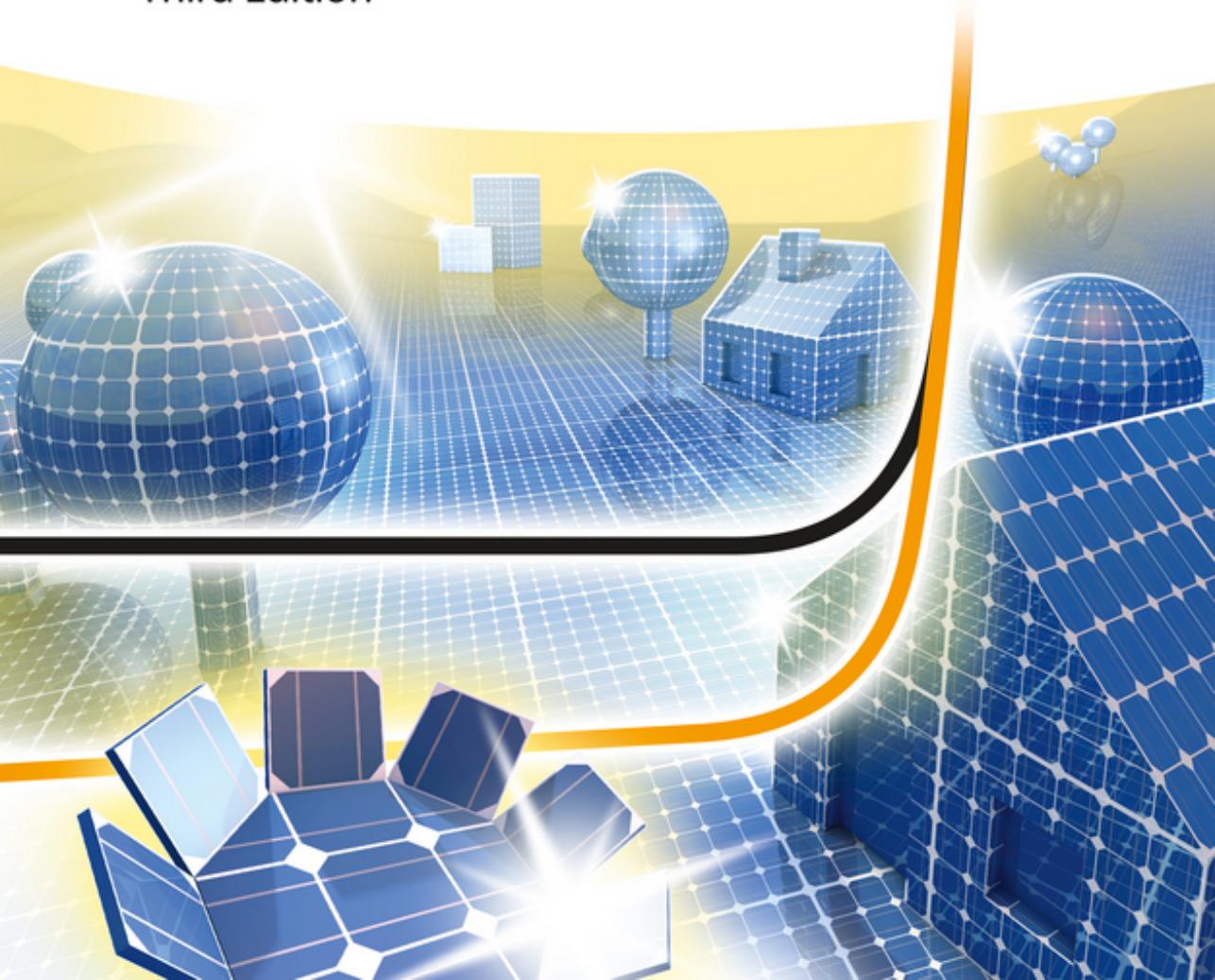


Peter Würfel and Uli Würfel

Physics of Solar Cells

From Basic Principles to Advanced Concepts
Third Edition



Peter Würfel and Uli Würfel

Physics of Solar Cells

Peter Würfel and Uli Würfel

Physics of Solar Cells

From Basic Principles to Advanced Concepts

3rd Edition

WILEY-VCH
Verlag GmbH & Co. KGaA

Authors**Peter Würfel**

Karlsruhe Institute of Technology (KIT)
Faculty of Physics
Engesserstr. 7
76131 Karlsruhe
Germany

Uli Würfel

Fraunhofer ISE
Heidenhofstr. 2
79110 Freiburg
Germany

All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.: applied for**British Library Cataloguing-in-Publication Data**

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

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <<http://dnb.d-nb.de>>.

© 2016 Wiley-VCH Verlag GmbH & Co.
KGaA, Boschstr. 12, 69469 Weinheim,
Germany

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Print ISBN: 978-3-527-41312-6

ePDF ISBN: 978-3-527-41311-9

ePub ISBN: 978-3-527-41309-6

Mobi ISBN: 978-3-527-41310-2

Cover Design Formgeber, Mannheim,
Germany

Typesetting SPI Global, Chennai, India
Printing and Binding

Printed on acid-free paper

Contents

List of Symbols *IX*

Preface *XI*

1 Problems of the Energy Economy *1*

1.1 Energy Economy *1*

1.2 Estimate of the Maximum Reserves of Fossil Energy *4*

1.3 The Greenhouse Effect *6*

1.3.1 Combustion *6*

1.3.2 The Temperature of the Earth *7*

1.4 Problems *9*

2 Photons *11*

2.1 Black-body Radiation *11*

2.1.1 Photon Density n_γ in a Cavity (Planck's Law of Radiation) *12*

2.1.2 Energy Current Through an Area dA into the Solid Angle $d\Omega$ *16*

2.1.3 Radiation from a Spherical Surface into the Solid Angle $d\Omega$ *19*

2.1.4 Radiation from a Surface Element into a Hemisphere
(Stefan–Boltzmann Radiation Law) *20*

2.2 Kirchhoff's Law of Radiation for Nonblack Bodies *22*

2.2.1 Absorption by Semiconductors *24*

2.3 The Solar Spectrum *24*

2.3.1 Air Mass *26*

2.4 Concentration of the Solar Radiation *28*

2.4.1 The Abbé Sine Condition *29*

2.4.2 Geometrical Optics *30*

2.4.3 Concentration of Radiation Using the Sine Condition *31*

2.5 Maximum Efficiency of Solar Energy Conversion *33*

2.6 Problems *39*

3 Semiconductors *41*

3.1 Electrons in Semiconductors *42*

3.1.1 Distribution Function for Electrons *43*

3.1.2 Density of States $D_e(\epsilon_e)$ for Electrons *43*

3.1.3	Density of Electrons	48
3.2	Holes	50
3.3	Doping	52
3.4	Quasi-Fermi Distributions	57
3.4.1	Fermi Energy and Electrochemical Potential	59
3.4.2	Work Function	63
3.5	Generation of Electrons and Holes	65
3.5.1	Absorption of Photons	65
3.5.2	Generation of Electron–Hole Pairs	69
3.6	Recombination of Electrons and Holes	71
3.6.1	Radiative Recombination, Emission of Photons	72
3.6.2	Nonradiative Recombination	74
3.6.3	Lifetimes	85
3.7	Light Emission by Semiconductors	87
3.7.1	Transition Rates and Absorption Coefficient	88
3.8	Problems	92
4	Conversion of Thermal Radiation into Chemical Energy	95
4.1	Maximum Efficiency for the Production of Chemical Energy	98
4.2	Shockley–Queisser Limit	103
4.3	Problems	104
5	Conversion of Chemical Energy into Electrical Energy	105
5.1	Transport of Electrons and Holes	105
5.1.1	Field Current	106
5.1.2	Diffusion Current	107
5.1.3	Total Charge Current	109
5.2	Separation of Electrons and Holes	111
5.3	Diffusion Length of Minority Carriers	113
5.4	Dielectric Relaxation	115
5.5	Ambipolar Diffusion	116
5.6	Dember Effect	117
5.7	Mathematical Description	120
5.8	Problems	120
6	Basic Structure of Solar Cells	123
6.1	A Chemical Solar Cell	123
6.2	Basic Mechanisms in Solar Cells	128
6.3	Dye Solar Cell	131
6.4	The pn-Junction	133
6.4.1	Electrochemical Equilibrium of Electrons in a pn-Junction in the Dark	133
6.4.2	Potential Distribution across a pn-Junction	134
6.4.3	Current–Voltage Characteristic of the pn-Junction	138
6.5	pn-Junction with Impurity Recombination, Two-Diode Model	143

6.6	Heterojunctions	147
6.7	Semiconductor–Metal Contact	150
6.7.1	Schottky Contact	152
6.7.2	MIS Contact	153
6.8	The Role of the Electric Field in Solar Cells	154
6.9	Organic Solar Cells	160
6.9.1	Excitons	160
6.9.2	Structure of Organic Solar Cells	163
6.10	Light Emitting Diodes (LED)	167
6.11	Problems	169
7	Limitations on Energy Conversion in Solar Cells	171
7.1	Maximum Efficiency of Solar Cells	171
7.2	Efficiency of Solar Cells as a Function of Their Energy Gap	174
7.3	The Optimal Silicon Solar Cell	175
7.3.1	Light Trapping	176
7.4	Thin-film Solar Cells	181
7.4.1	Minimal Thickness of a Solar Cell	182
7.5	Equivalent Circuit	183
7.6	Temperature Dependence of the Open-circuit Voltage	185
7.7	Intensity Dependence of the Efficiency	186
7.8	Efficiencies of the Individual Energy Conversion Processes	186
7.9	Problems	188
8	Concepts for Improving the Efficiency of Solar Cells	189
8.1	Tandem Cells	189
8.1.1	The Electrical Interconnection of Tandem Cells	193
8.2	Concentrator Cells	194
8.3	Thermophotovoltaic Energy Conversion	196
8.4	Impact Ionization	197
8.4.1	Hot Electrons from Impact Ionization	200
8.4.2	Energy Conversion with Hot Electrons and Holes	200
8.5	Two-step Excitation in Three-level Systems	203
8.5.1	Impurity Photovoltaic Effect	203
8.5.2	Up- and Down-conversion of Photons	208
8.6	Problems	211
9	Characterization of Solar Cells	213
9.1	Spectral Response and Quantum Efficiency	213
9.2	Quasi-Steady-State Photoconductance	218
9.3	Luminescence	220
9.3.1	Homogeneous Distribution of Electrons and Holes	221
9.3.2	Inhomogeneous Distribution of Electrons and Holes	222
9.3.3	Electroluminescence	223
9.3.4	Photoluminescence	224

9.3.5	Diffusion Length	225
9.3.6	Series Resistance	228
9.4	Thermography	229
9.5	Light-Beam-Induced Current (LBIC)	230
9.6	The Suns- V_{OC} Method	233
9.7	Transient Techniques	237
9.7.1	Photovoltage Decay	237
9.7.2	Transient Absorption	238
9.7.3	Charge Carrier Extraction	240
9.7.4	CELIV – Charge Extraction by Linearly Increasing Voltage	241
9.7.5	Impedance Spectroscopy	241

Solutions 245

Appendix 263

References 267

Index 271

List of Symbols

$h, \hbar = h/(2\pi)$	Planck's constant	eVs
$\hbar\omega$	photon energy	eV
$a(\hbar\omega)$	absorptance	
$r(\hbar\omega)$	reflectance	
$t(\hbar\omega)$	transmittance	
$e(\hbar\omega) = a(\hbar\omega)$	emittance	
$\alpha(\hbar\omega)$	absorption coefficient	cm^{-1}
k	Boltzmann's constant	eV K^{-1}
σ	Stefan–Boltzmann constant	$\text{W m}^{-2}\text{K}^{-4}$
T	temperature	K
n_j	concentration of particle type j	cm^{-3}
e	electron	
h	hole	
γ	photon	
Γ	phonon	
n_e, n_h	concentration of electrons, holes	cm^{-3}
n_i	intrinsic concentration of electrons and holes	cm^{-3}
N_C, N_V	effective density of states in conduction band, valence band	cm^{-3}
ϵ_e, ϵ_h	energy of an electron, hole	eV
ϵ_C	energy of an electron at the conduction band minimum	eV
ϵ_V	energy of an electron at the valence band maximum	eV
μ_j	chemical potential of particle type j	eV
η_j	electrochemical potential of particle type j	eV
χ_e	electron affinity	eV
φ	electrical potential	V
e	elementary charge	As
ϵ_0	dielectric permittivity of free space	As (V m)^{-1}
ϵ	relative dielectric permittivity	
V	voltage = $[\eta_e(x_1) - \eta_e(x_2)]/e$	V

X | *List of Symbols*

ε_{FC}	Fermi energy for electron distribution in conduction band	eV
ε_{FV}	Fermi energy for electron distribution in valence band	eV
m_e^*, m_h^*	effective mass of electrons, holes	g
b_e, b_h	mobility of electrons, holes	$\text{cm}^2 (\text{Vs})^{-1}$
D_e, D_h	diffusion coefficient of electrons, holes	$\text{cm}^2 \text{s}^{-1}$
τ_e, τ_h	recombination life time of electrons, holes	s
R_e, R_h	recombination rate of electrons, holes	$\text{cm}^{-3} \text{s}^{-1}$
G_e, G_h	generation rate of electrons, holes	$\text{cm}^{-3} \text{s}^{-1}$
σ_e, σ_h	cross-section for the capture of an electron, hole by an impurity	cm^2
j_j	current density of particles of type j	$(\text{cm}^2 \text{s})^{-1}$
j_Q	charge current density	A cm^{-2}

Preface

Mankind needs energy for a living. Besides the energy in our food necessary to sustain our body and its functions (100 W), 30 times more energy is used on average to make our life more comfortable. Electrical energy is one of the most useful forms of energy, since it can be used for almost everything. All life on earth is based on solar energy following the invention of photosynthesis by the algae. Producing electrical energy through photovoltaic energy conversion by solar cells is the human counterpart. For the first time in history, mankind is able to produce a high quality energy form from solar energy directly, without the need of the plants. Since any sustainable, i.e. long term energy supply must be based on solar energy, photovoltaic energy conversion will become indispensable in the future.

This book provides a fundamental understanding of the functioning of solar cells. The discussion of the principles is as general as possible to provide the basis for present technology and future developments as well. Energy conversion in solar cells is shown to consist of two steps. The first is the absorption of solar radiation and the production of chemical energy. This process takes place in every semiconductor. The second step is the transformation into electrical energy by generating current and voltage. This requires structures and forces to drive the electrons and holes, produced by the incident light, through the solar cell as an electric current. These forces and the structures which enable a directional charge transport are derived in detail. In the process it is shown that the electric field present in a pn junction in the dark, usually considered a prerequisite for the operation of a solar cell, is in fact more an accompanying phenomenon of a structure required for other reasons and not an essential property of a solar cell. The structure of a solar cell is much better represented by a semiconducting absorber in which the conversion of solar heat into chemical energy takes place and by two semi-permeable membranes which at one terminal transmit electrons and block holes and at the second terminal transmit holes and block electrons. The book attempts to develop the physical principles underlying the function of a solar cell as understandably and at the same time as completely as possible. With very few exceptions, all physical relationships are derived and explained in examples. This will provide the nonphysicists particularly with the background for a thorough understanding.

Emphasis is placed on a thermodynamic approach that is largely independent of existing solar cell structures. This allows a general determination of the efficiency limits for the conversion of solar heat radiation into electrical energy and also demonstrates the potential and the limits for improvement for present-day solar cells. We follow a route first taken by W. Shockley and H. J. Queisser.¹⁾

In some respects this book is more rigorous than is customary in semiconductor device physics and in solar cell physics in particular. The most obvious is that identical physical quantities will be represented by identical symbols. Current densities will be represented by j and the quantity that is transported by the current is defined by its index, as in j_Q for the density of a charge current or j_e for the density of a current of electrons. In adhering to this principle, all particle concentrations are given the symbol n , with n_e representing the concentration of electrons, n_h the concentration of holes and n_γ the concentration of photons. I hope that those who are used to n and p for electron and hole concentrations do not find it too difficult to adapt to a more logical notation.

The driving force for a transition from exhausting energy reserves, as we presently do, to using renewable energies, is not the exhaustion of the reserves themselves, although oil and gas reserves will not last for more than one hundred years. The exhaustion does not bother most of us, since it will occur well beyond our own lifetime. We would certainly care a lot more, if we were to live for 500 years and would have to face the consequences of our present energy use ourselves. The driving force for the transition to renewable energies is rather the harmful effect which the byproducts of using fossil and nuclear energy have on our environment. Since this is the most effective incentive for using solar energy, we start by discussing the consequences of our present energy economy and its effect on the climate. The potential of a solar energy economy to eliminate these problems fully justifies the most intensive efforts to develop and improve the photovoltaic technology for which this book tries to provide the foundation.

Peter Würfel and Ulli Würfel

1) W. Shockley, H. J. Queisser, *J. Appl. Phys.* **32**, (1961), 510.

1

Problems of the Energy Economy

The energy economy of nearly all and, in particular, of the industrialized countries is based on the use of stored energy, mainly fossil energy in the form of coal, oil, and natural gas, as well as nuclear energy in the form of the uranium isotope ^{235}U . Two problems arise when we use our reserves to satisfy our energy needs. A source of energy can continue only until it is depleted. Well before this time, that is, right now at the latest, we have to consider how life will continue after this source of energy is gone and we must begin to develop alternatives. Furthermore, unpleasant side effects accompany the consumption of the energy source. Materials long buried under the surface of the Earth are released and find their way into air, water, and into our food. Up to now, the disadvantages are hardly perceptible, but they will lead to difficulties for future generations. In this chapter, we estimate the size of the fossil energy resources, which, to be precise, are composed not only of fossil energy carriers but also of the oxygen in the air that burns with them. In addition, we examine the cause of the greenhouse effect, which is a practically unavoidable consequence of burning fossil fuels.

1.1

Energy Economy

The amount of chemical energy stored in fossil energy carriers is measured in energy units, some more and some less practical. The most fundamental unit is the joule, abbreviated J, which is, however, a rather small unit representing the amount of energy needed to heat 1 g of water by a quarter of a degree or the amount of energy that a hair drier with a power of 1 kW consumes in 1 ms. A more practical unit is the kilo Watt hour (kWh), which is 3.6×10^6 J. 1 kWh is the energy contained in 100 g of chocolate. The only problem with this unit is that it is derived from the watt, the unit for power, which is energy per time. This makes energy equal to power times time. This awkwardness leads to a lot of mistakes in the nonscience press such as kilowatt per hour for power, because most people

mistake kilowatt to be the unit for energy, which they perceive as the more basic quantity. The energy of fossil fuels is often given in barrels of oil equivalents or in (metric) tons of coal equivalents (t coal equiv.).

The following relations apply:

$$1 \text{ kWh} = 3.6 \times 10^6 \text{ J} = 1 \text{ kWh}$$

$$1 \text{ t coal equiv.} = 29 \times 10^9 \text{ J} = 8200 \text{ kWh}$$

$$1 \text{ kg oil} = 1.4 \text{ kg coal equiv.} = 12.0 \text{ kWh}$$

$$1 \text{ m}^3 \text{ gas} = 1.1 \text{ kg coal equiv.} = 9.0 \text{ kWh}$$

$$1 \text{ barrel oil} = 195 \text{ kg coal equiv.} = 1670 \text{ kWh}$$

The consumption of chemical energy per time is an energy current (power) taken from the energy reserves. Thus, the consumption of one ton of coal per year, averaged over one year amounts to an energy current of

$$1 \text{ t coal equiv./a} = 8200 \text{ kWh/a} = 0.94 \text{ kW}$$

We look at Germany as an example of a densely populated industrialized country. Table 1.1 shows the consumption of primary energy in Germany in the year 2002, with a population of 82.5×10^6 . These figures contain a consumption of electrical energy per year of

$$570 \text{ TWh/a} = 65 \text{ GW} \Rightarrow 0.79 \text{ kW/person}$$

The energy consumption per capita in Germany of 5.98 kW is very high compared with the energy current of $2000 \text{ kcal/d} = 100 \text{ W} = 0.1 \text{ kW}$ taken up by human beings in the form of food, representing the minimum requirement for sustaining life.

Table 1.2 shows the consumption of primary energy in the world in 2002, with a population of 6×10^9 . This energy consumption is supplied from the available reserves of energy with the exception of hydro, wind, and biomass. The current remaining reserves of energy are shown in Table 1.3. This is the amount of energy that is estimated to be recoverable economically with present-day techniques at current prices. The actual reserves may be up to 10 times as large, about 10×10^{12} t coal equiv.

Table 1.1 Primary Energy Consumption in Germany in 2002.

Type	Consumption ($10^6 \text{ t coal equiv./a}$)	Per capita consumption (kW/person)
Oil	185	2.24
Gas	107	1.30
Coal	122	1.48
Nuclear energy	62	0.75
Others	17	0.21
Total	494	5.98

Table 1.2 World Primary Energy Consumption in 2002.

Type	Consumption (10 ⁹ t coal equiv./a)	Per capita consumption (kW/person)
Oil	4.93	0.82
Gas	3.19	0.53
Coal	3.36	0.56
Nuclear energy	0.86	0.14
Others	0.86	0.14
Total	13.2	2.19

Table 1.3 The World's Remaining Energy Reserves.

Type	Reserves in 10 ⁹ t coal equiv.
Oil	210
Gas	170
Coal	660
Total	1040

The global energy consumption of 13.2×10^9 t coal equiv. per year appears to be very small when compared with the continuous energy current from the Sun of

$$1.7 \times 10^{17} \text{ W} = 1.5 \times 10^{18} \text{ kWh/a} = 1.8 \times 10^{14} \text{ t coal equiv./a}$$

that radiates toward the Earth.

In densely populated regions such as Germany, however, the balance is not so favorable if we restrict ourselves to the natural processes of photosynthesis for the conversion of solar energy into other useful forms of energy. The mean annual energy current that the Sun radiates onto Germany, with an area of $0.36 \times 10^6 \text{ km}^2$, is about $3.6 \times 10^{14} \text{ kWh/a} = 4.3 \times 10^{10} \text{ t coal equiv./a}$. Photosynthesis, when averaged over all plants, has an efficiency of about 1% and produces around 400×10^6 t coal equiv./a from the energy of the Sun. This is insufficient to cover the requirements of primary energy of 494×10^6 t coal equiv./a for Germany. What is even more important is that it also shows that over the entire area of Germany plants are not able to reproduce, by photosynthesis, the oxygen that is consumed in the combustion of gas, oil and coal. And this does not even take into consideration that the biomass produced in the process is not stored but decays, which again consumes the oxygen produced by photosynthesis. This estimate also shows that solar energy can cover the energy requirements of Germany over its area only if a substantially higher efficiency for the conversion process than that of photosynthesis can be achieved. The fact that no shortage in the supply of oxygen will

result in the foreseeable future is owed to the wind, which brings oxygen from areas with lower consumption. Nevertheless, well before we run out of oxygen we will be made aware of an increase in the combustion product CO₂.

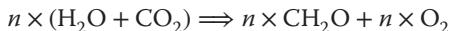
1.2

Estimate of the Maximum Reserves of Fossil Energy

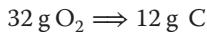
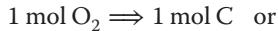
For this estimate [1] we assume that neither free carbon nor free oxygen was present on the Earth before the beginning of organic life. The fact that carbon and oxygen react quickly at the high temperatures prevailing during this stage of the Earth's history, both with each other to form CO₂ and also with a number of other elements to form carbides and oxides, is a strong argument in support of this assumption. Since there are elementary metals on the surface of the Earth even today, although only in small amounts, it must be assumed that neither free carbon nor free oxygen was available to react.

The free oxygen found in the atmosphere today can therefore only be the result of photosynthesis occurring at a later time. The present-day amount of oxygen in the atmosphere thus allows us to estimate a lower limit to the size of the carbon reserves stored in the products of photosynthesis under the assumption that all oxygen produced by photosynthesis is still present as free oxygen.

During photosynthesis, water and carbon dioxide combine to form carbohydrates, which build up according to the reaction



A typical product of photosynthesis is glucose: C₆H₁₂O₆ ≡ 6 × CH₂O. For this compound and also for most other carbohydrates, the ratio of free oxygen produced by photosynthesis to carbon stored in the carbohydrates is



The mass of the stored carbon m_C can therefore be found from the mass of free oxygen m_{O₂}:

$$m_C = \frac{12}{32} m_{O_2}$$

The greatest proportion of the oxygen resulting from photosynthesis is found in the atmosphere and, to a lesser extent, dissolved in the water of the oceans. The fraction in the atmosphere is sufficiently large to be taken as the basis for an estimate.

From the pressure p_E = 1 bar = 10 N cm⁻² on the surface of the Earth resulting from the air surrounding our planet, we can calculate the mass of air from the relationship m_{air} × g = p_E × area:

$$m_{\text{air}} / \text{area} = p_E / g = \frac{10 \text{ N cm}^{-2}}{10 \text{ m s}^{-2}} = 1 \text{ kg cm}^{-2}$$

Multiplying by the surface of the Earth gives the total mass of air

$$m_{\text{air}} = 1 \text{ kg cm}^{-2} \times 4\pi R_{\text{Earth}}^2 = 5 \times 10^{15} \text{ tons of air}$$

Since air consists of 80% N₂ and 20% O₂ (making no distinction between volume percent and weight percent), the mass of oxygen is: $m_{\text{O}_2} = 10^{15} \text{ t O}_2$. The amount of carbon produced by photosynthesis and now present in deposits on the Earth corresponding to the amount of oxygen in the atmosphere is therefore:

$$m_{\text{C}} = \frac{12}{32} m_{\text{O}_2} = 375 \times 10^{12} \text{ tons of carbon}$$

There may be even more. The large deposits of iron oxide in Australia are probably the result of the reaction of iron ions with oxygen produced in the early ages of photosynthesis in shallow sea water.

Up to now 10.4×10^{12} t coal equiv. has been found.

Thus, there is reason to hope that the reserves of fossil energy will continue to grow as a result of continued exploration. In fact, in recent years the known reserves have grown continuously because more has been found than was consumed. There are rumors that very large reserves of methane hydrate can be found in moderate depths on the ocean floor. This compound dissociates into methane and water when it is heated or taken out of the ocean. The prospects of possibly large reserves, however, must not distract our attention from the urgency of restricting the mining of these reserves. If we actually use up the entire reserves of carbon for our energy requirements, we will in fact reverse the photosynthesis of millions of years and in doing so eliminate all our oxygen. Even if more than the estimated 375×10^{12} tons of carbon should exist, we cannot burn more than 375×10^{12} tons of carbon because of the limited amount of oxygen.

If we examine oil and gas consumption as an example for the consumption of fossil energy reserves over a long period of time, for example, since the birth of Christ, we obtain a frightening picture (Figure 1.1). Up to the beginning of the twentieth century, the consumption of reserves was practically negligible. From then, it has been rising exponentially and will reach a maximum value in one or two decades. Consumption will then fall off again as the reserves are gradually

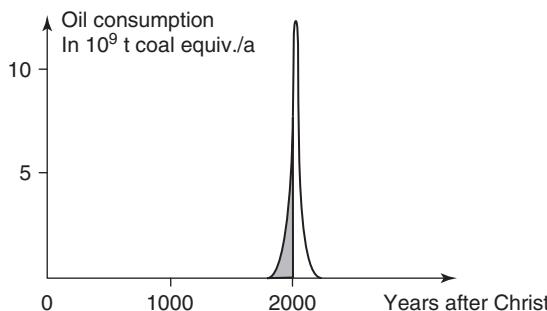


Figure 1.1 Annual consumption of oil. The area under the curve gives the estimated total oil reserves.

used up. The maximum consumption is expected when the reserves have fallen off to one half of their original levels. The reserves that have accumulated over millions of years will then literally go up in smoke over a period of only about 100 years. Here, the elimination of energy reserves is the lesser problem. Much worse will be the alteration of the atmosphere as a result of the products of combustion. These effects will last for a long time. Even if later generations change over to supplying energy from regenerative sources, they will still suffer from the heritage that we have left them.

1.3

The Greenhouse Effect

CO_2 is produced in the combustion of fossil energy carriers. The increase in the concentration of carbon dioxide in our atmosphere will have serious consequences on our climate. Currently, the atmosphere contains a fraction of 0.04% of CO_2 . This corresponds to 3.5×10^{12} t of CO_2 .

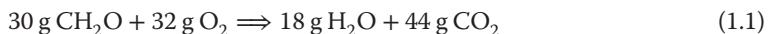
1.3.1

Combustion

Pure carbon is consumed according to the reaction $\text{C} + \text{O}_2 \Rightarrow \text{CO}_2$. Accordingly $12\text{ g C} + 32\text{ g O}_2 \Rightarrow 44\text{ g CO}_2$. The mass of CO_2 produced by combustion is given by the mass of carbon consumed according to $m_{\text{CO}_2} = 44/12m_{\text{C}}$. The combustion of 1 t of carbon thus results in 3.7 t of CO_2 .

For different compounds of carbon other relationships apply:

- Carbohydrates:



resulting in $m_{\text{CO}_2} = 1.47 m_{\text{CH}_2\text{O}}$. This is the chemical reaction for the combustion of food in the human body.

- Methane (main component of natural gas):



resulting in $m_{\text{CO}_2} = 2.75 m_{\text{CH}_4}$.

The present global consumption of the 10^{10} t coal equiv./a produces globally $\Rightarrow 2.2 \times 10^{10}$ t of CO_2 per year. Half of this is dissolved in the water of the oceans and half remains in the atmosphere. If the annual energy consumption does *not* continue to rise, the amount of CO_2 in the atmosphere will double only after about 200 years. However, it is necessary to take into account that energy consumption continues to increase. Currently the growth of 1% per year is relatively low. In the developing countries, energy consumption even decreased in 1999, because of the inability of these countries to pay for more energy. If global energy consumption continues to increase at about 1% per year the CO_2 concentration in

the atmosphere will have doubled after about 100 years. This increase is less the result of a per capita increase in energy consumption than that of the increasing global population. The increasing CO₂ concentration in the atmosphere will have consequences on the temperature of the Earth.

1.3.2

The Temperature of the Earth

The temperature of the Earth is stationary, i.e. constant in time if the energy current absorbed from the Sun and the energy current emitted by the Earth are in balance. We want to estimate the temperature of the Earth in this steady state condition. For this purpose we make use of radiation laws, not derived until the following chapter.

The energy current density from the Sun at the position of the Earth (but outside the Earth's atmosphere) is

$$j_{E,\text{Sun}} = 1.3 \text{ kW m}^{-2}$$

For the case of complete absorption, the energy current absorbed by the entire Earth is the energy current incident on the projected area of the Earth:

$$I_{E,\text{abs}} = \pi R_e^2 j_{E,\text{Sun}} \quad \text{with} \quad R_e = 6370 \text{ km} \quad (\text{radius of the Earth})$$

According to the Stefan–Boltzmann radiation law, the energy current density emitted by the Earth into space is given by

$$j_{E,\text{Earth}} = \sigma T_e^4 \quad \text{where} \quad \sigma = 5.67 \times 10^{-8} \text{ W m}^{-2}\text{K}^{-4}$$

is the Stefan-Boltzmann constant. The energy current emitted by the entire Earth is

$$I_{E,\text{emit}} = 4\pi R_e^2 \sigma T_e^4$$

From the steady state condition $I_{E,\text{abs}} = I_{E,\text{emit}}$, it follows that the estimated mean temperature of the Earth is $T_e = 275 \text{ K}$.

The mean temperature of the Earth is in fact around 288 K. The approximate agreement is, however, only coincidental. Taking into account that about 30% of the incident solar radiation is reflected back into space by the atmosphere of the Earth and thus only about 70% (1 kW m⁻²) reaches the surface of the Earth, a temperature of 258 K then results. The actual temperature of the Earth is in fact greater, because the radiation emitted by the Earth is partly absorbed in the atmosphere. The atmosphere then becomes warmer and emits heat back to the Earth. The same effect occurs in greenhouses, where the glass covering absorbs the thermal radiation emitted from within the greenhouse and emits some of it back into the greenhouse.

We can understand the greenhouse effect of the atmosphere using a simple model. Owing to a temperature of 6000 K of the Sun, the solar radiation spectrum (expressed as energy current per wavelength) has a maximum at a wavelength of about 0.5 μm, at which the atmosphere is transparent. As a result of the lower

temperature of the Earth, the emission from the Earth has its maximum at a wavelength of about $10\text{ }\mu\text{m}$ (in the infrared region). All triatomic molecules, including CO_2 , are good absorbers in the infrared region. Consequently, while most of the incident solar radiation reaches the surface of the Earth, a great part of the energy emitted from the surface is absorbed in the atmosphere. This causes warming of the atmosphere, which in turn leads to heat emission back to the Earth. The temperature of the Earth is at maximum when the radiation from the surface of the Earth is completely absorbed by the atmosphere, a situation that will be faced if the atmospheric concentration of CO_2 continues to rise.

In our model, we assume that the Earth's surface absorbs all radiation incident upon it from the Sun and the atmosphere. In the steady state, it must also emit the same energy current. All energy emitted in the infrared region is assumed to be absorbed by the atmosphere. This leads to the condition of Figure 1.2:

$$I_{E,\text{Earth}} = I_{E,\text{Sun}} + \frac{1}{2}I_{E,\text{atm}}$$

Since radiation emitted by the Earth's surface is fully absorbed by the atmosphere, the solar energy current incident on the surface of the Earth can only be emitted into space by emission from the atmosphere. This leads to the following conditions:

$$\frac{1}{2}I_{E,\text{atm}} = I_{E,\text{Sun}} \quad \text{and} \quad I_{E,\text{Earth}} = 2I_{E,\text{Sun}}$$

It then follows that

$$I_{E,\text{Earth}} = 4\pi R_e^2 \sigma T_{e,\text{greenhouse}}^4 = 2\pi R_e^2 \times 1.3 \text{ kW m}^{-2}$$

This yields a temperature of

$$T_{e,\text{greenhouse}} = \sqrt[4]{2} T_e = 1.19 \times 275 \text{ K} = 327 \text{ K} = 54^\circ\text{C}$$

With this mean temperature, the Earth would be virtually uninhabitable.

What makes the problem of the increased absorption of infrared radiation in the atmosphere due to human influences even worse is that only one half of the present greenhouse effect is caused by the increasing CO_2 concentration. The other half results from methane, fluorinated hydrocarbons, and nitrogen oxides.

To estimate the greenhouse effect, we have treated the atmosphere as a fire screen, allowing the solar radiation to pass through but absorbing the radiation from the Earth. In view of the fact that the temperature of the atmosphere is not

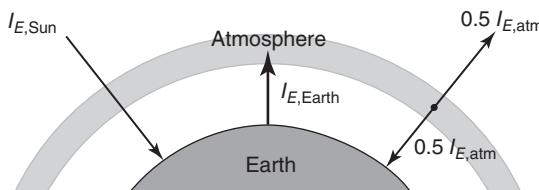


Figure 1.2 Balance of the absorbed and emitted energy currents on the surface of the Earth.

uniform, a better description would be in the form of several fire screens placed behind each other. Extending the fire-screen model to n fire screens, the temperature of the Earth's surface becomes

$$T_{e, \text{greenhouse}} = T_e \sqrt[4]{n + 1}$$

For large values of n , the temperature of the Earth could become arbitrarily large and, in particular, even greater than the temperature of the Sun. This result is of course not correct, because at such high temperatures an essential condition of the model is no longer satisfied. If the temperature of the Earth were equal to that of the Sun, the spectrum emitted from the Earth would be identical with the solar spectrum. It would then no longer be possible for the solar radiation to pass through the atmosphere, while the radiation emitted from the Earth would be absorbed, as assumed for the fire-screen model.

As long as the requirements are satisfied, with more than one fire screen a higher temperature results. The conditions prevailing on Venus clearly show this. The distance of Venus from the Sun is 0.723 times the distance of the Earth from the Sun. The energy current density of the solar radiation at the position of Venus is therefore twice that reaching the Earth. The atmosphere of Venus is composed almost entirely of CO₂. Treating the atmosphere of Venus as a single fire screen yields a temperature of 116 °C for Venus, a factor of $\sqrt[4]{2}$ greater than that of the Earth. The actual temperature of Venus is, however, 475 °C.

The number of fire screens required to describe the atmosphere depends on the mean free path of the emitted radiation after which it is absorbed in the atmosphere. This determines the spacing of the fire screens in the model. Because of the high density of Venus' atmosphere, with a pressure of 90 bar, this mean free path is much shorter on Venus than on Earth.

Fortunately, we do not have to worry about temperatures on the Earth reaching those found on Venus. Even if the entire supply of oxygen on the Earth were consumed, resulting in a CO₂ pressure of about 0.2 bar, the temperature of the Earth could never reach the temperature of Venus. Nevertheless, serious alterations are already occurring even with much lower temperature increases, which are not only possible but in fact very probable on Earth. We also have to consider that there are numerous feedback effects. The most dangerous would be the release of large quantities of methane, an even more effective greenhouse gas than CO₂, when methane hydrate melts in a warmer ocean.

1.4 Problems

- 1.1 Describe in a few sentences the greenhouse effect
 - (a) in a greenhouse for crops
 - (b) in the Earth's atmosphere.
- 1.2 Why does burning methane lead to a higher greenhouse activity than burning (the same mass of) carbohydrates?

- 1.3 How long may a computer (50 W) be operated with an amount of energy of 1 kg of coal equiv.?
- 1.4 Assume the global human energy consumption to be 15×10^9 t coal equiv./a by now and the Earth's remaining consumable energy reserves to be 2×10^{12} t coal equiv. How long will these reserves last with the energy consumption
 - (a) increasing linearly, doubling in 35 years?
 - (b) increasing exponentially, doubling in 40 years?
- 1.5 Calculate the amount of methane present on Earth if all the oxygen in the atmosphere would have resulted from the production of methane from CO₂. Consider a total mass of oxygen of $m_{O_2} = 10^{15}$ t.
- 1.6 Why has Venus such a high surface temperature? Venus is the brightest object in the sky (except for the Sun and Moon) due to its high reflectance. This so-called albedo of Venus is $r_v = 0.7$. Calculate the intensity reaching the surface of Venus. Knowing the intensity and the surface temperature (475 °C), determine the number of fire screens according to the model in Section 1.3.2.

2 Photons

Photons are particles of light. Photons to which the human eye reacts, that is, those that we see, have energies $\hbar\omega$ between 1.5 and 3 eV. They always move with the velocity of light, in a vacuum with a velocity of $c_0 = 3 \times 10^8 \text{ m s}^{-1}$, and in a medium with refractive index n with a velocity of $c = c_0/n$. The fact that we can also describe light as an electromagnetic wave is in no way contradictory. The square of the field strength of the electromagnetic wave describes the location of the photons. In contrast to the behavior familiar from shot pellets, photons obey the laws of quantum mechanics and not those of the more common Newtonian mechanics. The differences from Newtonian mechanics become apparent only for particles of very low energy (including the mass). Since it is difficult to visualize particles that do not move along a straight line, in a dualistic approach the wave property is generally invoked to describe diffraction and interference phenomena and the particle property is used to describe the quantumlike transport of energy, as in this book.

2.1

Black-body Radiation

A *black body* is defined as a body that completely absorbs radiation of all photon energies $\hbar\omega$. Its absorptance is

$$a(\hbar\omega) = 1 \quad (2.1)$$

A possible example of a black body is a tiny hole in a cavity. Photons incident on the hole are not reflected, but are absorbed inside the cavity. The hole is described as black, because it appears black when the inner part of the cavity is at a low temperature and very few photons are emitted by the walls of the cavity. When the inner part is at a high temperature, the cavity is like a furnace. We then see a very bright hole, through which we can view the interior of the furnace. Nevertheless, regardless of whether it emits radiation, the hole absorbs all incident radiation and is therefore still described as a black body.

The Sun is a black body, as is every body that is sufficiently thick and nonreflecting.

2.1.1

Photon Density n_γ in a Cavity (Planck's Law of Radiation)

Planck's law of radiation describes the photon density $dn_\gamma(\hbar\omega)$ in a cavity for photon energies between $\hbar\omega$ and $\hbar\omega + d\hbar\omega$. As always for the definition of particle densities, we first determine the density D of states for these particles. The particle density is then found by applying an occupation or distribution function f . For photons, this takes the form

$$dn_\gamma(\hbar\omega) = D_\gamma(\hbar\omega) f_\gamma(\hbar\omega) d\hbar\omega \quad (2.2)$$

Here $D_\gamma(\hbar\omega)$ is the density (per volume and per energy interval $d\hbar\omega$) of the states that the photons can occupy and $f_\gamma(\hbar\omega)$ is the probability of occupation or the distribution function that determines the distribution of photons over the states as a function of the energy $\hbar\omega$.

Distribution Function $f_\gamma(\hbar\omega)$ for Photons

Because of their integral spin, the Bose–Einstein distribution (Figure 2.1), which describes the probability for the occupation of states with the energy $\hbar\omega$, applies for the photons:

$$f_\gamma(\hbar\omega) = \frac{1}{\exp[(\hbar\omega - \mu_\gamma)/kT] - 1} \quad (2.3)$$

Here, and always in the combination kT , k denotes the Boltzmann constant $k = 8.617 \times 10^{-5}$ eV K $^{-1}$, and μ_γ is the chemical potential of the photons, which

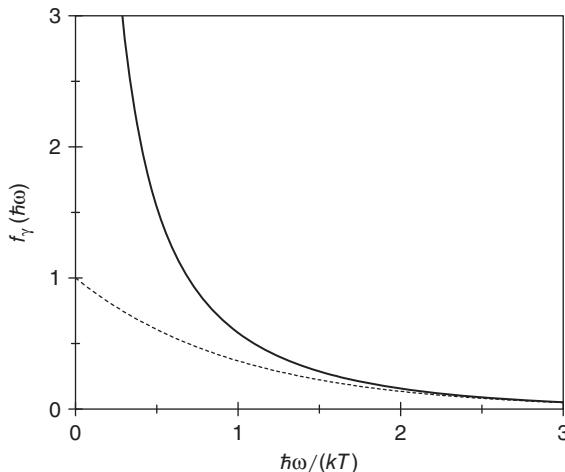


Figure 2.1 Bose–Einstein distribution function for $\mu_\gamma = 0$. The dashed line represents the Boltzmann distribution $\exp(-\hbar\omega/kT)$, a good approximation for $\hbar\omega > 2kT$.

has the value $\mu_\gamma = 0$ for solar radiation and in general for thermal radiation. Since the distribution function $f_\gamma(\hbar\omega)$ depends only on the energy $\hbar\omega$ of the photons, we must also know the density of the states $D_\gamma(\hbar\omega)$ as a function of the energy.

Density of States $D_\gamma(\hbar\omega)$ for Photons

Two particles can be distinguished only if they are in different states. Two states are different when their location x and momentum p differ by more than the minimum of the uncertainty Δx , Δp of one of these states. For one dimension, this Heisenberg uncertainty principle has the form

$$\Delta x \Delta p_x \geq h \quad (2.4)$$

Particles that cannot be distinguished as a result of the uncertainty principle are in the same state. In geometrical and momentum space, a state therefore has the "phase space" volume

$$\Delta x \Delta p_x \Delta y \Delta p_y \Delta z \Delta p_z = h^3$$

Since the photons are not localized in the cavity, the uncertainty in their position must be taken as the entire volume of the cavity, extending in the x , y , and z directions by L_x , L_y , and L_z . Thus,

$$\Delta x = L_x \quad \text{and consequently} \quad \Delta p_x = h/L_x$$

Possible values for the x component of the momentum lie in intervals with a width of h/L_x with

$$p_x = 0, \pm \frac{h}{L_x}, \pm \frac{2h}{L_x}, \dots$$

The relationships for Δp_y and Δp_z are analogous. States having the volume $V = L_x L_y L_z$ in geometrical space then have the volume $\Delta p_x \Delta p_y \Delta p_z = \Delta p^3 = h^3/V$ in momentum space. Figure 2.2 illustrates the volume of the states and their homogeneous distribution in momentum space. This result is based solely on the Heisenberg uncertainty principle and does not contain any specific particle properties and for this reason it is not only valid for photons but also for electrons, which will be treated later.

A property specific to light is its occurrence in two directions of polarization normal to each other with independent intensities. It is therefore necessary to consider two independent types of photons (with opposite spins). Each volume

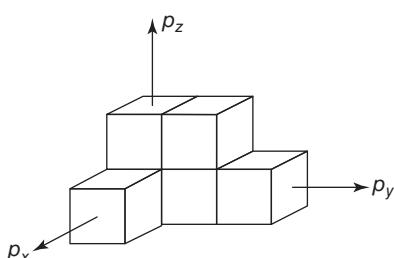


Figure 2.2 Volumes of the states in momentum space.

$\Delta p^3 = h^3/V$ in momentum space therefore contains two photon states, one for each type of photon.

Since the distribution function $f_\gamma(\hbar\omega)$ for the photons depends on energy and not on momentum, we need the density of photon states per energy interval, rather than the density of states in momentum space. For this we need the relationship between momentum p and energy $\hbar\omega$

$$\hbar\omega = c |p|$$

which is a specific property of photons.

All states in which the photons have energies $\hbar\omega' < \hbar\omega$ and thus $|p'| < |p|$ lie within a sphere of radius $|p| = \hbar\omega/c$ in momentum space. The number of such states is

$$N_\gamma = \frac{2 \times (4\pi/3)p^3}{(h^3/V)} = \frac{(8\pi/3)V(\hbar\omega)^3}{h^3c^3} \quad (2.5)$$

From Equation 2.5, we obtain the density of states per volume and energy interval as the increase dN_γ resulting from an increase in the energy of $d\hbar\omega$.

$$D_\gamma(\hbar\omega) = \frac{1}{V} \times \frac{dN_\gamma(\hbar\omega)}{d\hbar\omega} = \frac{(\hbar\omega)^2}{\pi^2 \hbar^3 c^3} \quad (2.6)$$

Here $\hbar = h/2\pi$. The velocity c of the photons is their velocity in the cavity. If the cavity is empty, the photons have the velocity of light in vacuum c_0 . With increasing index of refraction the velocity c_0/n of the photons decreases, resulting in an increasing density of states for the photons.

A problem results when photons pass from one medium to another with a lower index of refraction, i.e. from a higher to a lower density of states, such as from glass to air. To accommodate all photons in the low-index material with fewer photon states, the occupation function f_γ would have to increase. This, however, is not allowed by the second principle of thermodynamics, as will be shown later, since it would result in annihilation of entropy. For the same value of the distribution function $f_\gamma(\hbar\omega)$, to keep the entropy constant, only part of the photons can be accommodated in the low-index material. Nature chooses to totally reflect back into the high-index material the photons that cannot be accommodated in the fewer states of the medium with the lower index of refraction. Exchange of photons between two media is restricted to the same number of states in both media in such a way that all states in a restricted solid angle in the high-index medium communicate with all states in an unrestricted solid angle in the low-index medium. Snell's law of refraction is another consequence of this principle. In addition to total internal reflection, a discontinuity of the refractive index at an interface causes reflection even for photons within the solid angle range in which photons are exchanged, e.g. all photons incident from the low-index medium. This type of reflection, however, can be eliminated by an antireflection coating, in contrast to the total internal reflection of the photons incident from the high-index medium, but with momenta outside the restricted solid angle.

The density of states (Equation 2.6) accounts for all photon states, regardless of their direction of motion, which is the direction of their momentum p . At every location within the cavity, the motion of the photons is in all directions, i.e. their motion is isotropic. Of interest to us here are the photons that are able to leave the cavity through the hole. At any location within the cavity, these are the photons moving toward the hole, i.e. having a momentum vector pointing in this direction. Since the photons do not collide with each other and maintain their direction, we can find these photons by determining their density per solid angle at every location inside the cavity and multiply the density per solid angle by the solid angle element $d\Omega$ subtended by the hole. Imagine a sphere with a radius R given by the distance to the hole, around the location of interest of the photon density per solid angle $d\Omega$. This solid angle element $d\Omega$ includes all photons passing through the surface element dO of the sphere, which defines the hole. The solid angle is defined as $d\Omega = dO/R^2$.

The maximum value of Ω is 4π , for a solid angle, which includes all directions. For an isotropic distribution of states in momentum space, the density of states per solid angle is

$$D_{\gamma,\Omega}(\hbar\omega) = \frac{D_\gamma(\hbar\omega)}{4\pi}$$

Energy Distribution of Photons

The density of states and the distribution function lead us to Planck's law of radiation, which describes the number of photons per volume and per photon energy interval $d\hbar\omega$ in the solid angle element $d\Omega$.

$$\frac{dn_\gamma(\hbar\omega)}{d\hbar\omega} = D_{\gamma,\Omega}(\hbar\omega) f_\gamma(\hbar\omega) d\Omega = \frac{(\hbar\omega)^2 d\Omega}{4\pi^3 \hbar^3 (c_0/n)^3} \frac{1}{\exp(\hbar\omega/kT) - 1} \quad (2.7)$$

The photon density per photon energy interval $dn_\gamma/d\hbar\omega$ has its maximum value at $\hbar\omega_{\max} = 1.59 kT$.

With the energy per photon $e_\gamma = \hbar\omega$, we find the energy of the photons per volume and per photon energy interval $d\hbar\omega$ in the solid angle interval $d\Omega$

$$\frac{de_\gamma(\hbar\omega)}{d\hbar\omega} = D_{\gamma,\Omega}(\hbar\omega) f_\gamma(\hbar\omega) \hbar\omega d\Omega = \frac{(\hbar\omega)^3 d\Omega}{4\pi^3 \hbar^3 (c_0/n)^3} \frac{1}{\exp(\hbar\omega/kT) - 1} \quad (2.8)$$

The energy density per photon energy $de_\gamma/d\hbar\omega$ is called a spectrum. It has its maximum value at a photon energy of

$$\hbar\omega_{\max} = 2.82 kT \quad (2.9)$$

In the literature, another spectrum is often used, the energy density per wavelength $de_\gamma/d\lambda$ as a function of λ . With the relationship

$$\hbar\omega = h\nu = \frac{hc}{\lambda} \quad \text{from which follows} \quad d\hbar\omega = -\frac{hc}{\lambda^2} d\lambda \quad (2.10)$$

(Equation 2.8) is transformed into

$$\frac{de_\gamma(\lambda)}{d\lambda} = \frac{2hc d\Omega}{\lambda^5} \frac{1}{\exp[hc/(\lambda kT)] - 1} \quad (2.11)$$

We have left out the minus sign contained in Equation 2.10, which simply means that the photon energy $\hbar\omega$ decreases with increasing wavelength λ . The maximum value of $de_\gamma/d\lambda$ is at a wavelength

$$\lambda_{\max} = \frac{hc}{4.965 kT} = 0.2497 \frac{\mu\text{m eV}}{kT}$$

To convert photon energies $\hbar\omega$ to wavelengths λ , we make use of Equation 2.10:

$$\hbar\omega\lambda = hc = 1.240 \text{ eV } \mu\text{m} \quad (2.12)$$

From Equation 2.8 the total energy of radiation per volume in the cavity, or the energy density, is

$$e_\gamma = \int_0^\infty \frac{(\hbar\omega)^3 d\hbar\omega}{4\pi^3 \hbar^3 (c_0/n)^3 [\exp(\hbar\omega/kT) - 1]} \int_0^{4\pi} d\Omega \quad (2.13)$$

Setting $x = \hbar\omega/kT$, the energy density is

$$e_\gamma = \frac{(kT)^4}{4\pi^3 \hbar^3 (c_0/n)^3} \underbrace{\int_0^\infty \frac{x^3 dx}{e^x - 1}}_{\pi^4/15} 4\pi = \frac{\pi^2 k^4}{15 \hbar^3 (c_0/n)^3} T^4 \quad (2.14)$$

The same result is of course obtained by integrating Equation 2.11 over the wavelength.

With reference to Equation 2.7, the density of all photons in the cavity becomes

$$n_\gamma = \frac{(kT)^3}{4\pi^3 \hbar^3 (c_0/n)^3} \underbrace{\int_0^\infty \frac{x^2 dx}{e^x - 1}}_{2.40411} 4\pi = \frac{2.40411 k^3}{\pi^2 \hbar^3 (c_0/n)^3} T^3 \quad (2.15)$$

The average photon energy of blackbody radiation is thus given by

$$\langle \hbar\omega \rangle = \frac{e_\gamma}{n_\gamma} = 2.701 kT \quad (2.16)$$

2.1.2

Energy Current Through an Area dA into the Solid Angle $d\Omega$

The energy density per solid angle inside the cavity is $e_\Omega = e_\gamma/4\pi$ for the entire spectrum. It is transported by photons, which move with the velocity c . We would like to find the energy current through a small hole in the cavity with surface area dA , directed into a small solid angle element $d\Omega$ normal to dA . Figure 2.3 shows that those photons, which at time $t = 0$, are within the volume $dV = L^2 d\Omega dL$ and are moving toward the hole will fly through the hole at time $t = L/c$ for a time interval $dt = dL/c$. With them the energy flows through the hole dA and into the solid angle element $d\Omega$. For an isotropic distribution, the fraction $d\Omega/4\pi = dA/4\pi L^2$ of the energy $dE = e_\gamma dV$ contained in the volume element dV flows through the hole during the time dt , and the energy current is

$$dI_E = \frac{e_\gamma}{4\pi} \frac{dL}{dt} \frac{dA}{L^2} L^2 d\Omega = e_\Omega c dA d\Omega \quad (2.17)$$

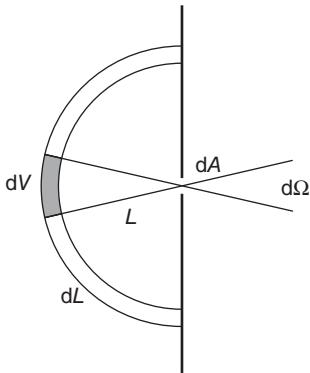


Figure 2.3 An energy current I_E originating from the volume element dV is flowing through the hole with area dA of the cavity, into the solid angle $d\Omega$.

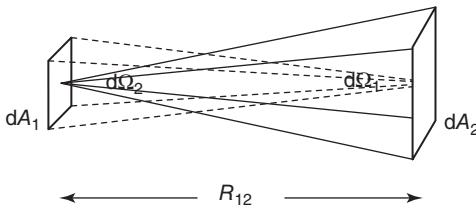


Figure 2.4 Solid angle $d\Omega_2$ under which the receiver area dA_2 is seen from the emitter dA_1 , and solid angle $d\Omega_1$ under which dA_1 is seen from dA_2 .

and is as large as one would expect from the homogeneous energy density at the location of the hole. This equation also tells us that the point source radiator ($dA = 0$), so popular in geometrical optics, would require an infinitely large energy density for a nonzero energy current and, therefore, cannot exist.

Let us now examine the propagation of the radiation toward a receiver. The energy current emitted from a radiating area dA_1 toward a receiver area dA_2 at a distance of R_{12} can be described from two different points of view (Figure 2.4):

- Viewed from the radiation source, the energy current dI_{E1} is emitted from all points of the surface area dA_1 into the solid angle $d\Omega_2 = dA_2/R_{12}^2$.
- Viewed from the receiver, the energy current dI_{E2} is incident from the solid angle $d\Omega_1 = dA_1/R_{12}^2$ onto all points of the surface area dA_2 .

Both of course represent the same energy current

$$dI_{E1} = e_{\Omega 1} c dA_1 \frac{dA_2}{R_{12}^2} = dI_{E2} = e_{\Omega 2} c dA_2 \frac{dA_1}{R_{12}^2} \quad (2.18)$$

We see that the energy density per solid angle $e_{\Omega 1} = e_{\Omega 2}$ is the same at dA_1 and at dA_2 and thus does not change during the energy transport through the vacuum. In the same way, the energy current density per solid angle

$$j_{E,\Omega} = e_{\Omega} c$$

also does not change along the path from the radiation source to the receiver. The fact that the energy current dI_{E2} incident on the receiver surface decreases with

increasing distance R_{12} from the emitter follows simply from the decreasing magnitude of the solid angle $d\Omega_1 = dA_1/R_{12}^2$ under which the emitter is seen.

Although we define solid angles by the size and distance of distant objects, it must be emphasized that the solid angle $d\Omega$ into which radiation is emitted at the location of the emitter (or from which radiation is received at the location of the receiver) is a local variable that describes the momentum distribution of the photons (with respect to direction) at this location.

According to Figure 2.5, the energy current from a surface element dA_s of the Sun incident upon a surface element dA_e of the Earth when dA_s and dA_e are both normal to a line connecting the Sun and the Earth over the distance R_{se} is given by

$$\begin{aligned} dI_{E,\text{Sun-Earth}} &= e_{\Omega,\text{Sun}} c d\Omega_e dA_s \\ &= e_{\Omega,\text{Sun}} c \frac{dA_e dA_s}{R_{se}^2} \\ &= e_{\Omega,\text{Sun}} c dA_e d\Omega_s \end{aligned} \quad (2.19)$$

It follows from the first expression that the energy current density per solid angle on the Sun at a point of the surface element dA_s is

$$j_{E,\Omega,\text{Sun}} = \frac{dI_{E,\text{Sun-Earth}}}{dA_s d\Omega_e} = e_{\Omega,\text{Sun}} c \quad (2.20)$$

In accordance with the third line of Equation 2.19 this is identical with the energy current density per solid angle on the Earth at a point on the surface element dA_e

$$j_{E,\Omega,\text{Earth}} = \frac{dI_{E,\text{Sun-Earth}}}{dA_e d\Omega_s} = e_{\Omega,\text{Sun}} c = j_{E,\Omega,\text{Sun}}$$

This is an important result. Even though the energy current per surface area, the energy current density, decreases with distance from the Sun, it is of course much larger near the Sun than on the Earth; the energy current density per solid angle remains unchanged. Close to the Sun, the Sun simply subtends a larger solid angle than when viewed from a location on the Earth.

A long, thin tube, which when you look through it defines a small solid angle, can be used as an instrument to measure $j_{E,\Omega}$ as long as the emitter fills a larger solid angle than defined by the tube.

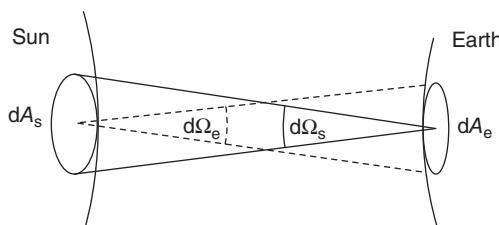


Figure 2.5 Solid angle $d\Omega_s$ under which the surface element dA_s on the Sun is viewed from the Earth, and solid angle $d\Omega_e$ under which the surface element dA_e on the Earth is viewed from the Sun.

To find the energy current emitted from the entire Sun toward the Earth, we must integrate over the surface of the Sun, taking into account that not all surface elements are oriented normally to the line connecting the Earth and the Sun.

2.1.3

Radiation from a Spherical Surface into the Solid Angle $d\Omega$

We can determine the energy current density dI_E that is radiated into the solid angle $d\Omega_e$ in a direction other than normal to the surface element dA_s by just looking at the Sun. In spite of its spherical shape, we see the Sun as a uniformly bright disc (apart from a slight drop in intensity very close to the edge). This enables us to conclude that each surface element on the Sun that *appears* to have the same size, when viewed from the Earth, in fact radiates the same energy current in our direction. The apparent size dA'_s of a surface element dA_s in Figure 2.6 is the projection of dA_s onto a plane normal to the line connecting the Earth and the Sun, that is,

$$dA'_s = dA_s \cos \vartheta$$

and the energy current emitted toward the Earth, (i.e. into $d\Omega_e$) is

$$dI_{E,e} = j_{E,\Omega} d\Omega_e dA_s \cos \vartheta$$

This dependence is called Lambert's law.

Figure 2.6 shows that all surface elements located on a ring around the line connecting the Earth and the Sun are positioned at the same angle in relation to this line. The radius of this ring is $r = R_s \sin \vartheta$, where R_s is the radius of the Sun. The area of the ring is

$$dA_s = 2\pi r R_s d\vartheta = 2\pi R_s^2 \sin \vartheta d\vartheta$$

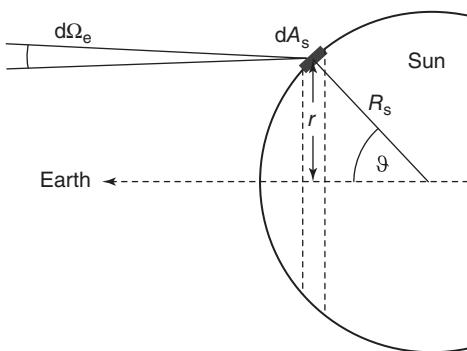


Figure 2.6 Energy current from the surface element dA_s on the surface of the Sun into the solid angle $d\Omega_e$ under which the Earth is viewed from the Sun.

Since only the hemisphere directed toward us radiates energy to the Earth, the total energy current that the Sun radiates to the Earth is

$$I_{E,e} = j_{E,\Omega} d\Omega_e \int_0^{\pi/2} 2\pi R_s^2 \sin \vartheta \cos \vartheta d\vartheta \quad (2.21)$$

Since $\cos \vartheta d\vartheta = d(\sin \vartheta)$ the integration is simple, and

$$I_{E,e} = j_{E,\Omega} d\Omega_e \pi R_s^2 \quad (2.22)$$

This is exactly the result that we had expected: the energy current emitted from the Sun to the Earth is the same as that of a flat disc with a radius R_s oriented perpendicularly to the line connecting the Earth and the Sun.

2.1.4

Radiation from a Surface Element into a Hemisphere (Stefan– Boltzmann Radiation Law)

Let us now imagine a hemisphere of radius R surrounding the surface element dA from which radiation is emitted. All surface elements lying on a ring on the hemisphere are then seen from dA at the same angle relative to the normal to the surface element. The rest follows in the same way as in Section 2.1.3. The area of the ring $dO = 2\pi R^2 \sin \vartheta d\vartheta$ defines a solid angle element $d\Omega = dO/R^2 = 2\pi \sin \vartheta d\vartheta$. Integration over all energy currents $dI_E = j_{E,\Omega} d\Omega dA \cos \vartheta$ then yields the total energy current emitted from the surface element dA into the hemisphere

$$I_E = j_{E,\Omega} \pi dA \quad (2.23)$$

The energy current density emitted from the surface element dA into a hemisphere is

$$j_E = j_{E,\Omega} \pi = e_{\gamma,\Omega} c \pi = ce_\gamma / 4 \quad (2.24)$$

Making use of Equation 2.14 for the energy density e_γ of black-body radiation then gives the Stefan– Boltzmann radiation law

$$j_E = \frac{\pi^2 k^4}{60 \hbar^3 c^2} T^4 = \sigma T^4 \quad (2.25)$$

The value of the Stefan– Boltzmann constant $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$. The Stefan– Boltzmann law for the emission of black bodies was originally discovered experimentally in 1879 by Stefan and later derived theoretically by Boltzmann. We keep in mind that because of the dependence of the energy current density on $\cos \vartheta$ at the location of the emitter, as shown in Figure 2.7, the total energy current density emitted into the hemisphere ($\Omega = 2\pi$) is obtained by multiplying the energy current density per solid angle by π , as the effective solid angle for the emission of a planar black body into the hemisphere.

Equation 2.24 also allows us to write the energy density per solid angle in the form

$$e_\Omega = \frac{\sigma}{\pi c} T^4 \quad (2.26)$$

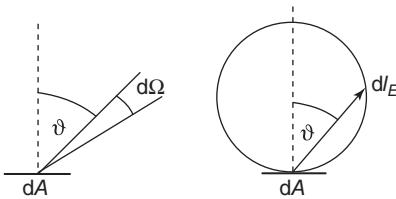


Figure 2.7 Dependence of the emitted energy current dI_E on the angle ϑ relative to the normal to the surface. The length of the arrow on the right is a measure of the magnitude of dI_E .

For a black body, the energy current density per solid angle is

$$j_{E,\Omega} = \frac{\sigma T^4}{\pi} \quad (2.27)$$

and the energy current density emitted normal to a surface into the small solid angle element $d\Omega$ is

$$j_E = \sigma T^4 \frac{d\Omega}{\pi} \quad (2.28)$$

As already mentioned, the dependence of the energy current on $\cos \vartheta$ means that a light-emitting surface appears equally bright at all angles.

As a consequence of this so-called Lambert's law for an emitting black, planar surface, the energy current density per solid angle observed from outside the surface $j_{E,\Omega}$ is independent of the angle ϑ at which the emitting surface is viewed. The dependence of the energy current density j_E on the angle of viewing results from a dependence of the solid angle subtended by the emitting area. This behavior precludes the recognition of any further details of a body by the emitted radiation, except for its contour.

This so-called Lambert behavior cannot be taken for granted. It applies strictly only for the surfaces of bodies that absorb all radiation incident upon them, i.e. black bodies. Weakly absorbing bodies such as a sheet of transparent plastic, in which low concentrations of dye molecules are dissolved, behave differently. These molecules emit light isotropically (e.g. as luminescence radiation when being excited by higher-energy photons). This light is not attenuated by absorption and remains isotropic. If we look onto the large surface of the pane, we see all molecules, just as if we look onto the edge of the pane, because as a result of their low concentration the dye molecules do not conceal each other. The eye perceives the same number of photons from each molecule, regardless of the direction from which we look at the pane. The edge therefore appears much brighter than the large surfaces. This property is common to luminescent plastic jewelry. Our interpretation is strictly valid only for materials with refractive index $n = 1$; for $n > 1$ the emission is less isotropic.

Very different designations are in use for energy transport by radiation. In this book, we will continue to use the terms:

Energy current	I_E	(measured in W)
Energy current density	j_E	(measured in W m^{-2})
Energy current density per solid angle	$j_{E,\Omega}$	(measured in $\text{W m}^{-2}\text{sr}^{-1}$).

These terms are physically clear and we think that the name for the transport of energy should not depend on the medium by which the energy is transported.

2.2

Kirchhoff's Law of Radiation for Nonblack Bodies

In the cavity of Figure 2.8 two plates are positioned opposite each other at a distance that is small compared to their lateral dimensions (different from the drawing). Plate 2 on the right is black, like the hole of a cavity, and absorbs all incident radiation completely ($\alpha_2 = 1$). Plate 1 on the left represents a real material. It reflects (according to its reflectance r_1), transmits (according to its transmittance t_1), and absorbs (according to its absorptance α_1) the respective part of the radiation incident from the black plate $dj_{E,2}$ or from the black walls of the cavity $j_{E,w}$. Since reflection, transmission, and absorption depend on the photon energy, we restrict ourselves to the exchange of photons with an energy between $\hbar\omega$ and $\hbar\omega + d\hbar\omega$, by inserting a filter between the plates. Radiation with other photon energies is perfectly reflected by the filter. The energy current density emitted by the black plate 2 in the energy interval $d\hbar\omega$ at the temperature T_2 with the energy current density $dj_{E,2}(\hbar\omega, T_2)$ follows from Planck's law of radiation in Equation 2.8 and is

$$dj_E(\hbar\omega, T_2) = \frac{d\Omega}{4\pi^3 \hbar^3 (c_0/n)^2} \frac{(\hbar\omega)^3}{\exp(\hbar\omega/kT_2) - 1} d\hbar\omega \quad (2.29)$$

The radiation incident on plate 1 is either reflected, transmitted, or absorbed, so that

$$r_1(\hbar\omega) + t_1(\hbar\omega) + \alpha_1(\hbar\omega) = 1 \quad (2.30)$$

Plate 1 also emits radiation, which is $\varepsilon_1(\hbar\omega)dj_{E,1}$ and differs from the energy current $dj_{E,1}$ emitted by a black body of the same temperature T_1 by its (still unknown) emittance $\varepsilon_1(\hbar\omega)$.

In a steady state, each plate must emit the same amount of energy as it absorbs. We examine the energy balance for the black body, plate 2. The energy current

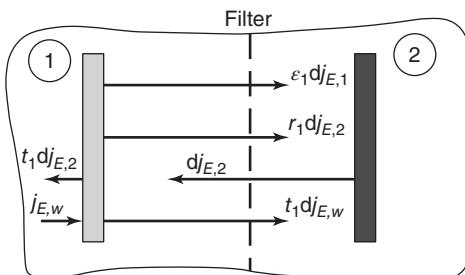


Figure 2.8 Exchange of radiation between two plates in a cavity with a filter in between, transmitting only photons with energy $\hbar\omega$.

incident onto the black body, plate 2, is composed of the energy reflected by plate 1, the energy that plate 1 emits at its temperature T_1 , and the energy from the black cavity walls at the temperature T_w transmitted by plate 1. Since the absorptance of plate 2 is $a_2(\hbar\omega) = 1$, all the incident energy is absorbed and must be balanced by the emitted energy current $dj_{E,2}$, which is transmitted by the filter:

$$\begin{aligned} dj_{E,2}(\hbar\omega, T_2) &= r_1(\hbar\omega) dj_{E,2}(\hbar\omega, T_2) + \varepsilon_1(\hbar\omega) dj_{E,1}(\hbar\omega, T_1) \\ &\quad + t_1(\hbar\omega) dj_{E,w}(\hbar\omega, T_w) \end{aligned} \quad (2.31)$$

Thermal equilibrium of course prevails in the cavity, and the exchange of radiation between the plates must not disturb this state. If the exchange of radiation were to cause differences in temperature, one could set up a thermoelectric generator for the generation of electrical energy, which would convert thermal energy completely into electricity. However, this would have to destroy entropy in contradiction with the second principle of thermodynamics.

With $T_1 = T_2 = T_w$ and consequently $dj_{E,1} = dj_{E,2} = dj_{E,w}$ in Equation 2.31, we obtain

$$r_1(\hbar\omega) + t_1(\hbar\omega) + \varepsilon_1(\hbar\omega) = 1 \quad (2.32)$$

We have left out the energy currents emitted by plate 2 toward the black cavity wall together with the energy current received by plate 2 from the wall, since they cancel each other. Comparing Equation 2.32 with Equation 2.30, we find

$$\varepsilon_1(\hbar\omega) = 1 - r_1(\hbar\omega) - t_1(\hbar\omega) = a_1(\hbar\omega) \quad (2.33)$$

The emittance $\varepsilon(\hbar\omega)$ of a body at the photon energy $\hbar\omega$ is equal to its absorptance $a(\hbar\omega)$ at the same photon energy. This relationship is Kirchhoff's law of radiation stating that the energy current density emitted by a nonblack body in the energy interval $d\hbar\omega$ into the solid angle element $d\Omega$ is

$$dj_E(\hbar\omega) = \frac{a(\hbar\omega)}{4\pi^3\hbar^3(c_0/n)^2} \frac{(\hbar\omega)^3}{\exp(\hbar\omega/kT) - 1} d\hbar\omega \quad (2.34)$$

It is transported by the emitted photon current density

$$dj_\gamma(\hbar\omega) = \frac{a(\hbar\omega)}{4\pi^3\hbar^3(c_0/n)^2} \frac{(\hbar\omega)^2}{\exp(\hbar\omega/kT) - 1} d\hbar\omega \quad (2.35)$$

The absorptance $a(\hbar\omega)$ is characteristic of a body and is a function of the body's geometry. It describes the fraction of the incident light that is absorbed by the body. It is related to the absorption coefficient $\alpha(\hbar\omega)$, which is a material property and independent of the geometry of a body. Neglecting multiple reflections, the transmittance of a plate of thickness d is

$$t(\hbar\omega) = [1 - r(\hbar\omega)] \exp[-\alpha(\hbar\omega)d] \quad (2.36)$$

From $a = 1 - r - t$, we obtain the absorptance of a plate of thickness d

$$a(\hbar\omega) = [1 - r(\hbar\omega)] \{1 - \exp[-\alpha(\hbar\omega)d]\} \quad (2.37)$$

2.2.1

Absorption by Semiconductors

In the next chapter, we learn that solar cells are made from semiconductors. The absorption of solar radiation by solar cells is therefore characterized by a threshold photon energy $\hbar\omega_G$. Only photons with energies greater than $\hbar\omega_G$ can be absorbed. Photons with lower energies are either reflected or transmitted. Let us assume that reflection has been eliminated by an appropriate antireflection coating and that the semiconductor absorber is thick enough so that no photons with energy $>\hbar\omega_G$ are transmitted. The absorptance is then approximated by $a(\hbar\omega < \hbar\omega_G) = 0$ and $a(\hbar\omega \geq \hbar\omega_G) = 1$.

The energy current density absorbed by such a semiconductor plate is

$$j_{E,\text{abs}} = \frac{d\Omega}{4\pi^3 \hbar^3 c^2} \int_{\hbar\omega_G}^{\infty} \frac{(\hbar\omega)^3}{\exp(\hbar\omega/kT) - 1} d\hbar\omega \quad (2.38)$$

This integral can be evaluated. With $\hbar\omega/kT = x$ and $\hbar\omega_G/kT = x_G$ and rewriting

$$\frac{1}{\exp(x) - 1} = \frac{\exp(-x)}{1 - \exp(-x)} = \exp(-x) \sum_{i=0}^{\infty} \exp(-ix) = \sum_{i=1}^{\infty} \exp(-ix)$$

which is valid for $x > 0$, Equation 2.38 becomes

$$j_{E,\text{abs}} = \frac{d\Omega (kT)^4}{4\pi^3 \hbar^3 c^2} \sum_{i=1}^{\infty} \int_{x_G}^{\infty} x^3 \exp(-ix) dx$$

Partial integration then gives

$$j_{E,\text{abs}} = \frac{d\Omega (kT)^4}{4\pi^3 \hbar^3 c^2} \sum_{i=1}^{\infty} \exp(-ix_G) \left(\frac{x_G^3}{i} + \frac{3x_G^2}{i^2} + \frac{6x_G}{i^3} + \frac{6}{i^4} \right) \quad (2.39)$$

For $\hbar\omega_G \gg kT$, that is $x_G \gg 1$, only a few terms of this series are required to obtain sufficient accuracy. Accordingly, for the absorbed photon current density we obtain

$$j_{\gamma,\text{abs}} = \frac{d\Omega (kT)^3}{4\pi^3 \hbar^3 c^2} \sum_{i=1}^{\infty} \exp(-ix_G) \left(\frac{x_G^2}{i} + \frac{2x_G}{i^2} + \frac{2}{i^3} \right) \quad (2.40)$$

2.3

The Solar Spectrum

Figure 2.9 shows the solar energy current density per photon energy interval as a function of the photon energy that is incident on the Earth before entering the atmosphere. It agrees well with a spectrum calculated from Equation 2.34 if we assume that the Sun is a black body and has a temperature T_S of 5800 K. At this temperature of the surface of the Sun, $kT_S = 0.5$ eV. The solar spectrum has a maximum at $\hbar\omega_{\text{max}} = 2.82 kT_S = 1.41$ eV, that is, in the infrared part of the spectrum, as seen in Figure 2.9. This is just outside the visible range extending from 1.5 to 3 eV. The average energy of the solar photons is $2.7kT_S = 1.35$ eV.

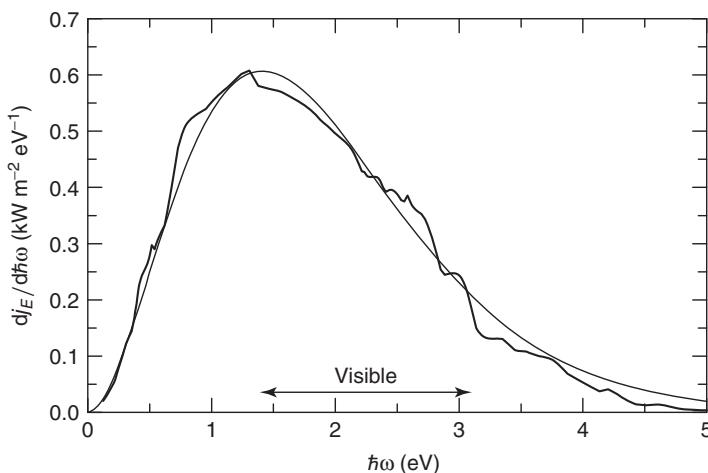


Figure 2.9 Energy current density per photon energy from the Sun as a function of the photon energy just outside the Earth's atmosphere (heavy line) compared with a black-body at a temperature of 5800 K (thin line).

Most often a different quantity is presented as the solar spectrum. It is the energy current density per wavelength interval as a function of the wavelength as shown in Figure 2.10.

This solar spectrum has a maximum at $\lambda_{\max} = 0.5 \mu\text{m}$, corresponding to a photon energy of $\hbar\omega = 2.48 \text{ eV}$ and distinctly different from the value $\hbar\omega_{\max}$ found

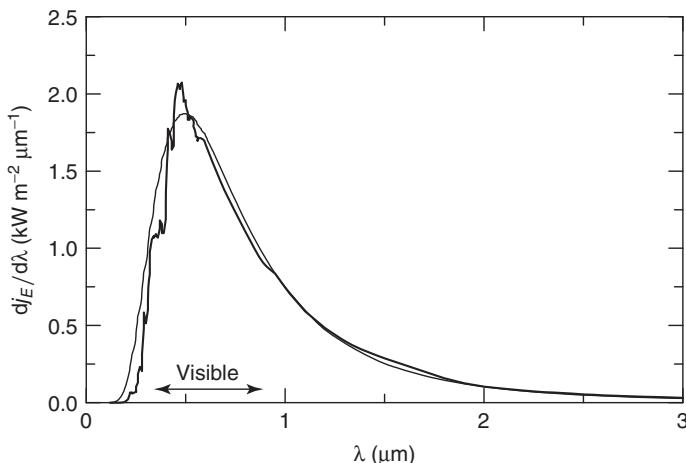


Figure 2.10 Energy current density per wavelength from the Sun as a function of the wavelength outside the Earth's atmosphere (heavy line) compared with a black body at a temperature of 5800 K (thin line).

from the photon energy spectrum. Although both spectra are called solar spectra, they are completely different quantities and have different units. We will, therefore, use different names and call the energy current density per energy interval $dj_E/d\hbar\omega$ as a function of the photon energy, shown in Figure 2.9, the energy spectrum, and the energy current density per wavelength interval as a function of the wavelength $dj_E/d\lambda$, shown in Figure 2.10, the wavelength spectrum.

The energy current density outside the atmosphere, that is, the integral over each of the curves in Figure 2.9 or 2.10, has a value of

$$j_{E,AM0} = 1353 \text{ W m}^{-2} \quad (2.41)$$

The form of the spectrum in Figure 2.9 by itself is not sufficient to assert that the source, namely the Sun, actually has a temperature of $T_S = 5800 \text{ K}$. Although the spectrum has its maximum at the photon energy expected for this temperature, it is also possible to produce an energy current density per photon energy with the form and magnitude as shown in Figure 2.9 with much colder light bulbs and appropriate filters. This is the basis for solar simulators used to test solar cells. The temperature of the radiation source, provided it is black, only follows from the spectrum after considering the solid angle $d\Omega$ subtended by the radiation source, which for the Sun is $\Omega_S = 6.8 \times 10^{-5}$.

Only the energy current density per solid angle (and per photon energy or per wavelength) is a measure of the temperature of a thermal emitter. For a solar simulator, radiating the same energy current density per photon energy onto an absorber, the radiation emerges from a much larger solid angle than for the Sun.

2.3.1

Air Mass

Solar radiation is partially absorbed during its passage through the atmosphere. The absorption is almost entirely caused by gases of low concentration in the infrared region of the solar spectrum, by water vapor (H_2O), carbon dioxide (CO_2), laughing gas (N_2O), methane (CH_4), fluorinated hydrocarbons, as well as by dust and, in the ultraviolet region of the spectrum, by ozone and oxygen. Absorption of course increases with the length of the path through the atmosphere and therefore with the mass of air through which the radiation passes. For a thickness l_0 of the atmosphere, the path length l through the atmosphere for radiation from the Sun incident at an angle α relative to the normal to the Earth's surface is given by

$$l = \frac{l_0}{\cos \alpha}$$

The ratio l/l_0 is called the *air-mass* coefficient. It characterizes the real solar spectrum resulting from the absorption of a layer of air of thickness l . The spectrum outside the atmosphere is designated by $AM0$ and that on the surface of the Earth for normal incidence by $AM1$. A typical spectrum for moderate climates is $AM1.5$,

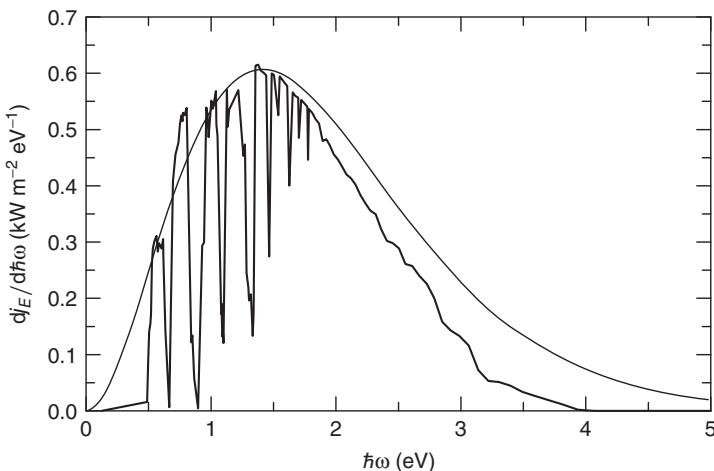


Figure 2.11 The AM1.5 spectrum (heavy line) compared with a black body at a temperature of 5800 K (thin line).

which corresponds to an angle of incidence of solar radiation of 48° relative to the surface normal.

The energy spectrum *AM1.5* shown in Figure 2.11 is regarded as the standard spectrum for measuring the efficiency of solar cells used terrestrially, i.e. on the surface of the Earth. The integral over this spectrum, the energy current density onto a surface normal to the Sun for a cloudless sky, is defined to be

$$j_{E,AM1.5} = 1.0 \text{ kW m}^{-2}$$

This maximum energy current density is only negligibly smaller than for the *AM1.0* spectrum. The maximum energy current density varies only slightly from the tropics to the moderate zones. The differences in the amount of energy incident onto a horizontal surface in one year are, however, much greater. In Germany this is around 1000 kWh m⁻²a⁻¹. We therefore speak of 1000 Sun-hours (with 1 kW m⁻²) per year.

Averaged over the year, the mean energy current density in Germany is 115 W m⁻², and roughly a factor of 10 less than the maximum energy current density of the *AM1.5* spectrum. The greatest annual amount of incident solar energy is found in Saudi Arabia, with a value of around 2500 kWh m⁻²a⁻¹. This corresponds to an average energy current density of 285 W m⁻². The average value over the entire Earth is 230 W m⁻².

For the *AM0* spectrum, outside the atmosphere, the energy current density averaged over the entire Earth is easy to calculate. We only have to divide the energy current $I_E = 1353 \times \pi R_{\text{Earth}}^2 \text{ W m}^{-2}$ incident upon the Earth by the total surface area of the Earth ($4\pi R_{\text{Earth}}^2$) to obtain the value $\langle j_e \rangle = 1/4 \times 1353 \text{ W m}^{-2} = 338 \text{ W m}^{-2}$ for the average energy current density, if there were no atmosphere.

2.4

Concentration of the Solar Radiation

Viewed from the Earth, the Sun has an angular diameter α_S of $32'$, corresponding to a solid angle of

$$\Omega_S = 2\pi \int_0^{\alpha_S/2} \sin \theta d\theta = 2\pi \left(1 - \cos \frac{\alpha_S}{2}\right)$$

$$= 6.8 \times 10^{-5} \quad (2.42)$$

It is due to this small value of the solid angle Ω_S that the energy current density on the Earth has a value of only 1 kW m^{-2} . When a greater energy current density is needed in order to obtain higher temperatures or increase the output of solar cells, it is necessary to focus the nearly parallel incidence of radiation using lenses or mirrors. This reduces the irradiated area and increases the solid angle from which the radiation impinges on the receiver surface. What happens then with the energy current density per solid angle? For an answer, we will make use of the deliberations of Helmholtz and Clausius from the year 1864.

Figure 2.12 shows two area elements dA_1 and dA_2 , projected onto each other by an imaging system. Because of the reversibility of the light path, dA_2 is the image of dA_1 and dA_1 is the image of dA_2 . We assume that both area elements are black bodies which, without any further illumination, emit thermal radiation toward each other through the transparent imaging system. Although the thermal radiation will be redirected by the imaging system, once again, the thermal equilibrium between dA_1 and dA_2 must not be disturbed in order to avoid violating the second principle of thermodynamics. Thus, the area dA_2 must receive exactly the same amount of radiation from the lens or the imaging system as it emits toward it. This equality of absorbed and emitted energy currents must exist not only for any photon energy, as in Section 2.2, but also for any direction, that is, any solid angle element, as defined by an area element on the surface of the lens. By redirecting the radiation from dA_1 , the imaging system (the lens) must therefore behave like a black body, emitting radiation toward dA_2 with the temperature of dA_1 . But in contrast to a real black body, it selectively radiates only in the direction of dA_2 and in no other direction. We have already seen that for the propagation of radiation in free space the energy current density per solid angle remains unchanged and is

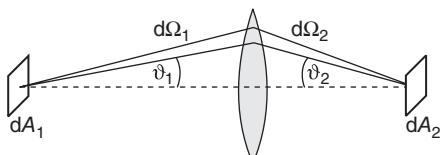


Figure 2.12 An imaging system projecting two area elements onto each other by redirecting their thermally emitted radiation must not disturb the thermal equilibrium between them.

therefore exactly the same at the location of the lens in the direction toward dA_2 as at the location of dA_2 from the direction of the lens. Owing to thermal equilibrium, this is of course exactly the same as the energy current density per solid angle at the location of dA_1 . *During passage through a nonabsorbing and nonemitting ideal imaging system, the energy current density per solid angle remains unchanged.*

The reversibility of the light path also implies thermodynamic reversibility. This in turn means that no entropy is created during passage through an imaging system. And in fact the conservation of the energy current density per solid angle in regions with the index of refraction n , with simultaneous conservation of the energy per photon, is identical with the conservation of entropy. More precisely, for the propagation of radiation not attenuated by absorption or scattering, even when the index of refraction n changes, the occupation probability of the photon states f_γ in Equation 2.3 remains the same.

