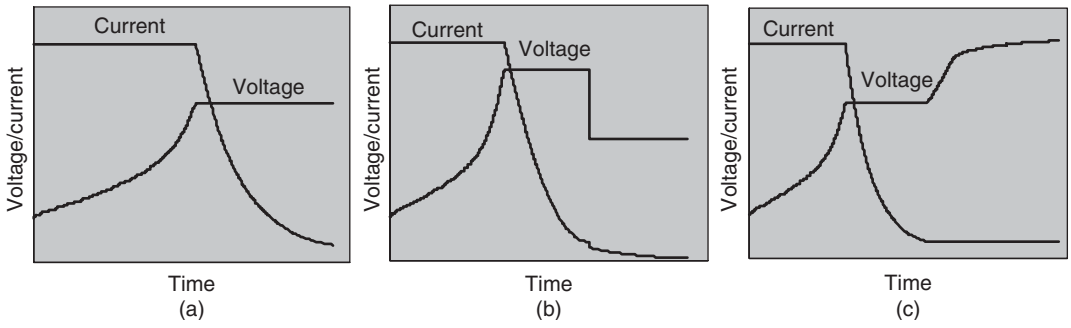


Figure 20.23 Schematics of different charge regimes. (a) Current and voltage during a constant current/constant voltage charge IU or cccv; (b) a constant current/constant voltage charge with two end-of-charge voltage limits IUU_0; and (c) a constant current/constant voltage charge followed by a limited constant current phase IUI_a are shown

time to a lower limit (IU_0U). This allows higher voltages during the first constant-voltage phase, but avoids negative effects like gassing and corrosion due to long durations of the high voltage. Therefore, this charging method allows overall faster charging but avoids hazardous conditions in the battery. More and more sophisticated charge controllers in the market use this charging procedure.

Especially for VRLA batteries, it turned out that longer lifetimes can be achieved with the constant current/constant voltage/constant current (IUI_a) charging as shown in Figure 20.23(c) [4]. After the current drops during the constant-voltage phase below the limit for I_a charging, the charging is continued for a limited time or amount of charge with a constant current. The voltage is not limited during this phase, but the I_a current must be limited in the range of I_{50} to I_{100}. No commercial device in PV applications uses such a scheme these days.

It is necessary to take into account that the batteries in PV systems hardly ever get fully charged due to the limited charging time per day [4]. Therefore, the term full charge has to be distinguished in a real full charge, defined by the point at which the complete active material is converted into charged material, and a practical or solar full state of charge (Figure 20.3). The latter is defined by the maximum state of material conversion that can be achieved during a sunny summer day or the maximum operation time of the back-up generator that is accepted by the system operator. A 'solar full charge' requires at least 5 h at a battery voltage of 2.4/cell.

In hybrid systems, a solar full charging can be achieved by operation of the back-up generator or from the PV generator. Full charging every four weeks is recommended. A detailed analysis of the operational data from systems showed that this has little impact on the overall energy balance, but is obviously enhancing the battery lifetime.

For batteries in systems of Class 1, according to Figure 20.6 an end-of-charge voltage of 2.4 V/cell is appropriate. However, the duration per day at this voltage should be limited to two hours. During the rest of the day (if the charging power is available), the battery voltage should be limited to 2.3 V/cell (charging regime as in Figure 20.23(b)).

In systems of Classes 2, 3 and 4, the end-of-charge voltage should be 2.45 V/cell also limited to 2 hours per day. An end-of-charge voltage of 2.35 V/cell for the rest of the day is appropriate (charging regime as in Figure 20.23(b)).

The values are valid for flooded and for VRLA batteries. In addition, for flooded batteries an increase of the end-of-charge voltage up to 2.6 V/cell periodically for a maximum of 5 hours per

14 days is appropriate. This causes gassing and therefore stirring of the electrolyte. Nevertheless, an active electrolyte-stirring system as shown in Figure 20.22 is the best and most efficient solution.

A significant lifetime extension can be achieved if the battery is charged to a really full SOC at least twice a year. This can be achieved by charging the battery normally to a solar-full SOC followed by a complete discharge with approximately I_{10}. The discharge must be followed by a recharge according to the IUI_a charging regime (Figure 20.23c) where the battery gets charged with 110 to 120% of the ampere hour capacity taken from the battery in the previous discharge or the nominal capacity (whatever value is higher).

The end-of-charge voltage limit depends on the battery-operating condition and on the temperature. All values for the voltage limits given here are for a battery temperature of 25 °C. At increasing temperatures, the voltage must be reduced by 4–5 mV/(K* cell) but not below 2.25 V/cell. At temperatures below 25 °C, the voltage must be increased accordingly but not above 2.6 V/cell. To protect DC loads or electronic devices connected directly to the DC bus bar, it might be necessary to limit the maximum voltage accordingly.

20.4.7.6.2 Deep-discharge protection

Lead–acid batteries suffer from deep discharge for several reasons. An increasing depth of discharge results in a decreasing acid concentration and due to the increased sulphate solubility in accelerated sulphation (Section 20.4.7.4.2), corrosion (Section 20.4.7.4.3) and higher sensitivity to freezing (Section 20.4.7.4.7). Further, the mechanical stress is increased because of the changes in the specific volume of the active materials and in long battery strings, the risk of reverse charging of single cells (Section 20.4.7.4.6) increases. Therefore, the maximum depth of discharge should be limited during normal operation.[12]

While choosing the appropriate DOD for the operation strategy, the data-sheet information given by the manufacturers should be analysed. They often give the number of cycles during the lifetime of a battery as a function of the depth of discharge. However, for the system design the number of cycles is not the most important parameter. The level of capacity throughput that can be realised during the battery lifetime is of more relevance. A cycle with 50% DOD means that only 50% of the capacity is used and therefore the overall capacity throughput for, for example, 200 cycles with 50% DOD is equivalent to 100 cycles with 100% DOD. However, from the point of view of the system design a battery which is limited to 50% DOD during normal operation must have double the size with respect to a battery with 100% DOD during normal operation. This is worthy of mention because batteries are always limited by the capacity throughput on one hand and by operation life on the other hand. Therefore, it makes no sense to operate a battery which is, for example, rated for 10 000 cycles at 20% DOD in autonomous power supply systems even though this might promise the highest capacity throughput. Assuming that on a daily basis 10 000 cycles take place, it would take more than 25 years to achieve this. However, the battery lifetime would not last that long due to other ageing processes.

Figure 20.24 shows for two different batteries the cycle life as a function of the DOD (data taken from data sheets) and the resulting capacity throughput. It is obvious that for the battery Type 2, the capacity throughput is almost independent of the DOD but for battery Type 1 there is a strong dependency leading to higher throughputs at lower DODs.

[12] As stated in Section 20.4.7.6.1, on charging, a complete discharge of the battery to 100% DOD twice a year is of benefit to the battery. This is not in contradiction to a limited DOD during normal operation. The defined discharge is done within a short time and is followed directly by complete recharging of the battery. In normal operation, discharge times and duration in deep states of charge can be very long and the next full charging may occur only weeks or month later.

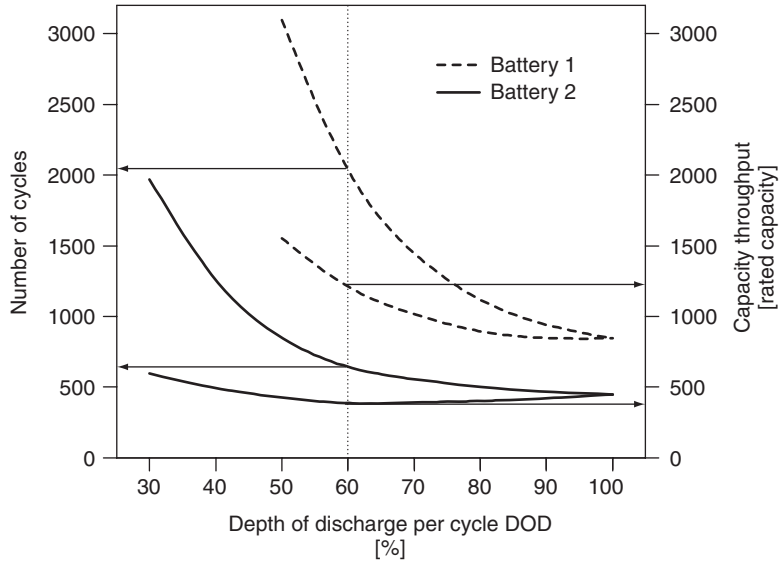


Figure 20.24 Number of cycles and overall charge transfer (capacity throughput) in units of the rated capacity during the battery lifetime as a function of the depth of discharge during cycling. Data from data sheets from battery manufacturers

For practical purposes, the following 'rules' can be used, which have proved their suitability in the field. In Classes 1 and 2 (Section 20.3.2), the maximum DOD should be 60–70% and in Classes 3 and 4, 80–90%. The lower values are for flooded batteries and the higher values are for VRLA batteries. Low-cost 'solar batteries' should be operated to a maximum of 50% DOD. It is very important to take into account that the mentioned values for the DOD are given on the basis of the C_{10} capacity. Using, for example, 80% of the C_{100} capacity means using more than 100% of the C_{10} capacity and this is hazardous.

Control of the maximum DOD can be realised either by deep-discharge disconnecting voltage or on the basis of the state of charge. Most commercial charge controllers control the maximum DOD by the voltage. The drawback of this method is that the discharged capacity up to a certain voltage limit depends very much on the discharge current. Table 20.5 shows typical end-of-discharge voltages up to which 100% of the C_{10} capacity has been discharged from the battery. This shows the problem of an efficient DOD control. If the maximum DOD is assured by a high

Table 20.5 Typical end-of-discharge voltages up to which 100% of the C_{10} capacity has been discharged from the battery at different discharge currents at room temperature

$I_{\text{discharge}}$	U [V/cell]
$1.0 \times I_{10}$	1.80–1.85
$0.5 \times I_{10}$	1.85–1.90
$0.2 \times I_{10}$	1.90–1.95
$0.1 \times I_{10}$	1.95–2.00

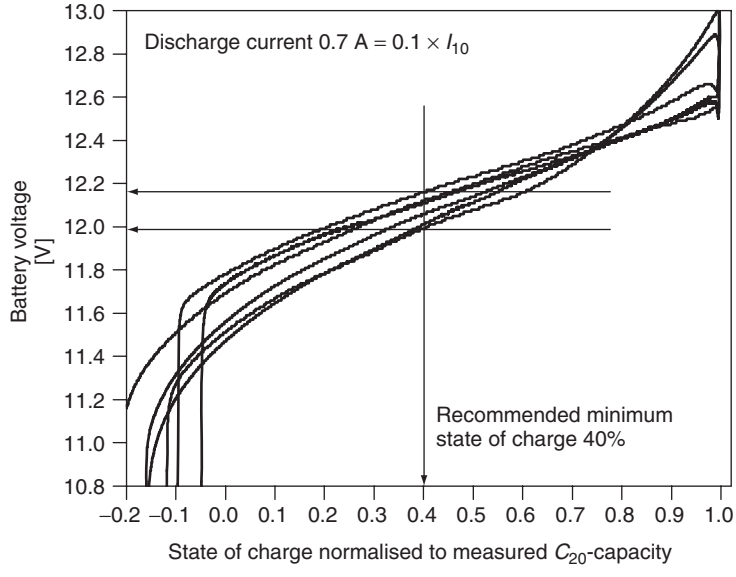


Figure 20.25 Discharge curves of six batteries rated for the same capacity, with the same technology (flat plate, flooded lead–acid and all batteries new). The voltage is shown as a function of the state of charge of the batteries. The state of charge is calculated on the basis of the discharge ampere hour and the nominal (not measured) capacity. Capacity tests were performed with $0.1 \times I_{10}$. The voltage is given for a 12-V block battery made from 6 cells connected in series (a typical design as used for SLI batteries for cars)

voltage limit even at small currents (e.g. 1.95 V/cell), the available capacity at higher currents is very limited [27]. Figure 20.25 shows a more pronounced version of the problem. A given voltage limit might be appropriate for one battery type, but for another with the same technology (flat plates, lead–acid and flooded electrolyte) the discharge curve looks quite different and the same voltage limit results in significant differences in the maximum DOD with respect to the minimum SOC defined to protect the battery as shown in the figure. The difference in SOC for the same voltage limit could be as high as 25%. Taking into account different battery technologies, the differences in SOC are even higher. In autonomous power supply systems, low and high currents occur (Figure 20.5, [9]).

Two solutions for the problem are available. One is a current-compensated end-of-discharge voltage threshold and the other is the use of state-of-charge determination. The latter solution is the most appropriate for autonomous power supply systems. There are various methods for state-of-charge determination in lead–acid batteries, which are appropriate for autonomous power supply systems [26].

20.5 SECONDARY ELECTROCHEMICAL BATTERY SYSTEMS WITH EXTERNAL STORAGE

The secondary batteries described in Section 20.4 use electrodes both as part of the electron-transfer process and to store the energy via electrode solid-state reactions. Consequently, both energy storage capacity and the power rating are intimately related to the electrodes' size and shape.

Electrochemical batteries with external storage overcome this drawback. The reaction occurs within an electrochemical cell and the energy is stored in two tanks separated from the electrochemical cell. The electrochemical cell has two compartments, one for each storage medium, physically separated by an ion-exchange membrane. This allows the designing of the battery power and the energy content separately.

Here, a distinction between the so-called redox-flow batteries in which salts are dissolved in liquid electrolytes and the hydrogen/oxygen storage systems based on the electrolyser and the fuel cell will be made.

The costs for the converters and therefore the power sizing are independent of the storage size. Therefore, these systems show an economy of scale concerning the energy storage. The larger is the storage, the lower are the specific storage costs. This makes the systems interesting for seasonal storage or other long-term storage applications.

20.5.1 Redox-flow Batteries

In redox-flow batteries, the active material is made from salts dissolved in a liquid electrolyte. The electrolyte is stored in tanks. As the solubility of the salts is typically not very high, the energy density is in the range of lead–acid batteries. The electrochemical charge/discharge reactions take place in the converter, which determines the power of the system. Therefore, redox-flow batteries belong to the group of batteries with external storage. Redox-flow batteries were already under investigation for stationary applications in the 1970s and 1980s. An overview of these activities can be found in reference [28]. Owing to problems with the materials, the investigations were almost stopped, but were again started in recent years.

Redox-flow batteries work with electrolytes in two circulations. Each circulation contains a redox system whose valence is changed during charging and discharging. The change in the valence of the two redox systems should take place at preferably high potential difference as this forms the equilibrium voltage of the battery. Figure 20.26 shows the principle of the redox-flow battery with the vanadium battery (Equation 20.14) as an example. The valence of all ions during each step can be seen in the figure.

Several different combinations of salts were and still are under investigation.

Fe-Cr	$Fe^{3+} + Cr^{2+} \rightleftarrows Fe^{2+} + Cr^{3+}$	(20.12)
Br_2Cr	$Br_2 + 2Cr^{2+} \rightleftarrows 2Br^- + 2Cr^{3+}$	(20.13)
Vanadium	$V^{5+} + V^{2+} \rightleftarrows V^{4+} + V^{3+}$	(20.14)
Regenesys	$3NaBr + Na_2S_4 \rightleftarrows 2Na_2S_2 + NaBr_3$	(20.15)

Several problems with redox-flow batteries have occurred and are still unsolved. The stability of the separator and the mixing of the electrolytes through the separators are severe problems. Therefore, the vanadium system became the centre of interest in the last few years as the materials and electrolytes are similar for the positive and the negative electrodes. Therefore, a crossing of ions through the separator just causes coulomb losses, but causes no deterioration of the electrolytes.

Defining the specific energy densities is difficult because of the independent sizing of the converter and the storage. Typical values for 20 kW/20 kW h vanadium redox-flow batteries are about 20 W h/kg and 50 W/kg. For mobile applications in electrical cars, this is not enough, but for stationary, especially in load levelling, applications it is an interesting option. Figure 20.27 shows a prototype of a redox battery at laboratory scale and Figure 20.28 shows a schematic of a redox-flow battery in the megawatt hour. As there have been no commercial products in operation for a long

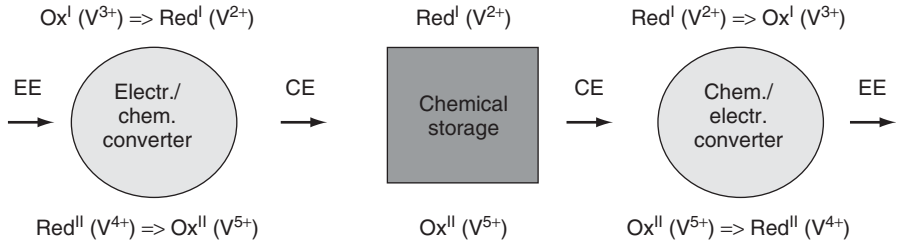


Figure 20.26 General concept of a redox-flow battery with two electrolyte/active mass circulations. The circulation in the upper row is equivalent to the negative electrode, the lower row denotes the positive electrode. In brackets, the vanadium battery is given as an example. The figure is based on an idea from [29] (EE: electrical energy, CE: chemical energy)

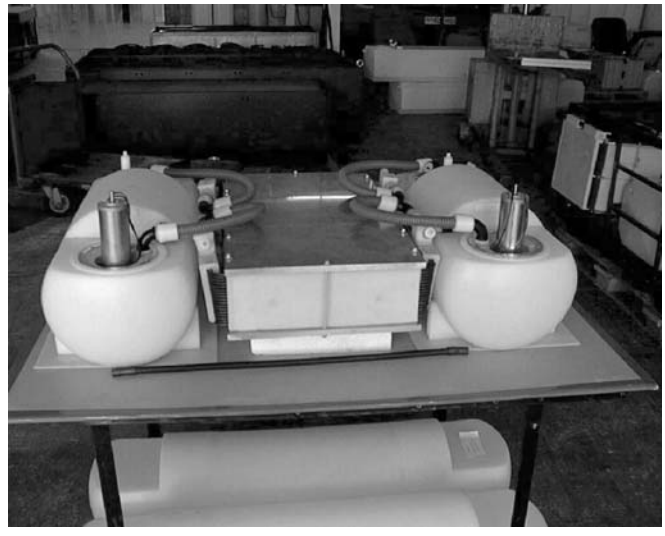


Figure 20.27 Prototype of a vanadium redox-flow battery with 32 cells and 14 Ah (Picture Courtesy by ZSW [29])

time, data on lifetimes are hardly available. Theoretically, long lifetimes can be expected as no part of the system undergoes structural changes as they occur in most other battery technologies. In the literature, data for a vanadium battery with more than 13 000 cycles have been reported [30]. In any case, a regeneration of the electrolyte/active mass is possible. The influence of vanadium batteries on the environment is described in [31]. No material loss or 'down cycling' of the electrolyte including the vanadium occurs.

What is true for the lifetime is also true for the costs. Rough estimations show, for the vanadium battery, costs of approximately 200 €/kW h for batteries with more than 20 h of discharge time at full power [29]. As no mass production is established, these data are subject to speculations on the mid-term achievable cost figures.

The energy efficiency for the vanadium battery has been demonstrated to be 80–85% for the cell itself. As the system requires peripherals (mainly pumps for the electrolytes), the system efficiency can be in the range of 75%, which is considerably better in comparison with the hydrogen

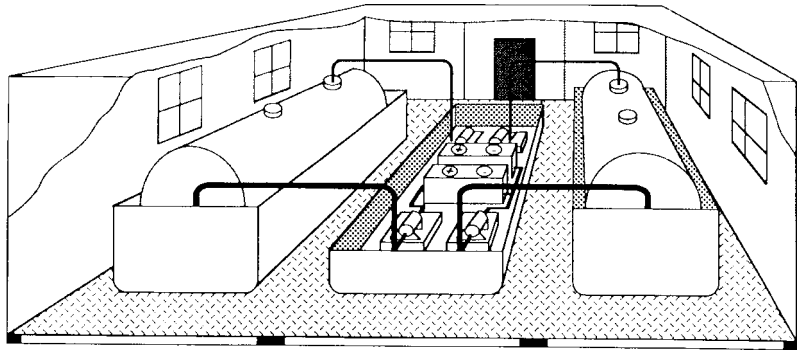


Figure 20.28 Schematic of a redox-flow battery in the megawatt hour range [29]

system described in Section 20.5.2. No self-discharge of the electrolytes in the tanks occurs. An optimum operation temperature for redox-flow batteries is defined by an optimum of the solubility of all the salts in the electrolyte while avoiding any recrystallisation of a salt.

20.5.2 Hydrogen/Oxygen Storage Systems

The other major technology with external storage is the use of liquid water (discharged state) and its gaseous components hydrogen and oxygen (charged state) [32].

Water can be split into hydrogen and oxygen gas by the fundamental reaction (Equation 20.16).

$$2H_2 + O_2 \rightleftharpoons 2H_2O \tag{20.16}$$

Electrolysis of water starts at 1.23 V/cell. A hydrogen storage system consists of three major components that are listed below:

1. An electrolyser for the production of the gases from electric power.
2. Gas storage for the hydrogen and, depending on the system's design, for the oxygen. Hydrogen is commonly stored either in pressure tanks or metal-hydride tanks.
3. A fuel cell to reconvert the gases into water and electric power. Hydrogen is taken from the storage and oxygen either from gas storage or from air.

Electrolyser/fuel cell systems have been demonstrated; the technology, however, still has to go through a long process of maturation until it reaches the market at acceptable costs and with high reliability.

20.5.2.1 Electrolyser

Electrolysers are needed to produce hydrogen and oxygen gas from water by using electric power. Electrolysis is used in many industrial processes apart from hydrogen production. Two technologies for low-temperature water electrolysis are available today:

- Alkaline electrolysers.
- Polymer electrolyte membrane (PEM) electrolysers.

In addition, high-temperature steam electrolysers are under investigation, which are in general good for higher efficiencies (see [33]).

Special electrolysers can release the gases already under pressure without using an additional compressor. PEM electrolysers can have efficiencies of 80–85%. They are commercially available, but are very expensive because only a few units are sold and expensive materials (membrane, catalysts) are needed. Presently, there is no important market for PEM electrolysers in the kW range. However, larger alkaline electrolysers for use with wind or water power plants are commercial and in operation. Reference [34] gives an overview on a research project on electrolysers operated by wind turbines.

20.5.2.2 Gas storage

Three major technologies for hydrogen gas storage are available today:

1. Pressure tanks (low pressure up to 30 bar, medium pressure up to 200 bar and high pressure up to 700 bar).
2. Metal hydrides with adsorption of hydrogen.
3. Liquid hydrogen storage (only for large scale applications).

Low-pressure tanks can be used in conjunction with pressure electrolysers. Specially designed electrolysers [35] produce hydrogen and oxygen at 30 bar without any compressor and therefore a minimum of energy loss due to compression. Medium-pressure tanks or bottles can be fitted to compressors. Today, hydrogen mechanical compressors with a high efficiency for small gas volumes are hardly available. Presently, no compressors for hydrogen with flow rates below approximately 10 Nm^3/h are available.[13] Another technology for gas compression is thermal compression with metal hydrides. The pressure in a metal-hydride storage unit increases significantly with an increase in temperature. As different metal alloys have different pressure/temperature curves, gas compression in a multi-stage process is possible. High-pressure gas bottles from composite materials are under development and in operation in R&D and demonstration projects.

Metal hydride is an interesting material for hydrogen storage. The hydrogen is adsorbed within the highly porous metal hydride. In fully loaded metal-hydride tanks, 1–2% of the overall weight is hydrogen. The volumetric energy density of metal hydrides is comparable to a 200 bar pressure bottle. The pressure in a metal-hydride tank depends on the temperature, the alloy and the state of charge. For outdoor applications, it is important to be aware that the pressure in the metal-hydride tank at constant hydrogen load approximately doubles with an increase of 20 K in temperature.

The energy content of 1 Nm^3 of hydrogen gas is approximately 3.5 kW h. Depending on the fuel cell system and power-converter efficiency, between 1 and 1.8 kW h of net energy can be drawn from 1 $N\,m^3$ of hydrogen gas.

A standard 200 bar pressure bottle contains 8.8 Nm^3 of hydrogen gas. The cost for metal-hydride storage systems is currently in the range of 500 to 1500 € per $N\,m^3$.

For oxygen storage, currently only pressure tanks are commercially available. Materials with adsorption properties for oxygen are under investigation. A reduction in volume by a factor of 3 has been achieved.

[13] $N\,m^3$ is the typical dimension for a gas amount. It is a gas in a volume of 1 m^3 at a pressure of 1 bar and a temperature of 0 °C.

Nanotubes for hydrogen storage are under investigation. After a very optimistic period some years ago, the optimism has been reduced, but meanwhile there are several activities to investigate this technology which promises very low costs.

20.5.2.3 Fuel cell

As fuel cells can only replace batteries in conjunction with the hydrogen gas generation, only the polymer electrolyte or proton exchange membrane fuel cell (PEMFC) is considered here. A very comprehensive overview of all the fuel cell technologies is given in reference [36].

The basic reactions are given in Equations (20.17) and (20.18).

$$\text{Anode reaction} \qquad H_2 \rightarrow 2H^+ + 2e^- \qquad (20.17)$$

$$\text{Cathode reaction} \qquad {}^1/_2O_2 + 2H^+ + 2e^- \rightarrow H_2O \qquad (20.18)$$

Figure 20.29 shows a schematic of a PEM fuel cell. If hydrogen and oxygen are stored, a closed-loop operation with water and gases can be realised. Then, the water demand for refilling is limited. If air is used instead of pure oxygen, the water produced in the fuel cell process gets lost with the air throughput. The water household of the membrane is one of the most challenging problems in fuel cell operation and control.

PEMFCs operate best at temperatures between 60 and 90 °C. The fuel cell stack itself can operate at an efficiency of 50–60%. The overall fuel system has additional components beside the stack like air compressors, electronics, valves and security devices. They cause a self-consumption of the fuel cell system and therefore reduce the overall efficiency to 35–50%. The efficiency is higher if pure oxygen is used instead of air, but this requires an additional oxygen tank. The stack efficiency is calculated from the ratio of the fuel cell voltage during power generation and the electrochemical potential of the reactants, which is 1.23 V. The coulomb efficiency is considered as 100% even though some gas losses synonymous with coulomb losses due to the penetration of gas through the membrane occur as well.

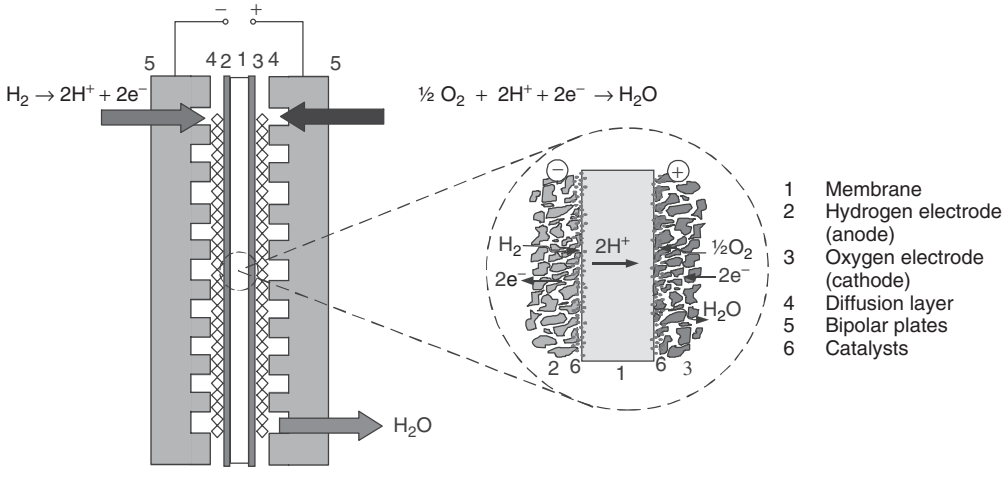


Figure 20.29 Schematic of a polymer electrolyte membrane fuel cell (PEMFC)

A big advantage of fuel cells in comparison with motor generators are the high efficiency, even at partial loads. Usually the stack efficiency increases with decreasing power output.

The current/voltage characteristic of a fuel cell is quite similar to the characteristic of a PV cell. Unlike secondary electrochemical batteries, the voltage depends very much on the current. To supply a load with a constant voltage, power electronics are necessary. The power electronics are also needed to adjust the point on the $I-V$ curve corresponding to the actual power demand. Therefore, the operation of a fuel cell in applications with varying load demands without power electronics is impossible. Charging of a battery with a fuel cell is in principle possible. However, the fuel cell must not exceed certain current limits for a safe operation. Power electronics that limit current and voltage according to the requirements are highly recommended. Power electronics are necessary in any case if a controlled charge of a battery according to one of the charging regimes given in Figure 20.23 is necessary.

Presently, no cost figures for marketable fuel cells can be given as all available PEM fuel cells are prototypes for R&D and demonstration. However, there is a huge bunch of activities in fuel cell research. They are mainly driven by the automobile industry and by the combined heat and power generation CHP space-heating applications. The target figures for fuel cell systems in these applications are approximately 100 €/kW for automobiles and 1000 €/kW for CHP space-heating systems.

20.5.2.4 Applications

Hydrogen storage systems have a low overall efficiency. Even under the assumption of a fuel cell system efficiency of 50%, an electrolyser system efficiency of 85%, no energy losses for the hydrogen storage and efficiencies of 97% each for the two power-converting steps, an overall storage system efficiency of 40% maximum is achievable. Compared with approximately 90% efficiency in lead–acid or lithium batteries, this is rather small and rules out the hydrogen system as the principle and only energy storage unit in autonomous power supply systems. This will be the case as long as the power production is as expensive as it is now. Secondly, the specific storage costs per kilowatt hour with the hydrogen system must come down to the values of today's lead–acid batteries and thirdly, the technical reliability of the complex hydrogen system must be as high as with conventional batteries. The hydrogen storage system is far away from all these goals and these goals will be hardly achieved within the next decade.

One interesting line of development is the reversible fuel cell RFC. RFCs fulfil the functionality of the electrolyser and the fuel cell at the same time. As the process is completely reversible, this is an obvious solution. Technical problems concerning the catalysts and the gas and water management within the cell have to be solved [37]. No commercial products for field applications are available in this technology today.

Applications in autonomous power supply systems will have a combined storage system made from a conventional battery and the hydrogen system. Figure 20.30 shows a principle system design as it is currently under development within an R&D project [38]. The hydrogen system here substitutes the conventional motor generator, which is used today in hybrid systems to bridge the energy gap between summer and winter. A conventional lead–acid battery (five days of autonomy) is assisted by the hydrogen storage systems. During summer, the energy excess from the PV generator is used to produce hydrogen in the electrolyser (oxygen is released to the atmosphere). The hydrogen is stored in a metal-hydride tank. During winter, in long periods with low energy supply from the PV generator the fuel cell takes the oxygen from the air and the hydrogen from the tank to produce electrical power.

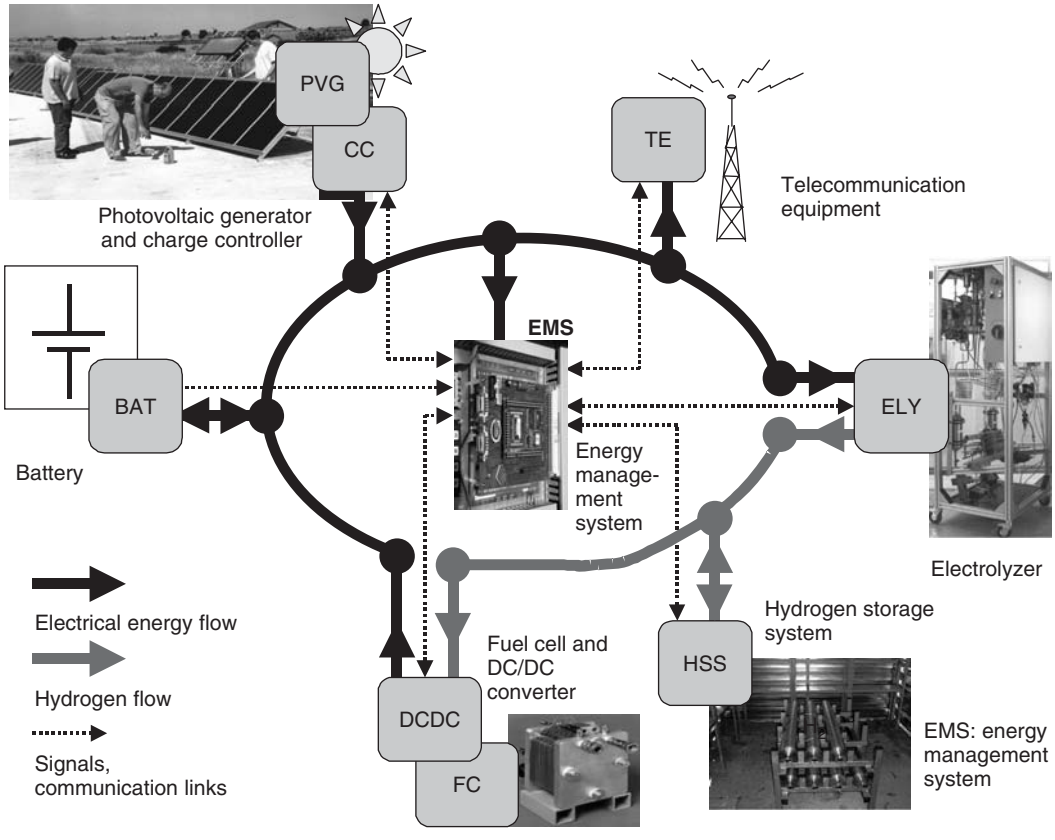


Figure 20.30 Example of an autonomous power supply system for telecommunication with a PVgenerator, a lead–acid battery for short-term storage (five days of autonomy) and a hydrogen storage system for seasonal energy storage. The system and components have been developed by partners in the EC co-financed project FIRST [38]

This concept allows operating the system with very high reliability with respect to the power supply throughout the year without any conventional fuel. The shift of energy from the summer to the winter is beneficial, as, otherwise, in Central Europe a larger PV generator or motor generator would be required to fulfil the energy requirements during winter. Numerous other systems of this type for different applications have been developed and installed by R&D projects within the last decade (e.g. [39–41]).

20.6 INVESTMENT AND LIFETIME COST CONSIDERATIONS

While designing an autonomous power supply system, it is essential to face the lifetime costs rather than looking only into the initial investment costs. The battery has major impacts on the system design, system operation and overall costs.

- The battery causes a considerable part of the initial investment costs.

- The size of the battery influences significantly the solar fraction of the power supply system.
- Typically, more than two-thirds of the energy flow in an autonomous power supply system goes across the storage system. Therefore, the battery acts as an important consumer of electrical energy due to its efficiency of less than 100%.
- The battery voltage influences the selection of the electronic components or vice versa.
- The battery is subject to ageing. Ageing depends very much on the operating conditions of the battery. Operating conditions depend on system sizing and control strategy.
- The lifetime of the battery determines the running costs through the replacement investment.
- The battery needs regular maintenance.
- Depending on the type of the battery, the different requirements of the battery room have to be considered. The requirements are defined in the standards.

These facts are valid for all battery technologies. For the system design, the characteristics of the chosen battery technology must be taken into account. For the following considerations, only lead–acid batteries are taken into account.

Investment costs for lead–acid batteries depend very much on the technology and the quality of the battery. Typical costs for end users are in the range of 75–250 €/kW h. Lifetimes are – depending on the operation conditions –3 to 8 years. Depending on the sizing of the system (days of autonomy) and the lifetime, the battery will be subjected to 100–1000 capacity throughputs. This results in electricity costs of 0.20 to 0.75 €/kW h dedicated to the storage unit. Additional costs occur for the peripherals, the charge controllers res. chargers and maintenance.

Lead–acid batteries need maintenance once or twice a year to check the cell connectors, to measure all cell or block voltages to identify weak cells, to clean the tops of the batteries to avoid creeping currents between the poles, to refill water for flooded batteries and to check the general state of the battery. The set points of the charge controllers or the battery management should be checked. For a 48 V battery system with single cells, 30–60 min for the maintenance are necessary.

Figure 20.31 gives an example of the cost share of the battery in a PV battery system. The system is designed to supply power with 100% reliability at a location in Mexico. The left-side graphs show a system that was designed to minimise the initial investment costs under the given boundary conditions. The right-side graphs show the results of a minimisation of the lifetime costs. The calculation includes the initial investment costs, maintenance, repair and replacement of the components, capital costs and other operating costs. It can be seen from the graphs that the lifetime costs are 288 or 248% of the costs for the initial investment. Sizing of the system with respect to lifetime costs resulted, in this example, in an overall cost reduction of 14%. It is interesting to see that this was achieved by approximately 20% higher investments in the PV generator. This allows on one hand a reduction of battery size by approximately 25% and on the other hand the larger PV generator allows a more frequent complete charging of the battery and therefore a lifetime extension. This altogether resulted in a cost reduction for the battery of more than 40% and – as mentioned above – for the overall system of 14%.

This example is just to illustrate the importance of an integrated system planning and design. Powerful tools are necessary and available [42]. The battery is the component with the highest cost share in autonomous power supply systems. Other than in photovoltaics, cost reductions on battery investment costs will be small within the next few years as lead–acid batteries are already a commercial product today with a market of more than €10 billion per year in the different applications. Cost reductions can occur from improved lifetimes and optimised operation strategies. These figures also show that little tolerance for higher battery costs is given. This is important for the estimation of the market entry of other or new storage technologies. Higher specific costs per kilowatt hour of energy supply and significantly lower efficiency than that achieved today with lead–acid batteries will hardly be accepted by a commercial market.

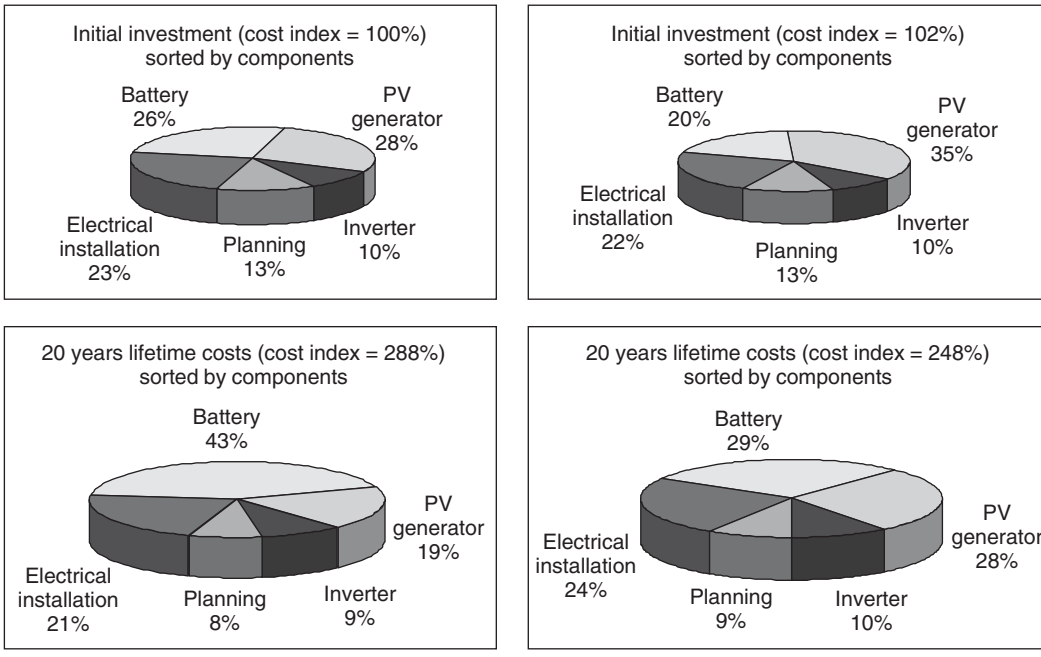


Figure 20.31 Comparison of costs for a PV battery system derived under different assumptions. Left-side graphs show the costs for the system optimised by initial investment costs, right-side graphs show the costs for a system optimised by lifetime costs. The overall costs for the initial investments (upper graphs) and the overall lifetime costs (lower graphs) calculated according to the annuity method are normalised to the initial investment costs for the system optimised by initial investment (cost index = 100%). Location, Mexico; annual power consumption, 1500 kW h; effective interest rate, 6%; lifetime of components: PV generator 20 years, electronic components 15 years, battery according to sizing and operating conditions, calculations and optimisation done with the simulation tool TALCO [42]

20.7 CONCLUSION

Even though there is a wide range of possible solutions for storage in autonomous power supply systems, the economic boundary conditions focus all solutions on lead €acid batteries. This will not change significantly for some years. Their electrical properties are very good. Nevertheless, they have a bad reputation among system designers and operators, which is mainly due to the fact that batteries have a limited lifetime. Owing to their electrochemical nature, they have very complex operation and ageing patterns. Technically speaking, batteries have several time constants. Rapid levelling out of microcycles occurs in the millisecond range, diffusion processes in the seconds and minutes range, state of charge in the hour and day range and ageing effects in the day, month and year range. Batteries have a memory with regard to operation conditions. Faults in the operation can often hardly be repaired, but might show their negative impact much later. Further, currently no method exists that allows a determination of the state of health of the battery within minutes. Nevertheless, it is most likely that lead–acid batteries will be the leading storage technologies in autonomous power supply systems for many years to come. This will be especially true if the still available improvements in the lead–acid technology concerning lifetime in autonomous power supply systems can be realised.

Mainly lithium batteries have enormous growth rates in the portable market segment. Today, it is difficult to forecast the costs for larger units in terms of capacity for lithium batteries – actual forecasts expect costs in the range of 150 €/kW h and from the technical parameters the lithium batteries are most suited to autonomous power supply systems.

Double-layer capacitors will have their market in applications with very high power demands in the range of a few seconds or less and most probably in combination with an electrochemical battery.

In general, a battery is considered to be 'used up' when it has less than 80% of the rated capacity guaranteed by the manufacturer. Nevertheless, this usually does not mean that the battery is completely non-functional. From field experience, it is well known that batteries can be used easily down to 50% of the rated capacity. However, the users must be aware that the days of autonomy are reduced according to the capacity loss and the power available from the battery is reduced due to ageing.

In hybrid systems, the solar fraction reduces with decreasing battery capacity. However, as ageing proceeds, the danger of the so-called fatal defect increases. This results in a more or less sudden breakdown of the battery, and can create considerable problems for the user. They are usually due to short circuits, caused by erosion sludge, corrosion flakes from the poles or dendrite growth between the electrodes. This risk should be considered when the decision is made about replacing a battery with clearly reduced capacity.

The operation strategy has a significant effect on the battery lifetime. It is necessary to take this into consideration when planning and designing an autonomous power supply system. Some additional investments in the battery peripherals will result in significant savings within the system lifetime. Frequent additional full charging and deep-discharge protection on the basis of a state-of-charge determination are highly recommended.

Hydrogen storage systems and redox batteries are options for the future, but the hydrogen storage systems have a fundamental drawback concerning the efficiency and the complexity of the system. Nevertheless, important impulses for their development will come from load-levelling applications in grids with a high penetration of renewable energy sources.

Further development on the conventional battery systems is necessary and will happen within the coming years. After almost one century (1880–1980) with no sustainable market entry of different fundamental battery technologies, we have seen within the last decade great achievements with nickel–metal hydride and lithium batteries, which were not expected. Further, improved battery technologies and a more integral system design and operation resulting in longer lifetimes will allow a decrease in the specific costs for electrical energy storage within the next decade by approximately a factor of two. Further cost reduction can hardly be seen from today's point of view.

REFERENCES

1. Garche J, Döring H, Harnisch P, *Workshop Fortschrittliche Back-up- und Speichersysteme für regenerative Energieversorgungsanlagen*, Forschungsverbund Sonnenenergie, pp 51–72 (Köln, 1996).
2. Linden D, *Handbook of Batteries*, 2nd edn, McGraw-Hill, New York (1995).
3. Berndt D, *Maintenance-Free Batteries – A Handbook of Battery Technology*, John Wiley & Sons, Ltd, Chichester, UK (1993).
4. Wagner R, Sauer D, *J. Power Sources* **95**, 141 (2001).
5. Sauer D *et al*., State of Charge – What do we Really Speak About? *INTELEC '99*, Electronic proceedings of the conference, Copenhagen (1999).

6. Bopp G *et al*., Hybrid Photovoltaic-Diesel-Battery Systems for Remote Energy Supply, *Proc. NORTH SUN '97* (Espoo, Finland, June 1997).
7. Sauer D *et al*., *J. Power Sources* **64**, 197–201 (1997).
8. Bopp G *et al*., *13th European Photovoltaic Solar Energy Conference*, Vol. 2, pp 1763–1769 (Nice, France, 1995).
9. Sauer D *et al*., *14th European Photovoltaic Solar Energy Conference*, pp 1348–1353 (Barcelona, Spain, 1997).
10. Ruddell A *et al*., *J. Power Sources* **112**, 531–546 (2002).
11. Reilly J, in Besenhard J (ed), *Metal Hydrid Electrodes in Handbook of Battery Materials*, 209, Wiley-VCH, Weinheim, Germany (1999).
12. Schmidt H, Siedle C, Anton L, Tuphorn H, *Proc. 30th ISATA Conference*, pp 581–588 (Florence, 1997).
13. Anton L, Schmidt H, 5. Design & Elektronik Entwicklerforum, pp 103–116 München, Germany, (1998).
14. Maxell Europe GmbH, *Lithium Ion Rechargeable Batteries*, Product explanation.
15. Bode H, *Lead-Acid Batteries*, John Wiley & Sons, Inc., New York (1977).
16. Hollenkamp A, *J. Power Sources* **59**, 87–98 (1996).
17. Preiser K *et al*., *14th European Photovoltaic Solar Energy Conference*, Vol. II, pp 1692–1695 (Barcelona, Spain, 1997).
18. Berndt D, *Blei-Akkumulatoren (Varta)*, VDI-Verlag, 11. Auflage, Düsseldorf (1986).
19. Sauer D, *J. Power Sources* **64**, 181–187 (1997).
20. Mattera F, Sauer D, Desmettre D, Rosa M, *Acid Stratification and Vertical Current Distribution: An Experimental and Theoretical Explanation of a Major Ageing Effect of Lead-Acid Batteries in PV Systems*, Extended Abstract for LABAT99, Sofia (1999).
21. McCarthy S, Kovach A, Wrixon G, Operational Experience with Batteries in the 16 PV Pilot Plants, *9th European Photovoltaic Solar Energy Conference*, pp 1142–1145 (Freiburg, Germany, 1989).
22. Döring H, Jossen A, Köstner D, Garche J, 10. *Symposium Photovoltaische Solarenergie*, pp 549–553 (Staffelstein, Germany, 1995).
23. Bohmann J, Hullmeine U, Voss E, Winsel A, *Active Material Structure Related to Cycle Life and Capacity*, Final Report, ILZRO Project LE-277 (1982).
24. Lander J, *J. Electrochem. Soc*. **103**, 1–8 (1965).
25. Garche J, *J. Power Sources* **53**, 85–92 (1995).
26. Piller S, Perrin M, Jossen A, Methods for State-of-Charge Determination and their Applications, *Int. Power Sources Symposium* (2001).
27. Kuhmann J, Paradzik T, Preiser K, Sauer D, 13. *Symposium Photovoltaische Solarenergie*, pp 97–101 Staffelstein (1998).
28. Bartolozzi M, *J. Power Sources* **27**, 219–234 (1989).
29. Garche J *et al*., *Study on New Battery Systems and Double Layer Capacitors*, Internal study in German language for German project EDISON, Financed by BMWi, Compiled by Centre for Solar Energy and Hydrogen Research Baden Württemberg (ZSW), Ulm, Germany (2000).
30. Tokuda N *et al*., *SEI Tech. Rev*. **45**, R22–1–R22-7 (1988).
31. Ryhd C, *J. Power Sources* **80**, 21–29 (1999).
32. Rzayeva M, Salamov O, Kerimov M, *Int. J. Hydrogen Energy* **26**, 195–201 (2001).
33. Pham A, High Efficient Steam Electrolyzer, *Proc. 2000 DOE Hydrogen Program Review*, NREL/CP-570-28890 (2000).
34. Menzl F, Wenske M, Lehmann J, XII. *WHEC Buenos Aires 1998 Proc*., pp 757–765 (1998).
35. Heinzel A, Ledjeff K, in Kreysa G, Jüttner K, Eds, *Elektrochemische Energiegewinnung*, DECHEMA- Monographien, Vol. 128, pp 595–601, Verlag Chemie, Weinheim, Germany (1993).

36. Fuel Cell handbook, 5th edn, Online version at http://216.51.18.233/fchandbook.pdf, Compiled for the U.S. Department of Energy by EG & G Services (2000).
37. Rau A, Heinzel A, *Reversibles Electrolyse-/Brennstoffzellen-System zur Energiespeicherung, Tagung der Gesellschaft Deutscher Chemiker e.V*., Fachgruppe Angewandte Elektrochemie, Ulm, Germany (2000).
38. Vegas A *et al*., The FIRST Project-Fuel Cell Innovative Remote Systems for Telecom, *12th Annual U.S. Hydrogen Meeting* (Washington, DC, 2001).
39. Armbruster A *et al*., *13th European Photovoltaic Solar Energy Conference*, pp 360–363 (Nizza, 1995).
40. Brinner A *et al*., *Int. J. Hydrogen Energy* **17**, 187–198 (1992).
41. Barthels H *et al*., *Int. J. Hydrogen Energy* **23**, 295–301 (1998).
42. Puls H, Sauer D, Bopp G, *Proc. 17th European Photovoltaic Solar Energy Conference and Exhibition*, pp 2673–2678 (Munich, Germany, 2001).

21

Power Conditioning for Photovoltaic Power Systems

Heribert Schmidt[1], Bruno Burger[1] and Jürgen Schmid[2]

[1]*Fraunhofer Institute for Solar Energy Systems ISE, Freiburg, Germany*
[2]*Fraunhofer Institute for Wind Energy and Energy Systems Technology IWES, Kassel, Germany*

In PV systems, power conditioning units are used to provide a match between the specific characteristics of the PV generator and the connected loads or balance of system (BOS) components. Furthermore, they take over the control of other BOS components, for example, batteries or back-up generators.

In general, the characteristic curve of a PV generator varying with solar radiation and temperature does not match the characteristic curve of the load. In those cases, the power conditioning unit effects a transformation of the load's voltage and current in such a way, that the PV generator is operated at its optimum operation voltage V_{MPP}, even under changing conditions.

In the following chapter, the characteristics of the most common power conditioning units – charge controllers, DC/DC converters and inverters – are described.

In almost every stand-alone system, a charge controller is required to optimally operate the storage battery within safe limits as prescribed by the manufacturer.

The matching of PV generator and the load can be achieved by means of DC/DC converters which can be integral part of a charge controller, an inverter or a DC pump, but can also be a separate BOS component.

If in stand-alone systems the load requires an AC voltage, inverters are used to convert the DC power supplied by the PV generator or the storage battery into AC power. Inverters are mandatory in grid-connected PV systems. Here, besides high efficiency, reliability and power quality, safety is an important issue and has to be dealt with.

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

21.1 CHARGE CONTROLLERS AND MONITORING SYSTEMS FOR BATTERIES IN PV POWER SYSTEMS

In PV-powered systems, batteries are still the component with the lowest average lifetime. Compared with solar modules, which in principle have an infinite economic lifetime and in many cases offer a guarantee period of 25 years, the lifetime of batteries is much lower. The maximum lifetime found in practice is around 8–10 years; in most cases it is in the range of 3–6 years and in some cases even lower. The upper limit will be determined by normal ageing; the shorter lifetimes are mostly caused by inappropriate treatment or unsuitable control strategies.

The cost of the batteries in a typical PV-diesel hybrid system is around 15% of the initial costs. Because of repeated replacement of exhausted batteries, this initial share grows to more than 35% or even 50% over the expected 25-year service life of the system. Compared with this, the costs for solar modules and other balance-of-system components become small.

To achieve minimum lifetime costs and satisfactory operation of the PV system, equipping the battery with appropriate peripherals is money well spent. Some examples, such as systems to automatically mix the electrolyte to prevent acid stratification or automatic water-top-up systems, are explained in the chapter dealing with batteries. Besides this, the application of appropriate operation modes is a crucial factor.

In this chapter, the technical realisation of the key component, the "charge controller", will be described. Furthermore, a system to operate long battery strings optimally will be introduced.

21.1.1 Charge Controllers

The fundamental task of a charge controller is to operate the battery within the safe limits defined with respect to overcharging and deep discharging by the battery manufacturer or by the operation mode.

Compared with conventional battery chargers powered by the public grid, the situation is much more complex in PV systems. Here, charging power, time and energy are limited and depend on the varying insolation and load demand. Well-known charging strategies such as constant-current–constant-voltage charging (CC/CV) or more complex charging strategies, cannot be applied one to one. For example, in PV systems the charging current varies according to the insolation. Nevertheless, the term used is "constant current charging". Also, regular full charging or equalisation charging of the battery – which is very important for a long service life – cannot be guaranteed.

Furthermore, very high energy efficiency is crucial for all balance-of-system components in PV systems. Most grid-powered battery chargers offer only unsatisfactory efficiency values.

The fundamental technical concepts of charge controllers as well as the associated control strategies have been developed in the beginning of the utilisation of photovoltaics [1, 2] and will be explained in the following sections. In addition, a list of criteria will be given that should be taken into account when developing or selecting charge controllers.

21.1.1.1 Self-regulating PV systems

In small systems, such as house-number illumination or power supplies for measurement systems, a charge controller can be avoided in special cases.

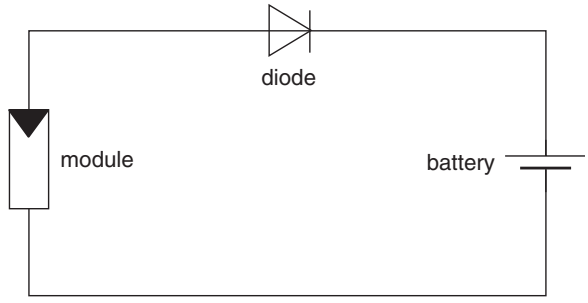


Figure 21.1 "Self-regulating" PV system without a charge controller

This is, for example, true for NiMH batteries, when the current provided by the PV module is lower than the continuous charging current accepted by the battery. Also, such "self-regulating" systems have been used in the past for systems in the 50-W range such as light buoys.

In these systems illustrated in Figure 21.1, only a series diode is needed to block reverse currents from flowing into the module at night. To prevent overcharging of the battery, specially tailored PV modules with, for example, 30 crystalline-silicon cells will be used for 12-V lead–acid batteries. With this low number of cells, the operating point will move into the steeply sloping part of the module's $I-V$ curve when the battery is fully charged.

To make this kind of system operate reliably, the load profile as well as the insolation and temperature at the place of operation must be known accurately. Because cheap and reliable charge controllers are available today on the market, such "self-regulating" systems should be avoided.

In all cases, deep-discharge protection according to Section 21.1.1.6 should be implemented.

21.1.1.2 Linear charge controllers

When PV applications commenced, the well-known principles of conventional linear charge controllers were adapted to photovoltaics. This mature technology offered a simple design and a smooth charging current, but with the drawback of high heat generation in case of regulation. Therefore, the initial linear charge controllers have been replaced by the switching charge controllers.

Nevertheless, in the small power range up to some watts, novel integrated low-drop-out (LDO) voltage controllers can be used.

In a linear charge controller, the charging current will be adjusted by a final controlling element that acts continuously and is located either in series or in parallel with the solar generator. By driving the controlling element appropriately, the battery voltage can be prevented from exceeding the end-of-charge limit.

Figure 21.2 shows the block circuit diagram of a linear series charge controller. In the constant-current (CC) phase, in which the battery voltage is below the end-of-charge voltage, the control element MOSFET t1 is fully conducting. The solar generator and the battery are directly coupled via the blocking diode d1. The operating point of the solar generator is determined by the instantaneous insolation and the battery voltage. The power losses inside the control element are almost negligible in this phase. Additional power losses are caused by the voltage drop across the blocking diode d1, which in most cases will be a Shottky diode with a very low forward voltage drop. To minimise the power losses, the blocking diode can be replaced by a second

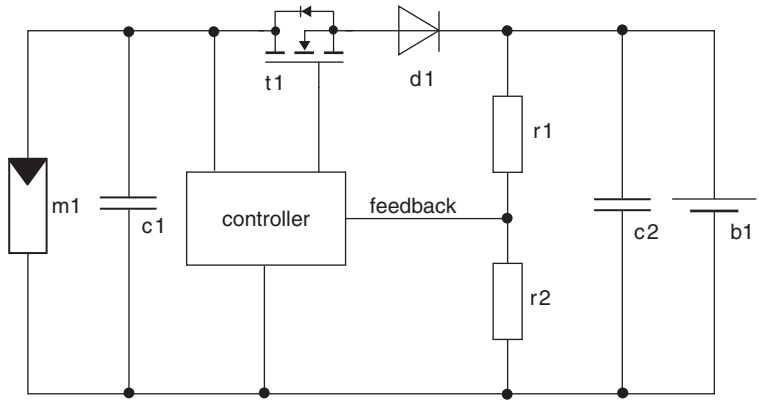


Figure 21.2 Linear charge controller based on the integrated LDO voltage controller and an external power MOSFET

MOSFET connected back-to-back in series with t1. With such designs, care has to be taken that both MOSFETs turn off during the night, and that no reverse current flows into the solar generator!

As soon as the end-of-charge voltage has been reached, the gate voltage for the MOSFET t1 will be reduced by the controller. Now, the output voltage is kept constant while the charging current will drop slowly according to the battery's increasing state of charge. In contrast to the first charging phase, now the difference between the solar generator voltage and the battery voltage will occur at the transistor terminals and cause some heat dissipation.

In Figure 21.3, the operating points for three different charging currents are shown as an example. The shaded areas are proportional to the power dissipated in the control element. As can be seen, this power is in the range between approximately one-fourth and one-tenth of the instantaneous maximum power point (MPP) power of the generator. The resulting heat must be dissipated by appropriate heat sinks. This is a clear disadvantage compared with the switching charge controllers described below. On the other hand, there are no problems with electromagnetic compatibility (EMC), and the battery will be charged with a life-extending constant current without micro-cycles.

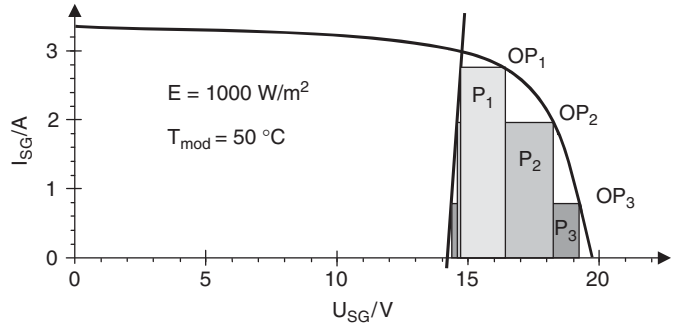


Figure 21.3 $I-V$ curve of a 36-cell PV module and characteristic curve of a 12-V lead–acid battery with 3 m input lead (1.5 mm^2 cross-section) for three different charging currents. The shaded areas show the power dissipated in the control element t1

21.1.1.3 Switching controllers

The disadvantage of intensive heat dissipation in linear charge controllers can be overcome with switching controllers. In switching controllers, the control element is either fully closed (conducting) or fully open (blocking). Under ideal conditions, the power losses are zero in both cases because either voltage or current at the control element is zero.