2.4.1

The Abbé Sine Condition

The energy current emitted in Figure 2.12 from dA_1 in the direction (ϑ_1, φ_1) into the solid angle element $d\Omega_1 = \sin \vartheta_1 d\vartheta_1 d\varphi_1$ is

$$dI_{E,1} = j_{E,\Omega,1} dA_1 \cos \vartheta_1 \sin \vartheta_1 d\vartheta_1 d\varphi_1 \quad (2.43)$$

This energy current is incident upon dA_2 and as a result of thermal equilibrium is identical with the energy current emitted from dA_2 in the direction (ϑ_2, φ_2) into the solid angle element $d\Omega_2 = \sin \vartheta_2 d\vartheta_2 d\varphi_2$

$$dI_{E,2} = j_{E,\Omega,2} dA_2 \cos \vartheta_2 \sin \vartheta_2 d\vartheta_2 d\varphi_2 = dI_{E,1} \quad (2.44)$$

To generalize the condition in Equation 2.44 by allowing for different media with different indices of refraction in front of and behind the lens in regions 1 and 2, we have to take into account Equation 2.34 for the energy current densities per solid angle,

$$\frac{j_{E,\Omega,1}}{j_{E,\Omega,2}} = \frac{c_2^2}{c_1^2} = \frac{n_1^2}{n_2^2} \quad (2.45)$$

The entire energy current absorbed by the area element dA_2 or emitted by it toward the lens is obtained by integrating over the entire lens or over the angle ϑ_2 from 0 to ν , where ν is the angle relative to the optical axis at which the edge of the lens is viewed from dA_2 , and over φ_2 from 0 to 2π

$$I_{E,2} = j_{E,\Omega,2} dA_2 \pi \sin^2 \nu \quad (2.46)$$

In exactly the same way, we find the entire energy current emitted by the area element dA_1 toward the lens or obtained from the lens to be

$$I_{E,1} = j_{E,\Omega,1} dA_1 \pi \sin^2 u \quad (2.47)$$

where u is the angle at which the edge of the lens is viewed from dA_1 .

Since the two energy currents must be equal due to thermal equilibrium, and using Equation 2.45, we obtain the Abbé sine condition for optical imagery

$$n_1^2 dA_1 \sin^2 u = n_2^2 dA_2 \sin^2 v \quad (2.48)$$

The basis of this condition is the conservation of the energy current density per solid angle in regions with the same index of refraction n , even when it is passing through imaging systems. This law of conservation is in no way trivial. As it turns out, it contradicts the simple imaging laws of geometrical optics.

2.4.2

Geometrical Optics

In accordance with geometrical optics, for the imaging of the Sun by a lens with radius r_L producing an image with area A_B as in Figure 2.13, we expect that the energy current incident upon the lens is also incident upon the image A_B of the Sun. The energy current incident upon the lens is

$$I_E = j_{E, \text{Sun}} \pi r_L^2$$

where $j_{E, \text{Sun}}$ is the energy current density of the solar radiation at the location of the lens. The energy current density in the image A_B is then

$$j_{E, \text{image}} = \frac{I_E}{A_B} = \frac{j_{E, \text{Sun}} \pi r_L^2}{A_B}$$

The concentration factor C is a measure of the enhancement of the energy current density by the lens and, according to geometrical optics, is therefore

$$C = \frac{j_{E, \text{image}}}{j_{E, \text{Sun}}} = \frac{\pi r_L^2}{A_B}$$

The size of the image A_B , formed in the focal plane because of the large distance from the Sun, depends only on the focal length f of the lens and not on its size. Since the aperture angle Ω_S of a bundle of rays incident upon the middle of the lens remains unchanged during passage through the lens, we have on the image side $\Omega_S = A_B/f^2$.

Expressing the radius of the lens as

$$r_L = f \tan v$$

in terms of the maximum angle v of incidence onto the image, we finally arrive at

$$C = \frac{\pi}{\Omega_S} \tan^2 v \quad (2.49)$$

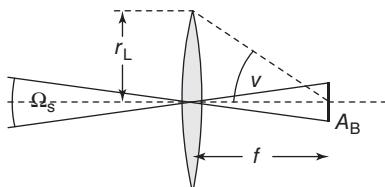


Figure 2.13 Image of the Sun with an area A_B in the focal plane of a lens with radius r_L and focal length f .

In agreement with the assumption that the energy current incident upon the lens is also incident upon the image of the Sun, the concentration factor becomes arbitrarily large when the lens is made arbitrarily large. As an example, let us imagine a lens with $\nu > 45^\circ$, that is $r_L > f$. For this lens $C > \pi/\Omega_S$.

The ratio $2r_L/f$ represents the f -number of lenses, expressed as the ratio $1:(f/2r_L)$. $1:1.4$ or $1:2.0$ or $1:2.8$ are typical values for fast camera lenses. In the example above, the f -number would be $> 1:0.5$. Certainly, no one has ever heard of such a lens, because it would in fact violate the second law of thermodynamics. For this lens, according to Equation 2.28 and Equation 2.49 the energy current density in the image of the Sun would be

$$j_{E,\text{image}} = C j_{(E,\text{Sun})} > \frac{\pi}{\Omega_S} \frac{\Omega_S}{\pi} \sigma T_S^4$$

so that

$$j_{E,\text{image}} > \sigma T_S^4$$

and therefore greater than on the surface of the Sun where it is σT_S^4 .

Under steady state conditions, a black body in the image of the Sun, with a perfectly reflecting rear surface and therefore unable to emit radiation from this side, would have to emit from the front side the absorbed energy current density $j_{E,\text{image}} = \sigma T_{\text{image}}^4 > \sigma T_S^4$. Its temperature T_{image} would then have to be greater than that of the Sun T_S . The emitted energy current $I_{E,\text{image}}$ also implies the emission of an entropy current. According to Boltzmann, the entropy current emitted by a black body of temperature T_{image} is related to the emitted energy current by

$$I_{S,\text{image}} = \frac{4}{3} \frac{I_{E,\text{image}}}{T_{\text{image}}}$$

This would be less than the entropy current absorbed together with the same energy current $I_{E,\text{image}}$

$$I_{S,\text{abs}} = \frac{4}{3} \frac{I_{E,\text{image}}}{T_{\text{Sun}}}$$

This would be possible only if entropy were continuously destroyed in the image of the Sun. However, the second law of thermodynamics forbids this.

2.4.3

Concentration of Radiation Using the Sine Condition

The correct determination of the enhancement of the energy current density by the use of concentrating optics takes into account the conservation of the energy current density per solid angle $j_{E,\Omega,\text{Sun}}$. Accordingly, viewed from the image of the Sun the entire lens appears to be as bright as the Sun (do not try this out!) and fills a cone with the angle ν . According to Equation 2.46, the energy current density in the image of the Sun is

$$j_{E,\text{image}} = j_{E,\Omega,\text{Sun}} \pi \sin^2 \nu$$

With $j_{E,\text{Sun}} = \Omega_S j_{E,\Omega,\text{Sun}}$ the concentration factor becomes

$$C = \frac{j_{E,\text{image}}}{j_{E,\text{Sun}}} = \frac{\pi}{\Omega_S} \sin^2 \nu \quad (2.50)$$

The maximum concentration is found for $\nu = 90^\circ$

$$C_{\max} = \frac{\pi}{\Omega_S} = 46200 \quad (2.51)$$

The contradiction between Equation 2.49, based on geometrical optics of an undistorted image, and Equation 2.50, based on thermodynamics and the impossibility of destroying entropy, can be resolved for imaging systems with curved principal "planes". The maximum possible *f*-number is 1 : 0.5 and requires a principal image-side plane in the form of a hemisphere, curved around the focal point. This is, however, not totally achievable.

The maximum concentration can nevertheless be obtained by other means. It ensures that an absorbing body reaches the temperature of the Sun, if it loses energy only by emission toward the Sun. In this situation it is in radiative equilibrium with the Sun. Let us assume that the absorber is already at the temperature of the Sun. It would then remain at this temperature if it were within a cavity with walls at the temperature of the Sun, that is, as though it could only see the Sun. It would also remain at the temperature of the Sun, if it were located in a cavity with perfectly reflecting walls, that is, as though it could only see itself. Figure 2.14 shows such an arrangement in which a body is in radiative equilibrium with the Sun. The lens placed in the wall of a reflecting cavity produces an image of the Sun that is at least as large as the cross-section of the absorber. Since the absorber then sees only the Sun or itself, it would reach the temperature of the Sun if there were perfect mirrors. According to this consideration, the structural principle of concentrators can now be defined very simply. A good concentrator must be designed so as to direct toward the Sun the greatest possible amount of thermal radiation emitted by the absorber.

The ideal arrangement shown in Figure 2.14 permits the determination of the maximum efficiency with which solar energy can be converted into electrical energy.

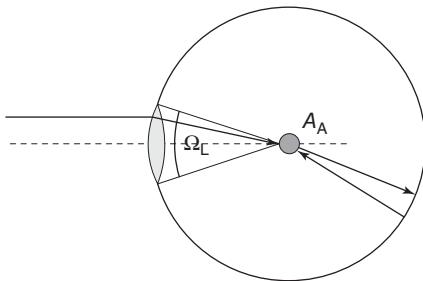


Figure 2.14 A lens projects an image of the Sun onto an absorber in a perfectly reflecting cavity.

2.5

Maximum Efficiency of Solar Energy Conversion

To use the incident solar energy in the arrangement of Figure 2.14, heat must be extracted from the absorber and conducted to a heat engine. Using an ideal heat engine, i.e. a Carnot engine, we obtain the greatest possible amount of electrical energy. Owing to the extraction of energy, the temperature of the absorber T_A will be lower than that of the Sun T_S . It must be lower so that it does not emit as much energy as it absorbs.

The useful net absorbed energy current $I_{E,\text{util}}$ that can be extracted from the absorber of cross-sectional area A_A is the difference between the absorbed energy current $I_{E,\text{abs}}$ and the energy current $I_{E,\text{emit}}$ emitted through the lens toward the Sun. Ideally, both fill the same solid angle Ω_L , at which the lens is viewed from the absorber. For a more general discussion, we will allow for different solid angles, Ω_{abs} for the incident radiation and Ω_{emit} for the emitted radiation:

$$I_{E,\text{util}} = I_{E,\text{abs}} - I_{E,\text{emit}} = \sigma \left(\frac{\Omega_{\text{abs}}}{\pi} T_S^4 - \frac{\Omega_{\text{emit}}}{\pi} T_A^4 \right) A_A$$

We can now define the efficiency for obtaining $I_{E,\text{util}}$ as

$$\eta_{\text{abs}} = \frac{I_{E,\text{util}}}{I_{E,\text{abs}}} = 1 - \frac{I_{E,\text{emit}}}{I_{E,\text{abs}}} = 1 - \frac{\Omega_{\text{emit}}}{\Omega_{\text{abs}}} \frac{T_A^4}{T_S^4}$$

This relation states that for a high efficiency of net absorption η_{abs} the temperature T_A of the absorber should be as low as possible to avoid the emission of radiation.

Along with the heat energy current $I_{E,\text{util}}$, the entropy current $I_S = I_{E,\text{util}}/T_A$ is transferred from the absorber to the heat engine. According to the second law of thermodynamics, entropy cannot be destroyed, but it can in fact be created. In an ideal Carnot engine, entropy is conserved. Since the absorbed entropy current I_S cannot be stored in the engine, it must be transported to another container, a heat reservoir. If this accepts the entropy at the temperature T_0 , the entropy current will be accompanied by a heat energy current $T_0 I_S$ from the engine to this heat reservoir. Only the difference between the absorbed energy current and that given to the reservoir is supplied by the Carnot engine as an entropy-free energy current, e.g. in the form of electrical energy $I_{E,\text{el}}$,

$$I_{E,\text{el}} = I_{E,\text{util}} - T_0 I_S = T_A I_S - T_0 I_S$$

The efficiency for this process is

$$\eta_C = \frac{I_{E,\text{el}}}{I_{E,\text{util}}} = \frac{T_A I_S - T_0 I_S}{T_A I_S} = 1 - \frac{T_0}{T_A} \quad (2.52)$$

The efficiency of such an ideal heat engine is known after its discoverer as the Carnot efficiency. For a high Carnot efficiency, the temperature T_A of the absorber should be as high as possible.

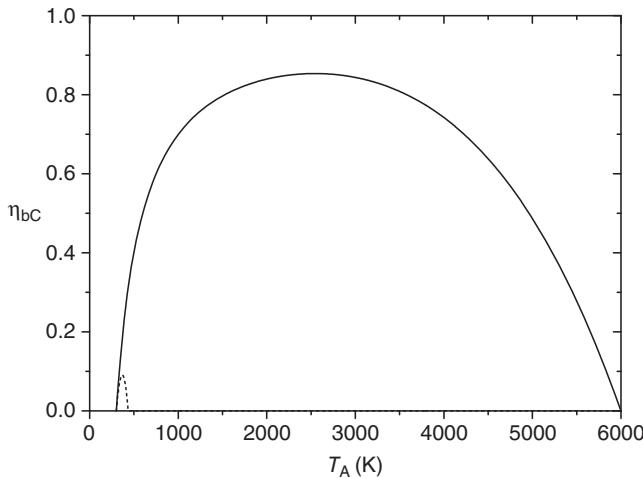


Figure 2.15 Efficiency η_{bC} for the conversion of solar heat energy by a black absorber and a Carnot engine as a function of the absorber temperature T_A , for maximum

concentration ($\Omega_{\text{emit}} = \Omega_{\text{abs}}$, heavy line) and for nonconcentrated radiation ($\Omega_{\text{emit}} = \pi$, $\Omega_{\text{abs}} = 6.8 \times 10^{-5}$, broken line).

The overall efficiency η_{bC} for the conversion of solar energy to entropy-free energy, e.g. electrical energy, by a black-body absorber combined with a Carnot engine is

$$\begin{aligned}\eta_{bC} &= \frac{I_{E,\text{el}}}{I_{E,\text{abs}}} = \frac{I_{E,\text{el}}}{I_{E,\text{util}}} \frac{I_{E,\text{util}}}{I_{E,\text{abs}}} = \eta_{\text{abs}} \eta_C \\ \eta_{bC} &= \left(1 - \frac{\Omega_{\text{emit}}}{\Omega_{\text{abs}}} \frac{T_A^4}{T_S^4}\right) \left(1 - \frac{T_0}{T_A}\right)\end{aligned}\quad (2.53)$$

Figure 2.15 displays η_{bC} as a function of T_A for $T_S = 5800$ K and $T_0 = 300$ K for maximum concentration, i.e. equal solid angles of incident and emitted radiation as seen in Figure 2.14. The overall efficiency has a maximum value of 0.85 for an absorber temperature $T_A = 2478$ K. This high efficiency value demonstrates that solar energy is very high-quality energy because of the high temperature of the Sun.

Since the thermal radiation from the Sun at a temperature $T_S = 5800$ K is absorbed by the absorber at a lower temperature T_A , the absorption process creates entropy. We will accept this, because we want to obtain the greatest possible energy current, and the heat radiated back to the Sun is lost to us.

An interesting aspect arises if we had to pay for solar energy, but could also get a refund for energy returned to the Sun (which would be justified, since it would actually prolong the Sun's lifetime). Under these conditions maximal

concentration must be chosen along with a temperature of the absorber that is only slightly (by dT) smaller than T_S . The net absorbed energy current then is

$$I_{E,\text{abs}} = \left[\frac{\Omega_{\text{abs}}}{\pi} \sigma T_S^4 - \frac{\Omega_{\text{emit}}}{\pi} \sigma (T_S - dT)^4 \right] A_A \approx 4 \frac{\Omega_{\text{abs}}}{\pi} \sigma T_S^3 dT A_A$$

Since this is the energy current for which we would have to pay, the absorption efficiency is $\eta_{\text{abs}} = 1$. The net absorbed entropy current is

$$I_{S,\text{abs}} = \frac{4}{3} \frac{\Omega_{\text{abs}}}{\pi} \sigma \left[T_S^3 - (T_S - dT)^3 \right] A_A \approx 4 \frac{\Omega_{\text{abs}}}{\pi} \sigma T_S^2 dT A_A$$

This time, the absorption process does not generate entropy, since the ratio of absorbed energy and entropy currents is equal to the Sun's temperature T_S . It is not surprising that the overall efficiency for this reversible process, including the Carnot engine, is $\eta = 1 - T_0/T_S$. The output of the engine is, however, rather small, since almost all of the incident solar radiation is emitted back to the Sun and an efficiency that relates the energy output not to the net absorbed but to the incident energy current is almost zero.

A slight improvement in efficiency over the combination of a black absorber and a Carnot engine is found for a combination of a monochromatic absorber and a Carnot engine. However, very (infinitely) many absorbers, each for a different photon energy interval, are needed to cover the whole solar energy spectrum. Each absorber would have its own Carnot engine and operate at its own optimal temperature.

We start with a single monochromatic absorber, absorbing all incident solar photons that are in a narrow energy interval $d\hbar\omega$ around a photon energy of, e.g. $\hbar\omega = 1.5 \text{ eV}$ and emitting photons in the same energy interval according to its temperature. Again, we assume maximal concentration of the incident solar radiation, i.e. equal solid angles for incident and emitted radiation. Using Planck's equation for the photon current density per photon energy in Equation 2.35, the net absorbed photon current, the difference between absorbed and emitted photon currents is

$$\begin{aligned} dj_{\gamma,\text{net}}(\hbar\omega) &= dj_{\gamma,\text{abs}} - dj_{\gamma,\text{emit}} \\ &= m \left[\frac{(\hbar\omega)^2}{\exp(\hbar\omega/kT_S) - 1} - \frac{(\hbar\omega)^2}{\exp(\hbar\omega/kT_A) - 1} \right] d\hbar\omega \end{aligned} \quad (2.54)$$

where the factor m contains all the constants including the solid angle. Each net absorbed photon carries its energy $\hbar\omega$ to the Carnot engine, which produces from it entropy-free energy of an amount of $\hbar\omega(1 - T_0/T_A)$ per photon, resulting in an energy current delivered by the Carnot engine of

$$dj_{E,\text{Carnot}}(\hbar\omega) = dj_{\gamma,\text{net}}(\hbar\omega) \hbar\omega \left(1 - \frac{T_0}{T_A} \right)$$

Both the net absorbed photon current and the free energy per photon depend on the absorber temperature T_A . By varying T_A we find pairs of absorbed photon

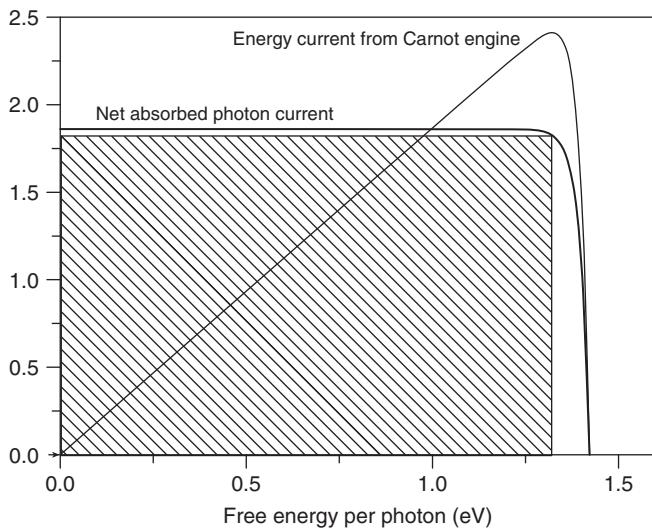


Figure 2.16 Net absorbed photon current of 1.5 eV photons as a function of the free energy per photon produced by a Carnot engine. The hatched rectangle, the largest

rectangle for any point on the photon current curve, belongs to the point of maximum power, which also follows from the maximum of the energy current curve.

current and free energy per photon, which are plotted in Figure 2.16 as the thick line. The current of free energy delivered by the Carnot engine is the area of a rectangle formed by the x - and y -values of each point on this curve. The hatched rectangle has the largest possible area and belongs to the point of maximum power on the curve. This point of maximum power follows from the maximum of the energy current $dj_{E, \text{Carnot}}(\hbar\omega)$ shown by the thin line.

The efficiency of energy conversion is found by dividing the energy current $dj_{E, \text{Carnot}}(\hbar\omega)$ by the incident monochromatic energy current $\hbar\omega dj_{\gamma, \text{abs}}(\hbar\omega)$. This efficiency is shown as a function of the photon energy in Figure 2.17 for maximum concentration. For the conversion of the whole solar spectrum, many monochromatic conversion engines have to operate in parallel. The entropy-free energy obtained from them is the integral over the energy current per photon energy interval weighted with the appropriate monochromatic efficiency. For the overall efficiency, this output energy current must be divided by the incident energy current, resulting in a value of 86%, which is slightly higher than the 85% obtained with a black absorber and a single Carnot engine.

So far, efficient energy conversion has had to avoid the emission of photons toward the Sun as far as possible. The temperature of the absorber, therefore, had to be substantially lower than the temperature of the Sun. This lower temperature of the absorber, however, gives rise to unwanted entropy generation, which could only be avoided when the absorber has the same temperature as the Sun, where the net absorbed energy current is zero. In this dilemma, we might ask ourselves how high the efficiency would be if it were possible for a black body at a temperature

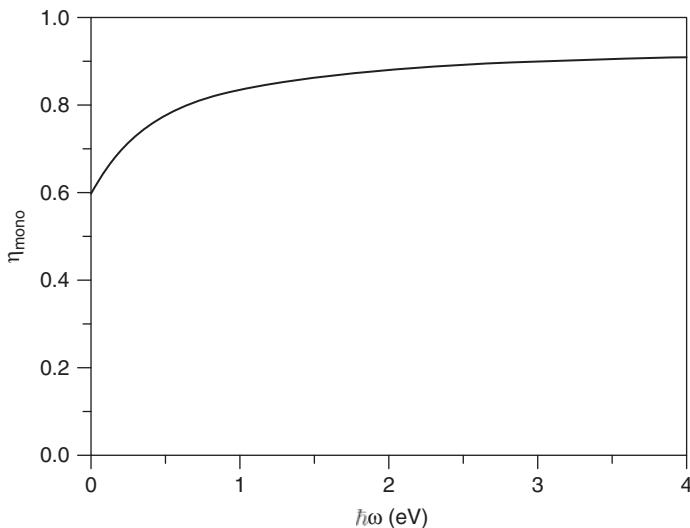


Figure 2.17 Efficiency of monochromatically absorbing Carnot engines as a function of the photon energy for fully concentrated black-body radiation of 5800 K.

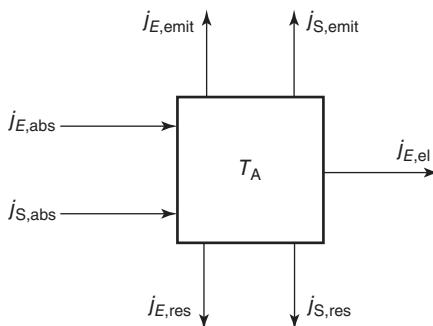


Figure 2.18 Balance of absorbed and emitted energy and entropy currents for a reversible heat engine working with radiation.

$T_A < T_S$ to absorb the radiation from the Sun without creating entropy. This efficiency is called the *Landsberg efficiency*. It follows from a balance of absorbed and emitted energy and entropy currents under the condition of reversibility as indicated in Figure 2.18.

The entropy absorbed with solar radiation is given off in two ways. One part is emitted back to the Sun, together with the thermal radiation at the temperature T_A , and the remaining part goes to a heat reservoir at the ambient temperature T_0 , implying a loss of energy.

Assuming that the absorbing engine sees only the Sun, the absorbed energy current density is

$$j_{E,\text{abs}} = \sigma T_S^4$$

For black-body radiation, the density of the entropy current absorbed is

$$j_{S,\text{abs}} = \frac{4}{3} \frac{j_{E,\text{abs}}}{T_S} = \frac{4}{3} \sigma T_S^3$$

The energy current density emitted by the absorber with temperature T_A (to the Sun) is given by

$$j_{E,\text{emit}} = \sigma T_A^4$$

and the entropy current density emitted is

$$j_{S,\text{emit}} = \frac{4}{3} \sigma T_A^3$$

In order to get rid of all the entropy absorbed, the entropy current density $j_{S,\text{res}}$ is transferred to the reservoir. With the entropy the energy current density

$$j_{E,\text{res}} = T_0 j_{S,\text{res}}$$

is therefore transported to the reservoir and lost for conversion. The condition of reversibility, that is, the conservation of entropy, takes the form

$$j_{S,\text{abs}} = j_{S,\text{emit}} + j_{S,\text{res}}$$

This leads to the relationships

$$j_{S,\text{res}} = \frac{4}{3} \sigma (T_S^3 - T_A^3)$$

and

$$j_{E,\text{res}} = \frac{4}{3} \sigma T_0 (T_S^3 - T_A^3)$$

The entropy-free, utilizable energy current density $j_{E,\text{el}}$ is

$$j_{E,\text{el}} = j_{E,\text{abs}} - j_{E,\text{emit}} - j_{E,\text{res}}$$

The Landsberg efficiency is therefore

$$\begin{aligned} \eta_L &= \frac{j_{E,\text{el}}}{j_{E,\text{abs}}} = 1 - \frac{\sigma T_A^4 + 4/3 \sigma T_0 (T_S^3 - T_A^3)}{\sigma T_S^4} \\ &= 1 - \frac{T_A^4}{T_S^4} - \frac{4}{3} \frac{T_0}{T_S} \left(1 - \frac{T_A^3}{T_S^3} \right) \end{aligned} \quad (2.55)$$

Figure 2.19 compares the Landsberg efficiency η_L with the efficiency η_{bC} of a single black absorber combined with a Carnot engine. The lower the absorber temperature T_A , the greater is η_L . However, it is highly questionable whether the reversible absorption of radiation of a temperature T_S by a black body of a temperature $T_A < T_S$ is permitted by nature. In this respect, the Landsberg efficiency sets an upper limit to any kind of solar energy conversion.

Later, when looking at semiconductors, we will discover absorbers capable of absorbing the incident radiation without creating entropy, even though their temperature is much lower than that of the radiation. Nevertheless, in these states

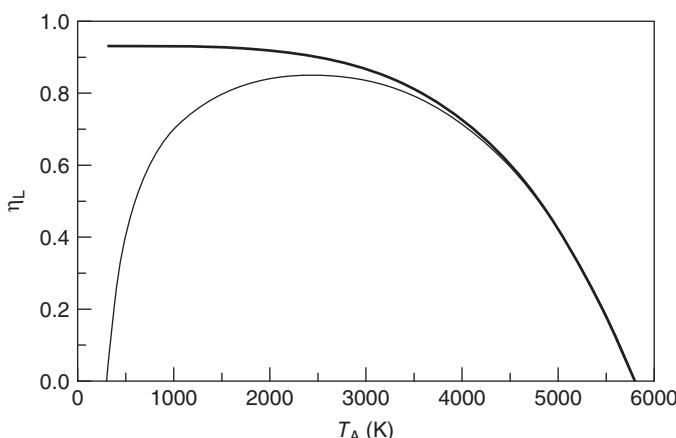


Figure 2.19 Landsberg efficiency η_L (heavy line) as a function of the absorber temperature T_A compared with the efficiency η_{bc} (thin line) of a solar heat engine with a single black absorber.

they emit photons with a chemical potential greater than zero. This is different from thermal radiation, for which photons have a chemical potential equal to zero. In spite of this, we cannot take advantage of this property of semiconductors because, even though their temperature is lower, in order to avoid the creation of entropy they emit photons with so great a chemical potential that the photon current density emitted toward the Sun has exactly the same magnitude as that obtained from the Sun and no net energy can be transferred to a Carnot engine without the production of entropy.

2.6

Problems

- 2.1** What is a black body?
- 2.2** Why does a light-emitting surface of a black body appear equally bright from all viewing angles?
- 2.3** How many photon states per volume exist in the visible energy range ($\hbar\omega$ between 1.5 and 3 eV)
 - (a) in vacuum?
 - (b) in silicon?
 - (c) At which temperature is the number of photons larger than 10^{-4} of all photon states in this energy range?
- 2.4** Determine the maximum photon energy $\hbar\omega$ for which the Boltzmann distribution differs from the Bose–Einstein distribution by less than 10^{-2}
 - (a) for $T = 300$ K
 - (b) for $T = 5800$ K.

- 2.5** Calculate the reflectance of a flat Si surface without antireflection coating, using a refractive index of $n_{\text{Si}} = 3.5$,
- (a) for normal incidence
 - (b) for incidence at an angle of $\alpha = 50^\circ$ to the surface normal. Assume the light to be unpolarized.
 - (c) as in case (b), but with light polarized perpendicular to the plane of incidence.
- 2.6** For normal incidence $AM1.5$ global corresponds to 1000 W m^{-2} .
- (a) A solar module is placed on a roof tilted at an angle $\beta = 25^\circ$ (with respect to a horizontal plane) toward the south. At which latitude in the northern hemisphere would the module receive 750 W m^{-2} at noon on a March 21?
 - (b) Imagine that the module from (a) with an (intensity-independent) efficiency of 15% is installed on a tracking system to gain more power. How many axes are necessary to optimally guide the module on that special day? What is the energy output from the module with and without the tracking system for the March 21 (assume $AM1.5$ global over the entire day)?
- 2.7** How large is the mean energy of photons from a black body of temperature T absorbed by a semiconductor with an absorptance of $a(\hbar\omega < \varepsilon_G) = 0$ and $a(\hbar\omega \geq \varepsilon_G) = 1$? Assume $\varepsilon_G \gg kT$.

3 Semiconductors

The thermal energy that the absorber in Figure 2.14 supplies to the Carnot engine is contained almost entirely in the random oscillations of the atoms in a solid about their rest positions. The energy of the oscillations is quantized similar to the electromagnetic modes of a cavity. For the electromagnetic field in the cavity, the vibrational quanta are the *photons*, and for the atomic oscillations of a solid, the vibrational quanta are called *phonons*. Phonon energies ε_{Γ} range from 0 to 0.05 eV. Only in exceptional cases can phonons be directly generated by the absorption of photons. The absorption of photons takes place through the excitation of electrons into states of higher energy. In order that photons of any arbitrary energy can be absorbed, i.e. for a body to be black, a continuous, uninterrupted range of excitation energies must be available to the electrons. For metals, this is in fact the case. Metals would represent the closest approximation to an ideal black body, if they did not reflect most of the incident light. This type of reflection, however, can be eliminated by roughening the surface, which renders a metal completely black. You will experience this yourself when you scrub an aluminum pot with some scouring powder and a cloth. The tiny aluminum particles scrubbed off the pot make the cloth totally black. Owing to the continuous energy range for electrons in a metal, an electron that is excited by absorbing a photon loses this extra energy easily, step by step, in small portions, by generating phonons. Although this may require many steps, the process typically takes place in times of the order of 10^{-12} s. Since in metals, the excitation energy remains for only such a short time in the electron gas, their direct utilization has a poor efficiency, although the emission of the excited electrons out of the metal is used in photomultipliers for the fast detection of single photons. Figure 3.1 is a schematic diagram for the absorption of a photon by an electron in a metal and the subsequent loss of the energy to phonons.

In semiconductors, this mechanism is different. Semiconductors are materials in which the range of excitation energies is interrupted by an energy gap of width ε_G . Figure 3.2 illustrates this schematically. The energy range below the gap, called the *valence band*, is nearly completely occupied with electrons. The energy range above the gap, called the *conduction band* is, however, nearly empty. To excite an electron by the absorption of a photon, the photon must have at least the energy $\hbar\omega = \varepsilon_G$. Photons with smaller energies cannot excite electrons. They are

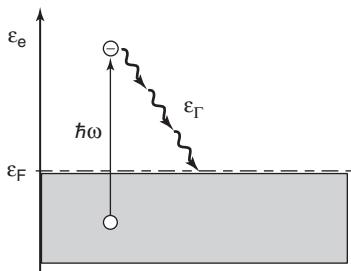


Figure 3.1 Excitation of an electron in the conduction band of a metal by the absorption of a photon with energy $\hbar\omega$ and the subsequent loss of the excitation energy by the generation of single phonons with energy ε_Γ .

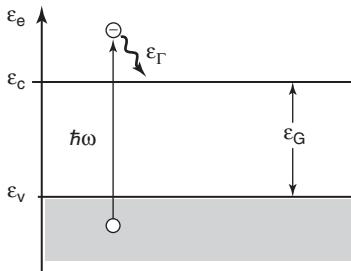


Figure 3.2 Excitation of an electron from the valence band to the conduction band of a semiconductor by the absorption of a photon with energy $\hbar\omega$.

not absorbed. They are either transmitted or reflected by the semiconductor. (We ignore the weak absorption by phonons, impurities, and free charge carriers.)

Electrons being excited well into the conduction band in a semiconductor, as shown in Figure 3.2, start losing their energy as quickly as in a metal by stepwise generation of phonons. However, once the electrons have reached the lower edge ε_c of the conduction band, losing energy in small steps by generating phonons is no longer possible, because there are no states for electrons with a little less energy. Returning to a state in the valence band requires the loss of the gap energy ε_G in a single step. Possible processes are the simultaneous generation of a large number of phonons or, alternatively, the emission of a photon. Both processes are, however, much less probable than the stepwise generation of phonons in a continuum of electron states. As a result, the electrons may “live” up to 10^{-3} s in the conduction band. It is this comparatively long time that allows the processes for the conversion of electron energy into electrical energy to take place.

3.1

Electrons in Semiconductors

As we have already seen for photons in a cavity, the density of the electrons $d n_e$ with energy ε_e distributed over the energy interval $d \varepsilon_e$ comprises the density of states $D_e(\varepsilon_e)$ and a distribution function $f_e(\varepsilon_e)$, which defines the occupation of the electron states

$$d n_e(\varepsilon_e) = D_e(\varepsilon_e) f_e(\varepsilon_e) d \varepsilon_e \quad (3.1)$$

3.1.1

Distribution Function for Electrons

The distribution of the electrons over the states must satisfy three conditions:

- 1) According to the Pauli exclusion principle for particles with nonintegral spin, there can never be more than one particle in the same quantum state. This applies for electrons, with a spin $\hbar/2$, in contrast to photons, which have a spin \hbar .
- 2) The occupation of the states depends only on the energy and not, e.g. on the momentum.
- 3) The occupation of the states must lead to a minimum of the free energy $F = E - TS$.

If the distribution of electrons over the states were to take place in such a way that the energy E is a minimum, the valence band would be completely occupied and the conduction band would be empty for all temperatures T . Under these conditions, the entropy S of the electrons would be zero, since there is only one possibility of producing a fully occupied valence band and an empty conduction band. Promoting a few electrons from the valence band to the conduction band would then lead to not only an increase in the energy, but to a still greater increase in the entropy as well. This is because there are now many possibilities for removing an electron from any of 10^{23} states per cubic centimeter of the valence band and exciting it to any of 10^{23} states per cubic centimeter of the conduction band. The “heat” TS increases and the free energy F is reduced. If there is already a certain number of electrons in the conduction band, the increase in entropy with another electron transition is correspondingly less. The increase in TS exactly compensates for the increase in the energy E when the free energy F reaches its minimum value.

The distribution function that satisfies all of these conditions is the Fermi distribution

$$f_e(\varepsilon_e) = \frac{1}{\exp [(\varepsilon_e - \varepsilon_F) / kT] + 1} \quad (3.2)$$

This contains the Fermi energy ε_F as a characteristic energy. Figure 3.3 shows the Fermi distribution function. For states with $\varepsilon_e \ll \varepsilon_F$, $f_e(\varepsilon_e) \approx 1$, so that the states are completely occupied. Conversely, for $f_e(\varepsilon_e \gg \varepsilon_F) \approx 0$, states with $\varepsilon_e \gg \varepsilon_F$ are not occupied. Half of the states with $\varepsilon_e = \varepsilon_F$ are occupied.

3.1.2

Density of States $D_e(\varepsilon_e)$ for Electrons

In isolated atoms, the electrons have discrete energy values, separated by large gaps on the energy scale. Decreasing the interatomic distance down to only a few Å in a solid introduces interaction among the atoms and causes the previously identical energy values to split into as many different values as the solid has atoms. This means that the originally discrete energy values now become energy

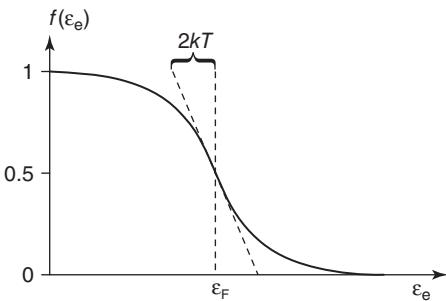


Figure 3.3 The Fermi distribution function $f(\varepsilon_e)$ defines the probability that an electron occupies a state with energy ε_e .

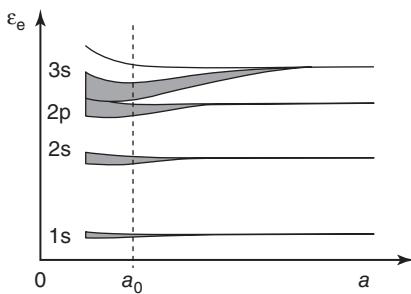


Figure 3.4 Energy of electron states in sodium as a function of the interatomic distance a between the sodium atoms. a_0 is the interatomic distance in solid sodium. States in the shaded energy ranges are occupied by electrons.

ranges in which the energy values lie so closely together that they appear to form a continuum. These regions of permitted electron energies are known as *bands*. The stronger the interaction of the electrons with neighboring atoms, the wider the bands become. The interaction is of course strongest for the outer shell valence electrons. Since these electrons have greater energies, the width of the bands increases with increasing energy, while the gaps between the bands decrease correspondingly until they disappear entirely at a certain energy. At higher energies the bands then overlap. Figure 3.4 illustrates this behavior for sodium atoms.

Sodium has the chemical valence one, and correspondingly only one of the two 3s electron states per atom is occupied. The band originating from the 3s states in solid sodium is therefore only half filled with electrons. In the continuous energy range of the 3s band, a great many unoccupied states are thus available to these “valence” electrons. This allows them to be given the additional energy, going along with the necessary additional velocity required for a charge current to flow. A partly occupied band is the prerequisite for metallic conductivity. A partly occupied band is therefore called a *conduction band*. Conversely, pure semiconductors and insulators are nonconductors, because (at least at $T=0\text{ K}$) the uppermost occupied energy band is completely occupied. This is called the *valence band*. The next higher band, separated by the energy gap ε_G , is the conduction band, which (at least at $T=0\text{ K}$) is unoccupied.

To determine the density of states, we proceed exactly as for photons. First, making use of the uncertainty principle, we determine the volume per state and the density of states in momentum space. Then, because the distribution function depends only on the energy, we have to convert the density in momentum space to the density of states per energy interval using the relationship between energy and momentum for electrons.

As for photons, two states differ in location and momentum by at least as much as the uncertainty principle defines, resulting in a phase space volume per state of

$$(\Delta x)^3 (\Delta p)^3 = h^3$$

If we treat the electrons as not localized, their uncertainty in position is $(\Delta x)^3 = V$, where V is the volume of the entire crystal. In momentum space, a state then has the volume

$$(\Delta p)^3 = \frac{h^3}{V}$$

All states with momentum $|p'| \leq |p|$ fill a spherical volume of $(4/3)\pi|p|^3$ in momentum space. The number N of states is found by dividing this volume by the volume of a state in momentum space:

$$N(|p|) = \frac{4\pi|p|^3 V}{3h^3} \quad (3.3)$$

As for photons, these momentum states apply for two electrons with opposite spins. The number of states available for electrons with momentum $|p'| \leq |p|$ is therefore

$$N_e(|p|) = \frac{8\pi|p|^3 V}{3h^3} \quad (3.4)$$

To determine the density of states as a function of the electron energy, we need the relationship between momentum and energy, which for electrons in a crystal might be quite different from that for a free electron. Here we recall that the probability for the occupation of states depends only on their energy. In order that no current flows in the equilibrium state, an equal number of states must be occupied in any arbitrary direction at positive and negative momenta. For this reason, the states must have the same density per energy at positive and negative momenta. Since the states are equidistant in momentum space, we conclude that the energy must be an even function of the momentum. To a first approximation, then

$$\varepsilon_e = \varepsilon_C + \alpha p^2 + \dots$$

By analogy with free electrons, we set $\alpha = 1/(2m_e^*)$ and call m_e^* the effective mass of the electrons. The kinetic energy of the electrons in the vicinity of ε_C is then given by

$$\varepsilon_{e,\text{kin}} = \varepsilon_e - \varepsilon_C = \frac{p^2}{2m_e^*} \quad (3.5)$$

If the energy depends on the direction of the momentum, as for noncubic crystals, the effective mass m_e^* is a tensor. The energies for oppositely directed momentum values are, however, always identical, so that $\varepsilon_{e,\text{kin}}(p) = \varepsilon_{e,\text{kin}}(-p)$.

Setting p from Equation 3.5 into Equation 3.4, assuming that the effective mass does not depend on the energy, gives the number of states between ε_e and ε_C

$$N_e(\varepsilon_e) = \frac{8\pi V(2m_e^*)^{3/2}}{3h^3} (\varepsilon_e - \varepsilon_C)^{3/2} \quad (3.6)$$

We obtain the density of states D_e in the conduction band as the number of electron states per volume and per energy interval at the energy ε_e by differentiating Equation 3.6.

$$D_e(\varepsilon_e) = \frac{1}{V} \frac{dN_e}{d\varepsilon_e} = 4\pi \left(\frac{2m_e^*}{h^2} \right)^{3/2} (\varepsilon_e - \varepsilon_C)^{1/2} \quad (3.7)$$

Figure 3.5 shows the density of states of the conduction and valence bands for germanium. For the following discussion it is only important that the densities of states at the upper and lower boundaries of a band show square-root behavior as a function of the energy ε_e in agreement with Equation 3.7, provided we choose m_e^* properly. For the upper boundary of a band, at ε_V for the valence band, for example, we must be prepared to choose $m_e^* < 0$. Treating electrons as quasi-free particles although they are in the field of the atoms ($\approx 10^{10} \text{ V m}^{-1}$) will lead to more peculiar properties. For electrons in the conduction band close to ε_C , we actually

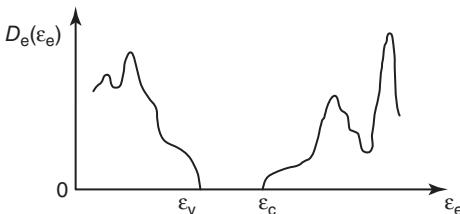


Figure 3.5 Density of states for electrons in the conduction and valence bands of the semiconductor germanium.

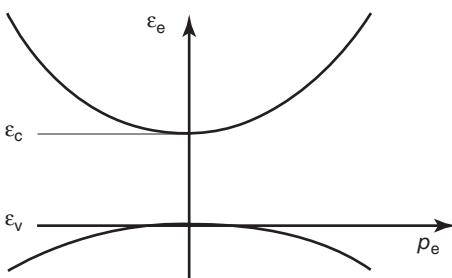


Figure 3.6 Energy ε_e of the electrons for a direct semiconductor as a function of their momentum p_e . The minimum of the conduction band and the maximum of the valence band occur at the same momentum $p_e = 0$.

expect the relationship between energy and momentum shown in Figure 3.6,

$$\epsilon_e - \epsilon_C = \frac{p_e^2}{2m_e^*} \quad (3.8)$$

and obtain the effective mass of the electrons by matching

$$\frac{1}{m_e^*} = \frac{d^2\epsilon_e}{dp_e^2}$$

to the actual energy–momentum relationship. At the upper boundary ϵ_V of the valence band $\epsilon_e - \epsilon_V < 0$ and according to Equation 3.8 and to $d^2\epsilon_e/dp_e^2 < 0$, a negative effective mass must be assigned to the electrons.

For the $\epsilon_e(p_e)$ relationship in Figure 3.6, the excitation of an electron from the valence band to the conduction band with the smallest possible energy, $\epsilon_C - \epsilon_V = \epsilon_G$, occurs without a change in momentum. This excitation is known as a *direct transition*, and semiconductors with this band structure are called *direct semiconductors*. An example of a direct semiconductor is GaAs (gallium arsenide).

Figure 3.7 shows the band structure of an *indirect semiconductor*. This behavior is unexpected from the idea of free particles. Here, the electrons have their lowest energy ϵ_C in the conduction band at a momentum different from zero, leading to the relation

$$\epsilon_e - \epsilon_C = \frac{(p_e - p_{e,0})^2}{2m_e^*} \quad (3.9)$$

Excitation from the maximum of the valence band to the minimum of the conduction band is only possible with a change of momentum. This type of transition is called an *indirect transition*.

Since in the absence of current the total momentum of the electrons vanishes, the band structure, i.e. the $\epsilon_e(p_e)$ relationship, must be symmetric with respect to the energy axis at $p_e = 0$. In particular, there must always be an even number of minima for the conduction band when these lie at $p_e \neq 0$. Examples of important indirect semiconductors are germanium (Ge) and silicon (Si).

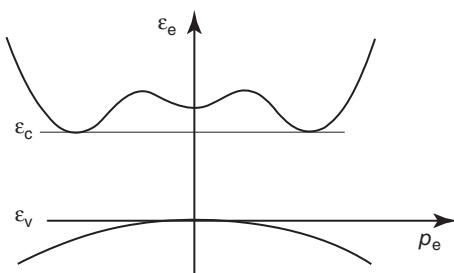


Figure 3.7 Energy of the electrons for an indirect semiconductor as a function of their momentum. The minimum of the conduction band and the maximum of the valence band occur at different values of the momentum p_e of the electrons.

3.1.3

Density of Electrons

The density of electrons in the energy interval $\varepsilon_e, \varepsilon_e + d\varepsilon_e$ is

$$dn_e(\varepsilon_e) = D_e(\varepsilon_e)f_e(\varepsilon_e)d\varepsilon_e \quad (3.10)$$

Integration over all energies in the conduction band gives the density of free electrons in the conduction band, or simply the density of electrons. From this point on, the term electrons will refer only to the electrons in the conduction band:

$$n_e = \int_{\varepsilon_C}^{\infty} D_e(\varepsilon_e)f_e(\varepsilon_e)d\varepsilon_e \quad (3.11)$$

For the integration, we can make use of the density of states from Equation 3.7, which is, however, valid only for the lower part of the conduction band, and apply it to the entire band (and even beyond this $\rightarrow \infty$) because the exponential decrease of $f_e(\varepsilon_e)$ leads to vanishingly small integrands for large values of ε_e . For $\varepsilon_F < \varepsilon_C - 3kT$, we can ignore the “+1” in the denominator of the expression for the Fermi distribution in Equation 3.2, and find from Equation 3.11

$$n_e = N_C \exp\left(-\frac{\varepsilon_C - \varepsilon_F}{kT}\right) \quad (3.12)$$

with

$$N_C = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} \quad (3.13)$$

N_C is called the *effective density of states* of the conduction band. For $m_e^* = m_e$ and $T = 300$ K, it has the value

$$N_C = 2.51 \times 10^{19} \text{ cm}^{-3} \quad (3.14)$$

The simplification to the integration obtained by neglecting the “+1” in the denominator of the Fermi function is permissible for $n_e \ll N_C$.

Since a completely occupied valence band does not permit charge transport, the few unoccupied states, known as *holes*, play an important role. Using the same approximation as for electrons, the density of the holes is

$$n_h = \int_{-\infty}^{\varepsilon_V} D_e(\varepsilon_e)[1 - f_e(\varepsilon_e)]d\varepsilon_e = N_V \exp\left(-\frac{\varepsilon_F - \varepsilon_V}{kT}\right) \quad (3.15)$$

where $N_V = 2(2\pi m_h^* kT/h^2)^{3/2}$ is the effective density of states of the valence band containing the effective mass of the holes m_h^* at $\varepsilon_e = \varepsilon_V$. It will be shown in the next section that at $\varepsilon_e = \varepsilon_V$ it is $m_h^* = -m_e^* > 0$ and $\varepsilon_e = -\varepsilon_h$.

Before examining the properties of holes more closely, we will state the important relationship

$$\begin{aligned} n_e n_h &= N_C \exp\left(-\frac{\varepsilon_C - \varepsilon_F}{kT}\right) N_V \exp\left(-\frac{\varepsilon_F - \varepsilon_V}{kT}\right) \\ &= N_C N_V \exp\left(-\frac{\varepsilon_C - \varepsilon_V}{kT}\right) = N_C N_V \exp\left(-\frac{\varepsilon_G}{kT}\right) \end{aligned} \quad (3.16)$$

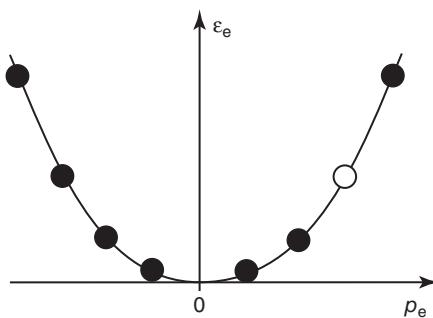


Figure 3.8 A band with a missing electron of $p_e > 0$ has a total momentum $\sum p_e < 0$ resulting in an electrical current $j_Q > 0$.

The product of the electron density and the hole density does not depend on the position of the Fermi energy and therefore it does not depend on the individual densities of either the electrons or the holes. Consequently, it cannot be influenced by doping. In a pure, so-called intrinsic semiconductor, the electrons in the conduction band originate from the valence band. The density of the electrons n_e is then equal to the density of holes n_h and both are known as *the intrinsic density* n_i .

$$n_e n_h = n_i^2 = N_C N_V \exp\left(-\frac{\epsilon_G}{kT}\right) \quad (3.17)$$

For an intrinsic semiconductor, the position of the Fermi energy is not exactly in the middle of the energy gap because of different effective densities of states in the conduction and the valence bands. It rather follows from the condition $n_e = n_h$ in Equations 3.12 and 3.15

$$\epsilon_F = \frac{1}{2}(\epsilon_V + \epsilon_C) + \frac{1}{2}kT \ln \frac{N_V}{N_C} \quad (3.18)$$

or expressed in terms of the effective masses contained in N_C and N_V

$$\epsilon_F = \frac{1}{2}(\epsilon_V + \epsilon_C) + \frac{3}{4}kT \ln \frac{m_h^*}{m_e^*} \quad (3.19)$$

Since we know the distribution of the electrons over the states, we can calculate their mean energy

$$\langle \epsilon_e \rangle = \frac{1}{n_e} \int_{\epsilon_C}^{\infty} \epsilon_e D_e(\epsilon_e) f_e(\epsilon_e) d\epsilon_e = \epsilon_C + \frac{3}{2}kT \quad (3.20)$$

For the electrons, ϵ_C is a potential energy. The mean kinetic energy of the electrons (in the conduction band) is $\langle \epsilon_e - \epsilon_C \rangle = (3/2)kT$, showing that the electrons in a nondegenerate semiconductor ($\epsilon_F < \epsilon_C - 3kT$), where $n_e \ll N_C$, form an ideal gas.

3.2 Holes

For each electron state with momentum p_e , there is also a state with the oppositely directed momentum. For a completely filled band, the total momentum is therefore zero, and no current flows.

We will illustrate the properties of holes for the example of a band with a single unoccupied state. For this purpose, we will first remove an electron with charge $-e$ and velocity \tilde{v}_e from the valence band and secondly, in an alternative description, create the missing electron by adding a hole to the completely filled valence band. The, as yet unknown, properties of the hole will then follow from a comparison of the two descriptions.

The current in a band results from summing over the velocities v_e of all occupied states. The removal of an electron with the velocity \tilde{v}_e from the completely filled band, results in

$$j_Q = -\frac{e}{\text{Vol}} \sum_{\text{occupied states}} v_{e,i} = -\frac{e}{\text{Vol}} \underbrace{\sum_{\text{all states}} v_{e,i}}_{=0 \text{ (filled band)}} - \left(-\frac{e}{\text{Vol}} \tilde{v}_e \right) = \frac{e}{\text{Vol}} \tilde{v}_e \quad (3.21)$$

The same current is obtained by introducing a hole with unknown charge q_h and unknown velocity \tilde{v}_h to a completely filled band:

$$j_Q = -\frac{e}{\text{Vol}} \underbrace{\sum_{\text{all states}} v_{e,i}}_{=0 \text{ (filled band)}} + \underbrace{\left(\frac{q_h}{\text{Vol}} \tilde{v}_h \right)}_{\text{hole}} \quad (3.22)$$

so that

$$j_Q = \frac{e}{\text{Vol}} \tilde{v}_e = \frac{q_h}{\text{Vol}} \tilde{v}_h \quad (3.23)$$

The charge of a band in which an electron is missing is the same as the charge resulting from adding a positive elementary charge to a completely filled band, that is, the charge of the hole is

$$q_h = +e$$

From Equation 3.23 it then follows that $v_h = v_e$ for a hole in a state in which the missing electron has a velocity v_e . This implies that the accelerations of the missing electron (in fact, the sum over all electrons in the band except the missing electron) and of the hole, as a reaction to an electric field E must be identical, so that

$$a_e = -\frac{eE}{m_e^*} = \frac{eE}{m_h^*} = a_h$$

The effective mass of the hole is therefore

$$m_h^* = -m_e^*$$

For the momentum of all electrons of a band in which an electron with momentum \tilde{p}_e is missing, we obtain

$$\bar{p} = \sum_{\text{filled band}} \bar{p}_i - \tilde{p}_e = \sum_{\text{filled band}} \bar{p}_i + \tilde{p}_h$$

$$\tilde{p}_h = -\tilde{p}_e$$

For the energy ε_h of the hole representing the missing electron of energy ε_e , we obtain in the same way

$$\varepsilon_h = -\varepsilon_e$$

The properties of a band result from summing either over all occupied states or over all unoccupied states. For a nearly empty band, such as the conduction band, a description in terms of the occupied states is simple because the few electrons present form an ideal gas. For a nearly filled band, such as the valence band, a description in terms of the occupied states is complicated, because the electrons at the upper edge of the band have a negative effective mass and their velocities and momenta are oppositely directed. A nearly filled band can thus be described more easily in terms of the few unoccupied states, i.e. the holes. These then carry a positive charge, have a positive effective mass, and, as with the electrons in the conduction band, form an ideal gas. Their mean energy is

$$\langle \varepsilon_h \rangle = -\varepsilon_V + \frac{3}{2}kT \quad (3.24)$$

The holes are not merely an imaginary concept, but are just as real as the electrons since the properties of the band can be equally well described in terms of the occupied or the unoccupied states. It is therefore entirely unnecessary first to explain phenomena in which the states of the valence band participate, in terms of electrons, before making the transition to the hole description. The description of holes in the valence band as positively charged particles with a positive effective mass is, in fact, so real that with sufficient kinetic energy they can knock electrons out of their chemical bond, that is, excite them by impact ionization from the valence band to the conduction band.

Figure 3.9 illustrates the transition of an electron from the valence band to the conduction band following the absorption of a photon γ with energy $\hbar\omega$. In this process, the semiconductor absorbs the energy and the momentum of the photon.

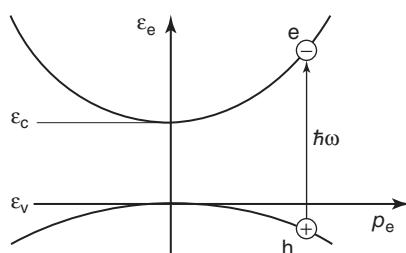


Figure 3.9 Creation of an electron–hole pair by the absorption of a photon.

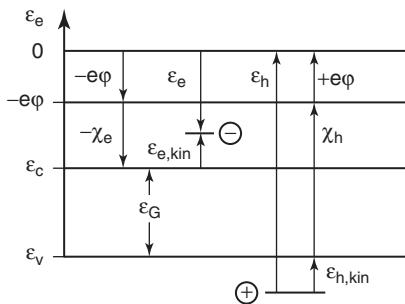
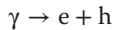


Figure 3.10 Energy scale for electrons and holes in a semiconductor. The binding energy of an electron in a state at the lower boundary of the conduction band is referred to as the *electron affinity* χ_e .

As a result of the excitation, the conduction band now has an extra electron with momentum \bar{p}_e and energy ϵ_e , whereas the valence band has an additional hole with momentum \bar{p}_h and energy ϵ_h . We therefore see the excitation as the creation of an electron (in the conduction band) and a hole (in the valence band) and describe this process as the chemical reaction



This process must of course satisfy the conservation of momentum $\bar{p}_\gamma = \bar{p}_e + \bar{p}_h$ and the conservation of energy $\epsilon_\gamma = \hbar\omega = \epsilon_e + \epsilon_h$.

Figure 3.10 shows an energy scale for electrons. Its zero point is defined by the energy of a free electron in vacuum in an electrical potential $\varphi = 0$ and having a kinetic energy $\epsilon_{e,\text{kin}} = 0$. In relation to this zero point, the electrons bound in the semiconductor have a negative energy ϵ_e . The energy ϵ_h of the holes is then positive. For them the boundaries of the bands ϵ_c , ϵ_v should really be drawn folded upward about the zero line. This representation is, however, complicated and is not commonly used. Instead, we enter the hole energies in the same way as the electron energies and take account of their magnitude and sign by the length and direction of arrows. Arrows pointing upward indicate positive energies. For electron energies the arrows point downward, since they are negative with respect to the zero line, and for hole energies the arrows point upward. With this representation, the sum $\epsilon_e + \epsilon_h$ is then equal to the difference of the distances from the zero line in Figure 3.10.

3.3 Doping

Doping of semiconductors means the introduction of impurity atoms. In the simplest case, these atoms replace the atoms of the semiconductor at their lattice positions. Figure 3.11 depicts this schematically for a lattice of atoms with a valency of four. Donors (D) are impurity atoms that, as a rule, have more valence electrons than is necessary for chemical bonding with the neighboring atoms. In most practical cases, there is just one electron not required for bonding. It is electrically bound to its atom by Coulomb forces as a negative charge in the field of a positive

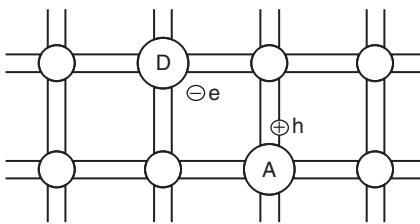


Figure 3.11 By doping lattice atoms are replaced by impurity atoms having a higher (D) or lower (A) valency.

charge. To a first approximation, we start with the binding energy of the electron to the proton in the hydrogen atom.

$$\varepsilon_H = \frac{m_e e^4}{2(4\pi\epsilon_0)^2 \hbar^2} = 13.6 \text{ eV}$$

Since the donor is, however, in a semiconductor and not in vacuum, the electric field binding the electron to the donor is somewhat weakened as a result of polarization of the neighboring atoms. We must therefore replace the dielectric permittivity of free space ϵ_0 in the binding energy of the electron for the H atom by the permittivity of the semiconductor $\epsilon\epsilon_0$. For typical semiconductors such as Ge, Si, and GaAs, the dielectric function ϵ has values > 10 for frequencies up to 10^{15} Hz. Such a large value leads to a drastic reduction of the binding energy down to < 0.1 eV. For the same reason, in fact, (the chemical bond of) NaCl dissolves in H₂O. In addition, we have to use the effective mass of the electron in the semiconductor instead of the mass of a free electron. This gives for the energy ε_D of the electron occupying a donor state

$$\varepsilon_D = \varepsilon_C - 13.6 \text{ eV} \frac{m_e^*}{m_e} \frac{1}{\epsilon^2} \quad (3.25)$$

The electron bound to the donor has the energy ε_D which, because of the weak Coulomb binding, is only slightly smaller than the lower edge ε_C of the conduction band, the lowest energy of free electrons within the semiconductor. Donors therefore donate their electrons easily to the conduction band, a property that gives them their name.

Acceptors (A) are impurity atoms that, as a rule, have one valence electron less than is necessary for chemical bonding with the neighboring atoms. An electron that fills this hole in the bond has no Coulomb attraction to the impurity atom. Consequently, it is not bound as strongly as an electron in the valence band. Since the missing Coulomb bond is, however, only a weak bond, as we have already seen for donors, the energy of an electron at the acceptor is only slightly greater than the upper edge ε_V of the valence band. Acceptors therefore accept an electron easily (i.e. with little excess energy) from the valence band, a property from which their name is derived.

In the hole picture, which, as we have seen, is simpler for the valence band, this means that holes are only weakly bound to acceptors and are easily (i.e. with little expenditure of energy) donated to the valence band.

For lattices of Ge and Si, with a valency of four, P (phosphorus) atoms or As (arsenic) atoms (both with a valency of five) are the most common donors, whereas B (boron) or In (indium) (both with a valency of three) are the usual acceptors.

For the GaAs lattice, consisting of Ga (with a valency of three) and As (with a valency of five), silicon (with a valency of four) is a donor when it is incorporated into a Ga position and an acceptor when it is incorporated into an As position. The type of incorporation in the lattice depends on the temperature. A more important acceptor is Zn (zinc), with a valency of two, at a Ga position, and a more important donor is Cl (chlorine), with a valency of seven, at an As position.

Homogeneous doping is achieved technically by adding the dopant to a semiconductor melt. Inhomogeneous doping, important for semiconductor devices, as we will see, is achieved at high temperatures by the diffusion of dopant atoms from the vapor phase or from the liquid or solid phase in contact with the semiconductor. Another possibility is the implantation of the doping material in ionic form into the semiconductor. With this ion implantation, the spatial distribution of dopant atoms can be more sharply defined. On the other hand, greater damage occurs to the lattice, which may be annealed at high temperatures.

Whether a doping material actually changes the density of electrons or holes depends on the energy of the electrons in the dopant atom, and also on the temperature. For the occupation of all states, the Fermi distribution applies for the bands and the dopant atoms. This defines the density of the positively charged holes and the ionized donors as well as that of the negatively charged electrons and the ionized acceptors. By doping, a semiconductor remains electrically neutral, so that the charge density is

$$\rho = e(n_h + n_D^+ - n_e - n_A^-) = 0 \quad (3.26)$$

Here n_e from Equation 3.12 and n_h from Equation 3.15 are just as much functions of the Fermi energy ε_F as

$$n_D^+ = n_D \left\{ 1 - \left[\exp \left(\frac{\varepsilon_D - \varepsilon_F}{kT} \right) + 1 \right]^{-1} \right\}$$

and

$$n_A^- = n_A \left[\exp \left(\frac{\varepsilon_A - \varepsilon_F}{kT} \right) + 1 \right]^{-1}$$

Equation 3.26 therefore defines the position of the Fermi energy ε_F and with it the values of all the concentrations.

Figure 3.12 shows the temperature dependence of the electron and hole concentrations in an n-type semiconductor with a band gap of 1 eV and effective densities of states of $N_C = N_V = 2 \times 10^{19} \text{ cm}^{-3}$, doped with 10^{16} donors per cubic centimeter with an energy level $\varepsilon_D = \varepsilon_C - 0.05 \text{ eV}$. In a wide temperature range around room temperature, the electron concentration is independent of the temperature and is equal to the donor concentration. According to Equation 3.17, the hole

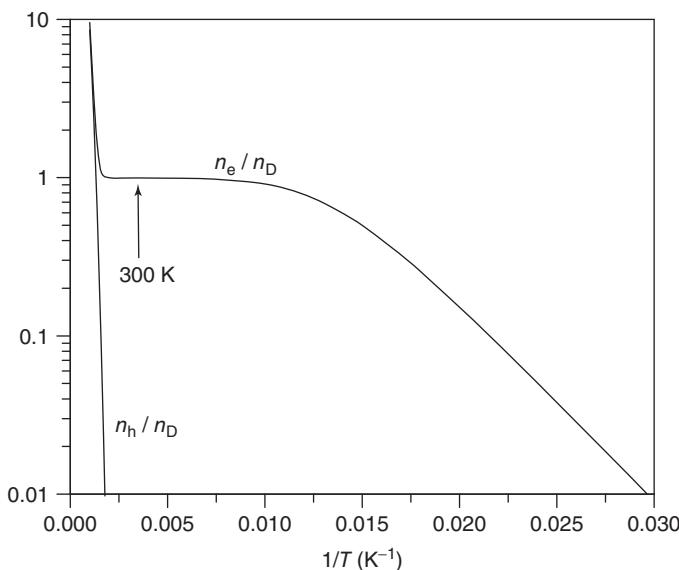


Figure 3.12 Temperature dependence of electron and hole concentrations in an n-type semiconductor with a band gap of 1 eV, containing 10^{16} donors per cm^3 with an energy level $\varepsilon_D = 0.05 \text{ eV}$ below the conduction band edge ε_C .

concentration falls off steeply with decreasing temperature. At lower temperatures not all of the donors are ionized. From the decrease of the electron concentration with decreasing temperature, the activation energy for exciting electrons from the donor level to the conduction band (0.05 eV in Figure 3.12) can be obtained. This works, however, only if the donor concentration is much larger than the acceptor concentration. In a more compensated but still n-type semiconductor, the temperature dependence of the electron concentration does not reveal the activation energy of the donors. At high temperatures, electron and hole concentrations rise steeply with increasing temperature. They become equal and much larger than the doping concentration and the semiconductor becomes intrinsic.

As a result of doping with shallow donors to which electrons are only weakly bound, the Fermi energy ε_F at room temperature is less than ε_D , since the donors are unoccupied. For doping with shallow acceptors to which holes are only weakly bound, the Fermi energy ε_F at room temperature is greater than ε_A . At room temperature, donors and acceptors are almost completely ionized.



$$n_e \approx n_D$$

$$n_h \approx n_A$$

With the incorporation of donors, a semiconductor becomes an electron conductor or n-type conductor, and with the incorporation of acceptors it becomes a hole conductor or p-type conductor. Table 3.1 gives the electron and hole densities and the position of the Fermi energy in an n-type and p-type semiconductor doped with either shallow donors or shallow acceptors at about room temperature.

Table 3.1 Electron and Hole Densities In n-type and p-type Semiconductors.

	n_e	n_h	ϵ_F
n-Type	$n_e \approx n_D$	$n_h = \frac{n_i^2}{n_e} = \frac{n_i^2}{n_D}$	$\epsilon_F = \epsilon_C - kT \ln \frac{N_C}{n_D}$
p-Type	$n_e = \frac{n_i^2}{n_h} = \frac{n_i^2}{n_A}$	$n_h \approx n_A$	$\epsilon_F = \epsilon_V + kT \ln \frac{N_V}{n_A}$

The Fermi energy is obtained from Equation 3.12 and is

$$\epsilon_F = \epsilon_C - kT \ln \frac{N_C}{n_e}$$

or from Equation 3.15

$$\epsilon_F = \epsilon_V + kT \ln \frac{N_V}{n_h}$$

Typical doping concentrations are in the range of 10^{15} cm^{-3} to 10^{19} cm^{-3} . This is relatively small compared with the lattice atom density, which is about 10^{23} cm^{-3} . This slight admixture does not significantly change the chemical nature of the semiconductor. The energies of the electron and hole states in the bands thus remain largely the same. In order for a slight doping to be effective, the semiconductor must be purified down to concentrations that are small relative to the doping concentrations.

Incorporation of both donors and acceptors in a semiconductor does not lead to increased concentrations of both electrons and holes. Equation 3.17, $n_e n_h = n_i^2$, applies to doped semiconductors as well. The electrons donated from the donors are instead largely taken up by the acceptors. (Alternatively, one could say that the holes that the acceptors leave behind are found in the donor atoms.) Donors and acceptors are both ionized without creating free electrons and free holes. For shallow impurities, only the difference between the donor and acceptor concentrations is effective. Intrinsic concentrations of electrons and holes can therefore be obtained in impure semiconductors as well by compensated doping, making the concentrations of donors and acceptors equal.

Donors have the property that they are neutral with respect to the charge at the unperturbed lattice site when they are occupied by an electron and carry a positive elementary charge when they are unoccupied. Acceptors, on the other hand, have a negative charge in the occupied state and are neutral in the unoccupied state. Donorlike impurities exist with energies not only in the vicinity of the conduction band. For deep donors, the hydrogen atom is not a good model. Deep donors are ineffective as donors, and for solar cells are, in fact, even harmful. If their electron energies are in the middle of the band gap, they act as recombination centers, and if they are close to the valence band they act as hole traps and remove holes from the valence band. Similarly, acceptor-like impurities with electron energies in the middle of the band gap also act as recombination centers, and if they are close to the conduction band they act as electron traps. For good solar cells, in particular,

impurities with electron energies in the middle of the band gap must be eliminated to avoid excessive recombination, as discussed in Section 3.6.2.

3.4

Quasi-Fermi Distributions

In a solar cell additional electrons and holes are produced by the absorption of photons from solar radiation. Actually, after their generation, the electron and the hole are bound to each other by Coulomb interaction to form an exciton. The exciton binding energy follows from Equation 3.25 after accounting for the polarizability of the semiconductor lattice by replacing ϵ_0 with $\epsilon\epsilon_0$. Polarization of the lattice and the smaller mass of the hole compared to the mass of the nucleus of an atom reduce the binding energy of the exciton to well below kT_0 (at room temperature). In materials with a large dielectric permittivity like in most inorganic semiconductors, excitons are unstable at room temperature and electrons and holes are essentially free. Figure 3.13 illustrates the generation process schematically. Immediately after the generation of additional electrons and holes, their energy distribution reflects the broad energy spectrum of the absorbed photons (indicating the high photon temperature) and is different from that in the dark state. Owing to collisions with the lattice by which phonons are emitted and absorbed, the energy distribution changes very rapidly. After about 100 collisions over a timescale of 10^{-12} s, the dark-state energy distribution is established where electrons and holes are each in thermal and chemical equilibrium with the phonons and have a mean kinetic energy of $\epsilon_{e,h,\text{kin}} = (3/2)kT_0$.

After this thermalization, the charge carriers exist for as long as their “lifetime” in their bands, which is long compared with their thermalization time, before they disappear by recombination. Since the thermalization time is so short, the steady-state energy distribution in the bands differs very little from a room temperature

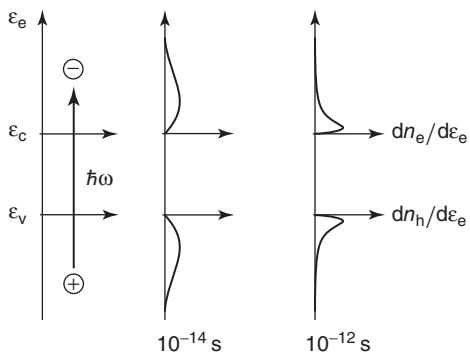


Figure 3.13 The broad energy distribution of electrons and holes, right after their generation by solar radiation, reflects the broad energy spectrum of the photons. After about 10^{-12} s, electrons and holes are thermalized into a narrow room temperature distribution.

distribution, even for continuous irradiation. Owing to the frequent collisions with the atoms, the electrons and holes are in a temperature equilibrium with the lattice, in which Pauli's principle is obeyed and a minimum of the free energy is established, just as in the dark state. The distribution of electrons and holes over the states in their bands must therefore be a Fermi distribution at room temperature.

By irradiation, both the electron and the hole densities are greater than in the dark, that is, $n_e > n_e^0$ and $n_h > n_h^0$, respectively. $n_e n_h$ is then greater than n_i^2 , which could not be achieved by doping. The temperature in the Fermi function required for describing the irradiated state must be the lattice temperature, because only this temperature gives a distribution with $\epsilon_{\text{kin}} = (3/2)kT_0$. Owing to the increased electron density, the Fermi energy describing their distribution in the conduction band must be closer to the conduction band than in the dark. Owing to the increased hole density, the Fermi energy describing their distribution in the valence band must be closer to the valence band.

The solution to this dilemma is as follows. There are (always) two Fermi distributions, the distribution f_C with the Fermi energy ϵ_{FC} , which applies for the occupation of the states in the conduction band and the (shallow) donors with electrons, and another Fermi distribution f_V with the Fermi energy ϵ_{FV} , which applies for the occupation of the states in the valence band and the (shallow) acceptors with electrons, and therefore also defines the hole density in the valence band.

The density of electrons (in the conduction band) is

$$n_e = N_C \exp\left(-\frac{\epsilon_C - \epsilon_{FC}}{kT}\right) \quad (3.27)$$

and that of the holes (in the valence band) is

$$n_h = N_V \exp\left(-\frac{\epsilon_{FV} - \epsilon_V}{kT}\right) \quad (3.28)$$

It then follows that

$$n_e n_h = N_C N_V \exp\left(-\frac{\epsilon_C - \epsilon_V}{kT}\right) \exp\left(\frac{\epsilon_{FC} - \epsilon_{FV}}{kT}\right) \quad (3.29)$$

or

$$n_e n_h = n_i^2 \exp\left(\frac{\epsilon_{FC} - \epsilon_{FV}}{kT}\right) \quad (3.30)$$

Figure 3.14 shows these two Fermi distributions. It also shows that especially in the energy range between ϵ_{FC} and ϵ_{FV} the distributions contradict each other. If f_C applies, the states in this energy range would have to be occupied, but if f_V applies they would have to be unoccupied. In fact, neither distribution applies in this range of energies. In an illuminated semiconductor the occupation between ϵ_{FC} and ϵ_{FV} is governed by kinetics, the preferential capture of holes or electrons. We will return to this point when recombination via impurities is discussed.

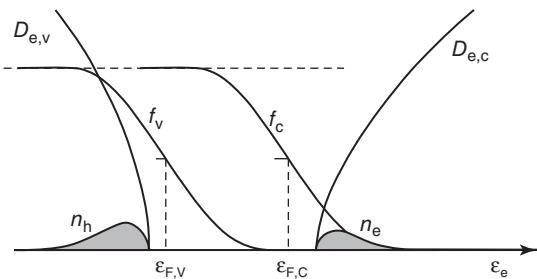


Figure 3.14 In an illuminated semiconductor, the occupations of the conduction band and the valence band are described by different Fermi distributions f_c and f_v .

3.4.1

Fermi Energy and Electrochemical Potential

On the energy scale shown in Figure 3.10, the energy of the electrons ϵ_e is divided into potential energy ϵ_C and kinetic energy ϵ_{kin} , where ϵ_C comprises the potential chemical energy $-\chi_e$ also known as the *electron affinity* and the electrical potential energy $-e\phi$. If an electron is removed from the semiconductor, as is the case when a charge current flows through a solar cell, the energy within the semiconductor decreases by the entire energy ϵ_e of the electron. But is this energy in fact the energy delivered to a load as electrical energy?

To answer this question we must consider that, according to Gibbs, the exchange of an amount of energy dE is linked to other quantities that are exchanged as well.

$$dE(S, V, N_i, Q, \dots) = T dS - p dV + \sum_i \mu_i dN_i + \varphi dQ + \dots \quad (3.31)$$

It is usual to refer to the energy exchange in terms of the quantity exchanged along with the energy. If only entropy S is exchanged with the energy, the energy exchanged is called heat. $-p dV$ is called *compressional energy*, $\mu_i dN_i$ is called the *chemical energy* of the particle species i , and φdQ is called *electrical energy*. There are many other forms of energy as well, e.g. magnetic energy, which are, however, of no interest in connection with solar cells. These energy forms are in all cases products of “intensive” variables and “extensive”, quantity-like, variables. To distinguish between these, let us imagine two identical systems, which we put together to form a new system that is twice as large. The intensive variables such as temperature T or pressure p then remain unchanged, while the values of the extensive variables – including the energy – become twice as large.

The extensive quantities, the entropy S , the volume V , the amounts N_i of the different particle species i , the charge Q and others of no importance for solar cells, are the energy carriers. For most of these, laws of conservation apply. For entropy, conservation applies only partially: it can be created, but can never be destroyed.

The intensive variables, temperature T , pressure p , the chemical potential μ_i of the particle species i and electrical potential φ define the amount of energy

H_2	O_2
T_{H_2}	T_{O_2}
p_{H_2}	p_{O_2}
μ_{H_2}	μ_{O_2}

Figure 3.15 Two compartments separated by a piston contain hydrogen H_2 and oxygen O_2 at different temperatures T , pressures p , and chemical potentials μ .

exchanged with the energy carriers. The gradients of the intensive variables drive currents of the respective energy carriers. Since the current of an energy carrier must vanish when its intensive variable is in equilibrium, the intensive variable then must have the same value everywhere.

When a system is in equilibrium with respect to *one* intensive variable, it is not necessarily in equilibrium with respect to other variables. Figure 3.15 explains this clearly. It shows two compartments of a vessel separated by a mobile piston immobilized for the moment. The two compartments contain different gases: in one there is hydrogen and in the other, oxygen, which therefore have different chemical potentials. In addition, the temperatures and pressures shall also be different. As a result of the temperature difference between the compartments, entropy flows through the still-immobilized piston until the temperatures in the two compartments are the same. Temperature equilibrium, often called simply *thermal equilibrium*, then prevails. However, the pressures and chemical potentials of the gases are still different and not in equilibrium. We will now remove the blockade of the piston, after which it will come to rest (possibly after several oscillations and dissipation of energy) at the position where the pressure is the same in both compartments. We now have equilibrium of temperature and pressure. Chemical equilibrium, that is, the same chemical potentials for hydrogen and oxygen in the two compartments, only occurs by the exchange of gas particles after opening a hole in the piston until each particle type has the same density or concentration in both compartments. Still another equilibrium, that of the chemical reaction between hydrogen and oxygen to form water, may then be attained by a small spark accompanied by a loud explosion. These equilibria are all independent of each other. That these equilibria never occur all at the same time is of paramount importance for our existence. Consequently, it is entirely meaningful to distinguish between different types of equilibria (and also nonequilibria) and therefore to name these separately as temperature equilibrium, pressure equilibrium, chemical equilibrium, etc. The common designation “thermodynamic equilibrium”, while it may sound more impressive, is however meaningless because it says nothing about the type of equilibrium. Furthermore, it also leads us to the false conclusion that “thermodynamic nonequilibrium” prevails and all equilibrium relations are invalid when only one of several intensive quantities is not in equilibrium. Fortunately, a general thermodynamic equilibrium as a state in which all conceivable intensive variables are in equilibrium does not exist in the real

world, considering that our very own existence would hardly be possible under such conditions.

In the following, we will make frequent use of the fact that some equilibria exist and allow the use of such equilibrium relations as the Fermi distribution and that there are also nonequilibrium states that are far more difficult to describe.

We will now return to the problem addressed above, namely, the question of how much electrical energy does a semiconductor deliver when an electron–hole pair is removed. Surprisingly, for solar cells just as for batteries, this is not just the energy form φdQ , since the charge Q of the solar cell does not change when electrons and holes are removed, again just as for a battery, in which exactly the same number of electrons flows in through one contact as flows out through the other contact. We know, however, that looking for electrical energy in fact means looking for that part of the energy of the electron–hole pairs not accompanied by entropy. This part is the free energy $F(T, V, N_i, Q, \dots) = E - TS$. With the removal of dN_e electrons, the free energy of a body changes by

$$dF_e(T, V, N_e, Q) = dE_e - d(TS_e) = -S_e dT - p_e dV + \mu_e dN_e + \varphi dQ$$

With the removal of dN_h holes, the free energy changes by

$$dF_h(T, V, N_h, Q) = -S_h dT - p_h dV + \mu_h dN_h + \varphi dQ$$

The total change in the free energy is $dF = dF_e + dF_h$

Even if no charge is removed from the solar cell at the end with an electrical current, we must still consider that the electrons and holes are charged particles, so that changes in their numbers are coupled to a change in the total charge.

$$dQ = z_i e dN_i$$

where $z_i = +1$ for holes and $z_i = -1$ for electrons. It then follows that

$$\mu_i dN_i + \varphi dQ = (\mu_i + z_i e \varphi) dN_i = \eta_i dN_i$$

Owing to this fundamental coupling of charge and particle number, only one coupled electrochemical equilibrium, and no separate chemical and electrical equilibria, exist for the electrons and holes. $\eta_e = \mu_e - e\varphi$ is the electrochemical potential of the electrons. For electrons in electrochemical equilibrium, it has the same value everywhere. $\eta_h = \mu_h + e\varphi$ is the electrochemical potential of the holes and for holes in electrochemical equilibrium, it has the same value everywhere.

For the steady-state operation of solar cells, the temperature T and the volume V occupied by the electrons and holes are constant. Furthermore, with the flow of an electric current, equal numbers of electrons and holes are always removed or added, i.e. $dN_e = dN_h = dN$. The change in the free energy, and thus the energy that a solar cell delivers to a load with dN electrons and holes is then

$$dF = dF_e + dF_h = (\eta_e + \eta_h) dN \quad (3.32)$$

The question now remains of how to relate the electrochemical potentials to the already known quantities. In Equation 3.20 we found the mean energy of the electrons to have the value $\langle \epsilon_e \rangle = \epsilon_C + (3/2)kT$ and, correspondingly, for the holes,

the value $\langle \varepsilon_h \rangle = -\varepsilon_V + (3/2)kT$. We arrive at the same expectation values for the energy per particle of one type by dividing the total energy for particles of one type by the number of such particles. For electrons, the total energy is

$$E_e = TS_e - p_e V_e + \eta_e N_e$$

and the mean energy per electron

$$\frac{E_e}{N_e} = \langle \varepsilon_e \rangle = T\sigma_e - \frac{p_e V_e}{N_e} + \eta_e = \varepsilon_C + \frac{3}{2}kT \quad (3.33)$$

Similarly, for holes

$$\frac{E_h}{N_h} = \langle \varepsilon_h \rangle = T\sigma_h - \frac{p_h V_h}{N_h} + \eta_h = -\varepsilon_V + \frac{3}{2}kT \quad (3.34)$$

The quantities σ_e and σ_h are the entropy per electron and per hole, respectively. Since electrons and holes are ideal gases, we can make use of the relationship between σ and the particle density n discovered by Sackur and Tetrode [2, 3] for ideal gases, after adapting it to particles with spin 1/2

$$\sigma = k \left\{ \frac{5}{2} + \ln \left[\frac{2 \left(\frac{2\pi mkT}{h^2} \right)^{3/2}}{n} \right] \right\} \quad (3.35)$$

Substituting the effective mass m^* of the electrons and holes from Equation 3.13 for m we find

$$\sigma_{e,h} = k \left(\frac{5}{2} + \ln \frac{N_{C,V}}{n_{e,h}} \right) \quad (3.36)$$

For ideal gases with particle number N the equation of state is

$$pV = NkT \quad (3.37)$$

With Equations 3.36 and 3.37, we find from Equation 3.33 for electrons

$$\begin{aligned} \langle \varepsilon_e \rangle &= \varepsilon_C + \frac{3}{2}kT = kT \left[\frac{5}{2} + \ln \left(\frac{N_C}{n_e} \right) \right] - kT + \eta_e \\ \varepsilon_C - \eta_e &= kT \ln \left(\frac{N_C}{n_e} \right) \end{aligned}$$

which leads to

$$n_e = N_C \exp \left[\frac{-(\varepsilon_C - \eta_e)}{kT} \right] \quad (3.38)$$

This is identical with Equation 3.27, and we can now identify the electrochemical potential of the electrons with their Fermi energy

$$\eta_e = \varepsilon_{FC} \quad (3.39)$$

For holes a similar result applies; thus

$$-\varepsilon_V - \eta_h = kT \ln \left(\frac{N_V}{n_h} \right) \quad (3.40)$$

and

$$n_h = N_V \exp\left(\frac{\eta_h + \varepsilon_V}{kT}\right) \quad (3.41)$$

Comparison with Equation 3.28 shows that

$$\eta_h = -\varepsilon_{FV} \quad (3.42)$$

These values are shown in Figure 3.16. The free energy delivered to a load by dN electrons and holes is

$$dF = dF_e + dF_h = (\eta_e + \eta_h)dN = (\varepsilon_{FC} - \varepsilon_{FV})dN$$

This result is in good agreement with the expectation that the electron–hole system of an unilluminated semiconductor, that is, one that sees only the 300 K ambient radiation and is in equilibrium with it, cannot deliver electrical energy because for this state $\varepsilon_{FC} - \varepsilon_{FV} = \eta_e + \eta_h = 0$. The electrochemical energy of an electron–hole pair, being the sum of the electrochemical potentials of an electron and a hole at the same location x at which the electron and the hole are in the same electrical potential φ , is equal to the chemical energy of the electron–hole pair because an electron–hole pair is uncharged:

$$\eta_e(x) + \eta_h(x) = \mu_e(x) - e\varphi(x) + \mu_h(x) + e\varphi(x) = \mu_e(x) + \mu_h(x) \quad (3.43)$$

This is important for the conversion of solar energy into chemical energy as in photosynthesis. This conversion takes place in every semiconductor exposed to light without any additional measures.