It must be pointed out that the additional energy gained in this way is normally not relevant for the functioning of the PV system. But, the reduction of heat generation leads to savings in component costs (e.g. heat sinks) and to an increased reliability due to lower temperature stress.

In a series controller as shown in Figure 21.4a, the charging current will be controlled by an element switched in series with the solar generator.

In early charge controllers, relays were used as switching elements, but today semiconductor switches like MOSFETs are used in almost every application.

One advantage of the series controllers is that in addition to PV generators, they can be used for other power sources that are not tolerant of short circuits, such as wind turbines. Furthermore, the voltage stress for the switch is lower compared with the shunt controller described below. With a fully charged battery, the solar generator operates in the open-circuit mode. In this operation mode, no module overloading due to partial shading can occur. On the other hand, the solar generator current is fully turned on and off, which can lead to greater EMC problems compared with shunt controllers.

In the past, the series controller was accused of having fundamentally higher losses. This is no longer true since low-resistive semiconductor switching elements have been used – the losses can be even lower than for shunt controllers. Furthermore, some early series controllers did not start with completely flat batteries because they did not have enough energy to activate the series switch. This problem can easily be solved by powering the controller from the PV module as well as from the battery.

A parallel or shunt controller as shown in Figure 21.4b makes use of the electrical characteristic of PV modules that they are inherently tolerant to short-circuits.

In the constant-current charging phase, the module current flows through the blocking diode d1 into the battery. When the end-of-charge voltage has been reached, the PV generator is short-circuited by the switch s1. The blocking diode now prevents the reverse current flowing from the battery into the switch. Furthermore, it suppresses the discharging currents into the PV generator during the night.

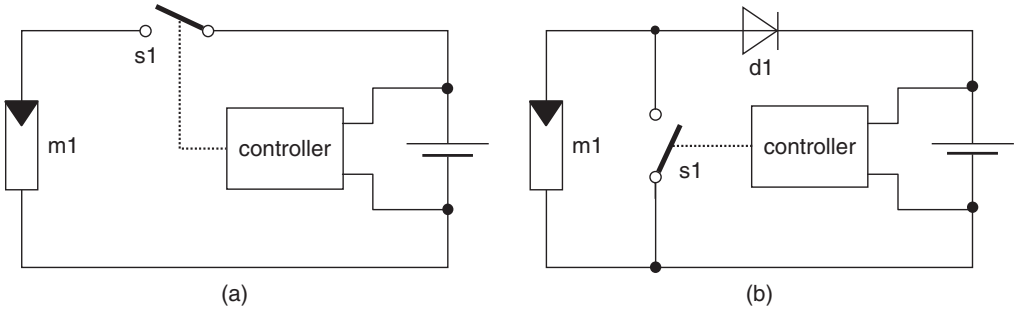


Figure 21.4 Principle of (a) series and (b) shunt switching controllers

In contrast to the series controller, this kind of charge controller will also reliably start charging with a fully discharged battery, because the switch has to be energised only when the battery is fully charged.

The hybrid charge controller is a modified shunt controller, in which the blocking diode is bypassed by a second transistor in the charging phase. This reduces the power dissipation in the controller, which leads to smaller heat sinks and lower thermal stress. Furthermore, the reduction of the voltage drop supports the use of cost-optimised PV modules with a smaller number of series-connected cells, for example, 30–33 crystalline-silicon cells for 12-V systems.

Another variant of the shunt controller is the partial-shunting controller according to Figure 21.5. Here, only some of the series-connected modules are shunted via taps. In consequence, the remaining generator voltage is too low to further charge the battery. The advantage of partial shunting is the reduced voltage stress for the switch in high-voltage systems. On the other hand, a tap and additional wiring are needed.

In high-power systems, sub-array switching according to Figure 21.6 is used for charge control. The PV generator is set up as a number of sub-arrays (strings), and each of these strings is connected to the battery via its own control element (e.g. blocking diode and switch).

Two different control strategies are used. The multiple switches can be driven in parallel, turning the full array current on and off in one step. A preferred strategy is to drive the switches in a certain sequence that allows stepwise adjustment of the charging current.

With all of the switching controllers described above, the designer or end user has to be aware that the solar generator is directly coupled with the battery and the load. If the battery connection is interrupted (maintenance work, lead breakage, blowing of the battery fuse, etc.), the full open-circuit voltage of the PV generator can be applied to the load and may destroy it. To prevent damage, either the loads must be able to withstand the high voltage or the controller must be designed to avoid voltages higher than the rated output. This can be achieved, for example, by rapid over-voltage detection and load cut-off.

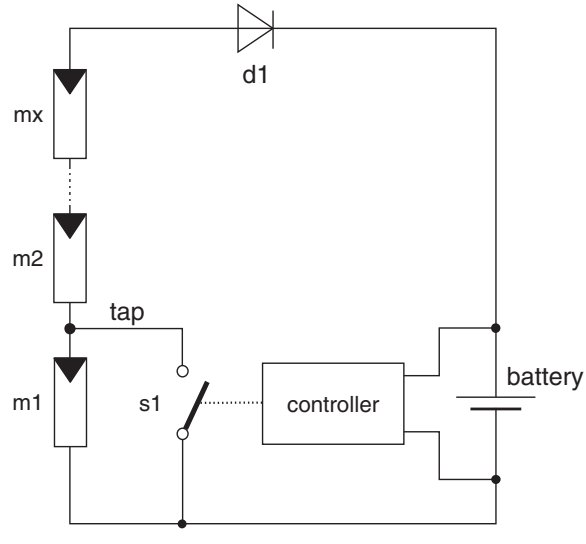


Figure 21.5 Partial-shunting controller for high-voltage systems

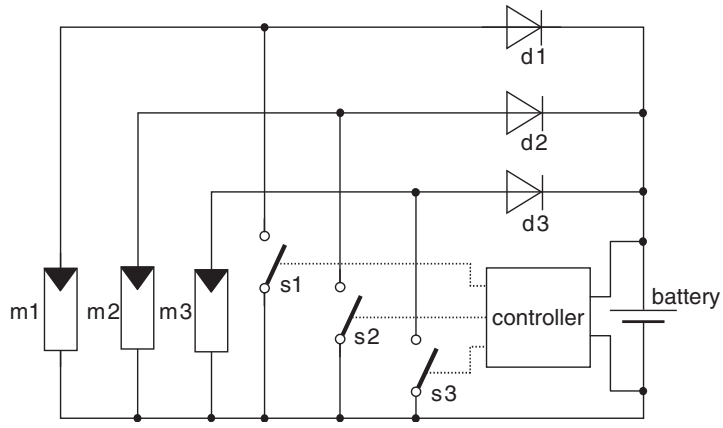


Figure 21.6 Sub-array switching controller for high-power systems

21.1.1.4 Control strategies

When the end-of-charge voltage is reached for the first time, the battery is not yet fully charged. The missing 5–10% of charge can be added to the battery by keeping it at the end-of-charge voltage level for a prolonged period. In this constant-voltage (CV) phase, the charging current will slowly decrease. How can such a charging regime be implemented by a series or shunt controller as described above, which can only switch all of the PV generator's current on and off? Two techniques are used in practice to approximate the ideal CC/CV charging mode.

In a two-level controller, the charging current is dropped to zero by either opening the series switch or closing the shunt switch as soon as the end-of-charge voltage has been reached. As a result, the battery-terminal voltage decreases. The charging current is enabled again when the battery voltage drops below a threshold that is between 5 and 50 mV/cell lower than the end-of-charge voltage.

This sequence gets repeated periodically and the charging pulses become shorter and shorter, while the intervals in between become longer, as the battery's state of charge increases as shown in Figure 21.7. The average charging current decreases, while the terminal voltage is more or less constant. The period of the cycle described above is not constant and depends on the battery capacity, the state of charge, the charging or discharging current as well as the voltage hysteresis. It can vary from milliseconds to minutes. If mechanical relays are used as switching elements, the period should not be shorter than 1–5 min.

The second control regime found in practice is pulse-width modulation (PWM). In principle, it works like the two-level controller described above, but the switching frequency of the control element is fixed and determined by a clock generator. The typical frequency is about 100 Hz. In the CC phase, the switch is permanently closed and the full charging current flows into the battery. When approaching the end-of-charge voltage, the duty cycle (the ratio between the charging time and the cycle period) will be reduced towards zero by the pulse-width modulator. As mentioned above, the average charging current will drop and the battery voltage is kept constant.

One advantage of PWM controllers is that the switching frequency is known and constant. EMC problems can be solved more easily, and also monitoring of the average charging current becomes simpler.

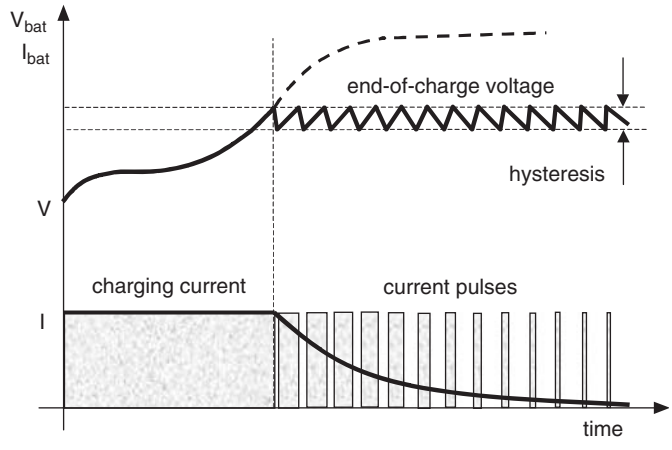


Figure 21.7 Battery voltage and current during charging

21.1.1.5 Matching DC/DC converters, MPP trackers

Both the battery voltage and the PV generator voltage vary over a wide range during operation due to the changing state of charge and boundary conditions such as temperature and insolation. When directly coupled, this leads to a certain mismatch between the actual and the optimum operation voltage (MPP voltage) of the solar generator and therefore causes energy losses.

The mismatch can be overcome by introducing a matching DC/DC converter (MDC) that decouples the characteristic curves of the PV generator and the battery. The power stage of these converters corresponds to the well-known topologies such as buck, boost or inverting converters. The control section is specially tailored to the PV conditions and consists of two control loops: one for the input and the other for the output. As long as the end-of-charge voltage has not been reached, the input voltage controller keeps the PV generator voltage at a constant level by appropriate adjustment of the DC/DC converter's switching regime (e.g. PWM). The controller's setpoint value can either be fixed (CV mode) or can track the actual MPP by an appropriate searching strategy (MPP tracking, MPPT). When the end-of-charge voltage is reached, the output voltage controller takes over and keeps the battery voltage at a constant level. The PV generator's operating point then shifts towards open-circuit conditions.

Numerous strategies and algorithms have been developed to find and track the MPP of a solar generator. They can be grouped into two categories:

- *Indirect MPP trackers*
  This type of MPP tracker estimates the MPP voltage by means of simple assumptions and measurements.
  Some examples from practice include the following:
  – The operating voltage of the solar generator can be adjusted seasonally. Higher MPP voltages can be expected in winter due to lower cell temperatures and vice versa.
  – The operating voltage can be adjusted according to the module temperature.
  – The operating voltage can be derived from the instantaneous open-circuit voltage by multiplication with a constant factor, for example, 0.8 for crystalline silicon solar cells. The open-circuit voltage is measured periodically (e.g. every two seconds) by disconnecting the load for a few milliseconds.

The advantage of these procedures is simplicity, but they only give an estimate of the optimum operating point. They are not able to adapt to changing solar generator characteristics due to ageing or soiling.

- *Direct MPP trackers.*
  In these systems, the optimum operating voltage is derived from measured currents, voltages or the power of the PV generator. Therefore, they are able to react to changes in the generator's performance.
  Some examples include the following:
  – Periodic scanning of a part of the $I-V$ curve. Here, the operating voltage of the module is varied within a given voltage window by means of the DC/DC converter. The maximum module power is determined and then the operating voltage is adjusted to the corresponding voltage level. In practice, it is much easier to measure the output current of the DC/DC converter and to maximise it. This leads to the same result as above.
  – "Mountain-climb" or "Perturb and Observe (P&O)" algorithm. Here, the operating voltage is periodically changed in small steps. The increment can either be constant or can be adapted to the instantaneous operating point as shown in Figure 21.8. If the module's power (and therefore the charging current) increases from one step to the next, the search direction is retained; otherwise it is reversed. In this way, the MPP is found and the operating point oscillates around the actual MPP.

The above mentioned losses due to mismatch are mostly overestimated. Particularly, if the components are chosen properly (e.g. modules with 30–33 cells for 12-V systems), the energy losses are on the order of a few percent when using direct coupling. This corresponds with the losses inside a matching DC/DC converter. Besides this fact, it should be considered whether the additional energy gained by optimum matching is relevant to the function of the system at all. For example, the battery in a typical Solar Home System will be fully charged before noon and the excess solar energy will be dissipated.

Nevertheless, there are some advantages in using charge controllers with matching DC/DC converter and MPP-tracking:

- Most of today's PV modules provide output voltages that do not fit to typical battery voltages of e.g. 12 or 24 V. This is in particular true for thin film modules. The problem can be solved by means of a matching DC/DC converter which leads to more flexibility in system design.
- More complex and battery-friendly charging current profiles can be realised when using a matching DC/DC converter or MPPT-charger.
- In the case of long wires from the PV generator to the battery, the generator voltage can be chosen much higher than the battery voltage, resulting in lower currents and therefore lower wiring losses.
- In very small applications, the PV module can consist of only a few large cells or even only one cell instead of numerous small cells connected in series. This reduces the module production costs, the impact of cell mismatch and the sensitivity to partial shading. Fully integrated DC/DC converters in the power range of a few watts, providing start-up-voltages of less than 0.5 V are appearing on the market.

21.1.1.6 Deep-discharge protection

To achieve a maximum service life, deep discharging of batteries as well as prolonged periods with a low state of charge should be avoided. Therefore, the load has to be disconnected automatically from the battery as soon as the state of charge falls below a certain level. The load should be reconnected to the battery only when a sufficient state of charge has been reached.

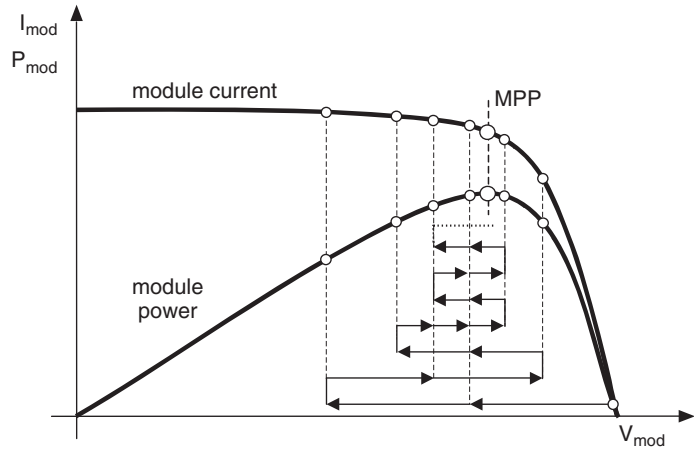


Figure 21.8 Working principle of the "mountain-climb" MPP-tracking algorithm

Different criteria for detecting the deep-discharge condition have been explained in the battery chapter. In commercial products, the battery voltage will be used as a criterion for load disconnection. As soon as the battery voltage drops below a determined level, the load will be disconnected via a (bi-stable) relay or a semiconductor switch. Also, a control signal can be output to shut down other balance-of-system components such as inverters.

More complex charge controllers are able to generate a warning signal when the deep-discharge condition is being approached. Also, different loads can be disconnected according to a given priority. Charge controllers including an energy-management system (EMS) are used to start back-up generators such as diesel or gas generators, depending on the battery's instantaneous state of charge. Additional parameters such as load demand, weather conditions and so on can be considered.

There should be an appropriate delay time $t_{d\,off}$ of 10–60 s between the undershooting of the end-of-discharge level and the actual disconnection of the load as shown in Figure 21.9. This ensures that undesirable disconnection of loads with large starting currents, for example, motors, refrigerators, washing machines and so on, can be avoided. As the end-of-discharge voltage

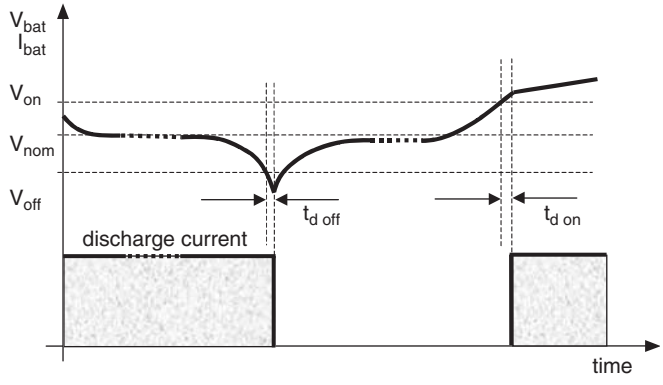


Figure 21.9 Course of the battery voltage and the load current during discharge

threshold depends on the instantaneous battery current, some advanced charge controllers offer a current-dependent adaptation of the disconnect threshold.

The ideal solution would be deep-discharge protection based on the actual state of charge (SOC) of the battery. As systems or algorithms for accurately measuring the state of charge under the complex boundary conditions in PV systems are still under development, this type of controller is not yet on the market. For safety reasons, deep-discharge protection systems based on complex algorithms should always be combined with a redundant hardware control system based on simple voltage thresholds as described above.

The voltage threshold for reconnection of the load must be adjusted properly. If it is too low, the battery's open-circuit voltage will overshoot the threshold and the load will be reconnected periodically. In spite of the protection system, the battery will be deeply discharged and damaged. A time delay $t_{\mathrm{d\,on}}$ as described above is also appropriate for guaranteeing a minimum state of charge before load reconnection.

In contrast to the end-of-charge threshold, the end-of-discharge threshold should not be temperature-compensated.

21.1.1.7 Monitoring systems and interfaces

Successful functioning of most PV systems strongly depends on the cooperation of the user. Reliable and meaningful information to the user about the current status of the system – the battery in particular – is therefore crucial. It would be ideal to have a "fuel gauge" that enables the system operator to plan the future energy demand. Such systems to monitor the state of charge have been discussed in the battery chapter.

As a minimum, the charge controller should be equipped with a display, for example, a light-emitting diode (LED) that indicates the conditions 'Battery can be further discharged' (green) and 'Battery is fully discharged' (red). Other conditions such as 'Battery close to end-of-discharge' (yellow) can be helpful. Furthermore, an experienced operator can draw a lot of information from a meter showing the battery's voltage and current.

All user interfaces should be ergonomic and provide only that information really required by the operator.

Energy-management systems are equipped with (potential free) inputs and outputs to communicate with external components, for example, to remotely start a diesel generator. In addition, standard interfaces such as RS 232, RS 485 or CAN-Bus can be integrated to parameterise the controller, to read out the status of the system or to download operation data from the built-in data logger. Another feature is to communicate with external balance-of-system components via power line transmission or wirelessly.

In Solar Home Systems, the charge controller can act as an energy meter or can be used to automatically debit energy costs from a pre-paid card.

21.1.1.8 Design criteria and appraisal factors for charge controllers

The following summary of requirements is intended to support a designer or a purchaser of a charge controller. Which of these requirements must be met has to be decided specifically for each individual application, see also [3].

The values of voltage thresholds and so on are related to lead–acid batteries.

1. *Charging phase*
   - The end-of-charge voltage threshold should be adjustable according to the battery in use (2.3–2.5 V/cell at 25 °C). To prevent extreme misadjustment, the setting range should not extend beyond these limits.
   - The end-of-charge voltage can be automatically adapted to the system voltage (e.g. 12 or 24 V).
   - If the battery temperature is expected to deviate more than ±10 °C from the average temperature under operation, the end-of-charge voltage should have a temperature compensation (approximately −4 to −6 mV/K per cell). If the temperature deviation is smaller, temperature compensation is not mandatory and the end-of-charge threshold should be set according to the average battery temperature. Fail-safe behaviour is crucial in case of a temperature-monitoring malfunction, for example, due to broken sensor wires.
   - The thresholds must be stable over temperature and time.
   - If relays are used as control elements, the minimum switching period should be 1–5 min.
   - The charge controller should be able to charge totally flat batteries. As a minimum requirement, charging should start from a cell voltage of 1.5 V.
   - The battery voltage can be monitored by separate sensor wires. This is recommended if the battery wiring is long and of low cross-section. Fail-safe behaviour is crucial in case of broken sensor wires.
   - The charge controller should be able to automatically perform gassing charging or equilibration charging according to the manufacturer's recommendations.
   - It must be possible to prevent gassing charging in case of valve-regulated batteries (Gel- or AGM-VRLA batteries).
   - The output voltage must be limited to safe values in the case of system operation without a battery, for example, due to unintended disconnection of the battery, wire breakage or blowing of the battery's fuse.
2. *Deep-discharge protection*
   - Deep-discharge protection is mandatory for a long service life. Only when the function of the system is more important than the battery life (e.g. in SOS telephones), deep-discharge protection can be omitted.
   - The threshold voltages should be adjustable in a range of 1.5–2.0 V/cell. The adjustment range should not extend beyond this range to prevent extremely incorrect settings.
   - The threshold should not be temperature-compensated.
   - The threshold must be stable over temperature and time.
   - The threshold can automatically adapt to the instantaneous battery current.
   - The threshold can be based on the battery's actual state of charge.
   - A time delay of 10–60 s should be implemented between undershooting of the threshold and the actual load disconnection.
   - A warning signal can be given out when the deep-discharge condition is approached, for example, switch on a yellow LED 30 min before load disconnection at full load.
   - Load disconnection can be performed according to load priorities.
   - After load disconnection, only minimum current ($< I_{10\,000}$) should be drawn from the battery. This can be achieved by appropriate design, for example, displaying of the current system status only on demand.
   - The threshold for load reconnection should be relatively high, for example, >2.1 V/cell.
   - A time delay of 10–60 s should be implemented between overshooting of the reconnection voltage threshold and the actual reconnection of the load.

3. *Efficiency*
   - The parasitic consumption of the charge controller should be less than 0.2% of the PV generator's power, for example, <8 mA in a 12-V/50-W controller.
   - The voltage drop measured between the PV input terminal and the battery terminal should be less than 4% at full charging current, for example, approximately 0.5 V in a 12-V system.
   - The voltage drop measured between the battery terminal and the load terminal should be less than 4% at full load current, for example, approximately 0.5 V in a 12-V system.
   - The above-mentioned figures lead to an efficiency of >96% under rated charging as well as discharging conditions.
4. *Safety aspects and compliance to codes*
   - The charge controller must be protected against the reverse polarity of the input voltage and the battery voltage, for example, by a combination of fuses and diodes. Also, an unintentional exchange of the inputs and outputs must not lead to damage.
   - The charge controller must permanently withstand the maximum possible open-circuit voltage of the PV generator. This occurs at maximum insolation and minimum module temperature.
   - Inputs and outputs must be protected against transient over-voltages by appropriate voltage arresters, for example, varistors.
   - The charge controller must be designed according to the ambient temperatures at the usage site.
   - The housing for the charge controller must withstand the environmental stress at the usage site, for example, protection level IP 00 (ingress protection) inside a control cabinet, or IP 65 for outdoor applications.
   - The electronic components should be protected by lacquer or encapsulation.
   - The terminals should be generously dimensioned and robust. Cage clamps are a preferred solution.
   - The charge controller must comply with relevant codes concerning electric safety and electromagnetic compatibility (EMC).

21.1.1.9 Examples of commercial charge regulators

PV Charge regulators are mass products that are manufactured by numerous companies worldwide. Comprehensive market overviews are compiled by various sources, for example the magazine PHOTON International [4]. In Figures 21.10–21.14, some examples of charge regulators from different power ranges and ingress protection (IP) level are shown.

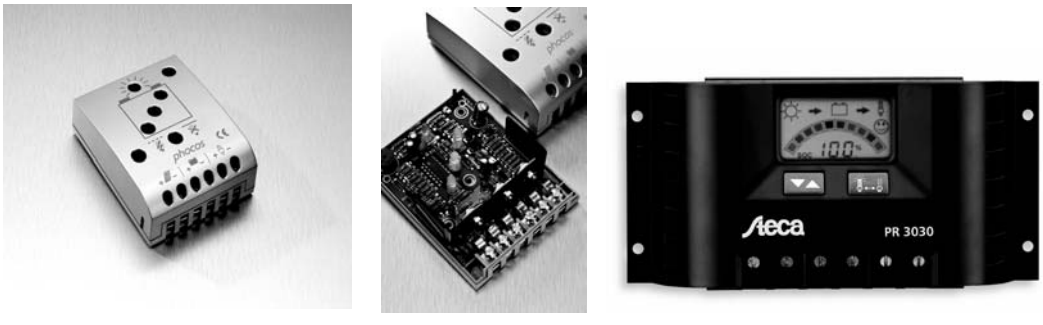


Figure 21.10 PWM charge controllers in the 15–30 A range (IP22). Courtesy of PHOCOS [5] and STECA [6]

Figure 21.11 20 A (left) and 140 A (right) PWM charge controllers for outdoor mounting (IP65). Courtesy of UHLMANN [7]

Figure 21.12 MPPT charge regulators in the 10–15 A range. Courtesy of STECA [6] and MORNINGSTAR [8]

Figure 21.13 Modular MPPT charge regulator system in the 30 A range and interior view of the MPP-Tracker unit. Courtesy of PHOCOS [5]

21.1.2 Charge Equaliser for Long Battery Strings

21.1.2.1 Introduction

All electrochemical storage systems react sensitively to operating states that do not conform to their specifications, such as overcharging, deep discharge or reversed polarity. The general observations are that the lifetime is shortened, the storage capacity and the efficiency are diminished and more maintenance is required. Further, detrimental conditions can arise with some batteries, for example, lithium-ion batteries, which lead to destruction or even explosion.

Figure 21.14 High power MPPT charge regulator. Courtesy of OUTBACK [9] and SCHNEIDER [10]

However, on proceeding from the individual cell to series connection of several cells or groups of cells, which are usually needed, a well-known phenomenon is observed. In all applications, from laptops through PV systems and electric vehicles to uninterruptible power supplies (UPS) and grid back-up systems: each individual cell in the series connection behaves differently. Conventional charge controllers are not able to recognise this variation in the behaviour of the cells, so that the undesirable operating conditions listed above arise. In practice, it is evident that the "weakest link in the chain" determines the quality of the whole string, and that the deviating performance of a single cell can lead to a chain reaction.

The problem of increasing divergence in individual cell properties within a battery has been known since the beginning of battery technology, so that over the years a number of different procedures to solve the problem have been developed. Most of them are based on the dissipation of the surplus energy of fully charged cells in a bypass element. This approach is not suitable for applications in which highest efficiency is crucial, such as PV systems. Furthermore, it is effective only with a fully charged battery – it has no impact when the battery is being discharged.

On the basis of experience with numerous PV systems, active, non-dissipative charge-equalising systems have been developed [11]. In contrast to conventional dissipative systems, here the surplus energy from cells having a higher state of charge is redistributed among the remaining

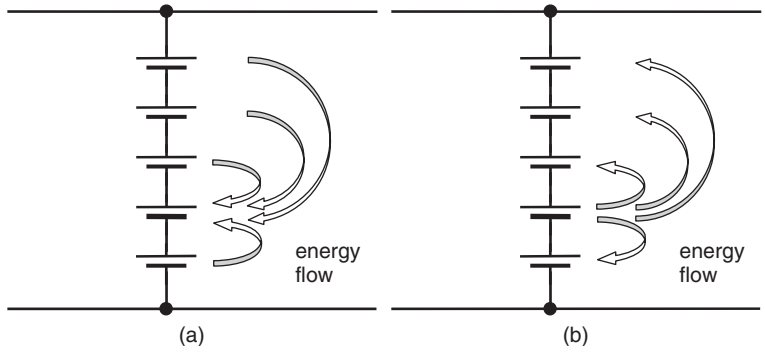


Figure 21.15 Support for a weak cell during discharge by redistribution of energy from stronger cells (left). Protection against overcharging (right)

cells. As indicated in Figure 21.15, this redistribution occurs not only during or at the end of charging but also while discharging.

As a result, cells with a lower capacity are supported by the other cells during discharge. Their relative state of charge decreases evenly with that of the cells with higher capacity. In this way, the entire available capacity of all cells can be used. During charging, some of the charging current is redistributed from weaker cells to stronger ones, so that these are charged with a higher current, resulting in faster charging.

The charge equalizer principle has been tested successfully in a number of remote PV systems and also electric vehicles as well as in uninterruptable power supplies.

21.2 INVERTERS

21.2.1 General Characteristics of Inverters

Solar cells and modules basically produce DC currents and DC voltages. Energy storage systems such as batteries are also based on DC values. On the other hand, most of the typical household loads are designed for an AC power supply. Therefore, in many off-grid applications an inverter is needed to convert DC power into AC power. To feed PV energy into the public AC grid, an inverter is mandatory. Accordingly, numerous different types of inverters specially tailored for these demands have been developed in the past. Inverters can be categorized into three main groups:

1. Inverters for grid connection.
2. Inverters for off-grid systems.
3. Inverters for special applications, for example water pumps.

The power range of commercial inverters varies in the range from tens of watts to megawatts. Nevertheless, in all three fields of applications and in all power ranges, similar hardware topologies are used that will be explained later in this chapter. The major differences between the applications are located in the structure of the control system.

In grid-connected applications, the aim of the inverter is to feed a sinusoidal current into the already existing and rigid public grid. The amplitude of the current is proportional to the present DC power – the inverter therefore has the characteristic of an electrical current source.

In a stand-alone application, the inverter itself must produce an output voltage which is stable in amplitude as well as in frequency – in this case, the inverter acts as a voltage source.

Inverters for water pumps produce an output voltage which varies in frequency as well as in amplitude according to the characteristic of the pump and the actual power provided by the solar generator.

For all PV applications, a high efficiency at all power levels is of key relevance due to the high generation cost of PV electricity. State-of-the-art grid connected inverters achieve efficiencies of $>95\%$, commercial high end units have more than 98%. In the laboratory, optimised topologies making use of novel semiconductor switches and diodes based on silicon carbide (SiC) with an efficiency of more than 99% for an entire inverter, have been demonstrated [12].

Besides high efficiency, there are numerous general requirements for inverters such as high reliability, ease of maintenance, low weight, low noise, electromagnetic compatibility (EMC), ease of use, interfaces/software and low cost.

21.2.2 Inverters for Grid-connected Systems

21.2.2.1 Basic structures of grid-connected PV systems

The basic structure of a typical grid connected PV system is shown in Figure 21.16. It consists of a PV generator directly coupled to the input side of an inverter without any means of storage. The output side of the inverter is coupled to the grid via safety devices, which typically are integrated into the inverter. A meter measures the energy fed into the public grid.

Based on this fundamental structure, three typical layouts of grid connected systems have emerged:

1. *Central inverter/master–slave systems*. This topology was used in the first grid- connected applications. A number of PV modules are connected in series building a string to achieve the input voltage required by the inverter. To attain the planned power, several of such strings are switched in parallel in a generator junction box. The collected DC power is then fed into a central inverter, converted to AC power and injected into the grid. As shown in Figure 21.17, the DC power can also be distributed to a number (typically two or three) of smaller inverters switched in parallel. In this so called master–slave configuration, the inverters are switched on and off according to the instantaneous DC power. Thereby, even at low input power, the inverters are operated in their optimum efficiency range which increases the overall efficiency of the system.

   Today, central inverters with or without master–slave configuration are mostly used in high-power systems in the several hundred kilowatt or megawatt range.
2. *String concept*. In this concept, each string has its own dedicated small inverter in the power range of several hundred watts up to several kilowatts. Typically, a grid-connected rooftop

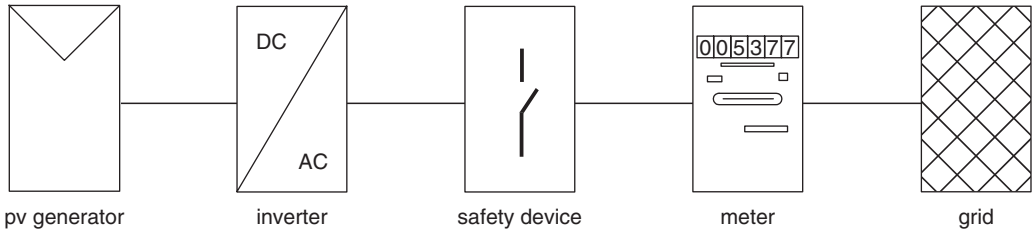


Figure 21.16 Fundamental structure of a grid connected PV system

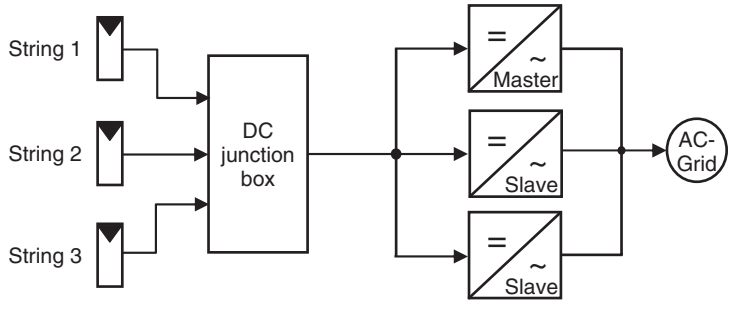


Figure 21.17 Central inverter in master–slave configuration

system on a single family house consists of one to three strings each with its own assigned inverter, thus, in principal forming independent subsystems as shown in Figure 21.18.

This broadly used system layout offers many advantages:

- generator junction boxes are avoided;
- the strings can be of different length and orientation which is especially advantageous for building integrated (BIPV) systems;
- each of the strings has its own maximum power point tracker (MPPT), leading to reduced losses in case of partial shading or different orientation of the strings;
- high production quantities of small inverters reduce the production costs;
- handling and maintenance is much easier compared to bulky central inverters.

These numerous advantages lead to the use of this system layout also in larger PV plants where, for example, in a 100 kW system, 20 inverters, each of 5 kW, work independently.

The layout described is often modified in such a way that two or three strings can be connected to one inverter unit. Inside the inverter, those strings are either simply switched in parallel, saving the costs for an external generator junction box at the expense of the major advantages of the original string technology. Or, each input has a dedicated DC/DC converter feeding an intermediate DC link, thus enabling individual MPP-tracking of each string. Typically, there are two or three inputs in those so called multi-string inverters.

3. *Module integrated power conditioning units*. From the very beginning of the use of terrestrial PV, there was the ambition to combine the power conditioning unit directly with the PV module as shown in Figure 21.19. Thus, each module has its own dedicated MPP-tracker reducing losses caused by deviant module parameters, different module temperatures or partial shading of the generator. Furthermore, it is claimed that it would be advantageous to have an AC cabling instead of DC cabling. A clear advantage of a PV plant based on AC-modules is the very simple system planning and later extension. Another approach to integrate power electronics is to build only a matching DC/DC converter into the junction box, thus enabling individual MPP-tracking. The DC energy is then collected on a (high voltage) DC bus and fed into a central inverter.

Besides some obvious advantages of the module integrated approach, there are some drawbacks.

The efficiency of a small inverter in the several hundred watt range will always be lower than that of a higher power unit. One of the main reasons is the self consumption needed to power the control systems, relays, monitoring and communication systems of the inverter, as well as, e.g. hysteresis losses in inductors and transformers. The self-consumption of a leading-edge inverter in the 5 kW range is less than 5 W, which means less than 0.1% of the nominal power.

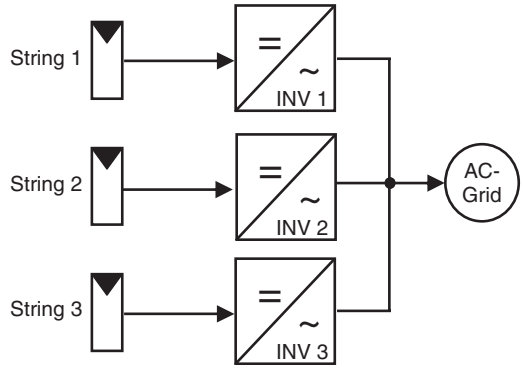


Figure 21.18 Grid-connected PV system based on string connected inverters

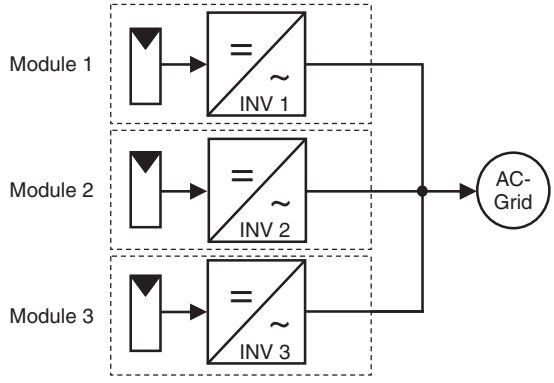


Figure 21.19 System based on module integrated inverters (AC-modules)

As the self consumption for a low power inverter will not decrease linearly, this leads to a poor efficiency especially under part load conditions. In an unshaded PV plant, a noticeable reduction of the annual yield can be expected compared with conventional string or central inverter concepts. In building-integrated PV systems, where regular recurring partial shading cannot be avoided, module integrated power conditioning may lead to an increase in annual yield.

Further sticking points are reliability and lifetime of such module integrated power conditioning systems. They are exposed to harsh environmental conditions such as high temperature levels and variations, humidity and electromagnetic fields in the case of lightning. At the same time, they have to have a lifetime comparable to that of the PV module, e.g. 25 years. To achieve this goal, design and manufacturing quality must be on an industry or military level at the costs of commercial products.

If these fundamental problems are overcome, building-integrated photovoltaics (BIPV) would be the first and rapidly growing market for module integrated power conditioning systems.

21.2.2.2 Safety aspects with grid-connected inverters

When talking about safety of grid connected inverters, two aspects must be considered: islanding and safety of the DC side.

Islanding occurs when the grid is shut down by the grid operator, and distributed energy sources still feed energy into this grid, keeping the wires live. This unintended situation may cause an accidental hazard if at the same time somebody gets in contact with the live wire.

In the beginning of grid-connected photovoltaics, islanding was a vigorously discussed subject. Meanwhile, after many years of experience with several hundred thousand grid-connected PV systems, it can be clearly stated that this risk has been overestimated and that the anti-islanding measures mentioned below are absolutely sufficient.

The prescribed anti-islanding measures differ from country to country, but the following techniques or a combination thereof are most commonly used:

- Supervision of the grid voltage (passive single phase or three phase systems);
- Supervision of the grid frequency (passive systems);
- shifting of the inverter's output frequency (active system)
- supervision of the grid impedance (active systems).

Passive monitoring of the grid voltage and frequency can easily be performed inside the inverter and leads to highly reliable results. Additional measures such as monitoring of the grid impedance require additional hardware and – as they are active systems - cause disturbances in the grid, often leading to malfunctions and unintended shutdowns of the inverter. Therefore, passive and low-complexity systems should be preferred.

As in transformerless inverters there is no galvanic isolation between the public grid and the PV generator, the risk of an electric shock exists when touching cells of a broken module or damaged DC cables. Also, it is claimed that there is a hazard caused by capacitive leakage currents when touching the surface of a module. Again, experience shows that this danger was overrated and can be overcome by observing the following rules:

- usage of protection class II (double isolated) modules and BOS-components;
- measurement of PV-generator leakage resistance before inverter turn-on;
- usage of specially tailored residual current detectors (RCD, AFI) inside the inverter;
- usage of two independent (redundant) grid monitoring systems;
- usage of two independent (redundant) relays for grid disconnection;
- grounding of the generator's support structure (equipotential bonding).

It can be stated, that due to these countermeasures, transformerless inverter systems are even safer than inverters with galvanic isolation. Meanwhile, more than 70% of the inverters installed in Germany are transformerless, and other countries such as the USA are changing their electric codes to permit the use of the advantageous transformerless inverters [13].

21.2.2.3 Active power control of the grid

As more and more decentralised power systems feed into the public grid, their role has changed from being tolerated by the grid operators to being an active part of the grid. In the past, the safety principle was based on 'shut down in case of incident'. In the near future, decentralized systems have to provide so-called ancillary services like stabilising the grid voltage, providing reactive power, active filtering of harmonics, balancing three-phase systems or fault ride-through in case of voltage dips. Also, in case of an overproduction of PV power, inverters have to reduce their output power by shifting the operating point of the solar generator towards open-circuit conditions or to shut down completely. All these additional features have a strong impact on the inverter design. According to the upcoming codes, each inverter must be able to provide reactive power to stabilize the grid voltage. To enable support of the grid in case of short-circuits or to bridge short brownouts, energy storage elements based on, for example, supercapacitors or Li-Ion batteries may be added to the system. Furthermore, depending on the power range, the inverters must be integrated into the overall control system of the grid by remote control, thus requiring access to the inverters by communication systems such as power line communication (PLC), ripple control or the internet. In the future, such multifunctional inverters together with measures such as demand-side management will lead to so-called smart grids [14].

21.2.3 Inverters for Stand-alone Systems

21.2.3.1 Requirements of inverters for stand-alone operation

Inverters for stand-alone applications typically are powered from batteries. Therefore, they should offer a good adaption to the battery voltage changing from – 30 to +25% relative to the nominal

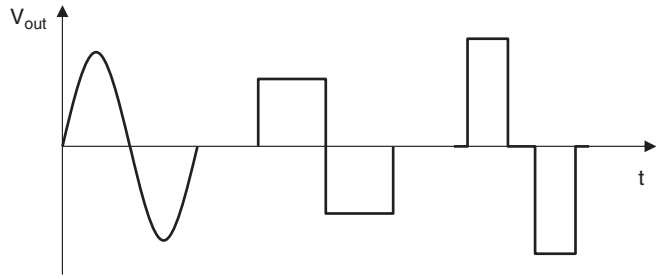


Figure 21.20 Sinusoidal (left), square wave (middle) and modified sine wave (right) output voltage of off-grid inverters

voltage, depending on the state of charge (SOC). To minimise the stress for the battery, the input current ripple should be low. In addition to the charge regulator's deep-discharge protection, an independent under voltage shut-down should be provided. The input should offer an overvoltage and a reverse voltage protection with diodes and/or fuses. Inverters integrated in hybrid systems can also operate bidirectionally, which means, they can charge the battery from AC power provided by other sources such as diesel generators or wind turbines. In this case, the inverter must take over the charging procedures for the battery such as CC/CV charging or equalisation charging at regular intervals. Also, a state of charge calculation may be performed by the inverter.

On the output side, off-grid inverters must provide an output voltage which is stable in effective (RMS) value and frequency irrespective of the actual load profile. In the case of reactive loads, they have to source and to sink reactive power. As typical appliances are designed for a sinusoidal supply voltage, the output voltage of off-grid inverters should be sinusoidal. Most of today's inverters provide high-quality sine wave outputs, but in the low-power and low-cost range, also square wave and so-called modified square wave inverters according to Figure 21.20, can be found. With those inverters, it has to be clarified if the loads operate reliably with such non-sinusoidal voltage waveforms.

To start loads with high inrush currents such as electric motors, off-grid inverters should be able to provide a high peak power for a short period of time. This feature is also important for tripping fuses in case of short-circuits on the load side. To power nonlinear loads such as dimmers of fluorescent lamps, the inverter must cope with high crest factors.

In the case of off-grid inverters, a very low self-consumption (which is equivalent to a very steep slope of the efficiency curve at low loads) is crucial because most of the time the inverter operates under part load conditions. If an extremely low self-consumption cannot be achieved by circuit design, two measures are common to reduce the energy losses. The inverter can either be shut down automatically if a no-load condition has been detected and vice versa. Or, a small, low-loss master unit provides an output voltage all the time and starts a high-power slave unit on demand (master–slave operation).

The inverter must comply with the electric safety codes and also regulations regarding electromagnetic compatibility (EMC).

21.2.3.2 Stand-alone systems with a DC bus

The original approach to build off-grid power supply systems was based on DC coupling of all components according to Figure 21.21. A stable DC bus provided by the battery builds the 'back-bone' of the system. In a so-called hybrid system, beside the PV generator, other generators such

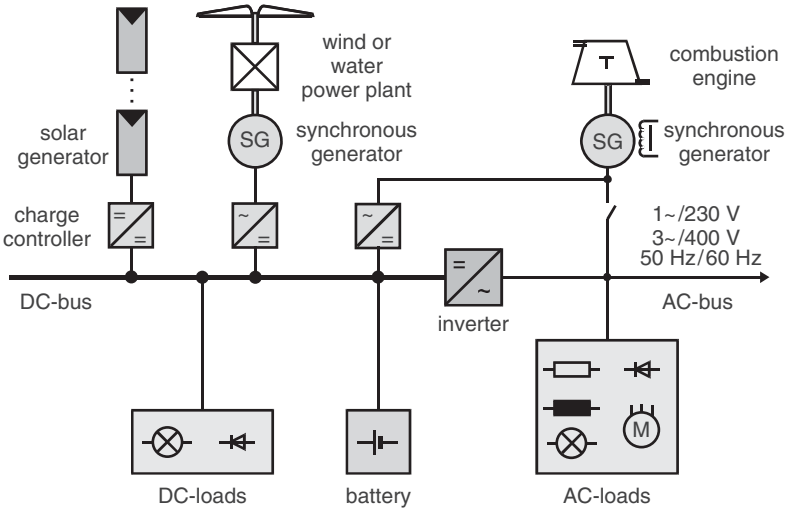


Figure 21.21 Block diagram of a DC-coupled hybrid system

as wind or hydro power plants or motor-generator units (gensets) are coupled to this DC bus, either via charge controllers or rectifiers/controllers. DC loads may be directly powered from the DC bus, AC loads are supplied via a single-phase or three-phase inverter. As an option, the AC loads can be served directly by the genset in case of high power demand or malfunction of the inverter.

The advantage of DC-coupled systems is simplicity and stability. Also, standard products from different suppliers can be combined easily.

21.2.3.3 Stand-alone systems with an AC bus

As shown in Figure 21.22, in an AC-coupled hybrid system generators providing AC power and AC loads can directly be connected to the AC bus. The AC bus is prepared by a bidirectional inverter fed by the system's battery. As in a grid connected system, the DC power produced by the solar generator must be converted into AC power by an inverter.

The design and also the extension of an AC-coupled system are easier as all components are coupled on a standardised voltage level of e.g. 230 V/50 Hz. Furthermore, the collection as well as the distribution of energy takes place on the same bus, which simplifies the integration of distributed energy sources. On the other hand, special inverters and generators are needed to allow parallel connection on the AC side.

In practice, both the DC and AC coupling are used.

21.2.4 Basic Design Approaches for PV Inverters

In the early days of PV utilisation, inverters from drive technology or uninterruptable power supplies (UPS) were taken as a basis, and have been adapted to the needs of photovoltaics. The results were poor, and especially the annual mean efficiency was very low due to the high self-consumption of these inverters that were designed for high efficiency at nominal power. Meanwhile, more than 50 different topologies specially tailored for the demands of photovoltaics have emerged, and the

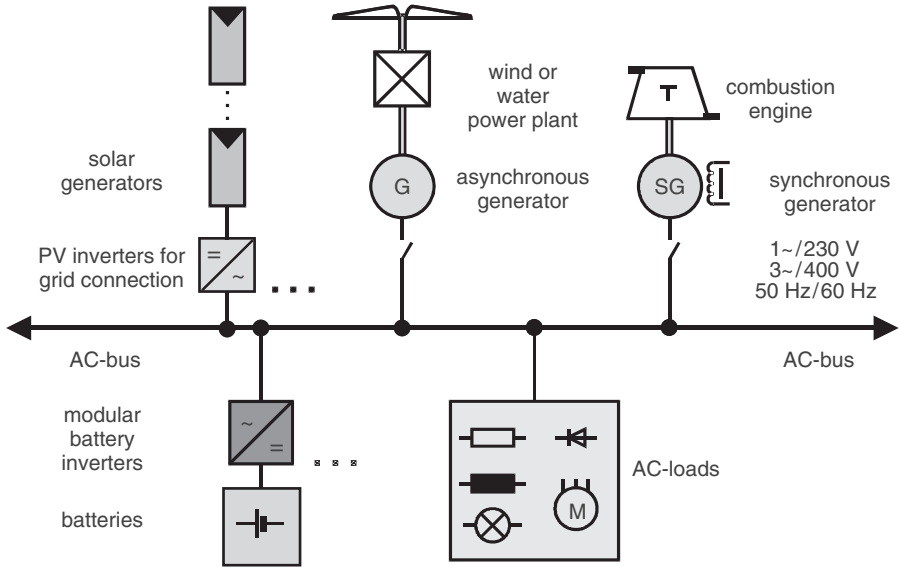


Figure 21.22 Block diagram of an AC-coupled hybrid system

maximum efficiency is close to 100%. Three commonly used topologies will be briefly described here, more detailed information can be found in e.g. [15].

21.2.4.1 Transformerless inverters

A very common design for single-phase transformerless inverters is shown in Figure 21.23. The DC power provided by the solar generator is buffered in an input capacitor C1 and then fed to a full bridge made of the switches S1–S4, typically MOSFETs or IGBTs. According to the switching regime generated by the inverter's controller, these switches are turned on and off with a switching frequency of 5–20 kHz and a varying pulse width (pulse width modulation, PWM). The rectangular output voltage of the bridge is then fed to the inductors L1 and L2, which smooth the pulses to form a sinusoidal output current which is finally fed into the grid. Additional filters on the DC and the AC side needed to fulfill the EMC requirements are not shown here. The main advantage of this design is simplicity, high efficiency and reliability. A disadvantage is the fact that the input voltage must always be higher than the grid voltage, thus leading to a minimum input voltage of

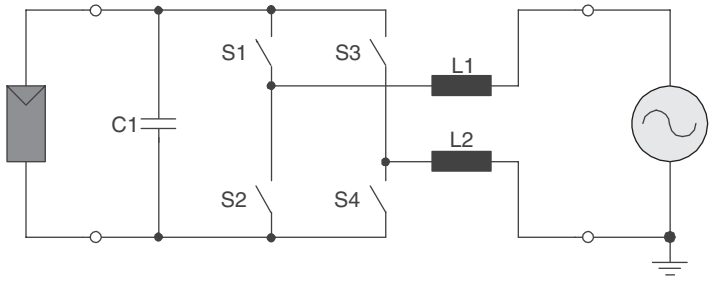


Figure 21.23 Basic transformerless inverter topology

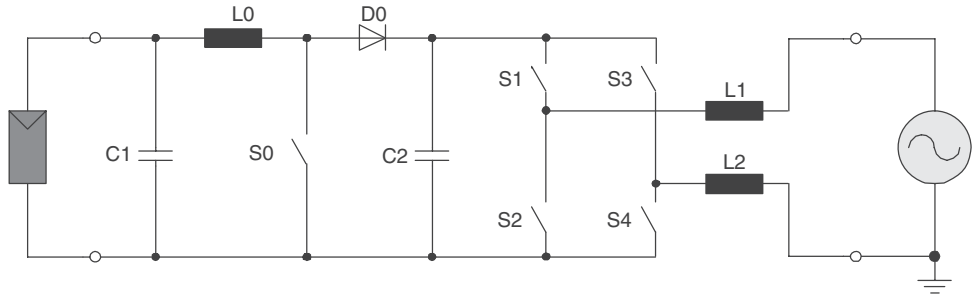


Figure 21.24 Basic transformerless inverter topology with input boost converter

around 350 V for 230 V grids. This drawback can be overcome by introducing one or more DC/DC converters stages according to Figure 21.24, which boost the input voltage to an intermediate DC voltage of 350 V. Due to this additional power stage, the overall efficiency will drop by 1–2%.

21.2.4.2 Low-frequency (LF) transformer inverters

The inverter design with low frequency shown in Figure 21.25 in principle consists of a transformerless inverter and a downstream transformer operating at grid frequency. The transformer offers a galvanic isolation between AC and DC sides, which allows shifting of the solar generator's potential relative to earth potential, which is relevant for some module technologies (see also Section 21.2.6). From a safety point of view, galvanic isolation is not required. Furthermore, the transformer helps to overcome the limitation of the input voltage range already mentioned. According to its turn ratio, even very low input voltages can be utilised.

Major drawbacks of low-frequency transformers are their high weight of 5–10 kg/kW, their high costs and their losses, which lead to reduced efficiency values, especially in the low and high power ranges.

21.2.4.3 High-frequency (HF) transformer inverters

The galvanic isolation and the voltage transformation can also be performed by a transformer operating in the high-frequency range, for example in the range of tens of kHz. As the volume of a transformer reduces linearly with the frequency, the weight and volume of a high-frequency transformer are significantly lower compared with a transformer operating at grid frequency. A typical

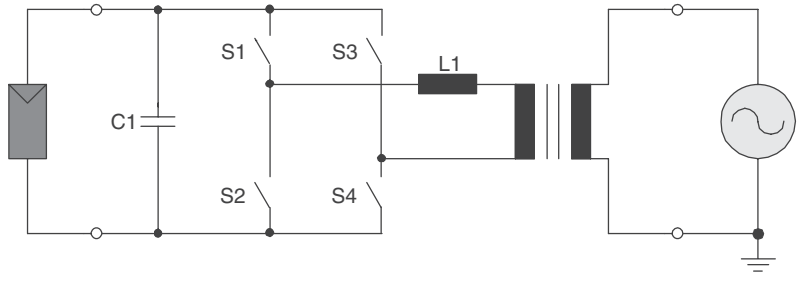


Figure 21.25 Basic design of a single-phase inverter with low-frequency (LF) transformer

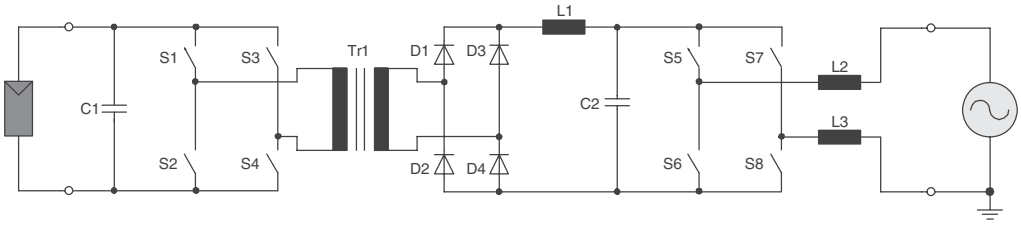


Figure 21.26 Basic design of a single phase inverter with high-frequency (HF) transformer

topology of an inverter with a high-frequency transformer is shown in Figure 21.26. The input power is buffered by the capacitor C1 and fed to the full bridge S1–S4, which is clocked with high frequency, e.g. 16 kHz. The high frequency transformer Tr1 performs a galvanic isolation and also a transformation of the rectangular output voltage of the bridge. On the secondary side, the voltage is rectified by the diodes D1–D4 and fed into the DC voltage link capacitor C2 via the inductor L1. The constant intermediate voltage is then converted into AC current by the full bridge S5–S8 and the inductors L2–L3, as described earlier.

As can be seen, to gain the advantage of a low weight transformer an extra effort has to be spent on semiconductors and control. Finally, due to the increased number of semiconductors the current has to pass through, the efficiency is around 2–4% lower compared with single stage-inverters.

21.2.5 Modelling of Inverters, European and CEC Efficiency

21.2.5.1 Efficiency modelling of inverters

Detailed simulation of inverters at the hardware or control system level with a high time resolution can be done by standard software such as P-Spice or MATLAB-Simulink. For prediction of the long-term behaviour, for example the estimation of the annual yield, a simple model of the efficiency versus the input or output power is adequate [16]. A standard model that is used in a number of PV simulation tools based on a three-parameter approach is shown in Figure 21.27.

The three parameters can be interpreted as a self-consumption p_{self}, which is independent of the actual transferred power, a voltage drop v_{loss} which causes losses proportional to the power and a loss resistor r_{loss} which produces losses proportional to the square of the power. With these parameters, the efficiency versus output power can be expressed as:

$$\eta_{pout} = \frac{p_{out}}{p_{out} + (p_{self} + v_{loss}\, p_{out} + r_{loss}\, p_{out}{}^2)}$$

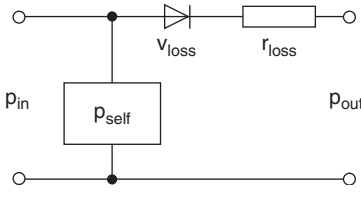


Figure 21.27 Simplified three-parameter efficiency model of power conditioning units

The three parameters can be derived from the measured efficiency curves by means of standard fitting algorithms (Gaussian least-mean-square algorithm) or by making use of the tailored equations utilizing the efficiency at 10, 50 and 100% of the nominal output power:

$$p_{\text{loss}} = 5/(36\eta_{10}) - 1/(4\eta_{50}) + 1/(9\eta_{100})$$
$$v_{\text{loss}} = -5/(12\eta_{10}) + 33/(12\eta_{50}) - 4/(3\eta_{100}) - 1$$
$$r_{\text{loss}} = 5/(18\eta_{10}) - 5/(2\eta_{50}) + 20/(9\eta_{100})$$

This simple model does not take into account the input voltage dependence of the inverters efficiency. To include this, the three parameters themselves are expressed as a function of the input voltage V_{DC}. The parameters are derived from at least three efficiency curves measured at different DC voltage levels, e.g. at 110% of the minimum input voltage, at the nominal and 90% of the maximum input voltage.

$$p_{\text{loss}} = p_{\text{loss0}} + p_{\text{loss1}} V_{\text{DC}} + p_{\text{loss2}} {V_{\text{DC}}}^2$$
$$v_{\text{loss}} = p_{\text{loss0}} + v_{\text{loss1}} V_{\text{DC}} + v_{\text{loss2}} {V_{\text{DC}}}^2$$
$$r_{\text{loss}} = r_{\text{loss0}} + r_{\text{loss1}} V_{\text{DC}} + r_{\text{loss2}} {V_{\text{DC}}}^2$$

Figure 21.28 shows exemplarily measured efficiency values (dots) of a transformerless inverter and a set of efficiency curves calculated by this procedure.