3.4.2

Work Function

The absolute value of the chemical potential for electrons μ_e is also known as the *work function*. This is the energy that must be expended in order to excite an electron from a bound state in a semiconductor or metal thermally, i.e. by thermionic

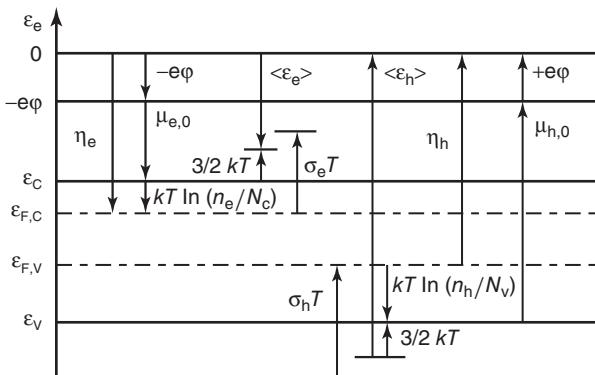


Figure 3.16 The energy forms of electrons and holes.

emission, to a state of a free electron in vacuum. We assume that while leaving the solid the electron is still in the same electrical potential so that its electrical energy $-e\varphi$ remains unchanged. For metals, in which all states up to the Fermi energy are occupied, it is evident from Figure 3.16 that the work function is equal in magnitude to the chemical potential μ_e . In semiconductors, however, there are no electrons at all with energies equal to the Fermi energy, since the Fermi energy is somewhere within the band gap.

The work function φ is determined experimentally from the temperature dependence of the electron current emitted by thermionic emission, which is found to be proportional to $\exp(-\varphi/kT)$. This emission current is proportional to the concentration of electrons $n_{e,\text{free}}$ not bound to the semiconductor material, but still belonging to the same Fermi distribution, which have an energy of at least $\varepsilon_e = -e\varphi$. As in the derivation of the concentration of electrons in the conduction band, which have an energy of at least $\varepsilon_e = \varepsilon_C$, we find from Equation 3.12

$$n_{e,\text{free}} \propto \exp\left(-\frac{-e\varphi - \varepsilon_F}{kT}\right) = \exp\left(-\frac{-e\varphi - \mu_e + e\varphi}{kT}\right) = \exp\left(\frac{\mu_e}{kT}\right)$$

for both semiconductors and metals. The thermal work function as an invariable property of a compound is characteristic for metals only. For semiconductors, it depends on the doping and is greater if a semiconductor is p-doped than if it is n-doped. Following Figure 3.16, we can decompose the chemical potential μ_e of the electrons in the semiconductor into a component $\mu_{e,0}$, determined by the chemical environment of the electrons, i.e. by the base material, and independent of their concentration, and into a concentration-dependent component

$$\mu_e = \mu_{e,0} + kT \ln\left(\frac{n_e}{N_C}\right) \quad (3.44)$$

The chemical potential is a well known quantity in chemistry. There, the standard concentration is usually taken as 1 mol/l and values for the concentration independent part of the chemical potential are given for this standard concentration. By using the effective density of states as the standard concentration in (3.44), the concentration independent part or the standard value of the chemical potential is the electron affinity. For semiconductors, the property that is characteristic for the base material is the concentration-independent component of the chemical potential of the electrons $\mu_{e,0}$ or its magnitude, the electron affinity $\chi_e = -\mu_{e,0}$, from which the work function is calculated according to Equation 3.44 for a known electron concentration.

Work functions can also be obtained experimentally from photoemission as the smallest photon energy for which the emission of electrons into the vacuum is observed. For metals, there is no significant difference compared with the thermionic work function. For semiconductors, however, photoemission does not measure the work function but the energy of the transition from the upper edge of the valence band into the vacuum. Owing to the much lower density of electrons in the conduction band, transitions from the conduction band into the vacuum are much more difficult to observe.

3.5

Generation of Electrons and Holes

Electrons and holes are produced by processes that can supply at least the minimum generation energy ϵ_G of an electron–hole pair. This includes impact ionization, in which an electron (or hole) with enough kinetic energy knocks a bound electron out of its bound state (in the valence band) and promotes it to a state in the conduction band, thereby creating an electron and a hole. The same process of exciting an electron from the valence band to the conduction band may take place with a giant lattice vibration supplying the energy or by the absorption of a photon. In the presence of impurities providing states with energies in the energy gap, the excitation can take place in several steps and the generation energy ϵ_G can be supplied in smaller portions by photons or even phonons. In all cases, in close proximity the electron and hole attract each other by their opposite charge. If their energy is insufficient to overcome the attraction, a bound state of electron and hole results, which is called an *exciton*. In inorganic semiconductors, the exciton binding energy is smaller than kT at room temperature and electrons and holes are essentially free particles. In contrast, organic semiconductors usually have exciton binding energies much larger than kT . The properties of the exciton and the problems that a large exciton binding energy presents for organic solar cells will be discussed in Section 6.9.

3.5.1

Absorption of Photons

For solar cells, the generation of electrons and holes by the absorption of photons is the most important process. The probability for the absorption of a photon of energy $\hbar\omega$ is defined by the absorption coefficient $\alpha(\hbar\omega)$, which is a material property, independent of the geometry of a body. Since absorption requires that an electron–hole pair is generated, $\alpha(\hbar\omega)$ is proportional to the density of occupied states in the valence band in which a hole can be generated, and unoccupied states in the conduction band in which an electron can be generated. The changes in occupation by the absorption of a photon must, of course, conserve momentum and energy.

Direct Transitions

Direct transitions are those in which the momentum of the electron–hole system does not change. The balance of momentum is then consistent with a reaction exclusively with photons, which have a negligible momentum as a result of the very large value of the velocity of light in the definition of their momentum $p_\gamma = \hbar\omega/c$,

$$p_\gamma = p_e + p_h \approx 0, \quad \text{and thus} \quad p_e = -p_h$$

Energy conservation, $\hbar\omega = \epsilon_e + \epsilon_h$ together with the energy–momentum relations of direct semiconductors are shown in Figure 3.6,

$$\varepsilon_e = \varepsilon_C + \frac{p_e^2}{2m_e^*} \quad \text{and} \quad \varepsilon_h = -\varepsilon_V + \frac{p_h^2}{2m_h^*} \quad (3.45)$$

leads to

$$\hbar\omega = \varepsilon_C - \varepsilon_V + \frac{p_e^2}{2m_e^*} + \frac{p_h^2}{2m_h^*}$$

Since from the conservation of momentum $p_e^2 = p_h^2 = p^2$, it follows that

$$\hbar\omega = \varepsilon_G + \frac{p^2}{2} \left(\frac{1}{m_e^*} + \frac{1}{m_h^*} \right) = \varepsilon_G + \frac{p^2}{2m_{\text{comb}}} \quad (3.46)$$

$m_{\text{comb}} = (m_e^* m_h^*) / (m_e^* + m_h^*)$ is the so-called combined mass.

In Equation 3.46 the dependence of the energy $\hbar\omega$ of a direct transition on the momentum is very similar to that of the energy–momentum relation of an electron in the conduction band (Equation 3.45). In the same way as the quantization of momentum together with this relation leads to the density of electron states in the conduction band in Equation 3.7, the energy–momentum relation in Equation 3.46 results in the so-called combined density of states for direct optical transitions,

$$D_{\text{comb}}(\hbar\omega) = \frac{4\pi}{h^3} (2m_{\text{comb}})^{3/2} (\hbar\omega - \varepsilon_G)^{1/2} \quad (3.47)$$

The probability of photon absorption, i.e. the change in the photon current density dj_γ over a distance dx by absorption is proportional to the absorption coefficient and to j_γ . This defines the absorption coefficient $\alpha(\hbar\omega)$, which is proportional to the combined density of states,

$$\frac{dj_\gamma(\hbar\omega)}{dx} = -\alpha(\hbar\omega) j_\gamma(\hbar\omega) \quad (3.48)$$

$$\alpha(\hbar\omega) \propto D_{\text{comb}}(\hbar\omega) \propto (\hbar\omega - \varepsilon_G)^{1/2} \quad (3.49)$$

Integration of Equation 3.48 leads to the well-known dependence of the photon current on the distance x from the surface at $x=0$ in an absorbing body:

$$j_\gamma(x) = j_\gamma(0) \exp(-\alpha x)$$

For semiconductors with the energy–momentum relation shown in Figure 3.6, in which direct transitions with almost no change in the momentum are possible we expect the square-root dependence on energy for the absorption coefficient. Figure 3.17 gives the absorption coefficient α of GaAs as an example.

Photons with $\hbar\omega < \varepsilon_G$ are not absorbed by electron–hole generation and $\alpha(\hbar\omega < \varepsilon_G) = 0$. Photons of this energy are either reflected or transmitted. For $\hbar\omega > \varepsilon_G$, α shows the steep rise in accordance with the theoretical expectations $\propto (\hbar\omega - \varepsilon_G)^{1/2}$ up to values of 10^4 cm^{-1} . For larger values of α the complexity of the real band structure of GaAs causes deviations from the square-root dependence. At the absorption edge at $\hbar\omega = \varepsilon_G = 1.4 \text{ eV}$, an exponential increase is superimposed over the square-root dependence. This is known as the *Urbach*

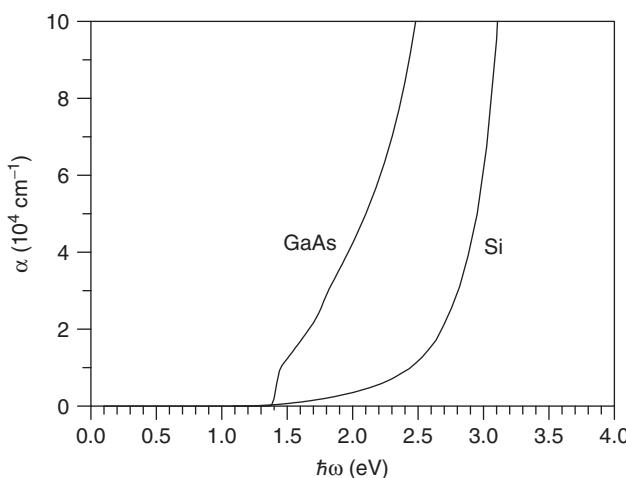


Figure 3.17 Absorption coefficient α of the “direct” semiconductor gallium arsenide and the “indirect” semiconductor silicon.

tail, which results from statistical fluctuations of the band gap caused by lattice vibrations. Since for a distance from the surface $x = 1/\alpha$ the intensity is attenuated by a factor of e , $L_\gamma = 1/\alpha$ is called the *penetration depth of the photons*. Owing to the large absorption coefficient and the small penetration depth of the photons in GaAs and other direct semiconductors, a solar cell made from these materials does not have to be thicker than only a few μm in order to absorb the absorbable part of the solar spectrum with photon energies $\hbar\omega > \varepsilon_G$.

Indirect Transitions

In an indirect semiconductor, a transition between the maximum of the valence band ε_V and the minimum of the conduction band ε_C is not possible with only the absorption of a photon, because the momentum $p_\gamma = \hbar\omega/c$ of the photon is too small. The momentum balance is satisfied through the participation of another “particle”, a lattice vibration or a phonon. Due to the large mass of the atoms, phonons with small energies $\hbar\Omega$ have a large momentum p_γ .

In an indirect transition a phonon Γ can either be absorbed with the absorption of a photon γ ,

$$\begin{aligned} \gamma + \Gamma &\rightarrow e + h \\ p_\gamma + p_\Gamma &= p_e + p_h \\ \hbar\omega + \hbar\Omega &= \varepsilon_e + \varepsilon_h \end{aligned} \tag{3.50}$$

or the phonon Γ can be emitted

$$\begin{aligned} \gamma &\rightarrow e + h + \Gamma \\ p_\gamma &= p_e + p_h + p_\Gamma \\ \hbar\omega &= \varepsilon_e + \varepsilon_h + \hbar\Omega \end{aligned} \tag{3.51}$$

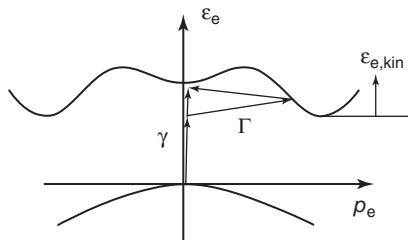


Figure 3.18 Energy ϵ_e of electron states in the conduction and valence bands between which transitions are possible with the absorption of a photon γ and the simultaneous absorption or emission of a phonon Γ .

The participation of the phonons allows photon-induced transitions from every state of the valence band to every state of the conduction band for which the energy balance shown in Figure 3.18 is satisfied. The transition probability into *one* state of the conduction band is then proportional to the number of all states in the valence band separated from this conduction band state by the photon energy $\hbar\omega$ plus or minus the phonon energy $\hbar\Omega$. To find the probability for the absorption of a photon of energy $\hbar\omega$ we have to integrate over all states of the conduction band with the energy $\epsilon_{e,\text{kin}}$ that are accessible from the valence band, with the simultaneous emission or absorption of a phonon. The smallest value for the kinetic energy of the electrons is $\epsilon_{e,\text{kin}} = 0$ for a transition to the minimum of the conduction band and the highest possible value is $\epsilon_{e,\text{kin}} = \hbar\omega \pm \hbar\Omega - \epsilon_G$ for a transition from the maximum of the valence band to the conduction band.

$$\alpha(\hbar\omega) \propto \int_0^{\hbar\omega \pm \hbar\Omega - \epsilon_G} D_C(\epsilon_{e,\text{kin}}) D_V(\hbar\omega \pm \hbar\Omega - \hbar\omega - \epsilon_G - \epsilon_{e,\text{kin}}) d\epsilon_{e,\text{kin}}$$

With the energy dependence of the densities of states of Equation 3.7 the integration yields

$$\alpha(\hbar\omega) \propto (\hbar\omega - \epsilon_G \pm \hbar\Omega)^2 \quad (3.52)$$

The plus sign applies for the simultaneous absorption of a photon and a phonon and the minus sign for the emission of a phonon with the absorption of a photon. Because of the requirement of phonon participation in order to satisfy the conservation of momentum, the absorption coefficient of an “indirect” semiconductor is small. Figure 3.17 shows the dependence of the absorption coefficient on the photon energy for silicon, a typical indirect semiconductor. The quadratic dependence on the photon energy is well developed. Near the band gap $\hbar\omega = \epsilon_G = 1.12 \text{ eV}$ the absorption coefficient of silicon is so small that the band gap cannot be seen in Figure 3.17. Also, the absorption and emission of phonons cannot be distinguished.

Because of the small value of α for indirect transitions, the penetration depth for photons in an indirect semiconductor is large. To absorb all the photons with $\hbar\omega > \epsilon_G$, which can be absorbed from the solar spectrum, an indirect semiconductor in the geometry of a plane-parallel plate must be more than $100 \mu\text{m}$ thick. Along with silicon, germanium is also an indirect semiconductor.

3.5.2

Generation of Electron–Hole Pairs

An important equation for the following considerations is the continuity equation. In its general form for a quantifiable variable, e.g. for the concentration of a certain type i of particle, it is

$$\frac{\partial n_i(x)}{\partial t} = G_i(x) - R_i(x) - \operatorname{div} j_i(x) \quad (3.53)$$

This relation expresses the simple consideration that the density n_i of particles of type i in a volume element at the location x increases with time when these particles are generated at a rate G_i , and that n_i decreases when particles at the location x are annihilated at the rate R_i or when they escape from the volume element at the location x because the particle current density j_i flowing away toward the right (to larger x) is greater than that flowing in from the left (from smaller x), expressed by $\operatorname{div} j_i > 0$.

In order to gain familiarity with the use of the continuity equation, we will apply this equation to a case for which we already know the solution, namely, to the photon current penetrating the semiconductor and being attenuated as a result of absorption. If an external photon current density j_γ is incident onto the semiconductor and no photons are created within the semiconductor ($G_\gamma = 0$), then in steady state, where no changes occur over time,

$$\frac{\partial n_\gamma}{\partial t} = -R_\gamma - \operatorname{div} j_\gamma = 0$$

With $R_\gamma = \alpha j_\gamma$ the expected proportionality between the rate of annihilation and the photon density (which moves at the velocity of light) is expressed. For light incident in the x direction this yields

$$\operatorname{div} j_\gamma = \frac{dj_\gamma}{dx} = -R_\gamma = -\alpha j_\gamma$$

Integrating, we find

$$j_\gamma(x) = j_\gamma(0) \exp(-\alpha x)$$

the absorption law with which we are already familiar. Here, $j_\gamma(0)$ is the photon current density penetrating the surface of the semiconductor. It is less than the incident photon current density $j_{\gamma,\text{in}}$ due to reflection at the surface

$$j_\gamma(0) = (1 - r) j_{\gamma,\text{in}}$$

With the transmittance of a body of thickness d

$$t = (1 - r) \exp(-\alpha d)$$

the total absorbed photon current becomes

$$j_{\gamma,\text{abs}} = (1 - r - t) j_{\gamma,\text{in}} = (1 - r)[1 - \exp(-\alpha d)] j_{\gamma,\text{in}} = \alpha j_{\gamma,\text{in}}$$

The absorptance $\alpha = (1 - r)[1 - \exp(-\alpha d)]$ applies for a plane-parallel plate of thickness d , if multiple reflection is neglected.

Since both a hole and an electron are produced for each absorbed (or annihilated) photon, they are both generated at the rate

$$G_h = G_e = R_\gamma = \alpha j_\gamma(x)$$

This assumes that an absorbed photon always generates only one electron and one hole. However, if the photon energy $\hbar\omega$ is at least twice as large as the band gap ε_G , one of the two charge carriers can then have a kinetic energy of $\varepsilon_{\text{kin}} \geq \varepsilon_G$. This charge carrier, whether an electron or a hole, is then able to knock an electron out of its chemical bond and thus produce another electron–hole pair by impact ionization. Owing to the symmetry of the band structure of the valence and conduction band, however, the energy of the photon is more or less equally divided among the electron and the hole. In this case multiple electron–hole pair generation by impact ionization would occur only for photon energies $\hbar\omega > 3\varepsilon_G$. This fundamental possibility requires a band structure for the semiconductor that ensures the conservation of energy and momentum. In real solar cells with band gaps ≥ 1 eV, however, it occurs only with very low probability and requires photon energies hardly found in the solar spectrum.

As an example for the production of one electron–hole pair per absorbed photon, we will calculate the generation rate of electrons and holes resulting from the absorption of the 300 K black-body radiation of the environment. As equilibrium rates they will be labeled with a superscript 0. Photons of different energies contribute according to the value of the absorption coefficient $\alpha(\hbar\omega)$:

$$\begin{aligned} G_e^0 = G_h^0 = R_\gamma^0 &= \int_0^\infty \alpha(\hbar\omega) dj_\gamma(\hbar\omega) \\ &= \frac{\Omega}{4\pi^3 \hbar^3 c^2} \int_0^\infty \frac{\alpha(\hbar\omega)(\hbar\omega)^2}{\exp\left(\frac{\hbar\omega}{kT_0}\right) - 1} d\hbar\omega \end{aligned} \quad (3.54)$$

c is the speed of propagation in the medium and is smaller than in vacuum by the refractive index n of the medium. This means that the photon current density per solid angle, and therefore the absorption rate, is proportional to n^2 . Since the photon current density does not change during the transition from the vacuum into the medium except by reflection, the larger photon current density per solid angle in the medium must be restricted to a range of solid angles into which photons flow from outside, which is smaller by a factor of n^2 than the solid angle range Ω_{vacuum} in the vacuum from which they come. This is the reason for Snell's law of refraction, according to which the photons make a smaller angle with respect to the surface normal in a medium than in vacuum. This is illustrated in Figure 3.19.

The generation rate of electron–hole pairs by photons incident from outside follows from Equation 3.54 by setting $\Omega = \pi$ and $c = c_0$, the vacuum velocity of light. These are, however, not all photons absorbed in the medium. Let us assume that the semiconductor is in thermal and chemical equilibrium with its environment and that no current flows. Then at every location within the semiconductor and in every solid angle element, exactly the same number of photons must be emitted as are absorbed. Since photon emission is generally isotropic, the radiation

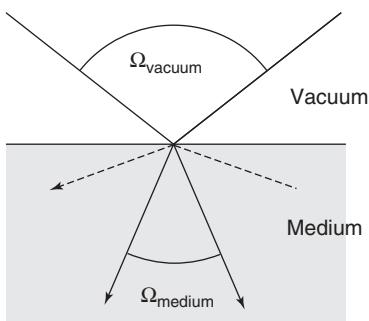


Figure 3.19 In accordance with the law of refraction at the interface between two media, only photons from a smaller solid angle in the medium with the larger refractive index are exchanged with the other medium. Photons for which the momentum lies in the solid angle range inaccessible from outside undergo total internal reflection (dashed light path).

in the medium is also isotropic with $\Omega_{\text{medium}} = 4\pi$. Furthermore, in this state of equilibrium the radiation in the semiconductor is also homogeneous, so that $j_\gamma^0 \neq j_\gamma^0(x)$. As we have just seen, certain solid angle ranges are not accessible to photons from outside. According to the principle of detailed balance the photons emitted inside the medium into the inaccessible solid angle range cannot leave the medium; they undergo total internal reflection. In equilibrium with the 300 K environment, these photons are absorbed and reemitted inside the medium and contribute to the generation and recombination rates, which then follow from Equation 3.54 by setting $\Omega = 4\pi$ and $c = c_0/n$, the speed of light in the medium.

For electrons and holes it is not meaningful to define a density per volume and per solid angle, because after a few collisions with phonons, i.e. in less than 10^{-13} s, an isotropic momentum distribution is established. The momentum of electrons and holes is then uniformly distributed over a solid angle range $\Omega = 4\pi$.

In addition to the generation of electrons and holes by the absorption of photons, they may also be generated by thermal transitions from the valence band to the conduction band in the absence of radiation, involving impurities. In thermal and chemical equilibrium with the environment, the nonradiative generation and recombination rates must also be balanced. We will come back to thermal generation of electron–hole pairs by the absorption of phonons in the next section, where we discuss nonradiative recombination.

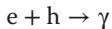
3.6 Recombination of Electrons and Holes

The different processes for the production of electrons and holes discussed in Section 3.5 also exist in reverse, that is in processes in which electrons and holes are annihilated in a reaction known as *recombination*. From the energy set free in this reaction, either photons or phonons are produced or both are produced simultaneously. In thermal and chemical equilibrium with the radiation in the environment, in which $n_e n_h^0 = n_e^2$, the rates of production and annihilation are exactly balanced for each of the different mechanisms. This is known as the *principle of detailed balance*.

3.6.1

Radiative Recombination, Emission of Photons

Radiative recombination, in which a hole reacts with an electron and produces a photon, is exactly the reverse of absorption (in the pure electron description it is the spontaneous transition of an electron from the conduction band to an unoccupied state in the valence band).



Since a free electron and a free hole must find each other, the rate of radiative recombination at which electrons and holes are annihilated and photons are generated increases with the concentration of electrons and the concentration of holes

$$G_\gamma = R_e = R_h = B n_e n_h \quad (3.55)$$

Here B is the coefficient for radiative recombination that remains to be determined. In equilibrium with the radiation in the environment, in which $n_e n_h = n_e^0 n_h^0 = n_i^2$,

$$G_\gamma^0 = R_e^0 = R_h^0 = R_\gamma^0 = G_e^0 = G_h^0 = B n_e^0 n_h^0 \quad (3.56)$$

It follows immediately from this equation that the product $n_e^0 n_h^0$ cannot depend on the doping and must therefore be constant, with a value of n_i^2 , as long as the absorption coefficient for band–band transitions and therefore the rate of absorption of photons R_γ^0 does not depend on the doping. Doping would have a significant effect on the absorption coefficient only if the density of impurities were in the same order of magnitude as the density of states in the bands.

The equality between G_γ^0 and R_γ^0 , which, in equilibrium, applies not only in integral form over the spectrum but also in each photon energy interval $d\hbar\omega$, is the result of Kirchhoff's law of radiation. Microscopically, this means that in thermal and chemical equilibrium with the radiation of the environment the rates of all generation processes are compensated individually by equally large rates of recombination processes between the same initial and final states: this is the principle of detailed balance. For the calculation of G_γ^0 we therefore take the expression for G_e^0 from Equation 3.54

$$G_\gamma^0 = R_\gamma^0 = G_e^0 = G_h^0 = \frac{\Omega}{4\pi^3 \hbar^3 c^2} \int_0^\infty \frac{\alpha(\hbar\omega)(\hbar\omega)^2}{\exp\left(\frac{\hbar\omega}{kT_0}\right) - 1} d\hbar\omega \quad (3.57)$$

Knowledge of the absorption coefficient $\alpha(\hbar\omega)$, determined from absorption measurements, and of n_i then allows the calculation of the coefficient B for radiative recombination in Equation 3.55.

The rate per volume at which photons in silicon in equilibrium with the 300 K black-body radiation of the environment are generated is given by

$$G_\gamma^0(\text{Si}) = 3 \times 10^5 \frac{\gamma}{\text{cm}^3 \text{s}} = B n_i^2$$

For silicon $n_i = 10^{10} \text{ cm}^{-3}$ and therefore

$$B(\text{Si}) = 3 \times 10^{-15} \frac{\gamma \text{ cm}^3}{\text{s}}$$

In a state of nonequilibrium between the electrons and holes in the semiconductor on one side and the photons in the 300 K radiation of the environment on the other side, resulting, e.g. from incident solar radiation or injection or extraction of electrons and holes with an electric current, $n_e n_h \neq n_i^2$. For extraction without additional production, in fact $n_e n_h < n_i^2$. The distribution of electrons and holes over the states is, however, still determined by Fermi distributions at the lattice temperature, with a mean kinetic energy of $\langle \epsilon_{\text{kin}} \rangle = (3/2)kT_0$. In a state of nonequilibrium with the 300 K radiation of the environment, photons are consequently emitted with the same energy distribution as in equilibrium with the 300 K radiation, but at the different rate

$$G_\gamma = G_\gamma^0 \frac{n_e n_h}{n_i^2} \quad (3.58)$$

With reference to the quasi-Fermi energies in Equation 3.29 we can now write

$$G_\gamma = G_\gamma^0 \exp\left(\frac{\epsilon_{\text{FC}} - \epsilon_{\text{FV}}}{kT}\right) \quad (3.59)$$

Owing to solar irradiation, the emission rate of photons increases in proportion to $n_e n_h$. Since the annihilated electron–hole pairs are lost for the electric current that a solar cell shall deliver, it is of interest to know the total current of photons emitted by a semiconductor of thickness d . This is not just the emission rate in Equation 3.58 multiplied by the thickness of the semiconductor, since some of the emitted photons are reabsorbed and generate electron–hole pairs again.

Although this would be a good exercise in the application of the continuity equation, the integration of the emission and absorption rates of the photons over the volume of the semiconductor and the consideration of total internal reflection at the surface is a rather confusing calculation. Instead, we will once again make use of the principle of detailed balance to determine the total loss of electron–hole pairs in a cell due to radiative recombination in a more elegant way.

In thermal and chemical equilibrium with the radiation of the environment, the photon current density $dj_\gamma^0(\hbar\omega)$ emitted from the surface of a body into the environment must be equal to the photon current density incident from the environment and absorbed by the body:

$$dj_{\gamma, \text{em}}^0 = dj_{\gamma, \text{abs}}^0 = a(\hbar\omega) dj_\gamma^0$$

In nonequilibrium of homogeneous electron and hole distributions, the emitted photon current density is larger by the same factor $n_e n_h / n_i^2$ by which the emission rate in Equation 3.58 is increased, so that

$$dj_{\gamma, \text{em}}(\hbar\omega) = a(\hbar\omega) \frac{n_e n_h}{n_i^2} dj_\gamma^0(\hbar\omega) \quad (3.60)$$

Here $a(\hbar\omega)$ is the absorptance, which, without considering multiple internal reflection, is $a(\hbar\omega) = [1 - r(\hbar\omega)] \{1 - \exp[-\alpha(\hbar\omega)d]\}$. The absorptance is always

$\leq 1 - r$ and for a thickness $d \gg 1/\alpha$, $\alpha = 1 - r$, independent of the thickness. This means that the emitted photon current density reaches a limit value for large thickness, which then no longer depends on the thickness. Although the rate of radiative recombination integrated over the volume increases linearly with the volume, for a large thickness, most of the emitted photons are reabsorbed and generate electron–hole pairs again. The loss of electron–hole pairs by radiative recombination is given only by the photons that are emitted through the surface of the semiconductor. The effective rate of radiative recombination is proportional to the surface area and does not depend on the volume, if the thickness is larger than the penetration depth of the photons.

Expressing Equation 3.60 in terms of the quasi-Fermi energies, we obtain

$$dj_{\gamma,\text{em}} = a(\hbar\omega) \exp \left[\frac{(\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}})}{kT} \right] \frac{\Omega}{4\pi^3 \hbar^3 c^2} \frac{(\hbar\omega)^2}{\exp(\hbar\omega/kT) - 1} d\hbar\omega \quad (3.61)$$

This result is based on the Boltzmann approximation of the Fermi distribution in deriving the relationship between the concentrations of electrons and holes and the Fermi energies in Equations 3.12 and 3.15. For this approximation, the Fermi energies had to be a few kT away from the band edges. The same restriction applies to the absorption coefficient in the radiative recombination rate (per volume) and to the absorptance in the emitted photon current (per surface area). The condition of a full valence band and an empty conduction band on which their derivation in Section 3.5 is based and for which the absorption coefficient is usually measured is no longer a good approximation if the Fermi energies come close to the band edges. Equation 3.61 is therefore only approximately true. An exact derivation is given in Section 3.7.

In place of Equation 3.61 the exact result is [4]

$$dj_{\gamma,\text{em}} = a(\hbar\omega) \frac{\Omega}{4\pi^3 \hbar^3 c^2} \frac{(\hbar\omega)^2}{\exp \left\{ \frac{\hbar\omega - (\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}})}{kT} \right\} - 1} d\hbar\omega \quad (3.62)$$

In this form, as the generalized Planck radiation law, it describes both the emission of thermal radiation for $(\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}}) = 0$ and the emission of luminescence radiation for $(\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}}) \neq 0$. The difference of the Fermi energies $(\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}})$ turns out to be the chemical potential μ_γ of the emitted photons. For the operation of solar cells, where $(\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}})$ is several kT smaller than ε_G , Equation 3.61 is, however, a very good approximation.

3.6.2

Nonradiative Recombination

With the recombination of an electron and a hole, the energy set free must always be taken up by other particles. For nonradiative recombination these are

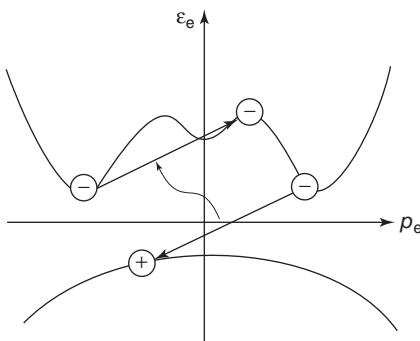


Figure 3.20 In Auger recombination, the energy set free by recombination of an electron–hole pair is absorbed by a free carrier (an electron in this figure) and is subsequently dissipated by generating phonons in collisions with the lattice.

primarily other electrons or holes (Auger recombination) or phonons (impurity recombination).

Auger Recombination

Auger recombination is the reverse of impact ionization, in which an electron or a hole with high kinetic energy knocks another electron out of its bond, thereby creating a free electron and a free hole. In the reverse process, the energy set free during recombination is transferred to an electron or a hole as kinetic energy that is subsequently lost to the lattice through collisions with phonons. Figure 3.20 illustrates the process of Auger recombination in which energy and momentum must initially be conserved.

If an electron absorbs the energy, two electrons and one hole participate in the reaction. The recombination rate is then

$$R_{\text{Aug, e}} = C_e n_e^2 n_h$$

and is large for strong n-doping.

If a hole absorbs the recombination energy, the recombination rate is

$$R_{\text{Aug, h}} = C_h n_e n_h^2$$

which is large in materials with strong p-doping. In general, the rate of Auger recombination is

$$R_{\text{Aug}} = n_e n_h (C_e n_e + C_h n_h) \quad (3.63)$$

In silicon the constants have the value

$$C_e(\text{Si}) \approx C_h(\text{Si}) \approx 1 \times 10^{-30} \text{ cm}^6 \text{ s}^{-1}$$

A pn solar cell requires doping. How strong the doping should be will be discussed later. For this structure, Auger recombination represents a practically unavoidable loss and largely determines the efficiency limits of the best silicon solar cells.

Recombination via Impurities

In equilibrium with 300 K radiation of the environment, for which $n_e^0 n_h^0 = n_i^2$ and the generation of electrons by photons is compensated by radiative recombination, nonradiative recombination processes are also compensated by the reverse

processes, i.e. nonradiative generation processes in which electrons are excited via states in the energy gap or, less frequently, as a result of the absorption of many phonons, even directly from the valence band to the conduction band.

In real solar cells, recombination via impurities is the predominant recombination process and will therefore be discussed in detail. Impurities that offer electron states with energies approximately in the middle of the forbidden zone play the most important role. They capture electrons and holes over a series of excited states, with successive dissipation of energy. Since the recombination energy can thus be imparted to the lattice in small portions through the production of individual phonons, the nonradiative transition of an electron from the conduction band to the valence band, which is tantamount to the capture of an electron and a hole by an impurity, can take place much more easily. In the following analysis, we will therefore neglect radiative recombination compared with nonradiative recombination. The rates of generation and recombination involving impurities and considered in the analysis, are shown in Figure 3.21. The analysis follows the calculations of Shockley, Read, and Hall [5].

As an example, we will examine a single type of acceptor-like impurities, having a concentration n_{imp} with a known electron energy ϵ_{imp} and a known capture cross section σ_e for electrons and σ_h for holes in a homogeneously excited semiconductor under open circuit conditions, i.e. without current flow. Under these conditions all electrons and holes must eventually recombine. During their lifetime they will be distributed homogeneously by diffusion. This allows to replace the actual and inhomogeneous generation rate of electron–hole pairs due to illumination by an average over the total volume of a semiconductor. Since for this consideration recombination via impurities is assumed to be dominant, no other recombination mechanism is considered.

The electron concentration n_e (in the conduction band) changes in time owing to the rate of generation $G_e = G_h = G = a(\hbar\omega) j_\gamma/d$ by the absorption of photons averaged over the thickness d and it changes by the rate $R_{e,\text{imp}}$ of capture into impurities occupied by holes and by the rate of thermal emission $G_{e,\text{imp}}$ from impurities occupied by electrons. The contributions of the 300 K radiation from the environment and radiative recombination are neglected. The continuity

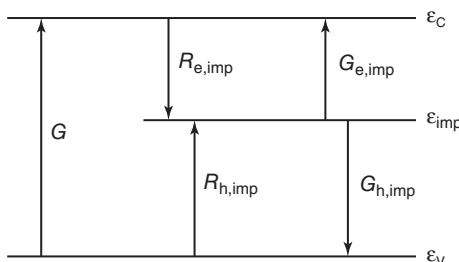


Figure 3.21 Electrons and holes produced by the absorption of photons with the generation rate G are captured by impurities with the rates $R_{e,\text{imp}}$ and $R_{h,\text{imp}}$ and they are emitted thermally with the rates $G_{e,\text{imp}}$ and $G_{h,\text{imp}}$ from the impurities back into the bands.

equations for the changes in the concentrations of electrons (in the conduction band) n_e , of holes (in the valence band) n_h and of occupied impurity states, i.e. of electrons trapped in impurities $n_{e,\text{imp}}$ as a function of the generation and recombination rates shown in Figure 3.21 are

$$\frac{\partial n_e}{\partial t} = G - R_{e,\text{imp}} + G_{e,\text{imp}} \quad (3.64)$$

$$\frac{\partial n_h}{\partial t} = G - R_{h,\text{imp}} + G_{h,\text{imp}} \quad (3.65)$$

$$\frac{\partial n_{e,\text{imp}}}{\partial t} = R_{e,\text{imp}} - R_{h,\text{imp}} - G_{e,\text{imp}} + G_{h,\text{imp}} \quad (3.66)$$

These are three equations for the three unknown concentrations n_e , n_h and $n_{e,\text{imp}}$. Unfortunately, these equations are not linearly independent, so that, e.g. the third equation follows from the difference of the first and the second equation. We obtain the missing third independent equation from the condition that charge neutrality of the semiconductor must prevail even when the different concentrations are changing.

$$e(n_{D^+} - n_e - n_A^- + n_h - n_{e,\text{imp}}) = 0 \quad (3.67)$$

In this equation we assume that the impurities are negatively charged when they are occupied by electrons, i.e. they are acceptor-type impurities. For donor-type impurities, positively charged when not occupied by electrons, $(-n_{e,\text{imp}})$ must be replaced by $(n_{\text{imp}} - n_{e,\text{imp}})$, where n_{imp} is the concentration of the impurities.

We now have to determine the various generation and recombination rates. The rate of recombination $R_{e,\text{imp}}$ at which electrons vanish from the conduction band follows from a purely kinetic argument. Whenever an electron that moves through the lattice at its thermal velocity v_e comes within a capture cross section σ_e of an impurity that is occupied by a hole, it will be captured. Impurities that are already occupied by an electron will not capture a second electron, because the energy of a second electron state is much higher and may lie in the energy range of the conduction band. The rate at which capture happens to all electrons is

$$R_{e,\text{imp}} = \sigma_e v_e n_e n_{h,\text{imp}} \quad (3.68)$$

where $n_{h,\text{imp}}$ is the density of impurities not occupied by an electron and thus occupied by a hole, which are the recombination partners for the electrons. The capture cross sections for electrons, σ_e , and for holes, σ_h , have values of the order of 10^{-15} cm^2 .

Similarly, the recombination rate for holes is given by

$$R_{h,\text{imp}} = \sigma_h v_h n_h n_{e,\text{imp}} \quad (3.69)$$

An impurity is likely to have different cross sections for electrons and holes. An unoccupied donor-type impurity is positively charged and consequently has a large capture cross section for electrons, while an unoccupied acceptor-type impurity is neutral, with a small capture cross section for electrons. Conversely,

the capture cross section for holes is large for an occupied, and therefore negative, acceptor impurity and is small for an occupied and neutral donor impurity.

In Equations 3.68 and 3.69 the occupation of impurity states, and as a result the concentration of electrons $n_{e,\text{imp}}$ or holes $n_{h,\text{imp}}$ in the impurities, is not known. If the energy ε_{imp} of the electrons in the impurities lies between the quasi-Fermi energies, the occupation of the impurities is not defined by either of the Fermi distributions but is rather determined by the kinetics.

For the generation rates of electrons or holes from the impurities into the bands, we know that they must be proportional to the concentration of electrons $n_{e,\text{imp}}$ or holes $n_{h,\text{imp}} = n_{\text{imp}} - n_{e,\text{imp}}$ in the impurities. The generation rate of electrons is

$$G_{e,\text{imp}} = \beta_e n_{e,\text{imp}} \quad (3.70)$$

In the same way, the generation rate of holes is

$$G_{h,\text{imp}} = \beta_h n_{h,\text{imp}} = \beta_h (n_{\text{imp}} - n_{e,\text{imp}}) \quad (3.71)$$

Here β_e and β_h are as yet unknown coefficients of emission. For their determination we make use of the principle of detailed balance in the dark, inherent in the thermal and chemical equilibrium with radiation from the environment. According to Equations 3.64, 3.68, and 3.70, in the absence of external excitation ($G=0$) and in a steady state

$$\frac{\partial n_e}{\partial t} = -\sigma_e v_e n_e (n_{\text{imp}} - n_{e,\text{imp}}) + \beta_e n_{e,\text{imp}} = 0 \quad (3.72)$$

From Equation 3.72 we can calculate the density of impurities occupied by electrons in the dark, which we already know is given by the equilibrium Fermi distribution,

$$n_{e,\text{imp}} = n_{\text{imp}} \frac{1}{\beta_e / (\sigma_e v_e n_e) + 1} = n_{\text{imp}} \frac{1}{\exp [(\varepsilon_{\text{imp}} - \varepsilon_F) / kT] + 1} \quad (3.73)$$

With $n_e = N_C \exp[-(\varepsilon_C - \varepsilon_F)/kT]$, we find for the coefficient of emission of electrons from the impurity into the conduction band

$$\beta_e = \sigma_e v_e N_C \exp \left(-\frac{\varepsilon_C - \varepsilon_{\text{imp}}}{kT} \right) \quad (3.74)$$

and similarly for the coefficient of emission of holes from the impurity into the valence band

$$\beta_h = \sigma_h v_h N_V \exp \left(-\frac{\varepsilon_{\text{imp}} - \varepsilon_V}{kT} \right) \quad (3.75)$$

The capture cross sections and the coefficients of emission are likely to depend on the energy distribution of the free charge carriers. Owing to the rapid thermalization, however, electrons and holes have the same energy and velocity distributions under illumination as in the dark. Capture cross sections and emission coefficients are therefore expected to have the same values both in the dark and under illumination.

We now have all the ingredients to solve Equations 3.64–3.67, which define the concentrations of the electrons, holes, and occupied (as well as unoccupied)

impurities as a function of the external rate of generation G of electrons and holes, and not only in the steady state. For non-steady-state processes this system of equations can, however, only be solved numerically.

For the steady-state case, we find from Equation 3.66 for the density of impurities occupied by electrons

$$n_{e,\text{imp}} = \frac{n_{\text{imp}} \left\{ \sigma_e v_e n_e + \sigma_h v_h N_V \exp \left[-(\varepsilon_{\text{imp}} - \varepsilon_V) / kT \right] \right\}}{\sigma_e v_e \left\{ n_e + N_C \exp \left[-(\varepsilon_C - \varepsilon_{\text{imp}}) / kT \right] \right\} + \sigma_h v_h \left\{ n_h + N_V \exp \left[-(\varepsilon_{\text{imp}} - \varepsilon_V) / kT \right] \right\}} \quad (3.76)$$

This result, which is not important by itself, inserted into Equation 3.64 gives

$$G = \frac{n_e n_h - n_i^2}{n_e + N_C \exp \left[-(\varepsilon_C - \varepsilon_{\text{imp}}) / kT \right] + \frac{n_h + N_V \exp \left[-(\varepsilon_{\text{imp}} - \varepsilon_V) / kT \right]}{n_{\text{imp}} \sigma_h v_h}} \quad (3.77)$$

The left side represents the rate of generation of both electrons and holes. For the steady-state case, then, on the right side we have the recombination rate, with the same value for both electrons and holes. Here $n_{\text{imp}} \sigma_h v_h$ is the rate per hole of hole capture if all impurities are occupied by electrons. Its reciprocal value $\tau_{h,\text{min}} = 1/(n_{\text{imp}} \sigma_h v_h)$ is the smallest mean lifetime of a hole in the valence band between generation and capture, since it implies that *all* impurities are occupied by electrons. Similarly, $\tau_{e,\text{min}} = 1/(n_{\text{imp}} \sigma_e v_e)$ is the mean lifetime of an electron in the conduction band when *all* impurities are occupied by holes, which is the smallest possible value of the electron lifetime. The minimal lifetimes defined in this way represent lower limits for the actual lifetimes when the impurities are only partially occupied.

Since separation of the Fermi energies is the main objective of a solar cell, it is worthwhile to express Equation 3.77 in terms of the Fermi energies. We use the relations

$$n_e = n_i \exp \left(-\frac{\varepsilon_i - \varepsilon_{FC}}{kT} \right) \quad (3.78)$$

$$n_h = n_i \exp \left(-\frac{\varepsilon_{FV} - \varepsilon_i}{kT} \right) \quad (3.79)$$

$$N_C \exp \left(-\frac{\varepsilon_C - \varepsilon_{\text{imp}}}{kT} \right) = n_i \exp \left(-\frac{\varepsilon_i - \varepsilon_{\text{imp}}}{kT} \right) \quad (3.80)$$

$$N_V \exp \left(-\frac{\varepsilon_{\text{imp}} - \varepsilon_V}{kT} \right) = n_i \exp \left(-\frac{\varepsilon_{\text{imp}} - \varepsilon_i}{kT} \right) \quad (3.81)$$

where ε_i is the value of the Fermi energy in an intrinsic semiconductor in the dark, and we find from Equation 3.77

$$G = \frac{n_i \left\{ \exp [(\varepsilon_{FC} - \varepsilon_{FV}) / kT] - 1 \right\}}{\tau_{h,\text{min}} \left\{ \exp [(\varepsilon_{FC} - \varepsilon_i) / kT] + \exp [(\varepsilon_{\text{imp}} - \varepsilon_i) / kT] \right\} + \tau_{e,\text{min}} \left\{ \exp [(\varepsilon_i - \varepsilon_{FV}) / kT] + \exp [(\varepsilon_i - \varepsilon_{\text{imp}}) / kT] \right\}} \quad (3.82)$$

For a good solar cell we want the difference in the Fermi energies, i.e. the numerator of Equation 3.82, to be large. For a given generation rate G , the denominator must then be large, too, requiring large minimal lifetimes. Conversely, a material produces a poor solar cell when the numerator and the denominator of Equation 3.82 are both small. If the minimum lifetimes of the holes $\tau_{h,\min}$ and electrons $\tau_{e,\min}$ are equal, the denominator will be minimal if, on the one hand, the impurity level ϵ_{imp} coincides with the position of the Fermi energy in the intrinsic state ($\epsilon_{\text{imp}} = \epsilon_i$) and, on the other hand, if the Fermi energies ϵ_{FC} and ϵ_{FV} are symmetrical about the intrinsic Fermi energy ϵ_i , which is approximately in the middle of the energy gap ($\epsilon_{\text{FC}} - \epsilon_i = \epsilon_i - \epsilon_{\text{FV}}$). In a homogeneous material the latter condition is fulfilled only if it is intrinsic. Doping, regardless of whether n-type or p-type, will increase the difference in the Fermi energies for a given generation rate compared with an intrinsic material.

For equal minimal lifetimes of electrons and holes the difference in the Fermi energies $\epsilon_{\text{FC}} - \epsilon_{\text{FV}}$ for a constant generation rate G is determined from Equation 3.82 and is shown in Figure 3.22 as a function of the impurity level ϵ_{imp} . As expected, $\epsilon_{\text{FC}} - \epsilon_{\text{FV}}$ is small when the electrons in the impurities have energies ϵ_{imp} in the middle of the energy gap. The damaging influence of the impurities is also greater in weakly doped material, where the Fermi energies are more nearly symmetrical about the middle of the energy gap than in strongly doped material.

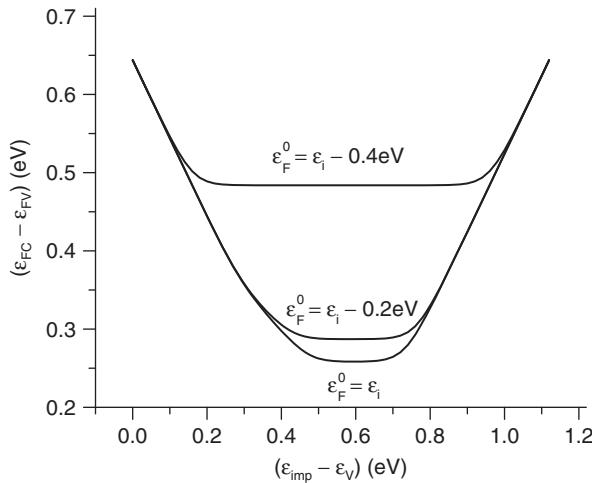


Figure 3.22 Difference of the Fermi energies as a function of the impurity energy ϵ_{imp} for a constant rate of generation $G = 10^{17} \text{ cm}^{-3}\text{s}^{-1}$ and for different doping levels in a material with $\epsilon_G = 1.12 \text{ eV}$ and densities of states $N_C = N_V = 10^{19} \text{ cm}^{-3}$. The doping concentration can be inferred from the difference between the Fermi energy in the dark ϵ_F^0 and the intrinsic

Fermi energy ϵ_i from Equation 3.78 or 3.79. The impurities have a concentration of $n_{\text{imp}} = 10^{14} \text{ cm}^{-3}$ with equal capture cross sections for electrons and holes of $\sigma = 10^{-15} \text{ cm}^2$, the velocity of electrons and holes is $v = 10^5 \text{ m s}^{-1}$, resulting in a minimal lifetime of $\tau = 10^{-6} \text{ s}$. The asymmetry is caused by the acceptor character of the impurities.

Whether the material is p-doped or n-doped, however, is irrelevant. The type and level of doping can be seen from the position of the Fermi energy ε_F^0 in the dark state. The greater the separation $|\varepsilon_F^0 - \varepsilon_i|$ of the equilibrium Fermi energy from the intrinsic Fermi energy, i.e. the closer the equilibrium Fermi energy is to one of the band edges, the stronger the doping of the material. For n-doped material $\varepsilon_F^0 - \varepsilon_i > 0$ and for p-doped material $\varepsilon_F^0 - \varepsilon_i < 0$.

The concentrations of additional electrons and holes due to the external generation rate G can be determined analytically from Equation 3.77 for small impurity concentrations n_{imp} and weak excitation. Furthermore, we will divide the charge carrier densities into those in the dark state and those produced by the external generation, $n_e = n_e^0 + \Delta n_e$ and $n_h = n_h^0 + \Delta n_h$, respectively. From the charge neutrality condition in Equation 3.67 it follows that for a small impurity concentration $n_{\text{imp}} \ll \Delta n_e$, Δn_h the additional concentrations of electrons and holes are equal, $\Delta n_e = \Delta n_h = \Delta n$. The condition of weak excitation is fulfilled, if $\Delta n \ll n_e^0 + n_h^0$. With these restrictions, it follows from Equation 3.77 for capture cross sections not too different for electrons and holes that

$$\Delta n = G \left\{ \tau_{h,\min} \frac{n_e^0 + N_C \exp [-(\varepsilon_C - \varepsilon_{\text{imp}})/kT]}{n_e^0 + n_h^0} + \tau_{e,\min} \frac{n_h^0 + N_V \exp [-(\varepsilon_{\text{imp}} - \varepsilon_V)/kT]}{n_e^0 + n_h^0} \right\} \quad (3.83)$$

Figure 3.23 shows the ratio $\Delta n/G$, which is also the mean lifetime τ of electrons and holes, as a function of ε_{imp} in p-doped silicon. $\tau_{e,\min} = \tau_{h,\min} = 10^{-6}$ s was

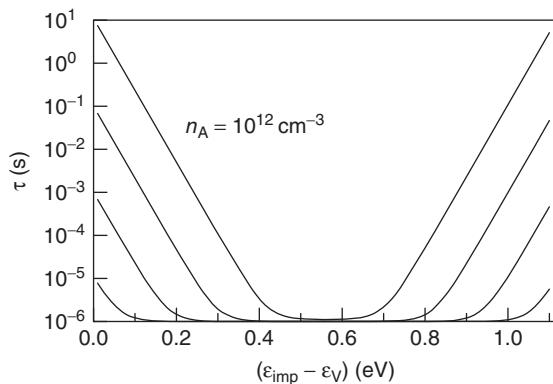


Figure 3.23 Lifetime τ of electrons and holes in p-doped silicon limited by impurity recombination as a function of the impurity energy ε_{imp} , measured from the upper edge of the valence band. With each curve from the inside toward the outside, the acceptor density increases by a factor of 100 from $n_A = 10^{12} \text{ cm}^{-3}$ to $n_A = 10^{18} \text{ cm}^{-3}$.

The generation rate is $G = 10^{17} \text{ cm}^{-3} \text{ s}^{-1}$, the impurities have a concentration of $n^{\text{imp}} = 10^{14} \text{ cm}^{-3}$ with equal capture cross sections for electrons and holes of $\sigma = 10^{-15} \text{ cm}^2$ and the velocity of electrons and holes is $v = 10^5 \text{ m s}^{-1}$, resulting in a minimal lifetime of $\tau_{\min} = 10^{-6} \text{ s}$.

used for the calculation, resulting from an impurity density $n_{\text{imp}} = 10^{14} \text{ cm}^{-3}$, a capture cross section for electrons and holes of $\sigma = 10^{-15} \text{ cm}^2$ and a velocity of $v = 10^5 \text{ m s}^{-1}$. We see in Figure 3.23 that the lifetime τ becomes shorter and the recombination accordingly more effective as the energy level of the impurity ε_{imp} approaches the intrinsic Fermi energy, which is in the middle of the forbidden zone, if the effective masses of electrons and holes are equal.

When the impurity level is close to the conduction band, the impurities are less effective, because most of the time they are not occupied by electrons. Captured electrons are much more frequently emitted back to the conduction band than annihilated by recombination with a hole captured from the valence band. When the impurity level is in the vicinity of the valence band, impurities are most of the time occupied by electrons, and captured holes are emitted back to the valence band before they can recombine by electron capture. In the middle of the forbidden zone, occupation by electrons and holes is equally probable. Both electrons and holes find enough free impurity states to be captured and reemission is less probable. We also recognize the extreme demands on the concentration of impurities with energies in the middle of the gap and therefore on the purity of the material in order to obtain lifetimes of a few milliseconds, as is achieved in good silicon solar cells.

A large difference of the Fermi energies and a large lifetime of electrons and holes are both indications of a good-quality solar cell material. A closer look at Figures 3.22 and 3.23 seems to reveal an inconsistency. While increasing doping concentrations lead to an increasing difference of the Fermi energies in Figure 3.22, they result in decreasing lifetimes in Figure 3.23. The explanation is simple. In a p-type semiconductor under low excitation ($\Delta n \ll n_h^0$) the difference of the Fermi energies is $\Delta\varepsilon_F = kT \ln(n_e n_h^0 / n_i^2)$. As long as the lifetime τ and with it the electron concentration $n_e = \Delta n = G\tau$ decreases less strongly than the doping concentration and with it n_h^0 increases, the difference of the Fermi energies will rise with the doping concentration, although the lifetime decreases. The question of whether the splitting of the Fermi energies or the lifetime is the more important quantity for a solar cell will be addressed later.

If, as in a real material, impurity states with different energies ε_{imp} and different capture cross sections are present, then contributions from each of them must be added to the recombination rate R_{imp} , the generation rate G_{imp} , and the charge neutrality condition in Equations 3.64–3.67. But as we have seen in Figures 3.22 and 3.23, impurity states in the middle of the band gap dominate the recombination properties.

Surface Recombination

States with energies in the forbidden gap, which are very effective mediators for nonradiative recombination are found in large concentrations on the surface of a material where neighboring atoms are no longer available to the lattice atoms and where impurity molecules (e.g. O₂ or H₂O) are adsorbed. These so-called surface states often have a continuous distribution over energy in the forbidden gap, as

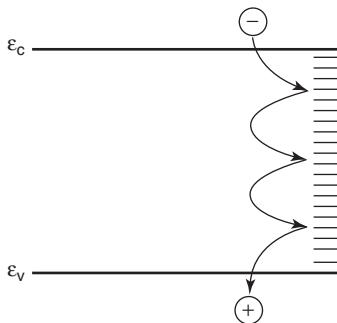


Figure 3.24 Recombination via surface states continuously distributed over energy in the energy gap of a semiconductor.

illustrated in Figure 3.24. The effect of the surface states on nonradiative recombination can be treated in the same way as the effect of the bulk states before. If the concentration of the surface states per energy and their cross section for capture of electrons and holes are known as a function of energy, the total recombination rate is the sum over the contributions from each energy interval.

In a simpler treatment, all the contributions are lumped into a single expression. For electrons, for example, the recombination rate via surface states is

$$R_{s,e} = \sigma_{s,e} v_e n_{s,h} n_e \quad (3.84)$$

Here $R_{s,e}$ is the recombination rate of electrons per area, $n_{s,h}$ is the density of surface states per area occupied by a hole, and n_e is the concentration of electrons at the surface. The product $\sigma_{s,e} v_e n_{s,h}$ has the dimension of velocity and is known as the *surface recombination velocity* $v_{R,e}$ of the electrons. It is characteristic for the surface quality, but it is not a material property alone, since it depends on the occupation of the surface states and thus on the excitation conditions,

$$R_{s,e} = v_{R,e} n_e \quad (3.85)$$

As with impurities in the bulk of a semiconductor, surface states are most effective with respect to recombination when their electron energies lie in the middle of the forbidden gap and in-between the Fermi energies. For this reason nonradiative recombination (at the surface) is reduced, if both Fermi energies are close to either the conduction band or the valence band at the surface, as is the case for heavy doping. Much less is known about the chemical nature of the surface states and their electron energies compared to impurities in the bulk of a semiconductor. In most cases, we must restrict their characterization to the surface recombination velocity.

Poor surfaces with high surface recombination velocities of $10^5 - 10^6 \text{ cm s}^{-1}$ are those freely exposed to air, which adsorb H_2O and O_2 or react chemically with these substances.

Metallic surfaces required as contacts for the supply of electric current are especially problematic. As can be seen in Figure 3.25, the forbidden gap of the semiconductor is adjacent to the continuous distribution of states in the conduction band of the metal. The loss of electron-hole pairs at a metal contact is not really a recombination process, in which an electron reacts with a hole by making a transition

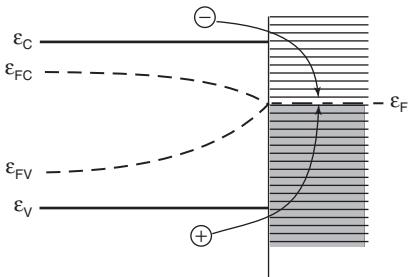


Figure 3.25 The continuous distribution of a high density of states around the Fermi energy in a metal results in a high interface recombination velocity at a semiconductor–metal contact.

into an empty state that is the hole as for instance an empty defect state. At a metal contact, electrons and holes rather move into the metal independently. In the metal, the electrons find plenty of unoccupied states above the Fermi-energy and holes find empty hole states (occupied electron states) below. Then electrons and holes independently thermalize towards the Fermi-energy. The surface recombination velocity at a metallic contact can thus be regarded as infinitely large. There is, however, a limit to the surface recombination rate, since it must be supplied by the current of electrons and holes reaching the interface with the metal. If the electrons are the minority carriers, their current density is limiting the recombination rate. At the metal it is $j_e = v_e n_e$. The electron velocity v_e is limited by the thermal velocity $v_e \leq v_{e,\text{th}} = \sqrt{3kT/m_e}$, which is about 10^7 cm/s at 300 K. As the recombination rate is usually written as $R_e = v_R n_e$, the limit of the recombination velocity v_R is identified with the limit of the transport velocity v_e . In the dark, the resulting large recombination rate is compensated by an equally large generation rate. Any additional generation rate at the semiconductor/metal interface due to light absorption is in most cases small compared with the dark rate. As a result, even with additional generation, the concentrations of electrons and holes do hardly differ from their equilibrium values in the dark state for which $n_e^0 n_h^0 = n_i^2$. In the same way, the Fermi energies are not separated at a metal contact. With the large recombination probability at a metal contact, large losses of charge carriers by surface recombination can only be avoided by allowing only one type of carrier to reach a metal contact.

Surface recombination at a free surface can be reduced by two different strategies. In both strategies, the semiconductor surface is covered by an insulator to prevent charge carriers from reaching the interface with the environment, which would always show high surface recombination. One way to reduce the recombination at the interface with the insulator is based on a high charge density incorporated in the insulator. This charge repels one type of carrier and reduces its conductivity near the interface. The surface recombination rate is then limited by the small possible current of this type of carrier, even if the density of surface states at the semiconductor-insulator interface is large. This happens at the interface of silicon with silicon nitride and aluminum oxide.

The other strategy employs a material, which by its lattice structure avoids broken bonds and thus surface states at the semiconductor-insulator interface.

With silicon, interfaces with a low interface state density can be prepared by carefully controlled oxidation under clean room conditions. Surface state densities of $D_s < 10^{10} (\text{cm}^2 \text{ eV})^{-1}$ can be obtained at Si/SiO₂ interfaces, with recombination velocities of $v_R \leq 10 \text{ cm s}^{-1}$.

Such perfect surface passivation is possible only for Si/SiO₂ and for the combination of a few III–V compounds having the same lattice constants. In spite of this, contact with another semiconductor in the form of a cover layer usually leads to lower surface recombination velocities than for a free surface or for the contact with a metal. In order for the recombination on the surface of the cover layer to remain insignificant, no electron–hole pairs should be generated within this layer. Semiconductors with a large band gap ($\epsilon_G > 3 \text{ eV}$) are therefore chosen for passivating cover layers. These are transparent in the visible range and are also known as *window layers*.

3.6.3 Lifetimes

Because of recombination, a charge carrier once produced does not “survive” arbitrarily long in its band. If electrons and holes are produced by the absorption of light and the light source is suddenly switched off, the charge carriers then vanish after a mean lifetime τ . Once again, the continuity equation tells us how:

$$\frac{\partial n_e}{\partial t} = G_e - R_e - \operatorname{div} j_e$$

We will restrict ourselves here to a spatially homogeneous system ($\operatorname{div} j_e = 0$) and divide the concentrations and rates into what is present in equilibrium, with $n_e^0 n_h^0 = n_i^2$, and what is caused by deviations from equilibrium. This division of the recombination rate, e.g. for electrons, is meaningful only if it changes in proportion to the electron density n_e , that is, when the density of the recombination partners, the holes, is constant:

$$\frac{\partial(n_e^0 + \Delta n_e)}{\partial t} = G_e^0 + \Delta G_e - (R_e^0 + \Delta R_e) \quad (3.86)$$

Since $\partial n_e^0 / \partial t = G_e^0 - R_e^0 = 0$ characterizes the equilibrium dark state, for the additional concentration of electrons

$$\frac{\partial \Delta n_e}{\partial t} = \Delta G_e - \Delta R_e$$

On suddenly switching off ΔG_e , we have for radiative recombination

$$\frac{\partial \Delta n_e}{\partial t} = -\Delta R_e = -B n_h \Delta n_e \quad (3.87)$$

This equation assumes that the generation of electrons and holes can be switched off, because it is caused entirely by photons incident from outside. It does not consider the reabsorption of the photons emitted by radiative recombination. The equation is therefore valid only for small semiconductor thicknesses. If we also

consider the reabsorption of photons produced in thick bodies, the electron density falls off more slowly by radiative recombination as the thickness of the body increases. The effective rate of radiative recombination is only caused by the photons leaving the semiconductor through the surface and is thus surface related in contrast to nonradiative recombination, which is volume related. For semiconductors thicker than the penetration depth of the photons the effective rate of radiative recombination saturates and becomes independent of the thickness. From then on the radiative lifetime increases with the thickness.

Equation 3.87 can be integrated easily for the electrons in a heavily doped p-conductor in which, because of the strong p-doping, the density of holes $n_h = n_h^0 + \Delta n_h \approx n_h^0$ is nearly independent of illumination. We refer to this case, in which the additional concentrations are small relative to the majority carrier concentration in the dark state, as “weak excitation” or “weak injection”. For solar cells in unfocused solar radiation, this condition is fulfilled. Equation 3.87 then yields

$$\Delta n_e(t) = \Delta n_e(0) \exp(-t/\tau_{e,\text{rad}}) \quad (3.88)$$

The characteristic time $\tau_{e,\text{rad}} = 1/(B n_h^0)$ is called the *lifetime* (of the minority charge carriers, here electrons) for radiative recombination.

In p-type silicon, with a typical hole density of $n_h^0 = 10^{16} \text{ cm}^{-3}$, the lifetime for electrons with respect to radiative recombination is

$$\tau_{e,\text{rad}} = \frac{1}{3 \times 10^{-15} \text{ cm}^3 \text{ s}^{-1} \times 10^{16} \text{ cm}^{-3}} = 0.03 \text{ s}$$

Making use of the lifetime, we can write the recombination rate in general as

$$\Delta R_e = \frac{\Delta n_e}{\tau_e}$$

Suddenly switching on ΔG_e starting from a state of equilibrium, we obtain

$$\Delta n_e(t) = \Delta G_e \tau_e \left\{ 1 - \exp\left(-\frac{t}{\tau_e}\right) \right\} \quad (3.89)$$

The additional steady-state concentration of electrons resulting from an additional generation rate by light absorption is

$$\Delta n_e = \Delta G_e \tau_e \quad (3.90)$$

For impurity recombination in a steady state, once again $\Delta R_e = \Delta G_e = \Delta n_e / \tau_e$ and thus

$$\tau_e = \frac{1}{\sigma_e v_e n_{h,\text{imp}}} \quad (3.91)$$

Figure 3.23 shows this lifetime as a function of the position of the impurity level.

If the density of the recombination partners varies with time, the lifetime also changes with time. We then speak of the momentary lifetime. Everything said here for the electrons in a p-conductor applies equally for the holes in a n-conductor and in general for the respective minority charge carriers.

The density of the free charge carriers can be monitored by the microwave absorption/reflection they cause. From the decay following the switching off of

the illumination, the lifetime is experimentally determined, which is an indicator of the quality of the semiconductor material.

As a result of transitions between the valence band and the conduction band, electrons and holes are created in pairs, so that $\Delta G_e = \Delta G_h$. If only donors or acceptors are present, completely ionized at room temperature, and otherwise there are only negligibly few impurities, the conservation of charge in Equation 3.67 then requires that $\Delta n_e = \Delta n_h$ and therefore $\tau_e = \tau_h$ as well. This is a surprising result in view of the large differences in the densities of electrons and holes in a doped semiconductor.

The different recombination processes take place in parallel to each other in a semiconductor. The total recombination rate is the sum of the rates for the different recombination mechanisms. If each of them is characterized by a lifetime τ_i , which the electrons or holes would have if no other recombination process were present, then the total lifetime follows from

$$R_{\text{tot}} = \frac{\Delta n_e}{\tau_{e,\text{tot}}} = \sum_i R_i = \sum_i \frac{\Delta n_e}{\tau_i}$$

to be

$$\frac{1}{\tau_{\text{tot}}} = \sum_i \frac{1}{\tau_i}$$

If no electron–hole pairs are extracted, i.e. at open circuit, all charge carriers produced must eventually recombine, and $\Delta G_e = \Delta R_e = \Delta n_e / \tau_{e,\text{tot}}$. The additional concentration of electrons Δn_e in steady state is proportional to their lifetime.

The upper limit for the lifetime of electrons is $\tau_{e,\text{rad}}$ if all recombination mechanisms are avoided, except for radiative recombination, which is unavoidable. Since the maximum obtainable concentrations of electrons and holes occur if there is only radiative recombination, the knowledge of radiative recombination and of the emitted photon currents is of great importance for the determination of maximum efficiencies.

3.7

Light Emission by Semiconductors

Light is emitted by radiative recombination of electron–hole pairs. This light emission can be used as a diagnostic tool for the quality of the semiconductor as explained in section 9.3. The generalized Planck equation in (3.62) describes the emission of light by any material as a function of its absorptance and of the chemical potential of its electron–hole pairs and these quantities can be deduced from the emitted light. The generalized Planck relation is so important that a rigorous derivation shall confirm its validity.

3.7.1

Transition Rates and Absorption Coefficient

The generalized Planck law for the emission of photons is derived in the same way as Planck's original formula, except that the body absorbing or emitting radiation contains electron distributions in separate energy ranges, which require different Fermi functions. This has consequences for the absorption coefficient and the probability of photon emission.

We analyze the absorption and emission rates per volume of photons by a semiconductor for radiative transitions between states in energy ranges of width $d\varepsilon$ at energies ε_1 in the valence band and ε_2 in the conduction band involving photons with energies in the energy range $\hbar\omega \dots \hbar\omega + d\hbar\omega$. The states are occupied according to two different Fermi functions $f(\varepsilon_1)$ and $f(\varepsilon_2)$. Other transitions between states with the same energy difference contributing to the interaction with photons of the same energy $\hbar\omega$ are disregarded for the moment.

The rate of upward transitions from states at ε_1 to states at ε_2 induced by an incident photon current density dj_γ is

$$dr_{\text{up}}(\hbar\omega) = |M|^2 D_{12} f(\varepsilon_1) [1 - f(\varepsilon_2)] \frac{d j_\gamma(\hbar\omega)}{d\hbar\omega} d\hbar\omega \quad (3.92)$$

where M contains the matrix element for the transition and D_{12} is the combined density of states between which the transitions occur.

The rate of stimulated emission of photons is likewise

$$dr_{\text{stim}}(\hbar\omega) = |M|^2 D_{12} [1 - f(\varepsilon_1)] f(\varepsilon_2) \frac{d j_\gamma(\hbar\omega)}{d\hbar\omega} d\hbar\omega \quad (3.93)$$

and the rate of spontaneous emission finally is

$$dr_{\text{spont}}(\hbar\omega) = |M|^2 \frac{c_0}{n} D_\gamma(\hbar\omega) D_{12} [1 - f(\varepsilon_1)] f(\varepsilon_2) d\hbar\omega \quad (3.94)$$

where the density of states for photons in the solid angle Ω in a medium with refractive index n is

$$D_\gamma(\hbar\omega) = \frac{\Omega n^3}{4\pi^3 \hbar^3 c_0^3} (\hbar\omega)^2 \quad (3.95)$$

and c_0 is the vacuum velocity of light.

Stimulated emission is a process by which a photon is duplicated, producing one additional photon in exactly the same state as the incident photon initiating the transition. These photons, therefore, cannot be distinguished from nonabsorbed photons. The net absorption rate at which photons disappear and which defines the absorption coefficient $\alpha_{12}(\hbar\omega)$ then is

$$\begin{aligned} dr_{\text{abs}}(\hbar\omega) &= dr_{\text{up}}(\hbar\omega) - dr_{\text{stim}}(\hbar\omega) \\ &= |M|^2 D_{12} [f(\varepsilon_1) - f(\varepsilon_2)] d j_\gamma(\hbar\omega) = \alpha_{12}(\hbar\omega) d j_\gamma(\hbar\omega) \end{aligned} \quad (3.96)$$

resulting in an absorption coefficient

$$\alpha_{12}(\hbar\omega) = |M|^2 D_{12} [f(\varepsilon_1) - f(\varepsilon_2)] \quad (3.97)$$

This relation is different from the absorption coefficient in Section 3.5.1, where an occupied valence band ($f(\varepsilon_1) = 1$) and an empty conduction band ($f(\varepsilon_2) = 0$) had been assumed. In contrast, the absorption coefficient in (3.97) becomes zero for $f(\varepsilon_1) = f(\varepsilon_2)$. This happens, when the Fermi-energies of the two Fermi-functions differ by the energy difference $\varepsilon_2 - \varepsilon_1 = \hbar\omega$. Note that this is valid independent of the position of the Fermi-energies with respect to ε_1 and ε_2 . As a consequence, the absorption coefficient is then zero for transitions between all pairs of states with an energy difference $\varepsilon_2 - \varepsilon_1 = \hbar\omega$. At this photon energy, the material becomes transparent. A process leading to this state is called bleaching. For $f(\varepsilon_1) < f(\varepsilon_2)$ the absorption coefficient in (3.97) even becomes negative leading to amplification of incident light as seen in a laser. $f(\varepsilon_1) < f(\varepsilon_2)$ is therefore called the lasing condition. The important relation (3.97) allows us to replace the factor M and the combined density of states D_{12} in the spontaneous emission rate by the absorption coefficient $\alpha_{12}(\hbar\omega)$ which makes this treatment applicable to real materials for which only the absorption coefficient is known:

$$dr_{\text{spont}}(\hbar\omega) = \alpha_{12}(\hbar\omega) \frac{c_0}{n} D_\gamma(\hbar\omega) \frac{[1-f(\varepsilon_1)]f(\varepsilon_2)}{f(\varepsilon_1)-f(\varepsilon_2)} d\hbar\omega \quad (3.98)$$

With the Fermi functions for the states at energy ε_1 in the valence band for which the Fermi energy is ε_{FV} and for ε_2 in the conduction band for which the Fermi energy is ε_{FC} :

$$f(\varepsilon_1) = \frac{1}{\exp[(\varepsilon_1 - \varepsilon_{\text{FV}})/kT] + 1} \quad \text{and} \quad f(\varepsilon_2) = \frac{1}{\exp[(\varepsilon_2 - \varepsilon_{\text{FC}})/kT] + 1}$$

and remembering that the energy difference $\varepsilon_2 - \varepsilon_1$ over which the transitions occur is equal to the photon energy $\hbar\omega$, Equation 3.98 becomes

$$dr_{\text{spont}}(\hbar\omega) = \alpha(\hbar\omega) \frac{c_0}{n} D_\gamma(\hbar\omega) \frac{d\hbar\omega}{\exp\{[\hbar\omega - (\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}})]/kT\} - 1} \quad (3.99)$$

Since Equation 3.99 does not depend on the energies ε_1 and ε_2 explicitly, but only on their difference $\hbar\omega$, each pair of states with energies ε_j in the conduction band and ε_i in the valence band with the same energy difference $\varepsilon_j - \varepsilon_i = \hbar\omega$ contributes to Equation 3.99 in the same way. All possible transitions between valence band and conduction band with the photon energy $\hbar\omega$ are accounted for and are contained in Equation 3.99, if $\alpha(\hbar\omega)$ now is the absorption coefficient for all these transitions.¹⁾ Note that the denominator in (3.99) becomes negative for $f(\varepsilon_1) < f(\varepsilon_2)$, when the absorption coefficient is also negative. The spontaneous emission rate in (3.99) remains positive as it must be.

A problem is that the emission rate of the photons cannot be observed. What can be observed and measured is the photon current emitted through a surface. To find the emitted photon current, we have to integrate the difference between photon emission and absorption rates over the thickness of the semiconductor

1) The spontaneous emission rate in (3.99) was derived for direct transitions. It is valid for indirect transitions as well, the derivation is a little more complicated because of the participation of phonons [6].

$$\operatorname{div} \mathbf{d}j_\gamma(x) = \mathbf{d}r_{\text{spont}}(x) - \alpha \mathbf{d}j_\gamma(x) \quad (3.100)$$

We will treat only the simple case of a homogeneously excited semiconductor, in which $\mathbf{d}r_{\text{spont}}$ is constant. Far from the surface in an infinitely thick semiconductor, emission and absorption rates balance each other and the photon current is constant for $x \rightarrow \infty$. This equilibrium photon current inside the semiconductor is, from Equation 3.100,

$$\mathbf{d}j_\gamma(\infty) = \frac{\mathbf{d}r_{\text{spont}}}{\alpha} \quad (3.101)$$

Integration of Equation 3.100 over a homogeneous semiconductor of finite thickness d yields

$$\mathbf{d}j_\gamma(\hbar\omega) = a(\hbar\omega) \frac{\Omega}{4\pi^3 \hbar^3 c_0^2} \frac{(\hbar\omega)^2 d\hbar\omega}{\exp\{[\hbar\omega - (\epsilon_{\text{FC}} - \epsilon_{\text{FV}})]/kT\} - 1} \quad (3.102)$$

where the integration gives

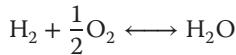
$$a(\hbar\omega) = [1 - r(\hbar\omega)] [1 - \exp(-\alpha d)] \quad (3.103)$$

for the absorptance of a homogeneous semiconductor of thickness d equal to its emittance. Equation 3.102 is the photon current emitted into the solid angle Ω outside the semiconductor after accounting for reflection at the surface. To determine the photon current density emitted by a real solar cell, the difference of the Fermi energies in Equation 3.99 must be known as a function of position to be able to perform the integration of Equation 3.100.

Equation 3.102 is a generalization of Kirchhoff's and Planck's laws and is valid for materials that are neither black nor have a single Fermi distribution over all states. It reduces to Planck's original emission law for a thermal emitter, in which all electrons belong to a single Fermi distribution, for which $\epsilon_{\text{FC}} - \epsilon_{\text{FV}} = 0$. The difference in the Fermi energies $\mu_{\text{eh}} = \epsilon_{\text{FC}} - \epsilon_{\text{FV}}$ is the free energy per electron–hole pair, also called the *chemical potential* of electron–hole pairs. It is free of entropy and we may therefore hope to transfer it into electrical energy without loss. Electron–hole pairs that recombine generate photons, which are emitted. These photons carry the free energy of the electron–hole pairs, and $\mu_\gamma = \mu_{\text{eh}} = \epsilon_{\text{FC}} - \epsilon_{\text{FV}}$ is recognized as the chemical potential of the photons.

It is often argued that the chemical potential of the photons is $\mu_\gamma = 0$ by nature [7], because their number n_γ is not conserved. According to this argument a minimum of the free energy F could only exist if in $dF = \dots + \mu_\gamma dn_\gamma + \dots = 0$ the chemical potential $\mu_\gamma = 0$, since nonconservation of n_γ would allow $dn_\gamma \neq 0$. This argument can hardly be correct. Photons carry other quantities like energy, momentum, and angular momentum, which are strictly conserved. Furthermore, photons do not react with each other and a photon cannot split into two by itself. Changes in the photon number are only possible in reactions with matter. The involvement of other particles, electrons or phonons, ensures the conservation of energy, momentum, and angular momentum when the photon number changes. The annihilation of a photon by the generation of an electron–hole pair with conservation of energy and momentum, discussed in this chapter, is an example. We

know other cases where the number of particles is not conserved without resulting in a zero chemical potential. It is rather characteristic for chemical reactions in general that the number of the particles involved is not conserved. As an example, we look at the reaction of hydrogen with oxygen to form water



Equilibrium exists if

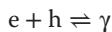
$$dF = \dots + \mu_{\text{H}_2} dN_{\text{H}_2} + \mu_{\text{O}_2} dN_{\text{O}_2} + \mu_{\text{H}_2\text{O}} dN_{\text{H}_2\text{O}} + \dots = 0$$

With the changes in the particle numbers in the reaction $dN_{\text{H}_2} = 2dN_{\text{O}_2} = -dN_{\text{H}_2\text{O}}$ we find the equilibrium condition

$$\mu_{\text{H}_2} + \frac{1}{2}\mu_{\text{O}_2} = \mu_{\text{H}_2\text{O}}$$

We conclude that nonconservation of the particle numbers does not lead to zero chemical potentials if a well-defined relation among the particle numbers exists as in every chemical reaction.

The absorption and emission processes of photons are chemical reactions as well, in which electron–hole pairs are generated or annihilated. The chemical reaction is



By the same argument as for the hydrogen–oxygen reaction, equilibrium between the electron–hole pairs and the emitted photons exists if

$$\eta_e + \eta_h = \mu_e + \mu_h = \varepsilon_{\text{FC}} - \varepsilon_{\text{FV}} = \mu_{eh} = \mu_\gamma$$

This equilibrium requires frequent interactions between electron–hole pairs and photons, i.e. frequent absorption and emission as will be present if the semiconductor is in a perfectly reflecting cavity and the temperature and chemical potential of its electron–hole pairs are kept constant (we will learn later how this is achieved). This equilibrium exists also in a homogeneously excited semiconductor that is thicker than the penetration depth of the photons, where photons are repeatedly emitted and reabsorbed before they reach the surface.

It must be emphasized that the validity of the generalized Planck equation (3.99) and (3.102) is not restricted to a situation of equilibrium between electron–hole pairs and photons. This equilibrium was not required in its derivation. The only condition for its validity is that the Fermi distributions for the energy ranges between which optical transitions occur (e.g. conduction band and valence band) must be well defined, which is the case if the recombination lifetimes of electrons and holes are much longer than the scattering times.

The above treatment is valid quite generally, even if $\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}} > \hbar\omega$, where the denominator in Equations 3.99 and 3.102 is negative. Under the same condition, the absorption coefficient in Equation 3.97 is negative too, and the spontaneous emission rate in Equations 3.98 and 3.99 remains positive. When the absorption coefficient is negative, stimulated emission overcompensates the rate of upward

transitions and the semiconductor amplifies the incident light as in a laser. Amplification results from a negative absorption coefficient, which leads to a negative absorptance a , increasing exponentially with the thickness d in Equation 3.103 $\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}} > \hbar\omega$ is also known as the *condition for lasing*. An equilibrium between the electron–hole pairs and the photons cannot exist if $\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}} > \hbar\omega$ because then the emission rate of the semiconductor is always larger than the rate of photon absorption, since the semiconductor is a photon current amplifier under this condition.

Equation 3.99 describes the rate of radiative transitions. In a real material, there are also nonradiative transitions occurring in parallel and in addition. This does not invalidate Equation 3.99 as long as the assumptions are justified that two separate Fermi distributions describe the occupation of states in the two energy ranges between which the transitions occur.

The generalized Planck law in Equation 3.99 is important for solar cells, because it allows one to determine the smallest, theoretically possible recombination rate, the longest possible lifetime, and the largest possible difference of the Fermi energies. But the generalized Planck law has practical importance as well.

Since Equation 3.102 is a quantitative law with no adjustable parameters, it allows the experimentalist to determine the difference in the Fermi energies from an absolute measurement of the emitted photon current and a knowledge of the absorptance $a(\hbar\omega)$. The knowledge of the difference in the Fermi energies allows a better characterization of a solar cell material than does the lifetime of electrons and holes.

Alternatively, the absorptance and from it the absorption coefficient can be obtained from Equation 3.102 by measuring the emitted photon current if the difference in the Fermi energies is known. Measuring the emitted photon current density in absolute units is not easy experimentally. If the difference in the Fermi energies is not known but the emitted photon current is known except for a constant instrumental factor, the measured photon current can be adjusted to Equation 3.102, if the absorptance is known at one single photon energy allowing one to obtain the absorptance $a(\hbar\omega)$ at all other photon energies in the measured luminescence spectrum [8].

3.8

Problems

- 3.1** What are the consequences of doping of a semiconductor?
- 3.2** Doping of a semiconductor with donors increases the concentration of electrons. Why does this reduce the hole concentration in the dark?
- 3.3** What is the condition for the validity of the quasi-Fermi energy concept?
- 3.4** Under which conditions is the occupation of impurity states given by one of the quasi-Fermi energies even if its energy level ε_{imp} is in between the two quasi-Fermi energies? Consider different capture cross sections for electrons and holes.

- 3.5** What is the difference between the temperature dependences of the electrical conductivity of a metal and of an intrinsic semiconductor?
Explain qualitatively how doping changes this temperature dependence for a semiconductor.
- 3.6** Why is the conduction band of a semiconductor not totally unoccupied at room temperature? Would that not minimize its energy?
- 3.7** The density of states for electrons in the conduction band was derived in Equation 3.7 How does the density of states depend on energy for a one-dimensional and a two-dimensional electron gas?
- 3.8** Which assumption leads to the dependence $\alpha(\hbar\omega) \sim (\hbar\omega - \epsilon_G)^{1/2}$ of the absorption coefficient α on the photon energy $\hbar\omega$?
- 3.9** Determine the minimum difference $\Delta\epsilon$ between the electron quasi-Fermi energy and the conduction band edge of a semiconductor in units of kT for which the concentration of electrons in the conduction band n_e calculated with the Boltzmann distribution approximates the result using Fermi–Dirac statistics within 1%
(a) for $T = 300$ K.
(b) for $T = 400$ K.
- 3.10** Determine quantitatively the free energy for electrons and holes per volume

$$F^* = F_e^* + F_h^* = \frac{[E_e + E_h - T(S_e + S_h)]}{V}$$

in intrinsic silicon ($N_C = 3 \times 10^{19} \text{ cm}^{-3}$, $N_V = 1 \times 10^{19} \text{ cm}^{-3}$ and $\epsilon_G = 1.12 \text{ eV}$) for $T = 300$ K. Set the zero point of F^* to the state where the conduction band is empty and the valence band is fully occupied. For which concentration of n_e does F^* reach its minimum?

- (a)** Use the Sackur Tetrode equation (3.35) for the determination of the entropy S .
(b) Calculate the entropy according to $S = k \ln \Omega$, with Ω being the number of realizations. The number of realizations is the number of possibilities to distribute n particles in N states. Assume the density of states in the conduction band to depend on energy according to Equation 3.7

$$D_e(\epsilon_e) = 4\pi \left(\frac{2m_e^*}{\hbar^2} \right)^{3/2} (\epsilon_e - \epsilon_C)^{1/2}$$

and for the valence band according to

$$D_h(\epsilon_h) = 4\pi \left(\frac{2m_h^*}{\hbar^2} \right)^{3/2} (\epsilon_h + \epsilon_V)^{1/2}$$

- 3.11** Calculate the binding energy of the ground state and the first two excited states (hydrogen atom approximation) of an electron of a donor with five valence electrons in a silicon lattice. Use a dielectric function $\epsilon(\text{Si}) = 11.9$ and an effective mass of
(a) $m_e^* = 1.08 m_e$.

- (b) $m_e^* = 0.02 m_e$.
- (c) How large is the Bohr radius of the ground states for case (a) and (b) and over how many atoms is the electron “smeared out”? Use $\rho_{\text{Si}} = 2.3 \text{ g cm}^{-3}$ and a molar mass of $M_{\text{Si}} = 28.1 \text{ g/mol}$.
- 3.12** Calculate the Fermi energy in n-doped Si with $n_D = 10^{16} \text{ cm}^{-3}$ for $T = 300 \text{ K}$,
- assume full ionization of the donors.
 - numerically, use $\epsilon_C - \epsilon_D = 0.2 \text{ eV}$.
 - At which temperature is $n_h \geq 0.5 n_e$? What fraction of n_e originates from the donors at that temperature?
- 3.13** A sample of n-doped Si is illuminated resulting in $n_e = 10^{17} \text{ cm}^{-3}$ and $n_h = 10^{10} \text{ cm}^{-3}$, which are assumed to be constant throughout the sample due to large diffusion lengths. Consider acceptor-type impurity states with $\tau_{h,\min} = \tau_{e,\min} = 10^{-6} \text{ s}$. Calculate the recombination rates as a function of the impurity level ($\epsilon_V \leq \epsilon_{\text{imp}} \leq \epsilon_C$) for temperatures $T = 300 \text{ K}$ and $T = 400 \text{ K}$. Plot a graph of the recombination rates as a function of the impurity level.

4

Conversion of Thermal Radiation into Chemical Energy

In Chapter 2 we have seen what a Carnot engine can do, which operates with the heat gained from a black or a monochromatic absorber of solar radiation. The absorber was characterized by a high temperature T_A and a chemical potential of its electron–hole excitations of $\mu_{eh} = 0$. A solar cell remains at about $T_0 = 300\text{ K}$ and if it had $\mu_{eh} = 0$ as well, it would not be able to convert heat into an entropy-free form of energy. But as was shown in Chapter 3, electron–hole pairs in an illuminated semiconductor have a chemical potential of $\mu_{eh} \neq 0$, which means that conversion of solar heat into chemical energy has already taken place in the semiconductor.

In order to examine this conversion in more detail, we will follow the generation and thermalization of electrons and holes step by step. These steps are shown in Figure 4.1. In the first step, photons absorbed from fully concentrated solar radiation establish an energy distribution of the electron–hole pairs, which is identical to the energy distribution of the absorbed photons reflecting the high temperature of the Sun, T_S . In this state, radiative recombination of electron–hole pairs results in the emission of photons with the same spectrum and filling (for maximal concentration) the same solid angle as the absorbed photons. This will not be changed, when, in a next step, scattering of electrons and holes between each other is allowed (the interaction with the lattice vibrations is still switched off), because scattering preserves the number of electrons and holes and the average energy per electron–hole pair. The temperature of the electron–hole pairs defined by their energy distribution is still $T_{eh} = T_S$. In this situation, the emitted photons and the electron–hole pairs in the semiconductor are in thermal and chemical equilibrium with the solar radiation and have a chemical potential of $\mu_\gamma = \mu_{eh} = 0$.

When concentration is not at maximum and emitted photons cover a larger solid angle than the absorbed photons, the emitted photon current density per solid angle is smaller than the absorbed, but still has the same spectrum. In this situation, $T_{eh} = T_S$ because of the spectrum but $\mu_\gamma = \mu_{eh} < 0$.

Returning to maximum concentration, the electrons and holes are in thermal and chemical equilibrium with the Sun with no difference of the Fermi energies.

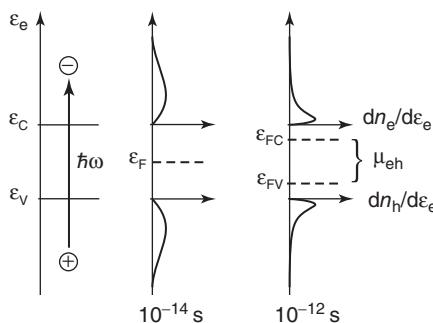


Figure 4.1 Electrons and holes generated by photons with energy $\hbar\omega$ lose energy by thermalization, which produces chemical energy per electron–hole pair μ_{eh} .

In a rough estimate, the average energy of the electron–hole pairs is then $\epsilon_G + 3kT_S$. From Equation 3.29 it follows that at $T_{eh} = T_S$

$$\mu_{eh}(T_S) = \epsilon_{FC} - \epsilon_{FV} = \epsilon_G - kT_S \ln \left(\frac{N_C N_V}{n_e n_h} \right) = 0$$

from which $kT_S \ln [N_C N_V / (n_e n_h)] = \epsilon_G$.

In the next step, the interaction with the lattice vibrations is switched on. This leads to a cooling of the electrons and holes, until after about 10^{-12} s the lattice temperature T_0 is reached, while the concentrations of electrons and holes remain constant. For $T_{eh} = T_0$

$$\mu_{eh}(T_0) = \epsilon_{FC} - \epsilon_{FV} = \epsilon_G - kT_0 \ln \left(\frac{N_C N_V}{n_e n_h} \right)$$

Because of the constant concentrations of electrons and holes and neglecting the temperature dependence of the effective densities of states N_C and N_V , this finally gives us an estimate of the chemical energy per electron–hole pair produced by the cooling process

$$\mu_{eh}(T_0) = \epsilon_{FC} - \epsilon_{FV} = \epsilon_G \left(1 - \frac{T_0}{T_S} \right)$$

This relation looks like a Carnot efficiency, but it is different. For the efficiency, we have to compare $\epsilon_{FC} - \epsilon_{FV}$ not with ϵ_G but with the energy $\epsilon_G + 3kT_S$ per electron–hole pair before the cooling. We can then see a large energy loss resulting from thermalization, which can be attributed to the production of entropy by the lattice vibrations. The cooling of electrons and holes at constant concentrations is therefore far from an ideal process for the production of chemical energy, if we start from a broad energy distribution. Although in this consideration no electron–hole pairs have been extracted from the semiconductor absorber, an equilibrium between the Sun and the semiconductor does not exist if the semiconductor is at a different temperature. A nonzero chemical potential of its electron–hole pairs reduces the entropy production compared with a zero chemical potential absorber, as it enables the absorber to emit as many photons as

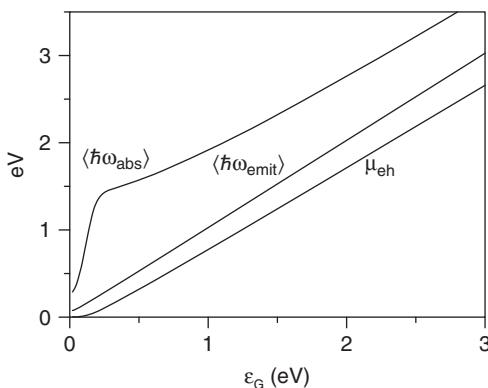


Figure 4.2 Average energy of absorbed photons $\langle \hbar\omega_{\text{abs}} \rangle$ from a nonconcentrated AM0 spectrum and from the 300 K background radiation, the average energy of emitted photons $\langle \hbar\omega_{\text{emit}} \rangle$ and the chemical potential of electron–hole pairs μ_{eh} of

a semiconductor as a function of its band gap ϵ_G for radiative recombination under open-circuit conditions, resulting from the generalized Planck radiation law. The difference between $\langle \hbar\omega_{\text{abs}} \rangle$ and $\langle \hbar\omega_{\text{emit}} \rangle$ is lost by thermalization of the electron–hole pairs.

it absorbs, but it does not fully compensate the temperature difference, because the emitted spectrum is different from the absorbed spectrum, and less energy is emitted than is absorbed.

Figure 4.2 shows the exact result of the thermalization of electron–hole pairs generated by nonconcentrated sunlight and by 300 K background radiation under open-circuit conditions, if there is only radiative recombination. Applying the generalized Planck law from Equation 3.102, the comparison of the average energy $\langle \hbar\omega_{\text{abs}} \rangle$ of the absorbed photons, which is invested for the generation of each electron–hole pair with the average energy $\langle \hbar\omega_{\text{emit}} \rangle$ of the emitted photons, shows the energy loss by thermalization, which, however, results in the chemical energy per electron–hole pair μ_{eh} . This chemical energy can, e.g., initiate further chemical reactions as in photosynthesis, where a long-term storage of energy is achieved. In a solar cell, the chemical energy is an intermediate product from which electrical energy can be obtained by a further step.

The chemical energy per electron–hole pair is the sum of the electrochemical potentials of the electrons and the holes, and therefore

$$\eta_e + \eta_h = \mu_e + \mu_h = \epsilon_{\text{FC}} - \epsilon_{\text{FV}}$$

According to Equation 3.29, this gives us

$$\mu_e + \mu_h = \epsilon_{\text{FC}} - \epsilon_{\text{FV}} = kT \ln \left(\frac{n_e n_h}{n_i^2} \right) \quad (4.1)$$

If we again separate the particle densities into dark-state and additional densities, so that $n_e = n_{e,0} + \Delta n_e$ and consider the case of a p-type semiconductor with $n_h \approx n_h^0$ (weak excitation), we then find

$$\mu_e + \mu_h = kT \ln \left(1 + \frac{\Delta n_e}{n_e^0} \right) \quad (4.2)$$

This means that for weak excitation the chemical energy per electron–hole pair is the result of only the change in the chemical potential of the minority charge carriers (here the electrons) or the change in their Fermi energy.

4.1

Maximum Efficiency for the Production of Chemical Energy

Is it imaginable that a material at room temperature is in equilibrium with the Sun, in which it emits toward the Sun as many photons, as much energy and as much entropy as it absorbs from the Sun? As we have seen in the previous paragraph, the main problem is the energy loss accompanying the entropy production during the thermalization process. Figure 4.3 shows that an energy loss cannot occur if the states available for electrons and holes are confined to a narrow energy interval. Such a material could only absorb almost monochromatic radiation. Alternatively, if in a broadband semiconductor the electrons and holes are generated by almost monochromatic radiation, and only into those states in which they are after the thermalization, an energy loss can be avoided. For maximum efficiency, we therefore make the following idealized assumptions:

- 1) Only radiative recombination takes place.
- 2) There is no extraction of electrons and holes.
- 3) There are no thermalization losses. Absorption and emission are monochromatic, with an energy of $\hbar\omega = \varepsilon_G$ and an absorptance of $a(\hbar\omega = \varepsilon_G) = 1$ over an interval $d\hbar\omega$. We can consider this condition fulfilled by a filter enclosing the semiconductor, which is transparent only for $\hbar\omega = \varepsilon_G$, reflecting all other photons.

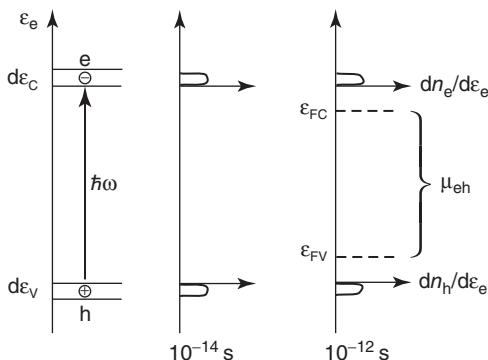


Figure 4.3 Electrons and holes generated in narrow energy ranges by monochromatic radiation have the same energy distribution after thermalization as before.

- 4) The rate of generation is at maximum. This condition is achieved by maximum concentration, where the solid angles for absorption Ω_{abs} and emission Ω_{emit} are equal.

From the first two conditions, it follows that in the steady state the emitted photon current is equal to the absorbed photon current. Referring to Equation 2.35 and Equation 3.102 this means that

$$\begin{aligned} dj_{\gamma,\text{emit}} &= \frac{\Omega_{\text{emit}}}{4\pi^3 h^3 c^2} \frac{\varepsilon_G^2 d\hbar\omega}{\exp\left(\frac{\varepsilon_G - \mu_{\text{eh}}}{kT_0}\right) - 1} \\ &= \frac{\Omega_{\text{abs}}}{4\pi^3 h^3 c^2} \frac{\varepsilon_G^2 \ln \hbar\omega}{\exp\left(\frac{\varepsilon_G}{kT_S}\right) - 1} = dj_{\gamma,\text{abs}} \end{aligned} \quad (4.3)$$

With condition 4, $\Omega_{\text{emit}} = \Omega_{\text{abs}}$, under which the semiconductor would reach the temperature of the Sun T_S unless we hold its temperature constant at $T = T_0$, we find the chemical energy per electron–hole pair

$$\mu_{\text{eh}} = \mu_e + \mu_h = \varepsilon_G \left(1 - \frac{T_0}{T_S}\right) \quad (4.4)$$

The efficiency for the conversion of solar heat into chemical energy is

$$\eta = \frac{\mu_e + \mu_h}{\varepsilon_G} = 1 - \frac{T_0}{T_S} \quad (4.5)$$

This is the Carnot efficiency, a limiting value that is obtained for the conversion of heat into an entropy-free form of energy when the conversion process is reversible, i.e., takes place without the production of entropy. We can now see that an ideal semiconductor, which only has radiative recombination, represents an ideal converter of heat into chemical energy in monochromatic operation.

The state defined by Equation 4.4 is a state of equilibrium between the semiconductor and the Sun, in which not only are the absorbed and emitted photon currents equal, as expressed by Equation 4.3, but also the energy currents that follow from the photon currents by multiplication with the photon energy, which for absorption and emission have the same value $\hbar\omega = \varepsilon_G$. For the absorbed and emitted radiation, the entropy per photon state is [9]

$$\sigma_\gamma = k \left[(1 + f_\gamma) \ln(1 + f_\gamma) - f_\gamma \ln(f_\gamma) \right] \quad (4.6)$$

where f_γ is the Bose–Einstein distribution function

$$f_\gamma = \frac{1}{\exp[(\hbar\omega - \mu_\gamma)/kT] - 1} \quad (4.7)$$

The exponent in the distribution function has the same value for the Sun and the semiconductor according to Equations 4.3 and 4.4 with $\mu_e + \mu_h = \mu_\gamma$ because of the chemical equilibrium between the electron–hole pairs and the photons by frequent absorption and emission in a material with absorptance $a(\hbar\omega = \varepsilon_G) = 1$. In addition, absorption and emission involve the same number of photon states

because of $\Omega_{\text{abs}} = \Omega_{\text{emit}}$. Under these conditions, the entropy absorbed with the photons from the Sun is equal to the entropy emitted with the photons toward the Sun. This all shows that an equilibrium is possible between the Sun and a material at room temperature, if the absorption and emission of photons is limited to the same narrow energy range. This equilibrium requires the right combination of temperature T and chemical potential μ as expressed by Equation 4.4, which is why we may call it a thermochemical equilibrium. The chemical energy produced by reversible, nondissipative cooling of electrons and holes is carried away by the emitted photons. The emitted photons, known as *luminescent radiation* from LEDs, look the same as the solar photons. For monochromatic radiation, there is no way to decide whether they have a high temperature T and a small chemical potential μ_γ or a low temperature and a large chemical potential.

The process in which all the chemical energy is emitted with the photons is in fact of no interest to us. We are instead interested in how much chemical energy can be harvested with the electron–hole pairs, if we knew how to extract them from the absorber. From the continuity equation for the electrons under steady-state conditions,

$$\frac{\partial n_e}{\partial t} = G_e - R_e - \text{div } dj_e = 0$$

we see that

$$\text{div } dj_e = G_e - R_e \quad (4.8)$$

gives just the rate with which electrons can be extracted from a volume element, since more flow out of this element than into it when $G_e > R_e$. We are considering a differentially small electron current dj_e , since it results from absorption and emission in a differentially small photon energy interval $d\hbar\omega$.

Integrating Equation 4.8 over the volume of the semiconductor gives the total extracted electron current dj_e

$$dj_e = \int G_e dx - \int R_e dx$$

If we again restrict ourselves to radiative recombination, the extracted electron current density is given by the difference between the absorbed and emitted photon current densities, which already includes photon recycling by reabsorption. Thus,

$$dj_e = dj_{\gamma,\text{abs}} - dj_{\gamma,\text{emit}} \quad (4.9)$$

The same relationship is true for the holes. The current of extracted electron–hole pairs is $dj_{eh} = dj_e = dj_h$.

For a spatially constant chemical potential of the electron–hole pairs $\mu_{eh} = \text{const}$, j_{eh} follows from the relations in Equation 4.3 for the absorbed and emitted photon currents.

In an approximation to the generalized Planck law, when $\mu_{eh} < \epsilon_G - 3kT$ allows us to neglect the “−1” in the denominator of the Planck law, we find according to Equation 4.1

$$dj_{eh} = dj_{\gamma,abs} - dj_{\gamma}^0 \frac{n_e n_h}{n_i^2} = dj_{\gamma,abs} - dj_{\gamma}^0 \exp\left(\frac{\mu_e + \mu_h}{kT}\right) \quad (4.10)$$

where dj_{γ}^0 is the photon current density absorbed and emitted in equilibrium with the 300 K background radiation.

The condition $\mu_e + \mu_h = \text{constant}$ is fulfilled, for example, when electrons and holes are produced uniformly throughout the volume as a result of weak absorption or when they are uniformly distributed throughout the volume as a result of large diffusion coefficients and lifetimes.

For a given current of absorbed photons $dj_{\gamma,abs}$, the current dj_{eh} of extracted electrons and holes is the difference between the absorbed and the emitted photon currents, which are shown in Figure 4.4 as a function of the chemical energy extracted along with the electron–hole current. We see that the extracted electron–hole current is nearly equal to the absorbed photon current for small values of $\mu_{eh} = \mu_e + \mu_h$, where almost no photons are emitted. The emitted photon current $dj_{\gamma,emit}$ rises exponentially with $\mu_e + \mu_h$, until the open-circuit situation is reached at $\mu_{eh,oc}$, where absorbed and emitted photon currents are equal and no electron–hole pairs are extracted. The current of chemical energy extracted along with the electron–hole pairs is

$$dj_{\mu} = dj_{eh}(\mu_e + \mu_h)$$

Its maximum value is the largest rectangle (dark grey in Figure 4.4), which can be fitted between the absorbed and emitted photon currents. The absorbed energy current in the interval $d\hbar\omega$ is $dj_{\gamma,abs} \hbar\omega$ and is given by the lightly shaded rectangle with the broken-line border in Figure 4.4. The efficiency with which chemical energy is obtained from the absorbed radiation is finally given by the ratio of the dark and lightly shaded rectangles. It must be emphasized that we just assume that electron-hole pairs can be extracted along with their chemical energy, although, at this point, we do not know, how to do it.

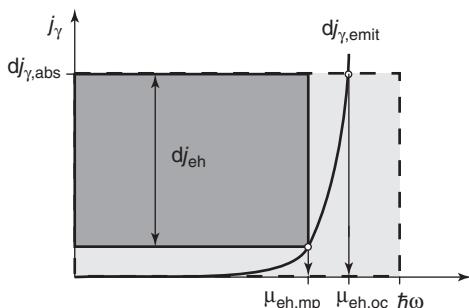


Figure 4.4 Current $dj_{\gamma, emit}$ of emitted photons as a function of the chemical energy $\mu_{eh} = \mu_e + \mu_h$ of the electron–hole pairs. The dark rectangle is the current of chemical energy extracted along with the

current $dj_{eh} = dj_{\gamma, abs} - dj_{\gamma, emit}$ of extracted electron–hole pairs. It is produced from the current of absorbed energy represented by the grey rectangle.

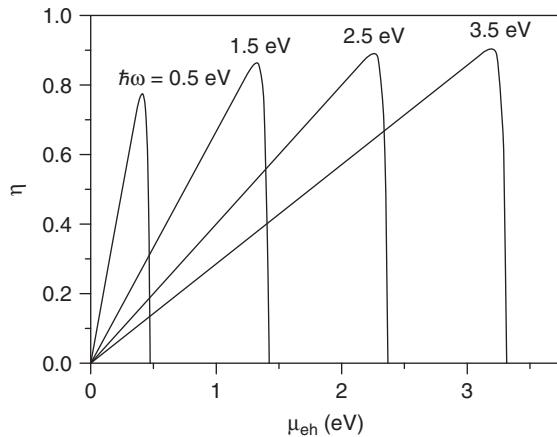


Figure 4.5 Efficiency η with which chemical energy is extracted as a function of the chemical potential μ_{eh} of the electron–hole pairs from monochromatic absorbers for different photon energies.

Figure 4.5 shows the efficiency η with which chemical energy is extracted as a function of the chemical potential of the electron–hole pairs. It starts with zero at $\mu_{eh} = 0$, when electrons and holes are so effectively extracted that none can accumulate, then it rises linearly and reaches a maximum at the point of maximum power (mp), after which it drops sharply down to zero at the open-circuit situation, where all electron–hole pairs recombine to emit photons.

In order to find the maximum efficiency with which chemical energy can be obtained from fully concentrated solar radiation, we proceed as follows. From the generalized Planck law in Equation 3.102, the absorbed and emitted photon currents are found as a function of the monochromatic photon energy. For each photon energy, the chemical potential μ_{eh} in the emitted photon current is varied until the value $\mu_{eh,mp}$ is found for which the current of extracted chemical energy is maximal. Dividing this chemical energy current by the absorbed monochromatic energy current gives the efficiency $\eta_{mono}(\varepsilon_G)$ as a function of the band gap.

The result of this calculation is shown in Figure 4.6 for nonconcentrated radiation and for fully concentrated radiation. The efficiencies are rather large and rise with the band gap ε_G , demonstrating the high value of solar energy. In order to make proper use of the total solar spectrum, very (infinitely) many monochromatic absorbers have to be employed. The overall efficiency then follows from an integration of the efficiency $\eta_{mono}(\varepsilon_G)$ over the spectrum after weighting each interval $d\hbar\omega$ with its share of the absorbed energy current.

$$\eta = \frac{\int_0^\infty \eta_{mono}(\varepsilon_G) \hbar\omega dj_{\gamma,abs}}{\int_0^\infty \hbar\omega dj_{\gamma,abs}}$$

The total efficiency is 86% for fully concentrated radiation from a 5800 K Sun and 67% for nonconcentrated radiation from a solid angle of 6.8×10^{-5} [10].

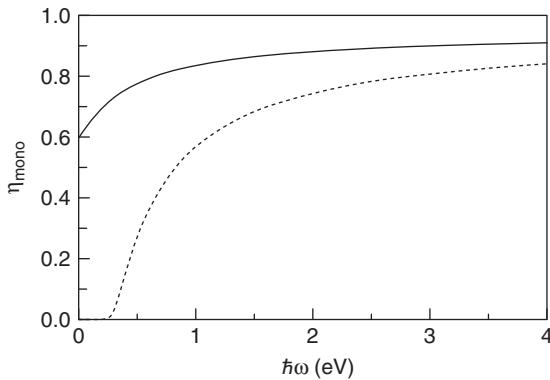


Figure 4.6 Monochromatic efficiency $\eta_{\text{mono}}(\varepsilon_G)$ for obtaining chemical energy $\mu_e + \mu_h$ as a function of the photon energy $\hbar\omega$ of fully concentrated (full line) and nonconcentrated (broken line) monochromatic solar radiation.

Although we were able to calculate the efficiency of the chemical energy that can be obtained from an illuminated semiconductor by the extraction of some of the generated electron–hole pairs, we still have no idea how a device should look like that allows this extraction of electron–hole pairs and their energy. This is the subject of the next chapter.

4.2

Shockley–Queisser Limit

Of more interest for real solar cells is the illumination of a semiconductor with the entire solar spectrum. For the AM0 spectrum the photon current absorbed by a semiconductor of bandgap ε_G with an absorptance of $a(\hbar\omega < \varepsilon_G) = 0$ and $a(\hbar\omega \geq \varepsilon_G) = 1$ is

$$j_{\gamma,\text{abs}}(\varepsilon_G) = \frac{\Omega_S}{4\pi^3 \hbar^3 c^3} \int_{\varepsilon_G}^{\infty} \frac{(\hbar\omega)^2 d\hbar\omega}{\exp[\hbar\omega/(kT_S)] - 1}$$

The emitted photon current from Equation (3.102) is

$$j_{\gamma,\text{emit}}(\varepsilon_G) = \frac{\pi}{4\pi^3 \hbar^3 c^3} \int_{\varepsilon_G}^{\infty} \frac{(\hbar\omega)^2 d\hbar\omega}{\exp[(\hbar\omega - \mu_{\text{eh}})/(kT_0)] - 1}$$

If there is only radiative recombination, the current of chemical energy, which can be withdrawn from the semiconductor, is $\mu_{\text{eh}} (j_{\gamma,\text{abs}} - j_{\gamma,\text{emit}})$. By dividing the chemical energy current by the incident energy current, the efficiency is obtained

$$\eta(\varepsilon_G) = \frac{\mu_{\text{eh}} (j_{\gamma,\text{abs}} - j_{\gamma,\text{emit}})}{\frac{\Omega_S}{4\pi^3 \hbar^3 c^3} \int_0^{\infty} \frac{(\hbar\omega)^3 d\hbar\omega}{\exp[\hbar\omega/(kT_S)] - 1}}$$

The maximum of this expression as a function of μ_{eh} is the maximum efficiency by which solar energy with the AM0 spectrum can be converted into chemical

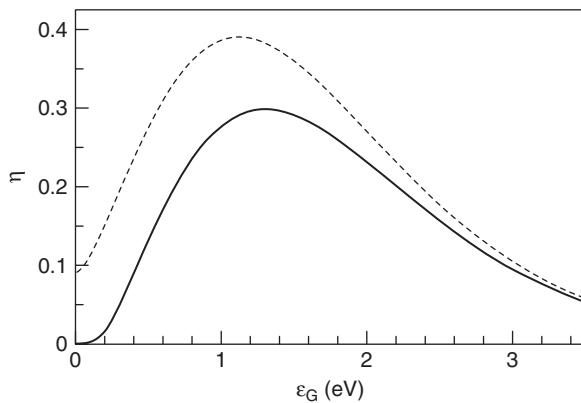


Figure 4.7 Efficiency of solar cells with radiative recombination only as a function of their energy gap for the AM0 spectrum, non-concentrated (solid line) and for full concentration (dashed line).

energy by a semiconductor with an absorptance that rises from 0 to 1 at the bandgap ε_G . This limiting efficiency is shown in Figure 4.7. It was found by Shockley and Queisser in 1961 [11]. It is based on the assumption that the chemical energy can be withdrawn from the semiconductor without any losses. A lossless transport of the electrons and holes out of the semiconductor would imply that the driving forces driving the electrons and holes are zero; otherwise the energy dissipated by them in the transport would be a loss.

4.3

Problems

- 4.1 What is the reason for the linear increase of the efficiency η of the extraction of chemical energy in Figure 4.5 for small values of μ_{eh} ?
- 4.2 Calculate the quasi-Fermi energy of the electrons in a p-type Si sample ($n_h \approx n_h^0 = 5 \times 10^{16} \text{ cm}^{-3}$) at $T_0 = 300 \text{ K}$ under open-circuit conditions, regarding
 - (a) only radiative recombination;
 - (b) only Auger recombination.
 Assume excitation by black-body radiation with $T_{bb} = 5800 \text{ K}$, from a solid angle of $\Omega = 6.8 \times 10^{-5}$ and an absorptance $a(\hbar\omega < \varepsilon_G) = 0$ and $a(\hbar\omega \geq \varepsilon_G) = 1$ with $\varepsilon_G = 1.12 \text{ eV}$. The charge carriers shall be distributed homogeneously over a thickness of $d = 300 \mu\text{m}$ because of large diffusion lengths.
- 4.3 Calculate the emitted photon current density per wavelength into an effective solid angle of π for $\lambda_G = 500 \text{ nm}$. At this wavelength, the absorber is characterized by an absorptance of $a(\lambda_G) = 1$. The emission occurs under conditions where the chemical potential of the electron–hole pairs is
 - (a) $\mu_{eh} = 0.5 \text{ eV}$;
 - (b) $\mu_{eh} = 1.8 \text{ eV}$.

5**Conversion of Chemical Energy into Electrical Energy**

When electrons and holes are removed pairwise from the same location and along the same path, there is no charge current connected with the extraction, because an electron–hole pair is electrically neutral. Electrical energy currents are bound to charge currents. It is therefore necessary to separate the electrons and the holes and extract them along different paths as shown in Figure 5.1. Extraction, in fact, is the wrong word. Electrons and holes must rather be driven out of an absorber by some internal force, since they must still be able to perform work in an external circuit.

In this chapter, we find out which are the forces that drive electrons and holes and how they may be provided by a solar cell structure.

5.1**Transport of Electrons and Holes**

Electrons have many “handles” where forces can be applied. The gradient of the gravitational potential acts on their mass, and the gradient of the electrical potential on their charge. The temperature gradient acts on their entropy, and the gradient of the chemical potential acts on their quantity. The gravitational force is negligibly small. Temperature gradients give rise to thermoelectric effects and we disregard them in connection with solar cells. Even though we already know that the charge and number or quantity are coupled during the exchange of electrons or holes, and consequently the forces acting on them are coupled, we first treat the effects of these forces separately. For this, we assume that only one of the two forces (gradient of the electrical potential and gradient of the chemical potential) is different from zero. At the same time, we must remember that the forces acting on the charge and those acting on the quantity, act on the same particles and must therefore be added to obtain a resultant force, which is then responsible for driving the particles.

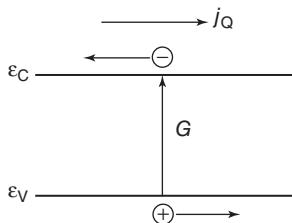


Figure 5.1 Electrons in the conduction band and holes in the valence band have to move in different directions to produce an electrical charge current j_Q .

5.1.1

Field Current

The electric field $E = -\text{grad } \varphi$ is the driving force acting on the charge. It is the only driving force for the charge current of electrons and holes if their concentration is spatially uniform. For electrons, then, their chemical potential

$$\mu_e = \mu_{e,0} + kT \ln\left(\frac{n_e}{N_C}\right)$$

is independent of position and

$$\text{grad } \mu_e = 0$$

the driving force for the diffusion current, still to be discussed, is zero. Figure 5.2 illustrates the dependence on position of the electron energies. The difference between the Fermi energy ε_{FC} and the conduction band edge ε_C does not depend on position, because the concentration of electrons is assumed to be uniform throughout. The gradient of ε_C is equal to the gradient of the electrical energy $-e\varphi$ because of the uniform electron affinity $\chi_e = -\mu_{e,0}$ in a uniform material.

The density of the charge field current for particles of type i with a charge per particle of $z_i e$ is

$$j_{Q,f,i} = z_i e n_i \langle v_i \rangle \quad (5.1)$$

The particles – with a concentration n_i – move at the mean velocity $\langle v_i \rangle$ and carry along the charge $z_i e$. Since the mean value of the velocity in the absence of an electric field, the mean thermal velocity $\langle v_{th} \rangle$, is zero, $\langle v_i \rangle$ is the mean value of the velocity in the presence of an electric field, or the drift velocity, which is small compared with $\langle v_{th} \rangle$.

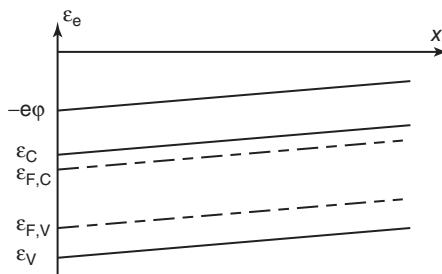


Figure 5.2 Electron energies in an electric field for uniform electron and hole concentrations.

Owing to their motion, which is mainly thermal, the charge carriers collide with obstacles, caused by any disturbance of the perfect periodicity of a crystal, such as phonons or impurities. The mean distance between collisions is called the *mean free path*.

The mean time $\tau_{c,i}$ between two collisions is called the *collision time*. In the electric field, the motion between collisions is characterized by the acceleration $a_i = z_i e E / m_i^*$. Because of the exponential distribution of the time intervals between collisions – a few particles undergo collisions only after a time much longer than $\tau_{c,i}$ – the mean velocity is given by

$$\langle v_i \rangle = \int_0^\infty a_i \exp(-t/\tau_{c,i}) dt = a_i \tau_{c,i}$$

thus

$$\langle v_i \rangle = z_i \frac{e}{m_i^*} \tau_{c,i} E \quad (5.2)$$

$b_i = e\tau_{c,i}/m_i^*$ is called the *mobility of the particles of type i*. The density of the field current is then

$$j_{Qf,i} = z_i^2 e n_i b_i E = \sigma_i E \quad (5.3)$$

Here $\sigma_i = z_i^2 e n_i b_i$ is the conductivity of the particles of type i . Making use of $E = -\text{grad } \varphi$ we can also express the charge current in terms of the gradient of the electrical energy per particle $z_i e \varphi$

$$j_{Qf,i} = -\frac{\sigma_i}{z_i e} \text{grad}(z_i e \varphi) \quad (5.4)$$

For electrons with $z_i = -1$, this is

$$j_{Qf,e} = \frac{\sigma_e}{e} \text{grad}(-e \varphi) \quad (5.5)$$

and for holes with $z_i = +1$,

$$j_{Qf,h} = -\frac{\sigma_h}{e} \text{grad}(e \varphi) \quad (5.6)$$

The direction of the charge field current is independent of whether it is carried by electrons or by holes.

5.1.2

Diffusion Current

When the electrical potential is the same at every location, that is, $\text{grad} \varphi = 0$, a pure diffusion current flows if the concentration is nonuniform. As Figure 5.3 shows, in a semiconductor made of uniform material the conduction band edge ε_C then has the same value everywhere. The difference between the Fermi energy ε_{FC} and the conduction band edge is concentration-dependent and therefore the nonuniform component of the chemical potential of the electrons

$$\mu_e - \mu_{e,0} = kT \ln \frac{n_e}{N_C}$$

where N_C is the effective density of states for electrons.

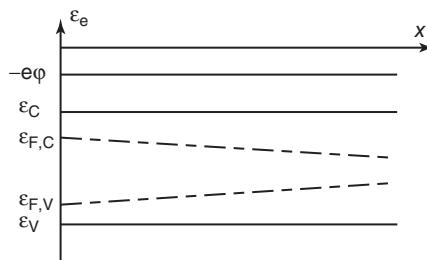


Figure 5.3 Electron energies for a position-dependent electron concentration in the absence of an electric field.

In the usual representation by Fick's Law, the charge current due to a nonuniform concentration of particles of type i is given by

$$j_{Q,d,i} = z_i e (-D_i \operatorname{grad} n_i) \quad (5.7)$$

where D_i is the diffusion coefficient. In this form, the current of the particles of type i is not proportional to their concentration n_i and thus does not allow us to recognize the driving force acting on this concentration of particles. We can make the current proportional to the concentration by transforming this expression

$$j_{Q,d,i} = -z_i e n_i D_i \frac{\operatorname{grad} n_i}{n_i}$$

Substituting $(\operatorname{grad} n_i)/n_i$ by $\operatorname{grad} \ln(n_i/N_i)$ and making use of the chemical potential we then obtain

$$j_{Q,d,i} = -\frac{z_i e n_i D_i}{kT} \operatorname{grad} \mu_i$$

In contrast to Fick's law, this expression is valid for an inhomogeneous chemical environment ($\operatorname{grad} \mu_{i,0} \neq 0$) as well, which would cause a diffusion current even if the concentration were homogeneous.

Using the so-called Einstein relation between diffusion coefficient D_i and mobility b_i

$$\frac{b_i}{D_i} = \frac{e}{kT}$$

we finally arrive at

$$j_{Q,d,i} = -\frac{z_i e n_i b_i}{e} \operatorname{grad} \mu_i = -\frac{\sigma_i}{z_i e} \operatorname{grad} \mu_i \quad (5.8)$$

For electrons with $z_i = -1$, the diffusion current is given by

$$j_{Q,d,e} = \frac{\sigma_e}{e} \operatorname{grad} \mu_e \quad (5.9)$$

and for holes with $z_i = +1$, the diffusion current is

$$j_{Q,d,h} = -\frac{\sigma_h}{e} \operatorname{grad} \mu_h \quad (5.10)$$

5.1.3

Total Charge Current

We have seen that there are two driving forces acting on all electrons and two other driving forces acting on all holes. If both forces, the gradient of the electrical energy and the gradient of the chemical potential are present simultaneously, they have to be added to give a resultant force, which is then multiplied by the conductivity to give the resultant current. We have written the field current and the diffusion current in such a way that the combination of the forces is easily possible. For the resultant current $j_{Q,i}$ of the particles of type i :

$$j_{Q,i} = -\frac{\sigma_i}{z_i e} \{ \text{grad } \mu_i + \text{grad } (z_i e\varphi) \} \quad (5.11)$$

or

$$j_{Q,i} = -\frac{\sigma_i}{z_i e} \text{grad}(\mu_i + z_i e\varphi) = -\frac{\sigma_i}{z_i e} \text{grad } \eta_i \quad (5.12)$$

The gradient of the electrochemical potential, $\text{grad } \eta_i = \text{grad} (\mu_i + z_i e\varphi)$, now gives us the combination of the individual forces acting separately on the quantity of the particles and on their charge. The gradient of the electrochemical potential is the total force, which, in a general case, drives the total current, when either the concentration of the particles or their chemical environment or the electrical potential, or all three, are inhomogeneous.

This procedure may not be very common. In most books, even in the general case, the argument is based on field and diffusion currents, as though there were electrons that "feel" only the electric field and others that only contribute to the diffusion current.

Since the only charged particles moving in a semiconductor are electrons and holes, the total charge current is

$$j_Q = \frac{\sigma_e}{e} \text{grad } \eta_e - \frac{\sigma_h}{e} \text{grad } \eta_h \quad (5.13)$$

or, making use of the quasi Fermi energies, for which $\varepsilon_{FC} = \eta_e$ and $\varepsilon_{FV} = -\eta_h$,

$$j_Q = \frac{\sigma_e}{e} \text{grad } \varepsilon_{FC} + \frac{\sigma_h}{e} \text{grad } \varepsilon_{FV} \quad (5.14)$$

This relationship is always true, i.e. when either an electric field or a concentration gradient, or both, exist.

Figure 5.4 gives an example in which both an electric field and concentration gradients for electrons and holes exist. According to Equation 5.14, the charge current j_Q is zero in spite of nonzero field and diffusion forces, because the Fermi energies are uniform. A separate description of the effects of the field and diffusion forces concludes that a field current and a diffusion current both flow and that both compensate each other exactly. The result, a zero total charge current, is identical. Nevertheless, the idea behind this result, i.e. all electrons and holes produce the field current as a result of the motion induced by the field, and, at the same time,

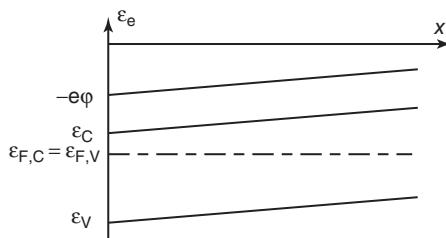


Figure 5.4 Constant Fermi energy in a semiconductor in which both an electric field and a concentration gradient exist.

all electrons and holes produce the diffusion current in the opposite direction, is incorrect.

That the field and diffusion currents do not exist separately can be seen very well if we examine the energy dissipation, that is, the generation of Joule heat associated with every current flow because of the scattering of the charge carriers. The rate of energy dissipation per volume element $d\rho_E/dt$ is proportional to the particle current density and the driving force. For the field current, it follows from Equation 5.4 that

$$\frac{d\rho_E}{dt} = \frac{j_{Q,f,i}}{z_i e} (-\text{grad}(z_i e\varphi)) = \frac{j_{Q,f,i}^2}{\sigma_i}$$

and for the diffusion current, it follows from Equation 5.8 that

$$\frac{d\rho_E}{dt} = \frac{j_{Q,d,i}}{z_i e} (-\text{grad} \mu_i) = \frac{j_{Q,d,i}^2}{\sigma_i}$$

As expected, in both cases the rate of energy dissipation $d\rho_E/dt > 0$, even if the individual currents would flow in opposite directions and compensate each other.

Figure 5.4 illustrates the case of thermal and electrochemical equilibrium of electrons and holes, as discussed in detail for the example of the pn-junction in Section 6.4. In this equilibrium state, dissipation of energy must not occur. In reality, the electrons and holes do not move in any preferred direction beyond their thermally random motion, because the forces acting on them add to a total force of $\text{grad} \epsilon_F = 0$, and no energy is dissipated.

The situation is the same as for the molecules in the air. There are also two forces acting on them: the gravitational force pulling them downwards and the pressure gradient or the gradient of the chemical potential pushing them upwards. We know the result. The distribution of the molecules is such that the two forces compensate each other. There is only Brownian motion of the molecules about an equilibrium position and there are no currents.

This discussion shows that field and diffusion currents are pure fiction and do not exist separately, and only the total current representation according to Equation 5.14 is correct.

However, it may not be concluded from this discussion that a mathematical treatment of charge transport in semiconductors in terms of field and diffusion

currents is incorrect. Mathematically, there is no difference, whether the driving forces are first added to give a resultant driving force, which is then multiplied by the conductivity to yield the total current, or whether the driving forces are first multiplied separately by the conductivity and then added to give the total current. The difference is in the physical picture that follows from the mathematical procedure. We obtain the physically correct description for the motion of the charge carriers only when we first add the driving forces to give a resultant driving force, which is the gradient of the Fermi energy.

5.2

Separation of Electrons and Holes

Now that we are familiar with the driving force for the motion of electrons and holes, we shall return to the original problem of defining a structure in which illumination produces a charge current, along with an electrical voltage V at its terminals, arising from a difference of the Fermi energies at the left and right terminals

$$eV = \varepsilon_{F,\text{left}} - \varepsilon_{F,\text{right}} \neq 0$$

We begin with a homogeneously exposed n-doped semiconductor. This is a structure in which, because of its symmetry, there is no preference for the transport of electrons in one direction and holes in the opposite direction, and no current or voltage is expected.

Figure 5.5 shows the Fermi energies between the bands. As a result of an assumed strong surface recombination, the concentrations of electrons and holes at the surfaces at the right and left do not differ from their values in the dark state, even though the semiconductor is illuminated. The Fermi energies for the conduction and valence bands, which are different inside the semiconductor, therefore merge into a single Fermi energy at the surface. This results in gradients for the two Fermi energies, which drive electrons and holes toward both surfaces, where they recombine. Since the absence of an external circuit implies that no

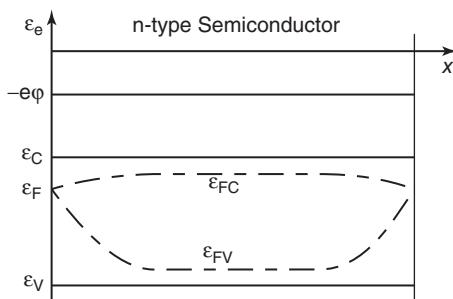


Figure 5.5 Distribution of the Fermi energies of a homogeneously illuminated n-type semiconductor with strong surface recombination on the left and on the right.

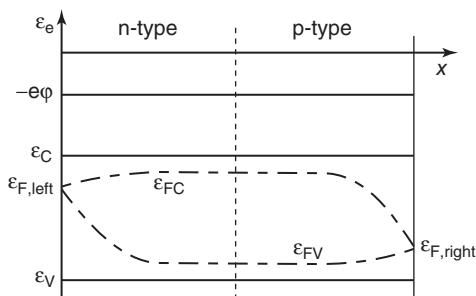


Figure 5.6 Distribution of the Fermi energies in a homogeneously illuminated pn-structure.

charge current can flow, the particle currents of electrons and holes flowing to the same surface must be equal. Owing to the larger conductivity of electrons for the n-conductor under discussion, it follows from Equation 5.14 that the gradient of the Fermi energy for the conduction band ϵ_{FC} is smaller than the gradient of ϵ_{FV} . In a p-conductor this would be just the opposite. This fact leads us to conclude that replacing the n-doping in the right half of Figure 5.5 by p-doping must give the distribution of Fermi energies in Figure 5.6, for a situation where no charge current flows. We see that, in an illuminated pn configuration without charge current, i.e. under open-circuit conditions, a difference in the Fermi energies results between two surfaces with strong surface recombination, fixing the carrier concentrations at the surfaces at different values.

The electrons flow toward the left when their electrochemical potential, the Fermi energy, $\eta_e = \epsilon_{FC}$, decreases toward the left. The holes flow toward the right when their electrochemical potential, $\eta_h = -\epsilon_{FV}$, decreases toward the right (when ϵ_{FV} increases toward the right). For the electron current toward the left and the hole current toward the right the gradients of the Fermi energies ϵ_{FC} on the left and ϵ_{FV} on the right are small, since the conductivities of the electrons in the n-conductor and of the holes in the p-conductor, are large. This is the important property that causes the Fermi energies at the two surfaces to be different.

The gradients of the Fermi energies for the valence band in the n-conductor on the left and for the conduction band in the p-conductor on the right, which are unavoidable owing to the convergence of the Fermi energies at the surface, can become large. This results from the very small concentrations, and consequently small conductivities, of holes in the n-conductor and of electrons in the p-conductor. The arrangement of Figure 5.6 is, however, not an ideal converter of chemical energy into electrical energy, because the difference in the Fermi energies between the left and right surfaces is less than the separation within the semiconductor. The chemical energy per electron–hole pair ($\epsilon_{FC} - \epsilon_{FV}$) resulting from exposure to light cannot be fully utilized by an external circuit. This is due to too small a concentration of electrons in the n-conductor and of holes in the p-conductor in the dark state, at least near the surfaces, which require a non-negligible gradient for the majority carrier Fermi energies, even under open-circuit conditions, to compensate the current of the minority carriers.

A charge current flows in an external circuit, when the voltage V between the terminals is reduced below the open-circuit value. This increases the gradients of the Fermi energies for the majority carriers and decreases the gradients for the minority carriers, resulting in a preferential movement of electrons to the left and of holes to the right in Figure 5.6. We return to this point in the following chapter.

The prerequisite for the existence of the gradients of the Fermi energies required to drive the currents (without applying a voltage externally) is the separation of the Fermi energies. Since the charge current densities of the electrons and holes are, at most, equal to the current density of the photons absorbed, multiplied by the elementary charge, (in Si $\leq 42 \text{ mA cm}^{-2}$), for the usual doping levels of $(10^{16} - 10^{17}) \text{ cm}^{-3}$ in both the n-conductor and the p-conductor, only a very small Fermi energy gradient is required for the majority carriers. In spite of the current flow, a large part of the Fermi energy separation is still present as the difference of the Fermi energies at the terminals on the left and right. This produces the voltage V at the terminals of the solar cell:

$$V = \frac{1}{e} \int_{\text{left}}^{\text{right}} \text{grad } \varepsilon_{\text{FC}} \, dx \quad (5.15)$$

Since $\varepsilon_{\text{FC}} = \varepsilon_{\text{FV}}$ in both terminals, one could also integrate over $\text{grad } \varepsilon_{\text{FV}}$.

That the voltage is given by the difference between the Fermi energies on the left and right with $eV = \Delta\varepsilon_F$, results from the fact that $\text{grad } \varepsilon_F$ is the driving force for the charge current in the external circuit as well, e. g. in a voltmeter with finite internal resistance (an infinite internal resistance does not exist). The voltmeter shows 0 V when there is no driving force for the electrons within the voltmeter, i.e. $\text{grad } \varepsilon_F = 0$ within the voltmeter or $\Delta\varepsilon_F = 0$ at its terminals.

We see that a voltmeter, in fact, always measures differences in the electrochemical potential of the electrons if these are the only mobile particles within the instrument. A voltmeter does not measure differences $\Delta\varphi$ of the electrical potential as, e.g. between the left and right sides of the semiconductor in Figure 5.4, where it would show 0 V although $\Delta\varphi \neq 0$.

5.3

Diffusion Length of Minority Carriers

Our goal is that of the charge carriers produced by the absorption of photons, as many as possible should flow toward the terminals of the solar cell. Unfortunately, however, the existence of the driving force for the charge carriers is not sufficient. Since the electrons and holes recombine after a lifetime, they must also be able to reach the terminals in this time. For the structure shown in Figure 5.6 the transport mechanism is diffusion. There is no gradient of the electrical potential and no field present in Figure 5.6.

To see how far an electron can travel by diffusion before it vanishes by recombination, we look at a simple example, where electrons are injected with a current density j_e into a p-conductor as minority charge carriers, as in Figure 5.7. The

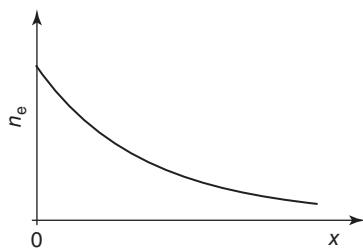


Figure 5.7 Distribution of electrons injected as minority charge carriers into a p-conductor.

electrons move in the x -direction; in the y - and z -directions the system is homogeneous. The charge injected with the electrons is quickly removed by a rearrangement of the many holes in the p-conductor without, however, eliminating the electrons as particles, as explained in Section 5.4.

The steady-state distribution of the additionally injected electrons in the x -direction is given by the continuity equation

$$\frac{\partial n_e}{\partial t} = G_e - R_e - \operatorname{div} j_e = 0 \quad (5.16)$$

For the particle diffusion current,

$$j_e = -D_e \frac{dn_e}{dx} \quad \text{and} \quad \operatorname{div} j_e = -D_e \frac{d^2 n_e}{dx^2}$$

In the p-conductor,

$$G_e = G_e^0 = \frac{n_e^0}{\tau_e} \quad \text{and} \quad R_e = \frac{n_e(x)}{\tau_e} = \frac{n_e^0}{\tau_e} + \frac{\Delta n_e(x)}{\tau_e}$$

It then follows from Equation 5.16 that

$$-\frac{\Delta n_e(x)}{\tau_e} + D_e \frac{d^2 \Delta n_e(x)}{dx^2} = 0 \quad (5.17)$$

The solution takes the form

$$\Delta n_e(x) = \Delta n_e(0) \exp\left(-\frac{x}{L_e}\right) \quad (5.18)$$

The characteristic length L_e is the diffusion length (here for the electrons). Substituting this result into Equation 5.17 we obtain

$$L_e = \sqrt{D_e \tau_e} \quad (5.19)$$

For electrons in pure silicon $D_e = 35 \text{ cm}^2 \text{ s}^{-1}$. For a lifetime of, e. g. $\tau_e = 10^{-6} \text{ s}$, $L_e = 60 \mu\text{m}$. In pure silicon, diffusion lengths of a few millimeters are obtained for electrons.

The diffusion length is the mean path length for the diffusion of a charge carrier during its lifetime. In the homogeneously illuminated pn structure in Figure 5.6 only those electrons that are generated in the p-conductor at a distance not larger than their diffusion length will reach the n-conductor.

Once the electrons are within the n-conductor, in a region in which the electrons are the majority charge carriers, their recombination has no effect on the charge

current. Even though they recombine in the n-conductor, recombination does not eliminate their charge. Since the charge is transported by the majority carriers, as is explained in the next section, other electrons continue to transport the charge. Recombination reduces the charge current only if it affects the minority carriers produced by the illumination.

5.4

Dielectric Relaxation

In our discussion of the diffusion of electrons in a p-conductor in the previous section, we did not consider the formation of space charge resulting from the charge of the injected electrons, and its effect on the transport of the electrons. We now show that any space charge of minority carriers is rapidly removed by a rearrangement of the majority carriers. In an otherwise completely homogeneous system with conductivity σ , a space charge $\rho_Q(x)$ is assumed to exist at the time $t = 0$. We now examine how rapidly it decays via the charge current j_Q driven by the field of the space charge. We assume that the majority carriers dominate the conductivity and that their rearrangement causes only negligible changes of their concentration, so that diffusion is not important.

The continuity equation for the charge takes the form

$$\frac{\partial \rho_Q}{\partial t} = -\operatorname{div} j_Q \quad (5.20)$$

Because charge is strictly conserved, there are no rates of production and annihilation. One of Maxwell's equations relates the space charge density ρ_Q to the electric field strength E

$$\operatorname{div} D = \epsilon\epsilon_0 \operatorname{div} E = \rho_Q \quad (5.21)$$

The electric field produces a charge current $j_Q = \sigma E$, from which we find

$$\operatorname{div} j_Q = \sigma \operatorname{div} E + E \operatorname{grad} \sigma \quad (5.22)$$

Because of the assumed homogeneity, $\operatorname{grad} \sigma = 0$. Substituting Equation 5.22 into Equation 5.20 and replacing $\operatorname{div} E$ from Equation 5.21,

$$\frac{\partial \rho_Q}{\partial t} = -\sigma \operatorname{div} E = -\frac{\sigma}{\epsilon\epsilon_0} \rho_Q \quad (5.23)$$

with the solution

$$\rho_Q(t) = \rho_Q(0) \exp\left(\frac{-t}{\epsilon\epsilon_0/\sigma}\right) \quad (5.24)$$

$\epsilon\epsilon_0/\sigma$ is known as the *dielectric relaxation time*. If it is small compared with the lifetime of the minority carriers, their motion is essentially unaffected by the initial space charge. In this example, the diffusion of electrons into the p-conductor of Figure 5.6 pulls compensating charge by dielectric relaxation of the majority carriers (holes) through the contacts into the p-conductor.

In p-silicon, with a doping of 10^{17} cm^{-3} , a hole mobility of $b_h = 480 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ and $\epsilon = 12$, the dielectric relaxation time is $\epsilon\epsilon_0/\sigma = 1.4 \times 10^{-13} \text{ s}$, and, consequently, several orders of magnitude smaller than the lifetimes.

5.5

Ambipolar Diffusion

For the treatment of diffusion in Section 5.3 we have assumed that the diffusing particles were minority carriers with a concentration much smaller than that of the majority carriers. As a result, the space charge accompanying the injected minority carriers had been completely eliminated by dielectric relaxation of the majority carriers.

Now we discuss the case of electrons and holes produced by light absorption close to the surface, both in large concentrations compared with the concentrations in the dark state. Owing to their concentration gradients they diffuse from the surface into the interior. We assume that the electrons have the larger mobility and larger diffusion coefficient, and, therefore, move faster. Electrons and holes therefore partly separate. This leads to the formation of a positive space charge near the surface caused by the holes left behind and a negative space charge due to the electrons in the interior. The electric field caused by this charge distribution is directed so as to compensate for the different mobilities of electrons and holes. The electrons are slowed down and the holes are accelerated. This coupled motion for a strong, nonhomogeneous excitation is referred to as *ambipolar diffusion*. Since this involves the motion of electrons and holes with the same velocity in the same direction, there is no charge current associated with ambipolar diffusion.

The particle currents of the electrons and the holes are

$$\begin{aligned} j_e &= -D_e \operatorname{grad} n_e - \frac{\sigma_e}{e} E \\ j_h &= -D_h \operatorname{grad} n_h + \frac{\sigma_h}{e} E \end{aligned} \quad (5.25)$$

Since the charge current vanishes,

$$j_Q = eD_e \operatorname{grad} n_e - eD_h \operatorname{grad} n_h + (\sigma_e + \sigma_h)E = 0 \quad (5.26)$$

we find that the electric field associated with ambipolar diffusion is given by

$$E = \frac{e}{\sigma_e + \sigma_h} (D_h \operatorname{grad} n_h - D_e \operatorname{grad} n_e) \quad (5.27)$$

It is because of this field that the electrons and holes move with the same velocity, so that $j_e = j_h$. Since very small differences in the concentrations already produce large space charges and high field strengths, the distributions of electrons and holes cannot be very different and $\operatorname{grad} n_e \approx \operatorname{grad} n_h = \operatorname{grad} n$. The resulting approximation for the field in Equation 5.27 substituted into Equation 5.25 gives the particle currents in the field E

$$j_e = j_h = -\frac{D_e \sigma_h + D_h \sigma_e}{\sigma_e + \sigma_h} \operatorname{grad} n \quad (5.28)$$

Since this has the form of the diffusion current equation, we can define

$$D_{\text{amb}} = \frac{D_e \sigma_h + D_h \sigma_e}{\sigma_e + \sigma_h} \quad (5.29)$$

as the ambipolar diffusion coefficient.

5.6

Dember Effect

To conclude this chapter, we present an example to show how important it is to distinguish between electrical potential differences and electrochemical potential differences.

We calculate the electrical potential difference resulting from the electric field of ambipolar diffusion. This means that the same assumptions of strong, inhomogeneous excitation apply, as a result of which the minority carrier concentration is greater than the majority carrier concentration in the dark. With the Einstein relation $D_i = b_i kT/e$, we express the diffusion coefficients in Equation 5.27 by the mobilities, apply the approximation that electrons and holes have the same spatial distribution ($\text{grad } n_e \approx \text{grad } n_h$), multiply both numerator and denominator in Equation 5.27 by $e(b_e + b_h)$, and find that

$$E = \frac{kT}{e} \frac{b_h - b_e}{b_e + b_h} \frac{\text{grad}(\sigma_e + \sigma_h)}{\sigma_e + \sigma_h} = \frac{kT}{e} \frac{b_h - b_e}{b_e + b_h} \text{grad} \ln(\sigma_e + \sigma_h) \quad (5.30)$$

Using $E = -\text{grad}\phi$, following integration, we find the electrical potential difference between the surface ($x = 0$) and the interior of the semiconductor ($x = \infty$) resulting from inhomogeneous exposure to light. This is the Dember voltage,

$$\Delta\phi_D = \phi(0) - \phi(\infty) = \frac{kT}{e} \frac{b_e - b_h}{b_e + b_h} \ln \left(\frac{\sigma_e(0) + \sigma_h(0)}{\sigma_e(\infty) + \sigma_h(\infty)} \right) \quad (5.31)$$

For the concrete case in which electrons and holes diffuse from the surface of a semiconductor, where they are produced, into the interior, with the electrons being more mobile ($b_e > b_h$), the surface becomes positively charged and assumes a potential, which is more positive by $\Delta\phi_D$ than the interior.

To correctly understand the meaning of the Dember effect, we examine the case in which the electrons are mobile, while the holes are immobile ($b_h = 0, \sigma_h = 0$). For this case, the Dember voltage reaches its maximum possible value,

$$\Delta\phi_D = \frac{kT}{e} \ln \frac{n_e(0)}{n_e(\infty)} \quad (5.32)$$

For this special case, however, we can easily see that this electrical potential difference cannot be measured as a steady-state voltage. For immobile holes, the electron particle current must vanish together with the charge current under the condition of open circuit during a voltage measurement. From Equation 5.13 the electrochemical potential η_e for the electrons must have the same value everywhere when the electron current is zero. Thus, no measurable voltage is found as

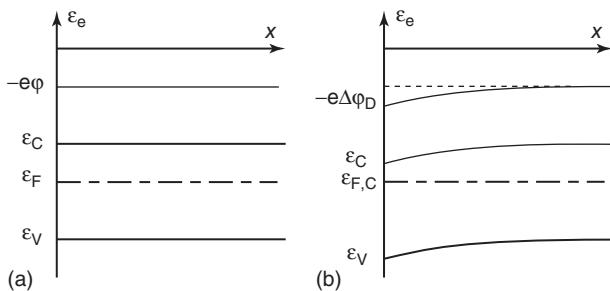


Figure 5.8 Distribution of potentials in the (a) dark state and (b) on illumination of the free left surface of a semiconductor in which only electrons are mobile.

the difference of the electrochemical potentials between the left and right sides. Figure 5.8 shows the distribution of the potentials. The electrical potential difference in Equation 5.32 can be read directly from the separation between the conduction band edge ϵ_C and the Fermi energy ϵ_{FC} in Figure 5.8. Its value in Equation 5.32 is thus based on a constant Fermi energy with the consequence that no voltage is measurable.

A Dember potential difference (we avoid the term voltage, which we reserve for measurable voltages) always arises from an inhomogeneous generation of electrons and holes when these have different mobilities. As expected, the Dember potential difference vanishes when the electrons and holes have the same mobilities and is a maximum when one of the carrier types is immobile.

The Dember effect becomes less important in a semiconductor with metal contacts. A concentration gradient, such as assumed in our example, can only develop at a surface where the recombination rate is not especially high. Taking into account that voltages are measured between metal contacts at which, due to the large recombination rate at this interface, the electron and hole densities remain unchanged even when exposed to light, we then find the potential distribution for immobile holes between two identical metal contacts shown in Figure 5.9. Here, the production of electron–hole pairs has been extended to a small region close to the surface in order to be able to see any changes in the concentration. Once again, we see electrical potential differences between the surface and the interior. Because of the unchanged concentrations at the metal contacts, however, the Dember effect does not produce either an electrical or an electrochemical steady-state potential difference between the metal contacts when only one charge carrier type is mobile.

From this discussion of the Dember effect, we conclude that no electrochemical potential difference arises as a result of inhomogeneous illumination when only one charge carrier type is mobile and when they are distributed according to Fermi statistics over the available states. This restriction to the Fermi distribution is necessary in order to exclude the photoemission of electrons from semiconductors and metals into vacuum (or into other media), which does result in measurable voltages.

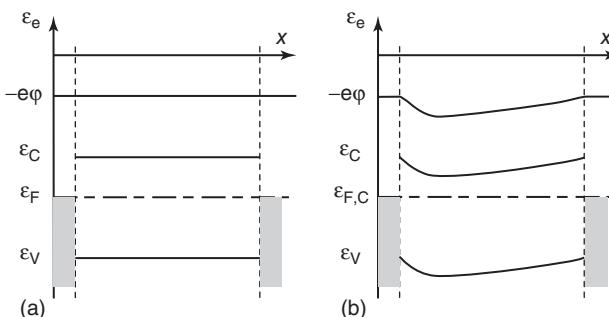


Figure 5.9 Distribution of potentials in a semiconductor with metal contacts in which only electrons are mobile, (a) in the dark state and (b) with the left surface exposed to light.

A voltage deriving from charge carriers behaving according to the Fermi distribution requires at least two different mobile carrier types. This is true, in general. For solar cells, these are electrons and holes; for a battery, these are electrons and ions.

For the preceding discussion, it was important to distinguish between electrical and electrochemical potential differences and we emphasize the essential differences once again.

Electrochemical potential differences are measurable as *steady-state* voltages and are associated with currents. The driving forces for the currents are the gradients of the electrochemical potentials and the voltage drop resulting from a charge current through a resistor in a steady state is an electrochemical potential difference.

By contrast, electrical potential differences, as seen already from their relation to the charge density in Poisson's equation, are related to the charge distribution. Temporal changes to the charge distribution can produce measurable voltages, which are, however, *nonstationary*. As an example, we can take the Dember effect for an illuminated, free surface of a semiconductor without contact, connected on the unilluminated rear side to one terminal of a voltmeter. The second terminal of the voltmeter is connected to a transparent, conductive layer in close proximity to the illuminated surface. In this arrangement, the voltmeter is capacitively coupled to the semiconductor surface. If the surface of the semiconductor becomes positively charged relative to the interior during illumination, the transparent electrode connected by the voltmeter to the rear contact of the semiconductor becomes negatively charged. The charge transport through the voltmeter then produces a voltage called *surface photovoltage*. It depends on the series connection of the capacitance between the transparent electrode and the surface, and the capacitance between the centers of positive and negative charge in the semiconductor. Here, it is important that the voltage shown by the voltmeter is *nonstationary* and decays after the RC time of the voltmeter measurement circuit although the semiconductor surface is still illuminated.

Just as electrochemical potential differences are distributed according to the values of series-connected resistances through which the same current flows, electrical potential differences are distributed according to the (reciprocal) values of series-connected capacitances, which carry the same charge.

5.7

Mathematical Description

We have now gathered all the equations, which, together with the appropriate boundary conditions, allow us to solve all problems of semiconductor devices. These equations are as follows:

- For the charge current

$$j_Q = -e j_e + e j_h \quad (5.33)$$

- For the particle currents

$$j_e = -\frac{\sigma_e}{e^2} \operatorname{grad} \eta_e \quad (5.34)$$

$$j_h = -\frac{\sigma_h}{e^2} \operatorname{grad} \eta_h \quad (5.35)$$

- For the carrier concentrations

$$\frac{\partial n_e}{\partial t} = G_e - R_e - \operatorname{div} j_e \quad (5.36)$$

$$\frac{\partial n_h}{\partial t} = G_h - R_h - \operatorname{div} j_h \quad (5.37)$$

- For the electrical potential

$$\nabla^2 \varphi = -\frac{\rho_Q}{\epsilon \epsilon_0} \quad (5.38)$$

This system of equations must be solved for every point in the solar cell subject to the boundary conditions. This is a complicated numerical procedure. In the discussion of the solar cell, we will not proceed in this direction, since it is rather abstract and conceals the physics behind the solar cell, rather than elucidating it.

For a particular structure, these equations can be solved numerically by semiconductor device simulators. Using them is like making an experiment. Predictions from a theoretical model can thus be examined.

5.8

Problems

- 5.1** Derive the so-called Einstein relation between diffusion coefficient and mobility for particles with concentration n_i , charge $z_i e$ and mobility b_i from Ficks law of diffusion, $j_{Q,i} = -z_i e D_i \operatorname{grad} n_i$ and from the general equation $j_{Q,i} = -(\sigma_i / z_i e) \operatorname{grad} \eta_i$.

- 5.2** A charge current density $j_Q = ej_h = 150 \text{ mA cm}^{-2}$ is flowing in a p-type silicon sample ($n_A = 10^{16} \text{ cm}^{-3}$, assume full ionization of the acceptors) having a thickness of $d = 200 \mu\text{m}$. The direction of the current shall be in the positive x direction from $x = 0$ to $x = d$.
- How large is the (hole) conductivity for a hole mobility of $b_h = 300 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$?
 - The concentration of the holes shall be homogeneous. Determine the electrical potential difference between the two contacts of the semiconductor.
 - If the current is caused by diffusion, how does the concentration profile of the holes look like if their total number remains unchanged? Compare the gradients of the electrochemical potential η_h for both cases.
- 5.3** Assume a mobility of electrons in a semiconductor of $b_e(T = 300 \text{ K}) = 300 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ depending on temperature as $T^{-3/2}$. An electric field of $E = 100 \text{ V cm}^{-1}$ is applied. At which temperature are the drift velocity $\langle v_{\text{drift},e} \rangle$ and the thermal velocity $\langle v_{\text{th},e} \rangle$ related as $\langle v_{\text{drift},e} \rangle = 0.001 \langle v_{\text{th},e} \rangle$? Assume $m_e^* = m_e$.
- 5.4** Intrinsic carrier densities of electrons and holes can be achieved in a doped semiconductor by compensation. Compensation means that the effect of the donors with regard to the carrier concentration in the valence and conduction band is compensated by a proper amount of acceptors. Has this material the same conductivity as an impurity free, intrinsic semiconductor?

6**Basic Structure of Solar Cells**

To properly understand the requirements for a solar cell, namely, to generate current and voltage, we first turn to a similar problem that may be easier to understand, and which demonstrates that the working of a solar cell has more of a chemical than an electrical nature.

6.1**A Chemical Solar Cell**

Figure 6.1 illustrates a hypothetical chemical solar cell in which water (H_2O) is decomposed into hydrogen (H_2) and oxygen (O_2) by the absorption of high-energy photons. As a result, the partial pressures of hydrogen and oxygen in the cell rise above their equilibrium values. If no hydrogen and oxygen are removed, that is, in the open-circuit state under steady-state conditions, the partial pressures reach values at which the reverse reaction, that is, the recombination of hydrogen and oxygen to water, occurs at the same rate as the decomposition reaction. Clearly, the partial pressures decrease with increasing probability of the reverse reaction. With a catalyst, such as a large platinum surface, the reverse reaction would have such a high probability that, even with the additional production of hydrogen and oxygen by decomposition, the partial pressures remain practically at the equilibrium values.

As with the electrons and holes in a solar cell, our aim is to extract the gases separately along with their chemical energy. How do we remove the gases separately from the mixture in the reaction cell? This requires membranes that selectively transmit only one of the gases. A semi-permeable membrane on the left in Figure 6.1 transmits the hydrogen and blocks the oxygen, while another membrane on the right is permeable for oxygen and blocks the hydrogen. If the partial pressure, and, therefore, the chemical potential, of one gas type is greater in the reaction cell than outside, this gas flows through its membrane and out of the cell. For a membrane with good permeability, only a slight difference in the partial pressures or the chemical potentials between the inside and the outside is sufficient for the transport of the gases. Similarly, a gradient of the partial pressure or the chemical potential is required inside the cell as the driving

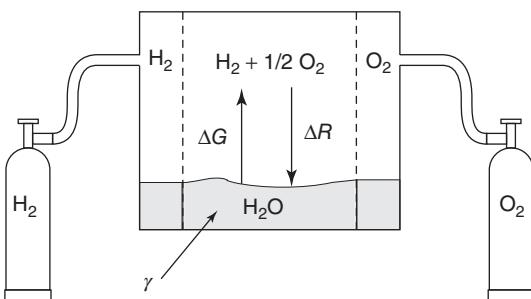


Figure 6.1 Hypothetical chemical solar cell in which water is decomposed into hydrogen and oxygen by the absorption of photons. Hydrogen and oxygen can be separately

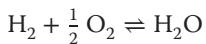
removed through membranes that selectively pass hydrogen on the left and oxygen on the right.

force for the transport of the molecules toward the appropriate membranes. How is this gradient established although the decomposition rate generating the gas molecules is homogeneous? The gradients are an automatic consequence of allowing the molecules to pass through the appropriate membrane, which reduces the partial pressure or chemical potential at the membrane and sets up the gradients in the right direction. The chemical energy obtained from the decomposition of the water is greatest in the open-circuit state, when the gases are not removed and recombine. But in the absence of gas flow, no chemical energy is removed. On the other hand, removing the gases decreases their partial pressures. For the other limiting case, when all of the gases produced are removed and the recombination rate is at its equilibrium value, the partial pressures and chemical potentials are also at their equilibrium values and consequently no chemical energy is supplied with the gas flow. To obtain the maximum chemical energy current, we must therefore accept a certain level of recombination. In the following, we analyze the performance of this chemical solar cell quantitatively.

From the minimum of the free energy for the equilibrium of all components in the chemical cell at constant temperature and volume and without illumination

$$dF = \mu_{H_2} dN_{H_2} + \mu_{O_2} dN_{O_2} + \mu_{H_2O} dN_{H_2O} = 0 \quad (6.1)$$

From the chemical reaction



changes in the number of the molecules are related by

$$dN_{H_2} = 2 dN_{O_2} = -dN_{H_2O}$$

and Equation 6.1 becomes

$$\mu_{H_2} + \frac{1}{2} \mu_{O_2} = \mu_{H_2O} \quad (6.2)$$

which, as we have seen before, could have been read directly from the reaction equation. Equation 6.2 is valid for chemical equilibrium. When the sum of the chemical potentials of the gases is larger than the chemical potential of the water, there is a net production of water from the gases. When the sum of the chemical potentials of the gases is smaller than the chemical potential of the water, there is a net production of the gases from the dissociation of water. The chemical potentials of the gases depend on their concentrations in the same way as for electrons and holes, whereas the chemical potential of the water molecules in the water does not change when the amount of water changes, since the concentration of water molecules stays constant.

$$\mu_{\text{gas}} = \mu_0 + kT \ln \left(\frac{p_{\text{gas}}}{p_0} \right) \quad (6.3)$$

where the proportionality between concentration and partial pressure p was used. The standard pressure p_0 is usually chosen as 1 bar under standard conditions. With the relation in Equation 6.3, a relation for the partial pressures of hydrogen and oxygen in chemical equilibrium with liquid water at 300 K, which is the equivalent of the relation between the concentrations of electrons and holes in an unilluminated semiconductor, $n_e^0 n_h^0 = n_i^2$, is found from Equation 6.2.

$$kT \ln \left(\frac{p_{\text{H}_2}^0 \sqrt{p_{\text{O}_2}^0}}{p_0^{3/2}} \right) = \mu_{0,\text{H}_2\text{O}} - \mu_{0,\text{H}_2} - \frac{1}{2} \mu_{0,\text{O}_2} = -2.46 \text{ eV} \quad (6.4)$$

The value on the right-hand side of Equation 6.4 is found in books on physical chemistry. As for the equilibrium concentrations of electrons or holes by doping, any one of the partial pressures can be chosen. When the partial pressure of oxygen is $p_{\text{O}_2}^0 = 0.2$ bar as in air, the equilibrium partial pressure of hydrogen is only $p_{\text{H}_2}^0 = 1.8 \times 10^{-41}$ bar. If the hydrogen partial pressure is reduced below this value, the recombination rate of the gases is smaller than the thermal rate of splitting of water molecules, until the equilibrium partial pressure is reestablished. Since 1 H₂ molecule reacts with $\frac{1}{2}$ O₂ molecule to give 1 molecule of H₂O, the recombination rate of hydrogen is

$$R_{\text{H}_2} \sim p_{\text{H}_2} \sqrt{p_{\text{O}_2}}$$

This recombination rate for nonequilibrium partial pressures can be expressed by the equilibrium recombination rate

$$R_{\text{H}_2} = R_{\text{H}_2}^0 \frac{p_{\text{H}_2} \sqrt{p_{\text{O}_2}}}{p_{\text{H}_2}^0 \sqrt{p_{\text{O}_2}^0}} \quad (6.5)$$

When the generation rate is increased from the equilibrium generation rate $G_{\text{H}_2}^0$ by ΔG_{H_2} by illumination, hydrogen may flow out from the reaction cell through

its membrane. The hydrogen current, arbitrarily counted negative if it is flowing out of the cell, is

$$\begin{aligned} j_{H_2} &= 2j_{O_2} = R_{H_2} - G_{H_2} \\ j_{H_2} &= 2j_{O_2} = R_{H_2}^0 \frac{p_{H_2} \sqrt{p_{O_2}}}{p_{H_2}^0 \sqrt{p_{O_2}^0}} - G_{H_2}^0 - \Delta G_{H_2} \\ j_{H_2} &= 2j_{O_2} = R_{H_2}^0 \left(\frac{p_{H_2} \sqrt{p_{O_2}}}{p_{H_2}^0 \sqrt{p_{O_2}^0}} - 1 \right) - \Delta G_{H_2} \end{aligned} \quad (6.6)$$

The interpretation is simple. Hydrogen and oxygen flow into the cell in a steady state only if more molecules in the cell disappear by recombination than are generated by water splitting. On the other hand, molecules flow out of the cell only if more are generated in the cell than recombine. If we assume that the gradients of the partial pressures required to drive the currents are small, the partial pressures in the cell and in the gas bottles in Figure 6.1 are approximately the same, as is also shown in Figure 6.2. As a result, the currents of hydrogen and oxygen are limited only by the rates of their chemical reaction and not by transport resistances.

The hydrogen current flowing between the hydrogen bottle and the cell is shown as a function of the hydrogen pressure in the hydrogen bottle, while the oxygen pressure is kept constant (Figure 6.3). The current changes with the hydrogen pressure only because of the recombination rate. The generation rate, in the dark and under illumination does not depend on the hydrogen pressure. At zero hydrogen pressure in the cell and outside, there is no recombination and all molecules generated flow out of the cell. The open-circuit pressure at zero current, when the recombination rate equals the generation rate, is the equilibrium pressure in the dark and a larger pressure with illumination. For still larger pressures, more molecules disappear by recombination than are generated. They are replenished by a positive current flowing into the cell.

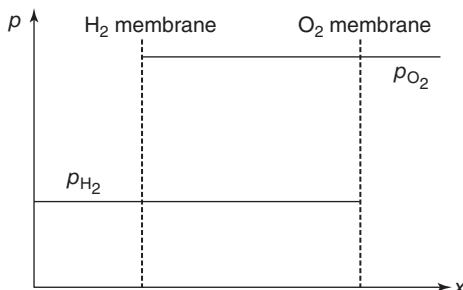


Figure 6.2 The conductivities for hydrogen in the hydrogen membrane on the left and for oxygen in the oxygen membrane on the right, are assumed to be large. The gradients

of the partial pressures required to drive the currents are therefore small and do not show up in the spatial distribution of the pressures.

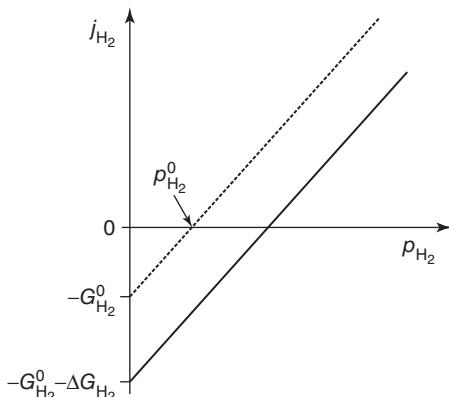


Figure 6.3 Current of hydrogen, positive if flowing from the hydrogen bottle in Figure 6.1 into the cell as a function of the hydrogen partial pressure, without

illumination (broken line) and with additional generation ΔG_{H_2} by illumination (solid line). The oxygen pressure is kept constant.

We could also plot the hydrogen current as a function of the chemical potential of the hydrogen while keeping the chemical potential of the oxygen constant. This plot has the advantage that the chemical energy exchanged between the gas bottles and the cell shows up as a rectangle as for the electron–hole current in Figure 4.4. With Equations 6.2 and 6.3, the recombination rate can be written as

$$R_{H_2} = R_{H_2}^0 \exp\left(\frac{\mu_{H_2} + 1/2 \mu_{O_2} - \mu_{H_2O}}{kT}\right) \quad (6.7)$$

and the dependence of the hydrogen currents on the chemical potentials is

$$j_{H_2} = 2j_{O_2} = R_{H_2}^0 \left[\exp\left(\frac{\mu_{H_2} + 1/2 \mu_{O_2} - \mu_{H_2O}}{kT}\right) - 1 \right] - \Delta G_{H_2} \quad (6.8)$$

which is shown in Figure 6.4.

From the discussion above, we draw an important conclusion. The driving forces for the transport of hydrogen and oxygen, the gradients of their partial pressures, are not selective, they drive both particles, hydrogen and oxygen, towards both outlets and in fact, the driving forces in the wrong direction are much larger than in the preferred direction. In order to prevent the flow of particles in the wrong direction, hydrogen through the oxygen membrane for instance, the oxygen membrane must have a very small conductivity for hydrogen, thereby effectively blocking the transport of hydrogen into the wrong direction, even in the presence of a large driving force. All these properties are important for solar cells, as well.

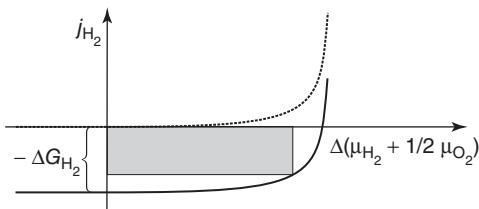


Figure 6.4 Current of hydrogen, positive if flowing from the hydrogen bottle in Figure 6.2 into the cell as a function of the deviation $\Delta(\mu_{H_2} + 1/2\mu_{O_2})$ of the chemical potentials of hydrogen and oxygen from their equilibrium values, without illumination (broken line) and with additional generation

ΔG_{H_2} by illumination (solid line). The chemical potential of water μ_{H_2O} is not changed by illumination. A smaller and more realistic equilibrium generation rate $G_{H_2}^0$ than in Figure 6.3 is assumed. The shaded rectangle is the largest current of chemical energy delivered by the cell.

6.2

Basic Mechanisms in Solar Cells

We now turn to a semiconductor solar cell in which, instead of hydrogen and oxygen, additional electrons and holes are generated by illumination. This is the only difference from the chemical cell in the last section. The thorough discussion of the chemical cell was intended to show the parallels and to emphasize the importance of the chemical reaction between electrons and holes in semiconductors.

What do we learn from the chemical solar cell? The most important point is that semi-permeable membranes are required on both sides of the absorber, such that electrons can only flow out to the left and holes to the right, forcing them to generate a charge current flowing from left to right in the cell.

A membrane that is permeable for electrons, and which blocks the holes, would be a material that has a large conductivity for electrons and a small conductivity for holes. We already know that an n-type semiconductor has this property. However, different from the hydrogen membrane, which has a very small mobility for oxygen, the hole conductivity in an n-type semiconductor is small because of a small concentration of holes. To function properly as an electron membrane, we have to make sure that holes are not injected from the absorber. A hole membrane is, of course, a p-type semiconductor, into which electrons must not be injected. Figure 6.5 shows a solar cell structure which has the required properties. The injection of holes into the electron membrane on the left is prevented by giving it a larger band gap, resulting in an energy barrier in the valence band for the holes. In the same way, a larger band gap of the hole membrane on the right, combined with a smaller electron affinity χ_e causes an energy barrier in the conduction band for the electrons. Owing to their larger band gap, the membranes transmit almost all of the photons to be absorbed in the absorber. As for the partial pressure distribution in Figure 6.3, transport resistances for electrons and holes are assumed to be small (which is realistic) and hence, gradients of the Fermi energies required to drive the electrons to the left and the holes to the right are negligibly small.

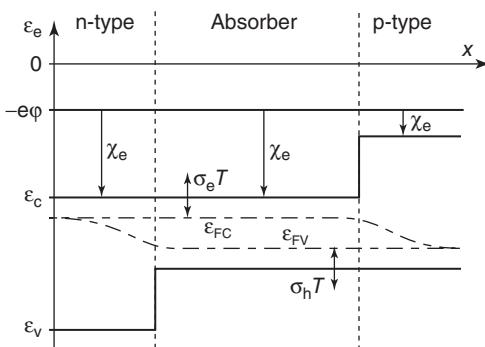


Figure 6.5 A n-type electron membrane on the left allows electrons generated in the absorber by illumination, to flow to the left, while blocking the holes. A p-type hole membrane on the right allows holes to flow

to the right, blocking the electrons. Electrons are driven by an invisibly small gradient of $-\epsilon_{FC}$, holes are driven by an invisibly small gradient of ϵ_{FV} .

Metal contacts are assumed to make contact with the membranes, which forces the Fermi energies to join at the surface. The large gradient of ϵ_{FV} in the electron membrane on the left does not lead to an appreciable hole current to the left because of the very small hole concentration in the electron membrane, obvious from the large distance of ϵ_{FV} from the valence band. Accordingly, the same holds for the electrons in the hole membrane.

We have to keep in mind that the driving forces existing in a solar cell are not selective in the sense that they would drive electrons towards one contact, the electron contact, and holes to the other, the hole contact. On the contrary, the driving forces, the gradients of the quasi-Fermi energies, drive electrons and holes towards both contacts. And even worse, the driving forces in the wrong direction driving electrons towards the hole contact and holes towards the electron contact are much larger than the driving forces for the desired directions. The only way to have the currents flowing in the desired directions needs manipulation of the conductivities, a large electron conductivity and small hole conductivity in front of the electron contact and a large hole conductivity and small electron conductivity in front of the hole contact. For an optimal solar cell the necessary degree of selectivity of the conductivities can be estimated with the following assumptions:

- 1) The difference of the quasi-Fermi energies at maximum power is on the order of 1 eV
- 2) The voltage drop for electron- transport to the electron contact, hole-transport to the hole contact shall not be larger than 1 meV, because that is a voltage loss over a series resistance in the current-voltage characteristic.
- 3) Under maximum power conditions, the current of electrons and holes into the undesired directions shall be smaller by a factor of 1000 than into the desired directions. This ensures that nearly all electrons and holes that leave the solar cell contribute to the charge current.

As a result of point 1 and 2, the changes of the quasi-Fermi energies on the order of 1 eV for the holes occur near the electron contact and for the electrons near the hole contact, whereas the change of 1 meV for the electrons towards the electron contact and for the holes towards the hole contact occur over the thickness of the solar cell. The driving forces in the wrong direction are thus larger by a factor of around 10^6 than the driving forces for the desired transport.

Although the driving forces for the undesired carrier transport are larger by a factor of around 10^6 , the currents in these directions shall be smaller by a factor of 1000 than the currents in the desired directions. In order to achieve this, the conductivities of the holes in front of the electron contact and of the electrons in front of the hole contact must be smaller by a factor of 10^9 than the conductivities of the electrons in front of the electron contact and of the holes in front of the hole contact.

Such a ratio of the conductivities is present in a pn - junction in the dark. If a region close to the electron contact is doped with donors resulting in an electron concentration of 10^{16} cm^{-3} , the hole concentration in this region is $n_h = n_i^2/n_e$, which would be 10^4 cm^{-3} in silicon. The same low concentration of the electrons follows from a doping of a region close to the hole contact with 10^{16} cm^{-3} acceptors. This is why microelectronic devices made from silicon work so well, in the dark. In a solar cell, the situation is much more severe. The solar cell is illuminated and the carrier concentration of electrons and holes should be as large as possible. Since a pn – junction is made of the same material everywhere between the contacts, the absorption properties are also the same everywhere. As a result, it is difficult to achieve the required large ratio of the conductivities close to the contacts. These difficulties are avoided by the structure shown in Figure 6.5.

In order to see, whether the structure in Figure 6.5 is an ideal solar cell, we calculate its performance by the semiconductor device simulator SENTaurus TCAD from Synopsys Inc. For the simulation we use a hypothetical non-doped absorber material with a bandgap of 1.34 eV, for which Figure 7.3 predicts the highest efficiency. The generation rate of electron-hole pairs is assumed to be homogeneous over the thickness of the absorber of 10 μm with a magnitude equivalent to the absorption of all absorbable photons from the AM1.5G spectrum. The electron and hole membranes on the left and right in Figure 6.5 both have a thickness of 100 nm, a bandgap of 3 eV and are doped with 10^{15} cm^{-3} donors or acceptors, respectively. Generation of electron-hole pairs in the membranes is neglected as is the recombination at the interface between the membranes and the absorber. In all three materials, electrons and holes have a mobility of $100 \text{ cm}^2/\text{Vs}$. The work functions of the contact metals are equal to the work functions of the membranes. The surface recombination velocity at the interface between contacts and membranes are assumed to be infinite to ensure that there is only one Fermi-distribution in the metal and at the interface.

The results of the calculation confirm our expectations. The short-circuit current density is 35 mA/cm², the open-circuit voltage is 1082 mV, the fill factor is 88.9 resulting in an efficiency of 33.7 % [12]. All these values agree with the theory developed in Chapter 7.