21.2.5.2 European and CEC efficiency

To enable a simple comparison of the annual mean efficiency of different inverters, a criterion based on weighting factors derived from the annual irradiation distribution has been established – the European Efficiency or ETA-EURO.

$$\eta_{\text{EURO}} = 0.03\,\eta_5 + 0.06\,\eta_{10} + 0.13\,\eta_{20} + 0.1\,\eta_{30} + 0.48\,\eta_{50} + 0.2\,\eta_{100}$$

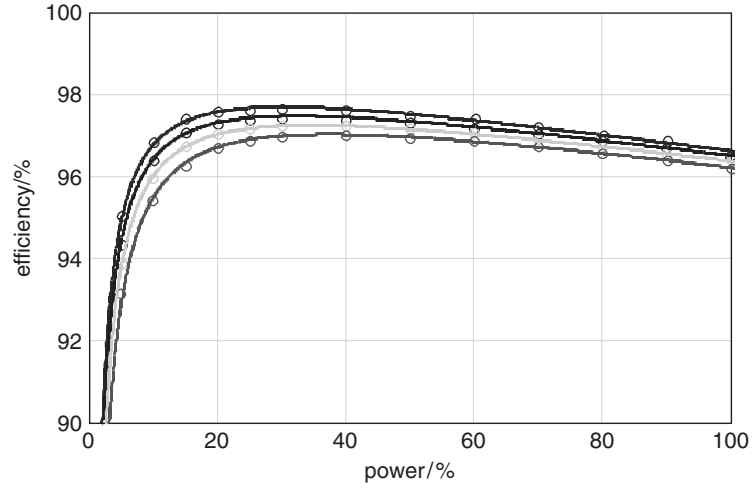


Figure 21.28 Measured efficiency (dots) and set of efficiency curves at $V_{\text{DC}} = 350$, 420, 500 and 600 V (from top to bottom)

For locations with high annual irradiation levels, the CEC efficiency has been defined by the California Energy Commission making use of other supporting points and weighting coefficients:

$$\eta_{CEC} = 0.04\,\eta_{10} + 0.05\,\eta_{20} + 0.12\,\eta_{30} + 0.21\,\eta_{50} + 0.53\,\eta_{75} + 0.05\,\eta_{100}$$

Both standards are used widely, but they are up for debate because:

- the weighting factors are derived from long-term averaged irradiation values, thus leading to an overestimation of medium irradiation levels;
- the PV generator's characteristics such as temperature effects, as well as the impact of typical system sizing (power ratio of PV-generator/inverter) have not been considered;
- the self-consumption of the inverter and other BOScomponents such as meters or monitoring systems has been ignored;
- the voltage dependence of the inverters efficiency has not been taken into account.

Paying attention to all these shortcomings leads to completely different weighting coefficients which weight the high power range of the efficiency curve much more strongly [17].

21.2.6 Interaction of Inverters and PV Modules

Besides 'classic' solar cells made of crystalline silicon, novel kinds of crystalline cells have established themselves on the market, such as back junction solar cells (BJC) and various thin film technologies. With this great variety, the question is whether all of these different cell technologies can be combined with any kind of inverter, or whether specific configurations reduce output, or even damage individual system components [15]. To answer this question, possible failure mechanisms and the impact of the inverter topology will be explained.

First, the potentials of the terminals and consequently of each cell of a string, will be defined according to Figure 21.29.

The solar generator voltage V_{SG} can be measured between the two terminals. The solar generator's voltages with respect to ground at the positive and the negative poles are V_{Plus} and V_{Minus} respectively. These voltages might be virtually pure DC voltages, but, depending on the inverters hardware topology and control strategy, major sinusoidal or rectangular AC voltages may also be present.

After several decades of experience with conventional crystalline Si solar modules it can be stated, that no correlation has been observed between a degradation of the module and the type of inverter used. However, the degradation of some modules based on special cells and also some thin

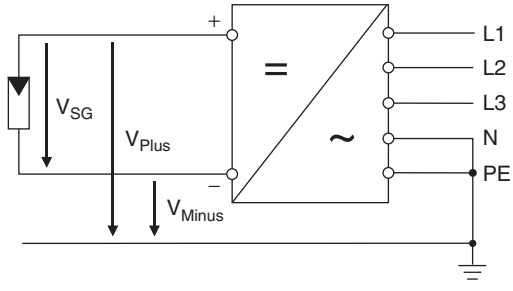


Figure 21.29 Definition of potentials at the input side of an inverter

film modules show such correlations. In the case of the back junction cells, a reversible degradation in efficiency has been observed due to the so-called polarization effect [18]. This effect is caused by small leakage currents flowing from the cells through the embedding material and through the front glass to the frame or support structure. These ultra-low currents cause an accumulation of charge on the antireflective coating of the cell, which has a strong impact on the cell efficiency. The time constant of the process is several hours or days, and it is fully reversible. The direction and the strength of the current depend on the potential of the cell relative to ground – with positive potential, negative charges congregate on the antireflective coating of the cell, leading to a reduction in efficiency of up to 30%. A negative potential leads to a positive surface charge, which can even increase the cell efficiency. The polarisation effect can therefore be prevented by positive grounding of the PV generator which means, the positive terminal of a string is connected to ground.

A similar effect has been observed with some string ribbon cells – here, a negative potential relative to ground caused a reversible degradation of the efficiency. In this case, negative grounding is appropriate [19].

Laboratory experiments and field experience shows, that some thin film technologies are also prone to leakage-current-induced degradations. The probable cause is a corrosion of the transparent front contact (TCO) due to sodium ions moving from the front glass to the interface between the glass and the TCO [20]. The migration of sodium ions can be avoided by negative grounding of the generator.

The input terminal potentials of an inverter are determined by the topology, the switching regime of the power switches and by an eventual internal or external grounding of the generator. It is very important to notice that inverters with and without galvanic isolation by a transformer may have exactly the same potentials. Therefore, galvanic isolation by a transformer without additional grounding measures does not solve these problems!

To clarify the behaviour of inverters, three groups, A–C, of typical potentials have been proposed [15].

Group A consists of inverters where the solar generator's potentials are 'quiet' to ground potential, while none of the poles has a fixed connection to the ground potential. Depending on the topology, the partial voltages V_{Plus} and V_{Minus} can be symmetrical or asymmetrical about the zero line (ground potential). Except for a small 100 Hz ripple in single-phase units, no other alternating voltage interferes with these voltages, as shown in Figure 21.30.

Group B consists of certain transformerless single-phase inverters, in which a sinusoidal alternating voltage interferes with the voltages to ground, as illustrated in Figure 21.31. The

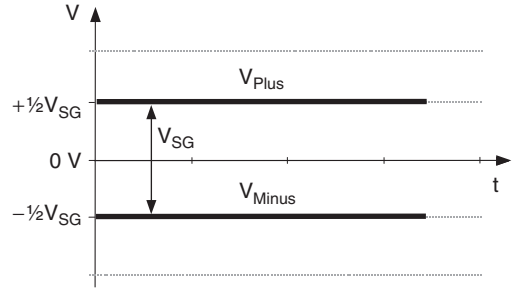


Figure 21.30 Basic potential curve for inverters belonging to group A. In this example, the potentials are symmetric to ground which is typical for inverters with galvanic isolation (transformer), but without additional grounding measures

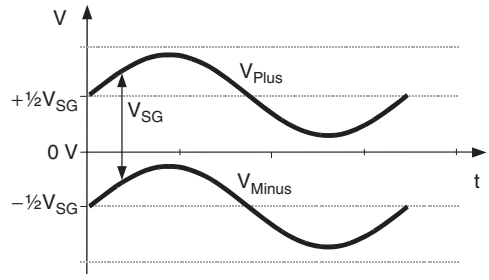


Figure 21.31 Basic potential curve for systems belonging to group B

alternating voltage corresponds to exactly half of the mains voltage, i.e. 115 V and 50 Hz in European grids.

Group C consists of systems, where one pole of the solar generator is connected to the ground potential (neutral conductor) within or outside the inverter; in other words, the solar generator has a fixed potential according to Figure 21.32. Certain transformerless topologies offer an inherent, thin film friendly grounding. In case of transformer-based inverters, special grounding kits are provided by some manufacturers.

The increasing adoption of thin film panels in systems with high system voltages has increased awareness of the problems described above, and panel manufacturers continue to improve their cell technologies, edge seals and frames. Manufacturers also perform highly accelerated lifetime tests (HALT) under exposure to high temperatures and humidity while applying the maximum allowable system voltage as a negative or positive bias voltage. Some manufacturers expect that, for their products, technological advances have eliminated these problems completely and that, considerating the restriction due to capacitive leakage currents, basically any inverter can be used.

If no clear information on inverter selection is provided by the panel manufacturer, the following recommendations should be observed for thin film panels:

- use transformerless inverters with an inherently grounded minus pole;
- use inverters with transformers and an internally grounded minus pole (grounding kit);
- use frameless panels; if necessary, mount with isolating clips or with clamps affixed on the back surface.

Ultimately, the panel manufacturers must specify whether their panels are suitable for operation with all inverter topologies or for which of the three classes A–C they can be used. Panel developments must ensure that future models can be combined with all inverter topologies.

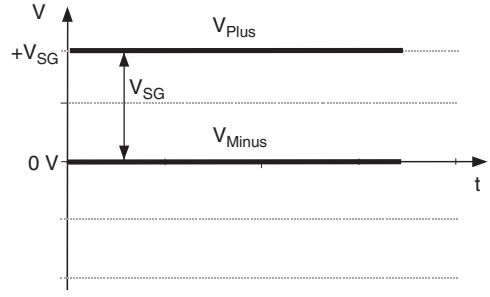


Figure 21.32 Basic potential curve with negative grounding for systems belonging to group C

REFERENCES

1. Nekrasov P, Partial Shunt Regulation, *Proc. 14th Annual Meeting of the Astronautical Society* (13–15 May 1968).
2. Salim A, A Simplified Minimum Power Dissipation Approach to Regulate the Solar Array Output Power in a Satellite Power Subsystem, *Proc. 11th Intersociety Energy Conversion Engineering Conference Proceedings*, *Vol. II* (Nevada, 12–17 September 1976).
3. Universal Technical Standard for Solar Home Systems, Version 2, 1998, updated 2001 Thermie B: SUP-995-96.
4. Brand B, Seeking the power peak *Market survey in PHOTON International 1-2008, Solar Verlag Aachen, Germany*, pp. 114–129.
5. PHOCOS: www.phocos.com.
6. STECA: www.steca.de.
7. UHLMANN: www.uhlmann-solar.de.
8. MORNINGSTAR: www.morningstarcorp.com.
9. OUTBACK: www.outbackpower.com.
10. SCHNEIDER: www.schneider-electric.com.
11. Siedle, Ch. *Vergleichende Untersuchungen von Ladungsausgleicheinrichtungen zur Verbesserung des Langzeitverhaltens vielzelliger Batteriebänke (Comparative investigations of CHarge EQualizers to improve the long-term performance of multi-cell battery banks)*, Doctoral thesis, ISBN 3-18-324521-3, VDI-Verlag Düsseldorf, Reihe 21, Nr. 245 (1998).
12. Burger B *et al*., Highly Efficient PV-Inverters with Silicon Carbide Transistors, *Proc. 24nd European Photovoltaic Solar Energy Conference, Paper no. 4BV.1.2*. (Hamburg, 21–25 September 2009).
13. Burger B *et al*., *25 Years Transformerless Inverters, Proc. 22nd European Photovoltaic Solar Energy Conference*, (Milan, 03–07 September 2007) pp 2473–2477.
14. European Technology Platform 'Smart Grids', http://www.smartgrids.eu.
15. Burger B (ed.) Power Electronics for Photovoltaics, *Proc. OTTI Professional Seminar, ISBN 978-3-934681-77-4*, (Munich, 10–11 June 2008).
16. Schmidt H *et al*. Modelling the Voltage Dependence of Inverter Efficiency (in German), *Proc. 23rd Symposium Photovoltaische Solarenergie*, (Bad Staffelstein, 5–7 March 2008)ISBN 978-3-934681-67-5, pp 158–163.
17. Burger B *et al*. Are we Benchmarking Inverters on the Bassis of Outdated Definitions of the European and CEC Efficiency? *Proc. 24nd European Photovoltaic Solar Energy Conference*, (Hamburg, 21–25 September 2009) Paper no. 4BV.1.10.
18. Sunpower: Sunpower Discovers the 'Surface Polarization' Effect, http://www.sunpower corp.com/html/Resources/TPpdf/polarization.pdf.
19. Evergreen: Evergreen releases updated Installation Manual, http://www.mail-archive.com/re-wrenches@lists.re-wrenches.org/msg00433/08-03-US_Installation_Manual_Update_Release_ 0108.pdf.
20. Osterwald CR, Accelerated Stress Testing of Thin-Film Modules with SnO2:F Transparent Conductors, http://www.nrel.gov/docs/fy03osti/33567.pdf.

22

Energy Collected and Delivered by PV Modules

Eduardo Lorenzo

Instituto de Energía Solar, Universidad Politécnica de Madrid, Spain

22.1 INTRODUCTION

This chapter deals with two different questions: 'How much solar radiation reaches the surface of a photovoltaic (PV) collector?' and 'How much of this solar radiation is converted into electricity for a given module configuration?' Both questions are analysed from the PV system designer's point of view, whose ultimate interests, depending on the type of application, are the determination of the PV collector surface needed to provide a certain required service, and the estimation of the average electricity yield from a certain PV array. The basic problem is that modules are rated by their output power (watts or kilowatts) under a single set of very specific testing conditions, but the engineer or owner is interested in the energy (kilowatt hours) produced over time under a widely varying set of conditions.

The input information available for the PV designer is usually restricted to the 12 monthly mean values of global horizontal irradiation and average temperature, that characterise the solar climate of the location, and to the electrical characteristics of the PV modules under Standard Test Conditions (STC), as provided by its manufacturers. This chapter contains a whole set of procedures applicable to most PV engineering practical problems, using only this input information. In particular, models are described to predict the daily, monthly or annual trend of the solar radiation incident over any arbitrarily oriented surface, and also models to anticipate the electricity output of PV arrays operating in other than STC. Looking for practical usefulness, the chapter is, as far as possible, self-contained and the formulae are presented in such a way, that they can be applied directly to practical problems, even if sometimes empirical or cumbersome expressions have to be introduced. Some examples intended to clarify the proper use of the formulae have also been included.

The two questions mentioned above differ intrinsically in nature. The random character of the solar radiation over the Earth implies that answers to the first question are never more than future predictions, unavoidably associated with a degree of uncertainty (even if very good past solar

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

radiation data are available). This uncertainty is more than is commonly assumed, and represents a serious limit to the accuracy (or, more strictly, to the significance) of the results of any PV design exercise, irrespective of the complexity of the radiation and PV array models supporting the particular design tool. This means that for stand-alone PV systems, reliability can only be properly estimated for loss of load probabilities greater than 10^{-2}. For grid-connected systems, it implies that yield predictions should be understood to have an uncertainty that can reach up to $\pm 30\%$, for monthly values, and $\pm 10\%$ for annual ones.

Regarding the behaviour of PV modules in arbitrary operating conditions, a current–voltage ($I-V$) model based on the incident irradiance and ambient temperature is presented. Particular attention is paid to the consideration of incidence angle effects, which can be significant in several real situations. Further refinements to incorporate the effects of wind speed, solar spectrum and low irradiance effects are also described. These second-order effects help to explain temporary differences between predicted and observed values, but are of minor importance when long-term energy production calculations are considered.

Finally, some problems related to relevant applications are particularly addressed: the reliability of stand-alone PV systems, the case of Solar Home Systems (SHS – characterised by a large variety of individual real energy consumption values in contrast to only a few standard ones for design purposes), and the energy yield of grid-connected PV systems.

22.2 MOVEMENT BETWEEN SUN AND EARTH

Although the movement of the sun relative to a fixed point on the Earth seems very familiar, because the sun is fortunately there everyday, the mathematics that governs it is surprisingly complex. In fact, the understanding of such movement was among the longest scientific adventures of mankind. It seems that the very first sundials were built in the Babylonian era (1800 BC), and that the attempts to explain the evolution of the 'gnomon' shadow led to the first proposed models of the Sun–Earth movement, which in turn led to the beginning of geometry, in Greece [1]. Figure 22.1 pays a fully deserved homage to such a glorious background. However, it was not until the seventeenth century,

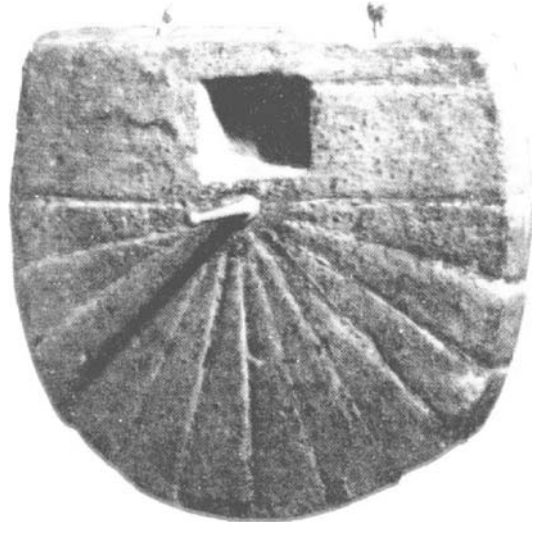


Figure 22.1 Sundial about 2000 years old

when Kepler published *Astronomia Nova* (1609) and *Harmonice mundi* (1618), that such movement was completely explained. And then the explanation spread only very slowly. For example, Galileo in his *Dialogo sopra i due massimi sistemi del mondo*, published in 1632 (just one year before the famous ecclesiastic sentence condemning him for publicly defending the heliocentric system formerly proposed by Copernicus), still promoted the idea of the planets revolving around the sun in circular orbits, fully ignoring Kepler's work.

Fortunately for us, such discussions ended a long time ago. Today, it is well established that the Earth goes around the sun in an elliptic orbit with the sun at one of the foci. The plane containing this orbit is called the *ecliptic plane* and the time that the Earth takes to complete this orbit leads to the definition of the year. The distance from the sun to the Earth r is given by

$$r = r_0 \left[1 + 0.017 \sin\left(\frac{360(d_n - 93)}{365} \right) \right] \tag{22.1}$$

where d_n is the day number counted from the beginning of the year. It is worth noting that the eccentricity of the ecliptic is only 0.017, that is, very small. Because of that, the deviation of the orbit from the circular is also very small, and it is normally adequate to express the distance just in terms of its mean value r_0, equal to 1.496×10^8 km, and is usually referred to as *one astronomical unit*, 1 AU. For most engineering applications, a very simple and useful expression for the so-called eccentricity correction factor is

$$\varepsilon_0 = (r_0 / r)^2 = 1 + 0.033 \cos\left(\frac{360 d_n}{365} \right) \tag{22.2}$$

The Earth also spins once a day on its own central axis, the *polar axis*. The polar axis orbits around the sun, maintaining a constant angle of 23.45° with the ecliptic plane. This inclination is what causes the sun to be higher in the sky in the summer than in the winter. It is also the cause of longer summer sunlight hours and shorter winter sunlight hours. Figure 22.2 shows the Earth's orbit around the sun, with the inclined polar axis; and Figure 22.3 adds some details for a particular day and a particular geographic latitude ϕ. It is important to note that the angle between the equatorial plane and a straight line drawn between the centre of the Earth and the centre of the sun is constantly changing over the year. This angle is known as the *solar declination,* δ. For our present purposes, it may be considered as approximately constant over the course of any one day. The maximum variation in δ over 24 h is less than 0.5°. If angles north of the equator are considered as positive and south of the equator are considered negative, the solar declination can be found from

$$\delta = 23.45^\circ \sin\left[\frac{360(d_n + 284)}{365} \right] \tag{22.3}$$

On the *spring equinox* (20/21 March) and the *autumn equinox* (22/23 September), the line between the sun and the Earth passes through the equator. Consequently, $\delta = 0$, the length of day and night is equal all over the Earth, and the sun rises and sets precisely in the east and west, respectively. On the *summer solstice* (21/22 June in the northern hemisphere) $\delta = 23.45^\circ$, the sun is situated directly above the Tropic of Cancer, and sunrise and sunset are displaced towards the north-east and north-west, respectively. In the Northern Hemisphere, the summer solstice is when the longest day and shortest night of the whole year occur. In the Southern Hemisphere, it is the opposite. On the *winter solstice* (21/22 December) $\delta = -23.45^\circ$, the sun is directly above the Tropic of Capricorn, and sunrise and sunset are displaced towards the south-east and south-west, respectively. In the Northern Hemisphere, this is the shortest day and longest night of the whole year, and again, the opposite is true in the Southern Hemisphere.

A classic way of representing the sky is as a sphere centred on a fixed point of the Earth, as indicated in Figure 22.4. This is known as the *celestial sphere*. Each of its points represents a

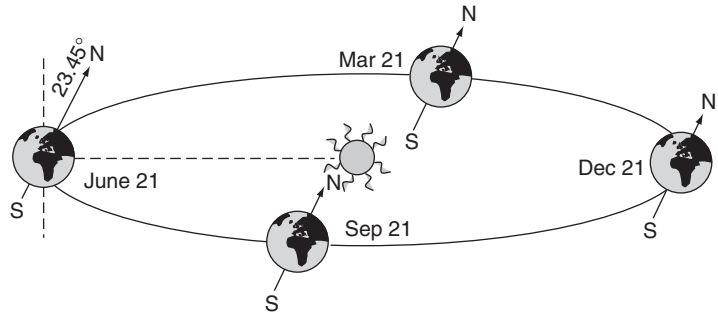


Figure 22.2 The orbit of the Earth around the sun

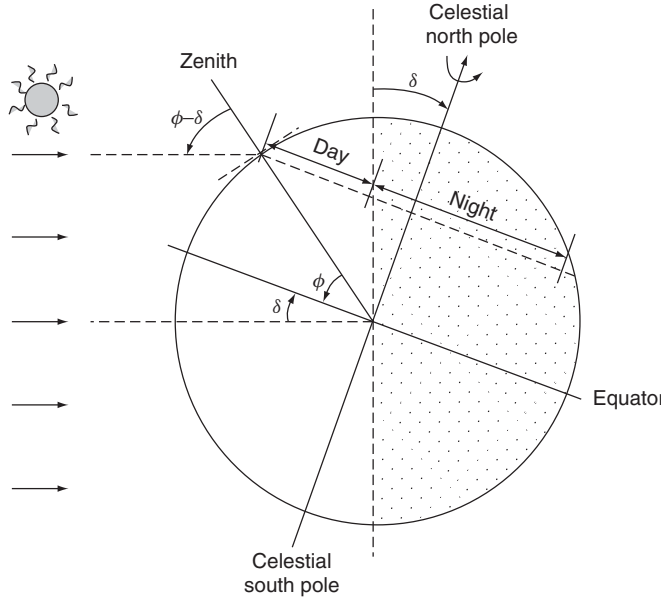


Figure 22.3 Relative Earth–Sun position at noon of a negative declination day (winter in the Northern Hemisphere, and summer in the Southern Hemisphere)

direction towards the sky as seen from the Earth. The intersection of the celestial sphere with the equatorial plane defines the *celestial equator*. The points of intersection with the polar axis are called the *celestial poles*. Using this form of representation, the movement of the Earth around the stationary sun may instead be seen as a movement of the sun with the Earth taken as fixed. The sun then travels through a great circle of the celestial sphere, the *ecliptic*, which forms an angle of 23.45° with the celestial equator. The sun completes this circuit once a year while the celestial sphere rotates once a day around the Earth (regarded as fixed). In this way, the sun marks out a circle around the Earth. The diameter of the circle changes daily, reaching a maximum on the equinoxes and a minimum on the solstices. The rotation of the sun around the ecliptic is in the opposite direction to that of the celestial sphere around the Earth.

Now, landing on a particular location on the Earth's surface, where a PV system is going to be used, it is convenient to specify the position of the sun by means of two angles that refer

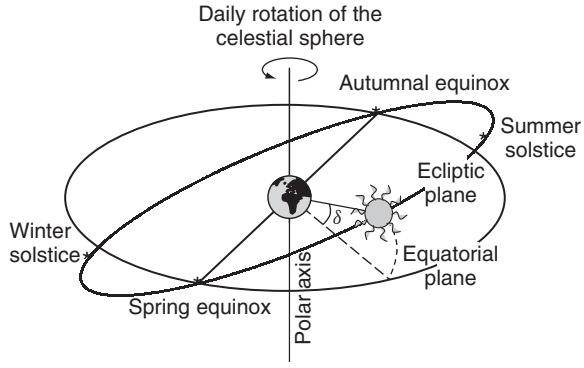


Figure 22.4 The celestial sphere and the ecliptic plane

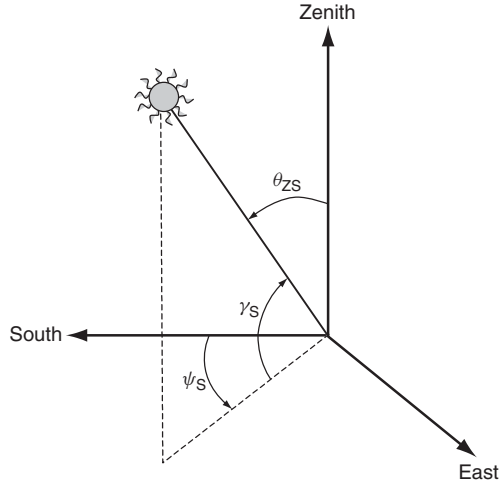


Figure 22.5 Position of the sun relative to a fixed point on the Earth defining the two critical angles ψ_S (azimuth) and θ_{ZS} (solar zenith). The complement of the last, γ_S (altitude) is also shown

to the horizontal plane and to the vertical, respectively. Figure 22.5 attempts to visualise these concepts. The *solar zenith angle* θ_{ZS} is between the vertical and the incident solar beam, i.e., the angle of incidence of beam radiation on a horizontal surface; and the *solar azimuth* ψ_S is between the meridians of the location and the sun, i.e.the angular displacement from noon of the projection of beam radiation on the horizontal plane. The complement of the zenith angle is called the *solar altitude*, γ_S. In the Northern (Southern) Hemisphere, the solar azimuth is referenced to true south (north) not magnetic south (north) and is defined as positive towards the west, that is, in the evening, and negative towards the east, that is, in the morning.

At any given moment, the angular coordinates of the sun with respect to a point of geographic latitude ϕ (north positive, south negative) are calculated from the equations:

$$\cos\theta_{ZS} = \sin\delta\sin\phi + \cos\delta\cos\phi\cos\omega = \sin\gamma_S \tag{22.4}$$

and

$$\cos\psi_S = \frac{(\sin\gamma_S \sin\phi - \sin\delta)}{\cos\gamma_S \cos\phi}[\text{sign}\,(\phi)] \tag{22.5}$$

where ω is called the true solar time, or local apparent time, or solar hour, and is the difference between noon and the selected moment of the day in terms of a 360° rotation in 24 h. $\omega = 0$ at the midday of each day, and is counted as negative in the morning and positive in the afternoon. [sign(ϕ)] means '1' for northern latitudes and '−1' for southern latitudes.

Although not really required for PV calculations, it is worth mentioning the true solar time ω is related to the local official time TO, also called local standard time (the time shown by a clock) by the equation

$$\omega = 15 \times (TO - AO - 12) - (LL - LH) \tag{22.6}$$

where LL is the local longitude and LH is the reference longitude (define this or give an example) of the local time zone (positive towards the west and negative towards the east of the Greenwich Meridian). AO is the time by which clocks are set ahead of the local time zone. In the European Union, AO is usually one hour during winter and autumn, and two hours during spring and summer. In this equation ω, LL and LH are given in degrees, while TO and AO are given in hours (360° means 24 h).

Figure 22.6 presents the sun's trajectory on the celestial sphere for: (a) a winter and a summer day; and (b) the corresponding plots of solar altitude versus azimuth. We will return to such plots later on.

Equation (22.4) may be used to find the *sunrise angle* ω_S since at sunrise $\gamma_S = 0$. Hence

$$\omega_S = -\arccos(-\tan\delta \tan\varphi) \tag{22.7}$$

In accordance with the sign convention, ω_S is always negative. Obviously, the sunset angle is equal to $-\omega_S$ and the length of the day is equal to $2 \times abs(\omega_S)$. In the polar regions, during the winter the sun does not rise ($\tan\delta\tan\phi > 1$) and Equation (22.7) has no real solution. However, for computing purposes, it is convenient to set $\omega_S = 0$. Similarly, during the summer, $\omega_S = -\pi$ is a practical solution for the continuous day. It is also interesting to note that just at noon $\omega = 0$ and the solar altitude is equal to the latitude complement plus the declination

$$\omega = 0 \Rightarrow \gamma_S = \pi/2 - \varphi + \delta \tag{22.8}$$

It should be noted that Equation (22.5) is indeterminate for $\gamma_S = \pi/2$ and for $\phi = \pi/2$. In the first case, the sun is just on the vertical, so that ψ_S is meaningless. In the second, the sun's position is given by $\gamma_S = \delta$ and $\psi_S = \omega$.

Another word of caution is necessary here. All these equations refer to a baseline year composed of 365 days, while the time length of the real year is 365 days, 5 h, 48 min and 45.9 s. As is well known, the difference is compensated by adding a day in the leap years. For most PV engineering applications, the same number corresponding to the precedent may be associated to this additional day, that is $d_n = 59$ for both 28 and 29 February. However, for some demanding applications, such as highly accurate trackers for concentrators, more precise equations may be necessary to calculate the relative Earth–Sun position. The interested reader is encouraged to consult reference [2]. It describes an accurate algorithm and provides a software file for its implementation.

Most practical applications require the position of the sun relative to an inclined plane to be determined. The position of a surface (Figure 22.7) may generally be described by its slope β

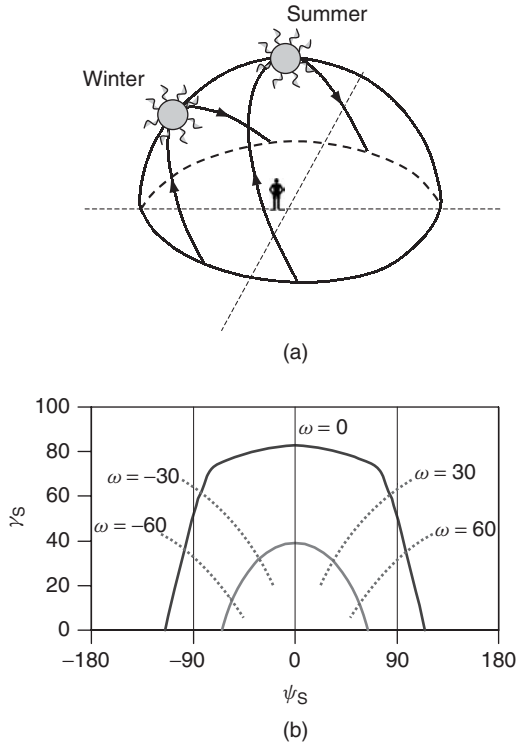


Figure 22.6 (a) Sun's trajectory corresponding to a winter and to a summer day represented over the celestial sphere. (b) The solar altitude γ_S is plotted against the solar azimuth ψ_S. The solar time ω is also shown

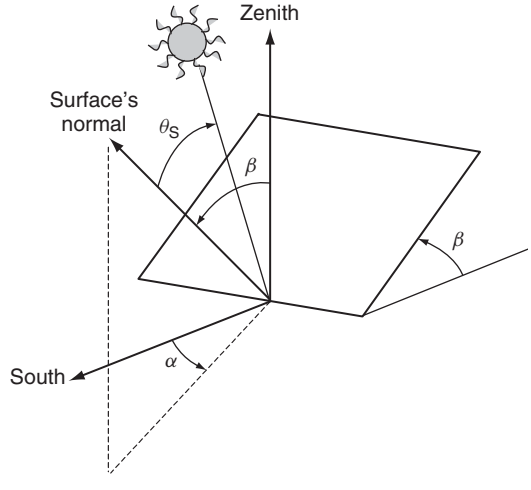


Figure 22.7 Receiver position (slope β and azimuth α) and sun's rays incidence angle θ_S

(the angle formed with the horizontal) and the azimuth α of the normal to the surface. The angle of solar incidence between the sun's rays and the normal to the surface may be calculated from

$$\cos\theta_S = \sin\delta\sin\phi\cos\beta - [\text{sign}(\phi)]\sin\delta\cos\phi\sin\beta\cos\alpha + \cos\delta\cos\phi\cos\beta\cos\omega + [\text{sign}(\phi)]\cos\delta\sin\phi\sin\beta\cos\alpha\cos\omega + \cos\delta\sin\alpha\sin\omega\sin\beta \quad (22.9)$$

Although this expression appears quite complicated, it is very convenient to use in most instances. In the case of surfaces tilted towards the equator (facing south in the Northern Hemisphere, or facing north in the Southern Hemisphere), $\alpha = 0$, and it simplifies to

$$\cos\theta_S = [\text{sign}(\phi)]\sin\delta\sin(abs\,(\phi) - \beta) + \cos\delta\cos(abs(\phi) - \beta)\cos\omega \quad (22.10)$$

22.3 SOLAR RADIATION COMPONENTS

Figure 22.8 helps to illustrate the following brief explanation of the different components of solar radiation that reach a terrestrial flat-plate PV surface

To a good approximation, the sun acts as a perfect emitter of radiation (black body) at a temperature close to 5800 K. The resulting power incident on a unit area perpendicular to the beam outside the Earth's atmosphere, when it is 1 AU from the sun, is known as the *solar constant*

$$B_0 = 1367\ \text{W/m}^2 \quad (22.11)$$

The radiation falling on a receiver situated beyond the Earth's atmosphere, that is, *extraterrestrial radiation*, consists almost exclusively of radiation travelling along a straight line from the sun. Since the intermediate space is almost devoid of material that might scatter or reflect the light, it appears virtually black, apart from the sun and faint points of light corresponding to the stars.

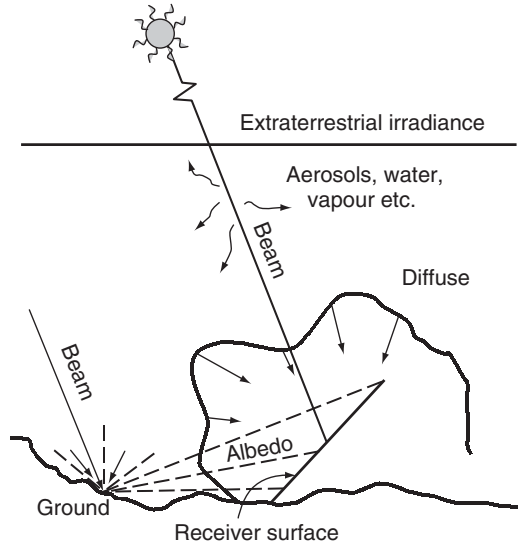


Figure 22.8 Different components of solar radiation

As the solar radiation passes through the Earth's atmosphere, it is modified by interaction with components present there. Some of these, such as clouds, reflect radiation. Others, for example, ozone, oxygen, carbon dioxide and water vapour, have significant absorption at several specific spectral bands. Water droplets and suspended dust also cause scattering. The result of all these processes is the decomposition of the solar radiation incident on a receiver at the Earth's surface into clearly differentiated components. *Direct or beam radiation*, made up of beams of light that are not reflected or scattered, reaches the surface in a straight line from the sun. *Diffuse radiation*, coming from the whole sky apart from the sun's disc, is the radiation scattered towards the receiver. *Albedo radiation* is radiation reflected from the ground. The total radiation falling on a surface is the sum of these (direct + diffuse + albedo) and is termed *global radiation*.

It is intuitively obvious that the directional properties of the diffuse radiation depend to a large extent on the position, form and composition of the water vapour and dust responsible for scattering. The angular distribution of the diffuse radiation is therefore a complex function that varies with time. Diffuse radiation is essentially anisotropic. The amount of albedo radiation is greatly affected by the nature of the ground, and a wide range of features (snow, vegetation, water, etc.) occur in practice.

In the following discussion, the word *radiation* will be used as a general term. To distinguish between power and energy, more specific terminology will be used. *Irradiance* means density of power falling on a surface, and is measured in W/m^2 (or similar); whereas *irradiation* is the density of the energy that falls on the surface over some period of time, for example, hourly irradiation or daily irradiation, and is measured in $W\,h/m^2$. Furthermore, only the symbols B_0, B, D, R and G will be used, respectively, for extraterrestrial, direct, diffuse, albedo and global irradiance, whereas a first subscript, h or d, will be used to indicate hourly or daily irradiation A second subscript, m or y, will refer to monthly or yearly averaged irradiation values. Furthermore, the slope and orientation of the concerned surface are indicated among brackets. For example, $G_{dm}(20,40)$ refers to the monthly mean value of the daily global irradiation incident on a surface tilted $\beta = 20^\circ$ and oriented $\gamma = 40^\circ$ towards the west. For surfaces tilted towards the equator ($\gamma = 0$), only the slope will be indicated. For example, $B(60)$ refers to the value of the direct irradiance incident on a surface tilted $\beta = 60^\circ$ and oriented towards the south (in the Northern Hemisphere).

An important concept characterising the effect of atmosphere on clear days is the *air mass*, defined as the relative length of the direct-beam path through the atmosphere compared with a vertical path directly to sea level, which is designed as AM. For an ideal homogeneous atmosphere, simple geometrical considerations lead to

$$AMI = 1/\cos\theta_{ZS} \tag{22.12}$$

which is generally sufficient for most engineering applications. If desired, more accurate expressions, considering second-order effects (curvature of the Earth, atmospheric pressure etc.), are available [3].

At the standard atmosphere $AM\,1$, after absorption has been accounted for, the normal irradiance is generally reduced from B_0 to $1000\,W/m^2$, which is just the value used for the standard test of PV devices (see Chapter 16). Obviously, that can be expressed as $1000 = 1367 \times 0.7^{AM}$. For general AM values, a reasonable fit to observed clear days data is given by [4].

$$G = B_0.\varepsilon_0 \times 0.74^{AM0.678} \tag{22.13}$$

A particular example can help to clarify the use of these equations, by calculation of the sun coordinates and the global irradiance on a surface perpendicular to the sun, and also on a horizontal

surface, over two geographic positions defined by $\phi = 30^\circ$ and $\phi = -30^\circ$, at 10:00 (solar time) on 14 April, being a clear day. The solution is as follows:

$$\begin{aligned}
14\ \text{April} \Rightarrow d_n = 104;\ \varepsilon_0 = 0.993;\ \delta = 9.04^\circ & \\
10:00\,\text{h} \Rightarrow \omega = -30^\circ & \\
\phi = 30^\circ \Rightarrow \cos\theta_{ZS} = 0.819 \Rightarrow \theta_{ZS} = 35^\circ \Rightarrow \cos\psi_S = 0.508 \Rightarrow \psi_S = -59.44^\circ & \\
\Rightarrow AM = 1.222 \Rightarrow G = 902.4\,\text{W/m}^2 & \\
\Rightarrow G(0) = G \cdot \cos\theta_{ZS} = 739\,\text{W/m}^2 & \\
\phi = -30^\circ \Rightarrow \cos\theta_{ZS} = 0.662 \Rightarrow \theta_{ZS} = 48.54^\circ \Rightarrow \cos\psi_S = 0.403 \Rightarrow \psi_S = -66.28^\circ & \\
\Rightarrow AM = 1.510 \Rightarrow G = 846.9\,\text{W/m}^2 & \\
\Rightarrow G(0) = G \cdot \cos\theta_{ZS} = 561\,\text{W/m}^2 &
\end{aligned}$$

When solar radiation enters the Earth's atmosphere, not only the irradiance, but also the spectral content is affected. Figure 22.9 shows the *AM* 1.5 spectrum, which is considered for standard test of PV devices. Figure 16.1 shows other spectra for comparison. In general, increasing air mass displaces the solar spectrum towards the red. This is why the sky becomes so nice at nightfall.

Of course, PV devices are sensitive to the spectrum, as discussed in Chapters 3, 9, 12 and 16. However, this is of little importance from the PV engineering point of view, compared with changes in total radiation incident on the PV modules. Because of that, in what follows, we will omit the detailed treatment of the spectral composition of sunlight. Additional comments will be given later on in this chapter.

22.4 SOLAR RADIATION DATA AND UNCERTAINTY

The amount of global radiation that reaches the receiver is extremely variable. On the one hand, even the extraterrestrial radiation experiences regular daily and yearly variations due to the apparent motion of the sun. These variations are predictable and can be theoretically determined just by geometrical considerations. For example, the extraterrestrial irradiance over a horizontal surface is given by

$$B_0(0) = B_0 \varepsilon_0 \cos\theta_{ZS} \tag{22.14}$$

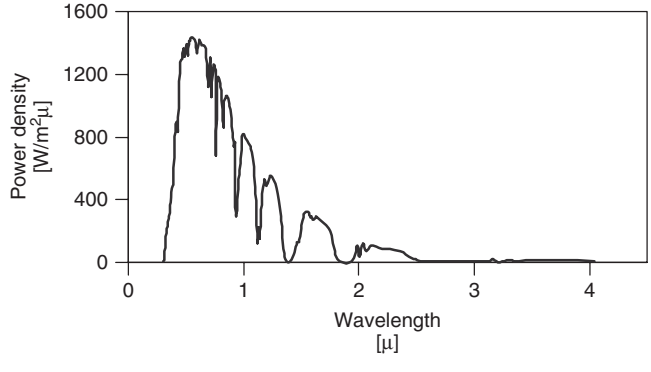


Figure 22.9 *AM* 1.5 solar spectrum, for an irradiance $G = 1000\,\text{W/m}^2$

which when integrated over the day, leads to [5]

$$B_{0d}(0) = \frac{T}{\pi} B_0 \varepsilon_0 \left[-\frac{\pi}{180}\, \omega_S \cdot \sin\delta \cdot \sin\phi - \cos\delta \cos\phi \sin\omega_S \right] \tag{22.15}$$

where T is the day length, that is, 24 h. The monthly average of this quantity, named $B_{0dm}(0)$, is of particular practical importance. It represents the average daily energy on a horizontal surface for that month. Obviously

$$B_{0dm}(0) = \frac{1}{d_{n2} - d_{n1} + 1} \sum_{d_{n1}}^{d_{n2}} B_{0d}(0) \tag{22.16}$$

where d_{n1} and d_{n2} are the day numbers of the first and last day of the month, respectively. It is useful to know that, for a given month, there is a day for which $B_{0d}(0) = B_{0dm}(0)$. It can be demonstrated that this day is the one whose declination equals the mean declination for the month. Table 22.1 shows the day number of this day and the corresponding value of $B_{0dm}(0)$ for each month of the year at several latitudes.

We have also shown that the effect of clear cloudless skies can be predictably accounted for by a single geometrical parameter, namely, the air mass (Equation 20.12). On the other hand, there are random variations caused by climatic conditions: cloud cover, dust storms and so on so that the PV systems design should rely on the input of measured data close to the site of the installations and averaged over a long time. This is routinely done by the National Meteorological Services (or similar services), which use a variety of instruments and procedures, from direct sunlight measurements using devices such as pyranometers, pyrheliometers, etc. to correlations with other meteorological variables (hours of sunshine, cloudiness, tone of satellite photographs, etc). Then, they are treated to derive some representative parameters, which are made publicly available by different ways: World radiation databases [6, 7], Radiation Atlas [8–11], web sites [12, 13], local information[14] and so on. The 12 monthly mean values of global horizontal daily irradiation, $G_{dm}(0)$, today represent the most widely available information concerning the solar radiation resource, and that is likely to remain in the years to come. It is important to note that solar radiation unavoidably represents a source of uncertainty for PV systems designers.

Table 22.1 Declination and extraterrestrial irradiation values for the characteristic day of each month

Month	Date	d_n	δ [°]	$B_{0d}(0) = B_{0dm}(0)$, in [W h/m^2]			
				$\phi = 30°$	$\phi = 60°$	$\phi = -30°$	$\phi = -60°$
January	17	17	−20.92	5 907	949	11 949	11 413
February	14	45	−13.62	7 108	2 235	11 062	9 083
March	15	74	−2.82	8 717	4 579	9 531	5 990
April	15	105	+9.41	10 225	7 630	7 562	3 018
May	15	135	+18.79	11 113	10 171	5 948	1 225
June	10	161	+23.01	11 420	11 371	5 204	605
July	18	199	+21.00	11 224	10 741	5 530	878
August	18	230	+12.78	10 469	8 440	6 921	2 294
September	18	261	+1.01	9 121	5 434	8 835	4 937
October	19	292	−11.05	7 436	2 726	10 612	8 226
November	18	322	−19.82	6 056	1 114	11 754	10 983
December	13	347	−23.24	5 498	613	12 174	12 177

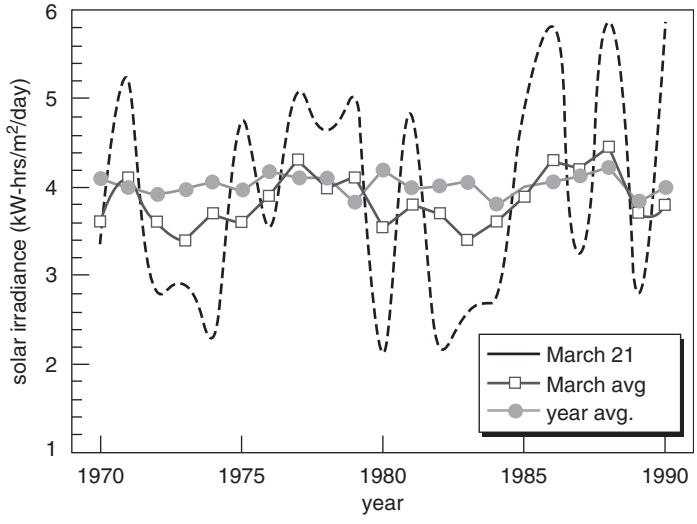


Figure 22.10 Daily global horizontal irradiation in Philadelphia, Pennsylvania (eastern USA) for the years 1970–1990, in kW h/m^2. 21 March, $G_d(0)$; March average, $G_{dm}(0)$; and year average, $G_{dy}(0)$ values are depicted.

As a representative case, Figure 22.10 presents daily global horizontal irradiation in Philadelphia (USA) for the years 1970–1990. The black line shows the $G_d(0)$ values corresponding to 21 March. That is properly described by the *average* and the *standard deviation* values, respectively, $\overline{G_a(0)} = 3.92\,\text{kWh/m}^2$ and $\sigma_d = 1.34$ kWh/m^2. We could then give the average as our estimate for the future. However, if someone now asked the question, 'What is the probability that, next 21 March, $G_d(0)$ will be exactly 3.92 kW h/m^2?', we would have to answer – somewhat uncomfortably – that the probability is no doubt close to zero. However, we would have to hasten to add that if we relax the prediction rigor, saying, for example, that the next (and the future) March 21$G_d(0)$ will be within the range 2–6 kW h/m^2, as in Figure 22.10, then the probability is high. Note that this range represents about ±50% of the average value, much larger than the accuracy of the original data! Then, our questioner's retort would probably be, 'How "high" is the probability and how "large" is the range?'.

Basic statistical theory gives proper answers to this question. If we desire a 95% probability (the *confidence coefficient*) we must keep a $\pm 2\sigma$ range (the *confidence interval*) around the average value. In our example, $\pm 2\sigma$ represents ±68% of the average value. Such large uncertainty becomes reduced when not individual but monthly averaged values are considered. Figure 22.10 also shows the $G_{dm}(0)$values corresponding to March average. This time, $\overline{G_{dm}}\,(0) = 3.84\,\text{kW h/m}^2$ and $\sigma_m = 0.3\,\text{kW h/m}^2$. Note than now $\pm 2\sigma$ represents ±15% of the average value. On the same lines, uncertainty is further reduced when not monthly but yearly averaged values are considered. $G_{dy}(0)$ values are also depicted in Figure 22.10. Now, $\overline{G_{dy}}(0) = 4.10\,\text{kW h/m}^2$ and $\sigma_m = 0.11\,\text{kW h/m}^2$, i.e. $\pm 2\sigma$ represents ±5.4% of the average value. It is worth mentioning that 21 March has close the same average as the annual because it is an equinox day, thus both 21 March and the entire year have on average 50% sunlight and 50% dark. But they have vastly different standard deviations, hence differences in confidence in prediction of their irradiation.

Uncertainty also becomes reduced by enlarging the number of years for which the prediction applies. That is, to give an estimate not for the next year, but for the average values of the following years. Then, it can be show that the confidence interval becomes reduced by a factor of

Table 22.2 Energy generation, in kW h, of a commercial PV system (peak power 6.12 kW; one vertical axis tracking) connected to the Spanish grid (reference [13])

	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Year
Estim.	651	800	970	1015	1187	1266	1299	1149	898	878	648	504	11 356
Exp04	318	640	821	1064	1066	1459	1318	1304	986	905	690	338	10 910
Dif(%)	*−104*	*−25*	*−18.1*	*4.6*	*−11.3*	*0.13*	*1.4*	*11.9*	*8.9*	*3*	*6*	*−49.1*	*−4*
Exp05	430	881	994	939	1197	1337	1573	1251	1160	777	581	566	11 686
Dif(%)	*−51.4*	*9.2*	*2.4*	*−8.1*	*0.8*	*5.3*	*17.4*	*8.1*	*14.7*	*−13*	*−11.5*	*10.9*	*2.8*
Exp06	355	805	697	1143	1179	1410	1141	0	1586	1120	674	585	10 695
Dif(%)	*−83.4*	*−0.6*	*−39.2*	*11.1*	*−0.8*	*10.2*	*−13.8*	*∞*	*43.4*	*21.6*	*3.9*	*13.8*	*−6.1*
Exp07	558	458	625	836	989	1102	1260	1364	1295	1259	961	849	11 556
Dif(%)	*−16.7*	*−74.4*	*−55.2*	*−21.4*	*−20*	*−14.9*	*−3.1*	*12.5*	*30.7*	*30.3*	*32.6*	*40.6*	*1.7*
Exp08	595	538	753	783	975	952	1257	1349	1410	1042	787	661	11.102
Dif(%)	*−9.4*	*−48.7*	*−28.8*	*−29.6*	*−21.7*	*−33.0*	*−3.3*	*14.8*	*36.3*	*15.7*	*−17.7*	*23.7*	*−2.3*
Ex-avg	*451*	*664*	*778*	*953*	*1081*	*1252*	*1310*	*1054*	*1287*	*1021*	*739*	*600*	*11.190*
Dif(%)	*−44.3*	*−20.5*	*−24.7*	*−6.5*	*−9.8*	*−1.1*	*0.8*	*−9.1*	*30.3*	*14*	*12.3*	*16*	*−1.5*

$1/\sqrt{N}$, N being the number of future years for which the prediction is extended. Coming back to our example, and considering a future length of 10 years, the correct predictions for 21 March, the March average and yearly average would be, respectively, $\overline{G_{\mathrm{d}}(0)} = 3.92\,\mathrm{kW\,h/m^2} \pm 22\%$, $\overline{G_{\mathrm{dm}}(0)} = 3.84\,\mathrm{kW\,h/m^2} \pm 5\%$ and $\overline{G_{\mathrm{dy}}(0)} = 4.1\,\mathrm{kW\,h/m^2} \pm 1.7\%$

As representative example of the practical consequences of solar radiation variability, Table 22.2 presents the monthly energy generation (as indicated in the utility bills) of a commercial PV system (peak power 6.12 kW; one vertical axis tracking) connected to the Spanish grid. The first line gives the initial estimation based on past averaged solar radiation values, as given in an available database13 (reference [13]). Lines 2–6 give the real production from January 2004 to December 2008. Line 7 gives the average production along these years. Italic numbers in lines 2 to 7 are the differences between estimation and realities (positive means overestimation). The corresponding error is very large for individual months (+104, – 43.4), still large for average months (+44.3, −30.3), low for individual years (+6.1, −2.8) and still lower for average year (1.5).

A particularly annoying consequence of solar radiation variability is large disparity found for the same location when different publications are consulted. Disparity derives from the random nature of solar radiation: different data sources mean different data recording periods for computing representative values. Moreover, practical difficulties on solar radiation measurements (dust on sensors, calibration errors, losses of data, geographical interpolations, etc.) also contribute to this confusing result. Disparity is larger for monthly that for yearly values. This is particularly troublesome for sizing the PV generator of stand-alone PV systems. For that, it is customary to select the so-called worst month, that is the month with the lowest value of $G_{\mathrm{dm}}(0)$. As an example, data for Madrid are given in references [6–9, 12–14]. $G_{\mathrm{dy}}(0)$ vary from 4.29 to 4.53 kW h/m^2, or ±3% around the mean. But $G_{\mathrm{dm}}(0)$ for December, which is the worst month vary from 1.55 to 1.8 kW h/m^2. Irrespective of the sizing methodology, variability in the choice of the particular data source would lead to PV array size differences up to 16% [(1.55–1.80)/1.55].

Finally, it is worth mentioning that solar radiation information based on satellite images is becoming increasingly available, and is improving on ground coverage and resolution. For example, NASA [12] makes freely available estimates of global irradiation for the world, on a grid of

cells, each 1° latitude $\times 1^\circ$ degree longitude. PVGIS [13] makes the same for Europe, Africa and Mediterranean regions, on a grid of cells, each 1×1 km.

22.4.1 Clearness Index

The relation between the solar radiation at the Earth's surface and the extraterrestrial radiation gives a measure of the atmospheric transparency. In this way, a *clearness index* K_{Tm} is calculated for each month:

$$K_{\mathrm{Tm}} = \frac{G_{\mathrm{dm}}(0)}{B_{0\mathrm{dm}}(0)} \tag{22.17}$$

Note that the clearness index is physically related not only to the radiation path through the atmosphere, that is, with the AM value, but also with the composition and the cloud content of the atmosphere. Liu and Jordan [15] have demonstrated that, irrespective of latitude, the fractional time during which daily global radiation is equal to or less than a certain value is directly dependent on this parameter. Because of that, K_{Tdm} can properly characterise the solar climate of a particular location. This provides the basis for estimating solar radiation on inclined surfaces.

22.5 RADIATION ON INCLINED SURFACES

To eliminate the effects of varying local features, such as obstructions that cast shadows and the specific ground covering, solar radiation is routinely measured on horizontal surfaces free of obstacles. Consequently, solar radiation data are most often given in the form of global radiation on a horizontal surface. Since PV modules are usually positioned at an angle to the horizontal plane, the radiation input to the system must be calculated from the data.

The assessment of radiation arriving on an inclined surface, using as input global horizontal data, raises two main problems: to separate the global horizontal radiation into their direct and diffuse components; and, from them, to estimate the radiation components falling on an inclined surface. In general, these problems may be posed for different time scales, for example, daily irradiation, hourly irradiation and so on. Individual or time-averaged values may be sought. Here, we will first focus on the monthly average daily irradiation values. This is not only convenient for presentation purposes, but also coherent with solar radiation data availability, and particularly suited to most PV engineering practical problems. Additional comments for other cases will also be given afterwards.

22.5.1 Estimation of the Direct and Diffuse Components of Horizontal Radiation, Given the Global Radiation

The underlying concept is the one originally proposed by Liu and Jordan [15]. It consists of establishing empirical correlation between the *diffuse fraction of horizontal radiation*, $F_{\mathrm{Dm}} = D_{\mathrm{dm}}(0)/G_{\mathrm{dm}}(0)$, (diffuse radiation/global radiation) and the *clearness index* (global radiation/extraterrestrial radiation) defined in Equation (22.17). Note that the clearer the atmosphere, the higher the radiation and the lower the diffuse content. Hence, F_{Ddm} and K_{Tdm} are expected to be negatively correlated. Actual analytical expressions are established from the comparison of simultaneous measurements of global and diffuse radiation performed in certain places. Liu

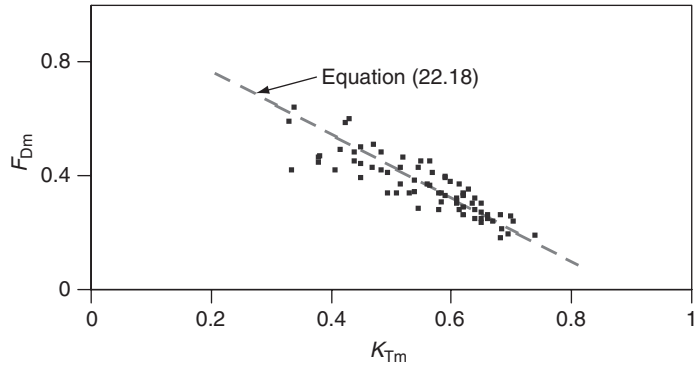


Figure 22.11 The diffuse fraction of the mean daily global irradiation F_{Dm} is plotted against the clearness index K_{Tm}. The cluster of points refers to measured values in Madrid from 1977 to 1988

and Jordan were, in fact, very clever in selecting the clearness index (which they called the cloudiness index) to characterise the solar climate at a particular location, because the division by the extraterrestrial radiation eliminates the radiation variations due to the apparent motion of the sun. This way, the correlation between F_{Dm} and K_{Tm} becomes independent of latitude effects, and, in principle, tends to be of universal validity. Figure 22.11 plots Equation (22.18) and a set of experimental points, obtained in Madrid, from 1977 to 1988,

Various empirical formulae are available. Using data from ten locations situated between 40 °N and 40°S, Page [16] recommended a liner equation that has been frequently identified as the one giving good results.

$$F_{Dm} = 1 - 1.13\,K_{Tm} \tag{22.18}$$

A little bit more complex, the correlation proposed by Erbs *et al.* [17] is now widely used: for $\omega_S \leq 81.4^\circ$ and $0.3 \leq K_{Tm} \leq 0.8$

$$F_{Dm} = 1.391 - 3.560\,K_{Tm} + 4.189 - 2.137$$

for $\omega_S > 81.4^\circ$ and $0.3 \leq K_{Tm} \leq 0.8$

$$F_{Dm} = 1.311 - 3.022\,K_{Tm} + 3.427 - 1.821 \tag{22.19}$$

Local correlations that are derived from the data of only one site can also be found. As an example, Macagnan [14] proposed the following for Madrid:

$$F_{Dm} = 0.758 - 0.428\,K_{Tm} - 0.503 \tag{22.20}$$

Faced with a diversity of correlations, the implications of using one or another is in fact very minor. As a representative exercise, we have calculated the monthly mean daily irradiation incident on a surface tilted to the latitude in Djelfa, Algeria ($\phi = 34.6^\circ$), for a winter and a for a summer month, January ($d_n = 17$) and June ($d_n = 161$), respectively, using the different correlations defined by Equations (22.18) to (22.20). $G_d(0) = 2778$ and 6972 W h/m^2 for the winter and for the summer month, respectively. This radiation data has been obtained from reference [6]. Details of

the calculating procedure will be given in the next section, but they are not relevant for the present discussion. The results are as follows:

Eq.(22.15); $\delta_n = 17 \Rightarrow B_{0dm}(0) = 5157\ \text{Wh/m}^2$

$G_{dm}(0) = 2778\ \text{Wh/m}^2 \Rightarrow K_{Tm} = 0.539$ Eq. (22.17)

22.20	0.382	1060	4463

$\delta_n = 161 \Rightarrow B_{0dm}(0) = 11525\ \text{Wh/m}^2$

$G_{dm}(0) = 6972\ \text{Wh/m}^2 \Rightarrow K_{Tm} = 0.605$

22.20	0.315	2196	6045

Up to 10% differences are observed in the estimation of the diffuse component of the horizontal irradiation. However, the key point is that these differences are significantly reduced (below 2%) in the estimation of the global irradiation.

Empirical correlations between the diffuse fraction of the horizontal irradiation and the clearness index can also be derived for an individual day (14 April is, for example, an individual day, while generic April is a monthly mean day). They should be mentioned here because they have been the object of extensive research in the general solar radiation community. However, their advantage for PV design purposes is far from clear. As a matter of fact, the electrical behaviour of PV generators is mainly governed by the linear dependence on incident irradiance. Thus, the ratio between daily delivered energy and daily incident irradiation tends to remain constant. Second-order effects, associated with ambient temperature or operation voltage fluctuations, do not significantly alter this idea, provided that they are adequately considered, as discussed later. Because of that, rather few benefits, in terms of long-term performance predictions, should be expected from analysing all the days of a month instead of only the mean day of such a month. Because of that, we will restrict here to mention that the most frequently referred correlation for daily values is that put forward by Collares Pereira and Rabl [18], using data from five stations located in the United States. It is expressed as

$$F_{Dd} = 0.99 \quad \text{for} \quad K_{Td} \leq 0.17$$

$$F_{Dd} = 1.188 - 2.272\,K_{Td} + 9.473\,K_{Td}^2 - 21.856\,K_{Td}^3 + 14.648\,K_{Td}^4 \quad \text{for} \quad 0.17 < K_{Td} < 0.8 \tag{22.21}$$

Finally, it should also be mentioned that not only daily, but also hourly based correlations have been proposed. These are correlations between the diffuse fraction of the hourly horizontal global irradiation, $F_{Dh} = D_h(0)/G_h(0)$, and the hourly clearness index, $K_{Th} = G_h(0)/B_{0h}(0)$. However, none of the hourly based correlations proposed so far is really satisfactory [55], so that the associated complexity is not justified. Hence, they will be not considered here.

22.5.2 Estimation of the Instantaneous Irradiance from the Daily Irradiation

In some cases, the treatment of solar radiation is more easily understood at the instant time scale, that is, at the radiance level. Because irradiation over an hour (in Wh/m^2) is numerically equal to the mean irradiance during this hour (in W/m^2), irradiance values can be, to a certain extent, assimilated to hourly irradiation values. However, since the availability of hourly irradiation data is limited, the problem is how to estimate the hourly irradiation, given the daily irradiation.

To introduce the solution to this problem, it is highly instructive to observe that, in terms of extraterrestrial horizontal radiation, the ratio between irradiance $B_0(0)$ and daily irradiation $B_{0d}(0)$ can be theoretically determined from Equations (22.4), (22.14) and (22.15).

$$\frac{B_0(0)}{B_{0d}(0)} = \frac{\pi}{T} \times \frac{\cos\omega - \cos\omega_S}{\left(\frac{\pi}{180}\,\omega_S \cos\omega_S - \sin\omega_S\right)} \tag{22.22}$$

where the sunrise angle ω_S is expressed in degrees and T, the day length, is usually expressed in hours.

From the examination of data from several stations, it has been repeatedly noted [18] that, considering long-term averages of terrestrial radiation, the correspondence between the measured ratio of diffuse irradiance to diffuse daily irradiation, $r_D = D(0)/D_d(0)$, and this theoretical expression for extraterrestrial radiation (Equation 22.22) is quite good, while the correspondence between the measured ratio of global irradiance to global daily irradiation, $r_G = G(0)/G_d(0)$, and this expression, although not perfect, is quite close, so that a slight modification is required to fit the observed data. The following expressions apply:

$$r_D = \frac{D(0)}{D_d(0)} = \frac{B_0(0)}{B_{0d}(0)} \tag{22.23}$$

and

$$r_G = \frac{G(0)}{G_d(0)} = \frac{B_0(0)}{B_{0d}(0)}(a + b\cos\omega) \tag{22.24}$$

where a and b are obtained from the following empirical formulae:

$$a = 0.409 - 0.5016 \times \sin(\omega_S + 60) \tag{22.25}$$

and

$$b = 0.6609 + 0.4767 \times \sin(\omega_S + 60) \tag{22.26}$$

Note that r_D and r_G have units of T^{-1}, and that they can be extended to calculate irradiations during short periods centred on the considered instant ω. For example, if we wish to evaluate the irradiation over one hour between 10:00 and 11:00 (in solar time), we set $\omega = -22.5^\circ$ (the centre time of the considered period is 10:30, i.e. one hour and a half, or 22.5°, before noon) and $T = 24$ h. If we wish to evaluate the irradiation over one minute, we just have to express T in minutes, that is, we set it to 1440, the number of minutes in a day.

An example can help in the use of these equations: the calculation of the irradiance components at several moments along the 15 April in Portoalegre, Brazil ($\phi = -30^\circ$), knowing the global daily irradiation, $G_d(0) = 3861$ W h/m^2. The results are as follows:

$$d_n = 105 \Rightarrow B_{0d}(0) = 7562\ \text{Wh/m}^2 \Rightarrow K_{Td} = 0.5106 \Rightarrow F_{Dd} = 0.423$$
$$\Rightarrow D_d(0) = 1633\ \text{Wh/m}^2$$
$$\omega_S = -84.51^\circ$$
$$\frac{\pi}{180}\omega_S \cos\omega_S - \sin\omega_S = 0.8542$$
$$a = 0.6172$$
$$b = 0.4672$$
$$r_D = 0.0922(\cos\omega + 0.0967)\ \text{h}^{-1}$$
$$r_G = r_D(a + b\cos\omega)$$

ω°	r_D [h^{-1}]	r_G [h^{-1}]	$D(0)$ in [$W\,m^{-2}$]	$G(0)$ in [$W\,m^{-2}$]	$B(0)$ in [$W\,m^{-2}$]
ω_s	0	0	0	0	0
±60	0.0618	0.0529	100.94	204.25	103.31
±30	0.1177	0.1211	192.24	467.58	275.34
0	0.1382	0.1508	225.73	582.24	356.51

Figure 22.12 plots r_D and r_G versus the solar time, along the day. It is interesting to observe that r_G is slightly more sharp-pointed than r_D. This is because, due to air mass variations, beam transmittance is higher at noon that at any other moment of the day. Obviously, on integrating, the areas below both, r_D and r_G must be equal to one. As already mentioned, for this calculation it can be assumed that the irradiation over one hour (in $W\,h/m^2$) is numerically equal to the mean irradiance during this hour and also equal to the irradiance at the instant half way through the hour. For example, the global irradiance at noon, $G(0) = 580.4\,W/m^2$ can be identified with the hourly irradiation from 11:30 to 12:30, $G_h(0) = 580.4\,W\,h/m^2$. This assumption does not introduce significant errors and it greatly simplifies the calculations by eliminating the need to evaluate integrals with respect to time, which, otherwise, can be quite tedious.

From these equations, it can be deduced that, on any day of the year and anywhere in the world, 90% of the total global horizontal irradiation is received during a period centred around midday and of length equal to two-thirds the total sunlight day length. Consequently, a stationary receiver tilted to the equator ($\alpha = 0$) captures all the useful energy in this period. The same need not be true, however, of receivers that track the sun.

It should be mentioned that the 'average day' profile given by Equations (22.23) and (22.24) preserves the observed monthly averaged irradiance values, but not the observed distributions of the instantaneous irradiances. Theoretically, this represents a potential source of error on energy estimations, at the extent of PV modules and inverter efficiency depend on the irradiance in a non-linear fashion. To overcome this problem, some authors [19] consider a weather pattern consisting only of instances of clear sky and instances of overcast conditions with almost zero irradiance. Depending on the frequency of the clear-sky conditions, the average could take any value. This

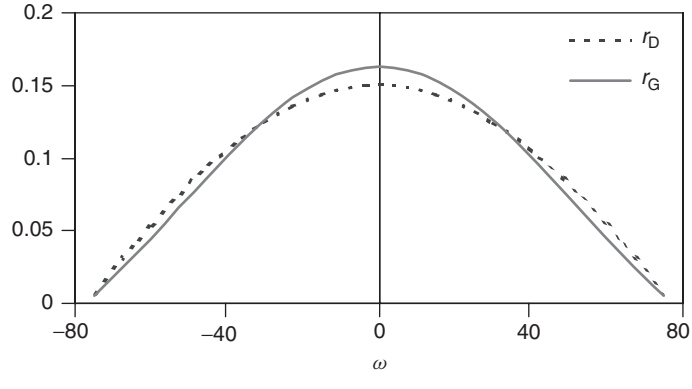


Figure 22.12 Plots of irradiance to daily irradiation ratios, for both diffuse and global radiation, r_D and r_G, during the day at a latitude $\phi = -30^\circ$ for 15 April

way, energy estimations for a month are performed on the basis of its 'clear-sky' profile and clear-sky index (ratio between the observed average daily irradiation $G_{dm}(0)$ and the daily irradiation given by the integral of the clear-sky profile). A very simple clear-sky model [20] consists of determining the direct component of the horizontal irradiance by means of Equation (22.13) and the diffuse component as 20% of the direct component. More accurate models [21] require information concerning the turbidity of the sky, which are today available for only some regions. The irradiance profile affects estimations of irradiation over inclined surfaces, inverter efficiency and PV module efficiency and spectral response. Nevertheless, it should be noted that nonlinearities on efficiency–irradiance relations are usually small, so the energy impact of the irradiance profile is also small, typically bellow 2% in terms of daily energy values,

22.5.3 Estimation of the Radiation on Surfaces on Arbitrary Orientation, Given the Components Falling on a Horizontal Surface

The most obvious procedure for calculating the global irradiance on an inclined surface $G(\beta, \alpha)$ is to obtain separately the direct, diffuse and albedo components, $B(\beta, \alpha)$, $D(\beta, \alpha)$ and $R(\beta, \alpha)$, respectively. Once these are known

$$G(\beta, \alpha) = B(\beta, \alpha) + D(\beta, \alpha) + R(\beta, \alpha) \tag{22.27}$$

22.5.3.1 Direct irradiance

Straightforward geometrical considerations lead to

$$B(\beta, \alpha) = B \max(0, \cos \theta_S) \tag{22.28}$$

where B is the direct irradiance falling on a surface perpendicular to the sun's rays, and $\cos \theta_S$ is the angle of incidence between the sun's rays and the normal to the surface, given by Equation (22.9). B can be obtained from the corresponding value on a horizontal surface

$$B = B(0)/\cos \theta_{ZS} \tag{22.29}$$

Note that when the sun is illuminating the back of the surface (for example, throughout the morning on a vertical surface oriented to the west) $|\theta_S| > \pi/2$. Then, $\cos \theta_S > 0$ and $B = 0$. In this way, the factor $\max(0, \cos \theta_S)$ reflects that the irradiance incident on the back surface of PV modules is not normally utilised.

22.5.3.2 Diffuse irradiance

We can assume that, when the sun is occulted, the sky is composed of elemental solid angles, such as $d\Omega$ (Figure 22.13) from which a diffuse radiance $L(\theta_Z, \psi)$ is emanated towards the horizontal surface. The term radiance is taken to mean the flux of energy per unit solid angle crossing a surface normal to the direction of the radiation. It is expressed in W/m^2 steradian.

Then, the horizontal irradiance is equal to the integral of the contribution of each solid angle, and can be written as

$$D(0) = \int_{sky} L(\vartheta_Z, \psi) \cos \vartheta_Z \, d\Omega \tag{22.30}$$

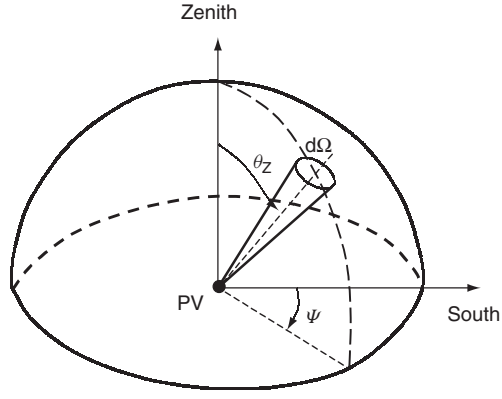


Figure 22.13 Angular coordinates θ_Z and ψ of an elemental solid angle of the sky

where the integral is extended to the whole sky, that is, $0 \leq \theta_Z \leq \pi/2$ and $0 \leq \psi \leq 2\pi$. When we are dealing with an inclined surface, a similar reasoning leads to

$$D(\beta, \alpha) = \int_{\propto} L(\vartheta_Z, \psi) \cos \vartheta'_Z \, d\Omega \tag{22.31}$$

where ϑ'_Z is the incident angle from the solid angle element to the inclined surface, and $\propto$ means that the integral is extended to the non-obstructed sky. The general solution of this equation is difficult because, under realistic skies, the radiance is not uniform and varies with the sky condition. For example, the form, brightness and position of clouds strongly affect the directional properties of the radiance.

The distribution of radiance over the sky is not measured routinely. Nevertheless, a number of authors [22–24] have developed instruments to measure it and have presented results for different sky conditions. Some general patterns may be discerned from these.

With clear skies, the maximum diffuse radiance comes from the parts of the sky close to the sun and to the horizon. The minimum radiances come from a region at an angle of 90° to the solar zenith (Figure 22.14). The diffuse radiation coming from the region close to the sun is called *circumsolar radiation* and is mainly due to the dispersion by aerosols. The angular extent of the sun's aureole depends mainly on the turbidity of the atmosphere and on the zenith angle of the sun. The increase in diffuse radiance near the horizon is due to the albedo radiation of the Earth and is called *horizon brightening*.

The radiance distributions associated with overcast skies are very well described by Kondratyev: 'for dense non-transparent cloudiness, the azimuthal dependence of diffuse radiation intensity is very weak. There is a slight monotonic increase of the radiance from the horizon upward towards the zenith' [25].

Some models can be derived from these general ideas. The simplest model makes use of the assumption that the sky radiance is isotropic, that is, every point of the celestial sphere emits light with equal radiance, $L(\theta_Z, \psi) = \text{constant}$. The solution of Equations (22.30 and 22.31) leads to

$$D(\beta, \alpha) = D(0) \frac{1 + \cos \beta}{2} \tag{22.32}$$

Because of its simplicity, this model has achieved great popularity, despite the fact that it systematically underestimates diffuse irradiance on surfaces tilted to the equator.

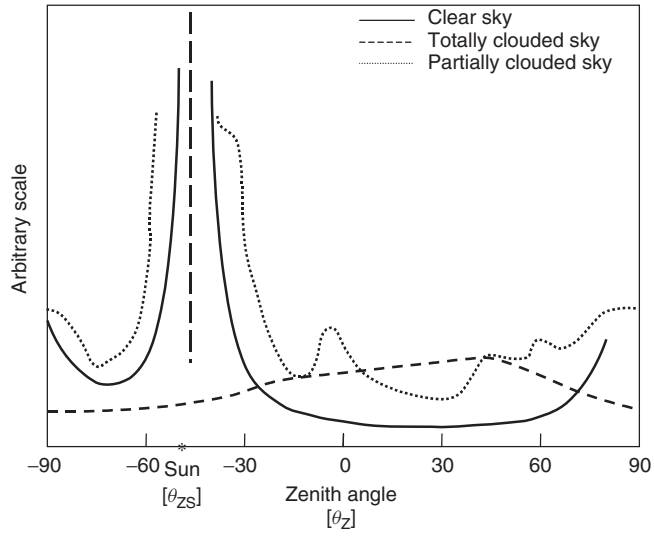


Figure 22.14 Typical angular distribution of the sky irradiance. The values are taken along the length of the meridian containing the sun, $\psi = \psi_S$

The opposite approach assumes that all the diffuse radiation is circumsolar, that is, from the sun. This is really a case of treating diffuse radiation as though it were direct, and leads to

$$D(\beta, \alpha) = \frac{D(0)}{\cos\theta_{ZS}} \max(0, \cos\theta_S) \tag{22.33}$$

This model also has the advantage of being very simple to use, but in general it overestimates diffuse irradiances.

In general, better results are obtained with so called anisotropic models. Hay and Davies [26] proposed considering the diffuse radiation as composed of a circumsolar component coming directly from the direction of the sun, and an isotropic component coming from the entire celestial hemisphere. Both components are weighted according to the so-called anisotropy index k_1, defined as

$$k_1 = \frac{B(0)}{B_0(0)} = \frac{B}{B_0 \varepsilon_0} \tag{22.34}$$

The solution of Equation (22.31) is now

$$D(\beta, \alpha) = D^I(\beta, \alpha) + D^C(\beta, \alpha) \tag{22.35}$$

where

$$D^I(\beta, \alpha) = D(0)(1 - \kappa_1)\frac{1 + \cos\beta}{2} \tag{22.36}$$

and

$$D^C(\beta, \alpha) = \frac{D(0)k_1}{\cos\theta_{ZS}} \max(o, \cos\theta_S) \tag{22.37}$$

respectively, define the contribution of the isotropic and of the circumsolar components.

Note than k_1 is just the ratio between a pyrheliometer's reading and the solar constant, once corrected by the eccentricity due to the ecliptic orbit of the Earth around the sun. In this way, k_1 can be understood as a measure of the instantaneous atmospheric transmittance for direct irradiance. When the sky is completely clouded over, $k_1 = 0$ and this equation becomes the same as that for the simple isotropic model. This anisotropic model is an excellent compromise between simplicity and precision. It has been well validated against measurements performed at different worldwide locations, and has been extensively used, for example, for the elaboration of the European Atlas of the Solar Radiation [8].

Also, very commonly employed, in particular with digital machines, is the model that has been put forward by Perez [27, 28]. It divides the sky into three zones acting as diffuse radiation sources: a circumsolar region, a horizontal band and the rest of the celestial hemisphere. The relative contribution of each component is modulated by means of empirical factors determined from the study of data from 18 measurement stations at 15 sites in North America and Europe. The Perez model used to perform slightly better than others [29], because the larger number of modulating factors allows for the consideration of a larger number of different sky conditions.

22.5.3.3 Albedo irradiance

The reflectivity of most types of ground is rather low. Consequently, the contribution of the albedo irradiance falling on a receiver is generally small. (An exception occurs in the case of snow.) There is therefore no point in developing very sophisticated models for albedo. It is usual to assume that the ground is horizontal and of infinite extent and that it reflects isotropically. On this basis, the albedo irradiance on an inclined surface is given by

$$R(\beta, \alpha) = \rho G(0) \frac{1 - \cos \beta}{2} \tag{22.38}$$

where ρ is the reflectivity of the ground and depends on the composition of the ground. When the value of ρ is unknown, it is common to take $\rho = 0.2$.

It is now opportune to go forward with the example of the previous section, by calculating the irradiance components over a surface tilted to the latitude, for 15 April in Portoalegre, Brazil. Using Equations (22.34–22.37) to deal with diffuse radiation and $\rho = 0.2$, the results, expressed in W/m^2, are as follows:

ω	k_1	$D^{\mathrm{I}}(\phi)$	$D^{\mathrm{C}}(\phi)$	$B(\phi)$	$R(\phi)$	$G(\phi)$
ω_s	0	0	0	0	0	0
$\pm 60^\circ$	0.2205	73.40	31.80	147.56	2.73	255.49
$\pm 30^\circ$	0.3082	124.14	76.94	357.14	6.26	564.48
0°	0.3403	138.97	98.09	455.31	7.80	700.18

It becomes clear that the albedo can be generally neglected in PV calculations.

22.5.3.4 Daily irradiation

The most accurate way of calculating $G_{\mathrm{dm}}(\beta, \alpha)$ from $G_{\mathrm{dm}}(0)$ is, first, to calculate the hourly horizontal irradiation components $G_{\mathrm{hm}}(0)$, $D_{\mathrm{hm}}(0)$ and $B_{\mathrm{hm}}(0)$; second, to transpose them to the inclined surface $G_{\mathrm{hm}}(\beta, \alpha)$, $D_{\mathrm{hm}}(\beta, \alpha)$ and $B_{\mathrm{hm}}(\beta, \alpha)$; and, finally, to integrate during the day.

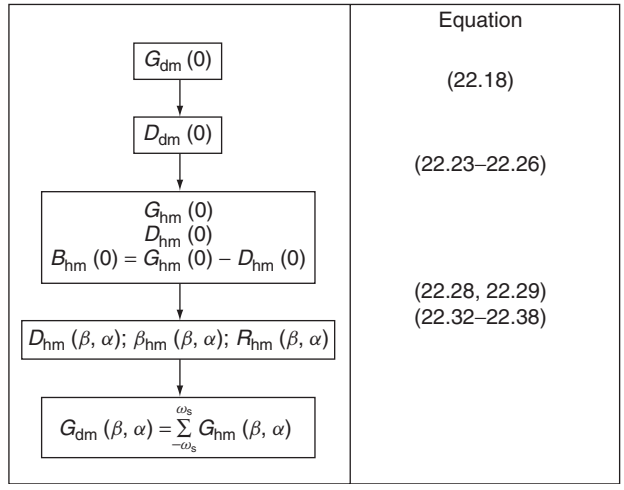


Figure 22.15 Diagram explaining the calculation of the daily irradiation on an inclined surface $G_{dm}(\beta, \alpha)$ from the corresponding horizontal value $G_{dm}(0)$

Such a procedure, summarised in Figure 22.15, allows us to account for the anisotropic properties of diffuse radiation, and leads to good results, whatever the orientation of the inclined surface. However, it is laborious to apply and a computer must be used. It is interesting to mention that, for the case of surfaces tilted to the equator ($\alpha = 0$), frequently encountered in photovoltaic applications, if the diffuse radiation is taken to be isotropic, the following expression may be applied

$$G_d(\beta, 0) = B_d(0) \times RB + D_d(0)\frac{1 + \cos\beta}{2} + \rho G_d(0)\frac{1 - \cos\beta}{2} \tag{22.39}$$

where the factor *RB* represents the ratio between the daily direct irradiations on an inclined surface and on an horizontal surface, and may be approximated by setting it equal to the corresponding ratio between daily extraterrestrial irradiations on similar surfaces. Hence, *RB* is given as follows:

$$RB = \frac{\omega_{SS}\frac{\pi}{180}[\text{sign}(\phi)]\sin\delta\sin(|\phi| - \beta) + \cos\delta\cos(|\phi| - \beta)\sin\omega_{SS}}{\omega_S\frac{\pi}{180}\sin\delta\sin\phi + \cos\delta\cos\phi\sin\omega_S} \tag{22.40}$$

where ω_{SS} is the sunrise angle on the inclined surface, which is given by

$$\omega_{SS} = \max[\omega_S, -\arccos(-[\text{sign}(\phi)]\tan\delta\tan(abs(\phi) - \beta))] \tag{22.41}$$

It is interesting to observe that for the equinox days, $\delta = 0 \Rightarrow \omega_S = \omega_{SS}$ and Equation (22.40) becomes $RB = \cos[abs(\phi) - \beta]/\cos\phi$.

Example. Estimate the average daily irradiation in January at Changchun, China ($\phi = 43.8^\circ$) over a fixed surface facing south and tilted at an angle $\beta = 50^\circ$ with respect to the horizontal, knowing that the mean value of the global horizontal irradiation is $G_{dm}(0) = 1861\,\text{W h/m}^2$ and the ground reflection $\rho = 0.2$. The solution is as follows:

January $\Rightarrow d_n = 17$; $\delta = -20.92^\circ$

$\phi = 43.8^\circ \Rightarrow \omega_S = -68.50$ and $B_{0d}(0) = 3586\,\text{Wh/m}^2 \Rightarrow K_{Tm} = 0.519$

$\Rightarrow F_{Dm} = 0.414$

$D_{dm}(0) = 770\,\text{Wh/m}^2$; $B_{dm}(0) = 1091\,\text{Wh/m}^2$

$\arccos(-\tan\delta\tan(\phi-\beta)) = -92.38^\circ \Rightarrow \omega_{SS} = -68.5^\circ \Rightarrow RB = 2.741$

$D_{dm}(50) = 633\,\text{Wh/m}^2$; $B_{dm}(50) = 2990\,\text{Wh/m}^2$; $R_{dm}(50) = 66\,\text{Wh/m}^2$

$G_{dm}(50) = 3689\,\text{Wh/m}^2$

It is worth mentioning that a more detailed calculation, following the procedure outlined in Figure 22.15, would lead to $G_{dm}(50) = 4111\,\text{W h/m}^2$. This means that the error associated with Equation (22.40) is about 10%. This difference is mainly due to the different consideration of the diffuse radiance distribution.

22.6 DIURNAL VARIATIONS OF THE AMBIENT TEMPERATURE

The behaviour of the photovoltaic modules depends on the ambient temperature, although much weaker than its dependence in irradiance. Just as it is for solar radiation, sometimes it is necessary to determine how this parameter varies throughout the day. The data available as a starting point for this calculation are, in general, the maximum and minimum temperature of the day T_{aM} and T_{am}, respectively.

A model that is simple but, nevertheless gives a good fit to the experimental values is obtained from the fact that the temperature evolves in a similar manner to the global radiation, but with a delay of about 2 h. This fact allows the following three principles to be deduced:

- T_{am} occurs at sunrise ($\omega = \omega_S$).
- T_{aM} occurs two hours after midday ($\omega = 30^\circ$).
- Between these two times, the ambient temperature develops according to two semi-cycles of a cosine function: one from dawn to midday, and the other between midday and sunrise of the following day.

The following set of equations based on those above, permits the ambient temperature throughout a day j to be calculated as follows.

For $-180 < \omega \leq \omega_S$

$$T_a(j,\omega) = T_{aM}(j-1) - \frac{T_{aM}(j-1) - T_{am}(j)}{2}[1 + \cos(a_T\omega + b_T)] \tag{22.42}$$

with $a_T = \dfrac{-180}{\omega_S + 330}$ and $b_T = -a_T\omega_S$

For $\omega_S < \omega \leq 30$

$$T_a(j,\omega) = T_{am}(j) + \frac{T_{aM}(j) - T_{am}(j)}{2}[1 + \cos(a_T\omega + b_T)] \tag{22.43}$$

with $a_T = \dfrac{180}{\omega_S - 30}$ and $b_T = -30a_T$

For $30 < \omega \leq 180$

$$T_a = T_{aM}(j) - \frac{T_{aM}(j) - T_{am}(j+1)}{2}[1 + \cos(a_T\omega + b_T)] \tag{22.44}$$

with $a_T = \dfrac{180}{\omega_S + 330}$ and $b_T = -(30a_T + 180)$

To apply these equations, it is necessary to know the maximum temperature on the previous day $T_{aM}(j-1)$ and the minimum temperature of the following day $T_{am}(j+1)$. If these data are unavailable, then it can be assumed, without introducing too much error, that they equal those for the day in question.

22.7 EFFECTS OF THE ANGLE OF INCIDENCE AND OF DIRT

The reflectance and transmittance of optical materials depends on the angle of incidence. The glass covers of solar collectors are no exception, and therefore the optical input of photovoltaic modules is affected by their orientation with respect to the sun, due to the angular variation of the glass reflection. Theoretical models, based on the well-known Fresnel formulae, have been developed for clean surfaces. The most popular formulation is from ASHRAE [30]. For a given incidence angle, θ_S, it can be described by the simple expression

$$FT_B(\theta_S) = 1 - b_0\left(\frac{1}{\cos\theta_S} - 1\right) \tag{22.45}$$

where $FT_B(\theta_S)$ is the relative transmittance, normalised by the total transmittance for normal incidence, and b_0 is an adjustable parameter that can be empirically determined for each type of photovoltaic module. If this value is unknown, a general value $b_0 = 0.07$ may be used. The effect of the angle of incidence on the successfully collected solar radiation can be calculated by applying Equation (22.45) to the direct and circumsolar irradiances, and by considering an approximate value $FT = 0.9$ for the isotropic diffuse and reflected radiation terms. Figure 22.16 shows a plot of $FT_B(\theta_S)$ versus θ_S. It presents a pronounced knee close to 60°. In practical terms, that means the effects of the angle of incidence are negligible for all the θ_S well below this value. For example, $FT_B(40°) = 0.98$.

The ASHRAE model is simple to use, but has noticeable disadvantages. It cannot be used for $\theta_S > 80°$, and, still worse, it cannot take into consideration the effects of dust. Dust is always present

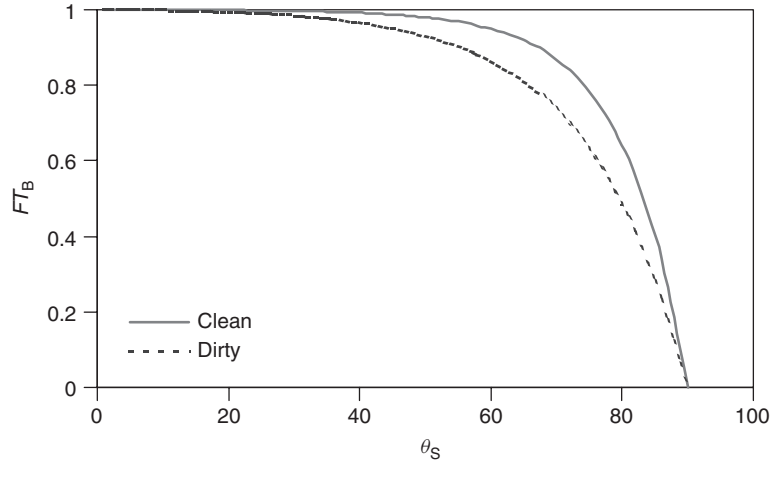


Figure 22.16 The relative transmittance FT_B is plotted against the angle of incidence θ_S, for a clean surface and also for a dust-covered surface

in real situations, and not only reduces the transmittance at normal incidence, but also influences the shape of $FT_B(\theta_S)$. Figure 22.16 shows that the relative transmittance decreases because of dust at angles from about 40 to 80°. Real $FT_B(\theta_S)$ are best described [31] by

$$FT_B(\theta_S) = 1 - \frac{\exp\left(-\frac{\cos\theta_S}{a_r}\right) - \exp\left(-\frac{1}{a_r}\right)}{1 - \exp\left(-\frac{1}{a_r}\right)} \tag{22.46}$$

where a_r is an adjustable parameter mainly associated with the degree of dirtiness, as shown in Table 22.3. Note that the degree of dirtiness is characterised by the corresponding relative normal transmittance, $T_{dirt}(0)/T_{clean}(0)$. Equation (22.46) applies for direct and circumsolar radiation components. Consistent equations for the angular losses for isotropic diffuse and albedo radiation components are given in reference [31]

It must be noted that $FT_B(0) = 1$. That means, this function does not include the effect of dirt on the relative normal transmittance, but only the angular losses relative to normal incidence. In other words, the 'effective' direct irradiance reaching the solar cells of a PV module, should be computed as

$$B_{eff}(\beta, \alpha) = B(\beta, \alpha) \times \frac{T_{dirt}(0)}{T_{clean}(0)} \times FT_B(\theta_S) \tag{22.47}$$

Following the example of 15 April in Portoalegre, Brazil, we can now calculate the effective irradiances over a surface tilted to the latitude, neglecting the albedo, supposing a medium dirtiness degree and by applying $FT_B(\theta_S)$ not only to the direct radiation, but also to the circumsolar component of the diffuse radiation.

ω°	FT_B (θ_S)	$B_{eff}(\phi)$, [W m^{-2}]	$D_{eff}(\phi)$ [W m^{-2}]	$G_{eff}(\phi)$ [W m^{-2}]	ΔG_{eff} [%]
ω_S	0	0	0	0	0
±60	0.913	80.84	126.39	207.23	−11.3
±30	0.991	249.63	249.74	499.37	−6.8
0	0.999	332.13	296.37	628.50	−6.1

The last column of this table describes the losses due to both dirt and angular effects. Taking into consideration that dirt reduces normal transmittance by a factor of 3% ($T_{dirt}(0)/T_{clean}(0) = 0.97$), it can be noted that pure angular losses dominate for $|\omega| > 30^\circ$.

Table 22.3 Recommended parameters for angular loss calculation

Dirtiness degree	$T_{dirt}(0)/T_{clean}(0)$	a_r
Clean	1	0.17
Low	0.98	0.20
Medium	0.97	0.21
High	0.92	0.27

Finally, it should be stressed that angular-dependent reflection is often neglected in PV simulations. However, they become significant in many practical situations, for example, where vertical (façade-integrated PV generators) or horizontal (N–S horizontal trackers) surfaces are concerned. Furthermore, they help to explain the observed low irradiance effects in PV module performance. This is because low irradiance just happens when the incidence angle is large or when solar radiation is mainly diffuse. In both cases, angular losses are particularly important. As a matter of fact, the failure to consider angular losses has been signalled as the main cause of error in some energy models [32].

22.8 SOME CALCULATION TOOLS

22.8.1 Generation of Daily Radiation Sequences

Long series (many years) of daily irradiation data are sometimes required for particular purposes, for example, when studying the long-term reliability of stand-alone photovoltaic systems. However, long series of historical data are scarce and hard to obtain. This leads to the need for methods that are able to generate a series starting from widely available information, such as the 12 long-term average monthly mean values of the daily irradiation $G_{dm}(0)$. The idea is that the generated series must preserve some statistical properties believed to be universal, as they are also found in historical data, when available. In particular, the persistence of solar radiation, that is, the dependence of today's irradiation on the irradiation of the preceding days, is adequately described by a first-order auto-regressive process [33]. Moreover, the probability function of the daily clearness index for any given period has a form associated with only its average value for the period. Several methods for the generation of daily irradiation sequences are available in the literature [34]. The method proposed by Aguiar [35] is the most widely used today.

22.8.2 The Reference Year

As already mentioned, the most widely available information related to the solar radiation resource at a given location is the set of 12 monthly mean values of global horizontal daily irradiation, $G_{dm}(0)$. The methods presented above allow estimation of all the radiation components incident on any surface of arbitrary orientation and at any moment of the average year, and even at any moment of a long sequence of years. This can be applied to all the problems related to the design of photovoltaic systems: sizing, prediction of energy yields, impact of shadowing, optimisation of tilt angles and so on.

Nevertheless, the solar radiation is still the object of systematic recording, and more and more irradiance and irradiation data are being accumulated and put at public disposal. Such data, whether in the form of crude recorded data or in the form of elaborate mathematical tools, attempt to properly represent the climate of the concerned location. The most widely used is the so-called reference year, also called the typical meteorological year *TMY*, or the standard year [36]. The *TMY* for a location is a hypothetical year in which months are real months, but are chosen from different years from the whole period for which data are available. In practice [37], the months are chosen such that the monthly mean of the daily global irradiation on the horizontal represents an average value for all values contained in the database. For example, January of 1986 was chosen for the *TMY* of Madrid, because it had a value of $G_{dm}(0) = 1.98\,\mathrm{kW\,h/m^2}$, the closest to the average value of $G_{dm}(0) = 1.99\,\mathrm{kW\,h/m^2}$ for all the months of January on record [14].

The most widely used *TMY* for photovoltaic applications is set in a one-hour time scale. Hence, it contains 4380 values[1] of global horizontal irradiation. Ambient temperature values are also specified for each hour. This huge number of initial data can lead to the impression that the corresponding results should be much more accurate than those obtained when simply using the $12\,G_{dm}(0)$ values as input. However, this impression is largely wrong. On the one hand, because the representativeness of any data – it should be again remembered – is limited by the random nature of the solar radiation, small differences in the results are scarcely meaningful. On the other hand, because the results obtained from the $12\,G_{dm}(0)$ values and from the *TMY* are very similar, provided the initial data are coherent (i.e. the monthly means in the *TMY* coincides with the $12\,G_{dm}(0)$ values) and that the selected correlations and diffuse radiation models to transpose from horizontal to inclined surfaces are the same. The physical reasons for this lie in the, already mentioned, quasi-linear power–irradiance relationship in most PV devices, and in the fact, initially shown by Liu and Jordan [15], that the solar climate of a particular location can be well characterised by only the monthly mean daily clearness index. As already mentioned, they have demonstrated that, irrespective of latitude, the fractional time during which daily global radiation is equal to or less than a certain value depends only on this parameter. Surely, to go into this question in-depth would increase the reader's boredom which is probably already large enough; hence, we will restrict ourselves to describing a representative case from our own experience:

In 1992, the Solar Energy Institute in the University of Madrid, IES-UPM, was involved in the design of the 1 MW PV plant in Toledo, Spain. It was the biggest European PV project at that time, so very careful studies were required at the initial project stage. Fortunately, a large historical database, containing 20 years of hourly irradiation data, was available from a nearby meteorological station, and was directly used to calculate the expected energy yields. Both static and sun-tracked photovoltaic arrays were analysed, while taking into account detailed features such as shadowing from adjacent rows, back-tracking features and so on. Moreover, the same calculation was also performed using as input the *TMY*, previously derived from the historical radiation sequence, and also using as the only input the $12\,G_{dm}(0)$ values and computing for just the mean day of each month. The results from the three calculation procedures never differed by more than 2%! As a matter of fact, the results were much more sensitive to the considerations of the solar angle of incidence effects [38] described below.

A clever friend, not involved in this project, but aware of this anecdote, posed the questions: Then why go into such exhaustive detail when they give similar results? Why not just be simple? The proper answers are best found in human psychology. Many people simply desire not to believe in some ideas. Hence, the ones daring to declare them are automatically impelled to provide strong arguments in favour of such ideas. To a large extent, this is usually the case when defending the argument that modern complex software tools do not necessarily yield better results than simple (but judicious) traditional methods. That was the position of the IES-UPM in 1992, and today for the author of this chapter the position remains the same.

22.8.3 Shadows and Trajectory Maps

Surroundings of photovoltaic modules can include trees, mountains, chimneys, walls, other PV modules and so on. Because of that, photovoltaic modules cannot always be positioned entirely free of shadows. This reduces their potential energy yield, and must be taken into account when

[1] There are 8760 hours per year, and the sun shines exactly half of the year in any location, hence there are 4380 hours of sunshine per year.

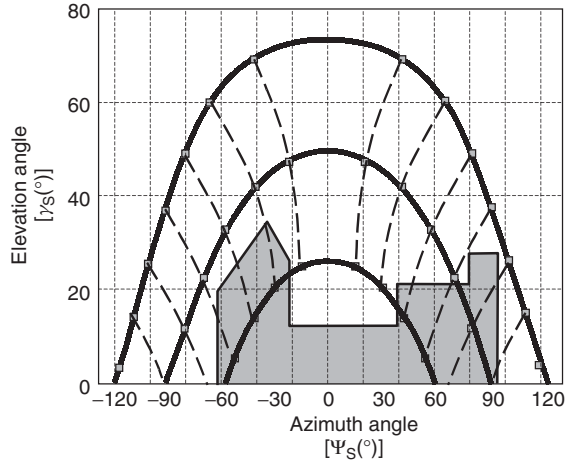


Figure 22.17 Sun's trajectory map corresponding to a latitude $\phi = 40.5^{\circ}$, with a skyscraper superimposed on the map. For example, at the winter solstice, shadows occur from sunrise to about 10:30 (solar time) and from 14:30 to sunset

designing photovoltaic systems. Equations (22.3–22.5) allow the plotting of the trajectories of the sun, in terms of elevation versus azimuth angles, as already explained in Figure 22.6.

These types of diagrams are called sun trajectory maps. They are a very useful tool for determining the duration and effect of shadowing cast by any obstacle. A correctly placed theodolite (the standard instrument for measurements of azimuth and elevation angles) can measure the azimuth and elevation angles of the most relevant points (corners, peaks, etc.) of any kind of obstacle. The *local horizon* can then be superimposed on the trajectory map, as Figure 22.17 shows. The effect of the shadow is calculated with the assumption that the direct and circumsolar radiations are zero when the sun is below the local horizon. Unless the shadows are very large, the effect of the local horizon on the diffuse radiation (other than the circumsolar component) can be neglected. Several tools have been proposed [39, 40] to simplify the employment of this calculation.

Mutual shading between adjacent support structures and trackers is particularly relevant when big PV plants are concerned. It can determine the shape and spacing of the arrays. Corresponding equations are developed in reference [41]

22.9 IRRADIATION ON MOST WIDELY STUDIED SURFACES

This section analyses some important features of the radiation available on commonly studied surfaces. As already mentioned, the methods presented before conform to a complete package, allowing the calculation of the irradiation incident over any arbitrary surface over any period of time, using the horizontal data as input. This can no doubt be a tedious task, so specific commercial software packages have been developed [42, 43]. However, for many practical engineering problems, more direct and simple tools can be developed. In particular, it is possible to develop analytical expressions that can be simply solved by hand calculations. In order to apply the discussion in the previous sections, let us analyse the particular case of the yearly mean daily irradiation collected at four different places on a fixed surface, tilted towards the equator ($\alpha = 0$) and inclined

at an angle β to the horizontal, $G_{dy}(\beta)$. Figure 22.18 plots, for each place, such a value in relation to its maximum and versus the inclination angle relative to the absolute value of the latitude, that is, $G_{dy}(\beta - |\phi|)/G_{dy}(\beta_{opt})$, where β_{opt} is the inclination angle associated with the maximum value of $G_{dy}(\beta)$. The calculation followed the procedure described in Figure 22.16. Solar radiation data has been obtained from reference [6]. Several aspects need to be outlined.

On the one hand, a great similarity between all the curves is noticeable. Despite large differences in latitude and clearness index of the selected locations, the shape of the curve and also the inclination angle maximising the collection ($\beta_{opt} - |\phi|$) of radiation are very similar for the four selected places; i.e. it is relatively close to the latitude. It is important to mention that the extension of this exercise to many other places all around the world verifies that this great similarity is nearly universal. In fact, we have performed a specific exercise covering 32 different places distributed from $\phi = 80^\circ$ to $\phi = -78.2^\circ$ (see list in Table 22.4). We limit Figure 22.18 to only four curves for presentation purposes. A physical explanation of this similarity can be argued, observing that, irrespective of the latitude, all the surfaces tilted towards the equator and inclined at an angle equal to the absolute value of the latitude are parallel all over the Earth, and also parallel to the Earth's rotation axis. Therefore, in the absence of the atmosphere, on the equinox days, they receive identical solar irradiation $B_{0d}(|\phi|)|_{\delta=0}$. And the same is true for surfaces equally tilted with respect to the latitude angle $B_{0d}(\beta - |\phi|)|_{\delta=0}$. For other than equinox days, this is not exactly true, because sunrise time (and, therefore, the length of daytime and, in turn, the daily extraterrestrial irradiation) depends on latitude, as described by Equation (22.7). However, the yearly ratio between extraterrestrial irradiation collected on both inclinations $B_{0dy}(\beta - |\phi|)/B_{0dy}(|\phi|)$ tends to remain constant. Now, when accounting for the Earth's atmosphere, such site independence is not necessarily maintained, because of the different climatic conditions, that is, the different annual clearness index K_{Ty}, from one site to another. Obviously, the collection of diffuse radiation is less sensitive to inclination angle variations than the collection of direct radiation. Because of this, a slightly less sharply peaked curve should be expected with the lower the K_{Ty} value (the higher the F_{dy}). Figure 22.18 reveals that this tendency is, in fact, very weak, so that the site independence is maintained and the function $G_{dy}(\beta - |\phi|)/G_{dy}(\beta_{opt})$ can be properly considered as being a universal invariant.

On the other hand, the smooth form of the curves of Figure 22.18 suggests that it may be possible to describe them analytically, thus avoiding the need to use a computer each time a particular value is required. It is worth noting that the larger the latitude, the larger the difference

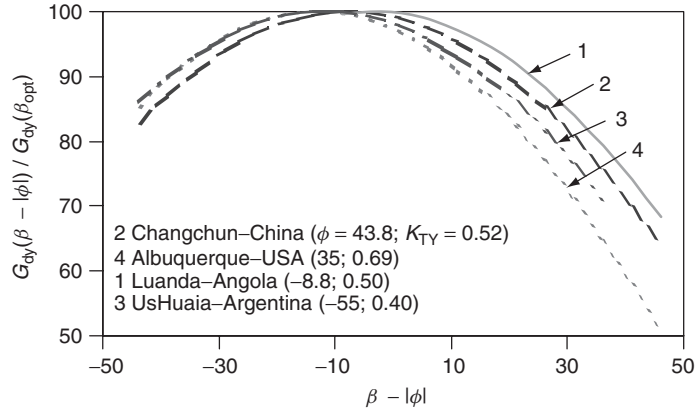


Figure 22.18 Yearly energy collection versus inclination angle for surfaces tilted towards the equator. The percentage of relative collection, with respect to the maximum, $G_{dy}(\beta - |\phi|)/G_{dy}(\beta_{opt})$ is plotted against the difference between the tilt angle and the latitude $\beta - |\phi|$

Table 22.4 Annual radiation availability on different surfaces, for 32 different places all around the world. Various tracking options (1-axis, 2-axis, etc.) described in Section 22.9.2 The global daily radiation on a horizontal surface Gdm(0) is obtained from the product of columns 3 and 4 [$B_{0dy}(0)$ and K_{Ty}]

Location	Lat. ϕ°	$B_{0dy}(0)$ [W h/m^2]	Clearness index K_{Ty}	Ratios to global horizontal yearly irradiation					
				2-axis	1 Az-axis	1 Ho-axis	1 Po-axis	Fixed β_{opt}	2-axis Conc.
Ice Island, Arctic	80	4180	0.591	2.92	2.86	2.19	2.70	1.82	2.41
St Petersburg, Russia	59.9	5616	0.460	1.85	1.80	1.55	1.75	1.27	1.33
Hamburg, Germany	53.6	6354	0.417	1.63	1.58	1.41	1.57	1.20	1.08
Freiburg, Germany	48	6998	0.433	1.54	1.49	1.37	1.49	1.14	1.03
Nantes, France	47.2	7088	0.473	1.63	1.56	1.42	1.58	1.19	1.13
Olympia, USA	46.6	7154	0.442	1.52	1.47	1.37	1.47	1.13	1.02
Changchun, China	43.8	7459	0.515	1.72	1.62	1.46	1.66	1.24	1.24
Sapporo, Japan	43	7543	0.425	1.55	1.47	1.35	1.49	1.18	1.01
Chicago, USA	41.5	7700	0.499	1.57	1.49	1.40	1.53	1.14	1.11
Madrid, Spain	40.4	7812	0.549	1.63	1.52	1.45	1.58	1.16	1.20
Washington, USA	38.6	7991	0.478	1.52	1.43	1.36	1.47	1.12	1.04
St Louis, USA	38.4	8011	0.523	1.57	1.46	1.41	1.52	1.13	1.12
Seoul, Korea	37.5	8098	0.433	1.51	1.42	1.34	1.46	1.15	1.00
Albuquerque, USA	35	8331	0.692	1.71	1.54	1.52	1.65	1.16	1.38
Djelfa, Algeria	34.6	8364	0.589	1.61	1.47	1.44	1.56	1.14	1.21
Shanghai, China	31.2	.8661	0.490	1.46	1.34	1.34	1.42	1.09	1.00
Cairo, Egypt	30.6	8710	0.641	1.60	1.42	1.45	1.54	1.11	1.24
Delhi, India	28.6	8869	0.633	1.62	1.41	1.45	1.56	1.13	1.26
Karachi, Pakistan	24.8	9144	0.603	1.55	1.34	1.41	1.49	1.10	1.17
Morelia, Mexico	19.7	9459	0.417	1.30	1.17	1.24	1.27	1.04	0.80
Daka, -Senegal	14.7	9702	0.601	1.44	1.19	1.37	1.40	1.03	1.08
Bangkok, Thailand	13.7	9743	0.491	1.36	1.15	1.29	1.32	1.03	0.91
Claveria, Philippines	8.6	9909	0.514	1.34	1.09	1.30	1.31	1.00	0.92
Colombo, Sri Lanka	6.9	9949	0.530	1.35	1.08	1.31	1.32	1.01	0.94
Medellin, Colombia	6.2	9963	0.470	1.29	1.06	1.26	1.26	1.00	0.84
Luanda, Angola	−8.8	9936	0.496	1.32	1.09	1.29	1.27	1.00	0.89
El Alto, Bolivia	−16.4	9685	0.577	1.45	1.21	1.36	1.28	1.05	1.07
Sao Paulo, Brazil	−23.5	9313	0.425	1.34	1.22	1.26	1.12	1.06	0.84
Porto Alegre, Brazil	−30	8863	0.501	1.46	1.33	1.34	1.14	1.08	1.01
Bariloche, Argentina	−41.1	7877	0.560	1.63	1.53	1.46	1.11	1.15	1.21
Ushuaia, Argentina	−55	6360	0.396	1.73	1.67	1.42	0.91	1.29	1.14
Little America, Antartic	−78.2	4438	0.560	2.63	2.58	1.98	1.17	1.69	2.04

between summer daytime and winter daytime and, in turn, the larger the difference between the summer and the winter irradiation. Therefore, it can be anticipated that as latitude increases, the optimal inclination angle should progressively give priority to the collection of summer over the collection of winter. This can be observed in Figure 22.19, where the values of $\beta_{opt} - |\phi|$ have been plotted against the latitude, for the 30 different locations mentioned above.

It is useful to fit a linear equation to these values

$$\beta_{opt} - |\phi| = 3.7 - 0.31|\phi| \quad \text{or} \quad \beta_{opt} = 3.7 + 0.69|\phi| \tag{22.48}$$

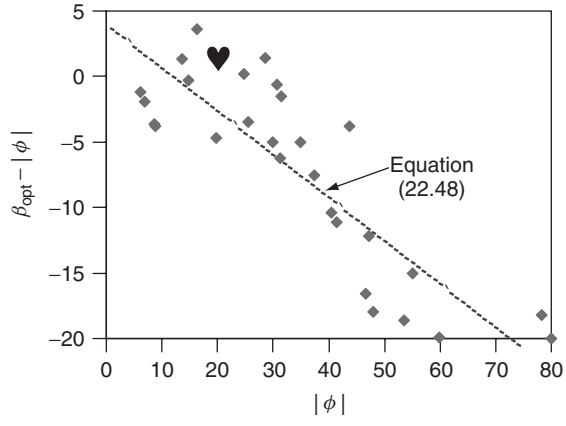


Figure 22.19 Optimal inclination angle versus latitude. The difference between $\beta_{opt} - |\phi|$ is plotted against the latitude $|\phi|$. The cluster of points corresponds to the different places listed in Table 22.4

where β and ϕ are given in degrees. It is worth mentioning that the observed dispersion of the cluster of points around Equation (22.48) (also depicted in Figure 22.19) is, in fact, of negligible importance, due to the very low sensibility of the energy collection to deviations from the optimal inclination angle. For example, the point corresponding to Delhi ($\phi = 28.6^\circ$) – marked as ♥ in the figure – indicates a value of $\beta_{opt} = 29.6^\circ$, while Equation (22.48) leads to $\beta_{opt} = 23.4^\circ$. The 6.2° of difference can appear relatively large ($\approx$20%), but a detailed simulation exercise would disclose that $G_{dy}(23.4^\circ)/G_{dy}(29.6^\circ)$ is 99%, that is, such a difference is irrelevant when translated into energy content. A similar reasoning also justifies the validity of the usual assumption $\beta_{opt} = \phi$. Now, a second-order polynomial describes very well the curves of Figure 22.18.

$$\frac{G_{dy}(\beta)}{G_{dy}(\beta_{opt})} = 1 + p_1(\beta - \beta_{opt}) + p_2(\beta - \beta_{opt})^2 \tag{22.49}$$

where p_1 and p_2 are the adjusted parameters for which values are $4.46 \times 10^{-4}(^\circ)^{-1}$ and $-1.19 \times 10^{-4}(^\circ)^{-2}$, respectively. The corresponding value of the so-called correlation coefficient R^2 is greater than 0.98, indicating that Equations (22.48) and (22.49) adjust very well to the simulated values. In this way, the results of a rather complex calculation, involving a lot of tedious steps, can be described by means of two simple mathematical expressions. Apart from simplicity and elegance, all the information has been condensed into just four numbers and an equation that has the advantage of having a continuous slope, which can be useful in many calculations. It is worth mentioning that these equations can also be used to calculate the value of $G_{dy}(\beta_{opt})$ from the data corresponding to the horizontal surface, which is the most usually available information.

Example: Estimation of the optimal tilt angle and the corresponding yearly irradiation in Sapporo, Japan knowing the latitude $\phi = 43^\circ$ and the annual average of the daily global horizontal irradiation $G_{dy}(0) = 3220\,\text{Wh/m}^2$. The solution is

$$\begin{aligned}
&\text{Equation (22.48)}, \phi = 43^\circ \Rightarrow \beta_{opt} = 33.37^\circ \\
&\text{Equation (22.49)}, \beta = 0^\circ \Rightarrow G_{dy}(0)/G_{dy}(\beta_{opt}) = 0.8526 \\
&\qquad\qquad \Rightarrow G_{dy}(\beta_{opt}) = 1.1729 \times G_{dy}(0) = 3776\,\text{Wh/m}^2.
\end{aligned}$$

The total yearly irradiation is $365 \times G_{dy}(0) = 1379$ kW h/m^2.

It is worth mentioning that a more careful calculation, using the 12 values of the monthly mean irradiation and following the procedure described in Figure 22.15, would lead to $G_{dy}(\beta_{opt}) = 3790\,W\,h/m^2$. This means, that the error associated with Equation (22.49) is below 0.4%.

Attempting to help in the following discussion, Table 22.4 presents the results of a detailed simulation exercise devoted to the calculation of the annual radiation availability on horizontal surfaces, optimally tilted fixed surfaces and several types of tracking surfaces. The exercise has been extended to 32 different places distributed around the world. The hope is that the readers could find here a relatively similar location, both in latitude and clearness index, to the location of their interest. The yearly means of the daily global horizontal irradiation, $G_{dy}(0)$, can be obtained by multiplying the corresponding values of the extraterrestrial irradiation and the clearness index (column 3× column 4). Then, these horizontal $G_{dy}(0)$ values are used as reference for the irradiation availability in all the other considered surfaces. In particular, column 8 of this table gives the ratio between the global irradiation on a fixed and optimally tilted surface to the global horizontal irradiation, $G_{dy}(\beta_{opt})/G_{dy}(0)$. Hence, the irradiation on the optimally tilted surface is given by the product (column 3× column 4× column 8).

Example: Estimation of the yearly irradiation corresponding to the horizontal and to the optimally tilted surfaces in Sapporo,-Japan. The solution is:

$$G_y(0) = 365 \times [\text{column } 3 \times \text{column } 4] = 1170\,kW\,h/m^2$$
$$G_y(\beta_{opt}) = G_y(0) \times \text{column } 8 = 1381\,kW\,h/m^2$$

Which differs in only 0.15% of the value calculated in the previous example.

22.9.1 Fixed Surfaces

The integration of PV generators in buildings, presently in vogue in many industrialised countries, led to the use of a large range of different orientations and tilt angles. PV module orientations from east to west, and tilt angles from horizontal to vertical are found in practice. Then, it is worth extending the previous exercise to surfaces other than those tilted towards the equator, that is, to $\alpha \neq 0$, and also to dirty surfaces. E. Caamaño [44] has proposed the following solution:

$$\frac{G_{effdy}(\beta,\alpha)}{G_{dy}(\beta_{opt})} = g_1(\beta - \beta_{opt})^2 + g_2(\beta - \beta_{opt}) + g_3 \tag{22.50}$$

where

$$g_i = g_{i1} \cdot |\alpha|^2 + g_{i2} \cdot |\alpha| + g_{i3};\ i = 1, 2, 3 \tag{22.51}$$

being the values of the coefficients as show in Table 22.5, for medium dirty surfaces. The subscript 'eff' in the first term of Equation (22.50) indicates that the dirt effect on the relative normal transmittance is also being included, in order to facilitate the direct application of the equation to real cases.

Let us continue with the example of Sapporo, Japan, by calculating the effective irradiation available on the following surfaces:

1. Optimally oriented, $\alpha = 0$, and optimally tilted, $\beta = \beta_{opt}$
   Equation (22.51), $\alpha = 0 \Rightarrow g_i = g_{i3}$
   Equation (22.50), $\beta = \beta_{opt} \Rightarrow G_{effdy}(\beta_{opt}) = 0.9314 G_{dy}(\beta_{opt}) = 3517\,W\,h/m^2$
   Note that the total losses due to the optical effects of the angle of incidence (≈7%) are larger than pure normal transmittance losses (≈3%)

Table 22.5 Coefficients used to solve Equation (22.50). Values are given for the representative case of medium dirtiness degree ($T_{dirt}(0)/T_{clean}(0) = 0.97$)

Coefficients	$T_{dirt}(0)/T_{clean}(0) = 0.97$		
	$i = 1$	$i = 2$	$i = 3$
g_{1i}	8×10^{-9}	3.8×10^{-7}	-1.218×10^{-4}
g_{2i}	-4.27×10^{-7}	8.2×10^{-6}	2.892×10^{-4}
g_{3i}	-2.5×10^{-5}	-1.034×10^{-4}	0.9314

2. Tilted 20° with respect to the horizontal, $\beta = 20°$, and oriented 30° towards the west, $\alpha = 30° \Rightarrow g_1 = -1.032 \times 10^{-4}$; $g_2 = 1.509 \times 10^{-4}$; $g_3 = 0.9057$
   $\beta - \beta_{opt} = -13.37° \Rightarrow G_{effdy}(20,30) = 0.8853 \cdot G_{dy}(\beta_{opt}) = 3343\,\text{W h/m}^2$
   It is worth mentioning that a similar calculation, but without considering the angular losses, would lead to $G_{dy}(20,30) = 0.936 \cdot G_{dy}(\beta_{opt}) = 3533\,\text{W h/m}^2$
3. A vertical facade, $\beta = 90°$, oriented towards the south-east, $\alpha = -45°$.
   $\alpha = -45° \Rightarrow g_1 = -0.885 \times 10^{-4}$; $g_2 = -2.065 \times 10^{-4}$; $g_3 = 0.8761$
   $\beta - \beta_{opt} = 56.63° \Rightarrow G_{effdy}(90,-45) = 0.5806 \cdot G_{dy}(\beta_{opt}) = 2192\,\text{W h/m}^2$

It is opportune to mention again the relatively weak sensitivity of the annual capture of energy to the inclination angle. A value of 0.2% loss for each degree of deviation from the optimum value is a rough approximation. This is true to an even greater extent in the case of azimuthal orientation, where only a 0.08% loss occurs for each degree of deviation from the south. This means that many existing surfaces (roofs, car parks, etc.) are suitable for PV modules integration, even if their orientation differs considerably from the optimum. This also means there is no need to carry out expensive civil works to level the site of PV arrays, despite it being an extended custom in large PV plants.

The case of stand-alone systems designed to feed equipment having a constant consumption throughout the year is an especially interesting one, and deserves particular mention. The design criterion here is to maximise the energy captured during the period of least radiation, rather than throughout the year. As might be expected, such receivers are positioned perpendicular to the winter sunlight, which leads to recommend a tilt angle $\beta \approx \phi + 10°$.

22.9.2 Sun-tracking Surfaces

Tracking the Sun by moving in prescribed ways to minimise the angle of incidence allow for increasing the incident radiation. Tracking mechanisms are often used in photovoltaics, mainly associated with relatively big grid-connected PV plants, where these mechanisms have already demonstrated very high reliability. As a particular example, the 100 kW_p tracking system at the Toledo PV plant is in routine operation with 100% of availability from 1994, and about half of the 1 GW_p installed in Spain during 2007 are mounted on some kind of tracker. Figure 22.20 shows some real examples. It must be mentioned that increased collection comes with an increased cost, because trackers are more expensive than static support structures. Roughly, it can be said that trackers will keep some economic sense meanwhile PV module cost remain larger than 1€/W_p.