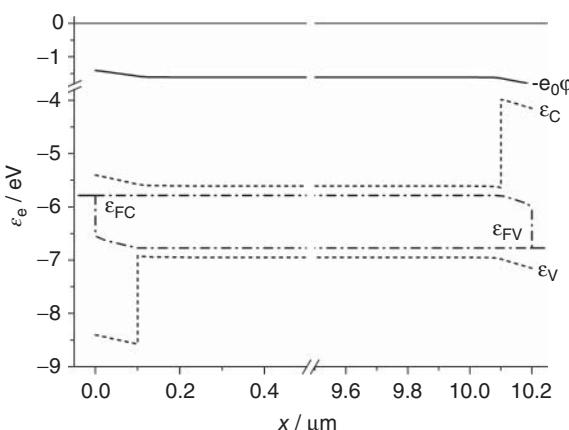


Figure 6.6 Energy diagram for the hetero-structure solar cell in Figure 6.5 under AM1.5 illumination and at maximum power. The bandgap is 1.34 eV, the voltage $V_{mp} = 987 \text{ mV}$ the current $j_{mp} = 34.1 \text{ mA/cm}^2$ and the efficiency $\eta = 33.7 \%$.

Figure 6.6 shows the energy diagram for the maximum power situation. A very remarkable feature is seen for the electrical energy per electron -eq in the membranes. It increases towards the electron contact and decreases towards the hole contact. This means that electrons in the electron membrane and holes in the hole membrane move against their electrical driving force.

For the conversion of chemical energy into electrical energy in a solar cell we have found a structure in which the electrons and holes flow outward through different contacts. This structure is virtually identical with the structure for the removal of hydrogen and oxygen from the chemical reaction cell. In the following, we discuss different arrangements that fulfill the same requirements with different degrees of perfection. Special emphasis is given to the pn-junction because of its technical importance and its model character for other devices, although it is not quite as ideal for a solar cell as the heterostructure in Figure 6.5.

6.3

Dye Solar Cell

A very good example of the solar cell structure required, with membranes for electrons on one side and for holes on the other side, is the electrochemical dye solar cell, shown in Figure 6.7 [13]. The “semiconductor” in which the absorption of photons produces electron–hole pairs is a dye layer. Since the electrons and holes in the dye layer have very small mobilities, this layer must be very thin for the charge carriers to reach the membranes within their lifetimes.

The dye is applied as a monomolecular layer to a good electron conductor in the form of TiO_2 . As Figure 6.7 illustrates, the electrons in the dye reach the conduction band of the TiO_2 without difficulty. However, owing to the large band gap of more than 3 eV of TiO_2 , the holes present in the dye encounter a high barrier for the transition to the valence band of the TiO_2 .

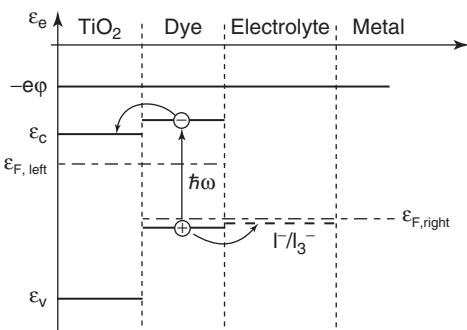


Figure 6.7 Dye solar cell in which the electron–hole pairs are produced in a ruthenium bipyridine dye. The electrons flow outward toward the left through the

n-conductor TiO₂ and the holes toward the right through the triiodide ions with which the acetonitrile electrolyte is doped.

On the other hand, the thin, only monomolecular dye layer required for efficient charge transfer has the disadvantage that the absorption of photons in one layer is very poor, since their penetration depth $1/\alpha$ into the dye is much larger than the thickness of the dye layer. In order to compensate for this disadvantage, the TiO₂ layer is composed of particles, only a few nanometers in size, in a porous structure. All TiO₂ particles are coated with the dye on their free surfaces, so that complete absorption of the photons is achieved with the many dye layers encountered by the photons.

The porous structure, however, greatly complicates the contact between the dye and a p-type hole membrane, through which the holes can flow outwards. This problem is solved with the use of an electrolyte that penetrates into all the pores. The iodide ions of the redox system (I^-/I_3^-) provide for charge transport. The energy of an electron in I^- differs only slightly from the energy of an electron in the ground state of the dye, so that the flow of holes from the dye to the electrolyte is unproblematic. The flow of the excited electrons from the dye into the electrolyte is, however, prevented since the electrolyte has no states at the energy of the excited electrons.

At first glance, this electrochemical cell would appear almost ideally to fulfill the requirements for the selective transport of electrons toward the left in the TiO₂ and of holes toward the right in the electrolyte. Besides direct recombination via the direct transition of an electron from its excited state in the dye to its ground state, there is also recombination via an indirect path, which we can regard as an internal shunt. Excited electrons that flow to the conduction band of the TiO₂ can also reach the ground state of the dye via surface states of the TiO₂ particles, either directly or, more probably, after a transition to the redox system. This back reaction of the electrons is facilitated by the large surface area of the TiO₂, which is, however, required in order to combine sufficient absorption with the poor transport properties of the dye. It is therefore not certain whether the dye solar cell represents a solar cell suitable for practical application. The absorption

of the dye needs improvement by extending it over a greater spectral region. Problems that remain to be clarified include its stability over a period of 20 years with respect to decomposition reactions and the possibility of the electrolyte leaking out or drying out. A solid-state hole conductor would be preferable, but a way has to be found to fill it into all the tiny pores.

6.4

The pn-Junction

A good, although not ideal, realization of the structure of a solar cell as discussed in the previous section can be found in commercially available solar cells made of crystalline silicon. A p-region about $300\text{ }\mu\text{m}$ thick, moderately doped with an acceptor concentration of $n_A = (10^{15} - 10^{16})\text{ cm}^{-3}$, as the absorber is sandwiched between a less than $1\text{ }\mu\text{m}$ thick, highly doped n-layer on the illuminated side, as the electron membrane and another thin, highly doped p-layer on the rear side, as the hole membrane. All of this consists of crystalline silicon. The pn-junction, formed by the electron membrane and the absorber, near which most electrons and holes are generated, is especially important for the solar cell and other devices, and is therefore the subject of a more detailed discussion.

6.4.1

Electrochemical Equilibrium of Electrons in a pn-Junction in the Dark

In thermal equilibrium with the environment, including the 300 K background radiation, no current may flow in the pn-junction if there is no external energy source. This means that

- 1) $j_Q = 0$
and as a result of chemical equilibrium with the 300 K radiation,
- 2) $\eta_e + \eta_h = \mu_\gamma = 0$
From Equation 5.13, j_Q is given by

$$j_Q = \frac{\sigma_e}{e} \operatorname{grad} \eta_e - \frac{\sigma_h}{e} \operatorname{grad} \eta_h = 0$$

From (2) above, it follows that $\operatorname{grad} \eta_e = -\operatorname{grad} \eta_h$ and therefore

$$j_Q = \frac{\sigma_e + \sigma_h}{e} \operatorname{grad} \eta_e = 0$$

Since $\sigma_e + \sigma_h \neq 0$, grade $\eta_e = 0$.

This tells us that, in the dark, the electrochemical potential η_e of the electrons (as well as the electrochemical potential η_h of the holes) has the same value everywhere in the pn-junction. This is the implication of electrochemical equilibrium between electrons in the n-region and electrons in the p-region. Indicating values in the p-region far away from the pn-junction by a superscript p and in the n-region far away from the pn-junction by a superscript n, we have

$$\eta_e^p = \mu_{e,0}^p + kT \ln \frac{n_e^p}{N_C} - e\varphi^p = \eta_e^n = \mu_{e,0}^n + kT \ln \frac{n_e^n}{N_C} - e\varphi^n \quad (6.9)$$

Since only a small fraction of the semiconductor atoms is replaced by doping atoms, the chemical environment for a free electron, which determines the value of $\mu_{e,0}$, remains unchanged. For $\mu_{e,0}^p = \mu_{e,0}^n$ the difference in the electrical potentials between the p-region and the n-region is, from Equation 6.9,

$$\varphi^n - \varphi^p = \frac{kT}{e} \ln \frac{n_e^n}{n_e^p}$$

Using $n_e^n = n_D$ and $n_e^p = n_i^2/n_A$, this potential difference, called the *diffusion voltage*, is

$$\varphi^n - \varphi^p = \frac{kT}{e} \ln \frac{n_D n_A}{n_i^2} \quad (6.10)$$

We can visualize this potential difference as arising in the following way. We start with the spatially separated and electrically neutral p-conductor and the n-conductor being at the same electrical potential. When making contact, the greater chemical potential of the electrons in the n-conductor (and of the holes in the p-conductor) drives a diffusion current of electrons from the n-conductor to the p-conductor and a hole diffusion current from the p-conductor to the n-conductor. This builds up a positive charge in the n-conductor and a negative charge in the p-conductor. The diffusion currents continue to flow until an electrical potential difference $\varphi^n - \varphi^p$ is established for which $\eta_e^p = \eta_e^n$, so that $\text{grad } \eta_e = 0$ and $\text{grad } \eta_h = 0$. The driving force then no longer exists, and the particle currents, and, with them, the charge current can no longer flow.

For electrochemical equilibrium, in which the gradient of the electrical energy is compensated by the gradient of the chemical energy, the electrons and holes do not experience any forces. Yet, they are mobile in random Brownian motion. It is important to recognize that the equilibrium situation of the pn-junction is an electrochemical equilibrium of the electrons and the holes. An electrochemical equilibrium of the donors and acceptors does not exist. In fact, the pn-junction is based on the non-equilibrium distribution of the donors and acceptors.

6.4.2

Potential Distribution across a pn-Junction

The potential difference $\varphi^n - \varphi^p$ is the result of the electrochemical equilibrium of the electrons in the n- and p-regions. The distribution of this potential difference over the n- and p-conductors follows from the relationship with the distribution of the charge density ρ_Q .

From the Maxwell equation

$$\text{div } D = \rho_Q$$

with $D = \epsilon\epsilon_0 E$ and with $E = -\text{grad } \varphi$, we find Poisson's equation

$$\text{div } E = -\text{div grad } \varphi = -\nabla^2 \varphi = \frac{\rho_Q}{\epsilon\epsilon_0} \quad (6.11)$$

The interface between the n- and p-regions is assumed to extend much farther in the y - and z -directions than the width of the space charge layer to be derived. This allows a one-dimensional treatment, giving

$$\frac{d^2\varphi}{dx^2} = -\frac{\rho_Q}{\epsilon\epsilon_0} \quad (6.12)$$

In the n-region the density of the space charge is

$$\rho_Q^n(x) = e(n_D^+ - n_e(x)) = e n_D^+ \left(1 - \exp \left\{ \frac{e[\varphi(x) - \varphi^n]}{kT} \right\} \right) \quad (6.13)$$

This relationship follows from Figure 6.8, which schematically depicts the potential distribution and the charge distribution.

Unfortunately, however, for this space charge density, Poisson's equation can only be solved numerically. According to Schottky, we can approximate the charge distribution by a spatially constant positive space charge density in a region up to w_n , as yet, undefined depth w_n on the n-side. A constant negative space charge density is assumed up to w_p in the p-region, which compensates the positive space charge in the n-region, as shown in Figure 6.8

$$\begin{aligned} \rho_Q^n &= e n_D^+ \approx e n_D & \text{for } -w_n < x \leq 0 \\ \rho_Q^p &= -e n_A^- \approx -e n_A & \text{for } 0 \leq x < w_p \end{aligned}$$

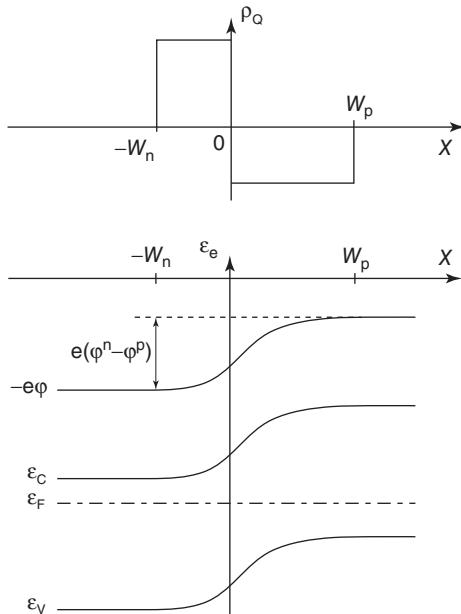


Figure 6.8 Distribution of the space charge density ρ_Q and of the electrical energy $-e\varphi$ per electron in a pn-junction.

The sum of the charges $Q_n = e n_D w_n$ in the n-region and $Q_p = -e n_A w_p$ in the p-region is $Q_n + Q_p = 0$, and the electric field strength differs from zero only between $-w_n$ and w_p . From the sum of the charges

$$w_p = \frac{n_D}{n_A} w_n \quad (6.14)$$

and the entire thickness of the space charge layer is

$$w = w_n + w_p = \left(1 + \frac{n_D}{n_A}\right) w_n \quad (6.15)$$

For the boundary conditions $E(-w_n) = 0$ for the electric field and $\varphi(-w_n) = \varphi^n$ for the electrical potential, integration of Poisson's equation 6.12 with $\rho_Q = \rho_Q^n$ in the range $-w_n < x \leq 0$ of the n-region yields

$$\varphi_n(x) = -\frac{en_D}{2\epsilon\epsilon_0} (x + w_n)^2 + \varphi^n \quad (6.16)$$

and for the boundary conditions $E(w_p) = 0$ and $\varphi(w_p) = \varphi^p$ integration with $\rho_Q = \rho_Q^p$ in the range $0 \leq x < w_p$ of the p-region yields

$$\varphi_p(x) = \frac{en_A}{2\epsilon\epsilon_0} (x - w_p)^2 + \varphi^p \quad (6.17)$$

For the given charge distribution the potential is continuous everywhere including at $x = 0$. From $\varphi_n(0) = \varphi_p(0)$ it thus follows that

$$\varphi^n - \varphi^p = \frac{e}{2\epsilon\epsilon_0} (n_D w_n^2 + n_A w_p^2) \quad (6.18)$$

With reference to Equation 6.14 and 6.15, we can now determine the total thickness w of the space charge layer, since $\varphi^n - \varphi^p$ is known from the electrochemical equilibrium according to Equation 6.10:

$$w = \sqrt{\frac{2\epsilon\epsilon_0}{e} \frac{n_A + n_D}{n_A n_D} (\varphi^n - \varphi^p)} \quad (6.19)$$

For the asymmetrical pn-junction of a Si solar cell, with $n_D = 10^{19} \text{ cm}^{-3}$ and $n_A = 10^{16} \text{ cm}^{-3}$, Equation 6.14 shows that the space charge layer in the p-region is much thicker than in the n-region. Its extension into the p-region w_p is essentially equal to the total thickness w which, for the doping levels above, has a value of $w = 0.35 \mu\text{m}$. Over this distance, the change in potential is $\varphi^n - \varphi^p = 0.9 \text{ V}$ as follows from Equation 6.10. At least for crystalline silicon, the extension of the space charge layer is therefore much smaller than the penetration depth of the photons and the diffusion lengths.

From the discussion of the recombination mechanisms in Chapter 3, we know that high doping levels favor Auger recombination and should be avoided in a solar cell. Are there minimum doping levels required for a pn-junction solar cell? In the discussion of the surface recombination in Section 3.6.2 we have seen that a semi-permeable membrane which functions properly should have a conductivity for the desired charge carriers to pass, which is much larger than for the carriers

to be blocked. If we assume that the mobilities of electrons and holes are roughly the same, this means that in the n-part of the pn-junction the concentration of electrons must be much larger than the hole concentration under illumination. Likewise, the hole concentration must be much larger than the electron concentration under illumination in the p-part.

Since the dark concentrations of the majority carriers are equal to the doping concentrations, this condition is expressed as

$$n_D \gg n_h^n \quad \text{and} \quad n_A \gg n_e^p$$

This is equivalent to the condition

$$kT \ln [(n_D n_A / n_i^2)] = e(\varphi^n - \varphi^p) > \varepsilon_{F,C} - \varepsilon_{F,V} = kT \ln [n_e^p n_h^n / n_i^2] \quad (6.20)$$

Suitable doping is required to correctly match the potential difference ($\varphi^n - \varphi^p$) in the dark to the expected chemical energy per electron–hole pair during illumination. In Figure 6.5 we have chosen $e(\varphi^n - \varphi^p) = \varepsilon_{F,C} - \varepsilon_{F,V}$, so that the electrical potential difference between the p-conductor and the n-conductor membranes just vanishes during illumination. By contrast in Figure 5.6 $e(\varphi^n - \varphi^p) < \varepsilon_{F,C} - \varepsilon_{F,V}$. The electrical potential difference vanishes in this case as well for the assumed illumination level, but the photo-voltage measured at the contacts is smaller than the difference between the Fermi energies inside the cell: hence chemical energy is wasted. Figure 6.9 gives the potential distribution for a more strongly doped cell, for which $e(\varphi^n - \varphi^p) > \varepsilon_{F,C} - \varepsilon_{F,V}$. The chemical energy per electron–hole pair is now completely converted into electrical energy. Due to the increasing probability of Auger recombination with increased doping, however, the doping density should not be larger than necessary.

In summary, there are two requirements for solar cells, which contradict each other in a pn-homo-junction. One requirement is a large increase of the minority carrier concentration by illumination, holes in the n-region and electrons in the p-region in order to generate a large difference of the quasi-Fermi energies and with it the potential for a large voltage. The second requirement is a large ratio of the conductivities to ensure a large current by preventing that charge carriers move

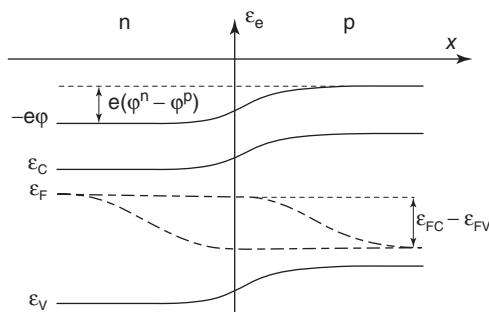


Figure 6.9 Potential distribution in an illuminated solar cell in which the diffusion voltage is $e(\varphi^n - \varphi^p) > \varepsilon_{F,C} - \varepsilon_{F,V}$.

into the wrong direction, which calls for small concentrations of minority carriers, in contradiction to the first requirement. For these reasons, the pn-junction is not an ideal structure for a solar cell. The condition (6.20) is a necessary requirement, but in view of the increasing probability for non-radiative recombination with increasing doping concentrations, it is not sufficient to guarantee good performance of a solar cell.

6.4.3

Current–Voltage Characteristic of the pn-Junction

For the charge current through a pn-junction we distinguish between the forward direction, in which the electrons of the n-region and the holes of the p-region flow toward the pn-junction, and the reverse direction, in which electrons and holes flow away from the pn-junction.

In the forward direction, shown in Figure 6.10a, both the electrons coming from the n-region and the holes coming from the p-region move as minority carriers into the oppositely doped region, where they recombine after an average path length of one diffusion length. More than a diffusion length away from the pn-junction, the minority carrier concentration is much smaller than the majority carrier concentration, both in the dark and with illumination (weak excitation), so that the charge current is carried only by the majority carriers, in the n-region by electrons and in the p-region by holes.

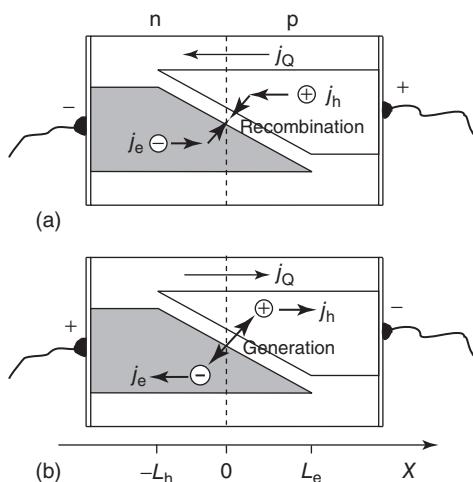


Figure 6.10 Electron and hole currents in a pn-junction. (a) For a negative polarity of the n-region with respect to the p-region, i.e. in the forward direction, electrons and holes flow toward the pn-junction, where they

recombine. (b) In the reverse direction, for a positive polarity of the n-region, electrons and holes flow away from the pn-junction where they are produced.

In the reverse direction, shown in Figure 6.10b, electrons come from the p-region and holes from the n-region. Since the charge current in the p-region is made up entirely of holes, no electrons are transported through the p-region. The electrons emerging from the p-region must have been produced there. However, only those electrons generated in the p-region that have not recombined before can reach the n-region. These electrons must have been generated at a distance not farther than a diffusion length from the n-region. For the same reason, only those holes generated at a distance not farther than a diffusion length from the p-region can reach the p-region. For both the forward and the reverse direction, the charge of the charge current is transferred from the electrons to the holes within a diffusion length on both sides of the pn-junction.

Outside an electron diffusion length L_e to the right or a hole diffusion length L_h to the left of the pn-junction, the charge current is a pure electron current in the n-region and a pure hole current in the p-region. The charge current is then given by integrating over the contributions to the hole current (alternatively, the contributions to the electron current). If the forward charge current is arbitrarily counted as positive (electrons and holes flow toward the pn-junction),

$$j_Q = -e \int_{-L_h}^{L_e} \operatorname{div} j_h \, dx \quad (6.21)$$

since $j_h = 0$ for $x < -L_h$, but $j_h = j_Q/e$ for $x > L_e$.

From the continuity equation for holes under steady state conditions

$$\frac{\partial n_h}{\partial t} = G_h - R_h - \operatorname{div} j_h = 0, \quad \operatorname{div} j_h = G_h - R_h$$

Again, we divide the generation rate into the components G_h^0 in the dark and ΔG_h from the illumination

$$G_h = G_h^0 + \Delta G_h$$

The recombination rate for radiative recombination is accordingly

$$R_h = R_h^0 \frac{n_e n_h}{n_i^2} = R_h^0 \exp\left(\frac{\eta_e + \eta_h}{kT}\right) \quad (6.22)$$

where, because of $\eta_e + \eta_h = 0$ in equilibrium with the 300 K photons and phonons,

$$G_h^0 = R_h^0$$

We can now write the charge current of Equation 6.21 as

$$j_Q = -e \int_{-L_h}^{L_e} \left\{ G_h^0 \left[1 - \exp\left(\frac{\eta_e + \eta_h}{kT}\right) \right] + \Delta G_h \right\} dx \quad (6.23)$$

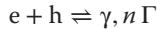
In principle, the sum of the electrochemical potentials $\eta_e + \eta_h$ is a function of the position and depends on the resistances that limit the charge current.

Resistances, however, occur as two different types:

- 1) the familiar type, when the transport resistance is limiting the current. Then

$$j_Q = -\frac{\sigma}{e} \operatorname{grad} \eta$$

- 2) the less familiar type, when the current is limited by the resistance of a chemical reaction as in the reaction of hydrogen and oxygen in the chemical solar cell in Section 6.1 and of electrons and holes with photons and phonons in a semiconductor cell



In this case,

$$\operatorname{div} j = G - R$$

If the chemical reaction limits the current, as it does in the chemical solar cell, no more electrons and holes can flow away from the pn-junction than are produced there, and no more can flow toward the pn-junction than disappear there as a result of recombination.

We can estimate the voltage drop across the transport resistance. The charge current through a solar cell is limited by the absorbed photon current and has a maximum of 42 mA cm^{-2} for silicon in nonfocused solar radiation. For a doping concentration of $n_A = 10^{16} \text{ cm}^{-3}$ and a mobility of $b_h = 470 \text{ cm}^2 (\text{Vs})^{-1}$, the conductivity is $\sigma_h = 0.75 \Omega^{-1} \text{ cm}^{-1}$. The voltage drop is then $1/e \operatorname{grad} \eta_h = J_Q / \sigma_h = 56 \text{ mV cm}^{-1}$, for a thickness of $300 \mu\text{m}$, thus less than 2 mV. This is negligible compared with $(\eta_e + \eta_h)/e$, which is of the order of 1 V. We conclude that the transport resistance is negligible. The current through a pn-junction is, instead, limited by the reaction resistance.

We therefore have

$$\operatorname{grad} \eta_h \approx 0 \quad \text{for } x > -L_h$$

and

$$\operatorname{grad} \eta_e \approx 0 \quad \text{for } x < L_e$$

so that

$$\eta_e + \eta_h \neq f(x) \quad \text{for } -L_h < x < L_e$$

It then follows that $\eta_e + \eta_h = eV$ where V is the voltage between the terminals for the n- and p-regions.

From the considerations above, it follows that the exponential function in the integrand in Equation 6.23 is constant over the range of the integration limits. This greatly simplifies the integration of Equation 6.23, and we obtain the current–voltage characteristic of the pn-junction as

$$j_Q = e G_h^0 (L_e + L_h) \left[\exp \left(\frac{eV}{kT} \right) - 1 \right] - e \int_{-L_h}^{L_e} \Delta G_h \, dx \quad (6.24)$$

An external short circuit ($V = 0$) defines the short-circuit current density j_{SC}

$$j_Q = -e \int_{-L_h}^{L_e} \Delta G_h \, dx = -e \int_{-L_h}^{L_e} \Delta G_e \, dx = j_{SC} \quad (6.25)$$

In the dark ($\Delta G_{e,h} = 0$) and for large negative voltages ($\exp(eV/kT) \ll 1$), we find the reverse saturation current j_S , which is independent of the voltage,

$$j_Q = -eG_{e,h}^0 (L_e + L_h) = -j_S \quad (6.26)$$

The short-circuit current j_{SC} and the reverse saturation current j_S are the essential elements of the current–voltage characteristic of a pn-junction

$$j_Q = j_S \left[\exp\left(\frac{eV}{kT}\right) - 1 \right] + j_{SC} \quad (6.27)$$

Using Equation 6.26, we can calculate the reverse saturation current for an ideal pn-junction, in which electron–hole pairs are produced only by the absorption of 300 K background radiation using the generation rate in Equation 3.54 and the lifetime from Section 3.6.3.

In real pn-junctions, we must also consider the electron–hole generation by nonradiative transitions as the reverse process of nonradiative recombination. The rate of generation, then, cannot be given in general form. However, it can be expressed in terms of the diffusion lengths, since the lifetimes that these contain represent a parameter that reflects the real recombination processes. For equilibrium between generation and recombination, the lifetime (of the minority carriers) together with the concentration of the minority carriers determines the generation rate, so that

$$G_{e,h}^0 = R_{e,h}^0 = \frac{n_e^p}{\tau_e} = \frac{n_h^n}{\tau_h}$$

From $L = \sqrt{D\tau}$ we obtain $\tau_e = L_e^2/D_e$ and $\tau_h = L_h^2/D_h$. Making the substitutions $n_e^p = n_i^2/n_A$ and $n_h^n = n_i^2/n_D$ then gives us the reverse saturation current

$$j_S = en_i^2 \left(\frac{D_e}{n_A L_e} + \frac{D_h}{n_D L_h} \right) \quad (6.28)$$

This expression for the reverse saturation current is valid even for real pn-junctions in which recombination is predominantly nonradiative, when empirically determined values are used for the diffusion lengths of the electrons and holes. However, recombination at interfaces, e.g. with the contacts, which is a problem in real solar cells, is not contained in this description.

For the short-circuit current, only the photons absorbed within the diffusion lengths are of interest. The pn-junction must therefore be no more than a distance L_h away from the surface. In fact, the n-layer at the surface is chosen to be very thin, and its absorption can be neglected. It must also be taken into account that the probability for the absorption of photons varies with the photon energy $\hbar\omega$. The generation rate resulting from light incident onto the n-layer at $x = 0$ is found by integration over the incident photon current spectrum $d_j(\hbar\omega, x = 0)$. This yields

$$\Delta G_{e,h}(x) = \int_0^\infty \alpha(\hbar\omega) (1 - r(\hbar\omega)) \exp(-\alpha(\hbar\omega)x) d_j(\hbar\omega, 0)$$

The contribution to the short-circuit current in the interval $d\hbar\omega$ is

$$\begin{aligned} dj_{SC}(\hbar\omega) &= -e [1 - r(\hbar\omega)] dj_\gamma(\hbar\omega, 0) \alpha(\hbar\omega) \int_0^{L_e} e^{-\alpha x} dx \\ &= -e [1 - r(\hbar\omega)] \{1 - \exp[-\alpha(\hbar\omega)L_e]\} dj_\gamma(\hbar\omega, 0) \end{aligned} \quad (6.29)$$

$[1 - r(\hbar\omega)] \{1 - \exp[-\alpha(\hbar\omega)L_e]\}$ is the familiar absorptance $\alpha(\hbar\omega, L_e)$ of a layer of thickness L_e for photons with an energy $\hbar\omega$. If the diffusion length is greater than the thickness of the solar cell, the absorptance for the actual thickness must be used.¹⁾ The short-circuit current is given by the photon current absorbed within the shorter of the two lengths, the diffusion length of the electrons, or the thickness of the solar cell. For $L_e < d$, we get

$$j_{SC} = -e \int_0^{\infty} a(\hbar\omega, L_e) dj_\gamma(\hbar\omega, 0)$$

Figure 6.11 shows the current–voltage characteristic for the pn-junction in the dark and with illumination. The similarity with the characteristic of the chemical cell in Figure 6.4 is of no surprise since, for both, the current is limited by chemical reactions. The similarity is even greater. We leave it to the reader to show that the same linear current characteristic as shown in Figure 6.3 for the chemical cell is found for the semiconductor cell as well, if the current is plotted as a function of the product of the concentrations $n_e n_h$.

In addition to the short-circuit current, the open-circuit voltage V_{oc} is important. From

$$j_Q = j_S \left[\exp \left(\frac{eV_{oc}}{kT} \right) - 1 \right] + j_{SC} = 0 \quad (6.30)$$

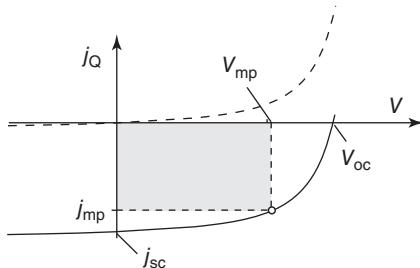


Figure 6.11 Charge current of the pn-junction in the dark (dashed line) and with illumination (solid line) as a function of the voltage. The sign of the voltage corresponds

to the polarity of the p-region. The shaded rectangle represents the maximum power delivered by the illuminated pn-junction.

1) In this case, a perfect hole membrane at the rear side of the cell must prevent the electrons from reaching the back contact where they would otherwise recombine, without contributing to the short-circuit current.

we obtain

$$V_{oc} = \frac{kT}{e} \ln \left(1 - \frac{j_{sc}}{j_s} \right) \quad (6.31)$$

It is essential for a large voltage that the reverse saturation current j_s is as small as possible. The generation rate in the dark $G_{e,h}^0$ has its lowest possible value, when electron–hole pairs are generated only by the absorption of 300 K radiation from the environment and therefore recombine only radiatively. In this case, the recombination rate is proportional to $\exp(eV/kT)$ resulting in the familiar current–voltage characteristic of the pn-junction. From the discussion of different recombination processes in Chapter 3, we know that the recombination rate has a different dependence on the difference of the Fermi energies for Auger recombination or for recombination via impurities, which leads to a different dependence of the current on the voltage.

6.5

pn-Junction with Impurity Recombination, Two-Diode Model

The current–voltage characteristic calculated for the pn-junction in the previous section is based on the assumption that only radiative recombination takes place. This is the ideal situation that enables us to determine upper limits for the open-circuit voltage and the efficiency of a solar cell. In real solar cells, however, recombination via impurities predominates, as discussed in detail in Section 3.6.2. From there, we already know that impurities with an electron energy in the middle of the forbidden gap at $\epsilon_{imp} = \epsilon_i$ contribute particularly strongly to recombination. To simplify the evaluation of the influence of impurity recombination on the current–voltage relationship, we assume that electrons and holes have equal thermal velocities v and are captured with the same cross sections σ . With this simplification, Equation 3.82 leads to the relationship

$$R_{imp} = n_{imp} \sigma v n_i \frac{\exp[(\epsilon_{FC} - \epsilon_{FV})/kT] - 1}{\exp[(\epsilon_{FC} - \epsilon_i)/kT] + \exp[(\epsilon_i - \epsilon_{FV})/kT] + 2} \quad (6.32)$$

For a given voltage V and thus a given difference $\epsilon_{FC} - \epsilon_{FV} = eV$ between the Fermi energies, Figure 6.12 shows the recombination rate R_{imp} as a function of the mean value $(\epsilon_{FC} + \epsilon_{FV})/2$ of the Fermi energies relative to the position ϵ_i of the impurity level assumed to be in the middle of the forbidden gap. It can be seen that the recombination rate has a pronounced maximum, when the Fermi energies lie symmetrically about the impurity level.

This behavior is understandable from Figure 6.13, which shows the potential distribution in a pn-junction at an applied voltage V and the position of the impurity levels relative to the Fermi energies. In the p-region, on the right, the impurity states are largely unoccupied, because their electron energies are above the Fermi energies, and in the n-region, on the left, they are largely occupied. In both regions, the recombination rate is small. The recombination rate is large only where the impurity level is in between the Fermi energies, that is, right at the pn-junction.

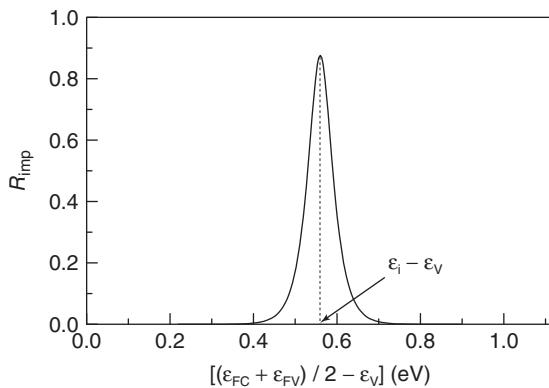


Figure 6.12 Impurity recombination rate R_{imp} as a function of the mean value of the Fermi energies $(\epsilon_{\text{FC}} + \epsilon_{\text{FV}})/2$ relative to the edge of the valence band ϵ_{V} in a semiconductor with an energy gap of $\epsilon_{\text{G}} = 1.12$ eV calculated for an applied voltage of $V = 0.4$ V.

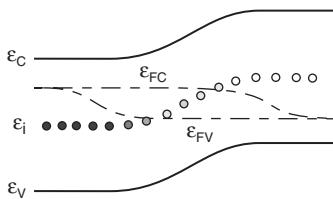


Figure 6.13 Potential distribution in a pn-junction with impurities at ϵ_i in the middle of the forbidden zone.

There $\epsilon_{\text{FC}} - \epsilon_i = \epsilon_i - \epsilon_{\text{FV}} = eV/2$. This means that the recombination rate in Equation 6.32 is

$$R_{\text{imp}} = n_{\text{imp}} \sigma n v_i \frac{\exp(eV/kT) - 1}{2[\exp(eV/2kT) + 1]} \quad (6.33)$$

Let us further assume that this recombination rate is constant over the thickness w of the space charge layers and also consider that $\exp(eV/kT) - 1 = [\exp(eV/2kT) + 1][\exp(eV/2kT) - 1]$. As a result, the impurity recombination leads to an additional charge current of

$$j_{Q,\text{imp}} = \frac{ew\sigma v n_{\text{imp}} n_i}{2} \left\{ \exp\left(\frac{eV}{2kT}\right) - 1 \right\} = j_{S2} \left\{ \exp\left(\frac{eV}{2kT}\right) - 1 \right\} \quad (6.34)$$

This current flows in addition to the current produced by band–band recombination, which determines the current over the region of the diffusion lengths. Because of the different voltage dependence, at small voltages, impurity recombination predominates and, at large voltages, band–band recombination predominates. The total current, including the short-circuit current due to illumination, is

$$j_Q = j_{S1} \left\{ \exp\left(\frac{eV}{kT}\right) - 1 \right\} + j_{S2} \left\{ \left(\exp\frac{eV}{2kT} \right) - 1 \right\} + j_{\text{SC}} \quad (6.35)$$

Except for the short-circuit current, we can visualize this total current as arising through two types of diodes connected in parallel. One diode with the reverse

current j_{S1} in which only band–band recombination takes place, is connected in parallel with two diodes in series both having a reverse current j_{S2} and in which only impurity recombination occurs in the space charge region. Because of the series connection, only half of the voltage is applied to each of these two diodes. This summarizes the two-diode model, which considers both band–band and impurity recombination and reproduces the current–voltage characteristic of real pn-junctions relatively well.

A remark may be appropriate at this point. The exponential dependence of the current on the voltage in a pn-junction is often seen as being caused by the potential barrier between the n- and p-sides, which the carriers in the forward direction of the current have to surmount. In this description, it is very hard to understand that the carriers “see” only half of this barrier when they subsequently disappear by impurity recombination. We know, however, that the rise of the forward current is, instead, caused by the dependence of the recombination rate on the difference of the Fermi energies and it is no surprise that these dependencies are different for different recombination processes.

We have seen that the properties of even a homojunction, a pn-junction consisting of the same base material, can be interpreted in terms of the membrane model. The contributions to the current originate from chemical reactions between electrons and holes on the one hand, and photons and phonons on the other hand, occurring within a reaction volume extending from a diffusion length left of the junction to a diffusion length right of the junction. The materials outside the reaction volume, n-type on one side, p-type on the other side, serve as membranes, in which charge transport involves only the majority carriers. pn-junctions are not ideal for solar cells for the following reason. For the best performance of a solar cell, the recombination probability should be made as small as possible and the semiconductor material as pure as possible, allowing only for radiative recombination. We have learned in Section 3.6.3 that the radiative lifetime of electrons is 30 ms in p-type silicon doped with 10^{16} cm^{-3} acceptors, typical for crystalline silicon solar cells. For this lifetime, the diffusion length is $L_e \approx 1\text{cm}$. In order to preserve the function of the hole membrane on the p-side that keeps the electrons away from the hole contact at the rear, the p-region of the solar cell would have to be thicker than that. The solar cell would then have to be much thicker than required to absorb the absorbable photons with $\hbar\omega \geq \epsilon_G$, a waste of precious material. In Section 7.3, we see how this problem is solved. In any case, the membrane function is necessary to prevent minority carriers from reaching the wrong metal contact, although gradients of their electrochemical potentials exist to drive them into the wrong direction.

The membrane – function of regions close to the contacts is important for nearly all electronic devices, not only for solar cells. The necessary property is that the electron contact exchanges only electrons and no holes and the hole contact exchanges only holes and no electrons. For devices working in the dark, rectifiers, transistors etc., the selectivity in the conductivities is very large. In an n-type region of silicon doped with $10^{16} \text{ donors/cm}^3$ the hole concentration is $10^4/\text{cm}^3$ and thus 12 orders of magnitude smaller, the same is roughly true for

the conductivities. In the dark, an n-type region is, therefore, a perfect selective membrane for electrons as is a p-type region for holes.

In solar cells, however, there is a problem.

- On the one hand, we want the minority carrier concentrations to be very small to preserve the membrane function and ensure that all electrons, which do not recombine, can only leave through the electron contact and all holes, which do not recombine, through the hole contact, thus establishing the largest possible current.
- On the other hand, we want the minority carrier concentrations to be as large as possible in order to get the largest possible difference of the quasi Fermi-energies, thus establishing the largest possible voltage

These two requirements contradict each other in a pn-junction made from the same base material.

To demonstrate this problem, we have made a numerical simulation of a model pn-junction solar cell. It has the same basic properties as the absorber of the hetero-junction solar cell in Figure 6.5, bandgap of 1.34 eV, mobilities of $100 \text{ cm}^2/(\text{Vs})$ for electrons and holes, only radiative recombination. The thickness of the cell is 10 μm with the pn-junction in the middle, consisting of a 5 μm thick n-type region doped with $10^{17} \text{ donors/cm}^3$ and a 5 μm thick p-type region doped with $10^{17} \text{ acceptors/cm}^3$. Full area ohmic metal contacts are applied to the surfaces with assumed recombination velocities of $10^7 \text{ cm}^2/\text{Vs}$. The generation rate is assumed to be homogeneous with the total generation rate equal to the absorption of all photons with $\hbar\omega \geq \epsilon_G$ from the AM1.5 solar spectrum. Figure 6.14 shows the energy diagram of this pn-junction solar cell at maximum power.

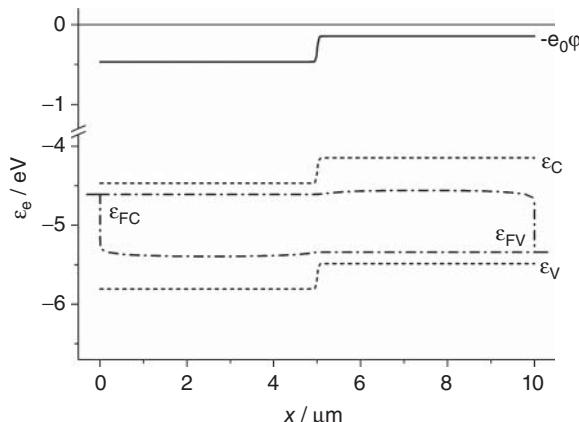


Figure 6.14 Energy diagram of a symmetrical pn-junction under AM1.5 illumination and at maximum power. The bandgap is 1.34 eV, the voltage $V_{mp} = 730.1 \text{ mV}$ the current $j_{mp} = 17.1 \text{ mA/cm}^2$ and the efficiency $\eta = 12.5 \%$.

Although there is only radiative recombination, this pn-junction solar cell has an open circuit voltage of 818 mV, a short circuit current of 17.8 mA/cm^2 a fill factor of 85.6 and an efficiency of only 12.5 %. These values are much less than what one expects for an ideal solar cell, namely $V_{oc} = 1082\text{mV}$, $j_{sc} = 35 \text{ mA/cm}^2$ and $\eta = 33.7 \%$, which are the values that were achieved with the hetero-junction solar cell in Figure 6.6. The reason for this bad performance, especially for the current, can be seen in Figure 6.14. The gradient of the quasi Fermi-energy ϵ_{FV} in the left half of the n-region drives holes towards the electron contact, in the same way as the gradient of ϵ_{FC} in the right half of the p-region drives electrons towards the hole contact. Due to the illumination, the minority carrier concentrations have increased by a factor of 10^{12} compared to the dark. This large increase of the minority carrier conductivities enables the gradient of the quasi Fermi-energies to drive all holes generated in the left half of the n-region and all electrons in the right half of the p-region towards the wrong contacts, where they recombine. This explains the strong reduction of the current, which is only half of the ideal value. This transport towards the wrong contacts happens even under open-circuit conditions, causing strong recombination of electrons and holes at the contacts and reducing the open-circuit voltage.

If this model pn-junction solar cell performs so poorly, although its base material shows only radiative recombination in the bulk, how can a real silicon pn-junction solar cell, which suffers from additional non-radiative bulk recombination, achieve a record efficiency of 25 %? The answer is that over many years of development a structure has been established, mainly by trial and error, which exhibits the properties of semi-permeable membranes. This structure is discussed in detail in Section 7.3. Here it may suffice that in real silicon solar cells, regions close to the contacts are more highly doped than chosen for Figure 6.14 and the contacts cover only a small part of the surface area in contrast to the model in Figure 6.14, where they cover the whole front and back surface.

6.6 Heterojunctions

In Figures 5.6 and 6.14 we see that there is transport in the wrong direction in a homo-junction, that is electrons flow to the contact on the p-side and holes to the n-side. The associated charge current by which the total current is reduced was neglected in the calculation of the current-voltage characteristic for the pn-junction. In fact, neglecting this contribution is justified only, when measures are taken to eliminate this current or its cause, which is recombination at the interface with the metal contact. One possibility is shown in Figure 6.5. The electrons flow through an n-conductor out of the cell, and the holes through a p-conductor, both with a large energy gap. Due to the resulting very low concentration of minority carriers in the membranes, minority carrier currents in the wrong direction are negligible, namely an electron current towards the hole contact on the p-side and a hole current towards the electron contact on the

n-side. The structure of Figure 6.5 requires three different materials; an absorbing semiconductor in the middle, between two semiconductors with a larger energy gap and different electron affinities χ_e . Such combinations of different materials are known as hetero-junctions.

In addition to preventing interface recombination at metal contacts, hetero-junctions are important if it is not possible to prepare a pn-junction from a single material, a so-called homo junction. There are, in fact, many materials that can only be doped either n-type or p-type. This includes nearly all materials with an energy gap greater than 2.5 eV. This is one of the reasons why it is so difficult to manufacture blue-emitting light emitting diodes (LEDs), with the exception of GaN.

For solar cells and other electronic components, it is not sufficient to bring semiconductors with suitable energy gaps and electron affinities into contact. It is more important that the interfaces are as free as possible of states with energies in the forbidden gap, in order to eliminate additional recombination and also to prevent electrical charging due to preferred trapping of one carrier type. Material combinations fulfilling these conditions are very rare. One suitable combination is silicon–silicon dioxide, in which the silicon dioxide cannot be doped and therefore serves as an insulator. Nevertheless, we will see later that this combination is of great importance for silicon solar cells. Most other known combinations with a low density of interface states consist of III–V compounds and are based on gallium arsenide.

The distribution of the electrical potential and of the band edges can be determined for a heterojunction just as easily as for a normal pn-junction, if the interfaces are not charged, since this implies that the dielectric displacement $D = \epsilon\epsilon_0 E$ and the electrical potential φ are continuous over the interface. Because of the continuity of D , the field strengths E to the right and left of the interface differ by the ratio of the dielectric constants ϵ .

In Figure 6.15 we see two different semiconductors, semiconductor 1 on the left, which is n-doped, and semiconductor 2 on the right, which is p-doped. To determine the potential distribution in the dark, we begin in Figure 6.16 with semiconductor 1, for which — except for the possible occurrence of a space charge layer — the electrical potential is held constant to the left of the interface. In electrochemical equilibrium with the p-conductor, the Fermi energy has the same value everywhere. To the right of a space charge layer, possibly occurring at the interface, the band edges ϵ_C and ϵ_V and the electron affinities χ of semiconductor 2 are entered relative to the Fermi energies, with the same values as in Figure 6.15.

The distribution of the electrical potential is determined by the charge distribution and is independent of the electron affinity or the energy gap. Owing to the continuity of the electrical potential across the interface, in the case without charge in interface states, the resulting potential distribution is, apart from the different slope arising from the ratio of the dielectric constants, the same as for the homo-pn-junction in Section 6.4.

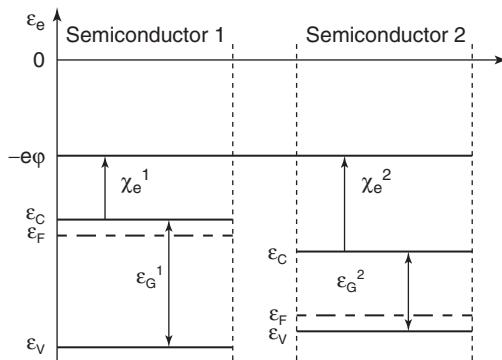


Figure 6.15 Two different semiconductors prior to making contact.

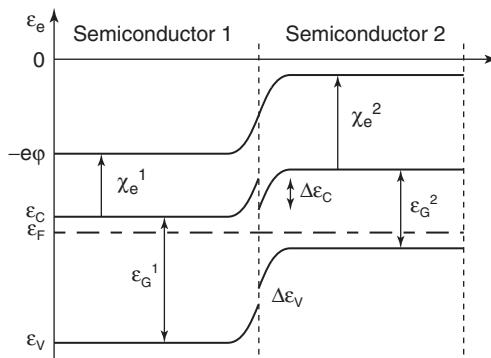


Figure 6.16 The two different semiconductors of Figure 6.16 in contact.

According to Equation 6.36, the electrical potential difference between the neutral regions of the n-conductor and the p-conductor is

$$e(\varphi^1 - \varphi^2) = \mu_{e,0}^1 + kT \ln \left(\frac{n_e^1}{N_C^1} \right) - \mu_{e,0}^2 - kT \ln \left(\frac{n_e^2}{N_C^2} \right) \quad (6.36)$$

where the standard values of the chemical potential $\mu_{e,0}$ of the electrons are now given by the different electron affinities χ_e of the two semiconductors:

$$\mu_{e,0}^1 - \mu_{e,0}^2 = \chi_e^2 - \chi_e^1$$

The difference in the electron affinities also determines the discontinuity $\Delta\epsilon_C$ in the conduction band edges and, together with the difference ϵ_G in the energy gaps, the discontinuity $\Delta\epsilon_V$ of the valence band edges directly at the interface.

$$\Delta\epsilon_C = \epsilon_C^1 - \epsilon_C^2 = \chi_e^2 - \chi_e^1$$

$$\Delta\epsilon_V = \epsilon_V^1 - \epsilon_V^2 = \chi_e^2 - \chi_e^1 + \epsilon_G^2 - \epsilon_G^1$$

By selecting materials with suitable electron affinities, χ_e , and energy gaps, ϵ_G , as in Figure 6.5, it is possible to avoid discontinuities in the band edges for the majority carriers and produce them for the minority carriers, and, in this way, control the charge transport.

If interface states are present on one or both materials, the interface may be charged or may have a dipole moment or both. While outside the space charge regions from the interface, the positions of the bands in the two materials relative to each other remain as shown in Figure 6.16, the distribution of the electrical potential φ within the space charge regions may change quite dramatically. A sheet charge of one sign would add a spike to the potential distribution in Figure 6.16, upward for negative charge and downward for positive charge. A dipole moment would add a step, e.g. upward in going from left to right if the positive charge sits on the surface on the left of the interface and the negative charge is on the right surface. Since the position of the bands in the space charge regions is affected in the same way, charging of the interface will have a profound influence on the transport of electrons and holes across the interface.

6.7

Semiconductor–Metal Contact

Metal contacts should allow an unimpeded charge transport between the solar cell and an external load. The contact with a metal must therefore not cause a depletion of the majority charge carriers in the adjoining semiconductor. The potential distribution at the semiconductor–metal contact follows the same rules as at a semiconductor–semiconductor contact. Figure 6.17 shows a semiconductor and a metal before making contact and in contact with each other.

The metal is characterized solely by the chemical potential $\mu_{e,m}$ of its electrons. The absolute value of the chemical potential is known as the *work function* of the metal. In contact, an electrical potential difference arises, as for the case of two

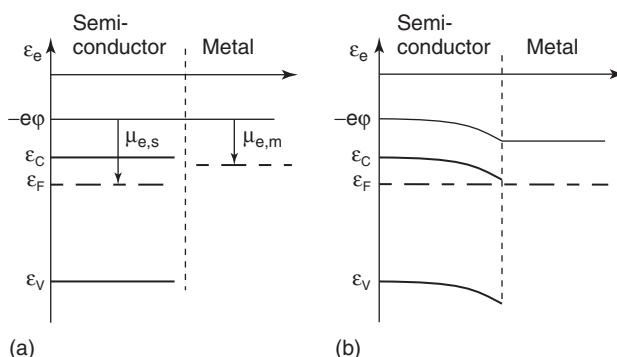


Figure 6.17 Schematic energy diagram of a semiconductor and a metal (a) before making contact and (b) in contact.

semiconductors, corresponding to the difference in the work functions. Owing to the high concentration of electrons in the metal, the charge distribution in the metal degenerates to a surface charge. The entire potential difference between the metal and the semiconductor, therefore, appears across the space charge layer of the semiconductor. In Figure 6.17, we have assumed that there is no charge at the contact interface other than the surface charge of the metal, so that the electrical potential is continuous across the interface. From this figure, we can see that metals having a smaller work function than the semiconductor cause an accumulation of electrons in the adjacent semiconductor. Metals with a small work function, therefore, are favorable for the exchange of electrons and make good, so-called ohmic, contacts to n-type semiconductors. Conversely, metals with a large work function make good, ohmic contacts to p-type semiconductors. Electrons see a potential barrier from the Fermi energy of the metal to the conduction band of the semiconductor, equal to the difference between the work function of the metal and the electron affinity of the semiconductor. Heat is consumed when electrons cross the barrier in traveling from the metal to the semiconductor, resulting in a Peltier cooling effect. If, according to the direction of the charge current, electrons travel from the semiconductor to the metal, a Peltier heating effect is observed.

For contacts on covalent semiconductors, silicon in particular, the band bending in the semiconductor is found to depend less strongly on the work function of the metal than would be expected from the difference in the work functions. This is probably caused by the presence of surface states on the silicon surface, which are charged in contact with the metal. Together with the surface charge on the metal, the surface charge on the silicon forms a charge double layer, over which a potential step occurs, which depends on the dipole moment of this layer. The potential difference between the interior of the semiconductor and the interior of the metal is still given by the difference in the work functions, but it is now divided between the band bending of the space charge layer and the, usually unknown, potential difference across the dipole layer.

Contacts with good carrier exchange can also be prepared on the basis of another principle. If the semiconductor is very highly doped, at least in the vicinity of the contact, a majority carrier depletion layer caused by an inappropriate work function of a metal is only very thin. It can, in fact, for suitably high doping levels, be so thin that the majority charge carriers can tunnel through this potential barrier between the semiconductor and the metal. Aluminum has a smaller work function than p-doped silicon and is, therefore, not expected to be a good contact material for it. If, however, after deposition, the aluminum is allowed to diffuse at high temperatures into the silicon, where it forms acceptor states, a strongly p-doped layer results with a potential distribution as shown in Figure 6.18. The thin barrier in the valence band permits good hole exchange between the valence band of the silicon and the conduction band of the Al contact by tunneling. Generating a strongly doped p⁺-layer in front of the metal contact also improves the membrane character of the p-region.

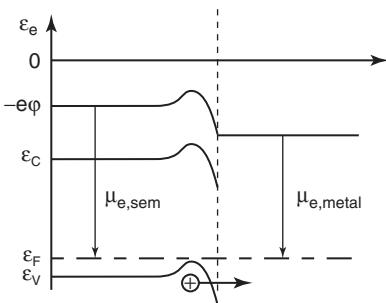


Figure 6.18 Holes from the valence band can tunnel through a thin potential barrier of a strongly p-doped depletion layer into the metal.

Table 6.1 Work Functions and Electron Affinities of Common Semiconductors and Metals.

	Si	GaAs	In	Ag	Al	Au	Pt
Work function (eV)				4.12	4.26	4.28	5.1
Electron affinity (eV)	4.01	4.07					5.65

6.7.1

Schottky Contact

The basic principle of a solar cell, namely, having two membranes, one exchanging electrons and the other exchanging holes, can be fulfilled simply by a homogeneously doped semiconductor with two different metal contacts, one ohmic contact for the exchange of the majority carriers and a second that causes depletion of the majority carriers, and thus accumulation of minority carriers. This second contact is called a *Schottky contact*. For an n-conductor, its work function must be much greater than that of the semiconductor, and for a p-conductor much smaller.

Figure 6.19 shows the potential distribution at a Schottky contact on a p-conductor in the dark. It must, however, be noted that Schottky contacts are

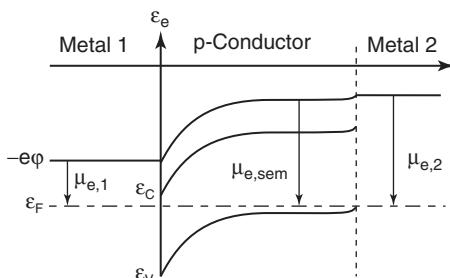


Figure 6.19 A metal with a small work function (left) forms a Schottky contact on a p-conductor with a larger work function. On the right side, a metal with a large work function forms an ohmic contact.

simple to prepare only for a few semiconductors. Furthermore, they have the disadvantage that the required unimpeded exchange of the minority carriers is unavoidably coupled with a high level of surface recombination at the metal contact. For solar cells, Schottky contacts play a role only in the testing of new materials. Past experience has shown that well-functioning contacts, assumed to be Schottky contacts, were in fact semiconductor heterojunctions formed by a chemical reaction of the metal with the semiconductor. An example is the contact between copper (Cu) and n-type cadmium sulfide (CdS), where p-conducting Cu_2S is formed.

6.7.2

MIS Contact

To reduce the surface recombination at the Schottky contact, one can insert an oxide layer between the metal and the semiconductor. Even though it is an insulator, this is no serious obstacle when it is sufficiently thin to allow electrons or holes to tunnel through the thin potential barrier. However, the depletion of majority carriers in the semiconductor is then not as pronounced, because a part of the difference in the work functions is now occurring across the oxide layer and not in the semiconductor.

In metal insulator silicon (MIS) structures, [14] this disadvantage is compensated by exploiting a particular property of silicon dioxide. Impurity atoms bound to silicon dioxide (e.g. sodium) are frequently ionized, i.e. charged. Their charge is neutralized by the charge in the metal and in the semiconductor. If the oxide is positively charged at the interface to the semiconductor, holes are pushed away from the interface and electrons are accumulated. Figure 6.20 shows that a potential distribution similar to that of a Schottky contact results in the p-conductor, but now without the disadvantage of a high level of surface recombination.

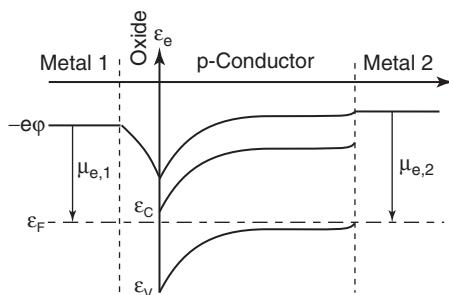


Figure 6.20 In a metal-insulator-silicon (MIS) structure, a very thin oxide layer between the metal and the silicon prevents surface recombination. The depletion of holes in the p-conductor is the result of a positive charge trapped in the oxide near the silicon.

6.8

The Role of the Electric Field in Solar Cells

The reader may find it confusing that the electric field that exists in the dark and, although somewhat reduced, also in the illuminated pn-junction, is of no significance for our understanding of the solar cell. The criterion for a solar cell structure is that electrons and holes are forced by membranes into different directions and that on their path their entropy is conserved. If this condition is fulfilled, in some structures, e.g. in a pn-junction of uniform material, an electric field will be present between the membranes. The direction of the short-circuit charge current in a pn solar cell agrees with the direction of this field. This seems to be sufficient to believe that it is also causing this current. To exaggerate somewhat, this is mere coincidence. It would be a completely unnecessary restriction to exclude structures for solar cells in which no electric field is present, but which have the membrane function incorporated and which fulfill the condition of conservation of entropy. The dye solar cell in Section 6.3 is a good example. The intimate mixture of electron membrane (TiO_2), dye, and hole membrane (electrolyte) on a nanometer scale, does not allow the formation of an extended space charge and of a field. Another example is given at the end of this section.

We frequently read that it is just the electric field of a pn-junction that supplies the driving force for the currents flowing during illumination. Let us take a closer look at this argument. If it were true that the charge carriers were driven by the field, then they would be continuously accelerated by the field and slowed down by collisions with the lattice. The field would perform work to keep them moving. There would have to be a source of energy present that continuously supplied energy to compensate for the energy dissipated during each collision of the charge carriers (see Section 5.1.3) in order to maintain a constant current. Such an energy source is, however, not present. We can see this best by comparison with a simple example.

Imagine a capacitor filled with a material that insulates in the dark and becomes conductive, as a result of the generation of electrons and holes, when illuminated. This capacitor is charged in the dark, and the voltage source is then disconnected. An electric field is present in the material between the plates of the capacitor. When this material is illuminated, a current will flow. Electrons flow to the positively charged plate and holes to the negatively charged plate. These particles gain kinetic energy from the field, which is subsequently dissipated by phonon scattering. However, the field is weakened by the charge transport, and the capacitor is discharged. The current vanishes after a dielectric relaxation time, and the electric field vanishes because the energy of the electric field stored in the capacitor during the charging process is used up. A steady state current driven by an electric field requires a continuous source of energy that maintains the field, in other words, a battery.

We have made some effort to distinguish between an electrochemical potential difference, which is measured by a voltmeter and an electrical potential difference, which is not. In fact, something must be wrong in our physics education,

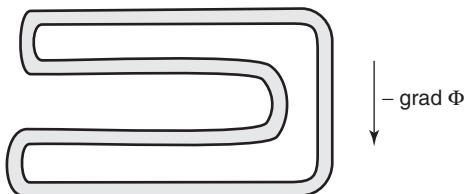


Figure 6.21 Water in a closed pipe does not flow, although the gradient of the gravitational potential Φ drives it downwards.

if we think that a DC current can at all be driven in a closed circuit by a purely electrical potential difference. The word potential alone should tell us that no energy can be gained by moving a charge along any closed path. When it is back at the starting point, it has the same electrical energy again and no energy could have been dissipated by a charge in moving through a resistive medium. As an example, Figure 6.21 shows water that can flow in a closed circuit. There is also a driving force, the gradient of the gravitational potential Φ , driving the water downwards. We know that the water in the pipe will not flow. If it would flow downhill on one side, it would have to flow uphill on the other side. But how does the water on the right “know” that it goes uphill on the left? A pressure gradient exists that compensates the gravitational force at every point. The same happens in a pn-junction. The gradient of the electrical potential is only maintained because it is compensated by the gradient of the chemical potential. Otherwise, the pn-junction would be discharged in the same way as the capacitor above, since there is no energy source to support the field.

A variation of the field as the driving force sees the driving force in the ability to dissipate energy. In Figure 6.22 we see that the energy of an electron in the conduction band of a pn-junction is larger on the p-side than on the n-side. By moving from the p-side to the n-side, the energy of the electron is reduced, that is, energy which came from the absorbed photon that created the electron–hole

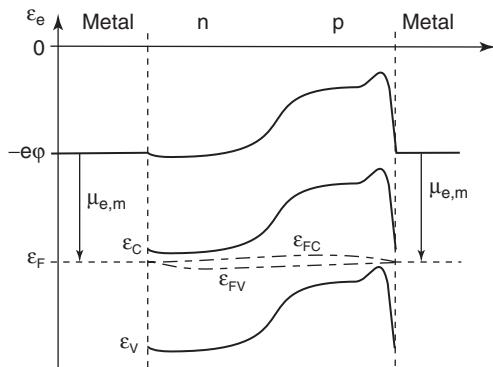


Figure 6.22 The distribution of the electrical potential φ in a pn-junction with metal contacts shows that a charge cannot gain energy from moving around a closed circuit.

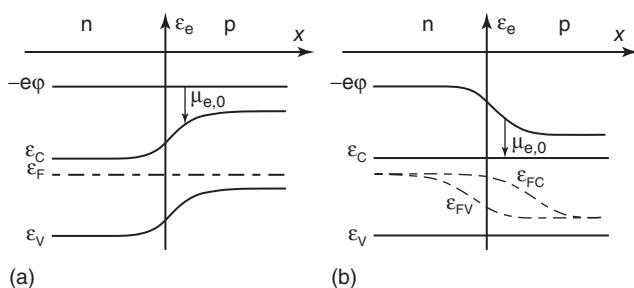


Figure 6.23 Potential distribution in a pn-junction, where the electron affinity $\chi_e(x) = -\mu_{e,0}(x)$ compensates the concentration-dependent part

$kT \ln[n_e(x)/N_C]$ of the chemical potential. (a) In the dark and in electrochemical equilibrium.(b) Illuminated and at open-circuit.

pair on the p-side. This energy reduction is largest under short-circuit conditions of an illuminated pn-junction and so, according to this argument, is the driving force. The error in this argument is that only free energy can be dissipated. We know that the free energy of an electron–hole pair is the difference of the two Fermi energies, which, in short circuit, is not large and is quite different from the energy change of an electron–hole pair when the electron is moved from the p- to the n-side and the hole from the n- to the p-side as is seen in Figure 6.23.

Energy, which is heat at the temperature of the solar cell cannot be dissipated, because the accompanying entropy is already at its maximum. The entropy per electron or hole from Equation (3.36) is

$$\sigma_{e,h} = k \left[\frac{5}{2} + \ln \left(\frac{N_{C,V}}{n_{e,h}} \right) \right]$$

This average entropy per particle, however, is not the entropy that is removed from the solar cell with the removal of an electron or a hole. Since the solar cell is operated in steady state, the removal of electrons and holes does not change their concentrations. We rather have to calculate the small change of the total entropy with a small change of the number of electrons or holes. For electrons the total entropy from (3.36) is

$$S_e = N_e k \left[\frac{5}{2} + \ln \left(\frac{N_C V}{N_e} \right) \right]$$

For a small change of the number of electrons dN_e in a state, where the number of electrons is N_e , the change of the entropy is $dS_e = (\partial S_e / \partial N_e) dN_e$, where $\partial S_e / \partial N_e$ is the amount of entropy that is removed with the removal of one electron. (The expression for the holes is analogous.)

$$\begin{aligned} \frac{\partial S_e}{\partial N_e} &= k \left[\frac{5}{2} + \ln \left(\frac{N_C V}{N_e} \right) \right] - k = \frac{3}{2} k + k \ln \left(\frac{N_C}{N_e} \right) \\ &= 1/T (\epsilon_{e,kin} + \epsilon_C - \epsilon_{FC}) \end{aligned}$$

The room-temperature heat which is removed with the removal of one electron-hole pair is

$$T \left(\frac{\partial S_e}{\partial N_e} + \frac{\partial S_h}{\partial N_h} \right) = 3kT + (\varepsilon_C - \varepsilon_{FC}) + (\varepsilon_{FV} - \varepsilon_V) = \varepsilon_G + 3kT - (\varepsilon_{FC} - \varepsilon_{FV})$$

By taking ε_{FC} at the electron contact and ε_{FV} at the hole contact, even the entropy production during the transport of the charge carriers towards the electrodes is accounted for.

We see that the reduction of the energy of an electron and a hole by separating them across a pn-junction is mainly a reduction of their heat along with a reduction of their entropy and has nothing to do with dissipation and a driving force. It is rather an exchange of heat and entropy with the phonon system. Under short-circuit conditions, the electrons and holes have gained their large entropy before moving across the pn-junction, because they were generated in an environment of low concentration. By being driven across the pn-junction by their proper driving forces (the gradients of their free energies), the entropy of the electrons and holes is given off to the phonon system and little entropy is generated when the gradients of the Fermi energies are small. This is the way by which entropy is given off to the environment, something that the solar cell has to do, in the same way as any other heat engine.

Figure 6.22 shows a pn-junction with contacts on the n- and p-sides. The contacts are made from the same metal as the wire (not shown) that connects them to ensure a short circuit. Outside the semiconductor, the electrical potential φ has the same value everywhere along the circuit, seen on the left and on the right in the figure. Inside the semiconductor it varies, owing to the variation in chemical composition (doping). A charge cannot benefit from this variation of the electrical potential by going around the circuit, since it goes uphill as much as it goes downhill in the electrical potential φ . Electrons and holes can, however, benefit from gradients of their electrochemical potentials, $\eta_e = \varepsilon_{FC}$ and $\eta_h = -\varepsilon_{FV}$, and the asymmetry of the membranes establishes preferred directions, for the electrons to go to the left and for the holes to go to the right.

To make perfectly clear that the electric field is irrelevant for the energy conversion in a solar cell, we construct a pn-junction in an admittedly cumbersome, but physically not forbidden, way where the p-side is uncharged relative to the n-side in the dark and no field is present. For the construction principle, we build on what we know about heterojunctions. According to Equation 6.36 the electrical potential difference between two bodies is determined by the difference in their work functions or by the difference in the chemical potentials μ_e of their electrons. Since the difference in the chemical potentials depends on the difference in the electron affinities and the ratio of the electron concentrations, it is conceivable to construct a pn-junction from continuously changing materials in such a way that all of them have the same energy gap and also have the same chemical potential for their electrons. In this structure, a spatial variation of the electron concentration, increasing from the p-side to the n-side, is compensated by a spatial variation of the electron affinity χ_e , decreasing from the n- to the p-region. In the dark, in

electrochemical equilibrium, this structure has the potential distribution shown in Figure 6.23a.

When this structure is illuminated, an electric field is produced, which is oppositely oriented to the field in the normal pn-junction and increases with the illumination intensity. It would drive electrons to the p-region and holes to the n-region, if it were the only force, as seen on the right in Figure 6.23 Does this then mean that the current in this pn-junction flows in a different direction from that in a normal pn-junction?

The actual driving forces, namely, the gradients of the electrochemical potentials, are identical with those in a normal pn-junction. The pn-junction illustrated in Figure 6.23 therefore behaves exactly as a normal pn-junction and has the same current–voltage characteristic, even though the electric field is opposing the charge current. We know that the exponential dependence of the current on the voltage is caused by the dependence of the recombination rate on the difference of the Fermi energies and that it is therefore the same for the hypothetical pn-junction in Figure 6.23 as for a normal pn-junction.

That charge currents are flowing against an electric field in an energy source is, by the way, nothing unusual. We find this in every battery. While electrons flow from the minus pole to the plus pole in the external circuit, the continuity of the charge current requires that ions within the battery move against the electrical force, i.e. from the plus pole to the minus pole, if they carry a negative charge or in the opposite direction, if they are positive. They are able to do so, because a gradient of their chemical potential overcompensates the gradient of the electrical potential. This results in a gradient of their electrochemical potential—their only driving force—pointing in the right direction.

In Figure 6.6, we have already seen as the result of an accurate numerical simulation that electrons and holes move against their electrical driving forces $-eE$ or eE in a hetero-junction solar cell at maximum power. The selective conductivities necessary for the functioning of this solar cell are achieved by an n-doped part in front of the electron contact and a p-doped part in front of the hole contact. Due to these different doping levels and the appropriately different work functions of the ohmic metal contacts, there is an electrical potential difference in the dark of $(kT/e) \ln(N_D N_A / n^2)$. This observation might suggest that the presence of an electric field in the dark is a necessary property of a solar cell structure, although it is not the driving force. In order to demonstrate that the electric field is only a by-product of one particular way of realizing the selectivity of the conductivities, we now look at a structure with no doping at all.

Figure 6.24 shows a $10\text{ }\mu\text{m}$ thick intrinsic absorber, again with a bandgap of 1.34 eV and electron and hole mobilities of $100\text{ cm}^2/(\text{Vs})$ sandwiched between two 100 nm thick undoped semi-permeable membranes with the same bandgap of 1.34 eV . Both contact metals are identical and have the same work function as the absorber and the membranes. Selective conductivities in the semi-permeable membranes in this case are assumed to be the result of different mobilities of electrons and holes. The electron transport layer has a large mobility for electrons and a small mobility for holes and the hole transport layer has a large hole

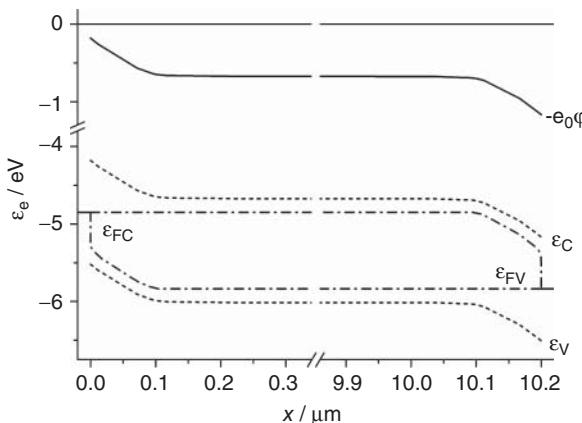


Figure 6.24 A $10 \mu\text{m}$ thick intrinsic absorber with a bandgap of 1.34 eV is sandwiched between two intrinsic transport layers of the same bandgap with a thickness of 100 nm . Selective conductivities of the transport layers are achieved by different mobilities. The electron transport layer has an electron

mobility of $10^8 \text{ cm}^2/(\text{Vs})$ and a hole mobility of $10^{-8} \text{ cm}^2/(\text{Vs})$. Opposite values are assumed for the hole transport layer. The energy diagram shows the situation at maximum power, where the efficiency is 33.7% . Absorption in the transport layers is neglected.

mobility and a small electron mobility. This structure has no electrical potential difference between its contacts in the dark. Under illumination the generation rate is assumed to be homogeneous, made up by all photons with $\hbar\omega \geq 1.34 \text{ eV}$ from a AM1.5 spectrum, evenly distributed over the thickness of the absorber, absorption by the transport layers is neglected. All recombination in the absorber is radiative. Calculations with the device simulator SENTaurus TCAD show that this structure achieves the same performance as the hetero-structure in Figure 6.6 with an efficiency of 33.7% in agreement with the Shockley-Queisser limit, if the ratio of the mobilities in the transport layers is large enough. For this purpose, the electron transport layer has an assumed electron mobility of $10^8 \text{ cm}^2/(\text{Vs})$ and a hole mobility of $10^{-8} \text{ cm}^2/(\text{Vs})$. The opposite values are assumed for the hole transport layer. Due to the small carrier concentrations at the contacts fixed by the metal electrodes, a surface recombination velocity of 10^{12} cm/s was assumed to allow the current to pass through the contact interface without voltage loss. Figure 6.24 shows an energy diagram at maximum power.

We know that the assumed mobility values are unrealistic. This example shall rather demonstrate that it is only the selectivity of the conductivities which makes a structure work as an efficient solar cell and nothing else. If this selectivity is achieved by doping in a practical case, an electrical potential difference is caused, but this is a side effect accompanying selective conductivities in front of the contacts [12].

6.9

Organic Solar Cells

We have already discussed the dye solar cell in Section 6.3, but we did not mention one very important difference between organic and inorganic semiconductors concerning the generation of mobile charge carriers. In principle, in both material classes, the absorption of a photon does not directly lead to a free electron and a free hole since they attract each other because of their opposite charge. This entity in which electron and hole are still bound to each other by Coulomb forces is called an *exciton*. It is mobile and electrically neutral.

6.9.1

Excitons

A model for the exciton is the hydrogen atom as it was for the binding of an electron to a donor in Equation 3.25

$$\epsilon_{\text{exc}} = \frac{m_{\text{red}} e^4}{2(4\pi\epsilon\epsilon_0)^2\hbar^2} \quad (6.37)$$

where the reduced mass of the exciton $m_{\text{red}} = m_e^* m_h^* / (m_e^* + m_h^*)$ is introduced to account for the movement of electron and hole around their common center of mass. Equation 6.37 gives the binding energy of the electron (or hole) in the ground state of the exciton. As known from the hydrogen atom, the exciton has a series of excited states above the ground state and below the ionization continuum where electron and hole are no longer bound to each other and are free to move. The binding energies following from Equation 6.37 are much smaller than kT at room temperature for most common inorganic semiconductors because of their large dielectric permittivity $\epsilon > 10$ and of effective masses of electrons and holes which are smaller than their real masses. As a consequence, excitons exist in inorganic semiconductors only at low temperatures, and at room temperature, electrons and holes are free to move after being generated by the absorption of a photon. No distinction is made between the minimum energy of absorbable photons and the minimum energy of free electron–hole pairs. This minimum energy is defined as the band gap for inorganic semiconductors.

This is different in organic semiconductors. They often consist of molecules that are only weakly bound to each other by van der Waals forces. As a result, they have a smaller dielectric permittivity and larger effective masses of electrons and holes. This raises the binding energy of the exciton and reduces its radius to a value comparable with an interatomic distance. For this size of the exciton, using bulk values for the dielectric permittivity and for the effective masses is no longer justified and the binding energies are even larger than predicted by Equation 6.37. The exciton binding energy in organic semiconductors may be as large as some tenths of an eV and therefore much larger than kT at room temperature. Even if excited above the ionization continuum by a photon with sufficient energy, electron and hole may quickly lose some of their kinetic energy by scattering with

lattice vibrations and may then be trapped in their mutual attraction and form an exciton. Most absorption processes in organic semiconductors end up in excitons and the generation of free electrons and holes is the exception. In any case, we have to keep in mind that absorbable photons with the lowest energy, defining the absorption band gap, will not generate free charge carriers. The minimum energy of a free electron–hole pair, defining the free-carrier band gap, is larger than the optical band gap by the binding energy of the exciton. The exciton is a two-particle state with an integer spin and as such does not follow Fermi statistics but Bose–Einstein statistics. Its energy cannot be well represented on a scale for electron energies. As a compromise, the exciton energy is often given with respect to the valence band edge. By measuring the exciton energy upward from the valence band edge as shown on the left of Figure 6.25, it is represented as the energy of an electron bound to a free hole. Since in a material, the hole has as much reality as the electron, the exciton energy may as well be represented by the energy of a hole bound to a free electron. In this way, the exciton energy is counted downward from the conduction band edge as on the right in Figure 6.25. This representation is uncommon, but is very helpful in discussing the dissociation of excitons in organic materials.

The terminology for organic semiconductors is often different from the terminology used for inorganic semiconductors: the transfer of an electron to a molecule is described as a reduction of the molecule, and the transfer of a hole is an oxidation. Instead of the conduction band edge ϵ_C , we encounter the lowest unoccupied molecular orbital (LUMO), and instead of ϵ_V , we find the highest occupied molecular orbital (HOMO).

Advantages and Disadvantages

We start with the disadvantages of organic semiconductors. The rather weak van der Waals forces between the molecules of an organic semiconductor result

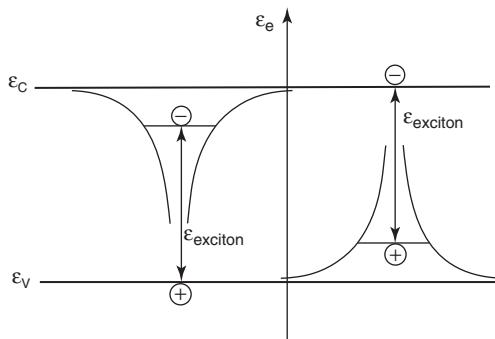


Figure 6.25 The exciton can be represented in two different ways. On the left, the electron is shown in the electric field of the hole and its energy (in the ground state of the exciton) is given relative to the energy of a free hole. The exciton binding energy is the

energy difference from the ground state of the electron in the exciton to the conduction band edge ϵ_C . On the right everything is reversed as from the perspective of the hole in the field of the electron and its energy is relative to that of a free electron.

in little overlap of electron wavefunctions from neighboring molecules. This causes narrow bands with large densities of states, large effective masses and, consequently, small mobilities of electrons and holes. The transport properties of organic semiconductors for electrons, holes and excitons are therefore rather poor.

On the other hand, organic semiconductors have advantages over inorganic semiconductors. The high densities of states in the bands result in large values of the absorption coefficient α on the order of 10^5 cm^{-1} for common direct transitions. To absorb most of the absorbable photons, layers of organic semiconductors do not have to be thicker than a fraction of 1 μm .

Another favorable property follows from the weak van der Waals forces between the molecules. States in the middle of the band gap that facilitate nonradiative recombination in the bulk of the semiconductor or on surfaces or interfaces have much lower densities than in inorganic semiconductors. Organic semiconductors, therefore, often have high fluorescence yields, which means that radiative recombination is more probable than nonradiative recombination. Organic semiconductors are therefore more efficient in transforming the energy from the Sun (heat) into chemical energy (of excitons) than inorganic semiconductors. And last not least, organic semiconductors need only moderate temperatures for their preparation, they are flexible, and may be cheap. They are, however, susceptible to oxygen and other gases and need tight encapsulation.

The favorable properties, however, do not make up for the poor transport properties. The exciton diffusion length is typically on the order of only 10 nm, i.e. much smaller than the penetration depth $1/\alpha$ of the absorbable photons.

For good organic solar cells, we have to solve two problems:

- 1) produce free electrons and holes from excitons
- 2) find a structure that is compatible with small transport distances and, nevertheless, complete absorption.

Dissociation of Excitons

It has been known for some time that even strongly bound excitons can dissociate into a free hole and a free electron at the interface with an appropriate different material. How this may work is explained by Figure 6.26. The exciton is represented as an electron bound to a free hole in material 1 on the left. The Coulomb interaction between electron and hole causes an electrical potential well $-e\varphi$ for the electron. (The hole would see a totally different potential distribution, which is suppressed since the reference is the free hole.) The Coulomb well for the electron extends into material 2 on the right. Caused by a larger electron affinity $\chi_{e,2}$ of material 2, the energy of a free electron in this material is smaller than that of the bound electron in material 1. If the exciton in material 1 comes very close to the interface with material 2, the electron may tunnel (at constant energy) out of its bound state into a free state in the conduction band of material 2. Since the electron affinity χ_e is so much larger in material 2 than in material 1, the electron has enough kinetic energy $\varepsilon_{e,\text{kin}}$ in material 2 to overcome the Coulomb attraction

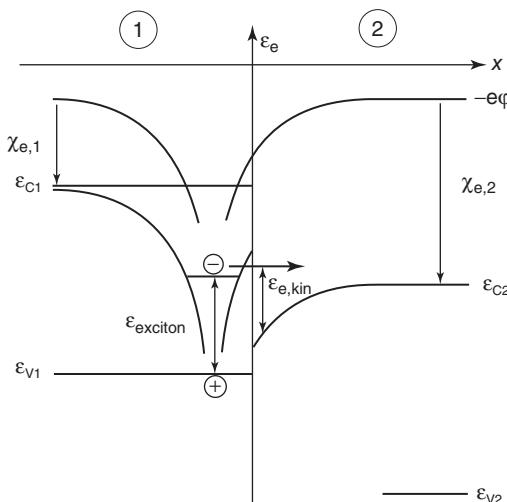


Figure 6.26 An exciton is shown in a material 1, the absorber, on the left as an electron bound to a free hole. The Coulomb potential well of the hole extends into material 2 on the right. This material has a larger electron

affinity $\chi_{e,2}$ than material 1 on the left. As a result, the energy of a free electron $\epsilon_e \geq \epsilon_{c,2}$ is smaller in material 2 than the energy of the electron bound to a free hole in material 1, even if it is far away from the hole.

of the hole and become a really free electron far from the free hole remaining in material 1.

This process of exciton dissociation has a problem. If the electron loses too much of its kinetic energy $\epsilon_{e,\text{kin}}$ by scattering with acoustic vibrations while it is still within the Coulomb well, it becomes trapped at the interface and will stay there until some nonradiative recombination process occurs. If, however, the mean free path of electrons in material 2 is similar to the extension of the Coulomb well, the electron can leave the attraction of the hole without losing much energy. If that is possible, after the dissociation of the exciton, we have a free electron in material 2 and a free hole in material 1 with a combined energy that is nearly the same as the energy of the photon by which the exciton was generated in material 1. Under these favorable conditions, very little energy is lost in the dissociation of the exciton.

6.9.2

Structure of Organic Solar Cells

The combination of two different materials that allows an electron to tunnel from a bound state in the absorber into a free state in the second material also has the properties of a semipermeable membrane for electrons, if the second material does not offer states for free holes at their energy in the absorber as shown in Figure 6.26. In this combination, only the (bound) electrons may pass from the absorber into the electron collector, but the passage of the holes is blocked. The

configuration in Figure 6.26 has the effect of the two different membranes that are necessary for solar cells with inorganic absorbers where free electrons and holes are directly generated by absorbed photons. After the dissociation of the exciton in Figure 6.26, the hole and the electron seem to be in media where there are no free recombination partners, no free electrons in the absorber, and no free holes in the electron collector. For good transport properties of a solar cell, the absorber should be p-doped for good hole conductivity and the electron collector should have good electron conductivity. While the electron collector can then be directly connected to a metallic contact, the absorber should not be in contact with a metal because that would result in nonradiative recombination of the excitons when they come close to the metal.

In a more schematic drawing, Figure 6.27 shows an ideal configuration of an organic solar cell employing an additional hole membrane on the left of the absorber, which allows the holes to pass toward the left to the metal contact, which is represented by its Fermi energy. This additional layer keeps the excitons away from the metal contact. In this ideal configuration, the sum of the energies of the free electron in the electron membrane on the right and the free hole in the hole membrane on the left in Figure 6.27 are almost equal to the energy of the exciton and of the absorbed photon by which it was generated. The use of organic materials in a solar cell does not mean that there is a principal energy loss associated with the dissociation of the excitons.

It would be consistent with the present discussion, if the electron collector on the right of Figure 6.27 would have the same bandgap as the absorber in the middle and therefore would absorb as well. Excitons in the electron collector could dissociate into a free electron in the electron collector and a free hole in the absorber in the middle. The energetic requirements for tunneling are just the same. The

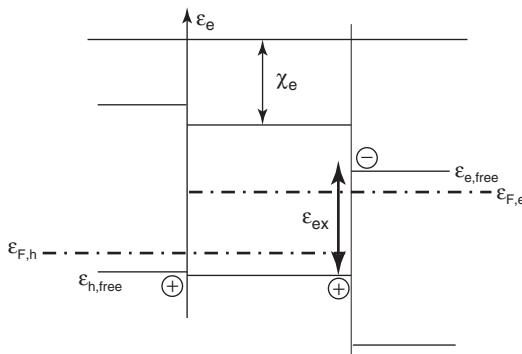


Figure 6.27 An exciton in the absorber in the middle dissociates into a free electron at the interface with the electron membrane on the right, which is then free to move to the metal contact on the right, represented by its Fermi energy $\epsilon_{F,e}$. The free hole that

remains in the valence band of the absorber after exciton dissociation can only leave the absorber through a hole membrane on the left and further on to a metal contact represented by $\epsilon_{F,h}$.

advantage would be that the material within an exciton diffusion length on both sides of the dissociating interface would contribute to the current, which could be twice as large as for absorption in the absorber only.

Although we have emphasized that illumination of an organic absorber generates excitons and no free carriers, Fermi energies for electrons ($\epsilon_{F,e}$) and holes ($\epsilon_{F,h}$) are shown for the absorber in Figure 6.27. It is true that the exciton has a large binding energy. Nevertheless, in steady state, there is chemical equilibrium between the excitons on one side and the electrons and holes on the other side in the absorber and this defines the values of the electron and hole Fermi energies. As always, one type of charge carrier can be fixed by doping, the concentration of the other type will adjust to make the sum of the chemical potentials of electrons and holes in the absorber (the difference of electron and hole Fermi energies) equal to the chemical potential of the excitons. The chemical potential of the excitons or of the electron–hole pairs determines the intensity of photon emission and the rate of radiative recombination by Equation 3.99 in exactly the same way as for inorganic semiconductors. Since in the limit of purely radiative recombination, the thermodynamic limits of the efficiency are approached, these limits are as valid for organic solar cells as for inorganic solar cells.