Tracking about two axes (more often, a vertical axis to adjust the azimuth and a horizontal axis to adjust the tilt, but two other axis combinations can also be found, provided both are perpendicular) maintains the receiver surface always perpendicular to the sun ($\beta = \theta_{ZS}$; $\alpha = \psi_S$). Hence, it allows collecting the maximum amount of energy possible. Mainly depending on the clearness index, the comparison with an optimally tilted fixed surface leads to the ratio $G_{dy}(2\text{ axes})/G_{dy}(\beta_{opt})$ varying from 1.25 to 1.55 (column 5 divided by column 9 of Table 22.4).

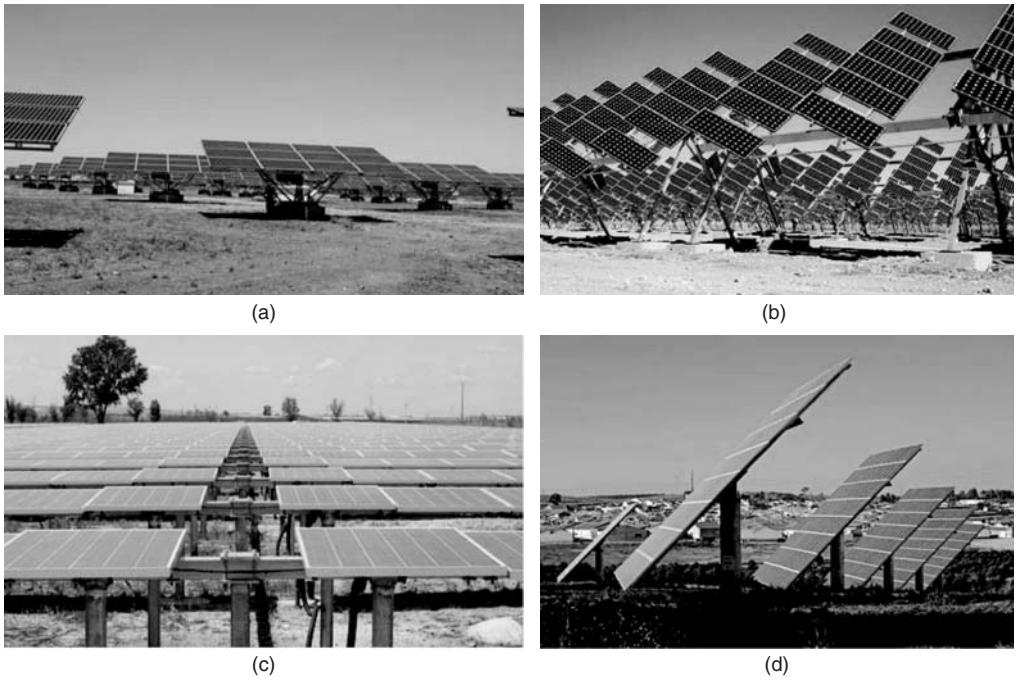


Figure 22.20 Different types of tracking: (a) two-axis, vertical axis for azimuth and horizontal axis for tilt; (b) two axis, NS horizontal axis turning from E to W and a perpendicular axis; (c) one NS horizontal axis turning from E to W; (d) one vertical axis turning from E to W, while tilt keeps constant

However, it is expensive to implement, because it uses relatively complicated mechanisms and takes up a great deal of space, due to the shadows cast. For these reasons, several types of one-axis trackers are sometime preferred. Single-axis trackers follow the sun from East to West during the day. They differ in their tilt angle.

Azimuthal one-axis trackers rotate around their vertical axis, in such a way that the azimuth of the receiver PV surface is always the same as that of the sun's azimuth. Meanwhile, the tilt angle keeps constant ($\beta = \beta_{cons}$ and $\alpha = \psi_S$). The incidence angle is given by the difference between the surface's tilt angle and the solar zenith angle ($\theta_S = \theta_{ZS} - \beta_{cons}$). Obviously, the amount of collected radiation depends on the inclination of the surface, being the maximum for a value close to the latitude. Again, the sensitivity of the annual capture of energy to this inclination angle is relatively low. A typical value of approximately 0.4% loss from each degree of deviation from the optimum inclination can be assumed. Note that an azimuthal tracker tilted to the latitude collects up to 95% of the yearly irradiation compared to two-axes tracking (column 6 divided by column 5 of Table 22.4).

Trackers turning around a single axis oriented N–S and tilted at an angle β_{NS} to the horizontal are also of great interest due to their mechanical robustness. It can be seen that, in order to minimise the solar incident angle, the rotation angle of the axis, ψ_{NS} – 0 at noon – must be

$$\tan \psi_{NS} = \frac{\sin \omega}{\cos \omega \cos \beta_\Delta - [\text{sign}(\phi)] \tan \delta \sin \beta_\Delta} \tag{22.52}$$

where $\beta_\Delta = \beta_{NS} - abs(\phi)$. The corresponding solar incident angle is given by

$$\cos \theta_S = \cos \psi_{NS} (\cos \delta \cos \omega \cos \beta_\Delta - [\text{sign}(\phi)] \sin \delta \sin \beta_\Delta) + \sin \psi_{NS} \cos \delta \sin \omega \tag{22.53}$$

A common configuration, called *polar tracking*, is when the axis is inclined just to the latitude. Then, the rotation axis is parallel to the rotation axis of the Earth, and Equation (22.53) becomes reduced to $\theta_S = \delta$. Because of the variation of the declination during the year, the cosines of the solar incident angle range between 0.92 and 1, having an annual mean value of about 0.95. This way, the polar tracker also collects about 95% of the energy corresponding to the two-axes case (column 8 divided by column 5 of Table 22.4). It is interesting to note that a polar tracker turns at just half the angular speed as that of a standard clock.

Another common configuration is when the axis is just horizontal. Horizontal one-axis trackers are of particularly simple construction and do not cast shadows in the N–S direction. This encompasses significant radiation reduction when compared with two-axis tracking (column 7 divided by column 5 of Table 22.4), but still significant radiation increase when compared with optimally tilted fixed surfaces (column 6 divided by column 8 of Table 22.4). Because of this, they are today a common tracking solution in large PV plants: PVUSA [45], Toledo [38]. And the same is true for solar thermal plants. We should remember that the very first solar tracker used in any significant way for power generation was just a N–S-oriented horizontal one-axis tracker associated to a parabolic reflecting trough, constructed in 1912 by Frank Shumann and C.V. Boys to power a 45-kW steam-pumping plant in Meadi, Egypt [46]. The tracking surface covered an area of 1200 m^2. The plant was a technical success, that is, reliable trackers already existed at the time, but it was shut down in 1915 due to the onset of World War I and cheaper fuel prices. The world's largest solar plant, at the well-known Luz solar thermal field, erected in California from 1984 to 1986, also employs this type of tracking and, again, with great technical success [47]. The rotation angle (which, in this particular case, coincides with the surface tilt) and the solar incident angle are given by

$$\psi_{NS} = \beta = \arctan\left|\frac{\sin\psi_S}{\tan\gamma_S}\right| \tag{22.54}$$

and

$$\cos\theta_S = \cos\delta \cdot [\sin^2\omega + (\cos\phi\cos\omega + \tan\delta\sin\phi)^2]^{1/2} \tag{22.55}$$

Finally, it is worth commenting that large PV generators have many, even hundreds, of rows of modules mounted above the ground. The distance between rows affects the energy produced by the PV generator. If the separation is increased, fewer shadows are cast by some rows on the others and more energy is produced. But it also affects the cost, as greater separations lead to more land being occupied, longer cables and more expensive civil works. Therefore, there is an optimum separation, giving the best trade-off between greater energy and lower cost. There is a widely held view that tracking generally requires much more land than static arrangements. The interested reader is encouraged to consult reference [41], which deals in detail with tracking and shadowing in large PV arrays. The ratio between the area of the PV array and the area of the land is called *ground cover ratio, GCR*. A rough idea for middle latitudes is $GCR \approx 0.5$ for static surfaces, $0.25 < GCR < 0.5$ for horizontal axis and $0.1 < GCR < 0.2$ for two axis trackers. As a relevant example, Figure 22.21 shows trackers at a 48 MW plant located in Amaraleja, Portugal.

22.9.3 Concentrators

PV concentrators are able to capture only direct-beam solar radiation and require tracking mechanisms to keep in focus the solar cells. Therefore, they are best suited to sunny sites with low scattering (or low diffuse radiation) and good solar resources. Figure 22.22 compares the amount of normal direct radiation for a PV concentrator mounted on a two-axis tracker, with the amount of global radiation, collected by a conventional flat-plate PV collector, at a fixed optimal tilt. The data represents the 30 locations in Table 22.4. The ratio between both quantities, $B_{dy}(\perp)/G_{dy}(\beta_{opt})$,

Figure 22.21 Trackers at Amaraleja 48 MW PV plant. Each tracker has a surface of 140 m^2, supporting a 182 kW_p PV array: (a) no shading occurs during most the day; (b) some shade occurs in early morning and late afternoon. See Plate 7 for the colour figure

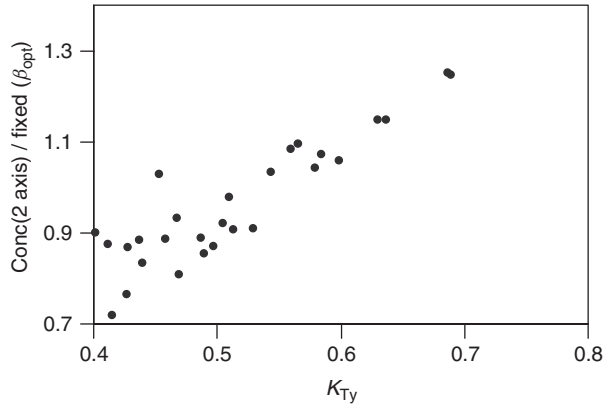


Figure 22.22 Comparison of annual average solar radiation available for fixed flat-plate conventional PV modules, and two-axis tracking PV concentrators. The y-axis is (column 10)/(column 9) and the x-axis column 4 of Table 22.4

(column 9 divided by column 8 of Table 22.4) is plotted against the annual clearness index. As a general rule, two-axis tracking concentrators collect more radiation than fixed flat-plate modules in places where $K_{Ty} > 0.55$. More information about PV concentrators is found in Chapter 10.

22.10 PV GENERATOR BEHAVIOUR UNDER REAL OPERATION CONDITIONS

A problem that the engineer frequently has to solve is the prediction of the electrical behaviour of a PV generator, given the information about the generator's construction, geographical location and the local weather. In particular this represents the base for predicting the generator energy delivery, which is a critical step of any PV system design. This leads to the question of establishing a PV module rating condition, at which power performance and other characteristics are specified and,

defining a method for calculating performance for a given set of environmental conditions such as solar irradiance, ambient temperature, wind speed and so on.

Traditionally, PV modules are being rated under the so-called standard test conditions (STC (irradiance 1000 W/m^2; spectrum *AM* 1.5; and cell temperature 25 °C). STC is also referred to as standard reporting conditions (SRC) in some publications. In the following, we will use an asterisk* to refer to parameters measured under these conditions. Typically, values of the short-circuit current, the open-circuit voltage, and the maximum power, which are always included in the manufacturer's data sheets. Furthermore, the characterisation of the PV module is completed by measuring the *nominal operating cell temperature* (*NOCT*), defined as the temperature reached by the cells when the PV module is submitted to an irradiance of 800 W/m^2, an ambient temperature of 20 °C and a wind speed of 1 m s^{-1}. However, performance predictions based on STC are being continuously questioned, mainly because the actual annual energy efficiencies are significantly lower than the power efficiencies predicted using STC. Certainly, PV users may be astonished to learn that the real efficiency of the PV module they have purchased, once installed and operating under real world conditions, has only about 70% of the STC efficiency they have read in the manufacturer's information. This could lead to a feeling of having been deceived. All together, these facts have stimulated several authors to propose other methods for PV module rating and energy performance estimation, more oriented to give buyers clear and more accurate information about the energy generation of PV modules (see Chapter 17 for a detailed description on rating module performance). This is still an open question, and it is difficult to predict the extent to which these new models would be incorporated in future PV engineering practices. Most of these methods require extending testing beyond STC, and rely on a relatively large number of parameters that should be empirically determined. Surely, the use of a large number of parameters would potentially allow for more accurate energy modelling. But it is not clear as to what extent the possible improvements would compensate for the associated increase of complexity derived from the experimental determination of such parameters. The proponents of new methods tend to argue that their procedures can be easily implemented. (Their papers used to include such phrases as... *the entire test procedure for outdoor measurements, including the set-up, takes approximately three hours* ... [48]). But, at the same time, the PV module manufacturers are very reluctant to incorporate troublesome novelties into their module characterisation procedures already established at factory PV. This dilemma is, in fact, easily understandable considering the inherent difficulties associated with the adoption of any innovation. This was magnificently explained by Machiavelli in his famous work *The Prince*, written in 1513: (*There is nothing more difficult to plan, more doubtful of success* ... *than the creation of a new order of things* ...). Together, claims of low PV module energy performance, and the flourishing of proposals for new rating methodologies have led, no doubt, to significant confusion in today's PV community, making this author's task, of selecting a particular methodology for recommending to his PV colleagues, risky. However, it is this author's opinion that, at least in the case of crystalline silicon, energy performance modelling based on only a few parameters obtained at STC, and always included in the manufacturer's data sheets, can lead to adequate predictions, providing that some judicious considerations are made. Obviously, precautions to assure that actual STC power of the purchased PV modules correspond to the manufacturer's declarations are a different matter [49].

It must be remembered that crystalline silicon solar cells remain the workhorse for PV power generation, despite significant advances in other PV devices. For example, c-Si technology increased its world market share [50], from 87% in 1999 (over a total of 202 MW) to 89.6% in 2007 (over a total of 4.36 GW). This predominance means that actual information concerning in-field c-Si PV modules performance is well established, which is not typically the case where other materials are concerned. Because of this, dealing with c-Si PV modules is, by far, the simplest and easiest case for PV designers. This is why, despite the confusion, the author dares to detail here a rather simple methodology that allows the estimation of the $I-V$ curve of c-Si PV modules operating in any prevailing environmental condition, exclusively using as input the values of I_{SC}^*, V_{OC}^* and

P_M^*. In principle, such methodology can be extended to materials other than c-Si, and additional comments to do this are also included in the text that follows.

22.10.1 The Selected Methodology

In a previous chapter (see Chapter 3), it has been shown that the characteristic $I-$of a solar cell can be expressed with sufficient accuracy by

$$I = I_L - I_0\left(\exp\frac{V + IR_S}{V_t} - 1\right) - \frac{V + IR_S}{R_P} \tag{22.56}$$

where I_L, I_0, R_S and R_P are the photogenerated current, the dark current, the series resistance and the parallel resistance, respectively. The voltage V_t equals mkT/e (we recall that for $m = 1$, $V_t \approx 25\,\text{mV}$ at 300 K). This expression gives an adequate representation of the intrinsic behaviour of a typical c-Si solar cell. Nonetheless, it cannot be used directly to obtain the required predictions, because some parameters I_L and I_O in particular, cannot be established from the usually available information, often restricted to the values of and which are always included on the manufacturers' data sheets.

This difficulty is effectively overcome when the following assumptions, which are generally valid for c-Si PV cells and modules, are made:

- The effect of parallel resistance is negligible.
- The photogenerated current and the short-circuit current are equal.
- $\exp((V + IR_S)/V_t) \gg 1$.

This allows Equation (22.56) to be written as

$$I = I_{SC} - I_O \exp\left(\frac{V + IR_S}{V_t}\right) \tag{22.57}$$

which, with $I = 0$, leads to the following expression for the open-circuit voltage:

$$V_{OC} = V_t \ln\left(\frac{I_{SC}}{I_O}\right) \tag{22.58}$$

where

$$I_O = I_{SC} \exp\left(-\frac{V_{OC}}{V_t}\right) \tag{22.59}$$

Substituting Equation (22.58) in (22.57), we arrive at

$$I = I_{SC}\left[1 - \exp\left(\frac{V - V_{OC} + IR_S}{V_t}\right)\right] \tag{22.60}$$

This is a very useful expression. As we shall see, the values of all the parameters on the right-hand side are easily obtained allowing immediate application of the expression. An inconsistency arises in the sense that $I(V = 0) \neq I_{SC}$. Nevertheless, in all solar cells of practical use, we find that $V_{OC} \gg IR_S \Rightarrow I(V = 0) \approx I_{SC}$, which therefore makes this objection irrelevant. The expression can be inconvenient to use in the sense that I is implicit (it appears on both sides of the equation), theoretically making it necessary to solve the equation iteratively. However, for voltages close to the maximum-power point, a reasonably accurate solution can be obtained with only one iteration by setting $I_M = 0.9 \times I_{SC}$ in the second term.

The calculation of the maximum power can, in principle, be carried out by considering that the power is given by the product $P = VI$. The values of I_M and V_M, defining the maximum power operation point, can be obtained from the usual condition for a maximum, $dP/dV = 0$. However, the implicit nature of the resulting expression makes it very cumbersome to use. It would be better to look for a simpler method, based on the existing relationship between the fill factor and the open-circuit voltage. According to M.A. Green [51] an empirical expression describing this relation is

$$FF = \frac{V_M I_M}{V_{OC} I_{SC}} = \frac{P_M}{V_{OC} I_{SC}} = FF_0(1 - r_s) \tag{22.61}$$

where

$$FF_0 = \frac{v_{OC} - \ln(v_{OC} + 0.72)}{v_{OC} + 1} \tag{22.62}$$

and $v_{OC} = V_{OC}/V_t$ and $r_s = R_S/(V_{OC}/I_{SC})$ are defined as the normalised voltage and the normalised resistance, respectively. It is interesting to note that the series resistance at STC can be deduced from the manufacturer's data by the expression

$$R_S = \left(1 - \frac{FF}{FF_0}\right)\frac{V_{OC}}{I_{SC}} \tag{22.63}$$

The values of V_M and I_M are in turn given by [52]

$$\frac{V_M}{V_{OC}} = 1 - \frac{b}{v_{OC}} \ln a - r_s(1 - a^{-b}) \text{ and } \frac{I_M}{I_{SC}} = 1 - a^{-b} \tag{22.64}$$

where:

$$a = v_{OC} + 1 - 2\, v_{OC} r_s \text{ and } b = \frac{a}{1 + a} \tag{22.65}$$

This set of expressions is valid for $v_{OC} > 15$ and $r_s < 0.4$. The typical accuracy is better than 1%. Their application to a photovoltaic generator is immediate, if all their cells are supposed to be identical, and if the voltage drops in the conductors connecting the modules are negligible.

Now, for the prediction of the $I-V$ curve of a PV generator operating on arbitrary conditions of irradiance and temperature, a good balance between simplicity and exactness is obtained through the following additional assumptions.

- The short-circuit current of a solar cell depends exclusively and linearly on the irradiance. That is,

$$I_{SC}(G) = \frac{I^*_{SC}}{G^*} G_{eff} \tag{22.66}$$

  where G_{eff} is the 'effective' irradiance. This concept must take into consideration the optical effects related to solar angle of incidence, as described in Section 22.7.
- The open-circuit voltage of a module depends exclusively on the temperature of the solar cells T_c. The voltage decreases linearly with increasing temperature. Hence,

$$V_{OC}(T_c) = V^*_{OC} + (T_c - T^*_c)\frac{dV_{OC}}{dT_c} \tag{22.67}$$

  where the voltage temperature coefficient, dV_{OC}/dT_c is negative. The measurement of this parameter use to be included in PV modules characterisation standards [53] and the corresponding value must, in principle, be also included on the manufacturer's data sheets. For crystalline silicon cells, dV_{OC}/dT_c is typically -2.3 mV per°C and per cell.

- The series resistance is a property of the solar cells, unaffected by the operating conditions.
- The operating temperature of the solar cell above ambient is roughly proportional to the incident irradiance. That is,

$$T_c = T_a + C_t G_{eff} \tag{22.68}$$

where the constant C_t has the value:

$$C_t = \frac{NOCT(^\circ C) - 20}{800\ W/m^2} \tag{22.69}$$

The values of *NOCT* for modules currently on the market varies from about 42 to 46 °C, implying a value of C_t between 0.027 and 0.032 °C/(W/m^2). When *NOCT* is unknown, it is reasonable to approximate $C_t = 0.030\,^\circ C/(W/m^2)$. This *NOCT* value corresponds to mounting schemes allowing the free air convection in both sides of the PV modules, which cannot be the case on roof-mounted arrays, that restrict some of the airflow. Then, it has been shown [54] that the *NOCT* increases by about 17 °C if some kind of back ventilation is still allowed, and up to 35 °C if the modules are mounted directly on a highly insulated roof.

Example: To illustrate the use and the usefulness of the equations of the preceding sections, we shall analyse the electrical behaviour of a PV generator rated at 1780 W (STC), made up of 40 modules, arranged 10 in series × 4 in parallel. The conditions of operation are $G_{eff} = 700\ Wm^{-2}$ and $T_a = 34\ ^\circ C$. It is known that the modules have the following characteristics under STC: $I^*_{SC} = 3\,A, V^*_{OC} = 19.8\,V$ and $P^*_M = 44.5\,W$. Further, it is known that each module consists of 33 cells connected in series and that $NOCT = 43\ ^\circ C$.

The calculations consist of the following steps:

1. *Determination of the characteristic parameters of the cells that make up the generator under STC (Equations 22.61–22.65)*:
   33 cells in series ⇒ Per cell: $I^*_{SC} = 3\,A$, $V^*_{OC} = 0.6\,V$ and $P^*_M = 1.35\,W$ and
   assuming m = 1; $V_t(V) = 0.025 \times (273 + 25)/300 = 0.0248V \Rightarrow v_{OC} = 0.6/0.0248 = 24.19 > 15$
   then, $FF_0 = (24.19 - \ln(24.91))/25.19 = 0.833$; $FF = 1.35/(0.6 \times 3) = 0.75$
   and $r_s = 1 - 0.75/0.833 = 0.0996 < 0.4 \Rightarrow R_S = 0.0996 \times 0.6/3 = 19.93\,m\Omega$

$$a = 20.371;\ b = 0.953 \Rightarrow V_M/V_{OC} = 0.787 \text{ and } I_M/I_{SC} = 0.943$$

   It is worth noting that these values lead to a value of $FF = 0.742$, slightly different from the starting value. This error shows the precision available by the method, better than 1% in this instance. Sometimes values of $m = 1.2$ or 1.3 give a better approximation.
2. *Determination of the temperature of the cells under the operating conditions considered (Equations 22.68 and 22.69)*:

$$C_t = 23/800 = 0.0287\,^\circ C\,m^2/W \Rightarrow T_c = 34 + 0.0287 \times 700 = 54.12\,^\circ C$$

3. *Determination of the characteristic parameters of the cells under the operating conditions considered (Equations 22.66 and 22.67)*:

$$I_{SC}(700\,W/m^2) = 3 \times (700/1000) = 2.1\,A$$
$$V_{OC}(54.12\,^\circ C) = 0.6 - 0.0023 \times (54.12 - 25) = 0.533\,V$$

With R_S considered constant, these values lead to:

$$V_t = 27.26\,\text{mV};\ v_{OC} = 19.55;\ r_s = 0.0785;\ FF_0 = 0.805;\ FF = 0.742;\ P_M = 0.83\,\text{W}$$

4. *Determination of the characteristic curve of the generator, (I_G, V_G)*: Number of cells in series 330; Number of cells in parallel: 4. Then:
$I_{SCG} = 4 \times 2.1\ \text{A} = 8.4\ \text{A};\ V_{OCG} = 330 \times 0.533\ \text{V} = 175.89\,\text{V};\ R_{SG} = 1.644\Omega;$

$$I_G(A) = 8.4\left[1 - \exp\frac{V_G(V) - 175.89 + 1.644 \cdot I_G(A)}{9.00}\right]$$

To calculate the value of the current corresponding to a given voltage, we may solve this equation iteratively, substituting I_G for $0.9I_{SCG}$ on the first step. Only one iteration is required for $V_G \leq 0.8V_{OCG}$. By way of example, the reader is encouraged to do it for $V_G = 140$ V and $V_G = 150$ V. The solution is $I_G(140\,\text{V}) = 7.77$ A and $I_G(150\,\text{V}) = 6.77$ A.

5. *Determination of the maximum power point*:

$$a = 17.48;\ b = 0.9458;\ V_M/V_{OC} = 0.7883;\ I_M/I_{SC} = 0.9332$$
$$V_M = 138.65\,\text{V};\ I_M = 7.84\,\text{A},\ P_M = 1087\,\text{W}.$$

Note that the ratio $P_M \big/ P_M^* = 0.661$, while the ratio $G_{eff}/G^* = 0.7$. This indicates a decrease in efficiency at the new conditions compared to STC, primarily due to the greater solar cell temperature, $T_c < T_c^*$. An efficiency temperature coefficient can now be obtained by

$$\frac{1}{\eta^*} \cdot \frac{d\eta}{dT_c} = \left(\frac{P_M}{P_M^*} \cdot \frac{G^*}{G_{eff}} - 1\right) \cdot \left(\frac{1}{T_c - T_c^*}\right) = -0.004\,\big/^\circ\text{C}$$

This means the efficiency, hence the power decrease is about 0.4% per degree of temperature increase, which can be considered as representative for c-Si. This value, calculated for typical parameters with some assumptions, agrees very well with the range of values commonly reported on module data sheets from −0.45 to −0.5%/°C.

In fact, this relation between power and solar cell temperature can always be directly employed. In this way:

$$P_M(G_{eff}, T_c) = P^* \frac{G_{eff}}{G^*}[1 - \gamma(T_c - T_c^*)] \tag{22.70}$$

or

$$\eta(T_c) = \eta^*[1 - \gamma(T_c - T_c^*)] \tag{22.71}$$

where γ is the so-called temperature power coefficient. Note that these expressions allow for calculation of power in any operation conditions without the need of previous calculation of the full $I-V$ curve. This is very useful for simulation purposes, because the computing time is greatly reduced.

It should be noted that, depending on the input data availability, other PV generator modelling possibilities exist. For example, commonly and are given in the specifications in addition to their product Then, the series resistance can directly be estimated from Equation (22.60). This leads to

$$R_S^* = \frac{V_{OC}^* - V_M^* + V_t \ln\left(1 + \dfrac{I_M^*}{I_{SC}^*}\right)}{I_M^*} \tag{22.72}$$

It should be noted that a similar but different set of $I-V$ translations are found in Section 17.2 of Chapter 17 of this book.

22.10.2 Second-order Effects

The model presented in the previous section is based only on standard and widely available information, which is an undeniable advantage, in particular for PVsystem design. Furthermore, it is simple to use. However, it can be argued that such simplicity is at the price of neglecting the following:

- The effects of the parallel resistance.
- The influence of the cell temperature in the short-circuit current.
- The influence of the irradiance in the open-circuit voltage.
- The nonlinearity due to low irradiance.
- The spectral effects.
- The effects of wind.

It should be recognised that differences between expected and real energies delivered by PV modules are often mentioned in the literature [55]. Hence, it is worth reviewing the importance of each one of these previously neglected factors, with the aim of clarifying possible error sources. Many of these factors are discussed further in Chapter 16.

The influence of the parallel resistance is, to a great extent, compensated here, by the particular way of estimating the series resistance of a PV module, which ensures that the maximum power of the modelled curve coincides exactly with that corresponding to the real one. Because of this, the accuracy of the model tends to be very good just around the maximum power operation point, that is, just on the voltage region of interest.

The short-circuit current tends to increase slightly with increasing temperature. This can be attributed, in part, to increased light absorption, since semiconductor bandgaps generally decrease with temperature, and, in part, to increased diffusion lengths of the minority carriers. This can be considered by adding a linear term to Equation (22.66). Thus

$$I_{SC}(G, T_c) = I^*_{SC} \cdot \frac{G}{G^*} \cdot \left[1 + (T_c - T^*_c)\frac{dI_{SC}}{dT_c}\right] \tag{22.73}$$

where the temperature coefficient dI_{SC}/dT_c depends on the semiconductor type and on the manufacturing process, but it is always quite small. Typical experimental values are below 3×10^{-4} (A/A)/° C [56]. For a solar cell operating at 70 °C, that represents only 1.3% of I_{SC} increase. Hence, ignoring this dependence has little effects, in most the cases.

The open-circuit voltage tends to increase with the illumination level. The ideal diode equation (see Equation 22.58) predicts a logarithmic dependence. Then, this effect can be considered by adding a logarithmic term to Equation (22.67). Thus

$$V_{OC}(T_c, G) = V^*_{OC} + (T_c - T^*_c)\frac{dV_{OC}}{dT_c} + V_t \cdot \ln\left(\frac{G_{eff}}{G^*}\right) \tag{22.74}$$

Note that the relative influence of this new term increases with decreasing irradiance. For example, for $G_{eff} = 500$ W/m^2 and $G_{eff} = 200$ W/m^2, it represents about 3 and 7%, respectively. Hence, its importance when predicting the energy delivered by PV modules depends on the irradiance distribution of the irradiation content. Obviously, it is more important for northern than for southern countries. For example, in Freiburg, Germany ($\phi = 48°$) about 50% of yearly irradiation is collected below 600 W/m^2 and 18% below 200 W/m^2. Meanwhile in Jaen, Spain ($\phi = 37.8°$) about 30% of yearly irradiation is collected below 600 W/m^2 and only 6% below 200 W/m^2. Moreover, it should be understood that PV systems have a minimum irradiance threshold. For example, in

grid-connected systems, DC power from PV modules must be large enough to compensate for the inverter losses. Otherwise, the PV system becomes a net energy consumer.

While this logarithmic term does account for some variation in open-circuit voltage as irradiance changes, it does not adequately predict the rapid decrease observed at values of irradiance less than about 200 W/m^2, which causes a noticeable efficiency decrease below this value. The use of a second logarithmic term has been proposed [57] to also consider this low irradiance effect. Thus

$$V_{OC}(T_c, G) = \left[V^*_{OC} + \frac{dV_{OC}}{dT_c}(T_c - T^*_c)\right]\left[1 + \rho_{OC}\ln\left(\frac{G_{eff}}{G_{OC}}\right)\ln\left(\frac{G_{eff}}{G^*}\right)\right] \tag{22.75}$$

where ρ_{OC} and G_{OC} are empirically adjusted parameters. Values of $\rho_{OC} = -0.04$ and $G_{OC} = G^*$ have proven adequate for many silicon PV modules.

The sun spectrum shifts over time, due to changes in atmosphere composition, and changes in the distance the light has to travel through the atmosphere. This can affect the response of PV devices, especially if they have a narrow spectral response. Martin and Ruiz have proposed [58] a model based on the parameterisation of the atmosphere by means of the clearness index and the air mass, and this considers independently the spectrum of each radiation component: direct, diffuse and albedo. It can be described by modifying Equation (22.66). Thus

$$I_{SC} = \frac{I^*_{SC}}{G^*}(B_{eff} \cdot f_B + D_{eff} \cdot f_D + R_{eff} \cdot f_R) \tag{22.76}$$

where f_B, f_D and f_R obeys the general form

$$f = c \cdot \exp[a(K_T - 0.74) + b(AM - 1.5)] \tag{22.77}$$

and a, b and c are empirically adjusted factors, for each module type and for each radiation component. Note that 0.74 and 1.5 are just the values of the atmospheric parameters corresponding to the STC. Table 22.6 shows the recommended values of these parameters for crystalline, c-Si, and amorphous a-Si modules. The usefulness of this table can be extended to other PV materials, by linear interpolation on the energy bandgap, E_g, between E_g(c-Si) = 1.12 eV and E_g(a-Si) = 1.7 eV. For example, E_g(a-SiGe) = 1.4 eV $\Rightarrow$ estimated value of c for direct irradiance is equal to 0.8.

Spectral effects used to be small on a yearly basis. Spectral losses with respect to STC are typically below 2% with semiconductors with broad spectral sensitivity, and below 4% for the others. However, on an hourly basis, spectral effects up to 8% can be encountered.

The PV module cell temperature is a function of the physical variables of the PV cell material, the module and its configuration, the surrounding environment and the weather conditions. It results from the balance of energy inputs and outputs through radiation, convection, conduction and power generation. Today, the more widely extended model, based on the *NOCT* concept and described by Equations (22.68 and 22.69), lump the contributions together in an overall heat-loss coefficient, resulting in a linear relationship between module temperature and irradiance under

Table 22.6 Coefficient for spectral response modelling, for c-Si and a-Si modules

	B_{eff}		D_{eff}		R_{eff}	
	c-Si	a-Si	c-Si	a-Si	c-Si	a-Si
c	1.029	1.024	0.764	0.840	0.970	0.989
a	−3.13E-01	−2.22E-01	−8.82E-01	−7.28E-01	−2.44E-01	−2.19E-01
b	5.24E-03	9.20E-03	−2.04E-02	−1.83E-02	1.29E-02	1.79E-02

Table 22.7 Temperature coefficients for module and cell temperature estimation, for two typical module designs

Type	T_1 [°C]	T_2 [°C]	b	ΔT [°C]
Glass/cell/glass	25.0	8.2	−0.112	2
Glass/cell/tedlar	19.6	11.6	−0.223	3

steady-state conditions. This implies accepting that the heat transfer process between the solar cell and the ambient is essentially dominated by the conduction through the encapsulating materials, and neglecting the wind effects on convection. This model is simple to use and requires only standard available input information, which are undeniable advantages for the PV designer. But it can lead to significant errors in cell temperature estimation for non-steady-state conditions [59] (observed thermal time constant of PV modules is about 7 minutes), and for high wind speeds. That has stimulated several authors to develop new thermal models for PV systems, based not only on irradiance, but also on wind speed. For example, Sandia Laboratories proposed [60] a two-component thermal model given by

$$T_c = T_m + \frac{G_{eff}}{G^*}\Delta T, \text{ where } T_m = T_a + \frac{G_{eff}}{G^*}[T_1 \exp(b \cdot U) + T_2] \tag{22.78}$$

In this equation, T_m is the back-surface module temperature, in °C; U is the wind speed measured at standard 10 m height, in m.s^{-1}; T_1 is an empirical coefficient determining the upper temperature limit at low wind speeds; T_2 is an empirical coefficient determining the lower temperature limit at high wind speeds; b is an empirical coefficient determining the rate at which the module temperature drops as wind speed increases and ΔT is also an empirical coefficient related to the temperature gap along the back encapsulation material. Table 22.7 gives the parameters found to be in good agreement with measured temperatures, for two different module types.

However, the real usefulness of all these second-order corrections (Equations 22.74–22.77) for temperature and wind speed is far from clear, because wind speed is difficult to predict, and also because cell temperature errors becomes more tolerable when translated into PV module power generation. For example, for a solar cell operating around 50 °C, an error of 20% on the estimation of its cell temperature (≈10 °C) reflects in an error of only about 4% (≈10 °C × 2.3(mV/°C)/600 mV) in the estimation of the corresponding power.

22.11 RELIABILITY AND SIZING OF STAND-ALONE PV SYSTEMS

The merit of a stand-alone PV system depends on how reliably it supplies electricity to the load. It is customary to quantify this reliability in terms of the *loss of load probability* (*LLP*), defined as the ratio between the energy deficit and the energy demand, both referring to the load, over the total operation time of the installation. Thus

$$LLP = \int_t \text{energy deficit} / \int_t \text{energy demand} \tag{22.79}$$

It should be noted that, because of the random nature of solar radiation, the value of *LLP* is always greater than zero, even if the PV system never actually breaks down. The available literature shows a large consensus over the expression of reliability in terms of energy shortage probability, but different names can be found for the *LLP*: deficit of energy [61], loss of power probability [62]

and loss of power supply probability [63]. Moreover, the load coverage rate [64] or solar fraction *SF* [65], defined as the fraction of energy load covered by the PV system, is also used to quantify reliability. Clearly, $SF = 1 - LLP$.

The 'size' of a PV system means the capacity of both the generator (PV modules) and the accumulator (batteries or other storage device). It is useful to relate these sizes to the size of the load, in terms of average daily energies. Thus, the generator capacity, C_A, is defined as the ratio of the average daily energy output of the generator divided by the average daily energy consumption of the load. The accumulator capacity C_S, is defined as the maximum energy that can be extracted from the accumulator divided by the average daily energy consumption of the load. Thus,

$$C_A = \frac{\eta_G \cdot A_G \cdot \overline{G_d}}{L} \text{ and } C_S = \frac{C_u}{L} \tag{22.80}$$

where A_G and η_G are the area and conversion efficiency of the photovoltaic generator, respectively, $\overline{G_d}$ is the mean value of the daily irradiation on the surface of the generator, L is the mean value of the daily energy consumed by the load and C_u is the useful energy storage capacity of the accumulator. More strictly, η_G should be the path efficiency from the array to the load including storage, resistance and ac/dc conversion losses, and C_u is the product of the nominal capacity (which refers to the whole energy that can be extracted from the accumulator if no particular limitations were imposed) and the maximum allowable depth of discharge. We will deal with the practical meaning of such parameters later. Meanwhile, it is worth pointing out that C_A depends on the local solar climate conditions. Therefore, the same photovoltaic generator, connected to the same load, may seem big in one place and small in another where there is less radiation.

Figure 22.23 shows how the energy generation varies over an assumed period of j days, for a given location and load, and for two different sizes of the generator ($C_{A1} < C_{A2}$). The shaded areas underneath the line $y = 1$, illustrate the temporal deficits of energy generation that need to be compensated by extraction of energy from the accumulator. It can be observed that the larger the generator, the lower the deficit and, hence, the smaller the required accumulator. Two ideas are now intuitively apparent: the first is that it is possible to find different combinations of C_A and C_S that lead to the same value of *LLP*; the second is that the larger the photovoltaic system, the better the reliability, that is, the lower the value of *LLP*, but also the greater the cost.

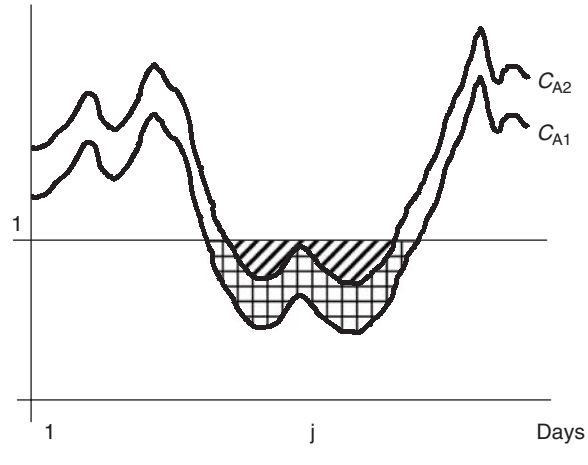


Figure 22.23 Time variation of the generated daily energy, relative to load, for two PV generators for which $C_{A1} < C_{A2}$. The shaded areas represent the corresponding energy deficits to be covered by the storage device

The degree of reliability needed depends on the type of load. For example, reliability requirements of telecommunication equipment are usually higher than that required by domestic appliances (higher reliability means lower *LLP* values). The problem confronting the PV engineer then takes the following 'theoretical' form: Which combination of C_A and C_S achieves the desired *LLP* with minimum cost? Since cost estimation is a classic economic problem discussed widely in the literature, the PV sizing problem is mainly rooted in the relationship between C_A, C_S and *LLP*. Later, C_A and C_S must be translated into the number and power of PV modules and battery capacity.

In essence, for a given load, any PV sizing method involves four different steps:

- obtaining solar radiation site information;
- preparation of global horizontal daily irradiation sequences;
- transposition from horizontal to inclined radiation values;
- simulation of PV system behaviour, in order to quantify *LLP* corresponding to pairs of C_A and C_S values.

The first three steps have already been discussed in previous sections of this chapter. We will now deal with the last step, with the following assumptions: first, the daily energy consumption is constant all through the year; second, all the daily consumption occurs at night (i.e. after the energy generation time ends); and, third, the components of the PV system are ideal and can be linearly modelled. This is adequate for analysing the 'pure' sizing problem, that is, the relation between C_A, C_S and *LLP*. Nonlinearities and non-ideal effects (for example, battery efficiency) are better taken into account by the use of proper correction factors when translating C_A and C_S values into nominal PV array power and battery capacity. It is interesting to note that short-term (hourly) variations of demanded energy have no effect on *LLP* [61], provided that $C_S > 2$, which is usually the case. The assumption described leads to particularly simple calculations. The state of charge, *SOC*, of the accumulator at sunset of day j is given by

$$SOC_j = \min\left\{SOC_{j-1} + \frac{C_A \cdot G_{dj}}{C_S \cdot \overline{G_d}} - \frac{1}{C_S}; 1\right\} \tag{22.81}$$

where G_{dj} is the total irradiation for day j, and

$$LLP = \sum_{1}^{N} E_{LACK}\, j/(NL) \tag{22.82}$$

where N is the number of days for which the simulation is carried out, and E_{LACK} is the daily energy deficit, given by

$$E_{LACK}\, j = \max\left\{\frac{1}{C_S} - SOC_j; 0\right\} \tag{22.83}$$

This equation implies that an energy deficit occurs only when the stored energy at the end of the day $SOC_j \cdot C_S \cdot L$ does not suffice to cover the daily load L. Note that $SOC_j \cdot C_S \cdot L < L \Rightarrow (1/C_S - SOC_j) > 0$.

Because information on the daily irradiation used to be expressed in terms of mean monthly values, a different value of C_A may be worked out for each month. In what follows, we shall use for C_A just the value corresponding to the worst month. Thus,

$$\overline{G_d} = \min\{G_{dm}(\beta, \alpha)\};\ (m = 1, \ldots, 12) \tag{22.84}$$

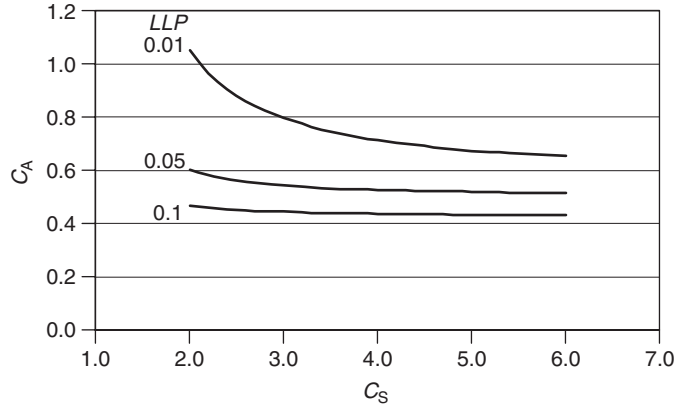


Figure 22.24 Reliability maps: generator capacity C_A versus storage capacity C_S with the reliability *LLP* as parameter

Computed results can be arranged and presented graphically. Figure 22.24 presents some examples of the so-called reliability maps. It shows that adding significant storage capacity is pointless without adding more PV to fill the storage.

The shape of such curves prompts the thought that it should be possible to describe them in an analytical form. Several attempts at proposing analytical methods based on this idea have been made [66–72] allowing sizing of PV systems by means of straightforward, simple calculations. For example, at the IES-UPM we have concluded [72] that all these curves conform to the relationship

$$C_A = f.C_S^{-u} \tag{22.85}$$

where f and u are two parameters that depend on the value of *LLP* and on the location. These parameters do not have particular physical meaning, and their determination previously requires a lot of simulations to be carried out. However, they allow a large number of simulated results to be condensed into just a few izing parameters.

It must be stressed that, whatever the detailed methodology, PV-system sizing relies on future prediction based on past observations of the solar radiation. Basic statistical laws imply that such prediction exercises are unavoidably associated with a degree of uncertainty, as mentioned before. This implies a basic limit of accuracy for PV sizing. We will try to clarify this aspect with an example.

We will suppose that 20 years of daily irradiation data measured in a certain location with a great level of accuracy are available. We will call them the 'historical sequence'. This allows us first, to establish the statistical characteristics of the radiation (mean value, standard deviation, etc.); and, second, to make detailed simulations of a PV system's behaviour over these particular years. Thus, we can map with high precision the reliability associated with different system sizes, for this particular historical sequence. As a simulation exercise, the accuracy of the result is limited only by the precision of the initial measurements, which have been assumed to be very good. However, when using such a sequence for sizing a future system, another limitation arises simply because the solar radiation in the future will not exactly repeat the same pattern as in the past. In fact, it is extremely unlikely in terms of daily sequences. All that can be expected is that the future solar radiation sequence will keep some statistical properties whose validity is known to be general, which opens the door for the generation of a vast collection of hypothetical solar radiation sequences

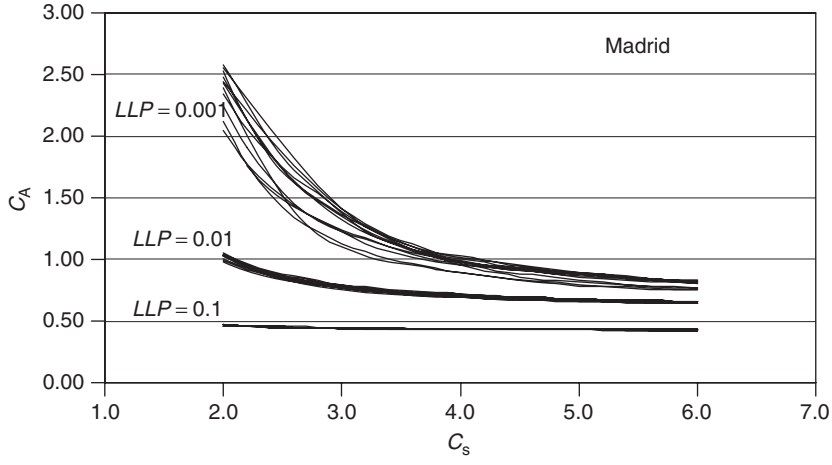


Figure 22.25 Simulated reliability maps, C_A versus C_S, with the historical and with 12 generated solar radiation sequences for 3 different reliability *LLP* values

with the same occurrence probability as the historical one. Then, a different reliability map can be associated with each of these radiation sequences, by means of the simulation exercise already discussed. Obviously, the similarity between the different maps can be understood as a measure of the uncertainty associated with the prediction. Figure 22.25 shows the result of superimposing such maps. The example is for Madrid, generating different solar radiation sequences following Aguiar's method [35]. It is clear that precautions must be taken with predictions for $LLP < 10^{-2}$. For example, for $LLP = 10^{-3}$ and $C_S = 3$, we can find C_A values from 1.1 to 1.5. Other authors [67, 73] have presented similar results. We must conclude that the validity of PV-sizing methodologies is generally restricted to the range $1 > LLP > 10^{-2}$, that is, to solar coverage below 99%. Beyond this limit, sizing results are statistically of doubtful quality, although unfortunately, they are often found in the literature and in simulation software tools marketed today.

It must be stressed that this basic uncertainty cannot be overcome either by reducing the simulation time-step (hourly instead of daily values) or by incorporating more complex models of the elements of the PV systems (nonlinear $I-V$ models of PV generators, battery efficiency dependence with *SOC*, etc.). In fact, the reduction of the considered simulation period can only worsen the situation. It can be shown that the validity of sizing results based only on the *TMY* (avoiding the generation of large-radiation sequences) is restricted to the range $1 > LLP > 10^{-1}$, independent of any other consideration [74]. Appropriately, Marion and Urban, when presenting USA *TMY* s, advise '*Because they represent typical rather than extreme conditions, they are not suited for designing systems to meet the worst-case conditions occurring at a location*'.

On the other hand, such basic uncertainty can help to explain why the result of the different PVsizing methods can be inconsistent; and also why the accuracy gains associated with the consideration of second-order effects when modelling the PV system are likely to be insignificant. In other words, such modelling can be useful for studying some PV system features (optimal number of solar cell per module, optimal charge regulation algorithms, etc.); but it is of little value when considering the pure size problem, for which a judicious and rather simple hypothesis suffices. A similar conclusion is presented in Kaiser and Sauer [75], from a comparison of the long-term energy yield of stand-alone PV systems simulated with both simple and detailed models. The authors observe that the range of uncertainty due to possible fluctuations of the radiation–time series is significantly larger than the those corresponding to changes in the PV system component

modelling, and conclude '*The results make it clear that exact simulation models can provide exact numbers, but do not automatically allow exact predictions*'.

This can also help explain why PV system sizing simply based on guesswork remains widely practiced in current engineering practices. This prevents any quantitative relationships between C_A, C_S and *LLP* being established. The size of the generator is instead chosen to ensure that the energy produced during the design period (most often, the worst month) exceeds the demand of the load by a margin that depends on the designer's experience. A similar procedure is used to size the accumulator. For example, $C_A = 1.1$ and $3 \leq C_S \leq 5$ are common values for rural electrification purposes [76]. $1.2 \leq C_A \leq 1.3$ and $5 \leq C_S \leq 8$ are common ranges on the so-called professional market [67]: telecommunication and so on.

It is worth pointing out that this rather unscientific way of proceeding does not necessarily give bad results, in terms of reliability and cost. As a matter of fact, a proper combination of expertise and common sense often leads to very good results. Today, PV systems have the reputation of being reliable, even in those sectors where high reliability is an established requirement, such as telecommunications and cathodic protection.

It must be mentioned that in the common case of PV arrays directly coupled to batteries, that is without maximum power-tracking devices, energy balance analysis can be done by means of simple ampere balances, initially supposing that the working voltage is always the nominal one V_{NOM} at which it equals the maximum power point voltage of the generator. Therefore,

$$L = V_{NOM} \cdot Q_L \text{ and } \eta_G \cdot A_G = \frac{V_{NOM} \cdot I_M^*}{G^*} \tag{22.86}$$

which leads to

$$C_A = \frac{I_M^* \cdot G_{dm}(\beta, \alpha)}{Q_L \cdot G^*} \text{ and } C_S = \frac{Q_B}{Q_L} \tag{22.87}$$

where Q_L is the amount of charge (expressed in ampere hours) drawn daily by the load, that characterises the required PV array, and Q_B is the useful ampere hour capacity of the battery, which is equal to the product of the nominal capacity by the maximum depth of discharge. This approximation, in spite of appearing oversimplified, gives very good results and simplifies the task of deducing the number of PV modules to be installed. Note the ratio $G_{dm}(\beta,\alpha)/G^*$ can be properly understood as time of full sunlight at $1\,kW/m^2$. This is often referred as 'sun-hours'. Thus, the amount of charge delivered by the PV array is just the current at STC multiplied by the 'sun-hours' of the period considered.

22.12 THE CASE OF SOLAR HOME SYSTEMS

Estimations, uncertain as they may be, [77] suggest that there are about 1.5 million off-grid PV home systems providing lighting, radio and television are currently in operation, totalling $40\,MW_p$, and large rural electrification programmes comprising some thousands of SHSs are increasingly becoming a part of the rural market. Thus, SHS represent the most widespread PV application nowadays (when only the total number but not the total installed power is considered), and the trend is likely to continue.

Standardisation of equipment and its mass production represent efficient ways to obtain low prices and high technical quality. In consequence, SHS designers have to assume a single 'standard' value of energy consumption for a large number of different families. It must be noted that such standardisation is a requirement imposed by the technology itself in order to reduce cost

and guarantee quality, but it does not correspond well with the idea of needs at the individual level. PV history shows some interesting cases [78, 79]. For example, in reference [79] it is stated that '... *an interesting aspect, clearly confirmed by the operating results of ENEL plants, is that the intake power by this type of user ... is rarely the same on any given day, and it is linked to the particular lifestyle of the people involved, to periods of absence and to the number of occupants of the houses being supplied, and so on*'. Apart from PV, other types of rural electrification also provide examples. Figure 22.26 shows the distribution of the individual monthly electricity consumption measured during 4 years in the 63 dwellings of Iferd, a Moroccan village where a small diesel generator set provides 3 h of electricity per day (consumers are metered and pay for their energy use). The large observed spread leads one to question the real meaning of reliability parameters such as *LLP*.

It appears that 'standard' *LLP* values derived from sizing methodologies are scarcely representative of the realities in the field. The relationship between reliability and load, that is, the function $LLP = LLP(L)$ for a given PV system, can be explored just by extending the previously described simulation procedure to a large number of cases. A certain baseline case has been, first, established by fixing the PV array power and the battery capacity values, C_A and C_S, for a given load, L_{BASE}, and a given reliability, $LLP = 0.1$. Then, the load has been varied from $0.8L_{BASE}$ to $1.2L_{BASE}$ and the corresponding reliability has been calculated. Figure 22.27 shows the result. Roughly speaking, we can say that an approximately logarithmic relationship exists in such a way that *LLP* decreases one order of magnitude for each 30% of load reduction. This result, together with the observation that real L values are generally found within the range -50% to $+100\%$ of the mean, let us conclude that real individual *LLP* values can vary more than three orders of magnitude (for example, from 10^{-1} to 10^{-4}) in the context of the same SHS project. This nullifies any attempt at finding a single representative *LLP* value. It is worth mentioning that the same is not true when centralised electricity generators are considered (PV or not), because the total energy consumed by all the families involved shows a much lower standard deviation than that corresponding to the individual consumptions (roughly, the standard deviation becomes reduced by a factor of N, the number of families), so that it is possible to find single L and *LLP* representative values for the whole population served.

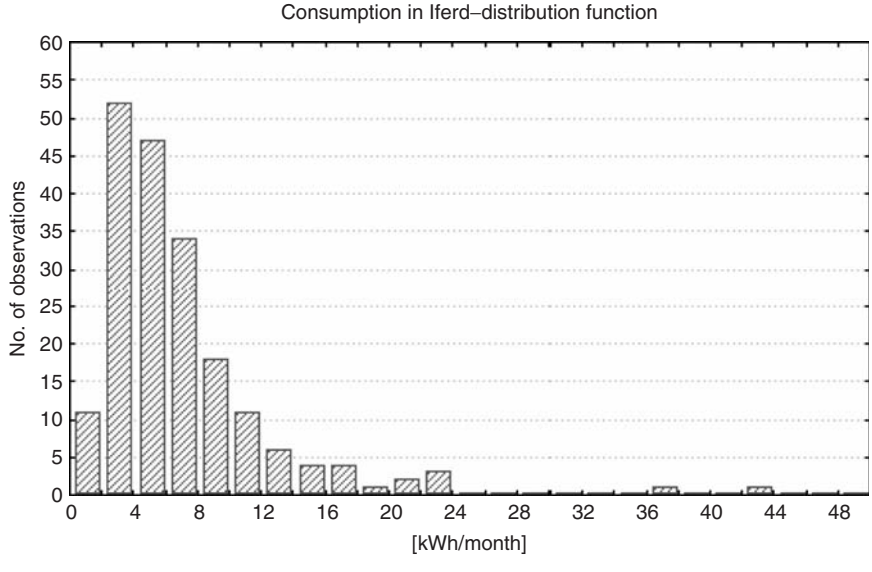


Figure 22.26 Distribution function of monthly electricity consumption in all the Iferd dwellings

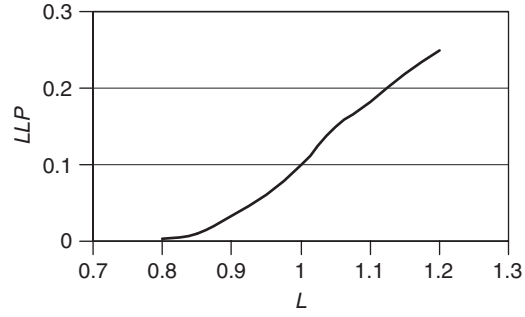


Figure 22.27 Reliability *LLP* as a dependent variable on the consumption *L* for a given PV array and capacity values

However, even in extremely varying applications, such as SHS, PV sizing methods based on reliability can be of great help if large-scale programmes become a future reality. This will probably require the development of rigorous engineering: standardisation of different levels of service, technical quality controls and so on. For example, PV sizing methods based on *LLP* represent an interesting possibility of comparing different alternatives (different offers from various manufactures) on an objective basis, as the *LLP* value respectively associated with each alternative, for the same considered energy service [74].

It is worth considering the question: 'How much electricity has to be provided to a rural house in a developing country to be socially and economically acceptable?'. Although this question is always at the origin of any PV rural electrification programme, its answer in terms of watt hour/day, is far from being clear. Energy consumption data, based on practical experience in developing countries, are scarce in the literature [79], which is paradoxical, considering that many thousands of SHS are currently operating in developing countries. Instead, there are a great number of consumption scenarios where, although starting from very different hypothesis concerning the number of appliances and the length of time they are in use, the SHSs finally selected have an installed power of about 40–50 W_p. This is because past in-field experience has shown PV designers that such systems are generally well accepted by the rural users, while the same is not always the case when small (20–30 W_p) PV modules are concerned. In this way, the SHS scenarios elaborated by PV designers must therefore be interpreted as explanation exercises, rather than as designs for systems starting from an evaluation of actual needs (see Chapter 23 for discussion of rural electrification programmes). So, we must conclude that energy scenarios for rural electrification purposes are still an open question, which need to be explored in depth.

22.13 ENERGY YIELD OF GRID-CONNECTED PV SYSTEMS

The output of grid-connected PV systems is the output from the PV array less the losses in the inverter. The output from the PV array has been considered in detail in this chapter, and the performance of inverters is described in Chapter 19. As far as inverters are concerned, it is important to account for the fact that the instantaneous efficiency η_y depends on the ratio between the actual power delivered to the grid P_{AC}, and the rated power of the inverter, P_{IMAX}. This dependence may be represented by [80]

$$\eta_i = \frac{p}{p + k_0 + k_{i1}p + k_{i2}p^2} \tag{22.88}$$

where $p = P_{AC}/P_{IMAX}$ and k_0, k_{i1} and k_{i2} are parameters characteristic of the inverter defining its electrical behaviour. k_0 is the quiescent power consumption, k_{i1} represents the losses that depend linearly on the current (voltage drop across diodes, etc.) and k_{i2} represents the losses that depend on the square of the current (resistive losses, etc.). These parameters can be obtained from the inverter efficiency curve. Depending on quality level, input voltage and rated power, the loss parameters of existing inverters have a spread of more than a factor of 10. As an example, $k_0 = 0.35$, $k_{i1} = 0.5$ and $k_{i2} = 1$ correspond to very good inverters leading to 95% typical energy efficiency.

Standard methods for performance analysis of PV grid-connected plants have been introduced in the JRC Ispra Guidelines [81], extended and improved by HTA Burgdorf [82]. Global performance is appropriately described by the so-called *performance ratio* (*PR*), which is the ratio of AC energy delivered to the grid, E_{AC}, to the energy production of an ideal, lossless PV plant with 25 °C cell temperature and the same solar irradiation. This gives a good indication of how much of the ideally available PV energy has actually been generated. It is given by

$$PR = \frac{E_{AC}}{\frac{G_y(\beta,\alpha)}{G^*} \cdot P_M^*} \tag{22.89}$$

Other interesting parameters are the *reference yield*, $Y_r = G_y(\beta,\alpha)/G^*$, the *array yield*, $Y_a = E_{DC}/P_M^*$, where E_{DC} is the DC energy generated by the PV array, and the *final yield*, $Y_f = E_{AC}/P_M^*$. All three have units of time, and allow us to distinguish between the losses due to the PV array, and the losses associated with the inverter and to the operation of the system. Capture losses, $L_C = Y_r - Y_a$, are defined as the energy losses, expressed in hours per day of PV array operation at STC power output, caused by: cell temperatures higher than 25 °C, losses in wiring and protection diodes, poor module performance at low irradiance, partial shading, snow and ice coverage, module mismatch, operation of the array at a voltage other than its maximum power point, and spectral and angular losses. *System losses*, $L_S = Y_a - Y_f$, are the losses due to inverter inefficiencies. It must be noted that $PR = Y_f/Y_r$. Again, it must be noted that Y_r can be understood as equivalent to 'sun-hours' of full sunlight at 1 kW/m$^{2.}$ This way, the energy yield of the PV array is just given by the product of these 'sun-hours' by the *PR*.