What we have discussed so far for a p-type absorber can be extended to an n-type absorber as well. In this case, exciton dissociation has to occur into a free hole in the valence band of a hole membrane with a free electron remaining in the conduction band of the absorber. This would be the right choice if the absorber exhibits good n-type conductivity, and if a hole membrane can be employed in which the free hole energy is smaller by the exciton binding energy than the free hole energy in the absorber.

Would it be a good idea to use the two concepts of exciton dissociation into a free electron in an electron membrane and into a free hole in a hole membrane simultaneously? The advantage would be that the absorber could be thicker. The excitons generated close to the front of the absorber would dissociate at one interface, while the excitons generated closer to the rear would dissociate at the other interface. Figure 6.28 shows such a configuration. Different from the previous concepts, both, free electrons and free holes exist in the absorber where they can recombine. Their concentrations, however, comply with the chemical equilibrium between electrons, holes, and excitons discussed earlier and the recombination rate is no larger than in the other concepts. In fact, the recombination rate is even smaller resulting from an inherent energy loss mechanism. The sum of the energies of free electrons in the electron membrane on the right and of free holes in the hole membrane on the left is now smaller than the exciton energy by the binding energy of the exciton. As a consequence, the difference of the electron and hole Fermi energies in the absorber, defining the voltage of the solar cell, is smaller than with only one exciton dissociating interface. Using two interfaces for exciton dissociation results in a principal loss of energy of at least the exciton binding energy and, as a consequence, in a loss of voltage in the solar cell.

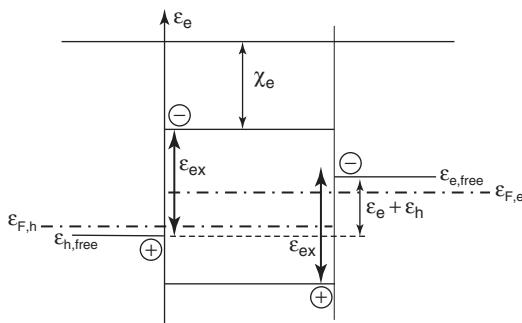


Figure 6.28 An exciton in the absorber in the middle may dissociate into a free electron at the interface with the electron membrane on the right or into a free hole at the

interface with the hole membrane on the left. The absorber must support transport of the remaining holes from right to left and electrons from left to right.

The biggest problem of most organic materials is their short exciton diffusion length of only some nm. This is shorter by a factor of 10–100 than the penetration depth $1/\alpha$ of the light. A planar layer, in which all excitons have an interface for dissociation within reach, must be so thin that it can absorb only a small fraction of the incident light. Figure 6.29(a) shows such a thin absorber layer between a hole and an electron accepting collector. It also shows that folding these layers increases the optical thickness but leaves the condition for efficient exciton

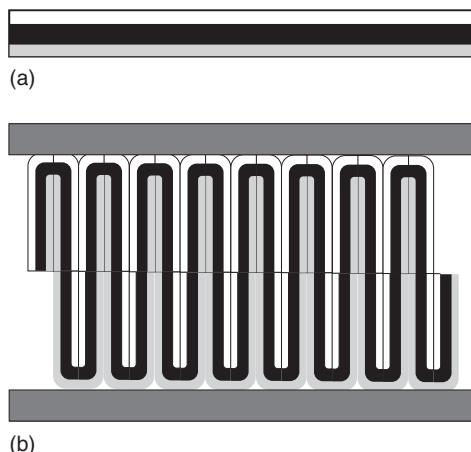


Figure 6.29 (a) In a planar configuration an organic absorber (black) is placed between a hole-conducting membrane (gray) and an electron-conducting membrane (white). Its thickness is less than the exciton diffusion length and therefore insufficient to absorb the absorbable light. (b) By folding the layer

stack from (a), the absorption is increased without increasing the distance, the excitons have to diffuse. Contacts are added to the front and rear side of the folded layer stack. The front contact as well as the electron and hole collecting layers have to be transparent, of course.

transport unchanged. Electrical contact is only made to the outside of the folded package provided the electron and hole collecting layers are sufficiently doped to warrant a small series resistance. With this geometry, a small exciton diffusion length and the required larger absorber thickness can be made compatible. Although more difficult to realize, efficient organic solar cells are not ruled out by physical principles.

Existing examples of organic solar cells incorporate the mechanisms described above. The first heterojunction solar cell where the exciton dissociation at the interface was incorporated in a planar bilayer configuration was presented by Tang [15]. More advanced versions contain electron and hole collector layers [16]. The bulk heterojunction solar cell is another realization of the model presented here. In an interpenetrating network, the p-type absorber (P3HT (Poly(3-hexylthiophene)) or similar) is mixed with an electron collector (PCBM ([6,6]-phenyl-C₆₁-butyric acid methyl ester) or other fullerenes) [17, 18]. As a result, the electron collector may not be everywhere within reach of the excitons or may be in direct contact with the hole collector which may cause shunts. Whether the energetic alignment of all the energy levels involved conforms to the model above is still in question. There seems to be room for improvement.

6.10

Light Emitting Diodes (LED)

We have seen that light emission from solar cells is a result of the difference of the Fermi energies for electrons and holes. This difference may result from the generation of electron–hole pairs by illumination or from the injection of electrons and holes by applying a voltage without illumination. The first case belongs to the operation of a solar cell, the second to the operation of an LED. Both devices are based on the same structure, ideally a semiconductor placed between semipermeable membranes, for holes on one side and for electrons on the other side. If the recombination in a semiconductor is only radiative, the same device is an ideal solar cell and an ideal LED as well. If the thickness of the absorber/emitter is much larger than the penetration depth $1/\alpha$ of the photons and provided that the series resistance is negligible, operation as a solar cell and as an LED would be with maximum efficiency.

The question then arises, why is silicon, one of the best materials for solar cells, unsuitable for LEDs and on the other hand, why are organic LEDs quite efficient while organic solar cells are not?

For a solar cell, the incident intensity is fixed and we are interested in the charge current and voltage from the cell. For an LED, the voltage is fixed and we are interested in the charge current and in the emitted intensity. Two properties of nonideal devices are important: the absorptance and the contribution of nonradiative recombination.

We first discuss the effect of the absorptance, but assume that the recombination is purely radiative. For a solar cell, the absorbed photon current and with it the short-circuit current is proportional to the absorptance, but this absorptance $a(\hbar\omega)$ has to be counted only for a layer from which the minority carriers are able to reach the contacts without recombining. The thickness of this layer is equal to their diffusion length. For an LED, according to Equation 3.102, the emitted photon current is also proportional to the absorptance and again this absorptance is counted for a layer that is filled by minority carriers, when they are injected through their selective contact. The thickness of this layer is equal to the diffusion length. Although, the photon current, either absorbed or emitted, is directly proportional to $a(\hbar\omega)$ for both devices, an absorptance $a(\hbar\omega) < 1$ affects solar cells and LEDs differently. For a solar cell, nonabsorbed photons are lost and result in a proportionate loss of the short-circuit current and of the efficiency. For an LED, $a(\hbar\omega) < 1$ means that less photons are emitted than would be possible for the given voltage, but the charge current is accordingly smaller, since only the electron–hole pairs that recombine have to be replenished by the charge current. There is no direct loss involved. The photon current can be raised to the ideal value expected for $a(\hbar\omega) = 1$ by supplying a little more energy, by raising the voltage in Equation 3.102 by an amount of $kT/e \ln(1/a)$. So, if for an organic absorber/emitter, $a(\hbar\omega) = 0.01$, a solar cell would have an efficiency that is roughly smaller by a factor of 100 than the ideal efficiency and is unacceptably small, whereas for a LED, the small absorptance of this material is fully compensated by an additional voltage of only 120 mV.

The effect of nonradiative recombination is quite different. This time, we assume an ideal absorptance $a(\hbar\omega) = 1$. The short-circuit current of a solar cell is then ideal. Additional recombination by nonradiative transitions, however, reduces the steady-state carrier concentration under open-circuit conditions from its ideal value by a factor that is given by the total recombination rate divided by the radiative recombination rate. For fairly good silicon solar cells, for instance, recombination is 1000 times more probable in nonradiative transitions than in radiative transitions. The minority carrier concentration is reduced by a factor of 1000 and the open-circuit voltage is smaller than its ideal value by $kT/e \ln(1000)$, i.e. by about 180 mV. In contrast, the emitted intensity of this solar cell operated as an LED with a fixed applied voltage is unaffected by nonradiative recombination and is the same as for ideal properties. The charge current, however, which is needed to supply all the carriers that recombine is larger by a factor 1000 than for an ideal device. The efficiency of this silicon-LED is therefore reduced by the same factor 1000 and is unacceptably small.

In conclusion, we see that for solar cells, the efficiency is proportional to the absorptance, but depends only logarithmically on the nonradiative recombination probability. Small absorptivities are prohibitive, whereas some nonradiative recombination can be tolerated. For LEDs, the charge current is directly proportional to the total recombination rate and a large proportion of nonradiative recombination is prohibitive, but small absorptivities can be tolerated, since they are easily compensated by raising the applied voltage by a small amount.

6.11 Problems

- 6.1** What happens to the reverse saturation current density j_S of a pn-junction if the lifetimes of electrons and holes are each improved by a factor of 4.
- 6.2** Calculate the width of the space charge region (depletion zone) of a pn-junction in the one-dimensional Schottky model for an applied voltage of $V_a = -2.3$ V. Use $n_D = 10^{19} \text{ cm}^{-3}$, $n_A = 10^{16} \text{ cm}^{-3}$, and a dielectric function $\epsilon = 11.9$.
- 6.3** Calculate the capacitance of the pn-junction in the last problem without external voltage and for an applied reverse bias of 4 V. From the experimental determination of the capacitance as a function of the applied voltage it is possible to derive the diffusion voltage and, under certain conditions, also the doping concentration. How can that be done?
- 6.4** Assume a Si sample with acceptor-type impurity states with $\tau_{h,\min} = \tau_{e,\min} = 10^{-6} \text{ s}$. Calculate the separation of the quasi-Fermi energies $\epsilon_{FC} - \epsilon_{FV}$ for a temperature of $T = 300 \text{ K}$ as a function of the impurity level ($\epsilon_V \leq \epsilon_{\text{imp}} \leq \epsilon_C$). Consider intrinsic Si and two different p-doping regimes with corresponding Fermi energies in the dark of $\epsilon_F^0 = \epsilon_i$, $\epsilon_F^0 = \epsilon_i - 0.2 \text{ eV}$ and $\epsilon_F^0 = \epsilon_i - 0.4 \text{ eV}$ for a generation rate
(a) $G = 10^{18} \text{ cm}^{-3} \text{s}^{-1}$
(b) $G = 5 \times 10^{20} \text{ cm}^{-3} \text{s}^{-1}$.
 Use $N_C = 3 \times 10^{19} \text{ cm}^{-3}$, $N_V = 10^{19} \text{ cm}^{-3}$ and $\epsilon_G = 1.12 \text{ eV}$. The electrons and holes shall be distributed homogeneously throughout the sample due to large diffusion lengths.
 Plot a graph of the results as a function of the impurity level.
- 6.5** Does the conservation of energy prohibit that a LED emits photons with an energy $\hbar\omega > \epsilon_G$ when a voltage V is applied with $V < \epsilon_G$?

7

Limitations on Energy Conversion in Solar Cells

In connection with the derivation of the current–voltage characteristic, we have neglected a voltage drop over the transport resistances as being negligibly small. With this approximation, the voltage V at the contacts of a sufficiently doped solar cell is given by the separation of the Fermi energies, $eV = \varepsilon_{FC} - \varepsilon_{FV}$. In addition, we did not consider currents of minority carriers flowing in the wrong direction, in spite of the large gradients of their Fermi energies, because of the small conductivity of the minority carriers in regions that function as semipermeable membranes for the majority carriers.

With this approximation, all electrons and holes produced by the illumination, which do not recombine, contribute to the charge current. For homogeneous excitation at short circuit, the electron–hole pairs contributing to the charge current are those which are produced within the diffusion lengths. This approximation is justified especially in structures such as in Figure 6.5, in which the discontinuities at the band edges prevent the flow of minority charge carriers to the wrong side.

A solar cell structure consisting of an absorber with n- and p-type membranes supplies a charge current at a voltage V with $eV = \varepsilon_{FC} - \varepsilon_{FV} = \mu_e + \mu_h$. It supplies electrical energy equal to the chemical energy $\mu_e + \mu_h$ per electron–hole pair. This structure is able to convert the chemical energy produced by illumination of a semiconductor absorber *entirely* into electrical energy.

7.1

Maximum Efficiency of Solar Cells

The maximum energy current delivered by a solar cell is given by the largest rectangle fitting under the current–voltage characteristic, as shown in Figure 6.10. It defines the “maximum power point” for the charge current density j_{mp} and the voltage V_{mp} . For a given current–voltage characteristic, it is therefore important to have an algorithm to find the maximum power point.

Independent of the form of the characteristic, the functional relationship between j_Q and V , the condition for maximum power yields

$$d(j_Q V) = dj_Q V + j_Q dV = 0$$

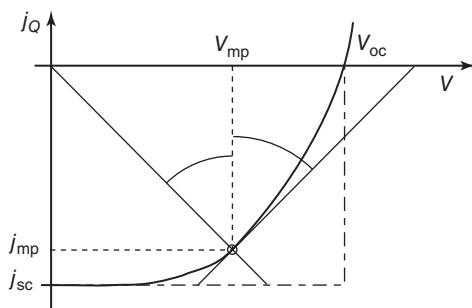


Figure 7.1 Geometrical construction of the maximum power point.

and thus

$$\left(\frac{dj_Q}{dV} \right)_{mp} = - \left(\frac{j_Q}{V} \right)_{mp} \quad (7.1)$$

This relationship is illustrated geometrically in Figure 7.1. From this construction, the maximum power point occurs where the tangent to the characteristic makes the same angle with a vertical line as the line connecting the coordinate origin and the maximum power point.

For a purely exponential current-voltage characteristic, as for radiative recombination, in Equation (6.27)

$$\frac{dj_Q}{dV} = j_s \frac{e}{kT} \exp\left(\frac{eV_{mp}}{kT}\right) = - \frac{j_{mp}}{V_{mp}} \quad (7.2)$$

With

$$j_{mp} = j_s \left\{ \exp\left(\frac{eV_{mp}}{kT}\right) - 1 \right\} + j_{sc}$$

and

$$\frac{j_{sc}}{j_s} = 1 - \exp\left(\frac{eV_{oc}}{kT}\right)$$

we find from Equation 7.2

$$V_{mp} = \frac{kT}{e} \left\{ \exp\left(\frac{e(V_{oc} - V_{mp})}{kT}\right) - 1 \right\} \quad (7.3)$$

Solutions of this equation, and thus the maximum power point, can be found by numerical techniques. With the fill factor

$$FF = \frac{j_{mp} V_{mp}}{j_{sc} V_{oc}} \quad (7.4)$$

we define a measure for how well the maximum power rectangle fits under the characteristic. For only radiative recombination $j_{sc} V_{oc}$ represents the chemical energy current emitted by the photons in the open-circuit state, in which all electrons and holes must recombine.

An approximate value for the fill factor can be obtained from Equation 7.3. This gives us

$$V_{\text{mp}} = V_{\text{oc}} - \frac{kT}{e} \ln \left(1 + \frac{eV_{\text{mp}}}{kT} \right) \approx V_{\text{oc}} - \frac{kT}{e} \ln \left(1 + \frac{eV_{\text{oc}}}{kT} \right)$$

Since the logarithm depends only weakly on its argument, we have substituted V_{oc} for V_{mp} in the logarithm. With this result for V_{mp} , we derive the current j_{mp} at the maximum power point from the characteristic and find an approximate fill factor

$$FF = \frac{eV_{\text{oc}}/kT - \ln(1 + eV_{\text{oc}}/kT)}{1 + eV_{\text{oc}}/kT} \quad (7.5)$$

This dependence of the fill factor on the open-circuit voltage V_{oc} also means that the fill factor increases with the bandgap (of the absorber material) of the solar cell. Values of the fill factor of good solar cells are between 0.8 and 0.9.

Maximum Short-circuit Current

A large short-circuit current requires a solar cell as thick as possible to maximize its absorptance. With antireflection coatings, we can theoretically reduce the reflectance to $r = 0$. For a cell with large thickness and at the same time a large diffusion length, the absorptance over the diffusion length is given by $a(\hbar\omega \geq \epsilon_G) \approx 1$. The short-circuit current produced by the absorbed photon current is then

$$j_{\text{sc}} = -e \int_0^{\infty} a(\hbar\omega) dj_{\gamma, \text{Sun}}(\hbar\omega) = -e \int_{\epsilon_G}^{\infty} dj_{\gamma, \text{Sun}}(\hbar\omega) \quad (7.6)$$

Maximum Open-circuit Voltage

The open-circuit voltage V_{oc} defines the separation $\epsilon_{\text{FC}} - \epsilon_{\text{FV}}$ of the Fermi energies at which recombination is in equilibrium with electron–hole generation throughout the entire cell. Owing to the exponential decay of the photon current density within the semiconductor, the rate of generation (per volume) is greatest at the surface. The electrons and holes produced are distributed by diffusion more or less uniformly over the thickness, depending on the diffusion length. The recombination rate is then equal to the averaged generation rate everywhere. When the thickness of the solar cell is reduced and surface recombination is prevented, the recombination rate (per volume) and with it $\epsilon_{\text{FC}} - \epsilon_{\text{FV}}$ must increase, because the averaged generation rate increases with decreasing cell thickness. The open-circuit voltage reaches a maximum when the thickness of the cell goes to zero. The open-circuit voltage increases only slightly, however, with decreasing thickness and this in no way compensates for the short-circuit current loss. In any case, a cell should not be made unnecessarily thick, and not merely to save on material. The optimum thickness is reached when a further increase in the thickness causes as much additional recombination at maximum power as it provides additional generation by more absorbed photons.

The greatest voltage is achieved if there is only radiative recombination, and then, surprisingly, the thickness of thick solar cells no longer plays a role. It is true that the total recombination rate increases in proportion to the volume. In thick

cells, however, a large part of the photons produced by recombination does not reach the surface and is reabsorbed, producing electron–hole pairs again. The effective recombination rate integrated over the entire cell is equal to the photon current $j_{\gamma, \text{emit}}$ emitted through the surface, which reaches a saturation value in thick cells when the absorptance reaches its maximum value $a = 1 - r$, which no longer depends on the thickness.

If the diffusion lengths are large compared with the thickness, the electrons and holes are uniformly distributed over the volume of the cell. For such a homogeneous distribution, we have in the approximation to the generalized Planck equation

$$j_{\gamma, \text{emit}} = \int_0^{\infty} a(\hbar\omega) d\gamma^0(\hbar\omega) \exp\left(\frac{\epsilon_{\text{FC}} - \epsilon_{\text{FV}}}{kT}\right) \quad (7.7)$$

For the maximum short-circuit current, $a(\hbar\omega \geq \epsilon_G) = 1$, and replacing $\epsilon_{\text{FC}} - \epsilon_{\text{FV}}$ by eV, the total rate of radiative recombination is

$$j_{\gamma, \text{emit}} = \exp\left(\frac{eV}{kT}\right) \int_{\epsilon_G}^{\infty} d\gamma^0(\hbar\omega) \quad (7.8)$$

independent of the thickness.

Defining the charge current from the p-region to the n-region (toward the left in Figure 6.5) as positive, the charge current delivered by the solar cell is

$$j_Q = ej_{\gamma, \text{emit}}(V) - ej_{\gamma, \text{abs}} \quad (7.9)$$

or

$$j_Q = e \left[\exp\left(\frac{eV}{kT}\right) - 1 \right] \int_{\epsilon_G}^{\infty} d\gamma^0(\hbar\omega) - e \int_{\epsilon_G}^{\infty} d\gamma_{\gamma, \text{Sun}}(\hbar\omega) \quad (7.10)$$

Since the spectra of the 300 K background radiation $d\gamma^0(\hbar\omega)$ and of the Sun $d\gamma_{\gamma, \text{Sun}}(\hbar\omega)$ (outside the atmosphere AM0 and on the surface of the Earth AM1.5) are known, the maximum energy current $(j_Q V)_{\max} = j_{\text{mp}} V_{\text{mp}}$ can be determined from Equation 7.10 and, from this, the efficiency

$$\eta = \frac{j_{\text{mp}} V_{\text{mp}}}{\int_0^{\infty} \hbar\omega d\gamma_{\gamma, \text{Sun}}(\hbar\omega)} \quad (7.11)$$

7.2

Efficiency of Solar Cells as a Function of Their Energy Gap

The short-circuit current of a solar cell depends on the absorbed photon current. It is a maximum for a semiconductor with an energy gap $\epsilon_G = 0$ and decreases with increasing ϵ_G . The open-circuit voltage V_{oc} is, however, zero for $\epsilon_G = 0$ and increases with increasing energy gap. The efficiency η is therefore zero at $\epsilon_G = 0$ and at $\epsilon_G \rightarrow \infty$. Somewhere in between is its maximum. From Equations 7.10 and 7.11, we can calculate the efficiency η as a function of the energy gap ϵ_G

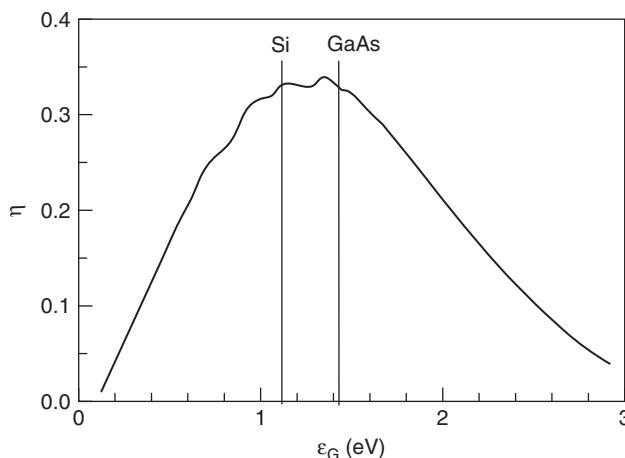


Figure 7.2 Efficiency of solar cells with radiative recombination only as a function of the energy gap for the AM1.5 spectrum.

if only radiative recombination takes place in the case of thick cells, in which $a(\hbar\omega < \epsilon_G) = 0$ and $a(\hbar\omega \geq \epsilon_G) = 1$.

The result for the AM0 spectrum outside the atmosphere was already shown in Figure 4.7. Figure 7.2 shows the maximum efficiency for the AM1.5 spectrum on the surface of the earth. We find a broad maximum, which indicates that semiconductors with an energy gap ϵ_G between 1 and 1.5 eV are suitable for solar cells. For the AM1.5 spectrum, the maximum efficiencies are greater than for the AM0 spectrum, because absorption in the atmosphere mainly eliminates photons with $\hbar\omega < 1$ eV, which cannot be utilized by solar cells with optimal band gaps. Silicon and gallium arsenide are especially well suited for the AM1.5 spectrum.

7.3 The Optimal Silicon Solar Cell

Silicon has many advantages. It is the second most abundant element in the Earth's crust and is thus available in practically unlimited amounts. Silicon is not toxic. On exposure to air, silicon forms an oxide surface layer that protects it fully and prevents any further corrosion. The interface between Si and SiO_2 when grown under clean-room conditions has a very low density of surface states resulting in a very low surface recombination velocity. With $\epsilon_G = 1.12$ eV, silicon has a favorable energy gap for the conversion of solar energy. Although silicon has all these advantages, with its indirect optical transitions it has the serious disadvantage of weak absorption. Consequently, silicon must be much thicker than a semiconductor with direct transitions. Moreover, owing to the weak absorption, the generation of electron–hole pairs is distributed over a large penetration depth $1/\alpha$ of the photons, and at least one kind of carrier must diffuse over a large distance in order

to reach the contact. This implies that this carrier type must also have a large diffusion length and lifetime. Because of the poor absorption, not only is more silicon needed but it must also have a higher purity than if the optical transitions were direct.

In the usual structure of a solar cell, the contacts are applied to opposite surfaces. For the illuminated surface, a nontransparent contact poses a problem. Metal contacts are therefore arranged in narrow strips, in comblike structures leaving most of the surface uncovered. The charge must then flow in the membrane layer toward the contact strips parallel to the surface. A high doping level is required in this layer to avoid an intolerably large series resistance. Adjacent to the thin membrane layer is a weakly doped absorber region over the greatest part of the solar cell thickness. Since electrons have a greater mobility than holes and thus a greater diffusion length for a given lifetime, they are chosen to be the minority carriers. The large middle region of the cell is therefore p-doped and the front surface is strongly n-doped, with the designation n^+ , to function as an electron membrane. To minimize the loss of electrons by surface recombination at the back contact, strong p-doping is used in front of the back contact to establish a hole membrane. The reduction of the recombination at the rear contact is commonly attributed to the so-called back surface field originating from the negative charge of the p^+ -doped region, thought to repel the electrons. This repulsion is, however, not recognizable in the total force ($\text{grad } \eta_e$) and the smaller recombination probability is instead due to the reduced concentration of electrons in the p^+ layer.

Owing to the high conductivity of the rear p^+ layer a metal contact is not required over the entire rear surface. The areas of the front and rear surfaces without metal contact are covered with a passivating layer, silicon dioxide or silicon nitride, in order to reduce the rate of surface recombination. On the rear surface, the oxide layer is covered by a metal mirror layer that reflects the photons not yet absorbed and thus enhances absorption in the cell. The oxide or nitride layer on the front side is in the form of a $\lambda/4$ layer to reduce reflection for the wavelength range around λ in the infrared and red part of the spectrum, where the absorbable photon current has its maximum. Reflection in the blue is less reduced, which gives silicon solar cells their characteristic blue appearance.

Figure 7.3 illustrates a cross section through this structure. The current–voltage characteristic in Figure 7.4 was calculated for the AM1.5 spectrum, assuming that the front side is nonreflecting and only the unavoidable radiative recombination and Auger recombination corresponding to the necessary doping concentrations are considered, but not surface recombination. This cell has an efficiency of 25%.

7.3.1

Light Trapping

The absorptance of a body increases as the reflectance is reduced and the path of the photons within the body becomes longer. This triviality enables us to consider

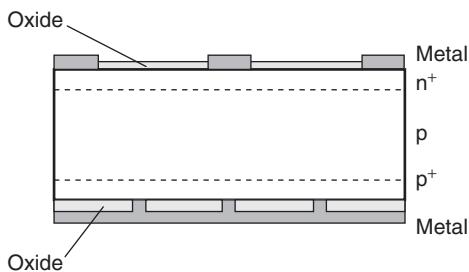


Figure 7.3 Cross-section of a silicon pn solar cell.

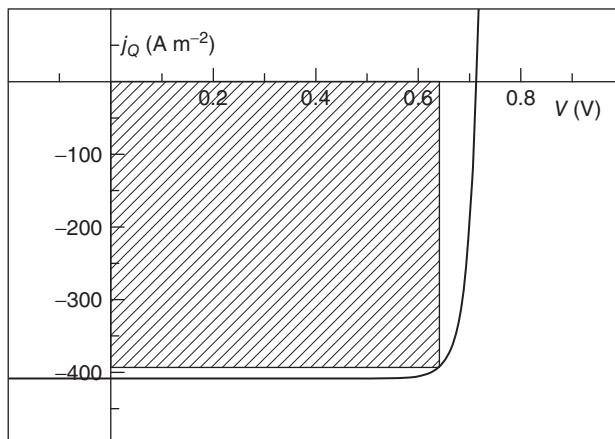


Figure 7.4 Charge current j_Q as a function of the voltage V for a Si solar cell with a thickness of 400 μm illuminated by the AM1.5 spectrum. The maximum power given by the rectangle corresponds to an efficiency of 25%.

still another way to improve the absorptance than by an antireflection coating and a large thickness. The reflectance of a body decreases when the reflected photons are deflected in such a way that they impinge on the body a second time. The pyramid-shaped structure of Figure 7.5 makes this possible. For light reflected twice, the total reflectance is given by

$$r_{\text{total}} = r_{\text{in}}^2$$

A surface with 10% reflection as a planar surface reflects only 1% in a structure, where each reflected photon hits the surface a second time.

In addition, as a result of the textured surface, together with a reflecting rear surface, the light path in a solar cell is considerably increased compared with normal incidence on a planar surface. The photons entering the cell are deflected in an oblique direction because of refraction at the textured surface. And, even more important, after reflection at the rear surface, there is a high probability of their impinging on the surface from within the cell at such an angle that they experience total internal reflection.

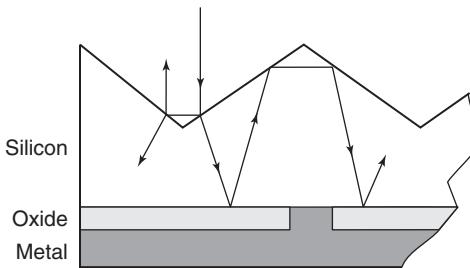


Figure 7.5 Surface texture for reducing reflection and increasing the length of the light path.

At the critical angle for total internal reflection

$$\sin \alpha_T = \frac{1}{n_{Si}}$$

Since the index of refraction for silicon with $n_{Si} = 3.5$ and for most other solar cell materials is very large, only those photons that strike the surface at an angle of less than $\alpha_T = 16.6^\circ$ to the local surface normal are not totally reflected. Most photons are therefore trapped and, if they are not absorbed, leave the solar cell only after multiple reflections when they strike the surface at an angle of less than 16.6° . We can easily estimate how much longer the mean light path will be, if we assume that the passage through the textured surface, together with reflection at the rear surface, leads to an isotropic distribution of the weakly absorbed photons in the solar cell. The photons that are emitted through the surface then have the same angular distribution with which a black-body Lambertian surface emits photons. For this angular distribution, the photons leaving through the front surface cover an effective solid angle of π outside the solar cell (for the assumed large index of refraction), as described in Section 2.1.4. The isotropic distribution internally fills the solid angle 4π , in which the photon current density per solid angle $j_{\gamma,\Omega}$ is greater by a factor of n^2 than outside the solar cell.

For a Si cell of thickness L with a surface area A the absorptance is defined as

$$a(\hbar\omega) = \frac{I_{E,abs}(\hbar\omega)}{I_{E,inc}(\hbar\omega)} = \frac{I_{\gamma,abs}(\hbar\omega)}{I_{\gamma,inc}(\hbar\omega)}$$

The absorbed photon current $I_{\gamma,abs}$ is found from the balance of the photon currents. Incident photons can either be reflected, absorbed, or scattered and redirected to emerge back out through the surface. This assumes that all photons are reflected from the rear surface. In this balance, we neglect the emission of photons within the absorber [19].

$$(1 - r)I_{\gamma,inc} = I_{\gamma,em} + I_{\gamma,abs}$$

For the assumed isotropic and homogeneous distribution of photons with the photon current density $j_{\gamma,\Omega}$ per solid angle, the photon current absorbed in the volume $V = AL$ is

$$I_{\gamma,abs} = 4\pi\alpha V j_{\gamma,\Omega}$$

The photon current leaving through the surface is

$$I_{\gamma,\text{em}} = A (1 - r) \frac{\pi}{n^2} j_{\gamma,\Omega}$$

The absorptance for perfect light trapping is therefore

$$\alpha_{\text{trap}} = \frac{I_{\gamma,\text{abs}}}{I_{\gamma,\text{inc}}} = (1 - r) \frac{I_{\gamma,\text{abs}}}{I_{\gamma,\text{em}} + I_{\gamma,\text{abs}}}$$

$$\alpha_{\text{trap}} = (1 - r) \frac{4\pi\alpha L}{(1 - r)\pi/n^2 + 4\pi\alpha L}$$

$$\alpha_{\text{trap}} = \frac{1 - r}{(1 - r)/(4n^2\alpha L) + 1} \quad (7.12)$$

For small values of the absorption coefficient α , we find the absorptance $\alpha_{\text{trap}} = 4n^2\alpha L$. For a reflectance r that is not too large, it is surprising that the absorptance does not depend on the reflectance. For silicon, α_{trap} is a factor of $4n_{\text{Si}}^2 \approx 50$ greater than the absorptance without light trapping. Relative to a single pass along the normal to the surface, the mean light path is then enhanced by this factor of $4n_{\text{Si}}^2 \approx 50$.

Although the above derivation is based on a homogeneous distribution of photons, which is a good approximation only when $\alpha \ll 1/L$, Equation 7.12 is also valid to a high degree of accuracy for larger values of α since the saturation value $\alpha_{\text{trap}} = 1 - r$ is already reached for $\alpha < 1/L$ and cannot increase further, even for larger values of α . Figure 7.6 shows that, by light trapping, high absorptance values can be obtained even for thin silicon layers in spite of silicon's small absorption coefficient.

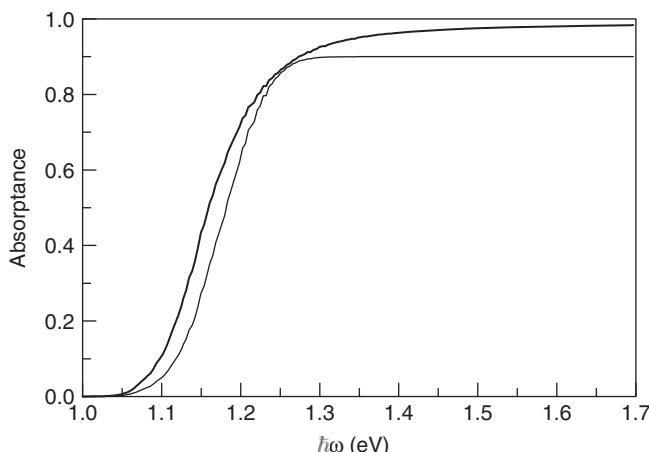


Figure 7.6 Absorptance α as a function of the photon energy $\hbar\omega$ for 20 μm thick silicon with light trapping (heavy line) and for 400 μm thick silicon without light trapping (thin line), assuming a reflectance of $r = 0.1$ in both cases.

The enhancement of the absorptance by light trapping can reach even greater values. In the derivation above, we have assumed that the photons leaving through the surface are emitted into the entire hemisphere, that is, into an effective solid angle π as a result of the Lambertian behavior of the scattering surface. Because of the reversibility of light paths, incident radiation from the entire hemisphere is able to enter the cell at the same time through the surface. For this type of light trapping, the solar cell does not have to track the Sun's position.

According to the discussion in Section 2.4.3, the maximum concentration of solar radiation is obtained when the photons emitted from the surface of an absorber are not emitted into the entire hemisphere but only toward the Sun. Theoretically, it is conceivable that with an appropriate surface structure, a photonic crystal structure, photons leaving through the surface will be directed only toward the Sun. The density of the photons in the semiconductor that are not absorbed or are only weakly absorbed then increases to a value that is larger by a factor n^2 than the photon density at the surface of the Sun. Such a structure would, however, not change the incident photon current density. It is a structure for minimum emission rather than for maximum concentration. A solar cell that only exchanges radiation with the Sun would of course have to track the position of the Sun. If we make this effort, the absorptance for absorbed photons with $\hbar\omega > \varepsilon_G$ would increase much more steeply than shown in Figure 7.6, and silicon solar cells could be made still thinner. For purely radiative recombination, a greater voltage would result, while the absorbed photon current, and with it the charge current of the solar cell, would only be slightly improved.

With the light-trapping technique, very thin solar cells can be realized with crystalline silicon, in spite of its low absorption coefficient. However, we must consider that the estimated increase in the mean light path is based on the incoherent scattering of photons, valid only when the dimensions of the surface structure and the thickness of the cell are large compared with the wavelength. For smaller structures, coherence and interference phenomena must be considered. In Figure 7.6, we see that a 20 μm thick silicon film with light trapping has a higher absorptance than a 400 μm thick wafer without light trapping. A higher absorptance results in a larger short-circuit current. In addition, the open-circuit voltage is larger for the thin cell, because impurity recombination and Auger recombination have a smaller probability in a smaller volume. The probability for radiative recombination, however, increases in comparison with the other recombination processes, since it depends only on the absorptance. Decreasing the thickness while maintaining a high absorptance makes the solar cell more ideal.

A high-efficiency, thin-film silicon solar cell would be a great achievement because of silicon's favorable properties for the environment and its chemical stability. One problem for future development is that thin silicon films require a substrate for support. The only known substrate on which single crystal silicon films can be grown is, however, crystalline silicon. Although only the crystal structure of the substrate is needed and not the purity required for a solar cell, it is still an expensive substrate. Recently, polycrystalline silicon films have been grown on a variety of substrates including glass, which is a very promising development.

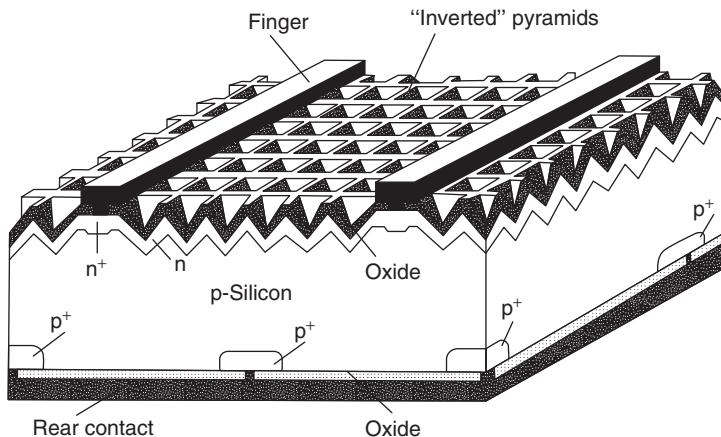


Figure 7.7 Structure of the best silicon solar cell manufactured to date with an efficiency of 24.4%, developed by M.A. Green's group at UNSW. (Courtesy of M.A. Green).

Light trapping by textured surfaces also improves the properties of thick cells by reducing reflection and enhancing the absorption of photons with $\hbar\omega \approx \epsilon_G$. The best silicon solar cell for the unconcentrated $AM1.5$ spectrum is made from very pure silicon and has all the properties mentioned above for an optimal Si solar cell [20]. Its efficiency is 24.4%. Figure 7.7 illustrates its structure.

7.4

Thin-film Solar Cells

Silicon has so many advantages for solar cells that other materials can compete only when they do not share its disadvantage, the poor absorption of light. In materials competing with silicon, the transitions between the valence and conduction bands must be direct. The absorption coefficient then has a large value. For the absorption of that part of the solar spectrum that can be absorbed, a thickness of only a few micrometers is sufficient for thin-film solar cells. For the same number of recombination centers as in a thick silicon cell, a higher impurity concentration and the presence of grain boundaries in the film can be tolerated. Because of the smaller distances to the membranes at the surfaces, the diffusion lengths can also be smaller. This allows the use of materials with lower mobility. All these advantages hold the promise of significant cost reduction in the production of solar cells.

Owing to the close proximity to the surface with its high surface recombination, a well-developed membrane is required at least on the front surface. Since electron–hole pairs should not be generated in this layer, it must have a large energy gap. It is called a window layer, through which photons pass unimpeded, but which protects electrons and holes from recombining at the front contact. The interface between the window layer and the absorber should have a low density of interface states in order to prevent recombination there.

A disadvantage of many materials with direct transitions and favorable energy gaps, with the exception of amorphous silicon, is that they cannot be

doped equally well n-type and p-type. The structure required for solar cells then demands heterojunctions. Examples include the combination of n-type CdS/p-type CuInSe₂ (abbreviated as (CIS)) or n-type CdS/p-type CdTe, where the window material CdS is the electron membrane in both cases.

It is interesting to note that amorphous silicon (a-Si) also belongs to the class of thin-film materials. Amorphous silicon is silicon without a crystalline structure. Because of the lack of long-range order, i.e. structural uniformity is found only over very small volumes, by the uncertainty principle of Equation 2.4, the momentum of electrons in bound states (valence band) and unbound states (conduction band) is largely undetermined. As a result, no phonons are required for transitions between these states in order to satisfy the conservation of momentum. The transitions are direct and have large absorption coefficients. However, the lack of order has the disadvantage that the states for electrons and holes are not confined to bands; the states fill the entire forbidden gap. The inclusion of about 10% hydrogen (a-Si:H) serves to saturate many of the dangling bonds of the silicon atoms in the amorphous structure. The density of states in the forbidden gap is drastically reduced, and the material can now be doped. However, the saturation of dangling bonds with hydrogen is not totally stable. During illumination, the bonds are again broken by the capture of holes. This property, known as *the Staebler–Wronski effect*, leads to a continuous decrease of the efficiency of solar cells made of a-Si:H. Hydrogenated amorphous silicon films have gained enormous importance as semi-permeable membranes on crystalline silicon solar cells. Their interface with crystalline silicon has a low surface recombination velocity, mainly because of the large amount of hydrogen incorporated in them, which saturates dangling bonds. Highly n-doped, they ensure good electron transport to the electron contact and block the holes more efficiently than possible with crystalline silicon, because of the larger bandgap of amorphous silicon. Highly p-doped, they are good selective transport layers for holes.

7.4.1

Minimal Thickness of a Solar Cell

The thickness of a solar cell is an important issue. It is not only that a larger amount of precious material is needed for a thicker cell but also that a thinner cell could tolerate less optimal material properties. Organic materials could be very useful for solar cells because of their good absorption properties and good luminescent quantum yields, indicating dominant radiative recombination. Their drawbacks are the very small mobilities and diffusion lengths of electrons and holes. In the common plane arrangement of an absorber between two membranes, shown in Figure 7.8(a), the diffusion lengths must be larger than the thickness of the absorber and the thickness must be larger than the penetration depth $1/\alpha$ of the photons. If this were a generally necessary condition, then many organic materials would never make good solar cells. In fact, these conditions are sufficient, but not really necessary. The necessary condition for the absorption is that there is enough absorber material to absorb the light; how the material is arranged is not important. Two separate thin layers absorb as much as a single layer that is twice

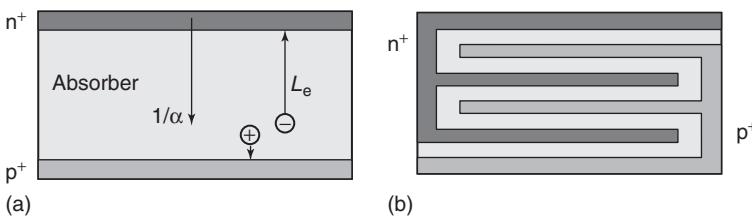


Figure 7.8 (a) In the plane arrangement of an absorber between electron and hole membranes, the diffusion lengths $L_{e, h}$ must be larger than the thickness of the absorber and the thickness must be larger than the

penetration depth $1/\alpha$ of the photons.
(b) Many absorbing layers in a meander-like structure combine good absorption with a small distance between the membranes.

as thick. The necessary condition for the transport of electrons and holes is that they must be able to reach the membranes, requiring the diffusion lengths to be larger than the distance between the membranes. The arrangement in Figure 7.8b fulfills the conditions of absorption and transport as well and would allow the use of absorber materials with arbitrarily small diffusion lengths. This principle of separating the absorber thickness from the distance between the membranes is successfully realized in the dye solar cell described in Chapter 5. There, the distance between the membranes, TiO_2 for the electrons and the I^-/I_3^- -redox system for the holes, has a minimal value, the thickness of the monomolecular dye layer. Nevertheless, many dye layers provide for sufficient absorption.

The organic or plastic solar cell is another example [21]. In this solar cell, a p-type polymer serves both as an absorber and as a hole membrane. The electrons tunnel from their bound exciton state in the absorber to a mobile state in n-type fullerene molecules while the holes stay in the polymer. The polymer and the fullerene are thoroughly mixed into a blend to facilitate the charge carrier separation. This makes the arrangement different from Figure 7.8b, where care was taken that the electron membrane makes contact with only one of the electrodes and the hole membrane only with the other electrode. With the mixing of electron and hole membranes in the plastic solar cell, as well as in the dye cell, a strong inhibition of the exchange of one carrier type must be present at each of the electrodes to prevent internal shunts. In the dye cell, hole exchange between the transparent electrode on the TiO_2 and the redox system is very poor, but is excellent with the small platinum islands on the counterelectrode.

With the small distance between the membranes and the resulting large interface area, interface recombination is enhanced and may be a problem.

7.5 Equivalent Circuit

In the current–voltage characteristic for the solar cell in Equation 6.35, we can regard the current I_Q as the sum of the current through the pn-junction in the dark

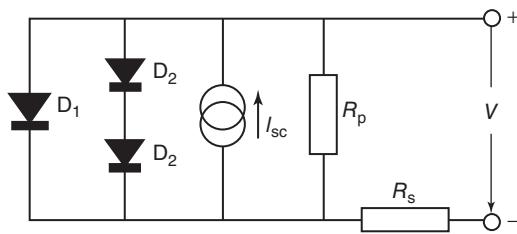


Figure 7.9 Equivalent circuit for a solar cell consisting of (from left) diode D_1 with direct recombination, diodes D_2 with impurity recombination, current source I_{sc} , parallel resistance R_p and series resistance R_s .

and the current I_{sc} from a current source, connected in parallel for the currents to add.

Figure 7.9 shows the equivalent circuit diagram, extended by two additional elements. The resistance R_p , in parallel with the two diodes of the two-diode model, represents the shunts that can occur in real solar cells across the surfaces, at pin-holes in the pn-junction or at grain boundaries. The series resistance R_s accounts for all voltage drops across the transport resistances of the solar cell and its connections to a load. The current–voltage characteristic then takes the form

$$I_Q = I_{S1} \left[\exp \left(\frac{e(V - I_Q R_S)}{kT} \right) - 1 \right] + I_{S2} \left[\exp \left(\frac{e(V - I_Q R_S)}{2kT} \right) - 1 \right] + I_{sc} + \frac{V - I_Q R_S}{R_p} \quad (7.13)$$

Figure 7.10 illustrates how the characteristic is affected by varying R_p and R_s separately. The effect of the series resistance alone is a displacement of the

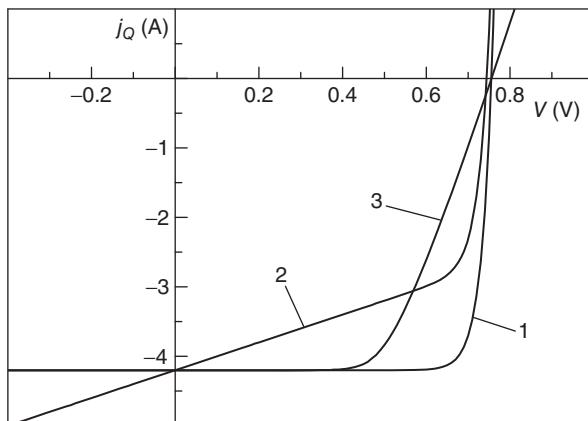


Figure 7.10 Current–voltage characteristic of a 100 cm^2 solar cell with (2) $R_s = 0\ \Omega$, $R_p = 0.5\ \Omega$; and (3) $R_s = 0.05\ \Omega$, $R_p = \infty$, compared with (1) $R_s = 0\ \Omega$ and $R_p = \infty$.

characteristic in the direction of lower voltages, in proportion to the current. The effect of the parallel resistance alone is a displacement of the characteristic to higher positive currents, in proportion to the voltage. Both effects, separately and combined, lead to a smaller fill factor FF .

7.6

Temperature Dependence of the Open-circuit Voltage

Solar cells deliver only a small part of the absorbed energy current as electrical energy to a load. The remainder is dissipated as heat; the solar cell must therefore have a higher temperature than the environment. For solar irradiation of 1 kW m^{-2} , the temperature difference to the environment may be some 10 K.

Heating reduces the size of the energy gap. The absorbed photon current increases, leading to a slight increase in the short-circuit current j_{sc} . The heating has a detrimental effect on the open-circuit voltage. From

$$V_{oc} = \frac{1}{e}(\eta_e + \eta_h) = \frac{kT}{e} \ln\left(\frac{n_e n_h}{n_i^2}\right) \quad (7.14)$$

we find for the temperature dependence

$$\frac{dV_{oc}}{dT} = \frac{k}{e} \ln\left(\frac{n_e n_h}{n_i^2}\right) + \frac{kT}{e} \left[\frac{1}{n_e} \frac{dn_e}{dT} + \frac{1}{n_h} \frac{dn_h}{dT} - \frac{1}{n_i^2} \frac{d(n_i^2)}{dT} \right] \quad (7.15)$$

Here

$$n_i^2 = N_C N_V \exp\left(-\frac{\epsilon_G}{kT}\right)$$

and

$$\frac{d(n_i^2)}{dT} = \frac{\epsilon_G}{kT^2} n_i^2$$

It follows that

$$\frac{dV_{oc}}{dT} = \frac{V_{oc} - \epsilon_G/e}{T} + \frac{kT}{e} \left(\frac{1}{n_e} \frac{dn_e}{dT} + \frac{1}{n_h} \frac{dn_h}{dT} \right) \quad (7.16)$$

No general statements can be made about the expressions in brackets, except that both are probably < 0 and that the first expression is negligible in a n-conductor and the second expression in a p-conductor. The temperature dependence is caused essentially by $(V_{oc} - \epsilon_G/e)/T$. The temperature dependence is more pronounced in bad cells, where V_{oc} is small.

For a silicon cell with $V_{oc} = 0.6 \text{ V}$ and $\epsilon_G = 1.12 \text{ eV}$ at $T = 300 \text{ K}$, $dV_{oc}/dT = -1.7 \text{ mV K}^{-1}$. This means that the open-circuit voltage decreases by 0.3% per degree of temperature rise. A temperature increase of 50 K reduces the open-circuit voltage by 85 mV, that is, by 14%. The efficiency is reduced accordingly.

7.7

Intensity Dependence of the Efficiency

We know from earlier discussions that maximum efficiencies are obtained for maximum concentration of the incident radiation. Since concentration means that a smaller area is required for a given energy current, it is an option for expensive solar cell materials. That an increase of the efficiency can be expected is a bonus.

The short-circuit current j_{sc} is simply given by the absorbed photon current and increases, as expressed by Equation 7.6, proportional to the intensity.

$$j_{sc} = -e j_{\gamma, \text{abs}} \quad (7.17)$$

The open-circuit voltage V_{oc} defines the difference between the Fermi energies at which the total recombination rate in the cell is equal to the total generation rate given by the absorbed photon current. With $j_{\gamma, \text{emit}} = j_{\gamma, \text{abs}}$, we find from Equation 7.8

$$V_{oc} = \frac{kT}{e} \ln \left(1 + \frac{e j_{\gamma, \text{abs}}}{j_s} \right) \quad (7.18)$$

where j_s is the reverse saturation current. For $j_{\gamma, \text{abs}} \gg j_s$, true for any observable light intensity, the open-circuit voltage V_{oc} increases logarithmically with the intensity.

The fill factor, finally, as given by Equation 7.5 can be approximated for $V_{oc} \gg kT$ roughly by

$$FF = 1 - \frac{kT}{V_{oc}} \quad (7.19)$$

and increases very slightly with the open-circuit voltage. To a first approximation, the increase of the fill factor with the intensity can be neglected.

In the efficiency

$$\eta = \frac{FF j_{sc} V_{oc}}{\int_0^\infty \hbar\omega dj_{\gamma, \text{Sun}}(\hbar\omega)} \quad (7.20)$$

the linear increase with the intensity of the denominator is compensated by the linear increase in the short-circuit current in the numerator, and the efficiency is seen to increase with the intensity by the logarithmic increase in the voltage. Transport resistances have not been considered in this discussion. For large currents generated by high intensities, this is certainly problematic.

7.8

Efficiencies of the Individual Energy Conversion Processes

With a theoretical limit for the efficiency of $\eta = 0.3$ for the AM0 spectrum, energy conversion with a solar cell is still well away from the theoretical limit of $\eta_{\max} = 0.86$ for the solar heat engine of Section 2.5. It is very instructive once again

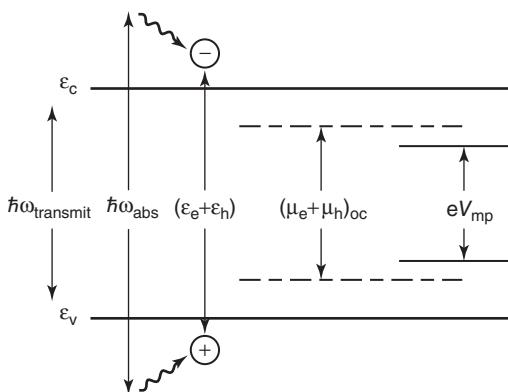


Figure 7.11 Individual processes in a solar cell.

to examine the processes in a solar cell individually and break down the overall efficiency into the efficiencies of the individual processes in order to recognize where the greatest losses occur. Figure 7.11 shows the individual processes schematically.

The first process is the absorption of the incident energy current. Its efficiency has to account for the photons with energy $\hbar\omega < \epsilon_G$, which are not absorbed by the solar cell. The absorbed energy current is given by the absorbed photon current times the mean energy $\langle \hbar\omega_{\text{abs}} \rangle$ of the absorbed photons. Let us assume that each absorbed photon generates just one electron–hole pair which, at short-circuit, contributes to the current. Then

$$j_{E,\text{abs}} = j_{\gamma,\text{abs}} \langle \hbar\omega_{\text{abs}} \rangle = -\frac{j_{sc}}{e} \langle \hbar\omega_{\text{abs}} \rangle \quad (7.21)$$

The absorption efficiency is

$$\eta_{\text{abs}} = \frac{j_{E,\text{abs}}}{j_{E,\text{inc}}} \quad (7.22)$$

The second process is the thermalization of the electron–hole pairs, each produced initially with the mean energy $\langle \hbar\omega_{\text{abs}} \rangle$, which by thermalization is reduced to $\langle \epsilon_e + \epsilon_h \rangle = \epsilon_G + 3kT$. The efficiency of this process is

$$\eta_{\text{thermalization}} = \frac{\langle \epsilon_e + \epsilon_h \rangle}{\langle \hbar\omega_{\text{abs}} \rangle} \quad (7.23)$$

The third factor defining the maximum chemical energy $(\mu_e + \mu_h)_{\text{oc}} = eV_{\text{oc}}$ that can be obtained from the energy $\langle \epsilon_e + \epsilon_h \rangle$ of the electron–hole pairs is the efficiency accounting for the difference between free energy and energy with a thermodynamic factor

$$\eta_{\text{thermodynamic}} = \frac{eV_{\text{oc}}}{\langle \epsilon_e + \epsilon_h \rangle} \quad (7.24)$$

This maximum chemical energy per electron–hole pair at open-circuit is, however, completely lost with the emitted photons. To gain energy, we have to go to

the maximum power point. This brings in the last factor, the fill factor FF , which determines how much of the maximum chemical energy current $-j_{sc} V_{oc}$ the solar cell delivers at the maximum power point as an electrical energy current $-j_{mp} V_{mp}$.

$$FF = \frac{j_{mp} V_{mp}}{j_{sc} V_{oc}} \quad (7.25)$$

The product of all these efficiencies gives the overall efficiency

$$\eta = \underbrace{\frac{j_{E, \text{abs}}}{j_{E, \text{in}}}}_{\eta_{\text{abs}}} \underbrace{\frac{\langle \epsilon_e + \epsilon_h \rangle}{\langle \hbar\omega_{\text{abs}} \rangle}}_{\eta_{\text{thermalization}}} \underbrace{\frac{eV_{oc}}{\langle \epsilon_e + \epsilon_h \rangle}}_{\eta_{\text{thermodynamic}}} \underbrace{\frac{j_{mp} V_{mp}}{j_{sc} V_{oc}}}_{FF} = \frac{-j_{mp} V_{mp}}{j_{E, \text{in}}} \quad (7.26)$$

For silicon, and in particular, for the 20 μm thick cell with light trapping, whose absorptance is shown in Figure 7.6, exposure to the AM1.5 spectrum gives the following values:

$$\begin{aligned} \langle \hbar\omega_{\text{abs}} \rangle &= 1.80 \text{ eV} \\ \langle \epsilon_e + \epsilon_h \rangle &= \epsilon_G + 3kT = 1.2 \text{ eV} \\ j_{sc} &= -413 \text{ Am}^{-2} \quad j_{mp} = -401 \text{ Am}^{-2} \\ V_{oc} &= 0.770 \text{ V} \quad V_{mp} = 0.702 \text{ V} \end{aligned}$$

The efficiencies are therefore

$$\begin{aligned} \eta_{\text{abs}} &= 0.74 \\ \eta_{\text{thermalization}} &= 0.67 \\ \eta_{\text{thermodynamic}} &= 0.64 \\ FF &= 0.89 \end{aligned}$$

The overall efficiency is then $\eta = 0.74 \times 0.67 \times 0.64 \times 0.89 = 0.28$.

The efficiencies for thermalization and for the conversion of the energy of the electron–hole pairs into chemical energy are particularly small and thus in need of improvement.

7.9 Problems

- 7.1** The current–voltage characteristic of a pn-junction shall be described by Equation 6.27 with $j_s = 10^{-12} \text{ mA cm}^{-2}$ and $j_{sc} = -35 \text{ mA cm}^{-2}$. Recalculate the characteristic adding
- (a) a series resistance normalized with respect to area $\tilde{R}_s = 10 \Omega\text{cm}^2$
 - (b) a normalized parallel resistance of $\tilde{R}_p = 50 \Omega\text{cm}^2$
 - (c) both.

8

Concepts for Improving the Efficiency of Solar Cells

We saw in the previous chapter that even avoiding all nonradiative recombination processes leaves us with a solar cell efficiency well below the theoretical maximum value of $\eta = 0.86$, derived in Chapter 2 as the upper limit for solar energy conversion. The main reasons were identified as losses by thermalization and the nonabsorption of low-energy photons. In order to improve the efficiency, we must focus primarily on reducing these losses. We now discuss different methods by which this can be accomplished, in principle. The underlying conditions are idealized, often to such an extent that it is difficult to imagine how they can be met, in practice. Nevertheless, it is still important to examine these methods in order to recognize the principles for possible improvements. A detailed discussion of these methods was presented by Green [22].

8.1

Tandem Cells

The reduction of thermalization losses and the improvement in the absorption efficiency can be simultaneously achieved by offering the solar cell only photons within the narrow energy interval $\varepsilon_G < \hbar\omega < \varepsilon_G + \Delta\varepsilon$ and processing the other photons by solar cells with a different band gap. Cells operated in this way are known as *tandem cells* or multi-junction cells.

For a black-body solar spectrum, a solar cell with the energy gap ε_G and the idealized absorptance $a(\hbar\omega < \varepsilon_G) = 0, a(\hbar\omega \geq \varepsilon_G) = 1$ has the short-circuit current

$$j_{sc} = -e \frac{\Omega_S}{4\pi^3 \hbar^3 c^2} \int_{\varepsilon_G}^{\infty} \frac{(\hbar\omega)^2}{\exp(\hbar\omega/kT_S) - 1} d\hbar\omega \quad (8.1)$$

Figure 8.1 gives the short-circuit current j_{sc} as a function of the energy gap ε_G/e . Following thermalization, the energy current flowing into the electron–hole pairs is

$$j_{E,eh} = \frac{-j_{sc} \varepsilon_G}{e} = j_{\gamma,abs} \varepsilon_G \quad (8.2)$$

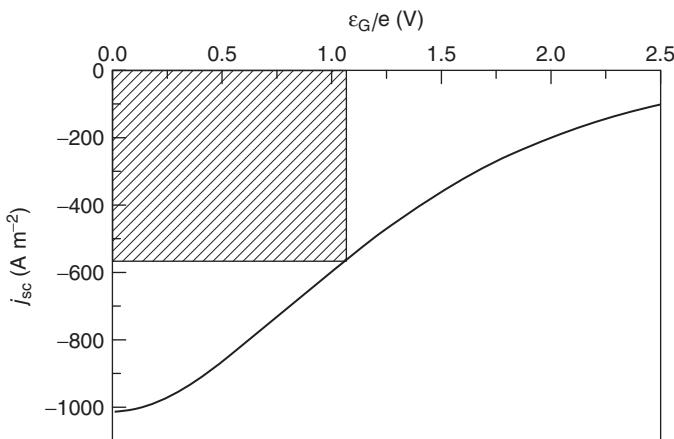


Figure 8.1 Short-circuit current of a solar cell as a function of the energy gap ε_G of its absorber for a black-body spectrum with $T_S = 5800$ K.

Using $\langle \varepsilon_e + \varepsilon_h \rangle = \varepsilon_G$ instead of $\varepsilon_G + 3kT$ here, we divide the overall efficiency η somewhat differently (and not quite correctly) into the thermalization and the thermodynamic efficiencies.

The shaded rectangle in Figure 8.1 indicates the energy current $j_{E,\text{eh}}$ transferred to the electron–hole pairs after thermalization, for the energy gap at which $j_{E,\text{eh}}$ has its maximum value.

The area under the curve $j_{sc}(\varepsilon_G/e)$ is the entire incident energy current $j_{E,\text{Sun}}$ coming from the Sun. This can best be seen if we determine the variation of dj_{sc} of the short-circuit current for a small variation $d\varepsilon_G$ in the band gap of the absorber

$$dj_{sc} = -e \frac{\Omega_S}{4\pi^3 \hbar^3 c^2} \frac{\varepsilon_G^2}{\exp(\varepsilon_G/kT_S) - 1} d\varepsilon_G \quad (8.3)$$

and then integrate over the variation in the absorbed energy current $-(\varepsilon_G/e) dj_{sc}$, caused by the variation in the band gap $d\varepsilon_G$; thus,

$$j_E = - \int_0^\infty \frac{\varepsilon_G}{e} dj_{sc} = \frac{\Omega_S}{4\pi^3 \hbar^3 c^2} \int_0^\infty \frac{\varepsilon_G^3}{\exp(\varepsilon_G/kT_S) - 1} d\varepsilon_G \quad (8.4)$$

Since the value of a definite integral does not depend on the names of the variables (we could also call it $\hbar\omega$ instead of ε_G), Equation 8.4 is recognized to describe the density of the incident solar energy current. The largest rectangle, shaded in Figure 8.1, showing the energy current transferred to the electron–hole pairs after thermalization, corresponds to 42% of the incident energy current. The area under the curve to the right of the rectangle gives the energy current, which is lost by thermalization. The area below the rectangle down to the $j_{sc}(\varepsilon_G/e)$ curve is the energy current, which is not utilized, because it is not absorbed by the solar cell. A solar cell with $\varepsilon_G = 1.1$ eV would thus have an efficiency of 42%, if all of the energy of the electron–hole pairs could be converted into electrical energy, i.e. if the thermodynamic efficiency were equal to 1. This value, called the *ultimate*

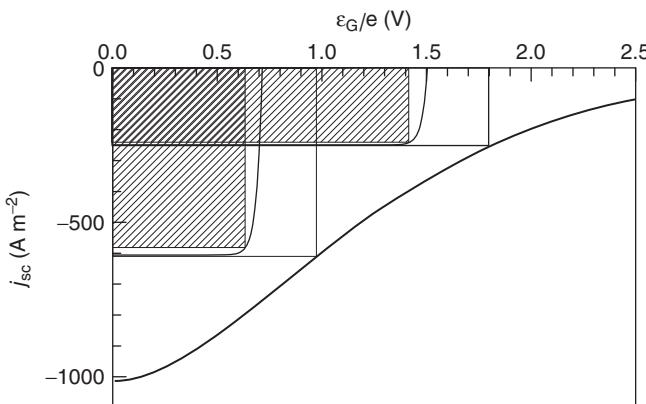


Figure 8.2 Current–voltage characteristics of two solar cells with energy gaps $\varepsilon_{G1} = 1.8$ eV and $\varepsilon_{G2} = 0.98$ eV.

efficiency by Shockley and Queisser [11] can, however, not be achieved at room temperature, because entropy must not be annihilated and some of the energy of the electron–hole pairs is room temperature heat.

Figure 8.2 shows the same curve as Figure 8.1, this time illustrating how two solar cells with different energy gaps ε_{G1} and ε_{G2} divide the incident energy current of the AM0 spectrum. The energy current first falls on the cell with the greater energy gap ε_{G1} , which absorbs all photons with $\hbar\omega \geq \varepsilon_{G1}$ and transmits all photons with $\hbar\omega < \varepsilon_{G1}$. The cell behind, with the lower energy gap, then absorbs the photons with $\varepsilon_{G2} \leq \hbar\omega < \varepsilon_{G1}$.

For the two cells depicted in Figure 8.2 the current–voltage characteristics are also shown together with the cross-hatched rectangles indicating the maximum electrical energy current that the cells deliver. Here again, we assume that only radiative recombination is present.

Figure 8.3 gives the efficiency obtained from two cells having the energy gaps ε_{G1} and ε_{G2} when their energy currents are added. For the AM0 spectrum, the optimal combination consists of $\varepsilon_{G2} = 1.0$ eV and $\varepsilon_{G1} = 1.9$ eV, yielding an overall efficiency for obtaining an electrical energy current of $\eta = 0.44$.

If we extend the division of the incident spectrum onto an infinite number of solar cells optically in series, with continuously decreasing energy gaps, the entire energy current from the Sun will then be transferred to energy of the electron–hole pairs, and no thermalization losses will occur at all. With this configuration, each cell absorbs those photons from the incident photon current that have energies in the interval $\varepsilon_G \leq \hbar\omega < \varepsilon_G + d\varepsilon_G$, assuming that the energy gaps of adjoining cells differ by $d\varepsilon_G$.

Under open-circuit conditions, with radiative recombination only, exactly the same number of photons is emitted as is absorbed, which is always true for radiative recombination. This time, however, the emitted spectrum is identical with the absorbed spectrum, since there are no thermalization losses. All emitted photons have the temperature of the cell, but they have nonzero chemical potentials that

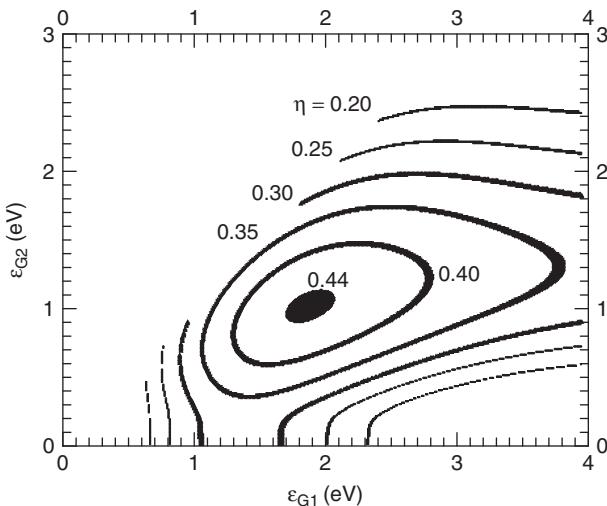


Figure 8.3 Efficiency for two solar cells in tandem operation, with energy gaps ε_{G1} and ε_{G2} for the AM0 spectrum when their energy currents are added.

depend on their energy. If we also assume maximum concentration of the incident solar radiation, in accordance with the generalized Planck radiation law (3.102), the chemical potential μ_γ of the emitted photons for one cell that absorbs and emits photons of energy $\hbar\omega = \varepsilon_G$ is then

$$\mu_\gamma = \hbar\omega \left(1 - \frac{T_0}{T_S}\right) = \varepsilon_G \left(1 - \frac{T_0}{T_S}\right) \quad (8.5)$$

In our treatment of maximum efficiencies in Chapter 2 we saw that the Landsberg efficiency requires an absorber capable of absorbing solar radiation without creation of entropy, although its temperature is that of the environment and not that of the Sun. The infinite tandem is one such absorber. However, in contrast to the infinite tandem, in which the production of entropy is prevented by the production of chemical energy, the Landsberg efficiency also requires that the absorber emits photons with $\mu_\gamma = 0$. The infinite tandem can only prevent the production of entropy in the open-circuit state, in which the absorbed energy current is completely reemitted back to the Sun, while the Landsberg efficiency process assumes that no entropy is created at maximum power.

At maximum power of the tandem stack of solar cells, at which each cell of the infinite tandem operates at a different voltage V_{mp} and emits photons with a different chemical potential $\mu_\gamma = eV_{mp}$, but always at the same temperature T_0 , the emission spectrum is identical with the emission spectrum of an infinite tandem of monochromatic thermal absorbers feeding energy into Carnot engines, as was discussed in Section 2.5. The upper limit for the efficiency that can be obtained with an infinite tandem is $\eta = 0.86$ and is identical with that for photothermal conversion discussed in Section 2.5.

8.1.1

The Electrical Interconnection of Tandem Cells

For the optimal absorption of the incident photons, the cells are arranged one after the other, optically in series. Electrical contacts between the cells, which absorb photons, must be avoided. This allows only for connection of the cells electrically in series. The voltages of the different cells must all have the same sign. This requires that the same type of membrane of all the cells, e.g. the p-type membrane, must be facing the Sun. It follows that the n-type membrane of the preceding cell is facing the p-type membrane of the following cell. The problem now is that holes flow out of the p-membrane and electrons out of the n-membrane. The transfer of the charge at the interface between these membranes requires that the holes recombine with the electrons. The difference between the Fermi energies, normally required for the recombination, however, would give rise to a voltage opposite in sign to the voltages of the solar cells. Figure 8.4 shows two cells that are series connected in such a way that electrons flowing out of the n-membrane of the left cell and holes flowing out of the p-membrane of the right cell recombine in a tunnel junction. In the tunnel junction, electrons and holes belong to the same Fermi distribution ($\epsilon_{F2} = \epsilon_{F3}$) and recombination proceeds without a difference in the Fermi energies. As is seen in the figure, the Fermi energy must lie in the conduction band of the n-membrane on one side and in the valence band of the p-membrane on the other side, requiring high doping concentrations ($n_D, n_A > N_C, N_V$). Since the tunnel junction is very thin, the absorption by its high free-carrier concentrations is very small.

Thermalization was found to be necessary for the conversion of solar heat into chemical energy in Chapter 2. Thermalization losses are reduced or even prevented when the generated electrons and holes populate only narrow energy

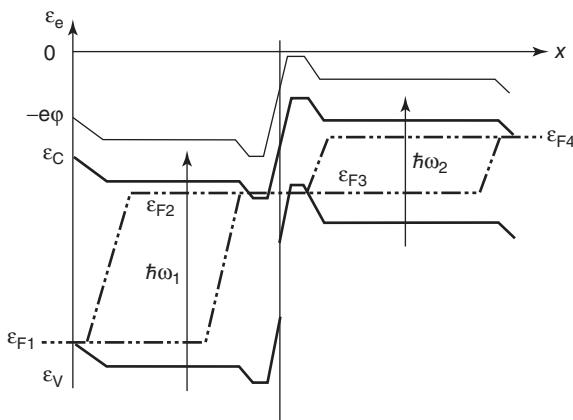


Figure 8.4 Electrical series connection of two solar cells by a tunnel junction provides efficient recombination of electrons and holes without requiring a difference of their Fermi energies.

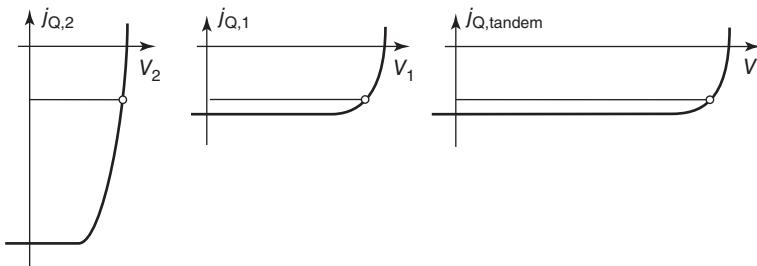


Figure 8.5 Current–voltage characteristic for the series connection of two solar cells with different short-circuit currents and open-circuit voltages. For each value of the current j_Q the voltages V_1 and V_2 are added to give the total voltage V .

ranges at the band edges, as explained above. This condition can be expressed in a different way. Prevention of thermalization losses requires many different Fermi distributions (one for each energy range) with many different Fermi energies. The tandem in Figure 8.4 has three different Fermi energies, one more than for a single-material solar cell.

Series connection forces the same charge current to flow through all the cells. Figure 8.5 shows how the current–voltage characteristic of a tandem results from the characteristics of two cells with different short-circuit currents and different open-circuit voltages. For a given current, the voltages V_1 and V_2 are determined from the characteristics of the individual cells and added to give the voltage V of the tandem configuration. The cell with the smaller short-circuit current determines the total current.

To prevent losses due to a series connection, as in Figure 8.5, the energy gaps must be chosen so that the currents j_{mp} at the maximum power points are the same for all cells. The solar radiation spectrum, however, changes over the day and the year due to different path lengths through the atmosphere, and the equality of the currents j_{mp} of different cells cannot always be maintained. For a tandem of three cells, a configuration where the series connection of the two cells with the smaller band gaps is connected in parallel to the third cell with the largest band gap is found to be less sensitive to spectral variations than a series connection of all three cells [23].

8.2 Concentrator Cells

In Chapter 2, we considered how strongly we can focus the incident solar radiation. For concentrated radiation, the same power is incident onto a solar cell with smaller area than for nonconcentrated radiation. Another advantage is that concentrated radiation can be processed with greater efficiency, as discussed in Chapter 7.

In areas with much direct, unscattered solar radiation, the additional expense for concentration is rewarded by a better efficiency from a smaller solar cell. When the radiation is concentrated with lenses or mirrors, the solar cell sees only a part of the hemisphere and, in the limiting case of maximum concentration, only the Sun. The greater the concentration, the more carefully the concentrating system must track the path of the Sun. The concentration of radiation also has disadvantages. The improvement in the efficiency assumes that, in spite of an increase in the incident radiation, the temperature of the cell remains the same. But in fact the temperature of the cell rises, and with poor cooling the efficiency can even decrease with increasing concentration. Another disadvantage is the result of the larger electrical currents causing larger voltage losses across the series resistance of the cell and in the leads.

Special solar cells known as *concentrator cells* have been developed for concentrated radiation. Because of the higher temperatures, semiconductors with a larger energy gap are always advantageous. However, they must be able to absorb a sufficiently large part of the solar spectrum. Cells made of GaAs are well suited for this. Because of their smaller surface areas, more expensive materials and more expensive constructions become economical. For concentrator systems, tandem cells based on III–V compounds have become cost effective. For silicon concentrator cells, a more complicated design was developed, where the n- and p-type membranes for the electrons and the holes are in the form of point contacts, all placed on the back side of the cell as shown in Figure 8.6.

The electrons and holes flow out of the n- and p-doped regions, placed alternately on the rear side. This is different from the structures discussed up to now and has the advantage that the contacts do not cause shadowing and can be kept relatively large to avoid series resistances. To reduce Auger recombination to a minimum, the greatest part of the solar cell is not doped. The highly doped regions required as membranes and to conserve the entropy per particle are limited to the contacts. For concentration by a factor of 100 an efficiency of 28% was obtained with a silicon cell. This was, however, for a temperature of 25 °C, which requires intensive cooling [24]. For a tandem, a record efficiency of 46% for a configuration of 4 cells operating at a concentration of the solar radiation by a factor of 508 was reported [25].

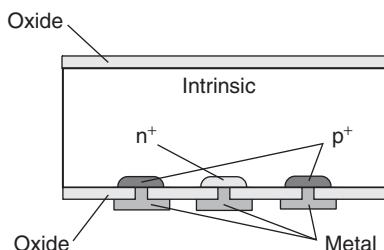


Figure 8.6 The p- and n-type membranes for holes and electrons and the metal contacts of the point-contact cell for concentrated radiation are both placed on the back side of the cell.

8.3

Thermophotovoltaic Energy Conversion

The solar–thermal conversion method of Section 2.5 can be modified to be applicable to solar cells. Figure 8.7 illustrates the principle.

A focusing optical system is used to concentrate the solar radiation onto an intermediate absorber that, as a result, is heated to the temperature T_A . Solar cells with an energy gap ε_G are placed concentrically around the intermediate absorber. They have an interference filter on their surface, which transmits all photons with $\varepsilon_G \leq \hbar\omega \leq \varepsilon_G + \Delta\varepsilon$ without loss and reflects all other photons, which cannot be used optimally, back to the intermediate absorber. These photons together with those emitted by the solar cell and transmitted by the filter help maintain the temperature T_A of the intermediate absorber. If the recombination is entirely radiative and the photons emitted by the cells are reabsorbed by the absorber and are therefore not lost, the cells can be operated close to their open-circuit voltage and thus have the efficiency derived in Equation 4.5 for the conversion of the absorbed photon energy, i.e.

$$\eta_{\text{cell}} = 1 - \frac{T_0}{T_A} \quad (8.6)$$

Since this is the Carnot efficiency, which was found for the efficiency of the heat engine of the solar–thermal conversion process described in Section 2.5, we find for the thermophotovoltaic conversion process the same efficiency as for the solar–thermal conversion; thus,

$$\eta = \left[1 - \left(\frac{T_A}{T_S} \right)^4 \right] \left(1 - \frac{T_0}{T_A} \right) \quad (8.7)$$

with a maximum value of $\eta = 0.85$ at an absorber temperature of $T_A = 2478$ K.

The implementation of this concept in practice is difficult for two reasons. At the optimal temperature of the intermediate absorber of $T_A = 2478$ K, all materials evaporate so strongly that the interference filter is quickly covered with an opaque layer. Moreover, in practice it is not possible to construct an interference filter transmitting only in a narrow energy interval and reflecting the rest

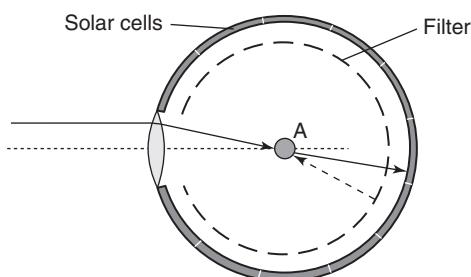


Figure 8.7 In the thermophotovoltaic converter, the intermediate absorber is surrounded in an evacuated cavity by solar cells illuminated by its thermal radiation.

of the spectrum while also being free of absorption. With the use of, e.g. silicon solar cells, only a very small portion of the photons emitted from the intermediate absorber have the required energy $\hbar\omega \geq \varepsilon_G$. Even very little absorption by the interference filter of all the other photons leads to a considerable loss. Smaller band gap materials like GaSb are more favorable for thermophotovoltaic conversion.

In the construction principle of Figure 8.7, another drawback is hidden. The absorber must be able to emit almost as much energy as it absorbs in the narrow energy interval transmitted by the filter. Since T_A is much smaller than the temperature of the Sun T_S , the emitting area must be much larger than the absorbing area. A factor of 4, as provided by the arrangement in Figure 8.7 is far from sufficient.

This problem can be solved by an even wilder idea called *thermo-photonics* [26]. Emission of only photons with $\hbar\omega \geq \varepsilon_G$ can be accomplished by placing a semiconductor on the emitting area of the intermediate absorber. (To prevent transmission by the semiconductor of smaller-energy photons emitted by the absorber, a mirror must be placed between the semiconductor and the absorber.) When the semiconductor is supplied with membranes such as a solar cell, it can be operated as a light-emitting diode (LED). A LED is the same engine as a solar cell, only operated in reverse, just as a refrigerator or heat pump is a reversely operated heat engine. If some of the power delivered by the solar cells is used to drive the LED on the absorber, the emitted intensity is enhanced enormously. Although some of the energy emitted by the LED is free energy supplied by the solar cells, most of it is heat supplied by the absorber. As a result, the area and the temperature of the emitter can be reduced, provided an LED can be made that works at about 1000 °C with close to 100% external quantum efficiency.

With the arrangement shown in Figure 8.7, however, the intermediate absorber does not have to be heated by the Sun. It is also possible to heat the intermediate absorber in another way, e.g. by burning gas. Radiation losses to the environment would then not have to occur, because the cavity can be completely closed, optically. The conversion of heat into electrical energy in this manner was first proposed in the Soviet Union for nuclear reactors. The basic concept was to surround incandescent reactor fuel elements with solar cells. Fortunately, no one had the courage to try this out.

8.4

Impact Ionization

Electrons and holes possessing large kinetic energies as a result of generation by high-energy photons can dissipate their kinetic energy in two ways. One is by elastic collisions with the lattice atoms, in which energy is transferred in small portions to the lattice atoms until thermal equilibrium with the lattice is established. The other is by inelastic collisions with the lattice atoms in which, by impact ionization, another electron is knocked off its chemical bond or, in other words,

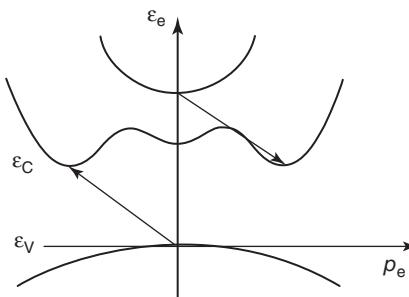


Figure 8.8 Transition of an electron from a higher band to the minimum of the conduction band by impact ionization in an indirect semiconductor, resulting in the additional generation of an electron and a hole at the band edges.

in which a free electron and a free hole are produced, as shown in Figure 8.8. Both processes take place in parallel and compete with each other. With elastic collisions, the excitation of lattice vibrations is at the expense of the energy of the electron–hole system, while the number of the electrons and holes remains constant. With impact ionization, the absorbed energy remains in the electron–hole system, but is more uniformly distributed over a larger number of electrons and holes than were originally generated by the absorption of the photons. Impact ionization, therefore, looks very promising for solar energy conversion because some of the energy removed from the electrons and holes during thermalization is used to generate additional electron–hole pairs [27].

To examine the efficiency of the impact ionization process, we will exclude the competing process, namely, the interaction of electrons and holes with the lattice vibrations, which leads to thermalization at constant concentrations. The electrons and holes are then isolated from the lattice vibrations. They do not “know” about the temperature of the lattice and cannot come into thermal equilibrium with the lattice. Collisions between electrons and holes are, however, permitted. This ensures that electrons and holes have a uniform temperature T_A , although this is not the same as the lattice temperature. Finally, according to the principle of detailed balance, we must expressly consider Auger recombination as the inverse process of impact ionization [28].

We will now examine the temperature and electrochemical potentials of the electrons and holes under these conditions. The simplest answer, unfortunately, may be the most difficult to understand. It is based on the difference between thermalization and impact ionization: whereas during thermalization by scattering with phonons, no electrons (or holes) are annihilated or created, so that their number remains constant, during impact ionization and its inverse process, Auger recombination, the number of electrons and holes changes. This has significant consequences for the values of the electrochemical potentials of the electrons and holes. The change of the particle numbers by impact ionization and Auger recombination is unrestricted, except for the condition to establish a minimum of the free energy of the electrons and the holes. Thus,

$$dF = \dots + \eta_e dN_e + \eta_h dN_h + \dots = 0$$

With impact ionization and Auger recombination, electrons and holes are always created or annihilated in pairs, that is, $dN_e = dN_h = dN$ and

$$dF = \dots + (\eta_e + \eta_h) dN + \dots = 0$$

Because the number of particles does not remain constant and is not tied to other particle numbers as in a chemical reaction, $dN \neq 0$, so that

$$\eta_e + \eta_h = 0$$

We will attempt to make this result more plausible. Let us assume that the free energy of the electrons and holes describes a state with $\eta_e + \eta_h > 0$. With a reduction in the number of particles due to Auger recombination, i.e. with $dN < 0$ and therefore $dF < 0$, the free energy can be further reduced. With the reduction in the number of particles, $\eta_e + \eta_h$ also decreases, until for $\eta_e + \eta_h = 0$ a further reduction of the particle number no longer reduces the free energy and equilibrium is established between impact ionization and Auger recombination. Since the total energy of the electron–hole system is preserved, by reducing the particle number, Auger recombination leads to an increase in the energy per particle, i.e. in the mean kinetic energy of the electrons and holes. This in turn leads to an increase in their temperature and in the number of electrons and holes capable of participating in impact ionization, until the rate of impact ionization is exactly the same as the rate of Auger recombination and $\eta_e + \eta_h = 0$.

If, however, scattering with the lattice atoms and Auger recombination and impact ionization all occur at high rates, the temperature of the electrons and holes will be the same as the lattice temperature, and in equilibrium with impact ionization and Auger recombination, the only possible state is one with $T_{eh} = T_0$ and $\eta_e + \eta_h = 0$, which does not permit the conversion of energy. It is therefore very important that thermalization and impact ionization, together with its inverse process, do not occur simultaneously and with similar probabilities. It would not improve the efficiency of solar cells, but in fact reduce it, if the probability for impact ionization (and Auger recombination, inevitably) were slightly increased while the interaction with phonons predominates.

We thus establish the fact that the interaction with the lattice vibrations maintains the temperature of the electron–hole system constant at the lattice temperature T_0 and, on exposure to light, produces a state with $\eta_e + \eta_h > 0$. Impact ionization and Auger recombination, on the other hand, maintain a state with no separation of the Fermi energies, $\eta_e + \eta_h = 0$, but in the absence of interaction with the lattice vibrations and on exposure to light produce a state with $T > T_0$.