Energy losses in very good PV grid-connected systems without significant shading are about $L_C = 15\%$ and $L_S = 7\%$, which lead to $PR \approx 0.78$. A value of $PR = 0.75$ is sometimes [13] recommended for quick estimation of yearly energy production.

Example: Estimate the energy yield of a grid connected PV plant at Albuquerque, USA, having a fixed position and optimally tilted generator. The solution is:

Equation (22.89) $\rightarrow E_{AC}/P_M{}^* = PR.G_Y(\beta_{opt})/G^*$, where $G_Y(\beta_{opt})$ can be obtained by the product 365 × [column 3 × column 4 × column 8]. In this way:

$G_Y(\beta_{opt}) = 2441$ kWh/m^2 (or 2441 sun-hours); $PR = 0.75 \rightarrow E_{AC}/P_M{}^* = 1831$ kW h/kW. It must be mentioned that this place, with G_y (0) = 2104 kW h/m^2 [365 × column 3 × column 4] is extremely fortunate, regarding solar energy availability. Because of that, tracking energy gains are very large. For example, a two-axis tracking free of shading collects up to 47% more radiation that the previously considered fix and optimally tilted surface.

Reported experimental *PR* values [83–86] range from 0.65 to 0.8. The main reason for *PR* reduction is that the actual power of installed PV arrays is sometimes below the rated power declared by the manufactures. As a representative case, Figure 22.28 describes the actual power measured by the IES-UPM at big grid-connected plants installed in Spain. The figure presents the histogram of corresponding peak power deviations (actual versus rated). Regrettably, actual power even below 90% of the rated values is still found at real installations.

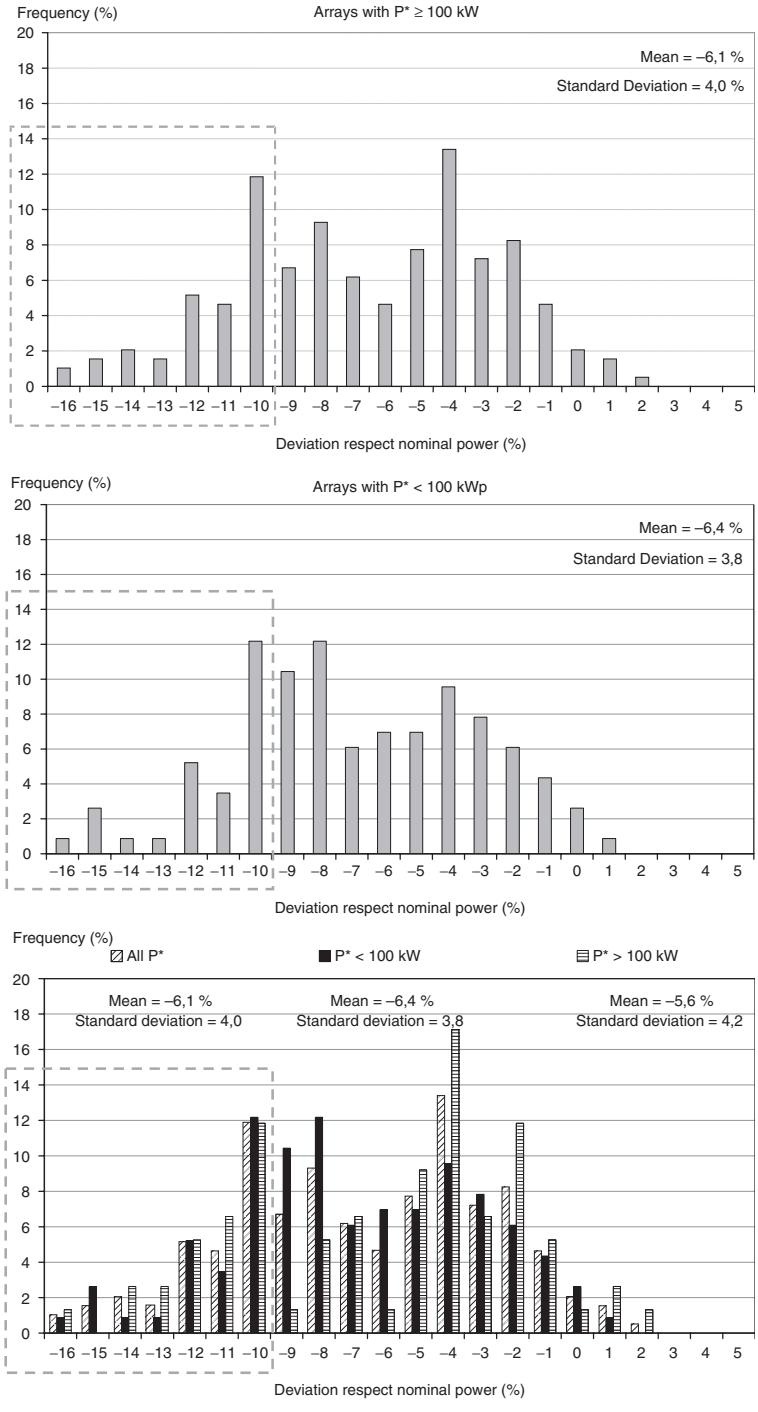


Figure 22.28 Real versus rated power at big PV plants installed in Spain between 2007 and 2009. Total peak power is about 200 MW

Table 22.8 Power distribution of the yearly irradiation for a low and a high latitude location

Location	$G_y(\beta_{opt})$	Percent of irradiation in different ranges of irradiance in [W/m^2]				
	[kW h/m^2]	<200	200–500	500–800	>800	Total
Jaen	2040	5.8	23.6	44.7	25.9	100
Copenhagen	1190	13.9	30.7	35.7	19.7	100

Equation (22.89) is given in yearly terms. However, the *PR* is a rather general concept, so that *PR* values can also be obtained for any other period. It is worth noting that, mainly because dependence on solar cell temperature and shading, *PR* is far of being an invariant. Hence, to measure the *PR* during a certain period of time, for example, a month or a week, and, then, to apply the corresponding value to a longer period, for example, a year, is simply wrong.

22.13.1 Irradiance Distributions and Inverter Size

The power distributions of solar irradiation are different for varied geographical latitudes. In places of high latitude, with very cloudy weather, the solar irradiation is almost evenly distributed over a wide range of power scale; while in low latitude places, with predominantly clear sky, the higher power range is enhanced. Table 22.8 presents the distribution in different irradiance classes of the total yearly irradiation over an optimally tilted surface, as obtained from the *TMY* of Copenhagen [87] ($\phi = 55.7^\circ$) and Jaen-Spain [88] ($\phi = 37.8^\circ$). Surprisingly, the energy content at low irradiances ($G < 200 W/m^2$) is relatively low in both places. This may appear counter-intuitive, but it is easily understood when considering the difference between time and energy distribution. For example, in Copenhagen, the low irradiance (< 200W) accounts for only 13.9% of the total annual irradiation, despite it occurring during 2461 h/year, which represents 55% of the total daily time. That leads one to question the idea, sometimes defended in PV literature [89], that PV module performance at low irradiances is very relevant for cloudy climates. As a matter of fact, empirical evidence that efficiency at low light levels is scarcely relevant is found in the literature [90].

However, because the most commonly occurring irradiation corresponds to medium irradiances, an energetic advantage can be obtained by selecting the inverter size smaller than the PV generator peak power, that is, $P_{IMAX} < P^*_M$. The corresponding reduction of relative inverter self-consumption and losses may compensate the possible energy loss by an inverter power limit lower than the maximum PV output power. Recommended values of P_{IMAX}/P^*_M range from 0.6 (high latitudes) to 0.8 (low latitudes) [91, 92].

22.14 CONCLUSIONS

Methods to estimate all the solar radiation components incident on any arbitrarily oriented surface and at any time of the year have been presented. The only required input information is the 12 monthly mean values of the daily global horizontal irradiation. Several sources of solar radiation data are available. The corresponding values, for the same location and the same month, can significantly differ from one source to another. The corresponding uncertainty does not derive primarily from a lack of precision in the measuring instruments, but from the random nature of the solar radiation. The intrinsic uncertainty of solar radiation represents an important limit to the significance of the results of any PV design exercise, irrespective of the complexity of the model supporting the particular design tool. Because of this, rather simple design methodologies can yield

results of similar confidence to those from sophisticated ones. The annual solar radiation availability on horizontal surfaces, optimally tilted fixed surfaces and several types of tracking surfaces has been calculated for 32 different places distributed all over the world. The hope is that the readers could find here a similar location, both in latitude and clearness index, to the location of their interest.

On the other hand, the PV module's conversion efficiency of solar radiation into electricity has also been considered. A model for predicting the $I-V$ characteristic of a PV generator at any prevailing condition such as solar irradiance, ambient temperature, wind speed and so on, have been described. The only required input data are the short-circuit current, the open- circuit voltage and the maximum power under the STC, and the nominaloperating cell temperature. All these data are commonly found in the manufacturer's standard information. Particular attention has been paid to the consideration of the dust and angle of incidence effects.

Some relevant aspects related to the design of stand-alone PV applications have been disclosed. Whatever the detailed methodology, stand-alone PV system sizing relies on future prediction (the expected system lifetime) based on past observations of the solar radiation data. The corresponding uncertainty, unavoidably associated with the random nature of the solar radiation, implies that simulation models can provide exact numbers, but not automatically exact predictions. Concerning SHS, by far the most widespread PV application nowadays (in terms of the number of systems currently in operation), the difficulties for deriving representative standard energy consumption values have been pointed out. Finally, energy performance ratios for grid-connected PV systems have been defined.

ACKNOWLEDGEMENTS

Montse Rodrigo has been extremely kind in preparing all the figures. The comments of the editors have been extremely valuable. The writing of this chapter has been at the expense of many hours away from my loved ones. I must thank the large tolerance and patience of Cristina, Leda and Celena.

REFERENCES

1. Serres M *et al*., *Éléments d'Histoire des Sciences*, Chap. 3, Ed Bordas, Paris, 77–117 (1989).
2. Blanco M, Alarcón D, López T, Lara M, *Sol. Energy* **70**, 431–441 (2001).
3. Kasten F, *Arch. Meteorol. Geophys. Bioklimatol*. **14**, 206–223 (1966).
4. Meinel A, Mainel M, *Applied Solar Energy, An Introduction*, Addison-Wesley, Reading, MA (1976).
5. Iqbal M, *An Introduction to Solar Radiation*, Academic Press, Ontario (1983).
6. International H-World Database, *Mean Values of Solar Irradiation on Horizontal Surface*, Ed Progensa, Sevilla, Spain (1993).
7. Meteonorm. Versión 6.0, www.meteotest.ch.
8. Palz W, Greif J, *European Solar Radiation Atlas*, Commission of the European Communities, Ed Springer, Germany (1996.)
9. Font Tullot I, *Atlas de la Radiación Solar en España*, Ed Instituto Nacional de Meteorología, Madrid, Spain (1984).
10. Capderou M, *Atlas Solaire de l'Algerie*, Ed EPAU, Alger (1985).
11. Colle S, Pereira E, *Atlas de Irradiaçao Solar do Brasil*, INM, Labsolar EMC-UFSC (1998).
12. NASA Surface Meteorology and Solar Energy Data Set. Available at *eosweb.larc.nasa.gov/sse*.
13. PVGIS, available at http://re.jrc.ec.europa.eu/pvgis/apps/radmonth.php.
14. Macagnan M, Lorenzo E, Jimenez C, *Int. J. Sol. Energy* **16**, 1–14 (1994).

15. Liu B, Jordan R, *Sol. Energy* **4**, 1–19 (1960).
16. Page J, *Proc. U.N. Conf. New Sources Energy*, 378–390 (1961).
17. Erbs D. G, Klein S. A, Duffie J. A, *Sol. Energy* **28**, 293–302 (1982).
18. Collares-Pereira M, Rabl A, *Sol. Energy* **22**, 155–164 (1979).
19. Huld T, Suri M and Dunlop E, *Prog. Photovolt*. **16**, 595–607 (2008).
20. Whenham, Green, Watt and Corkish, *Applied Photovoltaics*, 2nd edn, Chap. 1, pp 20–21, Ed Earthscan, London (2007).
21. Rigollier C, Baner O and Wald L, *Sol. Energy* **68**, 33–48 (2000).
22. Hopkinson R, *J. Opt. Soc. Am*. **44**, 455–459 (1954).
23. McArthur L, Hay J, *J. Appl. Meteorol*. **20**, 421–429 (1981).
24. Rossini E. G. and Krenzinger A, *Sol. Energy* **81**, 1323–1332 (2007)
25. Kondratyev K, *Radiation in the Atmosphere*, Academic Press, New York, NY (1969).
26. Hay J, McKay D, *Int. J. Sol. Energy* **3**, 203–240 (1985).
27. Perez R *et al*., *Sol. Energy* **36**, 481–497 (1986).
28. Perez R *et al*., *Sol. Energy* **39**, 221–231 (1987).
29. Siala F, Rosen M, Hooper F, *J. Sol. Energy Eng*. **112**, 102–109 (1990).
30. Standard ASHRAE 93-77, *Methods of Testing to Determine the Thermal Performance of Solar Collectors*, American Society of Heating, Refrigeration, and Air Conditioning Engineers, New York (1977).
31. Martin N, Ruiz J, *Sol. Energy Mater. Sol. Cells* **70**, 25–38 (2001).
32. Bottenberg W, Module Performance Ratings: Tutorial on History and Industry Needs, *PV Performance, Reliability and Standards Workshop*, pp 5–42, NREL, Vail, CO (1999).
33. Amato U *et al*., *Sol. Energy* **37**, 179–194 (1986).
34. Graham V *et al*., *Sol. Energy* **40**, 83–92 (1988).
35. Aguiar R, Collares-Pereira M, Conde J, *Sol. Energy* **40**, 269–279 (1988).
36. S. Wilcox and W. Marion. Users Manual for TMY3 Data Sets. *NREL Technical Report TP-581-43156*, 2008.
37. Benseman R, Cook F, *N Z J. Science* **12**, 296–708 (1960).
38. Lorenzo E, Maquedano C, *Proc. 13the Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2433–2436 (1995).
39. Quaschning V, Hanitsch R, *Proc. 13th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 683–686 (1995).
40. Skiba M *et al*., *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, 2402–2405 (2000).
41. Narvarte L, Lorenzo E, *Prog. Photovolt*. **16**, 703–714 (2008).
42. Castro M *et al*., *Era Sol*. **87**, 5–17 (1998).
43. Kaiser R, Reise C, *PV System Simulation Programmes*, Internal Report for the IEA SHCP Task 16, Fraunhofer-FISE, Freiburg (1996).
44. Caamaño E, Lorenzo E, *Prog. Photovolt*. **4**, 295–305 (1996).
45. Jennings C, Farmer B, Townsend T, Hutchinson P, Gough J, Shipman D, *Proc. 25th IEEE Photovoltaic Specialist Conf*., pp 1513–1516 (1996).
46. Butti K, Perlin J, *A Golden Thread: 2500 Years of Solar Architecture and Technology*, Ed Cheshire, Palo Alto, NY (1980).
47. de Laquil III P., Kearney D, Geter M, Diver R, in Johansson T, Kelly H, Reddy A, Willians R, (eds), *Renewable Energy: Sources for Fuels and Electricity*, Chap. 5, Island Press, Washington, DC, 213–236 (1993).
48. Eikelboom J, Jansen M, *Characterisation of PV Modules of New Generations*, ECN-C-00-067, Report to NOVEM Contract 146.230-016.1, Available at *www.ecn.com* (2000).
49. Caamaño E, Lorenzo E, Zilles R, *Prog. Photovolt*. **7**, 137–149 (1999).
50. *Photon Int*, **3**, 140–174 (2008).
51. Green M, *Solar Cells*, Chap. 5, Prentice Hall, Kensington, 95–98 (1982).

52. Araujo G, Sanchéz E, *Sol. Cells* **5**, 377–386 (1982).
53. King D, Kratochvil J, Boyson W, *Proc. 26th IEEE Photovoltaic Specialist Conf*., pp 1113–1116 (1997).
54. Fuentes M, *Proc. 17thEEE Photovoltaic Specialist Conf*., 1pp 341–1346 (1984).
55. Chianese D, Cerenghetti N, Rezzonico R, Travaglini G, *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2418–2421 (2000).
56. Scheiman D, Jenkins P, Brinker D, Appelbaum J, *Prog. Photovolt*. **4**, 117–127 (1996).
57. Smiley E, Stamenic L, Jones J, Stojanovic M, *Proc. 16th Euro. Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2002–2004 (2000).
58. Martin N, Ruiz J, *Prog. Photovolt*. **7**, 299–310 (1999).
59. Jones A, Underwood C, *Sol. Energy* **70**, 349–359 (2001).
60. King L, Kratochvil J, Boyson E, Bower W, *Proc. 2nd World Conf. Photovoltaic Solar Energy Conversion*, pp 1947–1952 (1998).
61. Gordon J, *Sol. Cells* **20**, 295–313 (1987).
62. Cowan W, *Proc. 12th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 403–407 (1994).
63. Abouzahr Y, Ramakumar R, *IEEE Trans. Energy Conversion* **5**(3), 445–452 (1990).
64. Negro E, *Proc. 13th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 687–690 (1995).
65. Kaiser R *Photovoltaic Systems*, Fraunhofer-FISE, Freiburg, Germany, (1995).
66. Chapman R, *Sol. Energy* **43**, 71–76 (1989).
67. Macomber H, Ruzek I, Costello F, *Photovoltaic Stand-Alone Systems: Preliminary Engineering Design Handbook*, Prepared for NASA, Contract DEN 3-195 (1981).
68. Barra L, Catalanotti S, Fontana F, Lavorante F, *Sol. Energy* **33**, 509–514 (1984).
69. Bucciarelli L, *Sol. Energy* **32**, 205–209 (1984).
70. Bartoli B *et al*., *Appl. Energy* **18**, 37–47 (1984).
71. Sidrach-de-Cardona M, Mora Ll, *Sol. Energy Mater. Sol. Cells* **55**, 199–214 (1998).
72. Egido M, Lorenzo E, *Sol. Energy Mater. Sol. Cells* **26**, 51–69 (1992).
73. Klein S, Beckman W, *Sol. Energy* **39**, 499–512 (1987).
74. Lorenzo E, Narvarte L, *Prog. Photovolt*, **8**, 391–409 (2000).
75. Kaiser R, Sauer D, *Proc. 12th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 457–460 (1994).
76. Universal Technical Standard for Solar Home Systems, Thermie B SUP 995-96, EC-DGXVII (1998).
77. Nieuwenhout F *et al*., *Prog. Photovolt*. **9**, 455–474 (2001).
78. Belli G, Iliceto A, Previ A, *Proc. 11th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1571–1574 (1992).
79. Morante F, Zilles R, *Prog. Photovolt*. **9**, 379–388 (2001).
80. Jantsch M, Schmidt H, Schmid J, *Proc. 11th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1589–1593 (1992).
81. Guidelines for the Assessment of PV Plants. Document B: Analysis and Presentation of Monitoring Data, Issue 4.1, JRC of the Commission of the European Communities, Ispra, Italy (1993).
82. Haeberlin H, Beutler C, *Proc. 13th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 934–937 (1995).
83. Jahn U, Nasse W, *Prog. Photovolt*. **12**, 441–448 (2004).
84. Alonso M, Chenlo F, Vela N, Chamberlain J, Arroyo R, Alonso F. J., *Proc. 20th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2454–2457 (2005).
85. Moore L. M., Post H. N. *Prog. Photovolt*. **16**, 249–259 (2008).
86. Garcia M, Vera J. A., Marroyo L, Lorenzo E, Pérez M, *Prog. Photovolt*. Available on-line (2009).
87. Katic Y, Jensen B, *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2830–2833 (2000).

88. Nofuentes G, *Contribución al desarrollo de aplicaciones fotovoltaicas en edificios*, Ph.D. thesis, Presented at the Polytechnical University of Madrid, Madrid (2001).
89. Mason N, Bruton T, Heasman K, *Proc. 14the Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2021–2024 (1997).
90. Wilk H, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 297–300 (1997).
91. Macagnan M, Lorenzo E, *Proc. 11th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1167–1170 (1992).
92. Macedo W. N., Zilles R, *Prog. Photovolt*. **15**, 337–352 (2007).

23
PV in Architecture

Tjerk H. Reijenga[1] and Henk F. Kaan[2]

[1]*BEAR Architecten, Gouda, the Netherlands*
[2]*ECN Energy Research Centre of the Netherlands, Petten, the Netherlands*

23.1 INTRODUCTION

23.1.1 Photovoltaics (PV) as a Challenge for Architects and Engineers

We are witnessing an essential change in thinking about renewable energy. The world is confronted with climate change due to the burning of fossil fuels. On top of that, Western countries want to become more independent from oil and gas delivered by politically not very stable areas. As a consequence, governments are spending hundreds of millions of dollars on research, development and the demonstration of renewable energy. Current developments show that renewables, such as solar energy systems, are incorporated more and more into our daily life, as conventional energy sources become depleted and environmental concerns grow [1].

Solar systems are becoming an integral part of our society and thus our environment. In various Western countries, but also in Japan, examples can be seen of large quantities of photovoltaics, incorporated in urban areas, as in Freiburg (Germany) and Heerhugowaard (the Netherlands). There are large incentives for urban planners and architects to incorporate photovoltaics into their design. Some countries such as France will give extra incentives for building integration of PV systems. New products are emerging, yet need further development to fully meet the architectural needs of sustainable buildings. Architects therefore need to start thinking about this new smart solar architecture.

The government's role in promoting and supporting sustainable energy (PV systems) strongly influences the extent to which these systems are used in buildings. The high interest in PV in Germany for instance is a result of the policy of the German government with regard

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

to PV and renewable energy in general. In countries with less government intervention, the power utilities play a bigger role. Even without financial support, the government can encourage sustainable energy, for example, by demanding better performance for buildings. By introducing certain performance goals, such as the Dutch building code, sustainable energy and solar energy PV systems might be considered.

There is still a growing interest in "green" products such as organic food, organic fibers as well as green buildings. Insurance companies and financial markets are becoming aware of "green" financing, which requires a different design approach from architects. "Green" design is the basic reason for integrating PV systems into buildings.

A large part of the future PV market will be associated with building applications, especially in Europe and Japan where the population density is high and the land is valuable [2]. The scale of building integration is increasing and goes up to over 11 MW_p for a building (mainly exhibition or factory halls) (Figure 23.1). In areas with less population, it will be possible to find land for ground-mounted PV structures (Figure 23.2) [3].

Figure 23.1 A 300 kWp roof-mounted system on a sporting hall in Wageningen (The Netherlands). Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.2 The 50 MWp ground-mounted system in Puertollano, Spain. Reproduced with permission by Prof. A. Luque IES

Figure 23.3 Building integration with modules between the tiles (physical integration) in Utrecht, The Netherlands. Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.4 Building integration with modules above the tiles (aesthetical integration) in Utrecht, The Netherlands. Reproduced with permission by BEAR Architecten T. Reijenga

23.1.2 Definition of Building Integration

A definition of building integration is hard to formulate, as it concerns the physical integration of a PV system into a building, but it also covers the overall image of the PV system in the building. For the architect, the aesthetic aspect, rather than the physical integration, is the main reason for talking about building integration. (Figures 23.3 and 23.4) The optimal situation is a physically and aesthetically well-integrated BIPV system. In fact, many examples of physical integration show a lack of aesthetic integration. Visual analysis of PV systems in buildings shows that the look of a poorly designed building does not improve, simply by adding a well-designed PV system. On the other hand, a well-designed building with a nicely integrated PV system will be accepted by everybody.

Building-integrated, grid-connected PV systems have the following advantages [4]:

- There is no additional requirement for land.
- The cost of the PV wall, and up to a certain point the roof, can be offset against the cost of the building element it replaces.
- Power is generated on site and replaces electricity that would otherwise be purchased at commercial rates.

- Connecting to the grid means that the high cost of storage is avoided and security of supply is ensured.

In addition, architecturally elegant, well-integrated systems will increase market acceptance, and building-integrated PV (BIPV) systems provide building owners with a highly visible public expression of their environmental commitment (Figure 23.4) [5].

The way people deal with photovoltaics in architecture differs from country to country. This depends on the scale, culture and type of financing for building projects. In countries such as Denmark, the Netherlands and the United Kingdom, where public housing is very common, serial production is strongly emphasized in housing projects. Professionals such as housing associations, project developers and architects implement the housing construction process, in which the main opportunities are for PV roof integration in single-family terraced houses and for façade and roof integration in apartment buildings.

Of course there are also many countries where the government or professional, non-profit housing associations have little influence on house building. In those countries, the process of developing and building houses is mainly a private initiative. Integration of PV systems in buildings can be carried out by professionals but, on the smaller scale of a single-family house, the motivation must mainly come from the private owner. In general, most building-integrated PV systems are found in buildings in which building professionals are involved. Consequently, in countries where there is little professional involvement in large-scale housing projects, PV can be found in the first place on commercial and industrial buildings [6]. With these types of buildings, PV systems are integrated both into façades and roofs. In addition, there is a significant market for private homeowners who buy small-scale (less than 500 W_p) PV systems and mount them somewhere on their house.

The aim of integrating PV systems into buildings is to reduce the requirement for land and the costs [7]. This could be the cost of the support construction and the cost of building elements, such as cladding elements. It is more efficient to integrate a PV system when constructing the building, rather than mounting it afterwards.

23.2 PV IN ARCHITECTURE

The following section aims to explain some basic thoughts about PV to non-designers, from an architectural and design point of view.

Note: All specified power of PV systems is the power under standard test conditions and tilt which may be greater than the power delivered when installed in non-optimum orientation required by the BIPV application.

23.2.1 Architectural Functions of PV Modules

For architects, the application of PV systems must be part of a holistic approach. A high-quality PV system can provide a substantial part of the building's energy needs if the building has been designed in the right way. In general, the energy consumption of buildings needs to be cut down by at least 50% compared with a typical, but inefficiently designed building.

In a holistic approach, integrating a PV system not only means replacing a building material, thus physically integrating the PV system, but also aesthetically integrating it into the design. The integration also takes over other functions of the building's skin. Mounted on a sloped roof for instance, profile systems mean that PV modules can be part of the watertight skin.

A distinction can be made between literal integration of PV in the building skin (PV as a cladding element or integrated into the roof) and integration of PV in building components (awnings, shading devices etc.).

23.2.1.1 Roof integrated PV

A PV system can be integrated into the roof in several ways. One choice is for the integrated system to be part of the external skin and therefore part of an impermeable layer in the construction. In the early days of BIPV (1990s), several building projects were constructed on the basis of this principle (Figure 23.5) [8]. As this is a solution which is debatable from the technical point of view, the system can also be mounted above an impermeable roof foil, thereby protecting the foil against UV light and direct sun. This extends the lifespan of the foil. Although this is a more secure option, it is also not without some risk, as the impermeable layer has to be pierced in order to mount the system on the roof. This kind of system is also available for flat roofs. The Powerlight Company from Berkeley, CA (USA) introduced a PV system into the market that is glued onto expanded polystyrene (XPS) insulation material. This type of warm roof construction (construction on the warm side of the insulation) system is very well suited to renovating large flat roofs. Using PV modules as roof covering reduces the amount of building materials needed, which is very favorable for a sustainable building and can help to reduce costs.

In addition to covering the complete roof with modules, there are also many products for small-scale use, for example, PV shingles and tiles. The small scale of these products (from 2 cells on a tile to around 20 cells on a look-alike tile) makes them very convenient for use in existing buildings or as do-it-yourself products.

Transparent PV modules used as roofing materials serve as water and sun barriers and also transmit daylight (Figure 23.6). In glass-covered areas, such as sunrooms and atriums, sun protection on the roof is necessary in order to avoid overheating in summer. The PV cells absorb 70–80% of the sun radiation. The space between the cells transmits enough diffused daylight to achieve a pleasant lighting level in the area. PV cells have been used in this way in numerous projects, for instance at the Centre for Sustainability De Kleine Aarde in Boxtel (NL) [10] and the

Figure 23.5 Roof integration in a renovation project with the Shell Solar/BOAL profiles in Leiden (The Netherlands), providing a 2.1 kW_p system per house. The PV roof is an impermeable layer. Reproduced from Maycock P *et al*., *Building with Photovoltaics*, pp 78–81, Ten Hagen & Stam, Den Haag (1995) with permission by NOVEM, R Schropp [9]

Figure 23.6 Solar office in Doxford Sunderland (UK) with a transparent double-glazing PV system integrated in the façade. The 73 kW_p PV system is manufactured by Saint Gobain and mounted in Schüco window frames. Reproduced with permission by Dennis Gilbert

Brundtland Centre in Toftlund (DK). Using transparent PV modules in the Solar Office in Doxford (UK), has resulted in a similar contrast in the façade [11] (Figure 23.6). In order to increase the usage of daylight in the workplaces, transparent PV modules have been used instead of glass.

Of course, PV cells convert sunlight into electricity (with typical efficiencies of 6–18%) with the remainder of the solar energy being converted into heat. At the project "Haus der Zukunft" in Linz (Austria), this residual heat is also used to warm the home [12]. An air cavity has been created underneath the PV modules, through which warm air (heated by PV modules) is exhausted. The hybrid collector provides warm air to the heating system in the home, which in this case, makes it a cost-effective use of the collector.

A relatively new application of PV combined with a thermal system is PVT: a PV module mounted on a solar thermal module. The residual heat is used to heat the water (or other liquid) in the thermal system. A demonstration project, which was financially supported by the European Commission, can be seen at the head offices of RES UK in Kings Langley, north of London, UK (Figure 23.7).

At the Energy Research Centre of the Netherlands (ECN) in Petten (NL), Building 42 has a conservatory with 43 kW_p BP solar roof-integrated transparent laminates that reduce light

Figure 23.7 PVT modules in the demonstration project at RES UK in Kings Langley, Great Britain. Reproduced with permission by H. F. Kaan

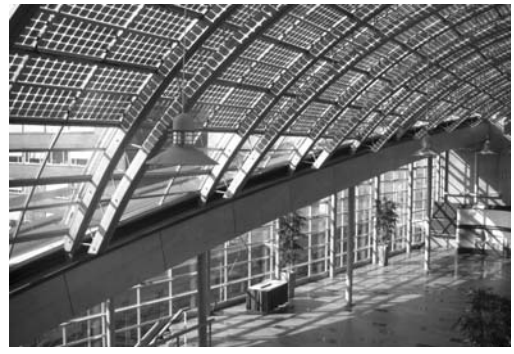


Figure 23.8 Interior view of the conservatory with integrated 43 kW_p transparent BP solar modules in ECN Building 42 in Petten (The Netherlands). Reproduced with permission by BEAR Architecten T. Reijenga. See Plate 8 for the colour figure

and sun transmission by around 70% compared with glass. The conservatory therefore acts as a big parasol over the offices, protecting them from the sun while still providing enough daylight (Figure 23.8) [13].

23.2.1.2 Façade-integrated PV

Façades are basically constructed using *in situ* bricklaying or concrete constructions, prefabricated elements or structural metal façades that are mounted in place. Concrete constructions form the structural layer and are covered with insulation and a protective cladding [14]. This cladding can be wood, metal sheets, panels, glass or PV modules. For luxury office buildings, which often have expensive cladding, cladding with PV modules is no more expensive than other commonly used materials, for example, natural stone and expensive special glass. This cladding costs around \$1500/m^2, which is considerably more than the cost of the PV module today (Figure 23.9).

Structural glazing or structural façades are constructed using highly developed profile systems, which can be filled with all types of sheeting, such as glass or frameless PV modules.

The development of transparent modules has gone further in the last 10 years. In the semi-transparent modules from the 1990s, the space between cells and the light transmission through the

Figure 23.9 Façade integrated PV at the University of Technology in Lisbon (Portugal). Reproduced with permission by H.F. Kaan

Figure 23.10 Suntech Power Headquarter in Wuxi, China has a large (2,552 semi-transparent Light-Thru™ modules) façade with an installed capacity of 1 MW_p. Reproduced with permission of Suntech Power Holdings Co., Ltd

Tedlar back foil stipulated the amount of light that came through. Starting with the Sunways cells, a complete new generation of light-through cells have been developed such as Schott ASI Glass and Suntech MSK's light-thru and see-thru cells that can be used in roofs and structural façades (Figure 23.10).

23.2.1.3 PV in building components

Façades are very suitable for all types of sunshading devices, louvers and canopies [15]. There is a logical combination between shading a building in summer and producing electricity at the same time. Architects recognize this and many examples of PV shading systems can be seen around the world (Figure 23.11). A terrace with a roof on the sunny side of a building is a good place for BIPV systems (Figure 23.12) thus providing shade, protection from rain, as well as electricity.

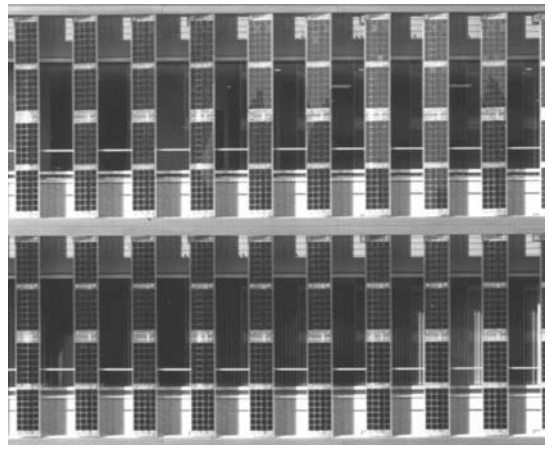


Figure 23.11 Solar shading (PV modules as part of a vertical louver system) on the west façade of the SBIC office building in Tokyo, Japan. The transparent vertical louvers, with a total capacity of 20.1 kW_p, were manufactured by Atlantis Switzerland. Reproduced with permission by Jiro Ono

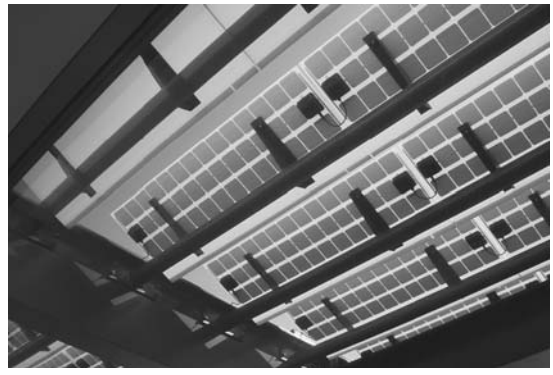


Figure 23.12 This system is a terrace roof of the new airport terminal in Zürich, Switzerland. The roof protects the terrace against direct sun but daylight is still allowed through via the transparent modules. Reproduced with permission by BEAR Architecten T. Reijenga

23.2.1.4 PV art in structures

Free-standing applications of PV power systems have been constructed in a variety of more or less creative designs. Well known are solar sails which are landmarks for companies and show their green involvement. Many other types of free-standing construction designed to support PV power systems have been constructed. Remarkable are the constructions that show PV as a flower and in some cases track the sun.

On the small scale there are many artistic modules both in cell color and pattern or in module form. Artist like Jürgen Claus (Germany) and Sarah Hall (Canada) became famous as PV artist, but many architects also use PV systems in an artistic way (Figure 23.13). The façade of

Figure 23.13 Natural ventilation tower of the underground library of Regents College, UBC, Vancouver, Canada. Reproduced with permission of Studio Sarah Hall [16]

Figure 23.14 Giant energy autonomous LED billboard with 2292 modules at the JingYa hotel at XiCui Lu in Beijing, China. Reproduced with permission by BEAR Architecten T. Reijenga

the JingYa hotel in Beijing is huge[1]. The façade is in fact a gigiantic 2200 m^2 billboard with 2292 modules that collect the energy to lite the LED billboard at night (Figure 23.14).

23.2.2 PV Integrated as Roofing Louvres, Façades and Shading Devices

The designer may well use building elements such as canopies and shading systems to integrate PV systems, but will need to look in detail at shading and PV technology to understand the details of how to design this PV integration. One of the first things that the designer will discover is the fact that an efficient PV system is not automatically a good shading system. In general, a PV system on louvres will need a certain mutual distance between the louvres to prevent shading of the cells, which may let too much sun through at a lower sun angle in spring and fall (Figure 23.15).

Heat load and daylight control systems can be combined with the integration of PV systems. Moreover, the designer who studies these aspects in detail, will discover that PV systems can also

Figure 23.15 Daylight control at the Kaiser fashion house in Freiburg (Germany). This 4 kW_p PV system on louvres is mounted in front of the glass façade and prevents glare inside. The PV louvres are in the center of the figure. Reproduced with permission by BEAR Architecten T. Reijenga

[1] Designed by Simone Giostra and ARUP for Greenpix.

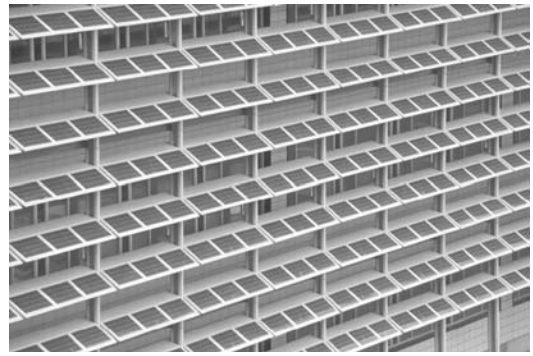


Figure 23.16 South façade of ECN Building 31 in Petten (The Netherlands). The shading structure supports the 35 kW_p Shell Solar modules. Reproduced with permission by BEAR Architecten T. Reijenga

be part of the thermal envelope or thermal system [18]. Another example is the refurbishment of Building 31 in Petten (NL) [19]. In this project, the PV system is integrated into a louvre system that supports the 35 kW_p Shell Solar modules, to keep out the summer heat and give less glare, and improve daylight conditions inside. To prevent shading of the modules by the upper louvre, the dimensions of the louvers have to be almost twice the size of the modules (Figure 23.16) [20].

Orientation is a major design issue for (green) buildings. The heat load of a building, the need for shading and the design of façades all depend on the orientation. Orientation is also important for PV systems. Façade systems might be suitable in certain countries, especially at a northern (above 50 °N) or a southern (below 50 °S) latitude. When shading of the façade cannot be prevented, and for the countries in between these latitudes, sloped surfaces facing the sun or even horizontal surfaces might be more suitable. The designer's final choice will be based on orientation, amount of total annual (sun)light on the PV module, shading from surrounding buildings and the aesthetics of the design (very usable tools for estimating and calculating the yields under various orientations and slopes have been developed by amongst others the Dutch consultancy firm Ecofys and the Polytechnics of Lausanne, Switzerland). An important issue for the designer is to appreciate the blue, grey or black cells and to become familiar with finding integration opportunities in the first draft design. Ideally, a PV system should not be added to a building but designed as part of the building.

23.2.3 Architectural Criteria for Well-integrated Systems

In order to decide whether BIPV systems are well integrated, we need to distinguish between the following:

- Technical quality of the integration of the BIPV system, that is, the technical aspects of PV, cables and inverters.
- Building quality of the BIPV system. Here we look for the quality of the integration of the system as a building element (part of the roof or the façade that is replaced by modules). The module and its integration must meet typical building standards, such as an impermeable layer or a structure strong enough to withstand wind or snow loads.
- Aesthetic quality of the BIPV system. This is the least scientific and most subjective part of judging BIPV systems. But the reality is that architecturally elegant, well-integrated systems will increase market acceptance.

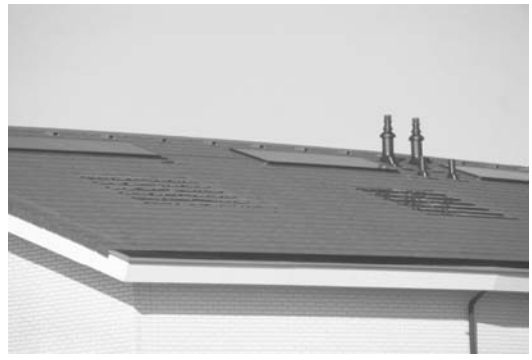


Figure 23.17 This sustainable WWF (World Wildlife Fund) project in Harderwijk, The Netherlands, has a roof with a thermal solar collector on top and a 460 W_p PV 700 system underneath. The PV 700 system fits almost invisibly in between the tiles. Reproduced with permission by BEAR Architecten T. Reijenga

Both the technical and building qualities of the PV system have been considered as preconditions. All installations in a building must function correctly.

Aesthetic quality is not a precondition. The discussion of architectural values is very broad. The average architect is not yet convinced of the "beauty" of a PV system on his building. Some architectural journals,[2] have criticized PV projects in, for example, the 250 kW_p project in Sloten, Amsterdam (NL) [25] and the 1.3 MW_p project in Nieuwland, Amersfoort (NL) [26], which are considered by many architects involved in PV projects as shining examples of good integration. All the more reason that this book should show appealing examples and critically judge PV products.

Manufacturers of building elements and products may have a different view on the aesthetics of PV. The Monier (Lafarge Braas) PV 700 roof tile system is a good example of how manufacturers look at their product. This system can be placed invisibly in between the flat Monier Stonewold tiles (Figure 23.17) [27].

However, in product advertisements, the manufacturer has chosen tiles with contrasting instead of harmonious colors, thus ignoring the fact that integration, in most situations, should be discrete. After commercial introduction, the system was prepared for use with a standard roofing tile. This corrugated tile is an even bigger contrast to the flat PV elements. Technically speaking, this high-quality product has been integrated. Aesthetically, however, the product has not been integrated because of the contrast. Therefore, the architect, building inspectors and clients might reject a PV system incorporating this product.

How can we discuss whether a BIPV system is well integrated [21]? A group of architects within the IEA PVPS (PV Power Systems) Task 7 workgroup discussed this subject and came up with several criteria for judging the aesthetic qualities of BIPV projects [22].

The criteria formulated by the IEA PVPS Task 7 workgroup for evaluating the aesthetic quality of building-integrated PV systems are [23, 24]:

- natural integration;
- designs that are architecturally pleasing;
- good composition of colors and materials;

[2] *Archis*, February 1998.

- dimensions that fit the gridula,[3] harmony, composition;
- PV systems that match the context of the building;
- well-engineered design;
- use of innovative design.

These architectural criteria need to be explained particularly to non-architects and manufacturers developing photovoltaic systems for integration into roofs and façades, who often believe that their systems fit perfectly.

- *Natural integration*. This means that the PV system seems to form a logical part of the building (Figure 23.18). The system adds the finishing touch to the building. The PV system does not have to be that obvious. In renovation situations, the result should look as though the PV system was there before the renovation.
- *Architecturally pleasing*. The design has to be architecturally pleasing (Figure 23.19). The building should look attractive and the PV system should noticeably improve the design. This is a very subjective issue, but there is no doubt that people find some buildings more pleasing than others.
- *Good composition of colors and materials*. The color and texture of the PV system should be consistent with the other materials (Figure 23.20).
- *Fit the gridula, harmony, and composition*. The dimensions of the PV system should match the dimensions of the building (Figure 23.21). This will determine the dimensions of the modules and the building grid lines used (grid = modular system of lines and dimensions used to structure the building).
- *Matching the context of the building*. The entire appearance of the building should be consistent with the PV system used (Figure 23.22). In a historic building, a tile-type system will look better than large modules. A high-tech PV system, however, would fit better in a high-tech building.
- *Well engineered.* This does not concern the waterproofing or reliability of the construction. However, it does concern the elegance of the details (Figure 23.23). Did the designers pay attention to detail? Has the amount of material been minimized? These considerations will determine the influence of the working details.
- *Innovative design.* PV systems have been used in many ways, but there are still countless new ways to be developed. This is all the more reason to consider this criterion as well (Figure 23.24).

Figure 23.18 A naturally integrated PV system that is clearly part of the building. This is the 202.8 kW_p system at the BP Solar (Solarex) facility in Frederick, MA (USA). Reproduced with permission by ECN J. Beurskens

[3] Gridula is not a common word outside architectural vocabulary. It means the grid that is used for the design that is a (sometimes hidden) part of the building.

Figure 23.19 Corridor in the Centre for Sustainability De Kleine Aarde in Boxtel (The Netherlands). The space is unheated and naturally ventilated. The 6.7 kW_p PV system with transparent modules has a double function and reduces the heat load by around 70%. Reproduced from Reijenga T, Böttger W, *Proc. 2nd World Conference on Photovoltaic Solar Energy Conversion*, pp 2748–2751 (1998) with permission by NOVEM, R Schropp [10]

Figure 23.20 The 80.5 kW_p Atlantis Sunslates on the roof of the historic horse stables in Bern (Switzerland). The color and texture matches so well that this PV system was allowed on a protected historic building. Reproduced with permission by Atlantis Solar Systems Ltd

23.2.4 Integration of PV Modules in Architecture

The preceding section has discussed in brief the architectural criteria for judging a PV system as such. The following section focuses on the way in which these systems can be integrated into the architectural concept of the building.

The integration of PV systems in architecture can be divided into five categories:

1. Applied invisibly.
2. Added to the design.
3. Adding to the architectural image.
4. Determining architectural image.
5. Leading to new architectural concepts.

These categories have been classified according to the increasing extent of architectural integration. However, a project does not necessarily have to be of a lesser quality just because PV

Figure 23.21 The cube at the Discovery Science Center in Santa Ana, Los Angeles, CA (USA). The 20 kW_p PV system fits the gridula of this huge structure and there is harmony between the PV modules and the structure behind. Some 494 Thin film Millienna photovoltaic modules from BP Solar were used on the solar cube. Reproduced from Eiffert P, Kiss G, *Building-Integrated Photovoltaic Designs for Commercial and Institutional Structures – A Sourcebook for Architects*, 48–49, NREL, Golden, CO (2000) with permission by BEAR Architecten T. Reijenga [17, 29]

Figure 23.22 The 180 kW_p Sofrel flat-roof system on the UBS Bank near Lugano (Switzerland) matches the context of the roof, which fits between different installations, high-tech chimneys and the tight rhythm of the flat-roof PV system on the roof. Here the BIPV is mainly aesthetically and not physically integrated. Reproduced from Maycock P *et al*., *Building with Photovoltaics*, 78–81, Ten Hagen & Stam, Den Haag (1995) with permisson by UBS Switzerland [9]

Figure 23.23 The small pavillions at the Beijing Olympic site are well engineered, Beijing, China. Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.24 The solar louvres for the Solar Declathon 2007 (USA) are very innovative. Reproduced with permission by Solar Declathon 2007, TU Darmstadt, Sebastian Sprenger

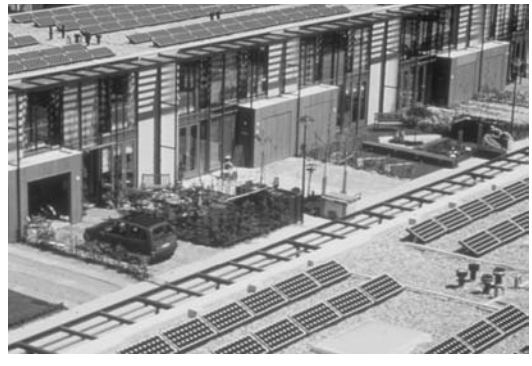


Figure 23.25 Modern style houses in the 1.3 MW project in Amersfoort, The Netherlands. The architect thought that the modules did not blend in with the design. He chose an "invisible" solution with a flat-roof system. The PV system is not visible from the surrounding streets. Reproduced with permission by REMU J. van IJken

Figure 23.26 Houses at the National Research Home Park in Bowie, MA, USA in historic style. The roofs are Unisolar standing-seam roofs with thin film (amorphous silicon). Reproduced with permission by NREL USA

Figure 23.27 The IES building in Madrid, Spain has a 6.6 kW_p PV system that was added shortly after the building was finished. The modules are mounted on the façade above the windows and keep out the sun. Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.28 These houses in Amersfoort Nieuwland, The Netherlands have a roof with a white foil on top. The architect chose a color contrasting with the color of the modules to show that the PV system is an addition to his design. Reproduced with permission by REMU J. van IJken

modules have been applied invisibly. A visible PV system is not always appropriate, especially in renovation projects with historic architectural styles. The challenge for architects, however, is to integrate PV modules into buildings properly. PV modules are new building materials that offer new designing options. Applying PV modules in architecture should therefore lead to new designs. In some of the selected projects, the design was based on this principle.

1. *Applied invisibly*. The PV system has been incorporated invisibly, and is therefore not architecturally "disturbing" (Figure 23.25). The PV system harmonizes with the total project. An example is the Maryland project in the USA (Figure 23.26), where the architect tried to integrate PV modules into the design invisibly. This solution was chosen because the entire project concerned historic architecture. A modern high-tech PV module look would not be appropriate for this architectural style.
2. *Added to the design.* The PV system is added to the design (Figure 23.27). Building integration is not really used here, but this does not necessarily mean that architectural integration is also lacking. The "added" PV system is not always visible either (Figure 23.28).

3. *The PV system adds to the architectural image.* The PV system has been integrated beautifully into the total design of the building, without changing the project's image (Figure 23.30). In other words, the contextual integration is very good (Figure 23.29).
4. *The PV system determines the architectural image.* The PV system has been integrated into the design in a remarkable and beautiful way and plays an important role in the total image of the building (Figure 23.31).
5. *PV system leads to new architectural concepts.* Using PV modules, possibly in combination with other types of solar energy, leads to new designs (Figure 23.32) and new architecture (Figures 23.33 and 23.34). The integration of PV modules was considered on a conceptual level, which gives the project extra value.

23.3 BIPV BASICS

23.3.1 Categories and Types of Building

Building-integrated PV systems can be divided into different categories according to:

1. cell and module type;
2. architectural integration;
3. type of building;
4. mounting technology;
5. the function of the integration, and possible additional building and architectural functions of the PV system.

It is important that architects know all these categories and their possibilities. The design process consists of translating the brief (program made by the client) into spaces and enclosures, as well as combining functions and materials into constructions. This process is mainly based on experience and knowledge of constructions and materials. New applications based on existing knowledge or techniques are very important in the creative process. New inventions play a minor role in the process.

Figure 23.29 The architect integrated the (in total 49 kW_p) PV modules into the façades above the windows and produced a combination with the roller blinds. The EMPA office building is located in Sankt Gallen, Switzerland. Reproduced with permission by Electrowatt Eng. Services

Figure 23.30 In this office building in Gouda, The Netherlands, the 6.2 kW_p PV system is mounted on top of the façade as a canopy that protects the wall. The architect's intention was to show the PV system to his clients when they visit the office. Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.31 The dwellings at the 5 MW project in the HAL district of Langedijk, The Netherlands, have a large PV roof without any perforations. This strengthens the architectural expression of the roof design. Each roof includes a 5.1 kW_p BP Solarex PV system. Reproduced with permission by BEAR Architecten T. Reijenga

23.3.1.1 Categories of cells and modules

There is a wide range of cells and modules in the market. There are various types of cell material, types of modules, framed or non-framed laminates, colors of the cells and colors of back sheets and frames; all provide a wide range of possible surfaces [30]. This is a very basic knowledge for architects. Architects will design BIPV systems with a certain image in mind. Their choice of monocrystalline or polycrystalline (multicrystalline) cells will depend on color and not on efficiency [31].

23.3.1.2 Categories of integration

BIPV systems in projects can be divided according to the location of the application: roof systems, façade systems, glass construction (conservatory/atrium) systems and building components such as shading and canopy systems. The main mounting locations are the roof and the façade. There are

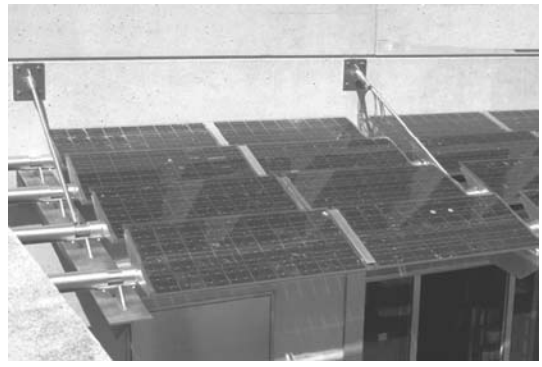


Figure 23.32 This shopping mall in Lausanne, Switzerland, includes a louver system with transparent modules mounted above the shop windows, which results in less reflection of sunlight in the windows, making it easier to look into the shops. Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.33 This zero-energy project in Etten-Leur, The Netherlands, combines different sustainable energy techniques. Each house has a roof structure with 6.2 kW_p BP Solar PV modules. The aesthetic integration of the PV system is a completely new concept in housing. Reproduced from Reijenga T, Energy-Efficient And Zero-Energy Building in the Netherlands, *Proc. International Workshop on Energy Efficiency in Buildings in China for the 21st Century*, CBEEA (Beijing, December 2000) with permission by BEAR Architecten [29]

also many creative solutions available in designing how PV systems can be integrated. All these solutions are grouped as building components.

23.3.1.3 Types of building

Different buildings obviously have very different functions, for example, apartments and family houses, public buildings, commercial and industrial buildings and non-occupied building structures. These can either be newly constructed or renovated buildings [32]. Non-occupied building structures include shelters (bus stops, canopies, parking structures), kiosks (newsstands, gazebos, pavilions, phone booths), public toilet buildings, car parks, streetlights, parking meters, screens and barriers (fences, noise barriers), road signs and commercial billboards.

Figure 23.34 Here the roof is facing north and the architect designed a "PV-spoiler". Each house has a structure with 0.4 kW_p Shell Solar PV modules. Reproduced with permission by BEAR Architecten T. Reijenga

All types of building or non occupied building structures are a potential location for BIPV systems. The various commercially available mounting applications[4] can be grouped into 10 types.

Location on a building:	Building component(s):	Manufacturers such as:
I. Sloped roof	1. Tiles and shingles	Monier (Figure 23.17), Atlantis Sunslates [33] (Figure 23.20), Unisolar/Bekeart Standing Seam panels [34] (Figure 23.35).
	2. Non integrated profiles	BP Sunflower (Figure 23.36), Econergy InterSole [35], Alutec profiles
	3. Integrated profiles	Solrif Solar profile system [36] (Figure 23.37)
II. Flat roof	4. Roofing element	SunPower PowerGuard®, Alwitra Evalon® roofing foil (Figure 23.38)
	5. Integrated profiles	Schüco profile system (Figure 23.39)
	6. Independent support structure [37]	Solgreen (Figure 23.40), Sofrel [38, 39] (Figure 23.28)
III. Façade	7. Integrated profiles	Schüco façade profile system (Figure 23.41)
	8. Cladding system	BP Solface [40] system
	9. Louvres or sun blinds	ADO louvre system (Figure 23.42), Colt Shadovoltaic louvre system (Figure 23.15)
	10. Canopy	Donjon canopy system (Figure 23.43)

23.3.1.4 Function of the integration

In addition to generating electricity, PV modules are also used as an integral part of the external skin (roof or façade), as sun protection (Figure 23.44) or as a daylight transmitter. Designing

[4] This overview just gives an impression and is not a total overview of all available systems. Also product specifications and brand names may be changed. A good digital overview is yearly updated by the German magazine *Photon*.

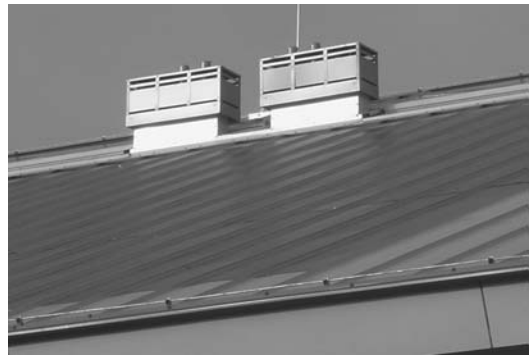


Figure 23.35 Bekaert/BESS Europe standing seam roof in a project in Leiden, The Netherlands. Capacity of the system is 21 kW_p. Integration between PV system, solar hot water modules and chimney. Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.36 BP Sunflower system at the bent roof at ECN building 31, Petten, The Netherlands. Capacity 35 kW_p. Reproduced with permission by BEAR Architecten H. Lieverse

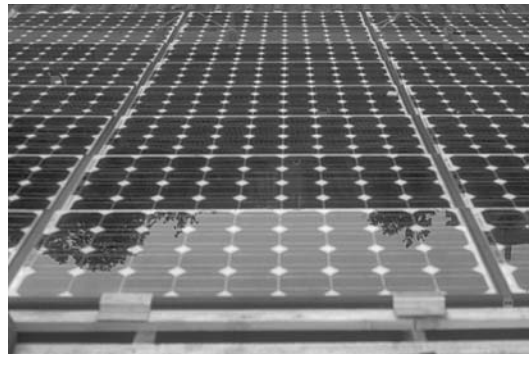


Figure 23.37 Solrif Solar profile system in a roof renovation in Zürich, Switzerland. Capacity of the system is 53 kW_p. Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.38 Alwitra Evalon® roofing foil with amorphous silicon cells. Reproduced with permission by O.Ö. Energiesparverband

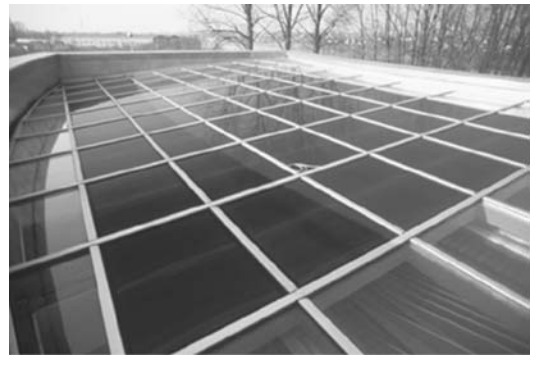


Figure 23.39 Schüco 1.9 kW_p roof profile system with transparent modules in the Energy Forum Center in Bad Oeyenhausen, Germany. This mounting system can also be used in (difficult) horizontal situations. Reproduced with permission by BEAR Architecten T. Reijenga

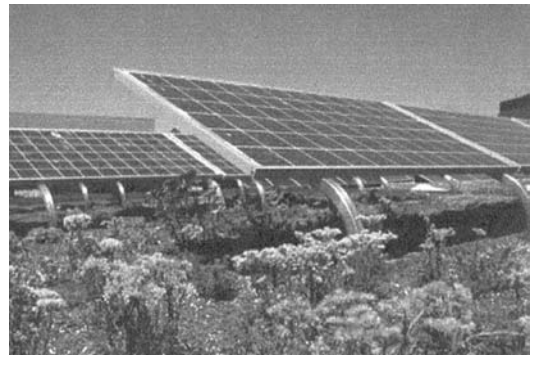


Figure 23.40 Solgreen flat-roof system for green roofs demonstrated on a roof in Switzerland. Reproduced with permission by SOLSTIS

Figure 23.41 Schüco façade profile system in the Doxford building, Great Britain. See also caption at Figure 23.6. Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.42 ADO louver system with integrated transparent PV modules at an apartment building in Amersfoort Nieuwland, The Netherlands. The automatic louvres can switch in two positions. The horizontal position at night (left on the figure) and tilted to the sun by day (right on the figure). Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.43 Canopy at the eaves of a roof. 6.2 kW_p PV system at an office building in Gouda, The Netherlands. Reproduced with permission by BEAR Architecten T. Reijenga

Figure 23.44 This building, at ECN in Petten, The Netherlands has a 43 kW_p PV system integrated into the conservatory roof. The conservatory acts as a parasol in front of the offices, thus eliminating the need for air-conditioning in a moderate climate. The building on the right side has a PV shading system (see caption to Figure 23.16). Reproduced with permission by BEAR Architecten T. Reijenga. See Plate 8 for the colour figure

double functions and then integrating PV modules into buildings results in cost reductions on the investment in the building and in a higher acceptance of PV as an integrated component. Several buildings have been built that demonstrate PV systems as part of a passive cooling strategy [13].

The advantages of this integration are

- the PV modules replace building elements;
- the PV modules are very well ventilated at the back;
- a separate mounting construction is not necessary;
- the air-conditioning system is eliminated.

23.3.2 Cells and Modules

The PV modules and mounting system are the elements of a PV system that can determine the image of a building. A PV system, particularly the cell material, framing materials, soldering, shape of the modules and the color of cells and back sheets, all influence the image of a building. For architects and designers, these aspects are more important than the electrical efficiency of a system. Typical efficiencies of today's commercially available solar modules are shown in Table 23.1.

Table 23.1 Typical efficiencies for modules. These values are obtained under standard test conditions. Different orientation in BIPV applications may result in lower performance

Cell type	Typical efficiencies [%]
Monocrystalline silicon	14–18
Multicrystalline silicon	12–15
Thin film silicon	6–10
Single amorphous	6–7
Cadmium telluride	10–12
Copper indium gallium diselenide	8–12

23.3.2.1 Solar cell materials

There are several types of solar cell materials: monocrystalline (single-crystal) silicon, polycrystalline silicon also called multicrystalline silicon (both discussed in Chapters 6 and 7), amorphous silicon (Chapter 12), copper indium gallium diselenide ($CuInGaSe_2$) (Chapter 13), and cadmium telluride (CdTe) (Chapter 14). Their characteristics that affect their implementation in BIPV applications are briefly presented here.

Monocrystalline Si cells are initially grown as long cylinders, then sliced into thin disks called wafers (~300 μm thick). Initially round, the wafers often have their edges cut to create a nearly square wafer with slightly rounded corners. This increases their packing density on the module. Typical wafers are presently $12.5 \times 12.5\,cm^2$ and $15.6 \times 15.6\,cm^2$ and may increase as technology develops. Monocrystalline cells have a very uniform appearance. Their color can be varied (see below) but are typically either dark blue or black since this gives the highest efficiency. The color of the cell is determined by the wavelengths that are reflected. The darker the cell looks, the less light that is being reflected. Therefore: darker means more absorption of sunlight by the solar cell. However it is possible to design cells with little losses of efficiency in a large variety of colors because the color can be caused by the reflection of a narrow band of wavelengths.

Polycrystalline or multicrystalline silicon wafers are manufactured with a lower-cost process per unity of area than monocrystalline silicon wafers. They are cast in square ingots. After slicing, polycrystalline wafers are already in the desired square shape. Compared with monocrystalline cells, polycrystalline cells also typically have a bluish color and are the same size, but are slightly less efficient and slightly of lower cost. The main difference in the appearance between mono and poly wafers, where one could see clearly the silicon crystallites on the poly material, has disappeared after the 1990s by the application of acidic texturization. This now standard process step results in uniform etching of the silicon surface, improves the reflectance of the surface and creates a uniform dark bluish color after applying the antireflection coating[5]. Both mono and poly cells of the older types have metal grids on the front in a rectangular pattern to collect the electricity and to connect to the next cell. These grids are typically not visible from beyond a few meters away. As large progress is made in so called back contact solar cells, having the necessary contacts at the back, these cells do not show the metal grid.

Amorphous silicon cells (a-Si) are composed of silicon atoms that are in a thin (~1 μm) layer and lack crystalline properties. They are commonly referred to as thin film Si PV technology. Amorphous silicon cells are deposited onto substrates such as glass window panes or flexible rolls of stainless steel or plastic, giving a wide range of mechanical strength, weight, and flexibility. The substrate is not visible since it is behind the solar cell. The cells have a uniform typically dark appearance. These cells have no grids. Flexible substrates are ideal for curved surfaces and rollable "fold-away" modules. Amorphous silicon modules have lower efficiency than mono or poly silicon (Table 23.1), but better performance at higher temperatures [41] (see Chapter 12] as often occurs in BIPV applications.

Other thin film PV materials presently include $CuInGaSe_2$ and CdTe. They have a uniform, nearly black appearance, indistinguishable visually from amorphous Si modules. They also have lower efficiency compared with mono- or polycrystalline silicon. $CuInGaSe_2$ cells can be deposited on flexible plastic or metal foils.

Semi-transparent cells for BIPV can be manufactured in two ways. Mono silicon wafers can have a series of deep grooves on the front and back which are perpendicular to one another. Where they intersect, light will be transmitted through the holes. Polycrystalline silicon cells with

[5] Personal communication from Jaap Hoornstra, Solar Energy Department of ECN.

2% transmission have been reported [42], but higher values should be possible with larger holes. Another approach is to make very thin amorphous silicon layers on glass with transparent contacts so the entire module is semi-transparent. However, the transmitted light will have an orange or red tint because the blue and green portion of the spectrum is absorbed in the silicon layers. Such modules could only be used in applications where this color of light was acceptable such as sunroofs for automobiles. A better method is to selectively remove part of the amorphous silicon layer using laser ablation. Unfiltered white light transmission of 5–15% has been reported for laser-scribed amorphous silicon in BIPV applications [41]. But modules with cells not filling all the space can also have good properties for the light going through it (through the module not through the cell). This is the option most often used, together with leaving light transmitting space between ordinary modules.

23.3.2.2 Module temperature

Module efficiency, therefore electricity produced, decreases as the temperature increases for mono- and polycrystalline silicon cells, but less so for amorphous silicon cells. In many non-BIPV applications, modules are mounted on free-standing frames with ambient air on both sides, allowing for cooling on both sides. In contrast, some BIPV applications install the modules in close contact to building material such as roofs or wall insulation. The lack of circulating air increases the module temperature. Relative losses of >5% are possible [43]. A good design criterion for mono- or polycrystalline silicon applications is to allow as much cooling as possible by providing for air flow behind the module and minimizing effect of insulation. This is not an issue for amorphous silicon modules [43].

23.3.2.3 Color of the cells and modules

Solar cells are basically blue, dark blue or black after processing. Different colors are possible but these are not manufactured as standard. Some manufacturers sell tailor-made cells in special colors (e.g. gold, gray, green, red–orange and yellow). The cell color is varied by changing the thickness of various optical coatings on top of the cell, which changes their reflection. The blue color produces the highest efficiency solar cells. Current literature gives efficiencies for colored cells as 11.8% and 14.5% compared with optimized cell efficiency of 16.8% [44, 45], which corresponds to around 75% of the power of dark-blue cells [46].

Modules have several sections that can be colored. Besides the cells, the frame and the back sheet will also have a certain color. Older modules had natural aluminum frames and a white back sheet. The shape of the cell was very pronounced because of the contrasting color. However, modern modules have colors that are more in harmony: dark-blue cells with a dark back Tedlar® sheet and a dark-colored frame around the module, which produces a very uniform image. A roof or façade containing these uniformly colored modules will be seen as a single surface. The opposite effect is also possible by using modules in striking colors to attract attention and focus on the PV system.

23.3.2.4 The architecture of modules

Architects select modules based on their shape and composition possibilities. There is a big visual difference between framed and frameless modules.

Frameless modules look very similar to window glass. A surface with frameless modules with a “hidden” mounting system looks very uniform. The seams seem to be hidden and the individual module is hard to recognize. This smooth surface has a high aesthetic value.

Framed modules give a totally different effect. The frames can be heavy and therefore determine the total impression of the surface. The very visible frames divide the surface into modules and every individual module is very recognizable. This is not always the image envisaged by the architect.

To solve this problem, smaller frames in the same (dark or blue) color as the cells can be used and are less visible. The soldering between the cells is a small detail, but is an important part of the image for very visible PV systems. In the older techniques the soldering was very visible and not very smooth. However, new techniques mean that the soldering is better hidden and new types of soldering, for example Black Contact cells and cells using the ECN PUM approach, are applied already in some modern modules.

Modules vary significantly in size. Standard modules are less expensive than applications using tailor-made modules. However, almost every form, shape and dimension is possible with tailor-made modules. The glazing is available as single and double (insulating) glass. Thin film modules possibly allow greater freedom to select size and color than c-Si modules.

A new type of modules are the HCPV (high-concentration photovoltaics) solar modules. Several companies in Europe (SOL3g in Spain) and the USA (SUNRGI in California) develop this modules. The principle is based on rather cheap Fresnel lenses that concentrate 500 or more times direct solar light on a triple-junction cell (efficiency >35%). The cells need passive cooling to get rid of the excessive heat. However application of these systems are not common yet in architecture.

23.4 STEPS IN THE DESIGN PROCESS WITH PV

23.4.1 Urban Aspects

The aim of integrating PV systems into buildings is to reduce costs and to optimize of scarce ground in urban areas. To generate maximum power from building-integrated systems, certain urban and architectural aspects are important.

The main starting point is the maximum power that can be generated by a system. The primary hindrances can be the (partial) shadowing of a system by other buildings or objects, and the nonoptimum orientation relative to the sun. Reflection can also be a problem for the surrounding buildings.

23.4.1.1 Orientation and angle

The amount of irradiance depends on the latitude. The maximum irradiance corresponds to surfaces, tilted at an angle equal to about the latitude minus 10° (see Chapter 21 for calculations on solar irradiance). For instance, at 52° north good results (>90%) can be achieved by orienting the modules between southeast and southwest, with system angles between 30° and 50° from the horizontal. Orientations between east and southeast, and between southwest and west, are fairly reasonable with system angles between 10° and 30° from horizontal. The irradiance will be reduced by around 15% of the maximum.

Flat-roof systems with very low angles (between 5° and 10°) can be a good solution for difficult orientations.

23.4.1.2 Distance between buildings

Shadowing is a critical issue for BIPV. In general, designs in which the PV modules are shaded for much of the year should be avoided. For low-rise areas, the problem is easy to solve. The distance between individual houses can be calculated. For mixed-height neighborhoods, it will be more difficult. A high-rise apartment building in a low-rise neighborhood can cause a lot of unwanted shading.

The density of an area also has a lot of influence. In high-density areas (cities) the distances between buildings are often so small that there is significant shadowing throughout a large part of the year.

On a general note, it is worth mentioning that façade systems (vertical) are more sensitive to shading and need longer distances from other buildings than tilted systems (roofs). Horizontal systems have a lower irradiance, as previously mentioned, but will be the best solutions for avoiding shadow. Only neighborhoods with a mixture of low- and high-rise buildings might be unsuitable for horizontal systems.

23.4.1.3 Trees

Greening the area around buildings makes the area look very attractive and the microclimate more comfortable for the inhabitants.

The shadowing effect of trees is very important, as the trees will be very dense during the summer. Even during the winter, when trees lose their leaves, the branches give too much shade.

The aspect of growth is sometimes underestimated. Planning for the future growth of trees is very important and must be done carefully to avoid problems a few years after the building has been completed or the PV system has been installed.

Solutions can be to:

- only plant trees on the north side of buildings;
- plant only small trees up to two stories high;
- prune trees annually to keep them small.

23.4.1.4 Zoning

In future, a special solar zoning will be needed in urban areas with PV systems. The borders of building areas can be clearly marked on three-dimensional maps to prevent future problems. The amount of sunlight can also be determined on these maps.

23.4.1.5 Reflection

Although not a major problem, under certain circumstances, reflection can occur. In low-rise buildings, there are no significant problems, but in mixed low- and high-rise areas residents in high-rise buildings may experience some annoying reflections if all the surrounding houses have (glass-covered) PV modules. The fact that there are certain distances between buildings (for shadowing) may eliminate most of the potential problems.

23.4.2 Practical Rules for Integration

There are a few important rules for integrating modules into buildings. These rules concern the functioning and maintenance of the system, for example:

1. shadow is not allowed on the module;
2. ventilation is required at the back of the modules (not as important for thin film a-Si);
3. make it easy to mount and remove a module;
4. ensure that the module stays clean or can be cleaned;
5. make easy electrical connections;
6. ensure that wiring is sun-proof and weatherproof.

As previously mentioned, even partial shadow on the modules will strongly decrease the energy output. Profiled mounting constructions, in particular awnings, can produce shadows along the edge of the adjacent module that will result in loss of efficiency.

Modules with crystalline cells have a higher output when the temperature is lower. With ventilation at the back of the module it is possible to keep the temperature low and avoid a decreasing output. However, thin film amorphous silicon reacts differently. The higher temperature does not influence the efficiency as much as crystalline silicon.[6]

Although the lifetime of modules is proven to be over 20 years, it is better to know how to remove a single module in the middle without removing the whole system. Electrical connections should also be "plug and play".[7] Easy electrical connections are needed for fast installation and for easy replacement of modules. Depending on the local safety regulations, precautionary measures should be taken, for example, using lifelines or moveable safety ladders.

The surface of the modules should be clean. Tilted modules will be cleaned by rain in most regions. Modules mounted at low angles can be treated with PV-Guard, a treatment that makes the surface smooth and makes cleaning with rain easier. In dry areas, cleaning should be part of the regular maintenance schedule.

Protect wiring against the weather. Rain is not the main problem, though all connections must be waterproof. Long-term influence from water should be avoided. Protection against sun and UV light is needed to ensure that the insulation of the wiring does not deteriorate. Depending on the area, wiring may also need to be protected from small gnawing animals.

23.4.3 Step-by-step Design

A PV system consists of a number of modules with solar cells, an inverter, batteries or, in most cases, a connection to the grid. A single house with a small installation can be connected through the existing electricity meter. The electricity that is produced will be used primarily in the house. Any surplus will be fed into the grid and the meter will spin backwards (Figure 23.45). However, not every utility company will allow this and in some cases a second meter is installed. This often happens with larger systems (more than 2 or 3 kW_p). Larger systems or combined systems that are maintained by the utilities are connected directly to the grid.

[6] See Chapter 12.

[7] "Plug and play" refers to very simple wiring and components that fit together like a computer and can easily be replaced.

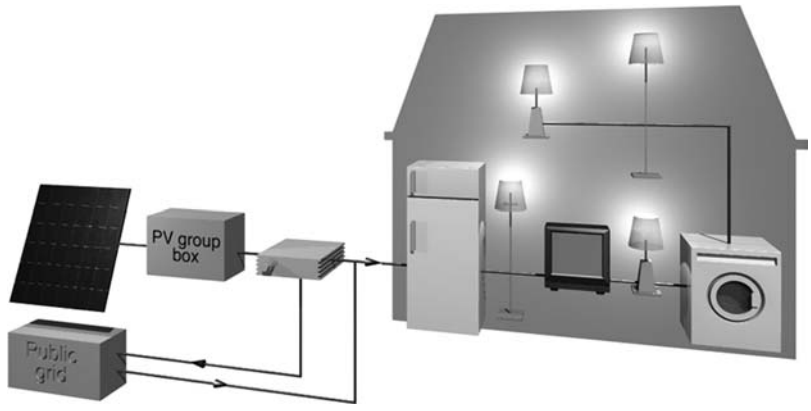


Figure 23.45 Scheme of gridconnection. Reproduced with permission by ECN M. van der Laan

23.4.3.1 Solar design

To design with photovoltaics the first set of questions are, "Why do I want to integrate PV into the building?" and "Is it for general energy supply, to make the building more independent, or to make a statement about the building's inhabitants?"

Large systems will be used for general energy supply. This means large surfaces that can be treated in an architectural way. Different types of modules, shapes, colors or textures can be used to design the look of the building.

The main issues for a more independent building are the efficiency of the system and the generated yearly output. The size of the PV system will depend on this and the designer has to allow for a certain number of modules. The designer will probably design the building around the integrated system, otherwise the system will be something that is connected, but not integrated, into the building.

23.4.3.2 Module placement and shadowing

The first step in the design process will be to look at the number of modules, their dimensions and the total dimensions of the system. All these aspects have to be integrated into the roof or façade. Shadowing of the modules is important. A module that is partly shaded will lose more efficiency than expected. If one row of cells in a module is covered or heavily shaded, this can block the output of the entire string.

Small objects that can cause shading, such as chimneys and fans, are less important. The shadow will move during the day and there may be indirect light available. Actually, modules have bypass diodes that allow a short break when a row of cells is covered or shaded.

23.4.3.3 Space required for balance of systems and interconnections

The modules have junction boxes at the back that are connected by cables to the inverters. Space is also required for a junction box at the back. Together with the ventilation required at the back of the module, this means a gap of 20–50 mm (depending on the size of the junction box) between the back of the module and the mounting surface that can be used for both functions.

Space is also required for the inverters. For better efficiency, the best place for these inverters is near the modules. An AC cable has to be fed from the inverters into the grid via the meter.

Safety switches are required near the inverters in order to work on the PV system safely. Space for a second utility meter may be required near the first meter, unless a double meter can be used.

23.4.4 Design Process: Strategic Planning

A few procedural steps may be necessary to ensure that the PV system is successfully integrated into the design. A common rule is to integrate the PV system into the building process without disturbing that process.

Step 1. The first step is consultation with the authorities about local regulations, building permits and the electrical connection to the grid.

Step 2. The second step is to consult the utility company about the grid connection, electrical diagrams and the metering system.

Step 3. The third step is the internal meeting with all building partners. A kick-off meeting very early in the process may be useful, to discuss the entire integrated PV system with the building contractor, the roofing company, the electrician and the PV supplier.

There are many unique issues to resolve in installing BIPV. The main points in this meeting concern the responsibilities of each party in the building process. Who is responsible for the water-proofing of the roof – the roofing company or the PV installer? Who is responsible for electrical safety – the electrician or the PV installer? Who is responsible for safety on the site – the general contractor or the PV installer? All these aspects must be clearly defined and noted in advance.

Many PV suppliers offer turnkey contracts. This is easy for clients because they receive a complete working system for their money. However, the client is then responsible for the coordination between PV supplier and building contractor. Placing all responsibility with the building contractor means an extra surcharge of perhaps 10% on the cost of the PV system. A good solution is to make the building contractor (general contractor) responsible for the PV system and negotiate a special fee for coordination and use of equipment (scaffolds and crane) from the building contractor.

23.5 CONCLUDING REMARKS

Building integration aims to reduce costs and minimize the requirement for land. To increase market acceptance it is important to show architecturally elegant, well-integrated systems. Moreover, building owners can show their environmental commitment with highly visible building-integrated PV systems.

This means that there is a large potential for BIPV in the built environment. The main factors for successful integration are suitable buildings [47] (i.e. suitable orientation and lack of shadow) and a reason for building integration. For newly constructed sustainable buildings, BIPV will be part of the energy strategy. However, for existing buildings there must be a valid reason for integrating PV systems. Building renovation, including the roof and façade, often provides an opportune time for selecting BIPV [48, 49].

The building or renovation process plays an important role in the success of BIPV. Can the building owner benefit from BIPV? If so, the owner will be willing to implement PV systems

in the building plans. The architect or designer needs to have a good basic knowledge of BIPV and be able to integrate PV into the design. If architects do not understand the basics of PV, they will make mistakes that eventually have to be resolved during the installation process. The worst cases involve mistakes that cannot be resolved at the end of the building process and that result in a lower efficiency and quality of the PV system.

Is the utility company willing to cooperate? If not, the building owner will probably try to avoid difficulties in an already complex construction process.

The architect or designer should use all visible opportunities to integrate PV into the design in a highly aesthetic way. The important issues are the architectural function of a PV module (replacing other building elements) and the visible aspects of modules, such as the dimensions, mounting system, form and color of cells, back sheet and frames.

To recognize these aspects, criteria have been formulated for judging building integration of PV. These criteria are useful for manufacturers and technicians who are involved with building integration from the engineering and technical aspects of the building process.

REFERENCES

1. Reijenga T, The Changing Cities of Europe, *Proc. Sustain '99* (Amsterdam, 1999).
2. Schoen T *et al*., *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 359–364 (1997).
3. Kurokawa K, Kato K, Paletta F, Iliceto A, *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2853–2855 (1998).
4. Thomas R, Grainger T, Gething B, Keys M, *Photovoltaics in Buildings – A Design Guide*, Report S/P2/00282/REP, ETSU, DTI, London (1999).
5. Strong S, Lloyd Jones D, *A Renewable Future*, IEA PVPS Task7, Final Task (Feb. 2001).
6. Kiss G, *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2452–2455 (1998).
7. Muller A, Roecker C, Bonvin J, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 889–892 (1997).
8. Reijenga T, *Prog. Photovolt*. **4**, pp 279–294 (1996).
9. Maycock P *et al*., *Building with Photovoltaics*, Ten Hagen & Stam, Den Haag, pp 78–81 (1995).
10. Reijenga T, Böttger W, *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2748–2751 (1998).
11. Lloyd Jones D, Matson C, Pearsall N, *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2559–2562 (1998).
12. Wilk H, *OKA-House of the Future*, IEA SHCP Task 19 (1997).
13. Reijenga T, Kaan H, *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1952–1959 (2000).
14. Hynes K, Pearsall N, Shaw M, Crick F, *Proc. 13th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2203–2205 (1995).
15. Hagemann I, Leppänen J, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 694–697 (1997).
16. Bijlsma, Joost, Lichtbaken voor een schone toekomst, NUON LUMEN 14, Amsterdam. pp 50–51 (2008).
17. Eiffert P, Kiss G, *Building-Integrated Photovoltaic Designs for Commercial and Institutional Structures – A Sourcebook for Architects*, NREL, Golden, CO, pp 48–49 (2000).
18. Pitts A, Tregenza P, Coutts R, *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1902–1905 (2000).
19. Kaan H, Reijenga T, *Proc. 2nd EuroSun*, pp V3.2–1–V3.2-6 (1998).

20. Reijenga T, Kaan H, *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2740–2743 (1998).
21. Reijenga T, Schoen T, *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2744–2745 (1998).
22. Schoen T *et al*., *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1840–1843 (2000).
23. Reijenga T, Photovoltaic Building Integration Concepts – What do Architects need? *Proc. IEA PVPS Task7 Workshop Lausanne Featuring A Review of PV Products*, IEA PVPS Task7, Halcrow Gilbert, Swindon (2000).
24. Reijenga T, Photovoltaics in the Built Environment, *Proc. 2nd World Solar Electric Buildings Conference*. ESAA, ANZSES (Sydney, 2000).
25. Cace J, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 698–700 (1997).
26. Vlek F, Schoen T, Iliceto A, *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1783–1786 (2000).
27. Reijenga T, *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1793–1796 (2000).
28. Eiffert P, Kiss G, *Building-Integrated Photovoltaic Designs for Commercial and Institutional Structures – A Sourcebook for Architects*, NREL, Golden, CO, 48–49 (2000).
29. Reijenga T, Energy-Efficient And Zero-Energy Building in the Netherlands, *Proc. International Workshop on Energy Efficiency in Buildings in China for the 21st Century*, CBEEA (Beijing, December 2000).
30. Ito T, Ishikawa N, Nii T, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 690–693 (1997).
31. Butson J *et al*., *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2571–2574 (1998).
32. Kaan H, Reijenga T, *Proc. ACEEE*, pp 5.025–5.014 (1998).
33. Posnansky M, Szacsvay T, Dütsch B, Stucki B, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1922–1924 (1997).
34. Nath P *et al*., *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2538–2541 (1998).
35. Scheijgrond P *et al*., *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2049–2050 (2000).
36. Toggweiler P, Ruoss D, Brügger U, Haller A, *16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1972–1975 (2000).
37. Böttger W, Schalkwijk M, Schoen A, Weiden T, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2288–2291 (1997).
38. Roecker C, Bonvin J, Toggweiler P, Ruoss D, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 701–704 (1997).
39. Bonvin J, Roecker C, Affolter P, Muller A, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1849–1850 (1997).
40. Schnaller F, Roecker C, *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1945–1947 (2000).
41. Arya R, Carlson D, *Prog. Photovoltaics* **10**, 69–76 (2002).
42. Willeke G, Fath P, *Appl. Physics Lett*. **64**, 1274–1276 (1994).
43. Doughtery B, Hunter-Fanney A, Davis M, *J. Solar Energy Engineering*, **127**(3), 314–323 (2005).
44. Mason N, Bruton T, *Proc. 13th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 2218–2219 (1995).
45. Ishikawa N *et al*., *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2501–2506 (1998).
46. Tölle R *et al*., *Proc. 16th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 1957–1959 (2000).
47. Frantzis L, Ghosh A, Rogers M, Kern E, *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2799–2801 (1998).

48. Gutchner M, Nowak S, *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2682–2685 (1998).
49. Reijenga T, Drok M, Oldegarm J, Kampen J van, *PV Systems and Renovation*, TNO, NOVEM, Delft (2001).

Further reading

Gaiddon, Bruno, Henk Kaan and Donna Munro (ed): *Photovoltaics in the Urban Environment: Lessons learnt from Large Scale Projects*, London (2009).
Hagemann, Ingo, *Gebäudeintegrierte Photovoltaik*, Verlag Rudolf Müller, Köln (2002).
Humm O, Toggweiler P, *Photovoltaics in Architecture*, Birkhäuser, Basel (1993).
Lloyd Jones D, Hattersley L, Ager R, Koyama A, *Photovoltaics in Buildings – BIPV Projects*, ETSU, DTI, London (2000).
Muller A, Roecker C, Bonvin J, *Proc. 14th Euro. Conf. Photovoltaic Solar Energy Conversion*, pp 889–892 (1997).
Prasad D, Snow M, *Designing with solar power – A source book for building integrated photovoltaics (BIPV)*, Images Publishing Sydney (2002).
Schoen T, Prasad D, Toggweiler P, Eiffert P, *Proc. 2nd WC Photovoltaic Solar Energy Conversion*, pp 2447–2451 (1998).
Strong S, *Photovoltaics in the Built Environment – A Design Guide for Architects and Engineers*, DOE/GO-10097-436, NREL, Golden, CO (1997).
Photovoltaics in an Architectural Context, *Prog. Photovoltaics* **12**(6), 395–408 (2004).
Photovoltaics in the Built Environment (Special Issue) *Prog. Photovoltaics* **4**(4), 237–320 (1996).

Internet

www.pvportal.com
www.iea-pvps.org
www.pvupscale.org
www.task7.org
www.iea-pvps-task10.org
www.pvdatabase.org
www.pvdatabase.com
www.demosite.ch

24

Photovoltaics and Development

Jorge M. Huacuz[1], Jaime Agredano[1] and Lalith Gunaratne[2]

[1]*Instituto de Investigaciones Eléctricas, Gerencia de Energías No Convencionales, Cuernavaca, México,* [2]*Solar Power & Light Co, Ltd, Colombo, Sri Lanka*

24.1 ELECTRICITY AND DEVELOPMENT

24.1.1 Energy and Early Humans

Survival has always been the main preoccupation of mankind. For thousands of years, food, shelter and protection against harsh weather and wild animals were the primary requirements of early man. Primitive societies spent most of their time hunting and gathering, and hence, their energy cycles were extremely simple: human energy for the pursuit of game and the manufacture of tools and weapons; wood collection for cooking, shelter heating and lighting, was also an important activity. Food from the hunt was the basic fuel for human energy while the leftover animal grease also served to heat and light the shelter to maintain a suitable microclimate.

Over the years, early humans incorporated new options to fulfill the basic requirements of survival, and at the same time altered the energy cycles of primitive societies. Gardening was added to fishing and hunting as a source of food. Additional human energy was put into planting, weeding and harvesting. Products from gardening provided a more predictable source of food energy for humans and the basic feedstock for raising domestic animals, which in turn became the source of high-quality protein and eventually an additional source of power. Gardening evolved into agriculture, which, along with animal husbandry, eventually displaced hunting as the main activity for survival. At some point in time, domesticated animals were incorporated to take away from humans the burden of load-pulling and back-carrying.

A time came when technology was developed to simplify everyday productive activities. Along with the dawn of progress came larger and larger requirements for energy. It is estimated that early humans had a daily energy consumption rate of around 2500 kilocalories [1]. Later on,

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

in primitive agricultural societies, which already had some domestic animals, this rate was around five times as large [2]. By the time of the low-technology industrial revolution in the mid-1800s, per capita daily consumption of energy reached 70 000 kilocalories in England, Germany and the United States [3]. During that period fuel wood and coal constituted the main sources of energy with smaller contributions from petroleum and hydropower. In the last quarter of the twentieth century, the dominant energy sources switched to petroleum, natural gas, nuclear energy and coal, while the average per capita consumption of energy in industrialized nations rose to over 230 000 kilocalories per day [3], two orders of magnitude larger than that of primitive humans!

24.1.2 "Let There Be Electricity"

With the early studies on electricity in the nineteenth century and the eventual development of the electric power industry, around 1882, the face of the Earth was changed forever. Electric lighting began flooding the cities at night and electric motors became the main source of power in factories. Earlier, in 1846, long-distance communication had been made easier and faster with the introduction of the electric telegraph. Over the years, many inventions based on electricity increased the productivity in factories and made life at home easier and more comfortable.

Modern life has become an endless chain of activities and events, fuelled by electricity. The world is full of electric gadgets, appliances and equipment. Electricity has also been the main element in improving the quality of basic services for the wellbeing of people, such as clean water, education, medical care, entertainment and modern means of information and communication such as the Internet.

24.1.3 One-third of Humanity Still in Darkness

Unfortunately, even today, not everybody has the fortune of enjoying all the benefits of progress: about one-third of humanity lacks access to electricity and, therefore, to a large number of electricity-based services and commodities. Around two billion people, mostly in the so-called developing countries, have remained in the earlier stages of human development, and still rely on wood fire, animal grease or kerosene lamps to light their paths and their homes at night. Modern means of communication and entertainment are either not known to them or are a distant possibility. Millions of people die every year from drinking polluted water, while others suffer from the lack of basic medical services. Illiteracy denies millions of people any possibility of gaining access to better opportunities. As hard as it may seem to believe, at the onset of the twenty-first century, with all the technology mankind has been able to create, survival is still the name of the game for millions and millions of people in remote rural areas of the world.

Reasons for such disparity are manifold, but certainly access to reliable, affordable and high-grade energy sources is one of them. The direct relationship between the per capita energy consumption and human development is well established, as can be observed in a number of studies [4, 5]. What is not too clear is what form of energy, how much of it and for what applications, is required to break the vicious circle of underdevelopment. However, experience shows that a little electricity properly applied could help resolve many ailments of society.

Photovoltaic technology has proven suitable to do at least part of the job. Hundreds of thousands of small individual PV systems are now supplying electricity to facilitate medical care and to light houses, schools and other communal buildings in small remote communities in all five continents. Means of electronic communications, including satellite telephones and the Internet, as well as clean water supply, are now being made accessible to growing numbers of individuals, thanks to the availability of PV. Similarly, productive activities in otherwise stagnant communities are

being facilitated through the use of PV-powered water pumps, small power tools and local shops. However, the rate at which this technology is being introduced to relieve other communities in developing nations is slow, and unfortunately decreasing, as multilateral and local government programs to this effect decline, and other, more profitable markets, such as large-scale grid-connected PV power plants and applications in the urban setting, lure product suppliers in that direction.

24.1.4 The Centralized Electrical System

Electricity is certainly the most sophisticated and flexible form of energy in use today around the world. But it has some drawbacks: electricity has to be used almost immediately after it is generated, as storing it may be expensive, time-limited and inefficient, and in the current scheme of supply, it has to be transported over long distances from the point of generation to the point of use, which can be inefficient and unreliable, especially when these two points are too far apart from each other in places lacking support infrastructure.

In the early days of the electric power industry, electricity was generated right at the point of use. Around 1880, even street lights in places such as Paris and London had their own individual generators [6]. Electricity was also then mostly generated using local and renewable sources of energy. Water wheels originally used in the factories to mechanically power process machinery, were later retrofitted into prime movers to turn electric generators, which in turn began powering electric motors in the late nineteenth century.

As the demand for electricity increased because of industrial and urban growth, and the distance between the point of generation and that of use became increasingly large, electric companies searched for new ways to deliver their services within good profit margins. Engineering research focused on alternatives to increase the power and hence the scale of the generating stations and the carrying capacity of transmission lines. Thus, the concept of economies of scale was introduced in the electric power industry, which for over one hundred years has influenced decision making for new investments in electric systems.

Since the generating units grew in size, electric companies often found themselves with excess generating capacity at the end of the construction of a new plant. The need to recover their investment in this excess capacity frequently motivated them strongly to look for new customers. Transmission and distribution lines were extended to reach the new customers, so an extensive grid was eventually created. Sometimes when demand was nonexistent, it was artificially created. For such purpose, donation of appliances was, at times, made by the electric company to the customer.

24.1.5 Rural Electrification

At some point in time, agricultural processes were identified as potential applications for electricity and, consequently, the lines began to be extended into the rural areas. Here, the density of potential clients and the intensity of electricity use was not as large as in the cities or the industrial centers; therefore investments in grid extensions became hard to recover, so new institutional and financing mechanisms were developed to support the operation. Official rural electrification programs were introduced in the most advanced nations, an initiative that eventually trickled down to the developing countries.

As rural electrification proved beneficial to developed societies, early policy planners felt that the same or similar benefits could be achieved in developing societies. Thus, a major effort

was undertaken in the 1960s and early 1970s to extend the electrical lines into the rural areas of developing countries. However, by the end of the twentieth century only a few developing nations had reached an acceptable degree of electric grid coverage in rural areas. The rest could not advance much as a result of a number of problems faced by the electric utilities, including lack of capital to finance additions of capacity and grid extensions. Thus, the process of rural electrification through grid extensions in many developing nations stalled, to the point that the problem of rural electrification again became a major political issue around the world.

24.1.6 The Rural Energy Scene

Life in many rural areas of the world is no different today than it was centuries ago. Even energy cycles resemble those of early humans, albeit with the introduction of some modern elements. Firewood remains the main source of fuel in most rural communities, in spite of the alarming deforestation, dangers to health and the amount of work needed to collect it. In most places firewood is used mainly for cooking, followed by shelter heating and lighting. A good reference point at hand is the case of Mexico, one of the most advanced economies in the developing world. Here, firewood consumption in 1997 represented around 2.7% of the total energy supply, a share that is larger than that of coal and nuclear electricity taken together and almost equal to that of hydroelectricity [7].

Kerosene lamps and votive candles are more convenient than firewood for lighting, as they provide a more steady and whiter light, and can easily be carried from place to place, thus serving as a portable means for lighting pathways at night. Candles are frequently available in nearby towns, are not too heavy to carry in moderate quantities for long distances, and can be stored for long periods of time. Almost the same can be said for kerosene, which can be used in rudimentary lamps for lighting, although its availability may be more geographically restricted than that of candles. Getting either candles or kerosene requires money, which imposes an economic burden on poor rural families, while their use poses the risk of fire due to the flammable nature of the materials oftentimes used in the construction of rural houses.

24.2 BREAKING THE CHAINS OF UNDERDEVELOPMENT

24.2.1 Electricity Applications in the Rural Setting

The current patterns of energy use in rural areas show that the provision of small amounts of energy, especially electricity, changes the lifestyles of the rural population significantly. Energy applied to improve quality of life of the population may be a good first step to break the chains of underdevelopment. Applications such as lighting, clean water supply, entertainment and communications, preservation of vaccines and other medical supplies, and means for modern education, are usually welcome by governments, aid development agencies and the rural communities themselves. A number of domestic tasks, mostly done by women, such as the provisioning of water, grain grinding, clothes sewing and others, could be made easier with the provision of small amounts of electricity.

Electricity for the household is at the top of the shopping list of rural communities. A house with electricity is a symbol of status. Beyond that, electrical lighting facilitates movement at night inside the house, helps prevent accidents, eliminates the need for kerosene and other combustible materials for illumination (thus avoiding the risk of fire and health-damaging fumes), helps the occupants to spot poisonous insects and other wild animals that could be a threat, and

helps respond instantly to critical situations in case of accidents or illness. Electricity in the house also makes modern means of entertainment a more realistic possibility.

Electricity is instrumental also in supporting communal life. Illumination of external areas helps promote social interaction and a number of after-hours outdoor activities. Social, civic and religious gatherings benefit from the availability of loudspeakers, VCRs and music players. The rural telephone system gives people the opportunity to keep in touch with their relatives in other parts of the world, and to call for help in case of emergency. In many instances, electricity for communal services is given higher priority by the community than domestic services.

24.2.2 Basic Sources of Electricity

With the introduction of the transistor radio and the handheld flashlight dry cells became a favorite means to provide light and entertainment in rural areas. Dry cells can be purchased in many places and are easy to carry. Thus, many rural families spend substantial amounts of money on them. Transistor radios play an important role in the life of remote communities, not only because they bring music and entertainment, but also because radio broadcasts in many places carry important messages such as warnings of floods, instructions on health practices and other valuable services, such as family to family message delivery. Some countries have even set up radio stations with regional broadcasting in the locally spoken native language or dialect, when it is different from the official national language.

New forms of entertainment, such as portable televisions, VCRs and CD players, have increased the demand for electricity in many rural communities. Because of this, dry cells prove to be expensive, and hence many users in rural areas not connected to the grid have resorted to the car battery as the power source for their needs, including home lighting. Car batteries are widely available in rural areas of many developing countries. They are rechargeable, and because of their relatively larger power capacity, they can be applied to a wider variety of services; they last longer and may turn out to be cheaper per unit of service delivered than dry cells. Recharging car batteries, however, requires a primary source of electricity. Where motorcars or tractors are available, people use them to recharge batteries. Otherwise, batteries are carried over long distances to the nearest source of electricity for recharging. However, batteries are heavy and burdensome to carry, so transport to the point of recharging is sometimes done on the backs of beasts of burden and sometimes on the backs of humans. There have been cases of local entrepreneurs setting up micro-businesses to offer battery-recharging services. In this model, batteries are collected, taken to the point of recharging, recharged and then returned to the owner. This operation however, requires some infrastructure such as roads and means of transportation, not always available in rural areas.

As the economic power of families increases, so does the need for electricity. Dry cells and car batteries are no longer sufficient, so people in many places put political pressure on electricity authorities to extend the grid to their communities, or individually resort to the use of small gasoline-fuelled generator sets. Small generator sets have the advantage of supplying alternating current of the right voltage, so that conventional appliances can be used. But their service is usually restricted to a few hours a day, because of the increasing cost of operation and maintenance of the equipment, plus the difficulties of getting and transporting the fuel.

Experience shows that, even when grid electricity is available, people in rural areas normally use it to light a few bulbs and perhaps to power small radios or typically black and white TV sets. Most people in rural areas lack the money needed to buy larger electric appliances, such as refrigerators and washing machines, or are simply not acquainted with them. Hence, because the number and size of appliances now in use is small, rural electrification represents an important niche of opportunity for the application of photovoltaic technology.

24.3 THE PV ALTERNATIVE

Photovoltaic (PV) technology nowadays is considered one of the most appropriate options to electrify dispersed population in remote places [8]. From an engineering point of view, modularity is perhaps the single most attractive feature of this technology. It allows designers to tailor electricity-generating systems as small in capacity as a few watts, or as large as many megawatts to suit specific needs, just following basic rules of electrical engineering. This feature combined with the suitability of the technology for autonomous operation, producing electricity with locally available sunshine, plus other characteristics such as lightweight, low-maintenance requirements and long useful life, has led people to consider photovoltaics as an attractive option for rural electrification. Ever since the technology was applied to power space satellites in the mid 1950s, the concept of reliable electricity generation for remote applications was firmly established. Terrestrial applications were developed with the basic idea of powering loads in remote places, where the cost of extending the grid was just too high. Today, hundreds of thousands of PV systems have been installed around the world to substitute for candles and kerosene lamps, gasoline- or diesel-powered generating sets, or even for unreliable grid extensions. The types of applications of PV technology in remote areas range from telecommunications, lighthouses and alarm systems in certain industries, to domestic applications and delivery of basic services. Leisure applications, such as sailboats and mountain cottages now carry PV panels to provide the required amounts of electricity, instead of noisy and smoky combustion engines. Even the petroleum industry, which is nowadays the basis for the world's energy supply, is using photovoltaics to supply electricity in offshore rigs or to power remote valve stations in oil and gas ducts.

The advantages of PV technology for rural electrification were demonstrated through a number of early projects in the period between 1968 and 1977 in Niger, Mexico and India. Applications included PV-powered educational television, telephones, medical dispensaries and boarding schools for native Indian children. This early work demonstrated not only the technical viability of the systems, but also the benefits to the user [9–11]. Some of these installations are still operational and in good condition, although with the limitations of a 30-or-so-year-old technology. Unfortunately, a critical mass of early projects was never achieved as to make a noticeable impact on society, and the lessons derived from the few projects on record have mostly fallen into oblivion. Only the notion that the technology was too expensive and not too reliable, has prevailed over the years among many decision-makers, in spite of the tremendous progress PV technology has made in recent years in terms of cost, efficiency and reliability.

Progress in materials technology, electronics and PV systems engineering, along with a drop in price of the main components of the PV system and a better understanding of the needs and expectations of the rural people, have resulted in a large variety of ideas, proposals and technological schemes to use photovoltaics as a source of electricity to promote human and economic development in rural areas of the developing world. A large number of applications have been identified for PV systems in rural areas, some more mature than others in technical terms, but all facing similar problems to enter the rural market on a massive scale.

24.3.1 PV Systems for Rural Applications

PV systems for remote applications typically include three basic elements: one PV panel that converts solar radiation into electricity, a means to store the electricity produced by the PV panel, which is normally an electrochemical battery, and an electronic device that helps control the flows of electricity within the system, thus protecting the battery by properly dispatching available energy. A variety of devices capable of using electricity to provide comfort, entertainment and other services for the benefit of the user are then attached to the PV system by means of the electronic charge

controller (ECC). Figure 24.1 shows a schematic diagram of a general PV system. Depending on the application, some elements of the PV system may not be necessary, as is the case of the battery in water-pumping systems.

PV systems can be designed and built in the stand-alone mode, in which photovoltaics is the only source of electricity, or as hybrids, in which photovoltaics is combined with other sources of electricity, such as wind generators, small hydropower stations or combustion generators. PV hybrids are normally built to take advantage of other locally available renewable energy resources while at the same time improving the economics of the application. Photovoltaics can be installed in the "disperse" mode, in which each individual application carries a full PV system as its source of electricity (just as it used to be in the very early stages of the electrification process in the late 1800s). Larger PV systems can be built to feed isolated electric minigrids, as is done today with diesel gen-sets in many parts of the world. Since PV panels produce relatively low-voltage direct-current electricity, a minigrid system requires additional components, such as DC/AC inverters and step-up transformers, to yield the right characteristics of the electricity on the user's side of the grid. Hybrid systems are also more complex to integrate, since different types of electricity can be produced by the different generating units that are included in the system. Figure 24.2 shows a number of possible configurations of PV systems, for most of which practical examples can be found.

Solar home systems (SHS) are perhaps the most popular of all the PV applications. An estimated 500 000–1000 000 such systems have been installed in rural communities around the world [12]. Systems for basic lighting generally include one small, 10–40 W, PV module and a small battery, enough for one to four points of light (normally compact fluorescent lamps). Larger, 50–100 W PV panels and batteries of around 100 A h open the possibility of feeding other electric appliances, such as transistor radios, CD players and small TV sets. Even larger PV systems can support complete sets of domestic electric appliances, just as in any urban house, but the price of such systems is at present prohibitively expensive for a poor family. The smallest SHS are direct substitutes for dry cells and other ancient means of lighting a house, and a form of using electrochemical batteries without the need of sending them elsewhere for recharging.

Quality and efficacy of communal services can substantially be increased when relatively small amounts of electricity are available on site. This can be done by photovoltaics. Schools can be supplied with modern audiovisual means, educational television, and even the Internet, as shown by the project *aldea solar* in Honduras, supported by UNESCO, the Honduras Ministry of Education and the Council for Science and technology (COHCYT) of Honduras. In Mexico, over 13 000 PV powered telephones now link tens of thousands of people in remote communities with the rest of the world, through terrestrial PV powered transmitting stations and satellite links [13]. In Cuba, the Ministry of Health has implemented a system of rural clinics powered by photovoltaics, which has been operational for a number of years [14]. PV-powered technologies to disinfect water have already been tested and field-demonstrated in countries of Africa and Latin America in a project financed by the European Commission [15]. A large number of PV water-pumping projects have been implemented around the world, and many examples of other PV powered communal services and applications can also be found (see for instance references [16–20]).

A large number of possibilities to apply photovoltaics for productive activities in rural areas of the developing countries can also be envisioned. Small irrigation, cattle watering, grain grinding, small handicrafts shops and other similar activities that require relatively small amounts of electricity can now be powered by photovoltaics to increase productivity and foster economic development.

24.3.2 Barriers to PV Implementation

For most urban people around the world, electricity comes into their homes just like magic: it is there, instantly and reliably at the touch of the switch. Few individuals make a conscious connection

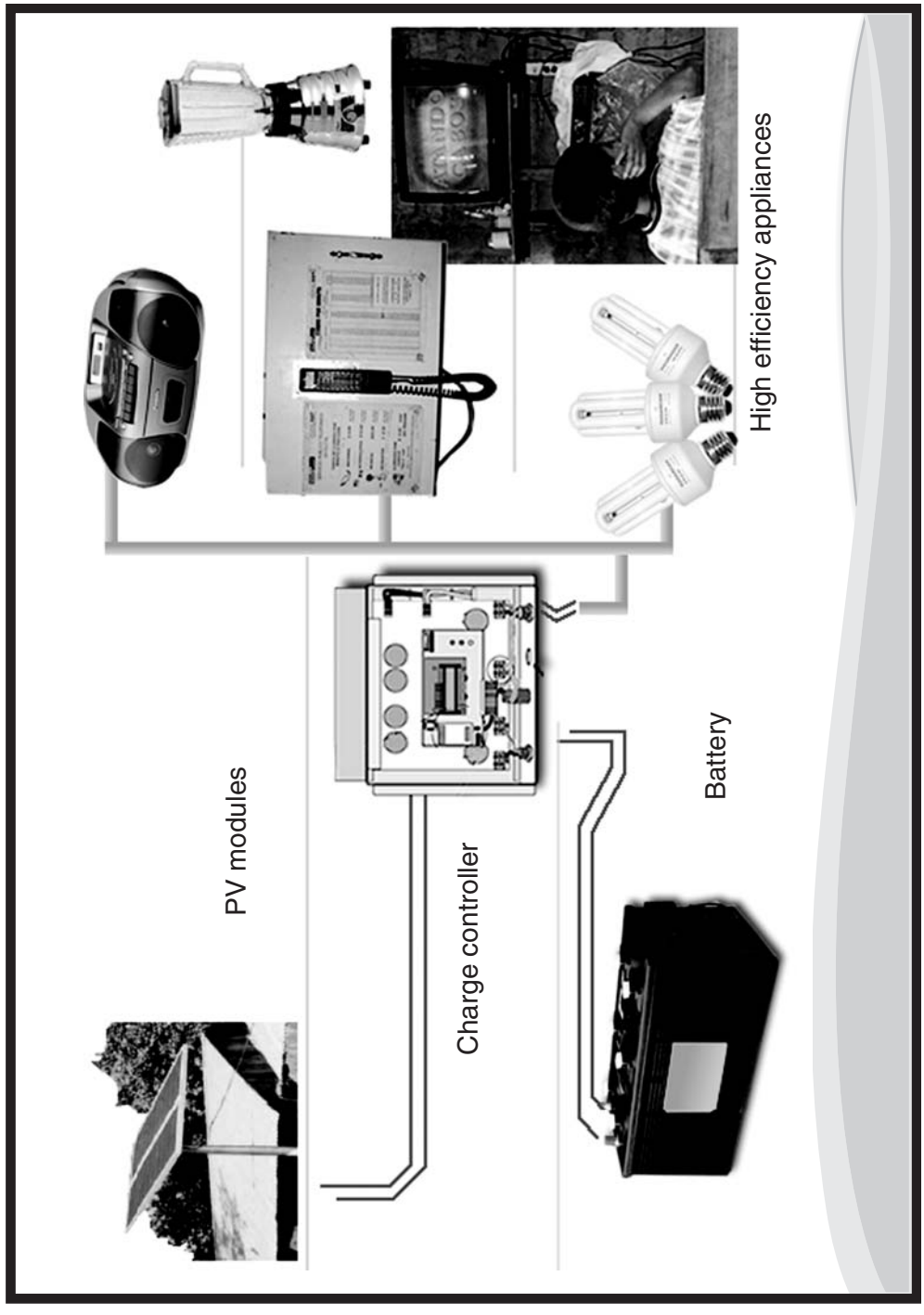


Figure 24.1 Schematic diagram of a general PV system

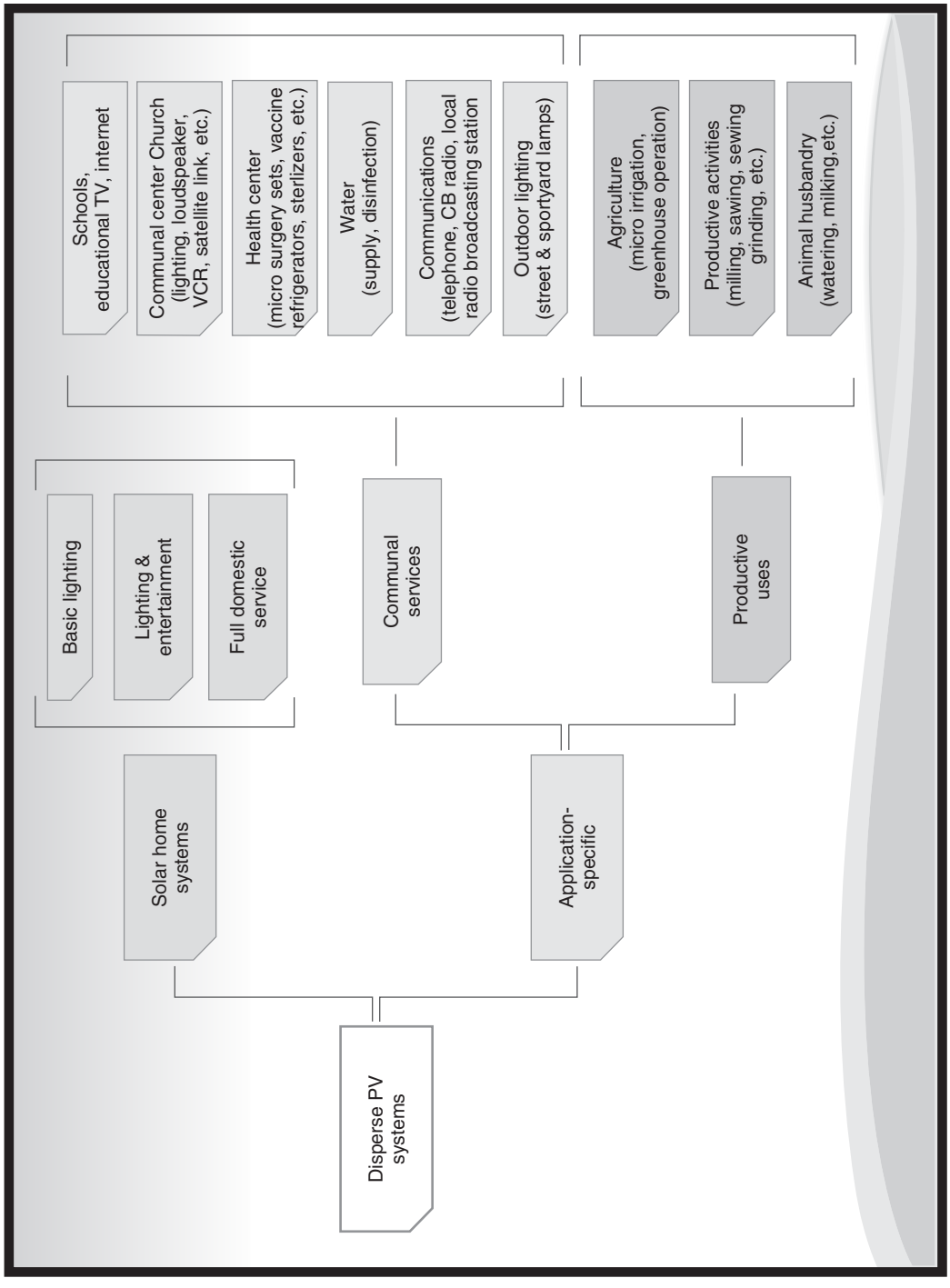


Figure 24.2 Possible configurations of PV systems

between their appliances and the electricity pole in the street, so paying their monthly bills is perhaps the closest they get to the electricity business. But even fewer people realize the technical and administrative complexities behind the process of generation, transmission and distribution of electricity that allows factories to run and people to enjoy the benefits of modern services.

Understanding the physical principles that turn primary energy into electricity belongs to a small group of technical elite in the universities, research centers and electric companies. For the layman, it makes little difference whether primary energy is hydropower, nuclear power, fossil fuels or solar energy. Thus, the fact that electricity can be locally produced using the sun's rays as the primary source of energy, with no wires connecting to remote and unknown places and facilities would be of significance only to the most knowledgeable people. It is interesting to know that PV users in native communities in many parts of the world make a cosmogonical connection between electricity and their ancient god, the Sun. Thus, for most of these people PV technology is an appealing means to get a long-awaited service.

However, after learning about photovoltaics people are inclined to ask why is it that, with so many virtues and so many advocates, PV technology has only reached one-tenth of a percent of the world's rural population with no access to the grid? The answer is found in the number of barriers a new technology such as this has to overcome to fully enter the market. In the case of photovoltaics, some such barriers are well known, others still unknown; some technical in nature and others having to do with institutional, social and financing issues. Just as conventional electricity is generated, in large and distant facilities, transmitted and then distributed to reach the individual consumer, so is PV technology produced in a small number of facilities in Europe, Japan and the United States, transported across the continents and distributed to reach the final user in very remote rural areas. And at each step a number of operations need to take place, which involve different degrees of complexity and cost. Therefore, bringing the PV solution to those sites where the grid has not been able to reach can be a very difficult task, unless such barriers are successfully removed.

24.3.3 Technical Barriers

PV systems are claimed to be reliable and long-lasting. This is true and proven insofar as the PV module is concerned, but not every component of the system's balance has the same degree of technological maturity. In both, stand-alone and hybrid systems, batteries are perhaps the weakest links. Batteries are exposed to overcharging and over-discharging, which usually reduce their useful lifetime. Batteries also demand a fair amount of attention and regular maintenance, albeit relatively simple. However, even the simplest technical tasks may prove to be complicated in rural areas where a high degree of illiteracy and a lack of familiarity with modern technology, and with electricity in particular, is more the rule than the exception. The problem with batteries arises in part because of the fact that the lead–acid technology now in use, available for over one hundred years, was not specifically designed for use in PV systems. Adaptations of the current technology into what is now called *solar batteries*, and development of new battery types, such as nickel–metal hydride [21], promise to ease many of the current problems.

SHS also face other problems, mostly in the charge controllers and the lamps, basically as a result of an industry that has until now had an uneven degree of development (for a detailed discussion on charge controllers, see reference [22]). Two schools of thought seem to underline the issue: simple, sturdy, low-tech, low-purchase-cost devices against high-tech, sometimes a bit complex and perhaps lower life-cycle-cost devices. The issue is not easy to resolve, especially with so little systematic information coming from the field on the performance of such devices. Taking into consideration that rural areas in developing countries are far from being mature markets for photovoltaics (and for many other goods) and accounting for the illiteracy of the rural population, consumer choice will hardly be a useful parameter to resolve the issue, especially if one considers

that a large number of PV rural electrification projects are still being carried out in the "technology push" mode.

The suitability of photovoltaics as a solution to the rural electrification problem is being taken for granted by many advocates of the technology. Unfortunately, few studies seem to have been carried out to assess the performance of SHS now in the field, in a systematic and comprehensive manner. At this stage of technology implementation, information from the field is vital as a feedback mechanism to gauge the efficacy of the PV solution, to improve the chances of overall success and to assure long-term sustainability. Field surveys, however, tend to be expensive, especially where the most remote and isolated communities are concerned, and availability of funds for monitoring and evaluating SHS projects in the field is not obvious.

A 35-man-month field study was done in the recent past in Mexico, in which 1740 SHS installations (out of around 60 000 installed with government financing) in most regions and communities included in government programmes were evaluated. The study had a threefold purpose: to assess the physical and operative condition of the systems, to probe the degree of satisfaction of the users and to evaluate the efficacy of measures previously implemented to make the projects sustainable. Preliminary analysis of the information gathered in the study shows that, from the technical point of view, things look good, with most SHS samples performing well (for more details on these results, see reference [23]). But there are reasons to believe that as systems age, the results may change, unless corrective measures are taken.

Introducing photovoltaics in rural areas of developing countries is an innovative exercise in society, with the particularity that a space-age technology is being adapted for operation in a sector of society living, in many cases, at least five hundred years in the past. From this perspective, it is hard (and even dangerous for project sustainability) to ignore the strong connection that must be established between the technology (hardware) and the user. For even the most sophisticated, well-designed and perfectly built piece of PV technology is bound to fail sooner than later, if the ground for seeding it is not properly prepared. This means information, training and local capacity building on the user side, as well as user involvement at every step of the process to make them aware of the important role they play in the solution of their own problems. Similar considerations can be made in connection with the environment (social and physical) in which the PV system is bound to be installed. For instance, anecdotal and written information from the field points to the fact that some PV components designed and built with technical criteria prevailing in advanced, cold countries are not performing well in the tropics where they are most needed. He reasons for this are many, but their discussion is outside the scope of this chapter.

Alternative technical schemes to the SHS as a means to provide basic electricity services in rural areas are being tested. Battery-charging stations, where a large central PV array is installed to recharge batteries brought in by the user, who pays a fee for the service, is one such scheme advocated by a number of people. The rationale behind this scheme is that peasants in many parts of the world already use battery-charging points distant from their homes. It is also argued that micro-businesses could develop in remote areas to enhance the chances of sustainability of the electrification process and so, several projects of this kind are underway in various countries. However, field studies reveal [24] that this alternative also has several shortcomings, which are forcing project officials and users to switch to SHS as a substitute to the previously used PV battery-charging stations.

The preceding discussion was meant to point out the fact that, even though engineers and industries will most likely solve the remaining technical problems facing PV technology for rural applications before this market enters into its mature stage, a number of more complex issues at the technology–user interface remain to be understood and dealt with. The cause–effect diagram of Figure 24.3 is an attempt to show the variety of factors leading to a failed system, and consequently an unsatisfied user. In the long run, the degree of user satisfaction will determine the level of acceptance of PV technology as the solution to the problem of rural electrification.