The problem of how to obtain electrical energy from hot electrons and holes then remains to be solved. Energy conversion by means of impact ionization first produces hot electrons and holes with no chemical energy. Chemical energy, and finally electrical energy, must be obtained in subsequent steps.

8.4.1

Hot Electrons from Impact Ionization

The temperature T_A of the electrons and holes in the absorber can be found very easily from the magnitude of the emitted energy current. In contrast to the interaction with phonons, where at open circuit the emitted *photon* current is equal to the absorbed photon current, for impact ionization at open circuit the emitted *energy* current must be equal to the absorbed energy current, since by impact ionization/Auger recombination the electron–hole system does not lose energy. The chemical potential of the emitted photons is then $\mu_\gamma = \eta_e + \eta_h = 0$, and we can calculate the temperature T_A of the electrons and holes by using Planck's law in Equation 2.34 for the emitted energy current. For maximum concentration of the incident solar radiation, we find, of course, $T_A = T_S$ at open circuit.

8.4.2

Energy Conversion with Hot Electrons and Holes

For conventional solar cells based on thermalization of electrons and holes in the absorber, complete conversion of chemical energy into electrical energy was achieved by membranes, n-type for the transport of electrons to the electron contact and p-type for the transport of holes to the hole contact. This type of membrane is not sufficient for hot carriers. In addition to the selective transport of electrons and holes, the membranes must now also serve the thermodynamic function of producing chemical energy from the heat of the electrons and holes by cooling them down to the temperature of the environment.

We will discuss this problem for the electrons; the solution can then easily be applied to the holes as well. If we would allow an exchange of electrons between absorber and membrane for all electron energies in the absorber, the thermalization of the electrons in the membrane would lead to a large energy loss, from $(3/2)kT_A$ in the absorber to $(3/2)kT_0$ in the membrane (less, however, than in a conventional solar cell). Secondly, with the unimpeded exchange of electrons between the absorber and the membrane, the electrons in the absorber would be cooled as well and would no longer be capable of impact ionization. This would result in a state with $\eta_e + \eta_h = 0$ and $T_A = T_0$. However, as we have seen earlier, the entire energy loss caused by thermalization can be prevented if the electrons in the membrane only occupy states over a narrow range $\Delta\epsilon_e$ at the energy ϵ_e , as shown in Figure 8.9.

For $\Delta\epsilon_e < kT_0$, the occupation of the electron states in the membrane cannot significantly change by interaction with the phonons. As a result, the entropy of the electrons also remains unchanged and thermalization takes place isentropically. Since the number of particles remains constant during thermalization in the membrane, the electrochemical potential of the electrons increases. Figure 8.9 demonstrates that the same process takes place with the holes at an energy ϵ_h in the hole membrane. The isentropic cooling therefore produces the chemical and electrochemical energy per electron–hole pair,

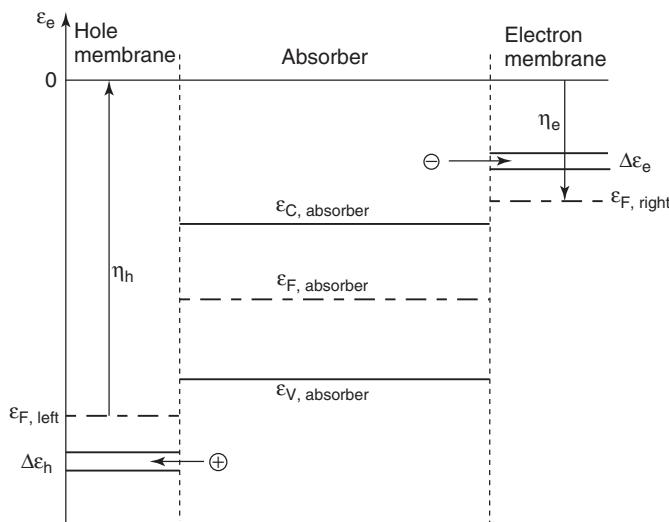


Figure 8.9 Energies of electrons and holes in the absorber, in which impact ionization and Auger recombination are in equilibrium at $T_A > T_0$, and in membranes through which

electrons and holes flow outward and where they are in temperature equilibrium with the environment at $T = T_0$.

$$\mu_e + \mu_h = \eta_e + \eta_h = (\epsilon_e + \epsilon_h) \left(1 - \frac{T_0}{T_A} \right) \quad (8.8)$$

The arrangement of Figure 8.9 is a working solar cell. The voltage is

$$V = \frac{(\eta_e + \eta_h)}{e} \quad (8.9)$$

and the current is

$$j_Q = e \frac{(j_{E, \text{absorbed}} - j_{E, \text{emitted}})}{(\epsilon_e + \epsilon_h)} \quad (8.10)$$

We can visualize its operation by increasing the current from zero, the open-circuit situation, where the emitted energy current equals the absorbed energy current. With increasing charge current, the energy in the electron–hole system decreases and with it the electron–hole temperature T_A in the absorber. Owing to the lower temperature, the emitted energy current is reduced and so is the voltage. The current rises until, at zero voltage, the short-circuit situation is obtained. A still larger current may be withdrawn by applying a negative voltage, spending energy from a battery, which would cool the remaining electrons and holes down to $T_A < T_0$.

It is interesting that the open-circuit voltage is determined by the energies at which the electrons and holes are removed and not by the absorber material. If the rates of removal of electrons and holes are small compared with the rate of impact ionization/Auger recombination and that of carrier–carrier scattering, which we assume to be the case, their equilibrium will hardly be affected by the removal

process. Electrons and holes removed at the energies ε_e and ε_h will quickly be replenished by impact ionization/Auger recombination and carrier–carrier scattering. Although the charge current and the voltage depend on the energies with which electrons and holes are removed, it is very surprising that the energy current delivered by the cell, obtained by multiplying Equations 8.9 and 8.10, is independent of the removal energies. Large removal energies give a large voltage and a small current and small removal energies give a large current and a small voltage, both resulting in the same energy current.

The efficiency is maximum at maximum concentration of the solar radiation, when the temperature T_A of the electrons and holes is equal to the temperature T_S of the Sun at open circuit. Since in the absence of interactions with the lattice vibrations the absorbed energy remains in the electron–hole system, it is advantageous to absorb as much as possible by reducing the band gap ε_G of the absorber material to zero. The electron–hole system is then a black body and, according to Equation 2.25, absorbs the energy current σT_S^4 and emits the energy current σT_A^4 at the temperature T_A . With Equations 8.9 and 8.10, the efficiency with which electrical energy is delivered is

$$\eta = \frac{j_Q V}{\sigma T_S^4} = \frac{\sigma(T_S^4 - T_A^4)}{\sigma T_S^4} \left(1 - \frac{T_0}{T_A}\right) = \left(1 - \frac{T_A^4}{T_S^4}\right) \left(1 - \frac{T_0}{T_A}\right) \quad (8.11)$$

and is therefore identical with the efficiency of the ideal solar heat engine in Equation 2.53 in Section 2.5 and that of the thermophotovoltaic conversion process discussed in the previous section. The efficiency has its maximum value of $\eta_{\max} = 0.85$ at a temperature of the electron–hole system of $T_A = 2478$ K if we assume a temperature of $T_S = 5800$ K for the Sun.

Figure 8.10 shows that the efficiency at full concentration falls off with increasing energy gap, because of the decreasing absorption. Without concentration, for $\Omega = \Omega_S$, that is, for the AM0 spectrum, however, a nonzero energy gap is

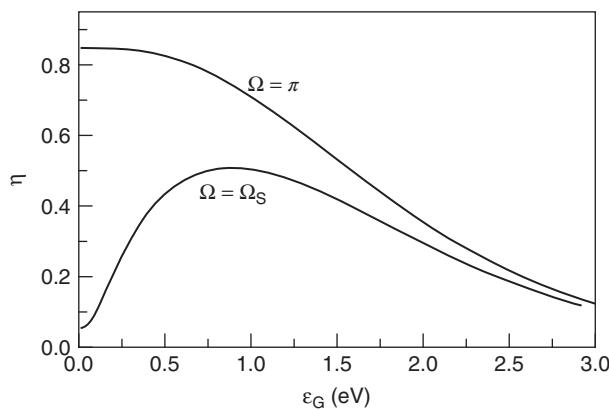


Figure 8.10 Efficiency for a hot-carrier cell with impact ionization for nonconcentrated incident solar radiation with $\Omega = \Omega_S$ and for maximum concentration with $\Omega = \pi$.

preferable, otherwise more photons would be emitted than absorbed at small photon energies. For $\epsilon_G > 0$ the balance becomes more favorable.

An earlier proposal for a hot-carrier solar cell by Ross and Nozik [29] that did not account for impact ionization and Auger recombination, finds even higher efficiencies for narrow-gap semiconductors under less than full concentration of the solar radiation. Similar to a conventional solar cell, the process of carrier–carrier scattering, while leading to a uniform temperature, is assumed to leave the number of electron–hole pairs unchanged, increasing by one for each absorbed photon and decreasing by one for each emitted photon. This assumption leads to high temperatures and negative chemical potentials of the electron–hole pairs for less than full concentration. Since electrons and holes are withdrawn through mono-energy contacts, scattering of the carriers with each other is necessary to replenish the carriers in the energy range from where they are withdrawn. A problem is that a distinction between scattering events that keep the carrier concentrations constant and impact ionization and Auger recombination which do not is physically impossible in narrow-gap semiconductors. The problem becomes obvious for the case where electron–hole pairs are withdrawn with an energy that is smaller than the average energy of the absorbed photons. The more electron–hole pairs are withdrawn, the more energy per pair piles up for the remaining electron–hole pairs. As a result, their temperature increases beyond any reasonable limit, far beyond the Sun's temperature. No such problems are encountered when impact ionization and Auger recombination are taken into account.

We thus see that impact ionization and Auger recombination allow ideal energy conversion, provided that interaction with the lattice vibrations is excluded. However, no material in bulk form is known in which these conditions are even approximately fulfilled. In a (very) thin film, however, one can imagine that electrons and holes can be removed in much less than 10^{-12} s, long before they are thermalized.

8.5

Two-step Excitation in Three-level Systems

If interaction with phonons cannot be prevented, thermalization losses can be reduced by dividing the incident spectrum over more than one transition as we have seen with tandem cells. In a three-level system, where the levels can be bands as well, three different transitions may occur in a single material: directly from the lower level to the upper level and in addition by a two-step process from the lower level to the intermediate level and from there to the upper level. In both ways, electrons are generated at the upper level and holes at the lower level.

8.5.1

Impurity Photovoltaic Effect

In Section 3.6.2, we have discussed nonradiative transitions between the bands and an impurity level. Impurities with energies for electrons in the middle of the

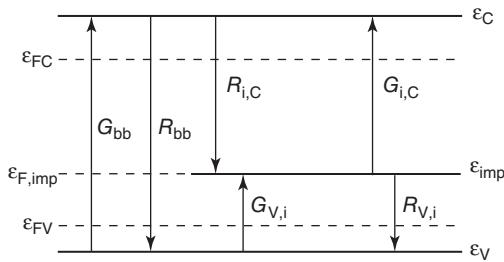


Figure 8.11 In addition to radiative band–band transitions with the rates G_{bb} and R_{bb} , radiative transitions between the bands and the impurity are taken into account. Nonradiative transitions are excluded.

energy gap were found to greatly enhance recombination, which is detrimental for the efficiency. In the analysis, generation of electron–hole pairs by optical transitions was neglected. Now we do just the opposite. Our model now permits only radiative transitions between the bands and to and from an impurity level. Thermalization of free charge carriers is considered, but not impact ionization or nonradiative recombination [30].

The model, as shown in Figure 8.11, has states in the valence band with $\epsilon_e \leq \epsilon_V$, at the impurity level with ϵ_{imp} , and in the conduction band with $\epsilon_e \geq \epsilon_C$. To optimally use the incident spectrum, the photons will be distributed over the different transitions in such a way that photons capable of a higher energy transition, e.g. band–band, are not wasted in lower energy transitions. For the impurity energy $\epsilon_{imp} - \epsilon_V \leq \hbar\omega < \epsilon_C - \epsilon_{imp}$ are exclusively absorbed in transitions from the valence band to the impurity. Photons having $\epsilon_C - \epsilon_{imp} \leq \hbar\omega < \epsilon_C - \epsilon_V$ are exclusively absorbed in transitions from the impurity to the conduction band, and photons having $\hbar\omega \geq \epsilon_C - \epsilon_V$ provide for the band–band transitions.

The absorption properties of the impurities are characterized by optical cross sections, $\sigma_{V,i}$ for transitions from the valence band to the impurity and $\sigma_{i,C}$ for transitions from the impurity to the conduction band. Optical cross sections are of the same order of magnitude as geometrical cross sections, 10^{-15} cm^2 being a typical value. Although optical cross sections vary with energy, we assume them to be constant over the energy range of absorbable photons.

The electrons and holes are assumed to have high mobilities resulting in their homogeneous distribution, even though they are generated inhomogeneously. The steady-state concentrations belonging to a given value of the charge current follow from the continuity equations for the particle densities, in which in addition to generation and recombination we consider the contribution to the charge current by the divergence of the electron and hole currents. From the concentrations of electrons and holes, the sum of their electrochemical potentials and thus the voltage can be derived. The continuity equations are

$$\frac{\partial n_e}{\partial t} = G_{bb} + G_{i,C} - R_{bb} - R_{i,C} - \operatorname{div} j_e = 0 \quad (8.12)$$

$$\frac{\partial n_h}{\partial t} = G_{bb} + G_{V,i} - R_{bb} - R_{V,i} - \text{div } j_h = 0 \quad (8.13)$$

$$\frac{\partial n_{e,imp}}{\partial t} = R_{i,C} - G_{i,C} + G_{V,i} - R_{V,i} = 0. \quad (8.14)$$

In the last equation, the divergence of the particle current is missing since electrons in the impurities are considered immobile, so they do not contribute to the current.

Since these three equations are not independent, we need the charge neutrality as an additional equation, as in Section 3.6.2. Because of the high absorption required, however, the impurity concentration n_{imp} is now no longer negligible compared with the densities of the electrons and holes. For this reason, most impurities must be electrically neutral, that is, either occupied, if they are donor-like, or unoccupied, if they are acceptor like. Both situations are unfavorable for the desired impurity absorption, since for transitions from the valence band to the impurity they must be unoccupied and for transitions from the impurity to the conduction band they must be occupied. The smaller of the two transition rates will determine the rate at which electrons are excited to the conduction band and holes to the valence band by two-step impurity transitions. For optimal absorption properties, we therefore choose half of the impurities to be donorlike and half acceptor like, and

$$\rho_Q = e \left(n_h - n_e + \frac{n_{imp}}{2} - n_{e,imp} \right) = 0 \quad (8.15)$$

Alternatively, if the impurities are donorlike, an occupation probability of 1/2 is achieved by doping the material additionally with half as many shallow acceptors as there are impurities. Either way, the Fermi energy in the dark has to coincide with the impurity level.

The generation rates, averaged over the thickness d , are given by

$$\begin{aligned} G_{bb} &= \frac{1}{d} \int_{e_C - e_V}^{\infty} a_{bb} \, dj_\gamma(\hbar\omega) \\ G_{i,C} &= \frac{1}{d} \int_{e_C - e_{imp}}^{e_C - e_V} a_{i,C} \, dj_\gamma(\hbar\omega) \\ G_{V,i} &= \frac{1}{d} \int_{e_{imp} - e_V}^{e_C - e_{imp}} a_{V,i} \, dj_\gamma(\hbar\omega) \end{aligned} \quad (8.16)$$

The absorptance for the band–band transitions is assumed to be $a_{bb} = 1$, whereas the absorptance for the impurity transitions depends on the concentration and occupation of the impurities,

$$a_{i,C} = 1 - \exp(-\alpha_{i,C} d); \quad a_{V,i} = 1 - \exp(-\alpha_{V,i} d)$$

where the absorption coefficients follow from Equation 3.97

$$\alpha_{i,C} = \sigma_{i,C} n_{imp} (f_i - f_C) \quad \text{and} \quad \alpha_{V,i} = \sigma_{V,i} n_{imp} (f_V - f_i)$$

These equations determine the generation rates not only due to illumination but also in the dark state with its incident 300 K background radiation. According to the principle of detailed balance, in this state of chemical equilibrium with the background radiation, the recombination rates must have the same value as the generation rates. The principle of detailed balance is accounted for if we write the recombination rates in terms of the generalized Planck radiation law. The rate of downward transitions per energy from a level j to a level i averaged over the thickness d is, from Equation 3.102,

$$\frac{dR_{ij}}{d\hbar\omega} = \frac{1}{d} a_{ij} \frac{1}{4\pi^3 \hbar^3 c_0^2} \frac{(\hbar\omega)^2}{\exp\{[\hbar\omega - (\varepsilon_{F,j} - \varepsilon_{F,i})]/kT\} - 1} \quad (8.17)$$

This rate is then integrated over the energy range associated with the band to band, valence band to impurity, and impurity to conduction band transitions to give the actual recombination rates.

The charge current j_Q finally results from the integral of the divergence of the electron current (or the hole current) over the thickness d of the cell and, because of the assumed homogeneous distribution of the electrons and holes, it is given by

$$j_Q = -e (\operatorname{div} j_e) d = -e (\operatorname{div} j_h) d \quad (8.18)$$

For a given charge current, the continuity equations 8.12–8.14 are solved for the positions of the Fermi energies subject to charge neutrality. This yields the current–voltage characteristic of the cell, since the voltage is given by $V = (\varepsilon_{FC} - \varepsilon_{PV})/e$. From its maximum power point the efficiency is finally determined.

The result for different energy gaps and optimized positions of the impurity level can be seen in Figure 8.12 for the AM0 spectrum. For this calculation, a high impurity concentration was assumed ensuring $a_{i,C} = a_{V,i} = 1$. The equations outlined above, however, allow us to account for smaller concentrations and inadequate occupation of the impurities, which may even vary in the course of the current–voltage characteristic.

The efficiency reaches a maximum value of $\eta = 0.46$ for an energy gap $\varepsilon_C - \varepsilon_V = 2.4$ eV and an impurity level at $\varepsilon_{imp} - \varepsilon_V = 0.93$ eV.

This result reminds us of the improvement of the efficiency by tandem cells because, here as well, the incident spectrum is divided over different transitions, leading to smaller thermalization losses. Figure 8.13 shows that the occupation of the various states is represented by more than two, namely, three, different Fermi energies, a condition for reduced thermalization losses known from the discussion of tandem cells. In fact, since the rate of recombination transitions is determined by the difference between the Fermi energies for the states involved, each transition can be represented by a current–voltage characteristic. This leads to the equivalent circuit shown in Figure 8.13. The solar cell representing the band–band transition is connected in parallel to a series connection of two solar cells representing the transitions involving the impurity level. Variations of the absorptance

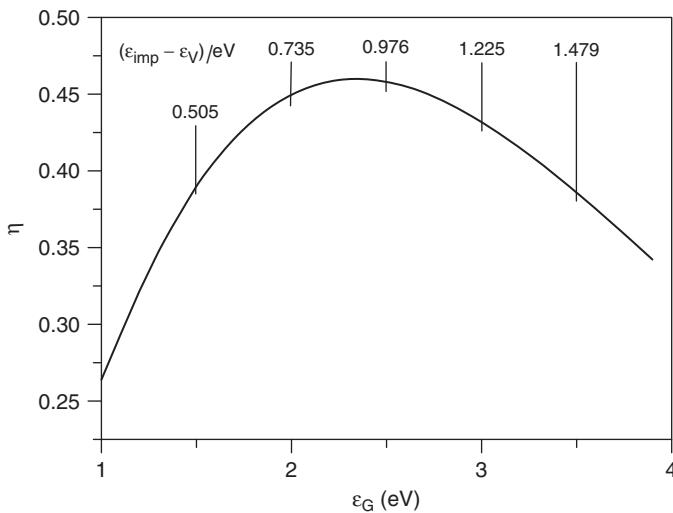


Figure 8.12 Efficiency as a function of the energy gap $\epsilon_C - \epsilon_V$ for radiative band–band transitions and radiative transitions between the bands and an impurity level at ϵ_{imp} .

Nonradiative transitions are excluded. The numbers at the curve give the optimal position of the impurity level with regard to the valence band for selected band gaps.

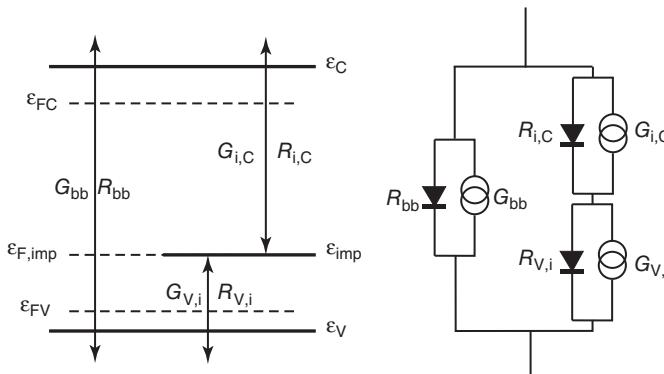


Figure 8.13 Equivalent circuit for a solar cell with an impurity level between valence and conduction bands as shown in Figure 8.11.

in the course of the current–voltage characteristic can, however, not be treated in an equivalent circuit model.

As for tandem cells, it is expected that the efficiency increases when more than one impurity level is present and the incident spectrum is divided into smaller portions over more transitions. Ensuring good absorption properties for all transitions, however, is a problem. Moreover, it must be emphasized that nonradiative recombination has been excluded. Although the optimal position of the impurity level for optical transitions is not in the middle of the energy

gap, where nonradiative recombination is most probable, including nonradiative recombination will certainly reduce the improvement expected from impurity transitions.

8.5.2

Up- and Down-conversion of Photons

A considerable loss of energy in a solar cell is due to the photons with energy $\hbar\omega < \epsilon_G$, which are not absorbed. It would be very convenient if two or more of these useless photons could be converted into one photon with energy $\hbar\omega \geq \epsilon_G$, which could then be absorbed by the solar cell. In the following discussion, ϵ_G defines the band gap of the solar cell for which small-energy photons will be up-converted. That such an up-conversion of the photon energy is not forbidden by thermodynamics is demonstrated by Figure 8.14, which shows a device consisting of a tandem of two small band gap solar cells connected to an LED. These solar cells absorb small-energy photons, and due to their series connection, deliver a voltage that is large enough to drive the LED with a large band gap to emit photons with energy $\hbar\omega \geq \epsilon_G$, which can be absorbed by the solar cell we have in mind. There is no doubt that this type of up-conversion will work. One may, however, ask why we do not use the electrical energy from the small-gap solar cells directly instead of investing it into an LED. We remember that the arrangement of two solar cells and an LED in Figure 8.14 is identical to the equivalent circuit of a three-level system represented by two bands and an impurity level in Figure 8.13 in the previous section.

A three-level system placed behind a solar cell could be used to convert small-energy photons transmitted by the solar cell into higher-energy photons supplied to the solar cell, in addition to the photons absorbed directly from the Sun as shown in Figure 8.15. A mirror behind the up-converter ensures that all emitted photons are directed toward the solar cell. A closer inspection of the impractical

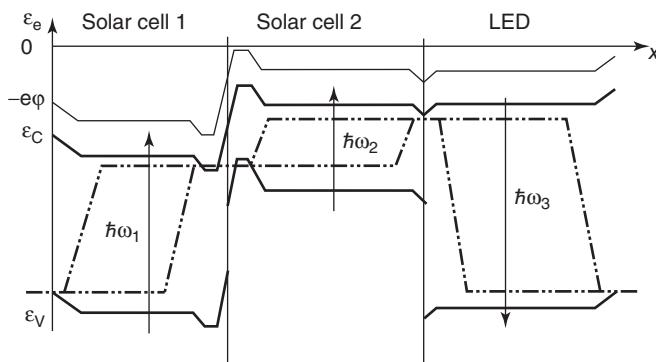


Figure 8.14 Two solar cells with small band gaps drive an LED with a large band gap, to emit photons useful for a large band gap solar cell, thereby up-converting two small-energy photons into one higher-energy photon.

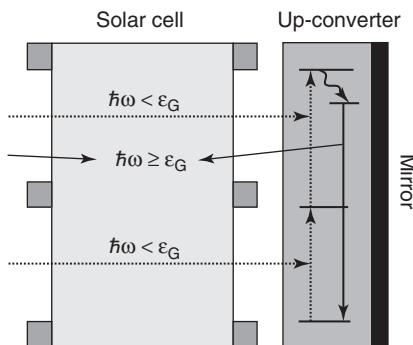


Figure 8.15 An up-converter behind a solar cell absorbs small-energy photons, transmitted by the solar cell, in a two-step excitation process. Higher-energy photons with $\hbar\omega \geq \varepsilon_G$ emitted by the up-converter generate additional electron–hole pairs in the solar cell.

but functioning up-converter in Figure 8.14 reveals that it is better represented by a four-level system than by a three-level system. The sum of the band gaps of the two solar cells is larger than the band gap of the LED. In fact, detailed calculations as outlined in the last section show that an energy loss, indicated in Figure 8.15 at the upper level, is necessary to prevent the recombination of the electron–hole pairs via the intermediate level with the reemission of two small-energy photons [31].

Figure 8.16 shows a substantial improvement in the efficiency of a solar cell for an incident 6000 K black-body spectrum. As always, the efficiency is larger for maximum concentration than for nonconcentrated radiation. The possible efficiencies of a solar cell with up-conversion and of the impurity photovoltaic effect are very similar. Both use two-step excitations to absorb otherwise non-absorbed photons. The up-conversion, however, has distinct advantages. First, the up-converter is a purely optical device and can consist of a material such as an organic dye, in which electrons and holes are virtually immobile, but which

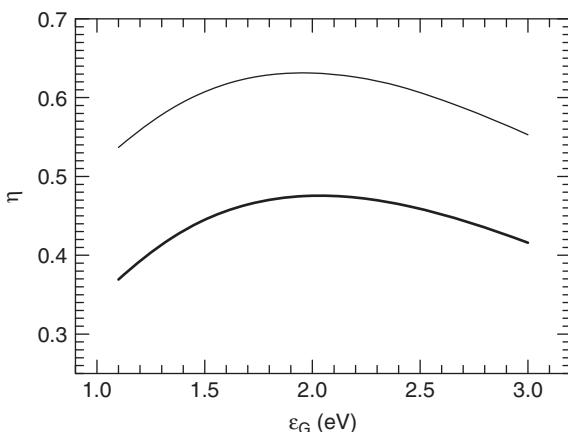


Figure 8.16 Efficiency of a solar cell as a function of its band gap ε_G operating with directly absorbed and with up-converted photons from a 6000 K black-body spectrum, from a solid angle $\Omega_S = 6.8 \times 10^{-5}$ (thick line) and with maximum concentration from $\Omega = \pi$ (thin line).

has a high quantum efficiency. Second, the up-converter is a separate device that can be applied to existing well-developed bifacial solar cells. Third, since the up-converter is separated from the solar cell, it would interfere very little with the recombination processes in the solar cell. Applying an up-converter to a solar cell would not do any harm but could only improve the solar cell's efficiency, even if it is not working quite as well as theoretically predicted.

Since in a three-level system the recombination from the upper level is more probable via the intermediate level instead of directly to the lowest level, a three-level system can be used for down-conversion [32]. In this process, a high-energy photon is absorbed in a transition from the lowest level to the upper level. By the back transition, via the intermediate level, two small-energy photons are emitted. Applied to a solar cell, a down-converter reduces the thermalization loss incurred by the absorption of photons with $\hbar\omega > 2 \varepsilon_G$ by splitting these photons into two photons with $\hbar\omega \geq \varepsilon_G$. Since high-energy photons are not transmitted by the solar cell, the high-energy part of the spectrum must be diverted by a dichroic mirror or other means to the rear side of the cell, where the down-converter is placed. In addition, another dichroic mirror transmitting high-energy photons but reflecting small-energy photons is applied to the back of the down-converter, where it prevents the loss of small-energy photons produced in the down-converter. The efficiency for a solar cell combined with a down-converter on its backside is shown in Figure 8.17 by the thin line as a function of the band gap of the solar cell.

A much simpler and more elegant method would be to place the down-converter on the front side of the solar cell. In this arrangement, all the incident

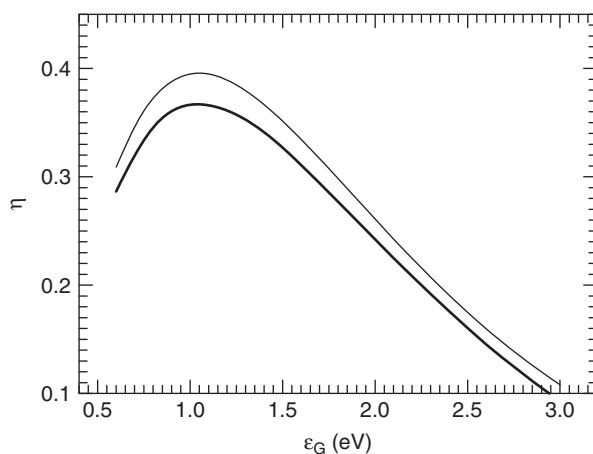


Figure 8.17 Efficiency of a solar cell as a function of its band gap ε_G operating with directly absorbed and with down-converted photons from a nonconcentrated 6000 K black-body spectrum for a down-converter

placed on the rear side of the solar cell (thin line). The efficiency for operation only with photons from the down-converter when it is placed on the front side is smaller (thick line).

solar photons, which the solar cell could absorb, are absorbed by the down-converter. Photons with $\varepsilon_G \leq \hbar\omega < 2\varepsilon_G$ are absorbed in transitions involving the intermediate level and larger-energy photons cause direct transitions from the lower to the upper level. Nearly all photons emitted by the down-converter have an energy $\hbar\omega > \varepsilon_G$ and could be absorbed by the solar cell. However, being on the front side of the solar cell, no mirror directing all the emitted photons into the solar cell can be applied and one might think that this deficiency leads to the loss of one half of the photons being emitted through the front surface toward the Sun. This is not necessarily the case. We remember that the probability for emission is proportional to the density of photon states. This makes the emitted photon currents proportional to the square of the index of refraction. If a material for the down-converter is chosen that has the same large index of refraction as the solar cell, all the photons emitted toward the solar cell enter the solar cell without reflection, whereas most of the photons emitted toward the front surface are totally reflected.

For the down-converter on the rear side of the solar cell, a maximum efficiency of almost 40% is found. A smaller efficiency of 36% is found, caused by the loss of down-converted photons, when the down-converter is placed on the front side of the solar cell. An index of refraction of $n = 3.6$ was used for the down-converter and the solar cell, which prevents the loss by total internal reflection of most down-converted photons.

Small band gap solar cells are advantageous in combination with a down-converter, whereas higher band gap solar cells are favorable with an up-converter. For both systems, the calculations consider only radiative transitions, which may, however, be closer to reality in materials that do not require good transport properties for electrons or holes.

8.6 Problems

- 8.1** List the different steps that take place in a semiconductor device that lead to the conversion of solar heat radiation into electrical energy. Comment on the phenomena that cause losses and make suggestions on how to overcome these if possible.
- 8.2** Consider a sample of n-doped bulk Si with a donor concentration of $n_D = 10^{15} \text{ cm}^{-3}$, fully ionized. Regard three different recombination mechanisms:
- Radiative recombination as defined in Equation 3.55 with $B = 3 \times 10^{-15} \text{ cm}^3 \text{ s}^{-1}$.
 - Auger recombination as defined in Equation 3.63 with $C_e = C_h = 10^{-30} \text{ cm}^6 \text{ s}^{-1}$.
 - Impurity recombination (via acceptor-like impurity states) with $n_{\text{imp}} = 10^{13} \text{ cm}^{-3}$, $\varepsilon_{\text{imp}} - \varepsilon_V = 0.7 \text{ eV}$, $\tau_{\min, e} = 10^{-3} \text{ s}$, and $\tau_{\min, h} = 10^{-5} \text{ s}$.

- (a) Calculate and plot the total recombination rate and the contribution of each of the three recombination mechanisms as a function of the minority carrier density n_h .
- (b) Calculate the difference of the quasi-Fermi energies $\varepsilon_{FC} - \varepsilon_{FV}$ as a function of the recombination rate. Consider the recombination as being caused by each of the three mechanisms alone and by the combination of all of them, respectively.

9**Characterization of Solar Cells**

In the previous chapters, we have described how solar cells work and how properties like diffusion length, lifetime, shunts, and series resistances affect their performance. It is of great importance to know these properties of the material and of the cell design to correctly interpret the behavior of solar cells and to determine the factors that limit their efficiency and need to be optimized.

In the following, several techniques are described and discussed by which properties of solar cell materials and solar cells can be determined experimentally.

9.1**Spectral Response and Quantum Efficiency**

When measuring the current density $j_Q(V)$ of a solar cell, information about the integrated generation and recombination rate is obtained: $j_Q(V) = -e \int_0^d \{G(x) - R(x, V)\} dx$. However, the wavelength dependence of these quantities is not resolved. In order to extract more information, the spectral response (SR) of the solar cell is measured. It tells how many charge carriers are extracted as a function of incident photons in a certain (narrow) wavelength interval $\lambda \pm \Delta\lambda$. The SR signal is therefore defined as

$$SR(\lambda \pm \Delta\lambda) = \frac{j_Q(\lambda \pm \Delta\lambda)}{j_E(\lambda \pm \Delta\lambda)} \quad (9.1)$$

with $j_E(\lambda \pm \Delta\lambda)$ being the energy current density in that specific wavelength interval. Note that SR can be measured at any voltage but in most cases the solar cell is placed under short-circuit conditions.

A typical measurement setup is shown in Figure 9.1. The light source is followed by a chopper wheel that interrupts the photon flux periodically (e.g., at 100 Hz). Behind the chopper wheel is a monochromator that transmits only photons of a certain (narrow) wavelength interval. The light is then used to illuminate both a so-called monitor cell (for which the SR is already known) and the actual solar cell to be measured. Because of the chopper wheel, the light shows a specific on–off pattern with time and therefore lock-in amplifiers are used.

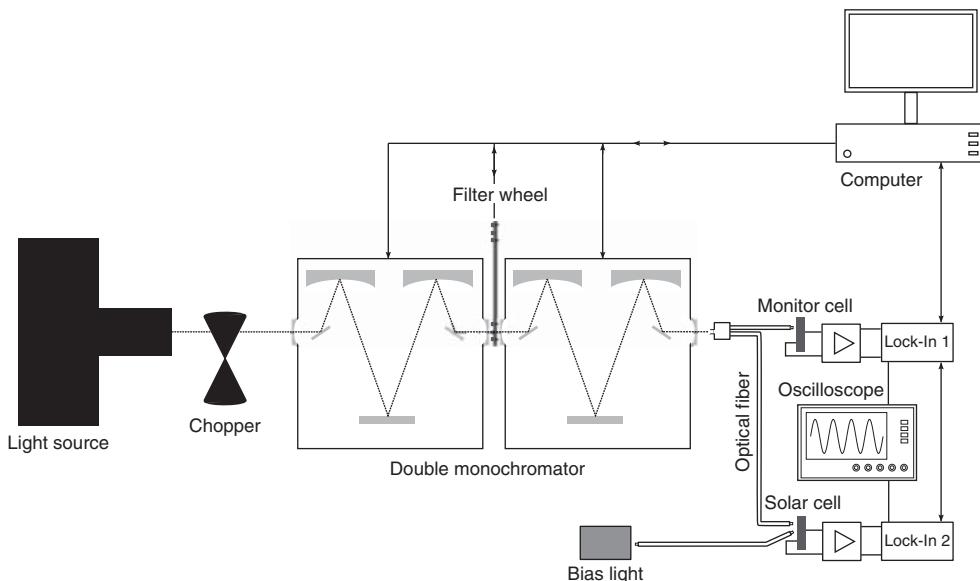


Figure 9.1 Measurement setup to determine the spectral response of a solar cell.

They are fed with the frequency of the on–off pattern and the signal from the solar cell (usually a voltage due to the solar cell current flowing through a defined ohmic resistor) is compared with that frequency. All signals that have a different frequency are very effectively suppressed by the lock-in amplifier. With the lock-in technique it is thus possible to enhance the signal-to-noise ratio by up to a factor of 10^5 – 10^6 . This is especially important if a bias light is used to subject the solar cell to a condition close to “1 sun.” In this case, the bias (white) light is typically a factor of ~ 1000 more intense than the monochromatic light. Hence, even a tiny fluctuation of the bias light intensity can cause a change in the short-circuit current of the solar cell, which is much larger than the current originating from the monochromatic light. With the lock-in technique, this small current can reliably be detected.

Once the SR is known as a function of wavelength, the so-called external quantum efficiency (EQE) can be calculated by

$$\text{EQE}(\lambda) = \frac{\text{SR}(\lambda)e_{\gamma}(\lambda)}{e} = \text{SR}(\lambda)hc/(e\lambda) \quad (9.2)$$

EQE gives information about how efficiently photons are converted into electron–hole pairs by the solar cell as a function of wavelength. Finally, if it is known how many photons are reflected or transmitted (i.e., how many photons are not absorbed in the photoactive material and do not therefore generate an electron–hole pair), the so-called internal quantum efficiency (IQE) can be derived with $R(\lambda)$ being the reflectance:

$$\text{IQE}(\lambda) = \frac{\text{EQE}(\lambda)}{(1 - R(\lambda))} \quad (9.3)$$

Thus, knowledge of the quantities EQE and IQE enables distinguishing the loss mechanisms due to nonoptimal optical absorption properties of the entire device from the photoconversion properties of the absorber. However, care has to be taken, as this does not mean that only the electrical properties of the solar cell, that is, the collection efficiency of generated charge carriers, are probed by the determination of its IQE. This can be seen clearly when considering that most solar cells comprise a highly reflective back surface such that no photons are transmitted in the wavelength range of interest. In this case, the reflected light contains photons that are reflected at the front surface of the solar cell (e.g., from a metal finger or from the absorber itself for nonideal antireflection coatings) as well as photons that are transmitted through the photoactive layer, reflected at the back electrode, and transmitted another time by the photoactive material, the so-called escape light. Of course, there may also be multiple courses through the photoactive material with multiple reflections at the back side for wavelengths for which the absorption coefficient is extremely small. However, as there is no unity reflection at the back side, that is, $R_b < 1$, a certain fraction of the photons will be absorbed by the back metallization. In addition, for weakly absorbed wavelengths, the light can pass several times through the absorber and therefore also through the strongly doped regions in the front and back. Therein, the photons can get absorbed by free carriers and excite these carriers within the conduction or valence band to states with higher energies. This excess energy is then quickly transferred to the lattice without any electron–hole pairs being generated. Every process in which photons are absorbed without generating an electron–hole pairs in the photoactive material is also referred to as *parasitic absorption*. As a consequence, the IQE contains information about how efficiently the generated electron–hole pairs are collected at the electrodes and it is influenced by parasitic absorption processes.

Another example is a silicon heterojunction solar cell with a monocrystalline absorber and carrier-selective membranes (see Section 6.2) made from amorphous silicon. In these devices, the c-Si absorber is sandwiched between thin layers of intrinsic a-Si followed by p- and n-doped a-Si layers, respectively. These membranes suppress surface recombination effectively, which is why a-Si/c-Si heterojunctions show the highest open-circuit voltages among all c-Si solar cells. A drawback of a-Si membranes is the fact that they absorb a small part of the incoming light. Although electron–hole pairs are generated in this material, their very small diffusion length impedes their efficient collection. Thus, these photons get lost, which leads to a reduction of the overall IQE even in the case of an ideal monocrystalline Si absorber. In Figure 9.2, the EQE of a c-Si/a-Si heterojunction solar cell is plotted together with the different losses due to reflection and parasitic absorption [33].

In addition to the relative losses indicated in Figure 9.2, the losses are also given in mA/cm^2 , which is obtained by multiplying the IQE values with the AM1.5G spectrum and subsequently integrating over the wavelength range 350–1200 nm. The parasitic absorption in the blue region is mostly due to interband transitions

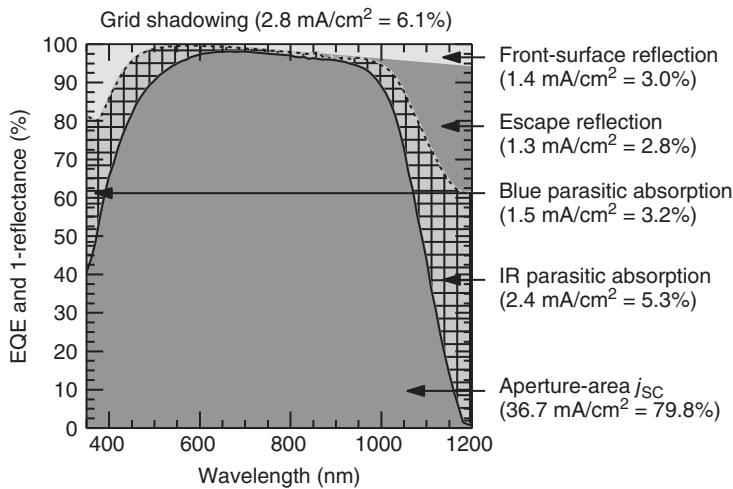


Figure 9.2 EQE and $1 - R$ (in %) of an a-Si/c-Si heterojunction solar cell with a power conversion efficiency of 21.3%. Shown are the j_{SC} and the different loss contributions, see text for details. Courtesy: Zachary C. Holman, Arizona State University.

in the a-Si layers, whereas in the infrared region it is mainly caused by free carrier absorption.

The probability that a generated electron–hole pair is collected at the contacts of a solar cell depends on where in the photoactive material it is generated. In a device with p-doped base, the electrons are the minority carriers and, on average, they can only travel one diffusion length through the p-doped region before they recombine. Thus their probability to be collected is lower the deeper in the p-doped region, that is, the farther away from the pn-junction they are generated. As the absorption coefficient decreases with increasing wavelength (Figure 3.17), relatively more charge carriers are generated closer to the back electrode for longer wavelengths. Thus, for the generated electrons, the average probability to be collected is lower at longer wavelengths. The exact behavior depends, of course, on the ratio between the thickness of the active material and the diffusion length of the minority carriers, that is, on the quality of the absorber material. In the following, we assume a pn-junction solar cell with thickness d with a thin n⁺-emitter at the front side and a p-doped base region wherein the electrons have a diffusion length $L_{e,B}$, and S is the surface recombination velocity at the back electrode. First, we have to solve the equation that describes the transport of the minority carriers through the p-doped region:

$$D_e \frac{d^2 n_e^p}{dx^2} = \frac{n_e^p - n_e^{p,0}}{\tau_e} - j_\gamma \alpha (1 - R) \exp[-\alpha x] \quad (9.4)$$

This is the continuity equation (see Equation 3.53) for the electrons in the p-doped region under the assumption that their transport occurs solely by diffusion and that they are generated predominantly in the p-doped region (which is fulfilled

for all wavelengths not too strongly absorbed, as the n⁺-region at the front surface is very thin). For $\alpha d \gg 1$, the IQE is given by [34]:

$$\text{IQE} = \frac{\alpha L_{e,B}}{\alpha^2 L_{e,B}^2 - 1} \left\{ \alpha L_{e,B} - \frac{(SL_{e,B}/D_e) + \tanh(d/L_{e,B})}{1 + (SL_{e,B}/D_e) \tanh(d/L_{e,B})} \right\} \quad (9.5)$$

We define an effective diffusion length $L_{e,eff}$ by

$$L_{e,eff} = L_{e,B} \frac{1 + (SL_{e,B}/D_e) \tanh(d/L_{e,B})}{(SL_{e,B}/D_e) + \tanh(d/L_{e,B})} \quad (9.6)$$

which takes into account additional recombination at the back surface, which reduces the diffusion length from its bulk value to the effective value and thus reduces the IQE. We will distinguish two cases: $L_{e,B} \ll d$ and $L_{e,B} \geq d$. In the case of $L_{e,B} \ll d$, $\tanh d/L_{e,B} \approx 1$ and therefore $L_{e,eff} \approx L_{e,B}$, which could be expected as in this case the minority carriers do not reach the back surface. Equation 9.5 then results in

$$\text{IQE} = \frac{\alpha L_{e,eff}}{\alpha L_{e,eff} + 1} \quad (9.7)$$

For $L_{e,B} \geq d$, it is also $\alpha L_{e,B} \gg 1$, which leads to

$$\text{IQE} = 1 - \frac{1}{\alpha L_{e,eff}} \quad (9.8)$$

Finally, from Equation 9.7 and using $\alpha L_{e,B} \gg 1$, for Equation 9.8 we get in both cases, that is, independent of the value of $L_{e,B}$

$$\text{IQE}^{-1} = 1 + \frac{1}{\alpha L_{e,eff}} \rightarrow \frac{1}{L_{e,eff}} = \frac{d\text{IQE}^{-1}}{d\alpha^{-1}} \quad (9.9)$$

Thus plotting the inverse IQE versus the inverse absorption coefficient, $L_{e,eff}$ is given by the inverse of the slope.

For wavelengths for which the penetration depth is larger than the device thickness ($\alpha^{-1} > d$), multiple reflections can occur, particularly at textured surfaces. This requires considering optical model equations, which was done by Basore [35].

In the last part of this section, an example is given to illustrate quantitatively the influence of $L_{e,B}$ on the IQE of a solar cell. Figure 9.3 shows the simulated IQE versus wavelength of three silicon solar cells with a thickness of 250 μm, which are identical except for their diffusion lengths $L_{e,B}$ in the weakly p-doped base, which are 100, 250, and 1000 μm, respectively.

The lower x -axis indicates the wavelength, and the upper x -axis indicates the absorption depth $1/\alpha$ (see Section 3.5.1), that is, the distance from the surface at which the incident light intensity drops to $1/e$. It can be seen that all three cells have an IQE of unity up to ~600 nm. Toward larger wavelengths, an increasing difference is noticeable in the IQE of the cell with the lowest $L_{e,B}$, which decreases most rapidly. The p-doped base makes up for most of the cell, and there is only a very thin n⁺-doped region in the front. Apart from the photons with very short

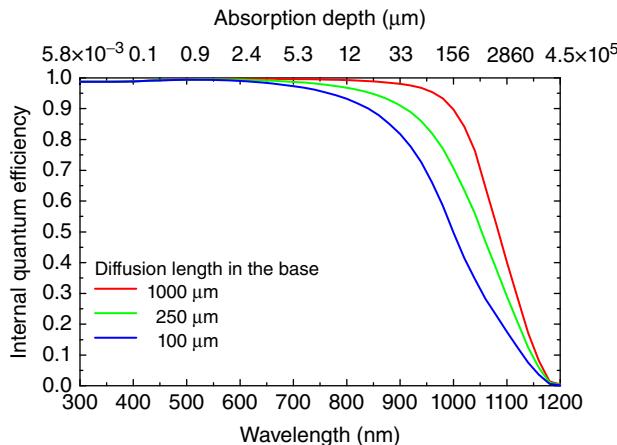


Figure 9.3 IQE curves of three silicon solar cells (all with a thickness of $d = 250 \mu\text{m}$) with different electron diffusion lengths in their p-doped base. Courtesy: Stefan Glunz, Fraunhofer ISE.

wavelengths, most of the light is absorbed somewhere in the p-doped region. Therefore, electrons, which are the minority carriers, have to reach the n⁺-doped region within their lifetime in order to contribute to the current of the solar cell. Only a fraction $1/e$ of the electrons that are generated at a distance of one diffusion length away from the junction (see Section 5.3) will reach the junction. For the cell with $L_{e,B} = 100 \mu\text{m}$, half of the electrons generated from photons with $\lambda = 1000 \text{ nm}$ (absorption depth $\approx 150 \mu\text{m}$) recombine before they reach the n⁺-region, whereas for the cell with $L_{e,B} = 1000 \mu\text{m}$ most of them still contribute to the current.

Note that the strong decrease of the signal for all three solar cells toward even larger wavelengths is caused by the rapid decay of the absorption in the photoactive layer and an increase of the parasitic absorption at the back electrode as well as by free carriers as the photon energy approaches the bandgap.

The evaluation of the IQE can, therefore, provide important information that is not directly accessible from the $I-V$ characteristic. In the present example, it shows that the material quality for the cells with diffusion lengths of 100 and 250 μm is too low to allow for a 250-μm-thick absorber. In this case, either the material quality has to be improved or – if this is not feasible – the solar cells would have to be made thinner, as this brings up the voltage. An improved light trapping would then be required in order to minimize transmission losses due to the reduced thickness.

9.2

Quasi-Steady-State Photoconductance

The quasi-steady-state photoconductance (QSSPC) measurement is a technique used to determine the lifetime of photogenerated charge carriers and was

introduced in 1996 by Sinton, Cuevas, and Stuckings [36]. In a QSSPC measurement, a solar cell is placed on top of a coil which is part of a high-frequency AC circuit. A light pulse generates excess charge carriers in the valence and conduction band of the semiconductor material. The charge carriers interact with the high-frequency electromagnetic field and generate eddy currents that are either detected with a second coil acting as receiver or monitored as changes in the phase and amplitude of the current flowing in the primary coil. The decay of the electrical conductivity (see Equation 5.3) of the sample due to recombination of electron–hole pairs causes a change in this eddy current and a corresponding change in the phase and amplitude of the current in the AC circuitry. As the time-dependent intensity of the flash light is known, the lifetime of the excess charge carriers can be derived from the decay time of the signal. Since the conductivity of a metal is much larger than of a semiconductor, the solar cell has to be measured before the deposition of any metal contacts.

Figure 9.4 shows the results from a QSSPC measurement of a high-efficiency crystalline silicon solar cell before the application of metal contacts [37]. The solid line shows the intensity of the optical excitation in multiples of 1 sun (left *y*-axis), and the dashed line corresponds to the photoconductance (right *y*-axis), both as functions of time.

The lifetime is calculated from the decay of the photoconductance; the first point used for the analysis is marked with a black dot. The results are shown in Figure 9.5, where the lifetime of the (minority) carriers is plotted against the excess carrier density Δn . Our example shows a p-doped sample, that is, the electrons are the minority carriers, and we know from Section 3.6.3 that in this case $n_e = n_e^0 + \Delta n \approx \Delta n$. We have used the fact that $\Delta n_e = \Delta n_h = \Delta n$.

In Equation 3.88, we have derived the lifetime of the minority carriers for radiative recombination in the case of low injection, for which the hole concentration

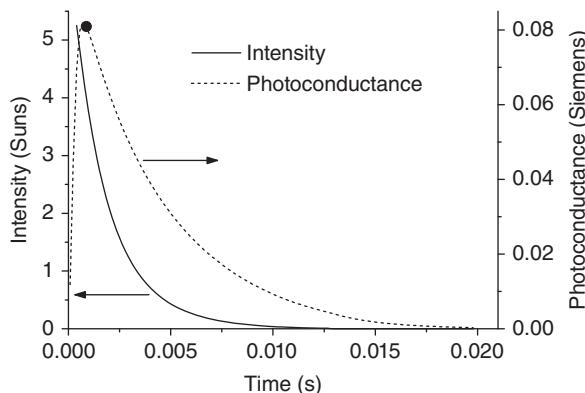


Figure 9.4 QSSPC data from a high-efficiency crystalline silicon solar cell without applied metal contacts [37]. The left *y*-axis shows the intensity in multiples of 1 sun (solid line), and the right *y*-axis the transient photoconductance (dashed line). As the lifetime is derived from the decay of the photoconductance, the data analysis is carried out starting from the point marked with the black dot.

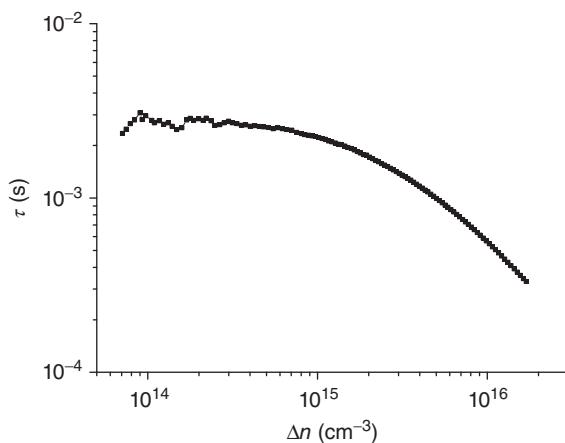


Figure 9.5 Lifetime versus excess carrier density derived from the QSSPC measurement shown in Figure 9.4 of a high-efficiency crystalline silicon solar cell without metal contacts. Courtesy: Frank Feldmann, Fraunhofer ISE.

of the p-doped sample is $n_h = n_h^0 + \Delta n \approx n_h^0$. For this reason, the lifetime of the electrons in Equation 3.88 is constant, as the number of their recombination partners does hardly change over time. In the curve shown in Figure 9.5, it can be seen that the lifetime does indeed depend on Δn . This is for two reasons: first, the sample was doped with $N_A = 1.5 \times 10^{16} \text{ cm}^{-3}$, therefore $n_h > n_h^0$, especially in the beginning of the measurement where Δn is large; second, it has to be taken into account that the decay of Δn with time is caused by different recombination processes, not only by radiative recombination as was assumed in Equation 3.88. For larger values of Δn , it is mostly Auger recombination that limits the lifetime of the minority carriers. For smaller values of Δn , it can be seen that the lifetime approaches a constant value. However, in this regime the signal becomes increasingly influenced by noise.

Hint: it is possible to correct the data for the impact of Auger recombination as the total lifetime is given according to $1/\tau_{\text{tot}} = \sum_i \frac{1}{\tau_i}$ (see Section 3.6.3).

A sophisticated parameterization of the Auger term for crystalline silicon was derived by Richter *et al.* [38]. The lifetime τ^* in absence of Auger recombination is given by $\tau^* = ((1/\tau_{\text{tot}}) - (1/\tau_{\text{Aug}}))^{-1}$, where τ_{tot} is the measured total lifetime.

9.3

Luminescence

The basis for the characterization of solar cells and solar cell materials by luminescence is the generalized Planck equation (3.99 and 3.102), which is shown again here in its local form describing the photon emission rate per volume inside a semiconductor at a position x into the solid angle Ω :

$$dr_\gamma(\hbar\omega, x) = \alpha(\hbar\omega) \frac{\Omega n^2}{4\pi^3 \hbar^3 c_0^2} \frac{(\hbar\omega)^2 d\hbar\omega}{\exp\{[\hbar\omega - \Delta\epsilon_F(x)]/kT\} - 1} \quad (9.10)$$

where the difference of the Fermi energies is $\Delta\epsilon_F(x) = \epsilon_{FC}(x) - \epsilon_{FV}(x)$.

On their way toward the surface at $x=0$, through which the photons are emitted, some of them are absorbed. Photons emitted in a slice of thickness dx at the distance x from the surface are attenuated by a factor $\exp(-\alpha x)$. Summing over all contributions from slices at different distances x from the surface up to the rear surface at $x=d$ gives the emitted spectral photon current density:

$$\begin{aligned} dj_\gamma(\hbar\omega) = & \int_0^d \alpha(\hbar\omega)(1 - R_f)\{\exp(-\alpha(\hbar\omega)x) \\ & + R_r \exp[-\alpha(\hbar\omega)(2d - x)]\} \frac{\Omega}{4\pi^3 \hbar^3 c_0^2} \frac{(\hbar\omega)^2 d\hbar\omega}{\exp\left\{\frac{\hbar\omega - \Delta\epsilon_F(x)}{kT}\right\} - 1} dx \end{aligned} \quad (9.11)$$

where reflection at the front (R_f) and rear surface (R_r) at $x=d$ have been included.

In order to simplify the analysis, we restrict ourselves to cases where the difference of the Fermi energies $\Delta\epsilon_F(x)$ is smaller by some kT than the bandgap of the semiconductor and thus smaller than all photon energies for which the absorption coefficient $\alpha(\hbar\omega) > 0$. This approximation is well justified in solar cells and allows us to neglect the “−1” in the denominator of the Planck equation and the “+1” in the denominator of the Fermi functions and to express the difference of the Fermi energies by the concentrations of electrons and holes by $\exp\{\Delta\epsilon_F(x)/kT\} \approx ((n_e(x)n_h(x))/(n_i^2))$.

In this approximation, Equation 9.11 is

$$\begin{aligned} dj_\gamma(\hbar\omega) = & \frac{\Omega}{4\pi^3 \hbar^3 c_0^2} (\hbar\omega)^2 \exp\left(-\frac{\hbar\omega}{kT}\right) d\hbar\omega \\ & \times \int_0^d \alpha(\hbar\omega)(1 - R_f)\{\exp(-\alpha(\hbar\omega)x) + R_r \exp[-\alpha(\hbar\omega)(2d - x)]\} \\ & \times \frac{n_e(x)n_h(x)}{n_i^2} dx \end{aligned} \quad (9.12)$$

We see how the difference of the Fermi energies or the distribution of electron and hole concentrations determines the intensity and the spectrum of the emitted luminescence radiation. It is, however, very difficult to deduce this distribution from the luminescence spectrum, except in special cases.

9.3.1

Homogeneous Distribution of Electrons and Holes

In a high-quality material, where the diffusion lengths of electrons and holes are much larger than the thickness of the material, the distribution of the charge carriers and of the difference of the Fermi energies is homogeneous. Taking the

carrier concentrations out of the integral in Equation 9.12, the remaining integral is the absorptance $a(\hbar\omega)$, and Equation 9.12 is

$$dj_{\gamma}(\hbar\omega) = a(\hbar\omega) \frac{n_e n_h}{n_i^2} dj_{\gamma,bb}(\hbar\omega) = a(\hbar\omega) \exp\left(\frac{\Delta\varepsilon_F}{kT}\right) dj_{\gamma,bb}(\hbar\omega) \quad (9.13)$$

where $dj_{\gamma,bb}(\hbar\omega)$ is the black body photon current density at 300 K. Equation 9.13 was the basis for determining the maximum efficiencies by Shockley and Queisser [11] for $a(\hbar\omega \geq \varepsilon_G) = 1$. If there is only radiative recombination, the integration over all photon energies $\hbar\omega \geq \varepsilon_G$ gives the total recombination rate per surface area, as presumed by Shockley and Queisser. Equations 9.10–9.13 are, however, generally valid even in the presence of dominant nonradiative recombination.

Since measuring the luminescence does not require contacts, Equation 9.13 can be used to determine the potential open-circuit voltage $\Delta\varepsilon_F/e$ for a material with large diffusion lengths without having to make a solar cell from it. It may thus be a valuable method to compare different surface passivation techniques. Equation 9.13 may also be used to determine the spectral absorptance and from it the absorption coefficient. If the absorptance $a(\hbar\omega)$ is known for one photon energy from a transmission experiment, Equation 9.13 can be adjusted and it is not even necessary to make absolute intensity measurements of the luminescence or to know the difference of the Fermi energies $\Delta\varepsilon_F$. It turns out that the absorption coefficient for band-to-band transitions can be obtained by this method, even though it is smaller than the contributions from other absorption processes such as free carrier absorption [40].

9.3.2

Inhomogeneous Distribution of Electrons and Holes

In many cases, the diffusion lengths are not much larger than the thickness, making the distribution of the charge carriers inhomogeneous. In this case, the carrier distribution may be obtained theoretically from a model, and Equation 9.12 predicts a theoretical luminescence spectrum for this model. A comparison with an experimental spectrum may then prove the validity of the model. The method is explained by the following example.

We take a slab of material that is homogeneously p-doped with a concentration of acceptors, which is large compared with the concentration of electrons, which are either injected through the electron contact by a forward bias or generated by strongly absorbed light. In the second case, contacts are not required. In both cases, the source of the electrons is restricted to close to the surface, from where they diffuse into the bulk of the material. Since the additional carrier concentration is small compared to the doping concentration (low injection), the concentration of the holes as the majority carriers is little affected and remains homogeneous. Drastic changes are observed for the electrons, which are the minority carriers. Their distribution follows from their equation of continuity under steady-state conditions:

$$\frac{\partial n_e}{\partial t} = G_e - R_e - \operatorname{div} j_e = 0 \quad (9.14)$$

For diffusion, the minority particle current density $j_e = -D_e \operatorname{grad} n_e$, where D_e is the diffusion coefficient of the electrons. Since there is no generation in the bulk, $G_e = 0$. The recombination rate $R_e = n_e/\tau_e$ contains contributions from radiative and nonradiative recombination. Using the relation $L_e = \sqrt{D_e \tau_e}$ between the diffusion length L_e and lifetime τ_e , we find from Equation 9.14 the diffusion equation

$$\frac{d^2 n_e(x)}{dx^2} - \frac{n_e(x)}{L_e^2} = 0 \quad (9.15)$$

A general solution is

$$n_e(x) = A \exp\left(\frac{x}{L_e}\right) + B \exp\left(-\frac{x}{L_e}\right) \quad (9.16)$$

The coefficients A and B follow from the boundary conditions.

- 1) At the front, at $x=0$, $n_e = n_e(0)$. This value accounts for injection or generation at the surface and fixes the recombination rate at the surface and in the bulk.
- 2) At the rear, at $x=d$, the remaining current of the electrons that did not yet recombine disappears by surface recombination and thus $-D_e \frac{dn_e}{dx}(d) = S_r n_e(d)$, where S_r is the surface recombination velocity at the rear surface. With these boundary conditions, the coefficients A and B in Equation 9.16 are

$$A = n_e(0) \frac{1 - r_d}{1 - r_d + (1 + r_d) \exp(2d/L_e)} \quad \text{and}$$

$$B = n_e(0) \frac{1 + r_d}{1 + r_d + (1 - r_d) \exp(-2d/L_e)}$$

$$\text{where } r_d = \frac{S_r \tau_e}{L_e}.$$

If the diffusion length L_e is small compared to the thickness d , we get the expected simple result

$$n_e(x) = n_e(0) \exp\left(-\frac{x}{L_e}\right) \quad (9.17)$$

With the spatial decay of the minority carrier concentration n_e in Equation 9.16, Equation 9.12 can be evaluated, revealing the influence of the diffusion length and of the rear surface recombination velocity on the emission spectrum.

9.3.3

Electroluminescence

A typical silicon solar cell consists of a p-doped part over almost all of the thickness except for a very thin highly n-doped layer at the front and a thin highly p-doped

layer at the rear. By applying a negative voltage to the front contact, electrons are injected, which diffuse into the p-type bulk of the solar cell. Neglecting the contribution to the emission from the very thin space charge layer at the front surface, we have the same situation as described in the previous section. The concentration of the electrons at the edge of the pn-junction, which we take as $x = 0$, from where the electrons diffuse into the p-type bulk resulting in the distribution in Equation 9.17 is now fixed by the applied voltage V and is

$$n_e(0) = \frac{n_i^2}{N_A} \exp\left(\frac{eV}{kT}\right)$$

9.3.4

Photoluminescence

Photoluminescence has advantages and disadvantages over electroluminescence. The advantage is that the electrons and holes are generated optically. Thus, contacts are not required. Photoluminescence can be observed in any stage of solar cell processing, starting from the pure absorber material through making the charge carrier separating structure and ending with the finished solar cell. By proper calibration, the absolute intensity of photoluminescence can be measured. The generalized Planck law then allows to obtain the difference of the Fermi energies in an absorber material from a purely optical measurement.

The disadvantage is that the low-intensity luminescence must usually be measured in the presence of a much more intense illumination to generate electron–hole pairs. Illumination must, therefore, be restricted to a narrow wavelength interval, which can be blocked by appropriate filtering from the measurement of the photoluminescence.

When the electron–hole pairs are generated by light absorption away from the surface, their spatial distribution is different from the distribution resulting from injection of electrons and holes as in electroluminescence. The generation rate in the continuity equation for the electrons, then, is

$$G_e(\hbar\omega, x) = [1 - R_f(\hbar\omega)]\alpha(\hbar\omega) \exp[-\alpha(\hbar\omega)x] \quad (9.18)$$

leading to the differential equation

$$\frac{d^2 n_e(\hbar\omega, x)}{dx^2} - \frac{n_e(\hbar\omega, x)}{L_e^2} + \frac{[1 - R_f(\hbar\omega)]\alpha(\hbar\omega)}{D_e} \exp[-\alpha(\hbar\omega)x] = 0 \quad (9.19)$$

A general solution for the distribution of the electrons, then, is

$$n_e(\hbar\omega, x) = A \exp\left(\frac{x}{L_e}\right) + B \exp\left(-\frac{x}{L_e}\right) + C \exp[-\alpha(\hbar\omega)x] \quad (9.20)$$

The coefficients A , B , and C are obtained from the boundary conditions. Finding an analytical solution for particular boundary conditions can, however, become rather complicated. A detailed description can be found in [41].

The analysis becomes much simpler if we restrict the illumination to short-wavelength light, which is strongly absorbed. In this case, the last term in

Equation 9.20 can be neglected and the minority carrier concentration $n_e(x)$ follows from Equation 9.17.

9.3.5

Diffusion Length

How do good materials with a large diffusion length and bad materials with a small diffusion length differ in the emitted photon current density described by Equation 9.12? Photons emitted through the surface must originate from within their penetration depth $1/\alpha(\hbar\omega)$ from the surface; otherwise they are reabsorbed. We, therefore, expect that the emission of high-energy photons with a small penetration depth, which originate from close to the surface, is identical for good and bad materials provided $n_e(0)$ is the same in both materials. Small-energy photons with a large penetration depth, however, suffer little from reabsorption. Their emission through the surface reflects the spatial distribution of the electrons and, thus, the value of the diffusion length. The photon current spectrum of a good material will show a larger contribution of low-energy photons than the spectrum of a bad material, whereas the contribution of high-energy photons is identical as long as $n_e(0)$ is the same. This results in larger overall photon currents emitted by good materials. In the following, we assume that the distribution of the minority carrier concentration follows Equation 9.17, either by injecting them from the surface as in electroluminescence or by generating them close to the surface by strongly absorbed light in photoluminescence.

Solar cells may consist of good and bad parts side by side, as is expected for multi-crystalline silicon solar cells with good and bad grains. A spatial resolution over the surface can be achieved by scanning a single detector with a narrow field of view over the surface of the solar cell. This is a time-consuming method, which can, however, be very sensitive by employing lock-in techniques. A much faster method with good spatial resolution employs a digital camera, which makes an image of the emission from millions of spots (pixels) simultaneously in a single measurement. The number of different detector materials for arrays in cameras is, however, more limited than for single detectors. Cameras with InGaAs or Si charge-coupled device (CCD) detector arrays are readily available. Figure 9.6 shows an electroluminescence image of a commercial multicrystalline silicon solar cell taken with a silicon CCD camera after correction for the spectral quantum efficiency of the camera. The exposure time for this image was 1 s.

To make use of the different information in the short- and long-wavelength part of the spectrum, two images were measured, both through short-pass filters, one transmitting light with wavelengths <900 nm and the other with wavelengths <1000 nm. The 1000-nm image was divided by the 900-nm image pixel by pixel. By this division, all factors are eliminated that have the same effect on different parts of the spectrum but may vary over the area of the solar cell, in particular voltage variations caused by the series resistance.

The same division is done for theoretical spectra from Equations 9.12 and 9.17 as a function of the diffusion length. Figure 9.7 shows the theoretical values of

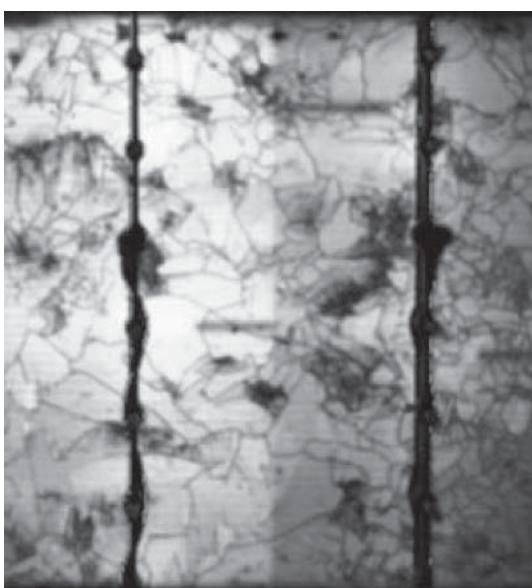


Figure 9.6 Electroluminescence intensity distribution from a multicrystalline silicon solar cell. In this cell design, the horizontal gridfingers in the left and right half of the cell are separately connected to the left and right

bus bars and not connected in the middle. The lower intensity of the right half is probably due to a lower voltage caused by a larger series resistance.

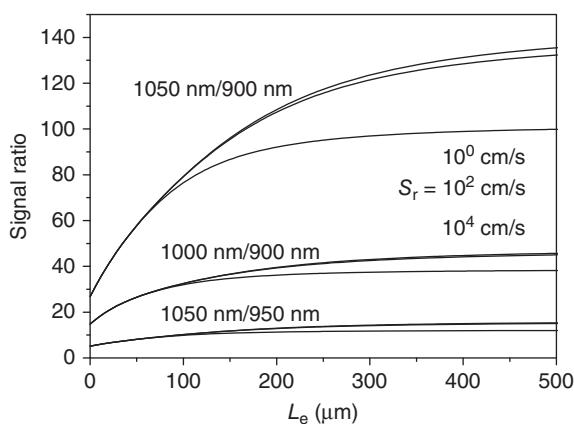


Figure 9.7 Theoretical intensity ratio of luminescence for several combinations of short-pass filters as a function of the diffusion length L_e . Three curves for each combination show the influence of the rear surface recombination velocity S_r .

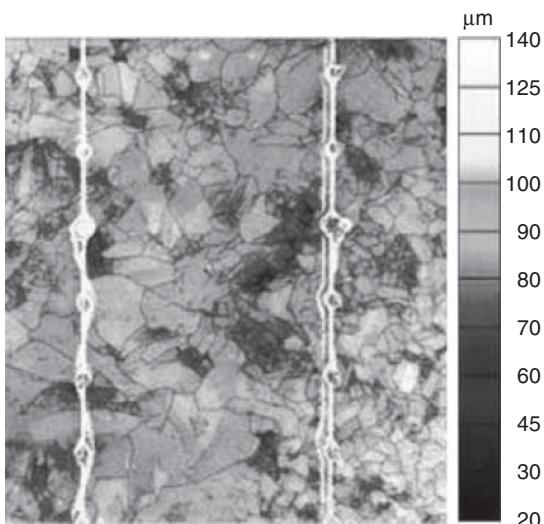


Figure 9.8 Diffusion length distribution for electrons obtained from the ratio of two luminescence images taken with 1000 and 900 nm short-pass filters.

intensity ratios as a function of the diffusion length. With this curve, the experimentally observed intensity ratios have been transformed into absolute values of diffusion lengths in Figure 9.8. As expected, the intensity difference between the left and right half of the image in Figure 9.6, caused by different voltages, is no longer present [42].

The distribution of diffusion lengths in a solar cell can also be obtained by light-beam-induced currents (LBIC, see Section 9.5), where one or several monochromatic light beams are scanned over the surface of a solar cell and the short-circuit current is measured. This current depends on the diffusion length, because electrons as minority carriers can contribute to the current only if they are generated within a diffusion length from the electron contact; otherwise they are lost by recombination.

The same technique for finding the diffusion length of the minority carriers can be used in photoluminescence. Since photoluminescence can be measured on a solar cell material before any processing, the method of obtaining the diffusion length from a ratio of photoluminescence signals at different wavelengths is even suitable to determine the quality of silicon blocks before they are sliced into wafers. Because of the variations in the front surface recombination velocity, the minority carrier concentration may vary over the front surface. This will affect the luminescence intensity but in the same way at all wavelengths of the emitted spectrum. Taking the ratio of intensities in different wavelength ranges removes the influence of the surface recombination velocity from the measurement of the diffusion length.

9.3.6

Series Resistance

Determining the series resistance of a solar cell is particularly simple with luminescence. From the generalized Planck equation 9.11, we know that the photon current density emitted by a solar cell is only a function of the difference of the Fermi energies, no matter how large the incident illumination intensity is or the charge current flowing through the solar cell. Therefore, if the same luminescence intensity is observed from a solar cell under different operating conditions, the difference of the Fermi energies, or the internal voltage of the solar cell, must be identical.

Figure 9.9 shows two current density versus voltage characteristics of a solar cell for two different illumination intensities. A first measurement of the luminescence intensity is made at point *A* at a given illumination intensity. For a second measurement, the illumination intensity is increased, resulting in the $j_Q(V)$ characteristic with the larger short-circuit current. If there were no series resistance in the solar cell, the same luminescence intensity as for measurement *A* would occur at the same terminal voltage *V*. With a series resistance, however, the same luminescence intensity as in *A* is observed at point *B* of the second characteristic. Because of the increase of the current density by Δj_Q from *A* to *B*, the voltage drop over the series resistance is increased by ΔV . The series resistance then is

$$R_s = \frac{\Delta V}{\Delta j_Q} \quad (9.21)$$

Using luminescence to determine the series resistance requires that the illumination of the solar cell, or rather the reflected part of it, does not interfere with the luminescence spectrum. A convenient illumination source would be a short-wavelength light-emitting diode (LED) or laser that is blocked from entering the detector by an appropriate filter.

This technique can even be applied to images of the luminescence intensity from a multicrystalline solar cell, as shown in Figure 9.6, allowing to obtain a spatial resolution of the series resistance [43]. In this case, a first image is taken at a given terminal voltage and illumination intensity, corresponding to point *A* in Figure 9.9.

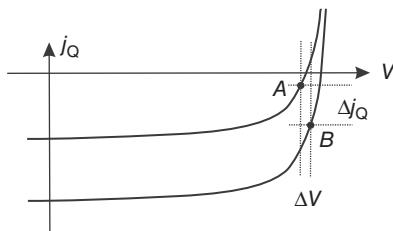


Figure 9.9 $j_Q(V)$ characteristics of a solar cell for two different illumination intensities. The luminescence intensity emitted by the solar cell is identical in points *A* and *B*.

The series resistance of the solar cell leads to an increase of the voltage by ΔV due to an increase of the current density by Δj_Q .

Since the terminal voltage at point B for a larger illumination intensity, which leads to the same luminescence intensity, may be different for each pixel of the image, several images are taken for the same larger illumination intensity but with different terminal voltages. By interpolation, the correct terminal voltage is found for each pixel that gives the same luminescence intensity as in the first image. Under the assumption of a homogeneous current density j_Q , a local series resistance is obtained for each pixel from Equation 9.21.

An overview of the many possibilities to determine the properties of solar cell materials and solar cells by luminescence was presented by Trupke *et al.* [44].

9.4 Thermography

In thermography, an image of the temperature distribution of an object is produced. Our objects are solar cells, and their temperatures are around 300 K. According to Equations 2.9 and 2.11, the blackbody spectrum is centered around a photon energy of 0.07 eV. Solar cells are no blackbodies, and due to their grid-like metalization, their surface has no uniform absorptance/emittance. Uniform blackbody behavior is achieved by placing an electrically insulating black foil in close contact with the surface. For an image, an array of sensors is needed. The sensors can be bolometers, where the temperature dependence of the resistance is used; or thermocouples, which produce a voltage proportional to a temperature difference; or pyroelectric detectors, which are only sensitive to temperature changes; or photoelectric sensors, in which the small energy of the incident photons produces electron–hole pairs. Materials with a small bandgap, such as indium antimonide or mercury cadmium telluride, are good candidates. To improve the signal-to-noise ratio, they are cooled to liquid nitrogen temperatures to be able to differentiate between the generation of electron–hole pairs by the radiation from the object and the thermal generation inside the sensor. The sensor array is placed in the focal plane of a camera lens. The lens must be made from a material that does not absorb in the infrared. Glass lenses cannot be used, but some plastics are a possibility.

With a direct infrared image, temperature differences on the order of 1 K can be detected. This is sufficient to detect defects in solar cell modules. Since the glass cover of the module is not transparent for the IR radiation from the solar cells, only the temperature distribution on the glass cover can be seen. Measurements can be made in the dark with an applied bias or under illumination. In the dark, shorts are revealed by their higher temperature, whereas badly connected cells and strings with little dissipated power remain at a lower temperature. Under illumination, it is almost the opposite. All cells receive the same energy current. Good and well-connected cells do not heat up as much as badly connected cells, since the former deliver more of the incident energy current as electrical energy to the outside load. Electrical shorts still appear as hot spots because they draw current from the surrounding cells and dissipate more power.

For the analysis of single cells, better spatial and temperature resolution is needed. Because of the good heat conduction of silicon or the glass substrates of thin-film solar cells, locally produced heat spreads out, which prevents good spatial resolution of local heat sources. For this reason, the energy current introduced into the cell either by applying a voltage or by illumination is pulsed. For a pulse frequency of 30 Hz, temperature modulations are present at a distance of 1 mm from a point heat source. To improve the temperature and spatial resolution, several images are taken between the pulses and are analyzed in an elaborate mathematical procedure emulating a lock-in technique [45]. The higher the frequency, the better the spatial resolution, but at the expense of a lower signal. In order to reduce the noise from which low signals suffer, hundreds or thousands of images are summed up, leading to measurement times in the order of tens of minutes. With these precautions, temperature amplitudes on the order of $10 \mu\text{K}$ are detected, and a spatial resolution down to $10 \mu\text{m}$ is achieved.

Since the local temperature is a direct result of the local power dissipation, shorts are most easily detected. Also for regions with large heat production due to large defect densities and enhanced nonradiative recombination, lock-in thermography is an appropriate technique. With sophisticated correlation techniques, the distribution of series resistances and lifetimes can be evaluated. These are properties that can more easily and in much shorter measurement times be detected by luminescence. A comparison of both techniques, namely lock-in thermography and luminescence, was made in [46].

9.5

Light-Beam-Induced Current (LBIC)

LBIC is an imaging technique for solar cells that delivers spatially resolved information. In this method, a light beam is scanned over the sample and the corresponding current is measured. The resulting image shows the contribution $\Delta I(x)$ of each point to the total current I of the solar cell. LBIC can be particularly used to identify spatial inhomogeneity due to, for example, grain boundaries in multicrystalline materials or defects induced by processing of the solar cell material. If the reflectivity $R(x)$ of the sample is inhomogeneous, it has to be measured in addition to the current to obtain the correct value: $\Delta I_{\text{corr}}(x) = \Delta I(x)/(1 - R(x))$.

This method can also be combined with bias light that illuminates the whole sample, and also at various applied bias voltages to keep the solar cell at a specific operating point, for example, its maximum power point. Furthermore, also the change in voltage $\Delta V(x)$ instead of the current can be measured. In this case, the method is called light-beam-induced voltage (LBIV). By performing LBIC measurements at different voltages, information about the local series and parallel resistance can be obtained [47].

If bias light is used, it generates a current that is much larger than the current generated by the scanning light beam with its small spot size. Therefore, the

scanning light beam has to be modulated in its intensity (either electronically or with a chopper wheel) in order to sensitively detect the small current it generates by lock-in amplification. The lock-in method is explained in Section 9.1.

A disadvantage of LBIC is the time it takes to scan over an entire sample. For a resolution of $100\text{ }\mu\text{m}$, it can take up to several hours to measure a complete $156\text{ mm} \times 156\text{ mm}$ cell.

Additional information can be obtained by using different wavelengths of the scanning light beam. As a lock-in amplifier selectively detects signals with a certain frequency, it is possible to perform a measurement with several different wavelengths simultaneously (see Figure 9.10). To do so, each wavelength is modulated with a different frequency such that the corresponding current can be identified. As different wavelengths have different penetration depths (see Section 9.1), it is thus possible to, for example, determine local diffusion lengths of the minority carriers.

LBIC allows the determination of the local EQE and IQE (see Equations 9.2, 9.3, and 9.9) and thus also of the local diffusion length of the minority carriers. This way, the impact of a spatial variation of specific material properties (e.g., grain boundaries or dislocations) can be analyzed thoroughly. LBIC images are very useful for comparison with photo and electroluminescence images.

Figure 9.11 shows the spatially resolved EQE of a $156\text{ mm} \times 156\text{ mm}$ multicrystalline Si solar cell through an LBIC measurement performed under short-circuit conditions using a wavelength $\lambda = 1064\text{ nm}$.

The EQE values are reduced because of reflected photons (reflection at the front surface and escape light, see Section 9.1) as well as of a limited diffusion

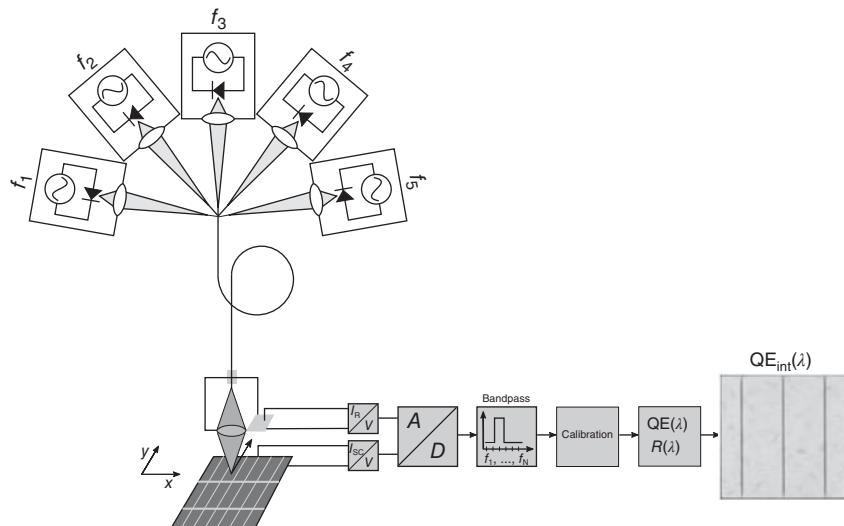


Figure 9.10 Scheme of an LBIC measurement setup with five different wavelengths (exemplarily).

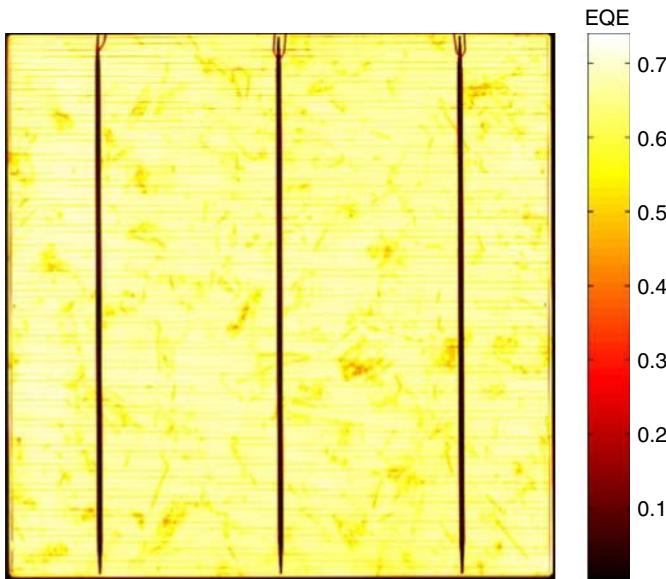


Figure 9.11 Spatially resolved EQE map derived from an LBIC measurement of a 156 mm × 156 mm multicrystalline Si solar cell detected at $\lambda = 1064\text{ nm}$. Courtesy: Martin Schubert, Fraunhofer ISE.

length. Furthermore, the randomly distributed dark structures originate from grain boundaries and dislocations.

A more detailed picture can be obtained by collecting LBIC data for various applied voltages. For example, Nguyen and coworkers have used LBIC to identify locally reduced shunt resistances [48]. The decisive feature in their analysis is a spot that shows, in a voltage range from small reverse bias to voltages near V_{OC} , a lower absolute current than its surrounding, whereas for a voltage significantly larger than V_{OC} the same spot shows a higher current than its surrounding.

To understand the origin of such a phenomenon, we have to imagine the solar cell as being composed of a large number of subcells connected in parallel (the number will be larger for higher resolution of the LBIC image). Equation 7.13 describes a solar cell with the two-diode model under explicit consideration of series and parallel resistances. Let us assume in the following that one of the mentioned subcells has a lower parallel resistance. Thus, we have to compare the current output of the subcell with the reduced R_p with a subcell with the nonreduced R_p . For the sake of simplicity, we use the one-diode model:

$$j_Q(V) = j_{S1} \left[\exp \left(\frac{V - j_Q R_S}{kT} \right) - 1 \right] + j_{SC} + \frac{V - j_Q R_S}{R_p} \quad (9.22)$$

For both subcells we use $j_{SC} = -40\text{ mA cm}^{-2}$ and $R_S = 5\Omega\text{ cm}^2$. For the “normal” subcell, we further assume an infinite value for the parallel resistance, that is, $1/R_p = 0$, and for the one with a reduced parallel resistance $R_p = 50\Omega\text{ cm}^2$.

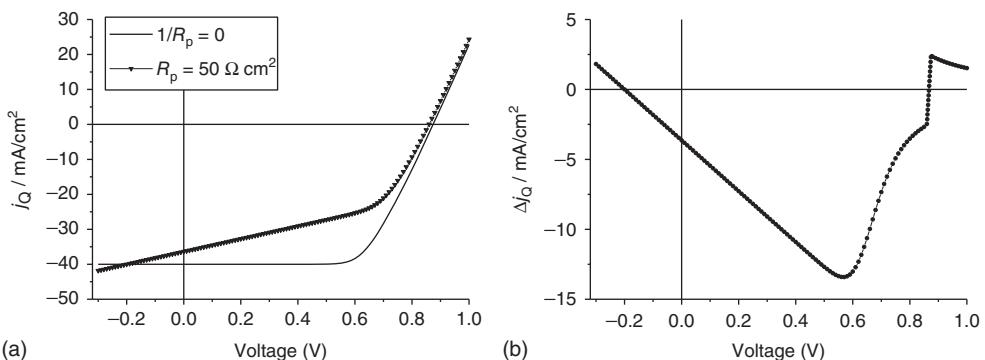


Figure 9.12 $j_Q(V)$ curves of two subcells with different values of their parallel resistance (a). Difference of the absolute values of the current densities: $\Delta j_Q = |j_Q(R_p = 50 \Omega \text{cm}^2)| - |j_Q(1/R_p = 0)|$ as a function of applied voltage (b).