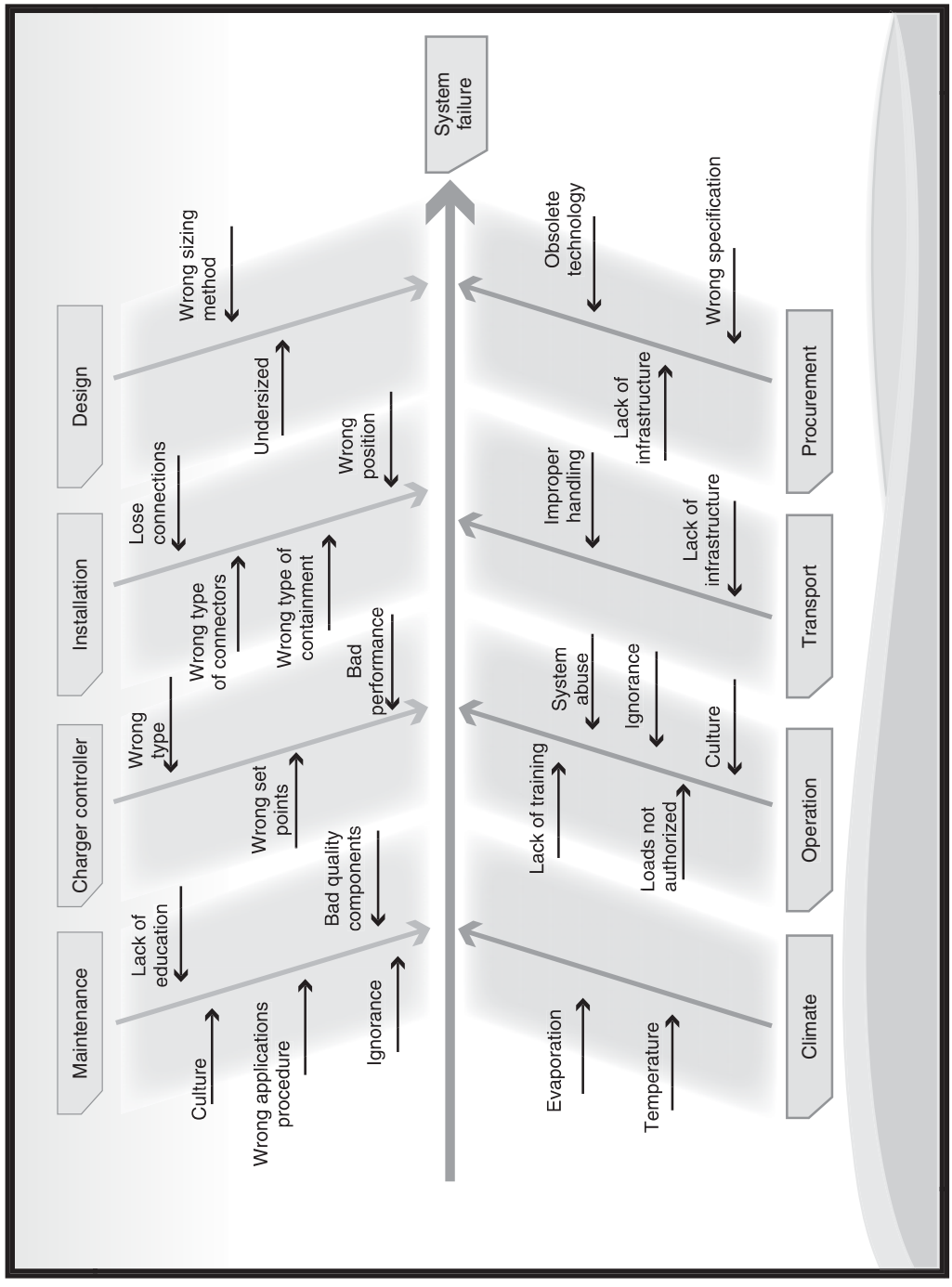


Figure 24.3 Cause–effect diagram showing a variety of factors leading to a failed system

Technical standards, design guidelines and other elements for quality assurance of PV systems are also important for the sustainability of PV rural electrification projects. A number of these elements are being developed in institutions, professional societies and international organizations around the world. Most of them, however, focus mainly on pure technical (hardware) issues, as is the case of the recommended specifications of the organization Global Approval Program for Photovoltaics (PV GAP) and the standards issued by the International Electrochemical Commission (IEC) Technical Committee 82, and little attention is being given to the soft and organizational aspects of the problem dealt with in other works [25, 26].

For a number of years rural electrification was the main market for PV, fuelled by assistance programs from governments, aid agencies and multinational organizations. However, as such programs come to an end, money for rural projects gets short. Clearly, lack of financing is an important barrier to market entry and sustainability. On the other hand, the growing demand for larger power modules for grid-connected applications has led to shortages of smaller modules for the rural market. Implementing products distribution and after-sales services chains in the rural setting is not an easy endeavor, especially in regions lacking roads and other basic infrastructure. Hence, in general terms, the rural market for PV faces a number of barriers germane to its own nature.

24.3.4 Nontechnical Issues

24.3.4.1 Initial cost

It is common knowledge that for a commercial operation to be sustainable in the modern economy, a flow of goods has to be properly matched by a counter-flow of money. Solar rays striking on the roofs of houses are free, but equipment to turn solar energy into electricity, and to transform this electricity into needed services, is not. PV manufacturing companies invest in factories and raw materials, pay wages to their workers and taxes to their governments, and are obliged to deliver revenues to their shareholders. All expenses plus profits essentially set the base price of their products. In a second step, PV components from different companies are transported to specific points for systems integration. In turn, packaged systems are fed into the distribution channels for retailing and final installation where the end user wants them.

By the time a PV system is installed on the user's premises, its base price has increased a number of times. So, people need money to get their systems installed. Considering the low capacity of the peasants in rural areas to pay for goods, a set of important questions emerge when one considers photovoltaics as the solution to the rural electrification problem. Are people willing to pay for the system? How much can they afford to pay? What mechanisms can be instituted to make systems more affordable? If people cannot pay, should they remain in darkness or should somebody come to their rescue? What roles can governments and development agencies play? And even when people can pay, should they bear the full system cost, even at this early stage of technology introduction when many companies are still building their infrastructure and learning how to manufacture, integrate and market the systems, so that their transactions costs are a lot higher than they should be? These are not trivial questions considering that, for the miracle of full-scale rural electrification to happen, around 300 000 million US$ must, in principle, flow from the poorest regions of the world to the modern sector of society so that PV systems can flow in the opposite direction.

24.3.4.2 Breaking the initial cost barrier

Consequently, since solar energy, the fuel used by PV systems, is free and systems are, at least theoretically, low-maintenance and long-lasting, system cost is commonly seen as the main stumbling

block for the introduction of photovoltaics. A number of schemes have been tried over the past decade in search of an effective way to remove this barrier. Most of them can be grouped into three categories: the *social route*, in which poverty alleviation programs and other socially driven mechanisms are used by governments and bilateral aid organizations to make funds available for the purchase of PV systems in favor of the least-privileged people; the *fiscal route*, in which taxes, import duties and other fiscal levies are removed to lower the local price of the PV system, thus making it somewhat more affordable for the final user, while at the same time facilitating the creation of a local market; and the *business route*, in which banks, private companies and entrepreneurs are devising and testing a number of schemes to make financing available for the purchase of the PV system, thus helping to create a market for photovoltaics at the same time. One way or another, currently each of these routes somehow benefit from the intervention of governments, multilateral organizations and lending institutions. On the other hand, boundaries between these routes are not necessarily clear-cut, so that combined or coexisting schemes are not uncommon.

Not everyone in rural areas is necessarily poor or totally dispossessed. Some people live in remote places for convenience, either because their source of income is attached to the natural resources locally available, or because they prefer living in a cleaner and quieter environment than in the city. They do not have electricity from the grid, simply because they are too far away, but they usually rely on gasoline or diesel-fuelled gen-sets for their electrical service. These people, however, could easily purchase a PV system without any financial assistance, if one was made available to them. This is already a good market that is being tapped in a number of countries, such as Spain, Colombia, Mexico and others. On the other hand, some estimates indicate that between 25 and 50% of people living in remote places could purchase a SHS provided some sort of financial assistance was available to them [27]. This is a substantial market in its early stages of development; to tap this market a number of schemes are being tried by private entrepreneurs and multilateral development organizations, as is the case in Kenya, Zimbabwe and the Dominican Republic. However, the poorest of the poor, representing the largest portion of the world's rural population, can hardly ever afford to buy their own PV systems.

24.3.4.2.1 The business route

For those in need of financial assistance, two alternative models are being tested. One focuses on the sale of the PV system (the sales model), the other on the sale of the electricity produced by the system (the service model). Both models have advantages and shortcomings and both involve flows of money beyond the control of the PV user.

In the sales model, the PV system is purchased on credit by the user, who becomes the owner and takes over the responsibility of system maintenance and replacement of parts. Money for the transaction is usually borrowed by the user from a set of different sources, which may include the system supplier, a finance institution, or any other type of credit organization such as a revolving fund or a local micro-finance operation. In any case, money is obtained at a cost, which adds to the cost of the PV system, and is paid back in periodic installments under prearranged terms and conditions. However, this additional cost is what entitles the user to the benefits of electricity, albeit in small quantities. This model is preferred by those who like the social status of system ownership and are willing to assume the task of maintaining the equipment and the responsibility for replacing damaged or worn-out parts.

People not willing to take any risks, may opt for the service model, in which the supplier retains ownership of the PV system, or of some parts thereof, and then charges some monthly fee for the electricity delivered to the user. The supplier maintains the system and becomes responsible for providing the user with electricity, according to system capacity. Following the current

practice, this kind of service can be provided through a regulated concession, an unregulated open market provider or a community-based provider. Needless to say, the risk assumed by the service supplier also has a cost, which reflects on the monthly fee charged. According to some estimates, the added monthly fee for service over a period of 10 years could double the life-cycle-costs of the same PV system when purchased on credit [27]. Current PV projects in the fee-for-service model show a large dispersion in the amount of the monthly fee charged to individual users. A number of factors may be responsible for this. For instance, some projects may involve some sort of subsidy, and in other cases, service suppliers may not follow the same rules to define the scope of their responsibility: some may retain responsibility over the PV module and charge controller only (the least troublesome parts of a SHS), while dumping on the user the responsibility of replacing the battery (the weakest link in the system) and the lamps.

Selecting a particular route for a given project is not a matter of personal preference only. The landscape is very important and has to be taken into account, since a number of its elements influence this choice, including local social and energy policies officially established in each country. For instance, some countries do not allow private sale of electricity, so that the fee-for-service model could not be applied, unless prevailing laws and regulations are changed as necessary. Low rural population density, difficult access to the communities, long distances from the supply centers and complex logistics to deliver goods and collect fees, can make both fee-for-service and financed sales schemes difficult to implement and geographically limited. In many cases, monthly fee/payment collection may be an expensive task, due to the time and effort it may take for the supplier or the service men to reach the customer. This fact has already been noted by PV companies operating in Sri Lanka, for instance [27]. Timing is another important factor, considering that people may not be home when the fee collector arrives, and that many peasants have money only during the post-harvest period. Such difficulties have led some utilities, for many years, to give up on monthly fee collection from remote clients connected to their grids, basically because of very small bills and comparatively larger billing expenditures. Thus, unless fee collection schemes better adapted to the local conditions can be found, one can theorise that the fee-for-service and purchase models will be applicable basically to the most accessible and higher density rural communities. The fraction of the rural population meeting these criteria in many countries seems to be inversely proportional to the degree of rural electrification by grid extensions.

Local culture and idiosyncrasy are two additional elements to be considered for the choice of a delivery model. The notion of common property embedded in many native communities or the lack of familiarity with commercial practices, or even with money, could lead to unsustainable operations in many rural areas. Some academics, studying the process of introduction of PV systems in rural communities, argue that freedom from any financial burden is one of the most cherished values for rural people, and hence could be an important barrier for them to take on any financial obligations. Hence, government intervention may be indispensable at some point to deliver the PV solution in which the business route cannot be applied.

Over the past ten years delivery models in the business route have been tested through a number of projects financed by the GEF and the World Bank in various countries. A review [12] of the GEF solar PV portfolio suggests the following emerging lessons, warning that it is still too early to draw definite conclusions:

- Viable business models must be demonstrated to sustain market development for solar photovoltaics.
- Delivery/business model development, evolution and testing require time and flexibility.
- Institutional arrangements for project implementation can greatly influence the value of the project in terms of demonstrating viable business models and thus achieving sustainability.
- Projects must explicitly recognize and account for the high transaction costs associated with marketing, service and credit collections in rural areas.

- Consumer credit can be effectively provided by micro-finance organizations with close ties to the local communities if such organizations already have a strong history and cultural niche in a specific country.
- Projects have not produced adequate experience on the viability of dealer-supplied credit under a sales model, and no project in the portfolio appears set to provide such experience.
- Rural electrification policies and planning have a major influence on project outcome and sustainability, and must be addressed explicitly in project design and implementation.
- Establishing reasonable equipment standards and certification procedures for solar home system components that ensure quality service while maintaining affordability is not difficult, and few technical problems have been encountered with systems.
- Substantial implementation experience is still needed before the success of the service approach can be judged.
- Post-project sustainability of market gains achieved during projects has not yet been demonstrated in any GEF project: it is too early in the evolution of the portfolio.

24.3.4.2.2 The social route

Governments can act in several capacities along the social route. Financing the purchase of SHS for poor people is perhaps the most critical one, although consumer protection and market regulation are also of importance. Government financing of SHS is seen by some as an unnecessary nuisance that distorts the market and creates dependence in the users' minds (users will not buy once the government has provided systems for free, goes the argument). Most critics of government intervention seem to forget that rural electrification has been historically subsidized by governments not only in developing countries, but also in some of the most advanced nations. In that sense, there is no reason to think that PV rural electrification has to be different altogether, as it is only the technology base that is changing; the rest of the landscape remains the same.

Direct government financing of photovoltaics for rural electrification is not necessarily a bad thing. Up to now, the largest volume of SHS installed around the world has been realized through government or donor-led programs. Government intervention may help aggregate markets, reduce transaction costs and, if properly done, can create a better setting for quality assurance and local industry development. At least, this has been the experience with government-financed PV projects in Mexico, where proper institutional mechanisms were implemented for this purpose, and as a result, a local industry has also emerged around government-financed projects, which now produces locally and exports balance-of-system components [13, 28–30]. In this case, money provided by the government is seen as an instrument to promote local development. Over 2500 rural communities have been electrified in Mexico with photovoltaics following this model, and it is interesting to note that in many of them people have instrumented a variety of cost-recovery and money-making mechanisms, which allows for system maintenance and additional communal projects.

A number of governments from developing countries are considering the social route to deliver PV-based electricity to rural areas. Policy setting and definition of financing mechanisms are usually the first steps in this direction. Examples can be found in several Latin American countries [31]. Some examples are discussed below.

In Bolivia, Article 61 of the Electrical Law establishes that the State is responsible for the electrification process in smaller localities and in rural areas that cannot be served by private interests. Resources to finance such projects must be provided by the State, through the National Development Fund. The Executive must also propose energy policies and strategies to foster the use of alternative energy sources.

A similar mandate can be observed in Colombia, where Law 143 urges the State to provide basic electrical services to lower-income families in rural areas and to make funds available to

cover necessary subsidies in this respect. The Energy and Gas Regulatory Commission is obliged to protect the rights of the lowest-income people, while the Colombian Institute for Electrical Energy has been mandated to prepare the National Energy Plan for regions not served by the grid, including the use of alternative energy systems in substitution of fossil-fuelled generating systems.

In Ecuador, the Law of Regime for the Electrical Sector addresses rural electrification and financing issues for the rural sector. It also assigns priorities for the application of renewable energy in rural electrification projects and describes the structure for project identification, approval, execution and operation. The National Fund for Rural and Urban Marginal Electrification (FERUM) is the body responsible for the administration of financial resources, and is directly regulated by the office of the president.

Similarly, in Nicaragua, the Law for the Electrical Industry holds the State responsible for developing rural electrification in remote areas in lieu of interest from other economic agents. For this, the State must provide the necessary resources through the Fund for Development of the National Electrical Industry. The State is also required by law to implement policies and strategies for the use of alternative energy sources for electricity generation.

As a last example, in Panama, Article 9 of Executive Decree 22 makes the State responsible for promoting rural electrification and for assigning an annual budget to carry out this task. Consequently, the Office for Rural Electrification was created within the Office of the President, and is in charge of promoting the use of renewable energy for rural electrification projects.

24.3.4.2.3 The fiscal route

Import of PV systems and components seems to be the only way of implementing this option for the foreseeable future in a number of developing countries. In many of them, however, import duties and taxes for electronics and other elements of PV systems are such that the cost of the technology at the user end becomes even higher. Laws and regulations to remove or lower import duties have been implemented in some countries as a way to foster the application of this technology. Such measures, however, are not always appealing, as they are considered detrimental to the government's fiscal revenue, or more to the benefit of the urban consumer instead of the rural user. Import duties are sometimes used as a protective measure to the local industry and there is reluctance in some countries to remove them.

24.3.5 Trained Human Resources

Adequate financing and institutional frameworks are necessary, but not sufficient conditions to remove the main barriers for PV rural electrification. Properly trained human resources to develop and operate programs, and to carry out projects, are equally important. PV systems and their implementation in rural areas are frequently looked upon in a very simplistic manner by a number of people. However, disregard for the complexities behind the process has resulted in a large number of failures. One would be surprised to learn how many PV projects in rural areas around the world have not lasted but a few years beyond the inauguration date; or how many others are still not complete because a variety of logistic aspects were not given due consideration at project inception; or even how many others are under-performing as a result of faulty engineering and construction practices. Unfortunately, little reliable information from the field is available in this respect as many projects are in their first years of operation so it would be too early to draw any definite conclusions about them.

PV systems packaging is increasingly being carried out by local companies in the developing countries. This practice has its merits, since it promotes the use of local labor and materials, which

benefits the local economy. However, workers are not always properly trained to carry out their duties, so that construction and installation guidelines, no matter how precise and elaborate, are not always followed. Sometimes instruction materials are not in the local language, or are translated incorrectly, and at other times they do not properly match the idiosyncratic framework of the local worker. The benefits of training local workers for industry support cannot be overstressed. An example of how far this training could go can be found in the project carried out by the Spanish Cooperation in Bolivia, where native Aymara Indians were successfully trained to assemble electronic charge controllers for SHS. This project is described in greater detail below [32].

Local distributors and vendors of PV systems frequently lack a proper understanding of the products they sell. This leads more often than not to undersized systems, to make the sale easier, or to overselling the attributes of the PV systems to be sold, creating customer expectations beyond the actual capabilities of the system. In any case, the result is customer dissatisfaction. Thus, for the PV business to grow on solid bases, the front lines that deal with end users must be properly trained in technical, marketing, selling and business practices. This is easier said than done, as many PV businesses in developing countries are small commercial ventures embedded in other lines of activity, such as hardware stores or cattle feed stores.

Photovoltaics being a novel technology for most people, developing and implementing programs for its massive deployment are not necessarily an easy task for program managers. Assistance is often required for program formulation, and to establish the proper mechanisms for project financing, implementation and monitoring. Multilateral agencies such as the World Bank, the GEF, UNDP and others, organize workshops and seminars around the globe to disseminate best practices that could help solve this problem. The Network for Rural Electrification with Renewable Energy (RIER) of the Ibero American Program of Science and Technology for Development (CYTED) has carried out a number of courses and workshops on strategies for PV rural electrification throughout Latin America. Attendees usually include officials from government agencies in charge of rural electrification, electric utilities, local financing organizations, PV distributors and salespeople, and university professors. Benefits from such training activities often translate into better-formulated programs and projects, and a more appropriate understanding of the critical elements to make PV rural electrification projects sustainable.

24.4 EXAMPLES OF PV RURAL ELECTRIFICATION

It is estimated that by the end of the twentieth century, anywhere between 500 000 and one million small PV systems had been installed to power rural homes in developing countries [12]. On top of that, tens of thousands of PV-powered water-lifting pumps and other communal services, such as health centers, schools, telephones, street lamps, had been deployed through government programs, donor-led initiatives and entrepreneurial activities. Many such programs have been subject to in-depth reviews as a means to better understand the process that is taking place in the field and to harvest the lessons learned that can be applied to other projects and programs. For details, the reader is encouraged to consult available reports (see for instance references [33–36]). In this section, a few examples of PV rural electrification projects and programs are reviewed.

24.4.1 Argentina

As part of the reform of the electrical sector in Argentina, the provincial electricity markets were divided into centralized and dispersed sectors. Each sector has its own structure, its own mission and its own operational form. Then, the program to supply electricity to the rural and dispersed population of Argentina (PAEPRA for its name in Spanish) was established by the government

in 1994. This program has the goal of providing electricity to 1.4 million people not yet served by the grid, and to electrify around 6000 public services, all this in areas where the low population density and the long distance to the electric grid makes it too costly to supply electricity by conventional means [37]. PAEPRA operates at the provincial level through concessions to private companies bound by contract to provide electricity services under the supervision and control of the provincial electricity regulatory body [38]. Provincial concessions are awarded through a public bidding process, in which the winning company is the one that requests the least amount of subsidy money, on a per customer basis, to operate.

After a period of stagnation, the PAEPRA program is now operational through the PERMER project of the national Energy Secretariat, launched in the year 2000 with funding from the GEF and the World Bank, complemented with funds from the provincial governments, the concessionaires and the users. This new project has the overall objective of supplying electricity to 70 000 rural households and 1100 communal installations, by means of grid extensions and off-grid installations, including PV, wind and micro-hydro facilities. Currently, around 3500 solar home systems plus 620 communal services, mainly schools, have been installed in 15 participating provinces. The implementation process includes establishing collaboration agreements between the provinces and the federal government, implementing contracts between the provincial government and the regional concessionaire, and carrying out market surveys to define the scope and structure of the concession. Currently the province of Jujuy is ahead of the rest in the number of systems installed (1500 SHS and 400 schools). Installation of individual SHS carries a subsidy to cover part of the initial system cost. This subsidy complements the funds collected from the user under a fee-for-service scheme. Other provinces are currently in the early stages of project implementation.

24.4.2 Bolivia

The National Rural Electrification Program (PRONER) was established to promote and support economic development in rural areas in order to improve the living conditions and the quality of life of the population. The original objective of the program was to provide electrical services to 100 000 households in a time span of five years by means of renewable energy. To accomplish this ambitious goal, a number of projects have been carried out since early 1977. It is estimated that the total number of SHS systems installed to date is over 25 000, over half of them implemented after the year 2000 at a rate close to 2000 PV systems per year. However, most installations have gone to meet the needs of rural people with higher purchase power. Among other things, the long process for getting projects financed, the lack of experience in handling large programs of the sort, and a young institutional and regulatory framework for project implementation, made it impossible to reach the original goal.

Most projects in Bolivia have been carried out with financial support from foreign aid agencies and multilateral organizations. Early initiatives include the Rural Electrification Program with Renewable Energy through Popular Participation Process, or BOL/97/G31 as it is commonly known. This program, launched in 1997, was supported by an 8.2 million US$ UNDP-GEF grant and had the goal of installing 3000 SHS through 22 projects in five municipalities. It was aimed at removing financial, institutional, technical and human resource barriers to the massive application of photovoltaics.

Almost at the same time, the Spanish Agency for International Cooperation and the Solar Energy Institute of the Polytechnic University of Madrid implemented a PV electrification project to bring electricity to the Aymara community in the high Bolivian plateau, with the purpose of fostering the development of a users' organization with capabilities to manage and maintain PV installations [39]. By 1993, a total of 1000 SHS had already been installed and the Association

for Solar Electrification (AES) had been created. This project has been studied thoroughly and important lessons derived therefrom [32].

Other relevant projects were carried out in this early period by organizations such as the GTZ with the PROPER program [43], NRECA, the local organization Rural Electricity Cooperative (CRE), who was a provider of complementary financing for the purchase of the PV systems and managed other aspects of project implementation, including the selection of users and providing technical training [40], the Cochabamba Electric Company (ELFEC), a privately owned electricity distribution company, who also financed the purchase of equipment and provided installation and maintenance services, and the local NGO Energetica.

A local model for project implementation has evolved over the years, based on this early experience. Alternative financing mechanisms have been tested around the micro-credit concept, having the PV module as collateral for the loan [41] and participation of government funds. A presidential decree was issued in 2001, establishing the official mechanism to finance rural electrification programs with PV. New initiatives are currently in place under the umbrella of the Under Ministry of Electricity and Alternative Energy. Among others, the programme "Electricity to Live with Dignity", launched in 2007, which calls for 100% electricity coverage by the year 2030, including the use of PV in stand-alone and wind-hybrid systems.

Collateral benefits of PV programs in Bolivia include savings in traditional fuels in rural areas, development of a larger human resource base for project development and implementation, and the creation of local industries (BATEBOL, PHOCOS Latin America) that produce PV system components for the local market and export [42].

24.4.3 Brazil

PV rural electrification projects in Brazil started in the period 1992–1993 [44]. Around 1500 SHS were installed as part of these projects in Northeast Brazil in cooperation with the local electricity distribution companies, who were responsible for systems installation, maintenance and performance monitoring. An initiative to promote the use of renewable energy was launched in 1995 with the goal of installing 50 MW_p of PV systems by the year 2005 [45]. PRODEEM, a program for energy development in municipalities and states was launched in 1994 to deliver electricity by means of renewable energy to rural communities not served by the grid [46]. By the year 2002 the fifth phase of this program was being completed, bringing the number of PV systems installed to a total of 8742, equivalent to 5.8 MW of power.

A general audit of the PRODEEM program was carried out by the corresponding federal authority in 2002, having found that the program had under-performed in a number of concepts. Among others, centralized decision making with low degree of involvement of the beneficiaries, not having prepared the user for the opportunities that electricity availability carries with it, responsibilities of the agents involved not clearly identified, lack of user training for system operation and maintenance, as well as lack of technical assistance and system sustainability measures. Hence, the PRODEEM program was restructured in 2003 and incorporated into the then recently created "Luz para Todos" (Light for All) program, which had the goal of implementing 300 000 rural systems by the end of 2008. A task force was created in 2004 to carry out field surveys so that legal ownership of already installed systems could be regularized. Training of services providers, communal agents and municipal technicians for system protection and maintenance was also carried out. The process for the implementation of PRODEEM is well established and documented in the literature cited here [48].

Several state governments, including Minas Gerais, Sao Paulo and Parana, launched their own PV rural electrification projects [47]. General criteria for eligibility in the CEMIG program

include localities farther than 5 km from the nearest electricity grid and a user density of less than 5 people per square kilometer. CEMIG provides financing for 64% of the project cost, the remaining 36% being covered by the community authorities [49]. A PV "pre-electrification" program was established in 1998 to stimulate growth of the electrical demand in remote places, to the point where grid extensions could become economically viable.

Over the years, a number of programs and projects have taken place in Brazil in cooperation with foreign agencies for the implementation of off-grid PV. Examples can be found in references [49, 50]. All in all, it is estimated that by the end of 2008 the total PV capacity installed in rural areas of the country is close to 20 MW, less than half the original goal established for 1995, which merits a thorough analysis of the causes that prevented the original goal from being met.

24.4.4 Mexico

The number of people in Mexico with no access to electricity has decreased in the past 10 years, mostly by grid extensions with some modest contribution from renewables, mainly PV. A highly dispersed rural population and a rough terrain make grid extensions technically difficult and economically unviable, so conventional electrification rates have decreased in recent years. On the other hand, former government programs for poverty alleviation, such as PRONASOL (1989), PROGRESA and OPORTUNIDADES, which were the source of funds for PV rural electrification, have been phased out or have declined to small local projects. Hence, the number of PV installations has remained practically unchanged in the last few years. Today, over 2500 rural communities have been fully supplied with SHS and other communal services through government programs, which means there are over 60 000 SHS in the whole territory. It is estimated that another 30 000 SHS have been installed outside government programs on a purely commercial basis. In addition, thousands of other PV-powered rural services have been provided, including rural telephones, schools, health centers and communal buildings. In a recent program by the Agriculture Ministry, partially financed by a GEF-World Bank grant, hundreds of small PV water pumps have been installed. The technology is becoming popular among farmers, and new programs to promote its use are in preparation. Also in preparation is another program to bring basic electricity services to 50 000 households in remote areas. In most instances PV will be the technology of choice. Total PV installed capacity in Mexico at the end of 2007 was reported at 21.7 MW, of which 15.5 MW belong to rural electrification.

A distinctive characteristic of the early Mexican PV rural electrification programs was the active participation of the national electric utility (CFE) as technical normative agency, a central element for quality assurance and sustainability [29]. Under contract to CFE, the Electrical Research Institute of Mexico developed, in the early 1990s, a set of technical standards and specifications for project implementation. Laboratory testing and field evaluation protocols were also developed and implemented. Systems procurement and installation were carried out according to the prevailing law for public works.

Finance to purchase the PV systems has been mostly provided by the Federal Government, with lesser contributions from the state and municipal governments. Communities are requested to contribute to the project according to their own economic capacity. In-kind contributions, such as carrying the PV equipment to the community from the nearest point where vehicles have access, is one of the most frequent services by the community in support of the project. Funds are provided by the government as part of the patrimony for the community and, hence, no *a priori* money repayment mechanisms are established. However, communities are free to implement any fund-raising activities that can help them maintain their systems and purchase additional equipment. A popular mechanism is the so-called communal fund, in which members of the community contribute money or man-hours or both to a common fund managed by the community representatives. The communal fund is then used to maintain the PV installations and/or to implement other projects

for the benefit of the community, such as waterworks repair, building a new schoolroom and the like. A more detailed description of the Mexican PV rural electrification program can be found in the literature mentioned in references [28, 51, 52].

24.4.5 Sri Lanka

An estimated two million households lack access to grid electricity in this country. Recent studies indicate that at least 10% of these households could afford a SHS at current prices, based on a monthly household income of Rs. 5000 (about €85) [53]. A market study commissioned by the National Development Bank of Sri Lanka in 1991 indicated that a market of 360 000 households could afford a PV system [54]. However, so far only about 15 000 such systems had been sold commercially for cash by 2002 through a retail network.

Many reasons have been cited for this low penetration level. As in other places, the biggest barrier for the exploitation of this market has been a lack of consumer financing. Political promises for grid extensions in rural areas and the government sponsored "Electricity for All by the Year 2000" campaign, which was perceived as *grid electricity* for all, have also been major barriers. Lack of government-level endorsement for solar photovoltaics, along with other renewables, was a hindrance to both private sector and NGO promoters. Nevertheless, four companies retail PV systems in Sri Lanka: Shell Renewables, Resco Asia (a subsidiary of the US Selco), Alpha Thermal Systems and Access International. Sarvodaya SEEDS, a local NGO, provides micro-financing in partnership with some of these companies.

In the late 1980s, solar PV modules, 12-V lamps and simple electronic charge controllers were assembled in Sri Lanka by Solar Power & Light Company (SPLC), a private venture established as Power & Sun in 1986. However, manufacturing of solar modules ceased as the advantage over imported products was eroded, in part because of the high import duties on raw materials. Thus, SPLC essentially developed into a marketing organization by creating an infrastructure to market, install and maintain PV systems in rural areas. SPLC uses retailers to stock solar home systems; trained technicians and individual agents, called *corresponding agents* (CA), are used to canvass sales, install systems and provide customer service. SPLC, which has just been bought over by Shell International Renewables, has sold over 3000 PV systems directly to consumers, mostly for cash.

Apart from cash sales of SHS, there are examples in Sri Lanka of both successful and "deemed as failed" projects, using loan repayments by the consumer as the measure of success. Projects implemented with total community involvement by NGOs, such as Sarvodaya and Solanka, have been quite successful. Sarvodaya's already successful micro-credit operation was adapted to market SHS, and implemented a pilot project with 250 installations through its Rural Technical Services branch, with assistance from the US-based NGO Solar Electric Light Fund (SELF, now the private company Selco) and a seed fund from a US foundation. Further activities by Sarvodaya using funds from a credit line for renewable energy provided to Sri Lanka by the World Bank, led to the installation of 300 more systems. Plans to install 5000 more within the next five years have already been drawn up. Sarvodaya SEEDS has become a Participating Credit Institution of the Energy Services Delivery Project. It can now access the fund directly and on-lend to customers. In January 2001, it has been reported that Sarvodaya SEEDS' lending portfolio exposure is Rs 89 million (about US $1 million) for solar photovoltaics.

Solanka Sun Associates, created community-level capabilities to do complete projects, including evaluation of potential customers, providing financing; designing, installing and maintaining the solar home systems; collecting repayment and managing the project with the inherent difficulties of managing projects from cities. This model has shown that success can be achieved

with proper training and incentives to the operators at the village level. With the lower than commercial rates of interest provided to its customers, Solanka has targeted the lower economic category of the market, having thus far installed two projects, one with 84 solar home systems in the village of Morapathawa, and the other with 77 systems in Thorawa. This has been possible since the funds for lending have been provided as grants. However, it will be difficult to sustain this scheme with commercial level funding, unless interest rates are increased and the loan repayment period is reduced. However, this will naturally exclude the current target market of this organization. Thus, the challenge for Solanka is to secure further grant funding to replicate such projects. Its biggest merit has been to prove that the village has the capacity to implement and manage PV electrification projects. Solanka had the provincial government's patronage and support when the selection of areas for implementation was made jointly. The project has an interesting feature in that the 12-V lamp units and the simple electronic controllers are manufactured at the village level. Also, a village level repair unit has been started where defective battery cells can be replaced to lengthen the life of the battery. Loan repayments to date for both projects are 100% due to the grass roots level service that is provided to users. For instance, even when a battery fails, the user immediately gets a replacement while the old one is being repaired or serviced. The Colombo-based head office focuses on long-term strategic planning and also imposes the accounting controls, with audits of all accounting operations in the process of selecting recipients and collecting repayments. On the basis of the experience of these two projects, its original promoter has established a commercial solar company called RESCO as a subsidiary of Selco-USA.

A counterexample is the Pansiyagama 1000 homes project, funded by the Sri Lankan and Australian governments, which has a very low repayment rate in spite of the very favorable finance scheme applied. Hence, it can be considered a failure, although technically over 90% of the systems are still operating. This project was politically motivated, and was implemented by the National Housing Development Authority (NHDA) of Sri Lanka, which attempted to implement a "grass roots" level program. However, the top-down manner in which it was done resulted in poor community level involvement and poor management infrastructure. The systems installed were more sophisticated than the normal SHS being installed elsewhere in the country and had a typical cost ranging from Rs 20 000 (US $571) to Rs. 32 000 (US $914), depending on the size of the module and number of lamps. This led to a monthly payment from Rs. 75 (US $2.14) to Rs. 135 (US $3.85), depending on the cost of the unit (US $1 = Rs. 35 in 1990), which turned out to be unrealistically low. According to a socioeconomic survey conducted by the Marga Institute of Sri Lanka, it was found that most households could afford to pay much more for the system. However, since the payments by households were set at nominal terms, their value has been eroding with inflation. At some point in time, the infrastructure for collecting repayments broke down as a result of bureaucratic problems, and some initial technical problems in the systems set a precedent for nonpayment, which was most difficult to break later. Some of these problems have been fixed after the NHDA handed over the maintenance and money collection duties to Power & Sun in 1991. However, the repayments remain around 50%, as the cost of collecting the money is much more than the actual collections due to the low monthly payment. This means that over-subsidizing solar photovoltaics has a negative effect on the dissemination of solar photovoltaics through the business route. This project has demonstrated that solar photovoltaics is an appropriate technology, but the implementing methodology used for the project was not sustainable and as a result was found to be detrimental to the commercial dissemination process of solar home systems in Sri Lanka.

24.4.6 Water Pumping in the Sahel

The Regional Solar Programme (RSP) was one of the early systematic programs to apply PV technology to solve pressing problems in the Sahel region of sub-Saharan Africa. Financed by a grant from the European Commission, to cover the cost of PV equipment and other procedures

such as training, information and public awareness activities, regional coordination and technical assistance, this program was launched in 1989. The goal of the RSP was to install almost 1.4 MW_p of PV modules (about 3.5% of the world market at that time) in water-pumping systems, vaccine refrigerators, community lighting and battery-charging stations. At the end of the program, a total of 626 pumping systems and 644 communal systems had been installed and a wealth of lessons learned.

The principal objectives of the RSP were [55] to improve the accessibility of water in both quantity and quality, to improve the economic condition of the villagers by development of complementary resources through gardening, to reduce the time spent in procuring drinking water, to train personnel for project management, to create management groups for the solar equipment and to develop and adopt a legal framework for operation of the equipment with contractual structure of the relations between users and private companies.

Not all specific goals and objectives originally planned were fully met, but important lessons were derived. The drinking water component of the program took more than 90% of the installed PV power, so the lessons learned apply basically to this application. Solar pumping was found to be more affordable than diesel motor pumping, by about a factor of two per cubic meter of water. Compared with the per inhabitant cost of manual water-raising pumps, including the borehole, investment in PV pumping systems was about 10% higher, but the service quality of PV pumps was superior. Of the total installed cost, 31% was for the supply of the PV system, 11% for installation, 12% for regional activities (including coordination, quality control, tests, monitoring and the like) and 46% for the distribution network, water tower and other reception infrastructure.

One mode of management of the water supply system used in some communities was directly inspired by the management of water points equipped with manual pumps through a village water committee. This management system, however, was not effective owing to a number of difficulties, including an imperfect mastering of the accounting tools and a quasi-systematic confusion of the responsibilities assigned to the principal members of the committee. Other communities preferred delegating the management of the entire system to a private-type body on a fee basis or to a communal-type body. These latter forms of management seemed better suited to the local conditions than the water committee scheme.

The idea of using PV technology to improve productivity of market gardening and farming was finally abandoned, hence the prospects of PV technology in Sahel were found in meeting specific domestic electrical needs (lighting, radios and TV sets), in pumping drinking water and in telecommunications.

24.5 TOWARD A NEW PARADIGM FOR RURAL ELECTRIFICATION

The problem of rural electrification has been traditionally handled by conventional means in a process of successive approximations. In this process, the most remote and dispersed population is attracted to larger population centers, which are then served by a mini-electrical local grid, fed by diesel gen-sets or small hydroelectric generators. As the load increases, a point is reached at which extensions of the main grid become economically viable. This process is known as *pre-electrification* among electric companies and has been the basic growth mechanism of the interconnected system in rural areas. Although effective from a purely technical/economic point of view, this pre-electrification process has several downsides. From the social point of view, it forces people to leave their place of origin to create larger population centers, which in turn induces the need for the central provision of other services and puts a larger stress on the environment.

The term pre-electrification is being used nowadays by some authors in reference to PV rural electrification and, in some cases such as the CEMIG example above, the operational scheme is also being transported. However, there are a number of reasons to think that using the term pre-electrification in the context of photovoltaics is not appropriate. Furthermore, transporting the concept is bound to cancel the advantages offered by photovoltaics to create a new path for rural electrification. First, from the purely technical point of view, photovoltaics offers the possibility of supplying high-quality electrical services, even to the most remote sites, without the need for relocating people or eventually having to resort to grid extensions. Because of its modular nature PV systems can grow in pace with the load. Furthermore, this vision is compatible with current trends in the electrical sector toward distributed generation systems, and is supported by the development of more efficient electrical appliances, the miniaturization of electrical technologies and the progress being made in electronic devices for system supervision and load management. There are also environmental reasons that lead to the notion that the local generation of electricity using renewable and nonpolluting energy resources is more convenient than building kilometers of electricity lines across ecologically sensitive areas. Local generation of electricity also offers the possibility for local management and, hence, for active community participation in the process of self-development.

On the other hand, there is evidence that the traditional electrification process based on the old paradigm of d*elivering electricity to rural people* has frequently lent itself mainly to fulfill the political need to improve the electrification statistics rather than to serve the real needs of people. Thus, it is not uncommon to observe rural communities with access to the grid, where a good portion of the households are not connected, usually because of lack of money to pay for the connection fee, or because the secondary distribution network only reaches the center of the community, leaving the rest of the population unserved. Therefore, if electricity supply is to be used as a tool for development, that is understood as increased life expectancy, more knowledge and a better standard of living [56], then the present rural electrification paradigm must be urgently changed to: "*increasing the access of the rural poor to electricity-based services*". This set of services includes health, clean water, education, food preservation, entertainment and the possibility of engaging in productive activities.

Under this perspective, current PV rural electrification activities represent a landmark of the transition period into a new rural electrification scheme to substitute the old one with better results. However, the *new electrification* requires among other things, a new culture for electricity supply and use in rural areas, an *ad hoc* legal framework, new business practices within the electrical sector, innovative financial mechanisms, new and better technologies and appropriate institutional schemes. Thus, the actual value of programs and projects currently underway has to be weighted not only in terms of the number of PV users in rural areas, but also in terms of the benefits PV installations are bringing to the population. Also important to gauge the effectiveness of current PV rural electrification projects is the relevance of the lessons learned from them, which will allow the implementation of improved schemes for successful project replication.

This new paradigm for rural electrification should in turn be embedded in the broader concept of rural energy supply, which calls for a timely supply of useful energy in a variety of forms, so that people can have better opportunities to improve their own quality of life, to foster local economic development and to protect the environment.

After almost twenty years of PV applications in rural areas of the world, the problem of providing electricity-based services to about one third of the global population still remains virtually untouched. Important lessons have been learned, but a number of critical elements need to be put in place if PV is to become the main technology option for rural electrification of remote communities. Of foremost importance is the institutionalization of the process for the large scale and sustainable deployment of the technology. Much has been written about this, but few programs or projects can in fact serve as good examples of long-term sustainability.

REFERENCES

1. Kemp W, *The Flow of Energy in a Hunting Society. Energy and Power*, A Scientific American Book, W.H. Freeman and Company, San Francisco, CA, pp 55–65 (1971).
2. Rappaport R, *The Flow of Energy in an Agricultural Society. Energy and Power*, A Scientific American Book, W.H. Freeman and Company, San Francisco, CA, pp 69–80 (1971).
3. Cook E, *The Flow of Energy in an Industrial Society. Energy and Power*, A Scientific American Book, W.H. Freeman and Company, San Francisco, CA, pp 83–91 (1971).
4. *World Energy Assessment: Energy and the Challenge of Sustainability*, United Nations Development Program, United Nations Department of Economic and Social Affairs, World Energy Council (Sept. 2000).
5. *The Challenge of Rural Energy Poverty in Developing Countries*, World Energy Council, Food and Agriculture Organization of the United Nations (Oct. 1999).
6. Bowers B, *A History of Electric Light and Power*, History of Technology Series 3, P. Peregrinus, The Science Museum, London (1982).
7. National Energy Balance 1997. Energy Secretariat of Mexico.
8. Foley G, *Photovoltaic Applications in Rural Areas of the Developing World*, World Bank Technical Paper Number 304, Energy Series (1995).
9. Lorenzo E, Photovoltaic Rural Electrification, *Prog. Photovolt.: Res. Appl*. **5**, 3–27 (1997).
10. Urbano J, Introduction to the Design, Operation and Application of Photovoltaic Systems – The CIEA Experience (in Spanish), *Proc. 4th Mexican National Solar Energy Society Meeting* (San Luis Potosi, Mexico, 1–3 October 1980).
11. Urbano J, *Rev. Sol*. **16**, 10–13 (1989).
12. Martinot E, Ramankutty R, Rittner F, *The GEF Solar PV Portfolio: Emerging Experience and Lessons*, Monitoring and Evaluation Working Paper 2, GEF pre-publication draft (August 2000).
13. Huacuz J, Agredano J, Beyond the Grid: photovoltaic rural electrification in Mexico, *Prog. Photovolt.: Res. Appl*. **6**, 379–395 (1998).
14. Ramos R *et al*., *Photovoltaic Solar Energy: An Option for Rural Electrification in Cuba (in Spanish)*, Solar Energy Research Centre of Cuba (1995).
15. *Clean Water with Clean Energy: Drinking Water Provision in Remote Regions with Decentralised Solar Power Supply*, INCO-DC Project ERBIC18CT960104, Final Report, European Commission, FhG-ISE, Freiburg, Germany (2000).
16. Sapiain R *et al*., *Solar Photovoltaic Pumping in Peasant Communities and New Productive Agricultural Applications in Arid Zones in the North of Chile (in Spanish)*, Renewable Energy Centre, Engineering Faculty, University of Tarapaca (Oct. 1997).
17. Ahm P, Small PV Powered Medical Equipment, *UNESCO World Solar Summit* (July 1993).
18. Muhopadhyay K, Sensarma B, Saha H, *Sol. Energy Mater. Sol. Cells* **31**, 437–446 (1993).
19. *De l'eau solaire pour la Somalie*, Systemes Solaires, No. 100 (1994).
20. *Les lampes portables solaires*, Systemes Solaires, No. 100 (1994).
21. Flores R *et al*., "Characterization and Evaluation of 30 PV Ovonic-Unisolar Solar Home Systems (in Spanish)". *Proc. ISES Millennium Solar World Forum* (Mexico City, October 2000).
22. Huacuz J, Urrutia M, Eds, *Proceedings of the International Workshop Charge Controllers for Photovoltaic Rural Electrification Systems*, Electrical Research Institute Cuernavaca, Mexico (1998).
23. Nieuwenhout F, *Monitoring and Evaluation of Solar Home Systems: Experiences with Applications of Solar PV for Households in Developing Countries*, Report ECN-C-00-089, Netherlands Energy Research Foundation, Petten (Sept. 2000).
24. dos Santos R, Zilles R, PV Residential Electrification: A Case Study on Solar Battery Charging Stations in Brazil. *Prog. Photovol.: Res. Appl*. **9**(6), 445–453 (2001).
25. European Commission, Directorate General for Energy; Universal Technical Standards for Solar Home Systems. Thermie B: SUP-995-96 (1998).

26. Electrical Research Institute, *Technical Specification for Rural Illumination Photovoltaic Systems (in Spanish)*, Revised Edition. Cuernavaca, Mexico, IIE (1999).
27. Martinot E, Ramankutty R, Rittner F, *Thematic Review of the GEF Solar PV Portfolio: Emerging Experience and Lessons*, Global Environment Facility Working Paper, pre-publication draft (June 2000).
28. Huacuz J, Martinez A, Renewable Energy Rural Electrification: Sustainability Aspects of the Mexican Programme in Practice, *Nat. Res. Forum* **19**, 223–231 (1995).
29. Huacuz J, Gonzalez C, Uria F, "The role of Commission Federal de Electricidad in the Mexican photovoltaic rural electrification program", *Proc. IERE Workshop Photovoltaic Rural Electrification and the Electric Power Utility*, IIE-EPRI (Cocoyoc, Mexico, May 1995).
30. Flores C, "Expanding PV Rural Electrification with Local Industry and Technology: The Mexican Experience", *Proc. Sustainable Development of the Rural World. Decentralized Electrification Issues* (Marraketch Marruecos, 13–17 November 1995).
31. Huacuz J, Sustainable Energy in Rural Zones within the Process of Modernization of the Energy Sector in Latin America and The Caribbean (in Spanish), *Proc. Enerlac '98, IV Energy Conference of Latin America and The Caribbean* (Dominican Republic, 1998).
32. Aguilera T, *Energía Solar Fotovoltaica en el Ambito de la Cooperación al Desarrollo*, Caso de Estudio: El Altiplano Boliviano, Tesis Doctoral, Universidad Politécnica de Madrid, Escuela Superior de Ingenieros en Telecomunicaciones. Madrid (1995).
33. Nieuwenhout F *et al*., *Monitoring and Evaluation of Solar Home Systems. Experiences with Applications of solar PV for Households in Developing Countries*, Report ECN-C-00-089, Netherlands Energy Research Foundation (Sept. 2000).
34. Hankins M, *Solar Rural Electrification in the Developing World*, Four Case Studies: Dominican Republic, Kenya, Sri Lanka and Zimbabwe. Solar Electric Light Fund (1993).
35. Cabraal A *et al*., *Best Practices for Photovoltaic Household Electrification Programs*, Lessons from Experiences in Selected Countries, World Bank Technical Paper Number 324, Asia Technical Department Series, The World Bank (1996).
36. Martinot E *et al*., World Bank/GEF Solar Home Systems Projects: Experiences and Lessons Learned 1993–2000. *Renewable and Sustainable Energy Reviews*, **5**, 39–57 (2001).
37. Fabris A, Sotelino E, Programas de Electrificación Rural en el Cono Sur de América Latina, *Los Recursos Energéticos Renovables y las Políticas de Electrificación Rural*, pp 109–123 (1997).
38. Frigerio A, Financiamiento del Programa de Abastecimiento Eléctrico a la Población Rural Dispersa de Argentina, *Seminario de Inversiones y Negocios para Energías Renovables en América Latina*, (Quito, Ecuador, 14–16 September 1998).
39. Castiella H, *Proyecto de Electrificación Rural Mediante Energía Solar Fotovoltaica en Bolivia*, Agencia Española de Cooperación Internacional (LaPaz, Bolivia, 1993).
40. Smith P, International Project CRE http://www.rsvp.nrel.gov (1996).
41. INTI K'ANCHAY, Informative Bulletin No. 1, 2. Energética (Sept. 1998, 1999).
42. Orellana Lafuente RJ, Morales Udaeta ME: *Energías Renovables y Alivio de la Pobreza en Bolivia*, Fundación Bariloche, GNESD 2006 http://www.fundacionbariloche.org.ar/idee/taller%20renovables/keynotes/Paper%20Orellana-Morales.pdf
43. *Renewable Energy Rural Electrification Projects for the Cochabamba-Bolivia Department (in Spanish)*, Comité de Coordinación Interinstitucional (Cochabamba, April 1995).
44. Barbosa E *et al*., *Toward a Sustainable Future for the Use of SHSs for Rural Electrification in Brazil*.
45. Ribeiro C, Bezerra P, Zilles R, Moskowics M, *Brazilian Strategy on PV Dissemination* (1998).
46. Quintans L, Lima J: *Prodeem: realizaçoes e progressos*, Cresesb Informa, Año III, No. 4 (December 1997).
47. Leonelli P, Borba A, *PRODEEM Aprendizados e Reflexoes*, II Simposio Nacional de Energia Fotovoltaica. Rio de Janeiro, Brazil, May 2005.

48. Silva S, *PRODEEM. IV Encontro do Forum Permanente de Energias Renovaveis* (Recife, Pernambuco, Brasil, 6–9 October 1998).
49. Diniz A, Programa de Implantado de Sistemas Fotovoltaicos da CEMIG, *IV Encontro do Fórum Permanente de Energias Renovaveis* (Recife, Pernambuco, Brazil, 6–9 October 1998).
50. dos Santos P, Programa de Implantaçao de Bombeamiento de Agua FV da COPASA, *IV Encontro Forum Permanente de Energias Renovaveis* (Recife, Pernambuco, Brazil, 6–9 October 1998).
51. Huacuz J, Martínez A, *ATAS Bull*. **8**, 177–194 (1992).
52. Huacuz J, Martínez A, *Mexico: Rural Electrification Program with Renewable Energy Systems*, EDG, No. 5, pp 15–20 (1996).
53. Gunaratne L, Funding and repayment management of PV system dissemination in Sri Lanka, *Proceedings of the Conference on Financial Services for decentralized Solar Energy Applications II* (Harare, Zimbabwe, 20–23 October 1998).
54. Gunaratne L, Solar PV Market Development in Sri Lanka. GEF Workshop: Making a Difference in Emerging PV Markets. Marrakech, Morocco, 24–28 September (2000).
55. Regional Solar Programme, *Lessons and Perspectives*, European Commission (DGVIII) (December 1999).
56. UNDP, *Human Development Report 2001* (2001).

Index

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus

 ISBN: 978-0-470-72169-8

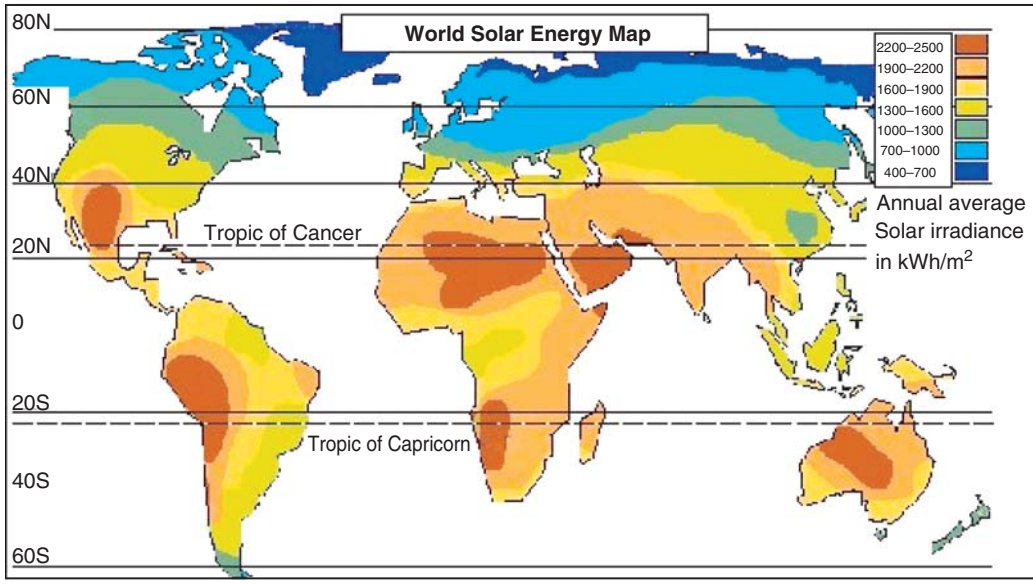


Plate 1 Figure 1.3 World distribution of the annual solar radiation (kWh/m^2) [obtained from www.rise.org.au/info/Applic/Array/image003.jpg]

Plate 2 Figure 10.2 Evolution and commercialization of the Fresnel lens technology initiated in Sandia Labs (1976): 750 kW_p plant, manufactured and installed by Guascor Fotón in Cáceres, Spain, 2007) (Courtesy Guascor Fotón)

Handbook of Photovoltaic Science and Engineering, Second Edition
Edited by Antonio Luque and Steven Hegedus
 ISBN: 978-0-470-72169-8

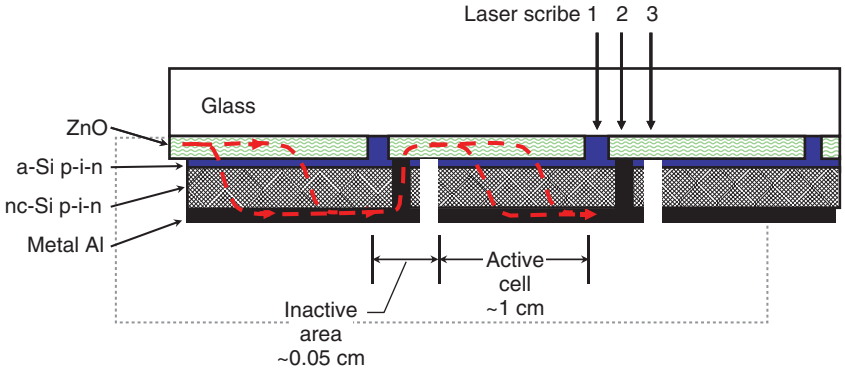


Plate 3 Figure 12.23 Cell interconnection (monolithic integration) of superstrate-type solar cells showing the three laser scribes

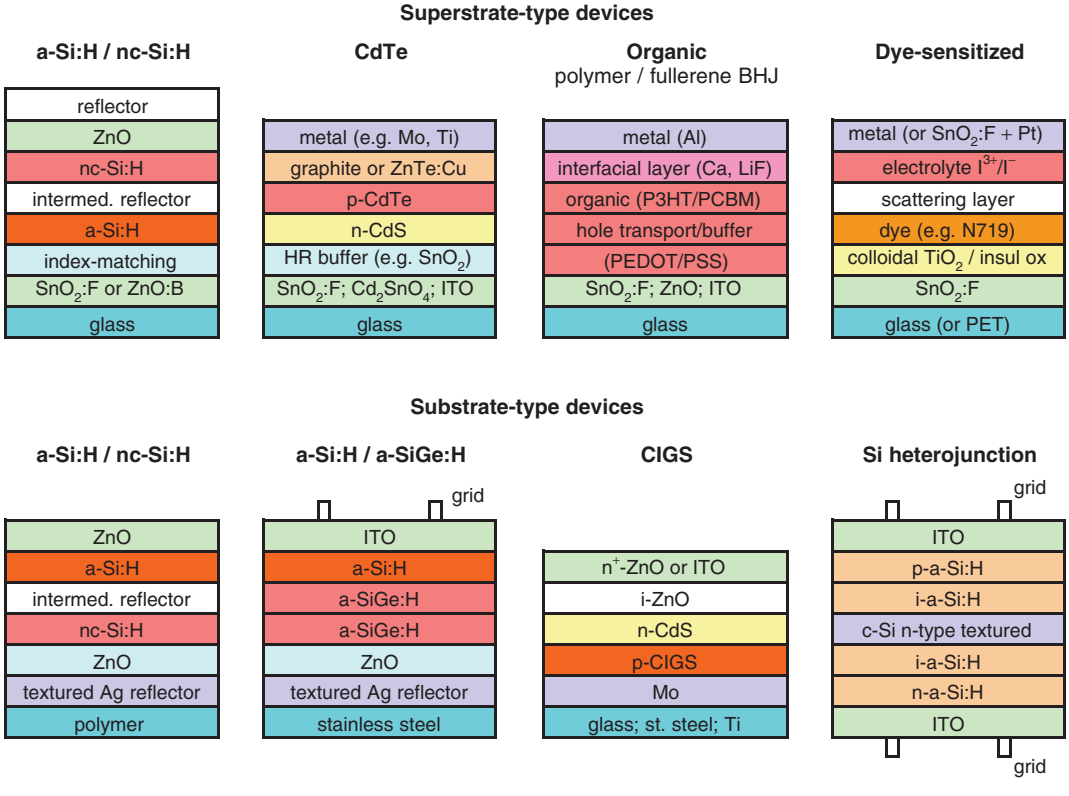


Plate 4 Figure 17.1 The principal types of solar cells that utilize one or more TCO layers in their construction. Top row: superstrate type devices; bottom row: substrate-type devices

Plate 5 Figure 19.2 Olmedilla, 2008: NOBESOL 60 MW_P (Courtesy of Nobesol)

Plate 6 Figure 19.10 Ota City, Japan distributed 550 on-grid residential systems, 2.2 MW (Courtesy of Hironobu Igarashi)

(a) (b)

Plate 7 Figure 22.21 Trackers at Amaraleja 48 MW PV plant. Each tracker has a surface of 140 m^2, supporting a 182 kW_p PV array: (a) no shading occurs during most the day; (b) some shade occurs in early morning and late afternoon

Plate 8 Figures 23.44 (L) and 23.8 (R). The ECN building 42 (Petten, The Netherlands) has a 43 kWp PV system integrated into the conservatory roof. The semi-transparent Si modules on the curved conservatory roof acts as a parasol by providing partial shading in front of the offices, thus eliminating the need for air-conditioning in a moderate climate. The attached building on the right side has a PV shading system over each window for the same purpose (see Figure 23.16). Interior view of the conservatory lobby shows the partial shading and diffuse interior light. Reproduced with permission by BEAR Architecten (T. Reijenga)