Figure 9.12 shows (a) the two $j_Q(V)$ curves for the two subcells and the difference of the absolute values $|j_Q(R_p = 50 \Omega \text{cm}^2)| - |j_Q(1/R_p = 0)|$ (b). Indeed, we can see that the subcell with reduced parallel resistance has a lower absolute current for voltages ranging from a small reverse bias to values around V_{OC} (the subcell with infinite parallel resistance has a V_{OC} of 874 mV), and a larger absolute current for voltages exceeding V_{OC} .

At first glance, it may be surprising that the short-circuit current of the shunted subcell is lower, as there is no parasitic current at zero voltage. However, the series resistance causes a positive voltage drop over the diode (representing the recombination in the photoactive material) and thus a positive parasitic current that reduces j_{SC} . This behavior remains the same even in the reverse direction as long as the voltage drop over the diode is positive, that is, as long as the voltage drop over the series resistance is larger than the applied (negative) voltage. We can see that, in our example, this is the case until ~ -0.2 V. For voltages larger than V_{OC} , the parasitic current $j_{Q,p} = \frac{V - j_Q R_s}{R_p}$ and the sum of the photogenerated and diode current (i.e., the first two terms in Equation 9.22) are both positive, which is why the current of the shunted subcell is larger. Lastly, the difference of the currents of the two subcells becomes smaller for even larger forward voltages. This is due to the fact that the parasitic current is a linear function of the voltage whereas the diode (recombination) current is an exponential one. Note that the model and the parameters used here are not meant to thoroughly describe a real solar cell but only to help understand the underlying processes.

9.6

The Suns- V_{OC} Method

The Suns- V_{OC} method was introduced in the year 2000 by Sinton and Cuevas [49]. It is based on the $I_{SC} - V_{OC}$ method presented formerly by Wolf and Rauschenbach

[50]. The Suns- V_{OC} method delivers a pseudo current–voltage characteristic that the solar cell would have if there were no series resistance. By comparing this data with the normal current–voltage characteristic, the solar cell's series resistance can be determined. A comprehensive overview about the different methods to determine the series resistance of a solar cell was given by Pysch *et al.* [39].

Background

The current–voltage characteristic of a solar cell of thickness d is given by

$$\begin{aligned} j_Q(V') &= -e \int_0^d G_e(x) dx + e \int_0^d R_e(x, V') dx \\ &= -e [G_{e,tot}(j_\gamma) - R_{e,tot}(V')] \end{aligned} \quad (9.23)$$

We note that the total generation rate $G_{e,tot}(j_\gamma)$ depends only on the incident light intensity, whereas the total recombination rate $R_{e,tot}(V')$ depends only on the internal voltage V' .

The voltage V measured at the contacts is usually smaller than the internal voltage $V' = \Delta\epsilon_F/e$ due to the series resistance (for area) R_S .

$$V = V' - j_Q(V')R_S$$

It is of interest for the interpretation of the current–voltage characteristic whether the effect of the series resistance could be eliminated from Equation 9.23. This would be possible if we could determine the total recombination rate R_{tot} as a function of the internal voltage. This is accomplished with the Suns- V_{OC} method where the solar cell is illuminated by a slowly decaying flash light while its open-circuit voltage V_{OC} is continuously monitored. At the same time, the intensity of the flash light is detected with a reference device. As there is no current flow during the measurement, the data obtained directly from the measurement (i.e., V_{OC} as a function of illumination intensity) is not affected by the series resistance of the device. The Suns- V_{OC} method is fast and simple with the further advantage that it can already be carried out at an early stage, for example, before a grid is applied to the solar cell. However, in this case, the value of j_{SC} has to be known, that is, it has to be determined from identical cells with a grid.

Under short-circuit conditions, the integrated recombination rate is much smaller than the integrated generation rate, even in the presence of a series resistance. The short-circuit current density can, therefore, in good approximation be written as

$$j_{Q,SC} = -eG_{e,tot}(j_\gamma)$$

This generation rate is a constant in a current–voltage characteristic, that is, it does not depend on the voltage. The total generation rate and with it the short–circuit current density depend, however, linearly on the incident photon current density

$$j_{Q,SC} = -eG_{e,tot}(j_\gamma) = \beta j_\gamma \quad (9.24)$$

The constant β is found at 1 sun illumination:

$$\beta = j_{Q,SC}^{1\text{sun}} / j_{\gamma}^{1\text{sun}}$$

Under open-circuit conditions, $j_Q = 0$ and $V = V'$, since there is no voltage drop over the series resistance. In this case

$$G_{e,tot}(j_{\gamma}) = R_{e,tot}(V) \quad (9.25)$$

By the Suns- V_{OC} method, we measure the open-circuit voltage V_{OC} as a function of the incident intensity j_{γ} and express this relation between voltage and intensity as $j_{\gamma}(V)$. From Equations 9.24 and 9.25, we know that

$$-eR_{e,tot}(V) = j_{Q,SC}^{1\text{sun}} \frac{j_{\gamma}(V)}{j_{\gamma}^{1\text{sun}}} \quad (9.26)$$

By inserting Equation 9.26 into Equation 9.23, we find the current–voltage characteristic at 1 sun illumination without the effect of the series resistance. Since this characteristic has not been measured explicitly, it is called the *implied current–voltage characteristic*.

$$j_{Q,impl}^{1\text{sun}}(V) = j_{Q,SC}^{1\text{sun}} \left(1 - \frac{j_{\gamma}(V)}{j_{\gamma}^{1\text{sun}}} \right) \quad (9.27)$$

An example from a real solar cell is shown in Figure 9.13, where the normal $j_Q(V)$ curve and the pseudo $j_Q(V)$ curve of a high-efficiency crystalline silicon solar cell are plotted. When the series resistance is determined at the maximum power point (as the series resistance contains also non-ohmic contributions, the result depends slightly on the value of the current – or voltage – at which it is determined), the result is $R_S(j_{Q,MPP}) = 0.63 \Omega \text{cm}^2$.

For crystalline inorganic solar cells, the determination of the series resistance is mostly used to evaluate the influence of different metalization procedures, as the

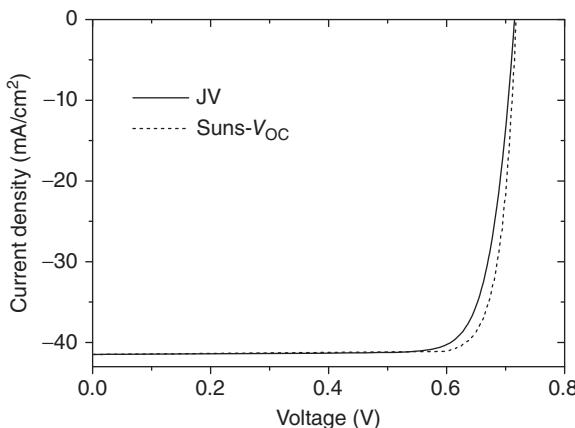


Figure 9.13 Normal $j_Q(V)$ curve (solid line) and pseudo $j_Q(V)$ curve obtained from the Suns- V_{OC} method (dashed line) of a high-efficiency crystalline silicon solar cell [37].

total series resistance is mainly determined by the metal grid (including contact resistances). The contribution of the photoactive layer is usually negligible, as the conductivity of the charge carriers is rather large. There are, however, exceptions when it comes to new cell concepts where the charge carriers have to cover larger distances before being collected, for example, in interdigitated back-contact solar cells [51].

The thickness of the absorber layer in organic solar cells is in the range of 100 nm and therefore roughly 3 orders of magnitude smaller than in crystalline silicon solar cells. However, this is overcompensated as the charge carrier mobilities are usually even 6–7 orders of magnitude smaller. As a consequence, the transport resistance (resulting from the transport of electrons and holes through the photoactive layer) can even be the dominant contribution to the overall series resistance of these devices [52]. Though this is, of course, an additional loss mechanism, the Suns- V_{OC} method can in this case be used to determine intensity-dependent charge carrier densities in the photoactive layer of such low-mobility solar cells [53]. The overall series resistance can, in general, be written as a (ohmic) resistance of the circuitry and a (non-ohmic) transport resistance of the photoactive layer:

$$R_S(j_Q) = R_{\text{circ}} + R_{\text{tr}}(j_Q) = \frac{V_{\text{IV}}(j_Q) - V_{\text{suns}}(j_Q)}{j_Q} \quad (9.28)$$

Note that there are also non-ohmic contributions that do not depend on the carrier density, such as contact resistances. However, they are usually rather small and their exhaustive description is beyond the scope of this section.

At high forward voltages, the charge carrier densities in the photoactive layer are very large, and hence R_{tr} can be neglected and R_{circ} determined. Once R_{circ} is known, R_{tr} can be calculated as a function of j_Q (and therefore also as a function of V). The voltage $\Delta V_{\text{tr}} = j_Q R_{\text{tr}}$ that drops over the transport resistance is related to the driving force for the electron and hole charge currents (Equation 5.14):

$$j_{e,h} = \frac{\sigma_{e,h}}{e} \text{grad } \epsilon_{\text{FC,FV}} \quad (9.29)$$

It can be written as follows for holes:

$$\Delta V_{\text{tr,h}}(x) = \int_0^x \frac{j_h(x')}{\sigma_h(x')} dx' \quad (9.30)$$

and for electrons:

$$\Delta V_{\text{tr,e}}(x) = \int_x^d \frac{j_e(x')}{\sigma_e(x')} dx' \quad (9.31)$$

Unfortunately, the latter two equations cannot be solved analytically since $\sigma_{e,h}$ are unknown functions of the position x in the absorber. In order to derive an approximate solution, we will assume the gradients of the quasi-Fermi levels to be constant and the charge carrier mobilities to be equal. Then $\Delta V_{\text{tr,h}}(x) + \Delta V_{\text{tr,e}}(x) = \text{const.} = V_{\text{tr}}$ and

$$V_{\text{tr}} = \frac{j_h(x')}{\sigma_h(x')} \int_0^x dx' + \frac{j_e(x')}{\sigma_e(x')} \int_x^d dx' = \left(\frac{j_h}{\sigma_h} + \frac{j_e}{\sigma_e} \right) d \quad (9.32)$$

It follows that the recombination rate is homogeneous, and for mobilities that are independent of carrier concentrations

$$\frac{j_h(x')}{n_h(x')} = \frac{j_e(x')}{n_e(x')} = \text{const.} \quad (9.33)$$

This means that we can replace these quantities by their spatial averages:

$$V_{\text{tr}} = \frac{d}{qb} \left(\frac{\bar{j}_h}{\bar{n}_h} + \frac{\bar{j}_e}{\bar{n}_e} \right) \quad (9.34)$$

If we use the fact that, per average, $j_e = j_h = 0.5j_Q$, we can thus write

$$V_{\text{tr}} = \frac{j_Q d}{2ebn_{\text{av}}} \quad (9.35)$$

and finally

$$n_{\text{av}}(j_Q) = \frac{d}{2ebR_{\text{tr}}(j_Q)} \quad (9.36)$$

Although the assumptions are not strictly fulfilled in real devices, very reasonable results are obtained from this analytical approach provided that surface recombination is negligible [53, 54].

9.7 Transient Techniques

9.7.1

Photovoltage Decay

In this method, a solar cell under open-circuit conditions is typically illuminated with a bias light which generates a constant photovoltage V_{OC} . Then the solar cell is subjected to an additional short light pulse that generates a transient voltage signal $\Delta V_{\text{OC}}(t)$. Its decay delivers information about the nature of the recombination process(es). The use of a bias light and the usually small amplitude $\Delta V_{\text{OC}} \ll V_{\text{OC}}$ ensures that the observed phenomena are relevant for the operation of the solar cell.

Different recombination mechanisms have different reaction orders β , that is, the corresponding recombination rates R depend in different ways on the concentrations of free electrons and holes, respectively:

$$R(n_{e,h}) \propto (n_e n_h)^{\beta/2} \quad (9.37)$$

For recombination via impurities, Shockley–Read–Hall or SRH recombination (see Section 3.6.2), the reaction order $\beta_{\text{SRH}} = 1$. For radiative recombination, $\beta_{\text{rad}} = 2$ and for Auger recombination $\beta_{\text{Aug}} = 3$. As a consequence, for a given V_{OC} and thus for given concentrations of electrons and holes, the transient decay of ΔV_{OC} shows different dynamics depending on which is the dominant

recombination mechanism. This often changes during a complete decay, for example, Auger recombination is usually dominant at high charge carrier densities whereas impurity recombination often governs the decay at low charge carrier concentrations.

Upon a variation of the bias light intensity, and thus of V_{OC} , the recombination dynamics can be investigated at different levels of the steady-state charge carrier concentrations.

$$\frac{dV_{OC}}{dt} \propto \frac{d}{dt}(\ln[n_e n_h]) \quad (9.38)$$

9.7.2

Transient Absorption

Transient absorption (TA) is in most cases performed in a so-called pump–probe experiment where the sample is optically excited with the pump beam. The generation of electron–hole pairs and subsequent processes cause changes in the absorption (or transmission) properties, which are observed with the probe beam. The pump beam is usually a short laser pulse at a wavelength at which the sample shows strong absorption. In order to obtain as much information as possible, the probe beam is some kind of white light continuum that passes through a monochromator and is split into two partial beams from which one hits the sample and the other does not. Then, the absorption (or transmission) of the sample can be derived at every wavelength by comparing the two partial beams. The usual experimental setups comprise a so-called delay stage, that is, a mechanism that allows varying the delay time between the pump and the probe beam. This way, the dynamics of the excitations can be observed. In the following, we will briefly discuss the different features of a TA measurement.

The TA signal is given by the quantity $\Delta A/A$, that is, the change in absorption caused by the pump beam. The pump beam generates electron–hole pairs that are – in most inorganic semiconductors – directly dissociated into free electrons and holes. As a consequence, the TA signal shows two dominant features. First, the fact that directly after the laser pulse there are more electrons in the conduction band and holes in the valence band leads to a reduction of the band-to-band absorption according to Equation 3.97, as absorption and hence transitions from valence to conduction band can occur only from occupied states in the valence band to unoccupied states in the conduction band. As the primary transition occurs from the ground state (corresponding to the valence band in inorganic crystalline semiconductors), this is also called *ground-state bleach*, shown schematically in Figure 9.14.

In addition, the TA signal is changed by the pump beam, as the electrons and holes generated by absorption of the laser beam $\Delta n_{e,h}$ do also absorb photons. This process is called *photoinduced absorption* and it promotes the charge carriers to higher energy states within the same band from where they quickly thermalize again (Figure 9.15). In the case of inorganic semiconductors, it is also called free carrier absorption, which can be used to record carrier density images.

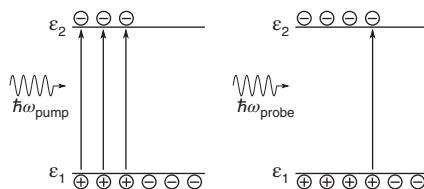


Figure 9.14 Visualization of the ground-state bleach. As the pump beam excites electrons from ε_1 to ε_2 , fewer electrons are available in ε_1 (and holes in ε_2) for subsequent excitations of the same type by the probe beam.

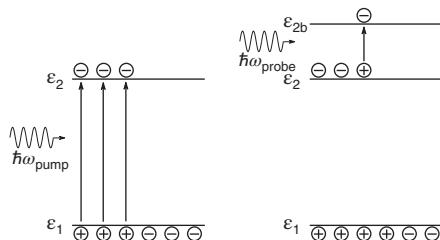


Figure 9.15 Photoinduced absorption. First, electrons are excited from ε_1 to ε_2 by the pump beam. Then they can be excited from ε_2 to higher states ε_{2b} within the conduction band (inorganic semiconductor) or to either higher vibronic states within the same electronic state or to a higher electronic state (molecules).

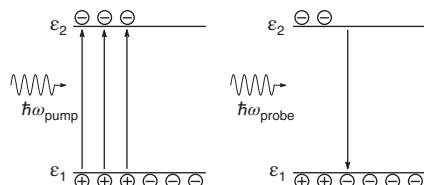


Figure 9.16 Stimulated emission. The incident probe beam can induce the emission of photons with the same phase, wavelength, polarization, and direction.

Whereas the TA signal for the band-to-band transitions is negative (the absorption becomes weaker after the pump beam), the photoinduced free carrier absorption produces a positive TA signal. In addition, another process that is observed is the stimulated emission, which refers to the emission of photons with the same phase, wavelength, polarization, and direction as the incident photons (Figure 9.16). As it increases the number of photons that are detected, it leads also to a negative TA signal.

Usually the measurement itself detects the wavelength-dependent transmission of the sample, from which the absorption can be derived according to $1 = R + A + T$. If the sample, for example, a complete solar cell, comprises a reflective back electrode, the TA is measured in reflectance.

In the case of an organic solar cell, the situation is more complex, as the electron–hole pairs initially generated by the pump beam do not directly form

free charge carriers but rather tightly bound excitons with typical binding energies $E_{\text{ex,B}} \gg k_B T$ due to the low dielectric permittivity in those materials. These excitons can dissociate at or near the interface between the donor and acceptor material (see Section 6.9) and thus form (quasi-)free electrons and holes in the acceptor and donor material, respectively. However, a certain fraction of the excitons may also recombine (either radiatively or nonradiatively) before dissociation. Hence, in addition to the signals observed in, for example, a crystalline silicon sample, in an organic solar cell the excitons can also contribute to the TA signal. Their (photoinduced) absorption usually has a different spectral dependence compared to the free charge carriers.

9.7.3

Charge Carrier Extraction

The concentration of free charge carriers, n_e and n_h , can in principle be derived from the value of the open-circuit voltage (Equations 3.30 and 7.14). However, there are cases where the shape of the density of states (as a function of energy) differs considerably from the square root approximation according to Equation 3.7. An example is the Gaussian distribution of states, which is often used to describe (organic) semiconductors with a high degree of spatial and energetic disorder [55]. In addition, disordered materials often also comprise a large number of trap states with energies in the forbidden gap. A useful information is therefore the actual density of charge carriers in the solar cell as a function of its voltage.

In a typical charge carrier extraction experiment, the solar cell is illuminated and kept under defined conditions, typically open-circuit conditions. Once it has reached the steady state, the light source is quickly switched off and the cell is – after a variable delay time t_d – subjected to reverse bias to extract the charge carriers. The number of charge carriers is derived by integrating the transient current $I(t)$. However, the current has two contributions: as the voltage is changed, that is, $dV/dt \neq 0$, there is a corresponding displacement current I_C . The geometric capacitance of the electrodes (with the photoactive material in between) has to be recharged:

$$I_C = dQ/dt = \frac{dV/dt}{C_{\text{geo}}} = \frac{dV/dt}{\epsilon_r \epsilon_0 A / d} \quad (9.39)$$

Here, ϵ_r is the relative dielectric permittivity and ϵ_0 the one of the vacuum, A is the electrode area, and d is the distance between the electrodes. In addition, there might be other contributions to the displacement current such as space charge regions (e.g., pn-junction). The displacement part of the transient current has to be subtracted before the quantitative determination of the amount of (photogenerated) charge carriers that had been extracted from the absorber can be carried out. Note that in the literature this technique is also referred to as *time-delayed collection field* or *TDCF* [56].

9.7.4

CELIV – Charge Extraction by Linearly Increasing Voltage

A typical voltage ramp in a CELIV measurement together with the current response of a cell is shown in Figure 9.17.

In CELIV, $dV/dt = \text{const.}$ and therefore the displacement current (j_0 in the lower graph) is constant as well. The extraction of charge carriers from the photoactive layer is the “bump” on top of the rectangle formed by the displacement current. The maximum extraction current is Δj , and the corresponding time is referred to as t_{\max} . Note that, in this example, there is no current flow during the pre-bias prior to the voltage ramp. This means that the cell is held under open-circuit condition, and the fact that $V_{\text{pre}} > 0$ means that it is illuminated.

It is obvious that the shape of the extracted current versus time is closely related to the charge carrier conductivities and thus to their mobilities. For this reason, there have been several attempts to derive a value for the charge carrier mobility from the quantities Δj and t_{\max} [57, 58]. A large degree of disorder and considerable amount of trap states, which are often present in organic solar cells, impose a large degree of uncertainty on the calculated values. For instance, Hanfland *et al.* could prove that the formulas cannot be applied in a reasonable manner if a considerable fraction of the charge carriers is trapped and thus rather immobile [59]. The reason is that the ratio between free and trapped charge carriers changes continuously during the voltage ramp and therefore the effective mobility as well.

9.7.5

Impedance Spectroscopy

As it is often not trivial to identify which factors limit the performance of a solar cell and to what extent, methods are required to distinguish between different processes occurring in the device. One method is impedance spectroscopy, which is typically used to detect electrochemical processes [60]. The technique is based on the fact that different processes occur in parallel in the solar cell under

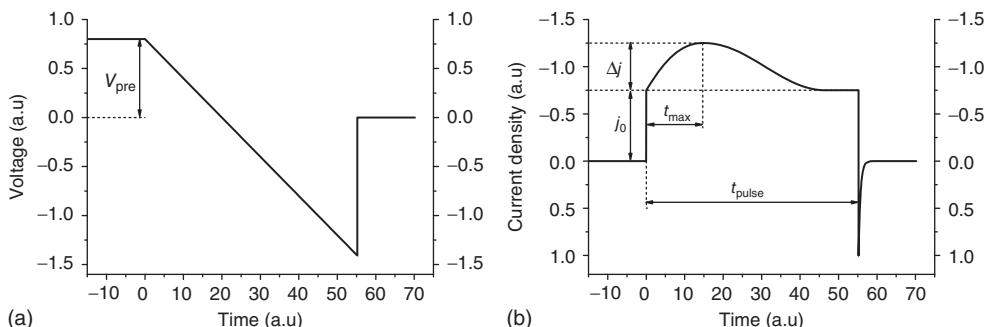


Figure 9.17 Scheme of a CELIV measurement. (a) The pre-bias V_{pre} and the subsequent voltage ramp in reverse direction. (b) The current response of the device.

steady-state conditions although they have different time dependencies. Again, to obtain meaningful results that are relevant for a solar cell under operation, the solar cell is kept under steady-state conditions, for example, under typical illumination and at open-circuit condition. Additionally, a small perturbation, usually a small amplitude AC voltage $\Delta V_{AC} = \Delta V e^{i\omega t}$, is applied, its frequency ω is varied over several orders of magnitude, and the current response of the device is measured. A lock-in amplifier (see Section 9.1) ensures that only the linear response of the device (i.e., the current that has the same frequency ω and not higher orders) is detected. The result is a (complex) resistance, which is called *impedance*:

$$Z(\omega) = \frac{\Delta V e^{i\omega t}}{\Delta I e^{i\omega t}} = \frac{\Delta V}{\Delta I} \quad (9.40)$$

Inductive elements are usually negligible. Hence the real part $Z_{re}(\omega)$ is related to the resistive elements of the system and the imaginary part $Z_{im}(\omega)$ to its capacitive elements. To further interpret the resultant impedance spectrum, the device is represented by an equivalent circuit, that is, a circuit comprising different resistors and capacitors. A typical element representing the interface of two materials is the parallel connection of a capacitor with a resistor (not necessarily ohmic), which has the following impedance:

$$Z(\omega) = \frac{R}{1 + (R\omega C)^2} - i \frac{R^2 \omega C}{1 + (R\omega C)^2} \quad (9.41)$$

At very low frequencies, the imaginary part can be neglected and the element behaves like a resistor representing the charge transfer from one material to another. In contrast, at very high frequencies, it behaves like a pure capacitance, as in this regime the time of a half period (i.e., where the voltage has a certain polarity) is too small in order for a complete charge transfer to happen. An example is the reduction of triiodide ions at the Pt back electrode of a dye solar cell (see Figure 6.6)



Here, the ions have to get rid of their solvation shell (i.e., solvent molecules that are weakly bound to ions in an electrolyte) before the charge transfer reaction, that is, the transfer of two electrons from the Pt can take place. Hence, for very high frequencies, only the double layer at the Pt electrode is charged according to the polarity of the AC voltage. At mid-frequencies, the charge transfer reaction can take place but the time it takes leads to a phase shift of the current.

Another example where impedance spectroscopy is used is the determination of the doping concentration via the capacitance of the space charge region that is formed when a p-doped (n-doped) semiconductor is in contact with a low (high) work function metal. The width of the space charge region and with it its capacitance depends on the doping concentration and on the applied voltage:

$$\frac{1}{C^2} = \frac{2(V_{bi} - V)}{\epsilon_r \epsilon_0 e n_{A,D}} \quad (9.43)$$

Thus, if the capacitance of the space charge region is determined with impedance spectroscopy at different DC voltages, a plot (the so-called Mott–Schottky plot) of the inverse of the square of the capacitance yields both the built-in voltage and the doping concentration.

Solutions

Chapter 1

- 1.1 The greenhouse effect is caused in both cases by a layer that functions as a selective absorber being transparent for the incoming sunlight but non-transparent for the thermal radiation from the Earth's surface. In (a), this is caused by the plastic or glass on the roof and in (b), by certain gas molecules (three-atomic molecules like CO₂, N₂O and others) in the atmosphere.
- 1.2 Burning methane leads to a higher amount of CO₂ because the (mass) amount of carbon dioxide m_{CO_2} produced corresponds to 2.75 m_{CH_4} in the case of methane and to 1.47 $m_{\text{CH}_2\text{O}}$ in the case of carbohydrates.
- 1.3 1 kg of coal equivalent corresponds to 8.2 kWh. On this amount of energy, the computer can be operated for 164 h or approximately 6.8 days.
- 1.4 For case (a) the function that describes the global human energy consumption is given by

$$\text{consumption}(t) = 15 \times 10^9 (1 + t/35a) \text{ t coal eq./a}$$

In case (b) it is

$$\text{consumption}(t) = 15 \times 10^9 \exp [t \ln(2)/40a] \text{ t coal eq./a}$$

In both cases,

$$\int_0^{t_{\text{end}}} \text{consumption}(t) dt = 2 \times 10^{12} \text{ t coal eq.}$$

has to be solved. This results in 272.4 years for case (a) and 183.7 years for case (b).

- 1.5 If all oxygen would have been produced by the reduction of CO₂ to methane, we have to consider the mass ratio $m_{\text{CH}_4} = \frac{16}{32} m_{\text{O}_2}$. The amount of methane would be $m_{\text{CH}_4} = 5 \times 10^{14} \text{ t}$.
- 1.6 Venus is located at 0.723 AU (astronomical units) from the Sun, that is, 0.723 times the average distance $d_{\text{Earth-Sun}}$ between Earth and Sun. Therefore, the intensity reaching the outside of Venus is given by:

$$j_{E, \text{Venus}} = j_{E, \text{Earth}} \left(\frac{d_{\text{Earth-Sun}}}{d_{\text{Venus-Sun}}} \right)^2 = 2.49 \text{ kW m}^{-2}$$

Owing to the high reflectance of $r_v = 0.7$ the intensity reaching the surface is $j_{E, \text{Venus}} = 0.746 \text{ kW m}^{-2}$.

The temperature at the surface of Venus as a black body without greenhouse effect would be

$$T_v = \sqrt[4]{\frac{j_{E, \text{Venus}}}{4 \sigma}} = 239.5 \text{ K} = -33.7^\circ\text{C}$$

In fact, the temperature at the surface of Venus is 475°C . Accordingly, the number of fire screens is

$$n = \left(\frac{T_{v, \text{greenhouse}}}{T_v} \right)^4 - 1 = 94.2$$

Chapter 2

- 2.1 A black body is characterized by the fact that its absorptance (and hence also its emittance) is $\alpha(\hbar\omega) \equiv 1$ for all photon energies $\hbar\omega$. It is best realized by a small opening in a large cavity.
- 2.2 According to Equation 2.27 the energy current density per solid angle from a black body is independent of the angle under which the black body is seen. The reduction of the received energy current caused by tilting the surface is thus compensated by an equal reduction of the apparent size of the surface area (i.e. the projection onto a plane perpendicular to the viewing direction). This phenomenon is observed as long as the emitting surface is resolved as an extended object. If the emitter is seen as a point source, only the variation of the energy current is registered but not the variation in apparent size of the emitting surface. For example, the different phases of Venus change in brightness when viewed with the naked eye, in contrast to observing it through a telescope, which resolves the bright and dark parts of the surface.
- 2.3 The number of states per volume for all photon energies $\leq \hbar\omega$ is given by Equation 2.5:

$$N_\gamma/V = \frac{(8\pi/3)(\hbar\omega)^3}{h^3(c_0/n)^3}$$

For $1.5 \text{ eV} \leq \hbar\omega \leq 3 \text{ eV}$ the number of states per volume is 10^{14} cm^{-3} for case (a) and $4.45 \times 10^{15} \text{ cm}^{-3}$ for case (b).

(c) From Equation 2.15 we know that the density of photons for *all* photon energies is

$$n_\gamma = \frac{(kT)^3}{4\pi^3 h^3 (c_0/n)^3} \underbrace{\int_0^\infty \frac{x^2 dx}{e^x - 1}}_{2.40411} 4\pi = \frac{2.40411 k^3}{\pi^2 h^3 (c_0/n)^3} T^3$$

Hence, in order to determine the density of photons for photon energies $1.5 \text{ eV} \leq \hbar\omega \leq 3 \text{ eV}$, we have to replace the limits of the integration by

$x_1 = 1.5 \text{ eV}/kT$ and $x_2 = 3 \text{ eV}/kT$, respectively:

$$n_\gamma = \frac{(kT)^3}{4\pi^3 \hbar^3 (c_0/n)^3} \int_{x_1}^{x_2} \frac{x^2 dx}{e^x - 1} 4\pi$$

As both the photon density n_γ and the number of photon states per volume N_γ/V depend on the refractive index n in the same way, the resulting temperature at which the total number of photons is larger than 10^{-4} of all photon states, does not depend on the particular material and is therefore the same for vacuum and for silicon.

For the photon energy range $1.5 \text{ eV} \leq \hbar\omega \leq 3 \text{ eV}$, this temperature is $T = 2582 \text{ K}$.

- 2.4** The Boltzmann distribution differs from the Bose–Einstein distribution by less than 10^{-2} for photon energies larger than
(a) $\hbar\omega = 0.119 \text{ eV}$ at $T = 300 \text{ K}$
(b) $\hbar\omega = 2.302 \text{ eV}$ at $T = 5800 \text{ K}$.

In both cases, this threshold photon energy is $\hbar\omega = 4.6 kT$. The Boltzmann distribution is thus a reasonable approximation of the Bose–Einstein distribution for $\hbar\omega \geq 4.6 kT$.

- 2.5** The reflectance for normal incidence is given by

$$r(\alpha = 0) = \left(\frac{n_{\text{air}}}{n_{\text{Si}}} \right)^2 = 30.9\%$$

The component polarized perpendicular to the plane of incidence is reflected according to

$$r_\perp(\alpha) = \left(\frac{\sin(\alpha - \beta)}{\sin(\alpha + \beta)} \right)^2$$

and the component polarized parallel to the plane of incidence according to

$$r_\parallel(\alpha) = \left(\frac{\tan(\alpha - \beta)}{\tan(\alpha + \beta)} \right)^2$$

with $\beta = \arcsin \left(\frac{n_{\text{air}}}{n_{\text{Si}}} \sin \alpha \right)$. Hence the reflectance is 31.1% in case (b) and 46.7% in case (c).

- 2.6** **(a)** On a March 21, at noon, the Sun is incident normal to the surface at the equator. To receive an intensity of 750 W m^{-2} , the angle between the surface normal of the solar cell and the incoming radiation has to be $\alpha = \arccos(750 \text{ W m}^{-2}/1000 \text{ W m}^{-2}) = 41.4^\circ$. With the roof tilted by $\beta = 25^\circ$ toward the south, that is, toward the direction of the incoming light, the latitude is $\gamma = \alpha + \beta = 66.4^\circ$.
(b) As the Sun moves in a plane called the ecliptic, guiding of the module around a single axis perpendicular to the ecliptic assures that the module is always perpendicular to the incident radiation. For a March 21, the axis of the guiding system is parallel to the axis of the Earth and therefore $90^\circ - 66.4^\circ = 33.6^\circ$ tilted against the surface normal at 66.4° latitude.

Without a tracking system, the incident intensity at noon is $1000 \text{ W m}^{-2} \times \cos(\alpha)$ and varies during the day according to $\cos[\delta(t)]$, where $\delta(t)$ is the so-called hour-angle, which (on a March 21) varies in 12 h from $-\pi/2$ at Sunrise to $\pi/2$ at sunset. The energy output W of the module is

$$W = 1 \text{ kW m}^{-2} \times 0.15 \cos(\alpha) 12 \text{ h} / \pi \int_{-\pi/2}^{\pi/2} \cos(\delta) d\delta = 0.86 \text{ kWh m}^{-2}$$

With the tracking system, the energy output for the whole day is simply given by $1000 \text{ W m}^{-2} \times 12 \text{ h} \times 0.15 = 1.8 \text{ kWh m}^{-2}$.

(To obtain an optimal output throughout the whole year, the guiding axis has to be adjusted according to the change of the inclination of the ecliptic with respect to the surface normal.)

- 2.7** For the mean photon energy of absorbed photons $\langle \hbar\omega_{\text{abs}} \rangle$, we have to divide the absorbed energy current density $j_{E,\text{abs}}$ in Equation 2.39 by the absorbed photon current density $j_{\gamma,\text{abs}}$ in Equation 2.40. $\varepsilon_G \gg kT$ allows to approximate these expressions by considering only the first element in the sum ($i = 1$).

$$\langle \hbar\omega_{\text{abs}} \rangle = kT \frac{x_G^3 + 3x_G^2 + 6x_G + 6}{x_G^2 + 2x_G + 2} = kT(x_G + 1 + 2/x_G + O(x_G^{-3}))$$

where $x_G = \varepsilon_G/kT$. Thus, $\langle \hbar\omega_{\text{abs}} \rangle \approx \varepsilon_G + kT$.

The condition $\varepsilon_G \gg kT$ is very well fulfilled for a typical semiconductor like Si or GaAs in equilibrium with the 300 K radiation from the environment. The reader may wonder how the absorbed photons with a mean energy of $\varepsilon_G + kT$ can be in equilibrium with electron–hole pairs with a mean energy of $\varepsilon_G + 3kT$, as derived in Equations 3.20 and 3.24.

Chapter 3

- 3.1** Doping with donors increases the concentration of electrons in the conduction band. A donor atom has a valency that is larger than the valency of the lattice atoms. It is important that the energy level (of one) of the electron(s) not needed for the chemical bond is close to the conduction band. An analogous argument applies for acceptors and the concentration of holes. Doping increases the conductivity, however, it reduces the mobility by introducing coulomb scattering by charged centers (as deviation from the periodic potential of the lattice ions).
- 3.2** Since the generation rate of electron–hole pairs by the 300 K radiation from the environment is not changed by doping, the equilibrium recombination rate cannot change either. As this radiative recombination rate is proportional to the product $n_e n_h$, an increase in n_e leads to a decrease in n_h .
- 3.3** The condition is that the thermalization time is much shorter than the lifetime, such that both the electrons in the conduction band and the holes in the valence band form ideal gases with a defined temperature.

- 3.4** Consider an impurity with a capture cross section much larger for electrons than for holes, $\sigma_e \gg \sigma_h$. In steady state, from Equation 3.66, we get

$$G_{e,\text{imp}} = R_{e,\text{imp}} - R_{h,\text{imp}} + G_{h,\text{imp}}$$

Because $\sigma_e \gg \sigma_h$, both exchange rates with the conduction band are much larger than the exchange rates with the valence band: $G_{e,\text{imp}}, R_{e,\text{imp}} \gg R_{h,\text{imp}}, G_{h,\text{imp}}$.

In this case, the occupation of the impurity is determined by the quasi-Fermi distribution for the conduction band (ε_{FC}), although the impurity level may lie in between ε_{FC} and ε_{FV} . For this reason, demarcation levels are introduced, which define the energy range in which the occupation of impurity levels is determined by the kinetics.

- 3.5** The electrical conductivity as resulting from Equations 5.3 and 5.4 is

$$\sigma = \sigma_e + \sigma_h = e n_e b_e + e n_h b_h$$

In an intrinsic semiconductor, the increase of the densities of free charge carriers, that is, electrons in the conduction band and holes in the valence band, rises exponentially with increasing temperature, according to Equations 3.12 and 3.15 (the dependence of the effective density of states on temperature is a minor correction). In a metal, there are only electrons and their concentration does not depend on temperature. On the other hand, the mobility is reduced with increasing temperature due to phonon scattering of the charge carriers. In a semiconductor ($\langle \varepsilon_{\text{kin}} \rangle = 3/2 kT$), the mean thermal velocity $\langle v_{\text{th}} \rangle$ rises with \sqrt{T} . The concentration of phonons n_Γ , by which the charge carriers are scattered depends on temperature according to

$$\begin{aligned} n_\Gamma &= N_\Gamma \left(\exp[\hbar\Omega/kT] - 1 \right)^{-1} \stackrel{kT \gg \hbar\Omega}{\approx} N_\Gamma (1 + \hbar\Omega/kT - 1)^{-1} \\ &\approx \frac{N_\Gamma kT}{\hbar\Omega} \end{aligned}$$

N_Γ is the density of states of the phonons and $\hbar\Omega$ is their energy.

The collision time τ_c is proportional to $1/(n_\Gamma \langle v_{\text{th}} \rangle)$ and thus the mobility in a semiconductor varies as $b_{e,h} \sim T^{-3/2}$.

If a semiconductor is doped, the temperature dependence of its conductivity exhibits three distinct temperature regimes. Figure 3.12 shows how the charge carrier densities depend on temperature. For temperatures relevant for solar cells (around room temperature), the concentration of the majority carriers is independent of temperature. The conductivity thus reflects the temperature dependence of their mobility.

The overall temperature dependence of the conductivity of an intrinsic semiconductor is, however, dominated by the exponential dependence of the charge carrier densities.

In a metal, $\langle v_{\text{th}} \rangle$ is independent of temperature, thus $b \sim T^{-1}$. Since the electron concentration is independent of temperature, the conductivity of a metal is proportional to T^{-1} as well.

- 3.6** An unoccupied conduction band would definitely minimize the energy E , but the quantity to minimize is the free energy $F = E - TS$. Since the entropy S is zero for an unoccupied conduction band and rises, for small concentrations, more strongly with increasing electron concentration than the energy, the minimum of F is found for a partially occupied conduction band. This topic is analyzed quantitatively in Problem 3.10.
- 3.7** As was derived in Equation 3.7 the density of states for a three-dimensional electron gas is

$$D_{e,3\text{-dim}}(\varepsilon_e) = 4\pi \left(\frac{2m_e^*}{h^2} \right)^{3/2} (\varepsilon_e - \varepsilon_C)^{1/2}$$

It rises with the square root of the (kinetic) energy of the electrons. In the one-dimensional case for a chain of length L , $\Delta p = h/L$ and the number of states with momentum $|p'| \leq |p|$ is $N(|p|) = 4|p|L/h$. This leads to

$$D_{e,1\text{-dim}}(\varepsilon_e) = 2 \left(\frac{2m_e^*}{h^2} \right)^{1/2} (\varepsilon_e - \varepsilon_C)^{-1/2}$$

In the two-dimensional case for an area of L^2 , we find $N(|p|) = 2\pi p^2 L^2 / h^2$. The density of states is

$$D_{e,2\text{-dim}}(\varepsilon_e) = 2\pi \left(\frac{2m_e^*}{h^2} \right)$$

- 3.8** The absorption coefficient $\alpha(\hbar\omega) \sim (\hbar\omega - \varepsilon_G)^{1/2}$ if the combined density of states varies as $D_{\text{comb}}(\hbar\omega) \sim \sqrt{\varepsilon_{\text{kin}}}$ and the transitions are direct.
- 3.9** The Boltzmann distribution differs from the Fermi–Dirac distribution by no more than 10^{-2} if the quasi-Fermi energy is at least $4.6 kT$ lower than the conduction band edge for all temperatures.
Compare with Problem 2.4.
- 3.10 (a)** We know that electrons and holes in a nondegenerate semiconductor form ideal gases with mean energies of $\langle \varepsilon_e \rangle = \varepsilon_C + 3/2 kT$ and $\langle \varepsilon_h \rangle = -\varepsilon_V + 3/2 kT$. For this reason, we can say that, starting from $F^* = 0$ (empty conduction band), the energy E per volume rises by $\varepsilon_C - \varepsilon_V + 3kT = \varepsilon_G + 3kT$ for every electron “brought” into the conduction band, leaving behind a hole in the valence band. The entropy per volume is calculated from the Sackur–Tetrode equation (3.35), so the free energy per volume is

$$\begin{aligned} F^*(n_{e,h}) &= \frac{3kT + \varepsilon_G}{2}(n_e + n_h) \\ &\quad - kT \left\{ n_e \left[\frac{5}{2} + \ln \left(\frac{N_C}{n_e} \right) \right] + n_h \left[\frac{5}{2} + \ln \left(\frac{N_V}{n_h} \right) \right] \right\} \end{aligned}$$

For an intrinsic semiconductor, $n_e = n_h = n_i$, which leads to

$$\begin{aligned} F^*(n_i) &= n_i \left\{ \varepsilon_G + 3kT - kT \left[5 + \ln \left(\frac{N_C}{n_i} \right) + \ln \left(\frac{N_V}{n_i} \right) \right] \right\} \\ &= n_i \left\{ \varepsilon_G - kT \left[2 + \ln \left(\frac{N_C}{n_i} \right) + \ln \left(\frac{N_V}{n_i} \right) \right] \right\} \\ &= n_i \left\{ \varepsilon_G - kT \left[2 + \ln \left(\frac{N_C N_V}{n_i^2} \right) \right] \right\} \end{aligned}$$

To find the minimum of F^* as a function of the electron (hole) concentration, we set the derivative to zero

$$\begin{aligned} 0 &= \frac{\partial F^*(n_i)}{\partial n_i} = \varepsilon_G - kT \left[2 + \ln \left(\frac{N_C N_V}{n_i^2} \right) \right] - n_i kT \left(-\frac{2}{n_i} \right) \\ &= \varepsilon_G - kT \ln \left(\frac{N_C N_V}{n_i^2} \right) \Rightarrow n_i^2 = N_C N_V \exp \left(-\frac{\varepsilon_G}{kT} \right) \end{aligned}$$

As expected, this result is exactly the same as the one for n_i^2 in Equation 3.17 and reproduces, therefore, the result obtained by Fermi–Dirac (or Boltzmann) statistics.

Inserting the result for n_i obtained above into the equation for F^* leads to a very simple expression (refer to Equation 3.17):

$$F^*(n_i) = n_i \left\{ \varepsilon_G - kT \left[2 + \ln \left(\frac{N_C N_V}{n_i^2} \right) \right] \right\} = -2n_i kT$$

For the given parameters the free energy per volume reaches its minimum $F^* = -3.52 \times 10^8 \text{ eV cm}^{-3}$ for an electron (hole) concentration of $n_{e,h} = 6.798 \times 10^9 \text{ cm}^{-3}$.

- (b) Here, we have to regard the electrons that are transferred from states with an electron energy of $\varepsilon_{e,1} = \varepsilon_V - \varepsilon_0$ to states with an energy of $\varepsilon_{e,2} = \varepsilon_C + \varepsilon_0$. We then have to integrate over all these states, i.e. using the approximation made in Equation 3.12 from $\varepsilon_0 = 0$ to $\varepsilon_0 = \infty$.

This is equivalent to the situation with many narrow-band semiconductors with different values of ε_G , but all with the same Fermi energy ε_F . For every semiconductor, we have to find the minimum for F^* as a function of n_e and n_h and sum over all these contributions to get the total value for F^* .

From statistics, we know that the number of possibilities Ω to distribute n undistinguishable electrons (or holes) in N states is given by

$$\begin{aligned} \Omega &= \binom{n}{N} = \frac{N!}{n!(N-n)!} \\ &\approx N \ln(N) - n \ln(n) - (N-n) \ln(N-n) \end{aligned}$$

The so-called Stirling approximation is a good approximation for large numbers N, n .

For the following calculation, we divide the bands into intervals $\Delta\varepsilon$ and determine the number of states per volume ΔN_l for intervals centered at an energy

$$\varepsilon_{e,l} = \varepsilon_C + (l + 1/2) \Delta\varepsilon$$

and

$$\varepsilon_{h,l} = -\varepsilon_V + (l + 1/2) \Delta\varepsilon$$

respectively, with

$$\Delta N_{e,h,l}(\varepsilon_{e,h,l}) = D_{e,h}(\varepsilon_{e,h,l}) \Delta\varepsilon$$

$N_C = 3 \times 10^{19} \text{ cm}^{-3}$ and $N_V = 1 \times 10^{19} \text{ cm}^{-3}$ correspond to effective masses of $m_e^* = 1.1264 m_e$ and $m_h^* = 0.5415 m_h$, respectively. These determine $D_{e,h}(\varepsilon_{e,h,l})$.

We obtain the free energy per volume for every $\Delta\varepsilon$ centered at ε_l

$$\begin{aligned} \Delta F_l^* &= \Delta E_l - T(\Delta S_{e,l} + \Delta S_{h,l}) = \Delta n_l [\varepsilon_G + 2(l + 1/2) \Delta\varepsilon] \\ &\quad - kT[\Delta N_{l,e} \ln(\Delta N_{e,l}) - \Delta n_l \ln(\Delta n_l)] \\ &\quad - (\Delta N_{e,l} - \Delta n_l) \ln(\Delta N_{e,l} - \Delta n_l)] \\ &\quad - kT[\Delta N_{l,h} \ln(\Delta N_{h,l}) - \Delta n_l \ln(\Delta n_l)] \\ &\quad - (\Delta N_{h,l} - \Delta n_l) \ln(\Delta N_{h,l} - \Delta n_l)] \end{aligned}$$

We again made use of $\Delta n_{e,l} = \Delta n_{h,l} = \Delta n_l$, being valid for an intrinsic semiconductor also in the case of $m_e^* \neq m_h^*$.

After having determined the minimum of $\Delta F_l^*(\Delta n_l)$, we have to sum over all l to get the correct results for F^* and for $n_e = n_h = n_i$.

$$F^* = \sum_{l=0}^{\infty} \Delta F_l^* \quad \text{and} \quad n_e = \sum_{l=0}^{\infty} \Delta n_l$$

Again, we can determine the minimum of F^* analytically:

$$\begin{aligned} 0 &= \frac{\partial \Delta F^*(\Delta n_l)}{\partial \Delta n_l} = \varepsilon_C - \varepsilon_V + (2l + 1) \Delta\varepsilon \\ &\quad - kT[-\ln(\Delta n_l) - 1 + \ln(\Delta N_{e,l} - \Delta n_l) + 1] \\ &\quad - \ln(\Delta n_l) - 1 + \ln(\Delta N_{h,l} - \Delta n_l) + 1 = \varepsilon_G + (2l + 1) \Delta\varepsilon \\ &\quad - kT \left[\ln \left(\frac{\Delta N_{e,l} - \Delta n_l}{\Delta n_l} \right) + \ln \left(\frac{\Delta N_{h,l} - \Delta n_l}{\Delta n_l} \right) \right] \end{aligned}$$

- For $\Delta N_{e,l} = \Delta N_{h,l} = \Delta N_l$ ($m_e^* = m_h^*$) follows

$$\begin{aligned} 0 &= \frac{\partial \Delta F^*(\Delta n_l)}{\partial \Delta n_l} \\ &= \varepsilon_G + (2l + 1) \Delta\varepsilon - 2kT \left[\ln \left(\frac{\Delta N_l - \Delta n_l}{\Delta n_l} \right) \right] \end{aligned}$$

and finally

$$\begin{aligned}\Delta n_l &= \frac{\Delta N_l}{\exp [(\varepsilon_G + (2l+1)\Delta\varepsilon)/2kT] + 1} \\ &= \frac{\Delta N_l}{\exp [(\varepsilon_C + (l+1/2)\Delta\varepsilon - \varepsilon_F)/kT] + 1}\end{aligned}$$

from which follows that $\varepsilon_F = (\varepsilon_C + \varepsilon_V)/2$, as expected. This is exactly the Fermi–Dirac formula for the electron concentration Δn_l at the energy $\varepsilon_{e,l} = \varepsilon_C + (l+1/2)\Delta\varepsilon$.

- For $\Delta N_{e,l} \neq \Delta N_{h,l}$ ($m_e^* \neq m_h^*$)

$$\Delta n_l = \Delta N_{e,l} \left(\exp [(\varepsilon_{e,l} - \varepsilon_F)/kT] + 1 \right)^{-1}$$

and

$$\Delta n_l = \Delta N_{h,l} \left(\exp [(\varepsilon_F + \varepsilon_{h,l})/kT] + 1 \right)^{-1}$$

it follows that

$$\varepsilon_F = \frac{1}{2}(\varepsilon_{e,l} - \varepsilon_{h,l}) + \frac{1}{2}kT \ln \left(\frac{\Delta N_{h,l} - \Delta n_l}{\Delta N_{e,l} - \Delta n_l} \right)$$

With the Boltzmann approximation, i.e. $\Delta N_{e,h,l} - \Delta n_l \approx \Delta N_{e,h,l}$ we get the result of Equation 3.18

$$\varepsilon_F = \frac{1}{2}(\varepsilon_{e,l} - \varepsilon_{h,l}) + \frac{1}{2}kT \ln \left(\frac{\Delta N_{h,l}}{\Delta N_{e,l}} \right)$$

Using the same approximation for the conditional equation for the minimum of F^* , we get

$$\begin{aligned}0 &= \frac{\partial \Delta F^*(\Delta n_l)}{\partial \Delta n_l} = \varepsilon_G + (2l+1)\Delta\varepsilon \\ &\quad - kT \left[\ln \left(\frac{\Delta N_{e,l} - \Delta n_l}{\Delta n_l} \right) + \ln \left(\frac{\Delta N_{h,l} - \Delta n_l}{\Delta n_l} \right) \right] \\ &\approx \varepsilon_G + (2l+1)\Delta\varepsilon - kT \left[\ln \left(\frac{\Delta N_{e,l} \Delta N_{h,l}}{\Delta n_l^2} \right) \right]\end{aligned}$$

Again, we see that Δn_l equals the intrinsic carrier concentration $\Delta n_{i,l}$

$$\Delta n_l = \sqrt{\Delta N_{e,l} \Delta N_{h,l}} \exp \left[-\frac{\varepsilon_G + (2l+1)\Delta\varepsilon}{2kT} \right] = \Delta n_{i,l}$$

With the above equation for the Fermi energy in Boltzmann approximation, this can be rewritten as

$$\begin{aligned}\Delta n_l &= \Delta N_{e,l} \left(\exp \left[\frac{\varepsilon_C + (l + 1/2) \Delta \varepsilon - \varepsilon_C/2 - \varepsilon_V/2}{kT} \right] \sqrt{\Delta N_{e,l}/\Delta N_{h,l}} \right)^{-1} \\ &= \Delta N_{e,l} \left(\exp \left[\frac{\varepsilon_C + (l + 1/2) \Delta \varepsilon - \varepsilon_C/2 - \varepsilon_V/2 - kT/2 \ln(\Delta N_{h,l}/\Delta N_{e,l})}{kT} \right] \right)^{-1} \\ &= \Delta N_{e,l} \exp [-(\varepsilon_{e,l} - \varepsilon_F)/kT]\end{aligned}$$

Without the Boltzmann approximation, the direct determination of the Fermi energy ε_F from the minimum of the free energy is a little complicated. Instead, we assume that we know the Fermi distribution for the holes, which gives

$$\varepsilon_{h,l} + \varepsilon_F = kT \ln \left(\frac{\Delta N_{h,l} - \Delta n_l}{\Delta n_l} \right)$$

we insert this into

$$\begin{aligned}0 &= \frac{\partial \Delta F^*(\Delta n_l)}{\partial \Delta n_l} \\ &= \varepsilon_{e,l} + \varepsilon_{h,l} - kT \left[\ln \left(\frac{\Delta N_{e,l} - \Delta n_l}{\Delta n_l} \right) + \ln \left(\frac{\Delta N_{h,l} - \Delta n_l}{\Delta n_l} \right) \right] \\ &= \varepsilon_{e,l} - \varepsilon_F - kT \ln \left(\frac{\Delta N_{e,l} - \Delta n_l}{\Delta n_l} \right)\end{aligned}$$

and find the exact Fermi distribution for the electrons

$$\Delta n_l = \frac{\Delta N_{e,l}}{\exp [(\varepsilon_C + (l + 1/2) \Delta \varepsilon - \varepsilon_F)/kT] + 1}$$

Figure S.1a shows the free energy per volume F^* as a function of the electron (hole) concentration of intrinsic silicon. The entropy is given by Sackur–Tetrode. The minimum of F^* determines the state of chemical equilibrium between the electrons and the holes.

Figure S.1b shows the sum over the contributions from each energy interval up to an energy $\varepsilon_{e,l}$ to the total free energy F^* and to the total electron (hole) concentration in the state of chemical equilibrium defined by a minimum of F^* . One sees that for energies $\varepsilon_{e,l} > \varepsilon_C + 0.25 \text{ eV}$ (and $\varepsilon_{h,l} > -\varepsilon_V + 0.25 \text{ eV}$) these contributions are negligibly small. The result in (b), involving the explicit calculation of the entropy, agrees well with the result in (a), using the entropy from the Sackur–Tetrode equation.

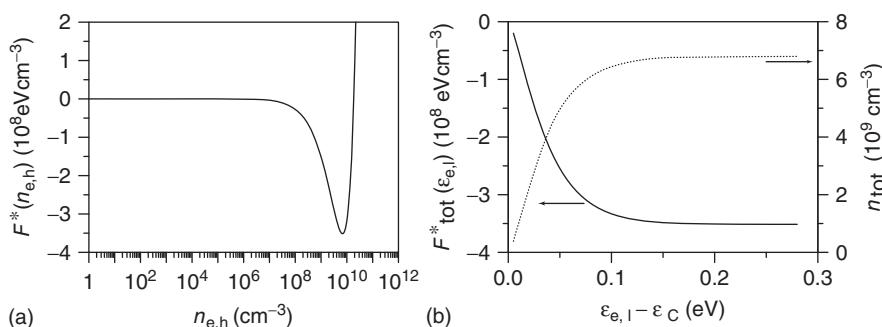


Figure S.1 (a) Free energy per volume F^* as a function of the electron (hole) concentration $n_{e,h}$ as a result of the Sackur–Tetrode equation (3.35). (b) F^* (solid line) and n_e (dotted line) as the sum of the contributions of all states with energies $\epsilon_e \leq \epsilon_{e,I}$

as a function of $\epsilon_{e,I}$. In graph (a), effective densities of states $N_C = 3 \times 10^{19} \text{ cm}^{-3}$ and $N_V = 1 \times 10^{19} \text{ cm}^{-3}$ were assumed and in graph (b) the corresponding effective masses $m_e^* = 1.1264 m_e$ and $m_h^* = 0.5415 m_e$. For both, the temperature is $T = 300 \text{ K}$.

3.11 The binding energy ϵ_{bind} for the ground and the excited states of an electron in a donor atom using the hydrogen approximation is

$$\epsilon_{\text{bind}}(n) = \frac{m_e^* e^4}{2 (4\pi\epsilon_0\epsilon)^2 \hbar^2} \frac{1}{n^2}$$

with n being the quantum number for the state, i.e. $n = 1$ for the ground state and $n = 2, 3 \dots$ for the excited states. For $\epsilon(\text{Si}) = 11.9$ we get

(a) $\epsilon_{\text{bind}}(1) = 0.104 \text{ eV}$, $\epsilon_{\text{bind}}(2) = 0.026 \text{ eV}$, $\epsilon_{\text{bind}}(3) = 0.012 \text{ eV}$.

(b) $\epsilon_{\text{bind}}(1) = 1.9 \times 10^{-3} \text{ eV}$, $\epsilon_{\text{bind}}(2) = 4.8 \times 10^{-4} \text{ eV}$, $\epsilon_{\text{bind}}(3) = 2.1 \times 10^{-4} \text{ eV}$.

(c) The Bohr radius r_B for the ground state ($n = 1$) is given by

$$r_B(1) = \epsilon_0 \epsilon_{\text{Si}} \frac{h^2}{\pi m_e^* e^2}$$

For case (a), it is $r_B(1) \approx 0.6 \text{ nm}$ and for case (b), $r_B(1) \approx 31.5 \text{ nm}$. For a mass density of crystalline Si of $\rho_{\text{Si}} = 2.3 \text{ g cm}^{-3}$ and a molar mass of $M_{\text{Si}} = 28.1 \text{ g mol}^{-1}$, there are about 4.95×10^{22} atoms in 1 cm^{-3} . It follows that in a spherical volume with the Bohr radius of case (a) calculated above, there are about 32 atoms over which the electron is then smeared out. In case (b), the number of atoms is 5.1×10^6 (!).

3.12 (a) Full ionization of the donors means $n_e = n_D$ and therefore $\epsilon_C - \epsilon_F = kT \ln \left(\frac{N_C}{n_e} \right) = 0.208 \text{ eV}$.

(b) We have three equations for the three unknown quantities n_e , n_{D^+} , and ϵ_F . Two express the dependence of n_e and n_{D^+} on ϵ_F and a third one results from the fact that the semiconductor remains electrically neutral upon doping ($e[n_D^+ + n_h - n_e] = 0$), see Equation 3.26 and following.

The results are $\epsilon_C - \epsilon_F = 0.218537 \text{ eV}$ using the Boltzmann approximation for n_e and $\epsilon_C - \epsilon_F = 0.218533 \text{ eV}$ with Fermi–Dirac.

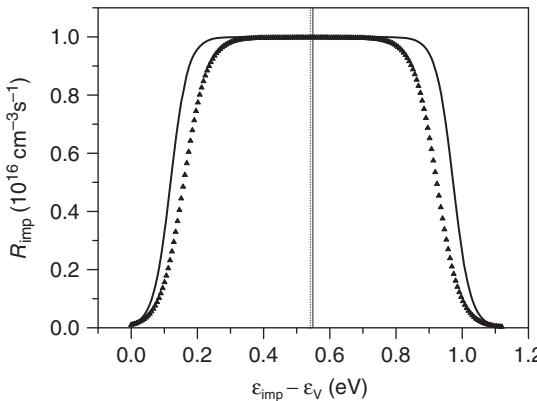


Figure S.2 Recombination rates R_{imp} from impurity states as a function of their energetic position for $T = 300 \text{ K}$ (solid line) and $T = 400 \text{ K}$ (triangles). The vertical solid line indicates the position of the intrinsic Fermi level with respect to the valence

band maximum $\varepsilon_i - \varepsilon_V$ for $T = 300 \text{ K}$, whereas the vertical dotted line represents $\varepsilon_i - \varepsilon_V$ for $T = 400 \text{ K}$. The difference in ε_i for the two temperatures is rather small: $\varepsilon_i(300 \text{ K}) - \varepsilon_V = 0.5458 \text{ eV}$; $\varepsilon_i(400 \text{ K}) - \varepsilon_V = 0.5411 \text{ eV}$.

- (c) The temperature is $T(n_h \geq 0.5 n_e) \approx 639.8^\circ\text{C}$. At that temperature, 50.2% of the electrons in the conduction band come from ionized donor atoms.
- 3.13** The recombination rates are calculated using Equation 3.77. The recombination rates as a function of the impurity level for temperatures $T = 300 \text{ K}$ and $T = 400 \text{ K}$ are shown in Figure S.2.

Chapter 4

- 4.1** The efficiency η is given by $dj_\mu/dj_{E,\text{abs}}$, that is, the ratio of the extracted chemical energy current density $dj_\mu = dj_{eh}(\mu_e + \mu_h)$ and the absorbed energy current density $dj_{\gamma,\text{abs}} \hbar\omega$. Using Equation 4.10 we get

$$\eta = \left[dj_{\gamma,\text{abs}} - dj_\gamma^0 \exp\left(\frac{\mu_e + \mu_h}{kT}\right) \right] (\mu_e + \mu_h) / (dj_{\gamma,\text{abs}} \hbar\omega)$$

As the first summand $dj_{\gamma,\text{abs}}$ does not depend on $(\mu_e + \mu_h)$, its contribution to the efficiency is linear in $(\mu_e + \mu_h)$. This linear dependence dominates the behavior of η as long as the radiative recombination rate $dj_\gamma^0 \exp\left(\frac{\mu_e + \mu_h}{kT}\right)$ is small.

- 4.2** For the change of the quasi-Fermi energy of the electrons ε_{FC} caused by illumination under open-circuit conditions, we first have to determine the absorbed photon current density

$$j_{\gamma,\text{abs}} = \frac{\Omega}{4\pi^3 \hbar^3 c^2} \int_{\varepsilon_G}^{\infty} \frac{(\hbar\omega)^2}{\exp(\hbar\omega/kT_{\text{bb}}) - 1} d\hbar\omega$$

The mean generation rate is $\bar{G} = j_{\gamma, \text{abs}}/d$. For $d = 300 \mu\text{m}$, we find $\bar{G} = 1.13 \times 10^{19} \text{ cm}^{-3} \text{ s}^{-1}$.

- (a) From $\bar{G} = R_{\text{rad}} = B n_h n_e$ with $B = 3 \times 10^{-15} \text{ cm}^3 \text{ s}^{-1}$ and $n_h = n_h^0 + n_e$ we find an electron concentration of $n_e = 4.13 \times 10^{16} \text{ cm}^{-3}$. By using $\epsilon_C - \epsilon_{\text{FC}} = kT_0 \ln \left(\frac{n_e}{N_C} \right)$ we find that the quasi-Fermi energy ϵ_{FC} lies 0.170 eV below the conduction band edge.
- (b) From $\bar{G} = R_{\text{aug}} = C n_h n_e (n_h + n_e)$ with $C = 10^{-30} \text{ cm}^6 \text{ s}^{-1}$ and $n_h = n_h^0 + n_e$ we find an electron concentration of $n_e = 3.67 \times 10^{15} \text{ cm}^{-3}$. Accordingly, the quasi Fermi-energy ϵ_{FC} lies 0.233 eV below the conduction band edge.

- 4.3 To calculate the emitted photon current density (into an effective solid angle of π) as a function of the chemical potential of the emitted photons (which equals the chemical potential of the electron–hole pairs), we have to make use of Equation 3.102. Since for the problem, this equation has to be expressed in terms of the wavelength λ , we need the relations in Equation 2.10 $\hbar\omega = hc/\lambda$ and $d\hbar\omega = -hc/\lambda^2 d\lambda$. The transformed equation is

$$\frac{dj_\gamma}{d\lambda} = \frac{2c\Omega}{\lambda^4} \frac{1}{\exp[(hc/\lambda - \mu_{e,h})/kT] - 1}$$

- (a) For $\mu_{eh} = 0.5 \text{ eV}$, the emitted photon current density per wavelength $dj_\gamma/d\lambda = 1.64 \times 10^{-12} \text{ cm}^{-2} \text{ nm}^{-1}$.
- (b) For $\mu_{eh} = 1.8 \text{ eV}$, $dj_\gamma/d\lambda = 1.14 \times 10^{10} \text{ cm}^{-2} \text{ nm}^{-1}$.

Chapter 5

- 5.1 We start from $j_{Q,i} = -\frac{\sigma_i}{z_i e} \text{ grad } \eta_i$, which for pure diffusive transport becomes

$$j_{Q,i} = -\frac{\sigma_i}{z_i e} \text{ grad } \mu_i = -\frac{\sigma_i}{z_i e} \text{ grad } kT \ln(n_i/n_0) = -\frac{\sigma_i}{z_i e} kT \frac{1}{n_i} \text{ grad } n_i$$

Using $\sigma_i = z_i^2 e n_i b_i$ from Section 5.1.1, we get

$$j_{Q,i} = -z_i b_i kT \text{ grad } n_i$$

As we know, this result is equivalent to Fick's law of diffusion, hence

$$-z_i b_i kT \text{ grad } n_i \equiv -D_i \text{ grad } n_i$$

which finally leads to

$$D_i = \frac{b_i kT}{e}$$

- 5.2 (a) It is $\sigma_h = z_h^2 e n_h b_h = e \cdot 10^{16} \text{ cm}^{-3} 300 \text{ cm}^2 (\text{Vs})^{-1} = 0.48 (\Omega \text{ cm})^{-1}$.
- (b) With $E = j_Q/\sigma_h = 0.312 \text{ V cm}^{-1}$, it follows that the electric potential difference between the two ends of the semiconductor is $\Delta\varphi = 6.24 \text{ mV}$.

- (c) If the current is caused by diffusion, the gradient of the hole density has to be constant throughout the sample, due to $\operatorname{div} j_Q = 0$. We have $D_h = b_h kT/e = 7.76 \text{ cm}^2 \text{s}^{-1}$. $j_Q = -eD_h \operatorname{grad} n_h$ leads to $\operatorname{grad} n_h = 1.2 \times 10^{17} \text{ cm}^{-4}$. The linear profile of the hole concentration is then given by

$$n_h(x) \approx (1.12 \times 10^{16} - 2.4 \times 10^{15} x/d) \text{ cm}^{-3}$$

The gradient of the electrochemical potential is independent of x only in case (a), where it trivially results in $\operatorname{grad} \eta_h = eE = 312 \text{ meV cm}^{-1}$. In case (b), we have to consider that the conductivity is a function of x and thus $\operatorname{grad} \eta_h$ as well:

$$\operatorname{grad} \eta_h = \operatorname{grad} \mu_h = \frac{kT}{n_h} \frac{\partial n_h}{\partial x} \approx (278.5 + 3.8 x/d) \text{ meV cm}^{-1}$$

The difference in chemical potential between $x = 0$ and $x = d$ is

$$\Delta\eta = \Delta\mu = kT \ln \frac{n_h(0)}{n_h(d)} = 6.27 \text{ meV}$$

- 5.3** The mean drift velocity of the electrons is given by $\langle v_{e,\text{drift}} \rangle = b_e(T)E$ and the mean thermal velocity by $\langle v_{e,\text{th}} \rangle = \sqrt{3kT/m_e}$. The temperature for which $\langle v_{e,\text{drift}} \rangle$ is 0.001 $\langle v_{e,\text{th}} \rangle$ is $T = 481 \text{ K}$.
- 5.4** The mobility of a compensated semiconductor is smaller than of a pure one because the electrons and holes collide with the impurities. Thus, the mean time between two collisions, $\tau_{c,i}$, is reduced, and, with it, the mobility b_i , see Equation 5.2. This effect is most pronounced if the impurities are charged.

Chapter 6

- 6.1** An improvement in the lifetimes of electrons and holes by a factor of 4 leads to a doubling of their diffusion lengths. In addition, the recombination rates are lowered by a factor of 4. As a result, the reverse saturation current density j_S is reduced by a factor of 2, see Equation 6.28.
- 6.2** It is

$$w = \sqrt{\frac{2 \epsilon \epsilon_0}{e} \frac{n_A + n_D}{n_A n_D} (\varphi^n - \varphi^p - V_a)}$$

with V_a being the applied voltage. The diffusion voltage is calculated with Equation 6.10 and is $\varphi^n - \varphi^p = (kT/e) \ln(n_D n_A / n_i^2) = 0.9 \text{ V}$ for the given doping concentrations. For an applied voltage of $V_a = -2.3 \text{ V}$, the width of the space charge region is $w = 0.65 \mu\text{m}$.

- 6.3** The space charge capacitance per unit area is given by $C_{sc}/A = C_{sc}^* = |\partial Q_{sc}^* / \partial V|$. We have

$$Q_{sc}^* = e n_D w_n = \sqrt{2e \epsilon \epsilon_0 \frac{n_A n_D}{n_A + n_D} (\varphi^n - \varphi^p - V_a)}$$

and hence

$$C_{sc}^* = \frac{1}{2} \sqrt{\frac{n_A n_D}{n_A + n_D} \frac{2e \epsilon \epsilon_0}{\varphi^n - \varphi^p - V_a}}$$

The capacitance for 0 V is $C_{sc}^* \approx 30 \text{ nF cm}^{-2}$ and for an applied voltage of -4 V , it is $C_{sc}^* \approx 13 \text{ nF cm}^{-2}$.

If C_{sc}^{*-2} is plotted against V_a (the so-called Mott–Schottky plot), it is possible to determine $\varphi^n - \varphi^p$ and the smaller of the dopant concentrations if either $n_D \gg n_A$ or vice versa. It is

$$\frac{1}{C_{sc}^{*-2}} = \frac{1}{2e \epsilon \epsilon_0} \left(\frac{4(n_A + n_D)}{n_A n_D} (\varphi^n - \varphi^p) - \frac{4(n_A + n_D)}{n_A n_D} V_a \right)$$

From the intercept $C_{sc}^{*-2} = 0$, we get $V_a = (\varphi^n - \varphi^p)$. From the slope, n_A , (n_D) can be determined if $n_D \gg n_A$, ($n_A \gg n_D$).

- 6.4** The separation of the quasi-Fermi energies $\epsilon_{FC} - \epsilon_{FV}$ as a function of the impurity level ϵ_{imp} can be calculated from Equations 3.67 and 3.77. For $n_{e,imp}$ in Equation 3.67 we have to use the expression given by Equation 3.76. We then have two equations for the two unknown quantities n_e and n_h . From the result we can easily derive ϵ_{FC} and ϵ_{FV} . The results for a) $G = 10^{18} \text{ cm}^{-3} \text{ s}^{-1}$ and b) $G = 5 \times 10^{20} \text{ cm}^{-3} \text{ s}^{-1}$ are shown in Figure S.3.
- 6.5** The rate at which photons are emitted from a LED is given by the generalized Planck law in Equation 3.102. If we neglect series resistances, the applied voltage V is related to the difference of the quasi-Fermi energies by $eV = \epsilon_{FC} - \epsilon_{FV}$. In Equation 3.102, the difference of the quasi-Fermi energies can have any value. However, light emitted by a LED is based on spontaneous emission in contrast to a Laser diode, which is based on stimulated emission. Therefore, in the bulk of a LED, it must be $eV < \epsilon_G$. If the quantum

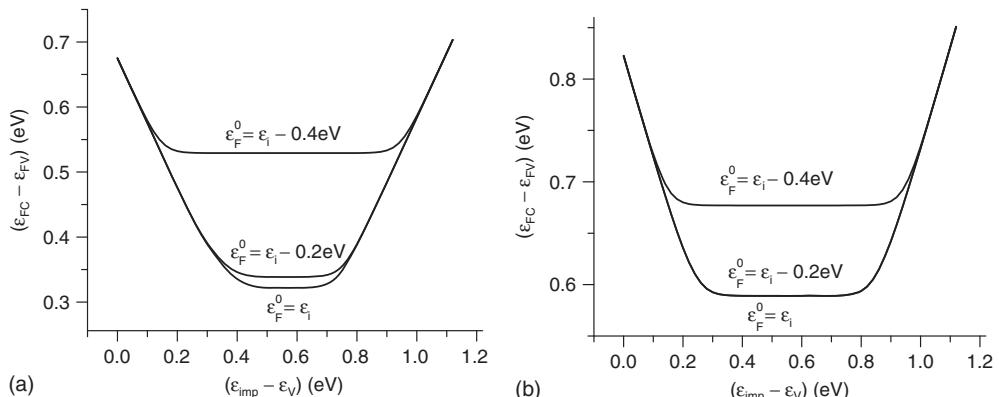


Figure S.3 Separation of the quasi-Fermi energies $\epsilon_{FC} - \epsilon_{FV}$ in Si as a function of the impurity level ϵ_{imp} for three different doping concentrations. (a) With a generation rate $G = 10^{18} \text{ cm}^{-3} \text{ s}^{-1}$. (b) With $G = 5 \times 10^{20} \text{ cm}^{-3} \text{ s}^{-1}$.

efficiency for the emission equals 1 (only radiative recombination), which is not excluded by the generalized Planck law, the LED emits more energy than is provided by the voltage source. This apparent violation of energy conservation is removed by noting that the energy provided by the voltage source is Free Energy (energy without entropy), whereas the energy emitted by the LED contains entropy (see Figure 3.16). This entropy and the heat energy going with it are provided by the phonon system. In fact, a LED with a quantum efficiency of 1 and zero series resistance is cooling its environment. That cooling is not observed for real LEDs, is due to the non-radiative part of the recombination, which consumes part of the current and to Joule heating at series resistances which consumes part of the applied voltage. For the same reason, voltages with $eV > \varepsilon_G$ applied to the contacts of a LED do not necessarily cause laser emission.

Chapter 7

7.1 Using Eq. (2.13) for the single diode model, we get

$$j_Q = j_s \left[\exp\left(\frac{e(V - j_Q \tilde{R}_s)}{kT}\right) - 1 \right] + j_{sc} + \frac{V - j_Q \tilde{R}_s}{\tilde{R}_p}$$

This transcendent equation is solved numerically and the results are shown in Figure S.4.

Chapter 8

8.1 A solar cell produces electrical energy from the absorbed solar energy. Ideally, a conversion system must be able to absorb the entire solar spectrum.

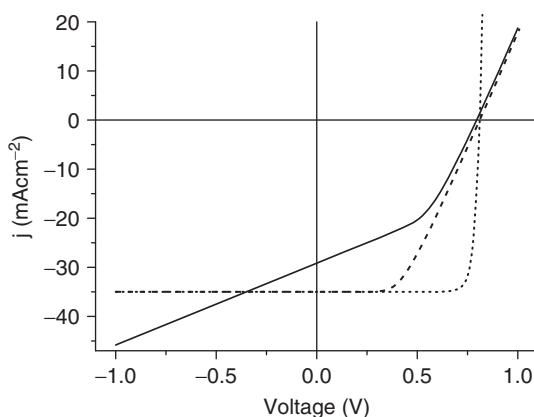


Figure S.4 Current–voltage characteristics of an ideal pn-junction with $\tilde{R}_s = 0$, $\tilde{R}_p = \infty$ (dotted line), one with $\tilde{R}_s = 10 \Omega\text{cm}^2$, $\tilde{R}_p = \infty$ (broken line) and one with $\tilde{R}_s = 10 \Omega\text{cm}^2$, $\tilde{R}_p = 50 \Omega\text{cm}^2$ (solid line).

Absorption losses are avoided in the thermophotovoltaic arrangement in Section 8.3, or in the hot carrier cell in Section 8.4.2 or in a tandem configuration of many cells with different band gaps in Section 8.1. The first step of the conversion process is the production of chemical energy of electron–hole pairs. This chemical energy is given by the difference of the Fermi energies $\varepsilon_{FC} - \varepsilon_{FV}$ of two different Fermi distributions describing separately the occupation of the valence and of the conduction band. For these distributions, the temperature must be defined. The first step in the conversion process is, therefore, the thermalization of the electrons and holes. If this thermalization occurs over a broad energy range, energy is lost. This loss is strongly reduced by restricting the thermalization to a narrow energy range, as by restricting already the generation of electron–hole pairs to a narrow energy range by a tandem configuration and by the filter in the thermophotovoltaic arrangement, or by thermalization in the energy-selective contacts of a hot carrier cell. The final conversion of chemical energy into electrical energy requires semipermeable membranes and a small resistance for the transport of the electrons and holes from where they are generated to the outside circuit. This ensures that the gradients of the quasi-Fermi energies, the driving forces for the transport, are small and little chemical energy is lost in the conversion to electrical energy.

- 8.2 (a)** To calculate the different recombination rates from the given quantity n_h the electron concentration n_e has to be determined. It follows from the condition of charge neutrality for the semiconductor. For radiative and/or Auger recombination, this condition leads to the equation $n_e = n_D + n_h$ according to Equation 3.26. If impurities are involved, the equation reads $n_e = n_D + n_h - n_{e,imp}$ according to Equation 3.67. The results are shown in Figure S.5.

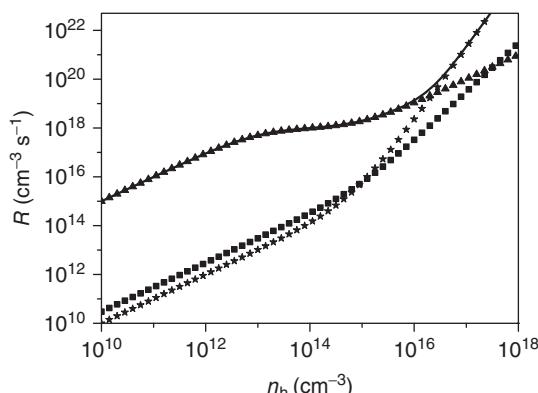


Figure S.5 The total recombination rate (solid line) as a function of the minority carrier concentration n_h and the different contributions from radiative (■), Auger (★), and impurity recombination (▲).

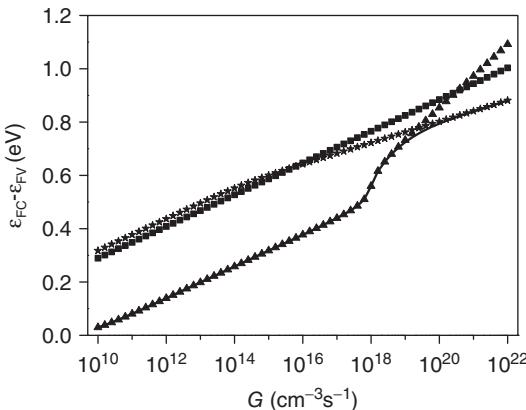


Figure S.6 Difference of the quasi-Fermi energies $\epsilon_{FC} - \epsilon_{FV}$ under open-circuit conditions as a function of the generation rate G . The recombination mechanism taken into

account is either radiative (■), Auger (★), or impurity recombination (▲) or a combination of all three types of recombination (solid line).

It can be seen from the graph that the impurity recombination rate determines the total rate for small values of n_h , corresponding to low illumination intensities. For higher values of n_h , the impurity recombination saturates, whereas the radiative and the Auger recombination become more important. Finally, for values $n_h \geq 3 \times 10^{16} \text{ cm}^{-3}$, Auger recombination dominates the total recombination rate. Note that for the given parameters, radiative recombination is not dominant at any intensity.

- (b) We can calculate the electron concentration n_e as a function of n_h and determine the quasi-Fermi energies and the recombination rates as a function of n_e and n_h . The other way is to determine n_e and n_h as a function of the recombination rates (which equal the generation rates) and determine ϵ_{FC} and ϵ_{FV} from n_e and n_h . The equation for charge neutrality is the same as in (a). The difference of the Fermi energies as a function of the generation rate for the different recombination mechanisms is shown in Figure S.6.

As is expected from (a), the open-circuit voltage is limited for small generation rates solely by impurity recombination, whereas for higher values of G , Auger recombination dominates. The graph shows the limitation of the open-circuit voltage by different recombination mechanisms.

Appendix

Fundamental Constants

Boltzmann constant	$k = 1.3807 \times 10^{-23} \text{ Ws K}^{-1} = 8.617 \times 10^{-5} \text{ eV K}^{-1}$
Planck constant	$h = 6.626 \times 10^{-34} \text{ Ws}^2 = 4.136 \times 10^{-15} \text{ eV s}$
Velocity of light	$\hbar = 1.0546 \times 10^{-34} \text{ Ws}^2 = 6.582 \times 10^{-16} \text{ eV s}$
Elementary charge	$c_{\text{vac}} = 2.998 \times 10^8 \text{ m s}^{-1}$
Electron mass	$e = 1.602 \times 10^{-19} \text{ As}$
Permittivity of free space	$m_e = 9.109 \times 10^{-31} \text{ kg}$
Stefan–Boltzmann constant	$\epsilon_0 = 8.85 \times 10^{-12} \text{ AsV}^{-1} \text{ m}^{-1}$
Solid angle of solar disc	$\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$
	$\Omega_s = 6.8 \times 10^{-5}$
	$\hbar\omega\lambda = hc_{\text{vac}} = 1.240 \text{ eV } \mu\text{m}$
	$\frac{1}{4\pi^3 \hbar^3 c_{\text{vac}}^2} = 5.04 \times 10^7 \frac{\text{W}}{(\text{eV})^4 \text{ m}^2}$

Units of Energy

$$\begin{aligned}1 \text{ eV} &= 1.602 \times 10^{-19} \text{ J} \\1 \text{ J} &= 1 \text{ Ws} = 1 \text{ N m} \\1 \text{ kWh} &= 3.6 \times 10^6 \text{ J}\end{aligned}$$

Material Constants at 300 K

	Ge	Si	GaAs
ϵ_G (eV)	0.66	1.12	1.42
χ (eV)	4.13	4.01	4.07
e	16	11.9	13.1
N_C (cm^{-3})	1×10^{19}	3×10^{19}	5×10^{17}
N_V (cm^{-3})	6×10^{18}	1×10^{19}	7×10^{18}
n_i (cm^{-3})	2.3×10^{13}	1×10^{10}	2.1×10^6
m_e^*/m_e	0.88	1.08	0.067
m_h^*/m_e	0.29	0.55	0.47
b_e ($\text{cm}^2 \text{ V}^{-1} \text{s}^{-1}$)	3800	1450	8500
b_h ($\text{cm}^2 \text{ V}^{-1} \text{s}^{-1}$)	1800	480	400

Standard Global AM1.5 Spectrumwith 1000 W m^{-2} total radiation

λ (μm)	$\frac{dj_E}{d\lambda}$ ($\text{W m}^{-2} \mu\text{m}^{-1}$)	λ (μm)	$\frac{dj_E}{d\lambda}$ ($\text{W m}^{-2} \mu\text{m}^{-1}$)	λ (μm)	$\frac{dj_E}{d\lambda}$ ($\text{W m}^{-2} \mu\text{m}^{-1}$)
0.3050	9.5	0.7400	1271.2	1.5200	262.6
0.3100	42.3	0.7525	1193.9	1.5390	274.2
0.3150	107.8	0.7575	1175.5	1.5580	275.0
0.3200	181.0	0.7625	643.1	1.5780	244.6
0.3250	246.8	0.7675	1030.7	1.5920	247.4
0.3300	395.3	0.7800	1131.1	1.6100	228.7
0.3350	390.1	0.8000	1081.6	1.6300	244.5
0.3400	435.3	0.8160	849.2	1.6460	234.8
0.3450	438.9	0.8237	785.0	1.6780	220.5
0.3500	483.7	0.8315	916.4	1.7400	171.5
0.3600	520.3	0.8400	959.9	1.8000	30.7
0.3700	666.2	0.8600	978.9	1.8600	2.0
0.3800	712.5	0.8800	933.2	1.9200	1.2
0.3900	720.7	0.9050	748.5	1.9600	21.2
0.4000	1013.1	0.9150	667.5	1.9850	91.1
0.4100	1158.2	0.9250	690.3	2.0050	26.8
0.4200	1184.0	0.9300	403.6	2.0350	99.5

λ (μm)	$\frac{dj_E}{d\lambda}$ ($\text{W m}^{-2} \mu\text{m}^{-1}$)	λ (μm)	$\frac{dj_E}{d\lambda}$ ($\text{W m}^{-2} \mu\text{m}^{-1}$)	λ (μm)	$\frac{dj_E}{d\lambda}$ ($\text{W m}^{-2} \mu\text{m}^{-1}$)
0.4300	1071.9	0.9370	258.3	2.0650	60.4
0.4400	1302.0	0.9480	313.6	2.1000	89.1
0.4500	1526.0	0.9650	526.8	2.1480	82.2
0.4600	1599.6	0.9800	646.4	2.1980	71.5
0.4700	1581.0	0.9935	746.8	2.2700	70.2
0.4800	1628.3	1.0400	690.5	2.3600	62.0
0.4900	1539.2	1.0700	637.5	2.4500	21.2
0.5000	1548.7	1.1000	412.6	2.4940	18.5
0.5100	1586.5	1.1200	108.9	2.5370	3.2
0.5200	1484.9	1.1300	189.1	2.9410	4.4
0.5300	1572.4	1.1370	132.2	2.9730	7.6
0.5400	1550.7	1.1610	339.0	3.0050	6.5
0.5500	1561.5	1.1800	460.0	3.0560	3.2
0.5700	1507.5	1.2000	423.6	3.1320	5.4
0.5900	1395.5	1.2350	480.5	3.1560	19.4
0.6100	1485.3	1.2900	413.1	3.2040	1.3
0.6300	1434.1	1.3200	250.2	3.2450	3.2
0.6500	1419.9	1.3500	32.5	3.3170	13.1
0.6700	1392.3	1.3950	1.6	3.3440	3.2
0.6900	1130.0	1.4425	55.7	3.4500	13.3
0.7100	1316.7	1.4625	105.1	3.5730	11.9
0.7180	1010.3	1.4770	105.5	3.7650	9.8
0.7244	1043.2	1.4970	182.1	4.0450	7.5

References

1. Falk, G. and Ruppel, W. (1976) *Energie und Entropie*, Springer-Verlag, Berlin.
2. Sackur, O. (1911) *Annalen der Physik*, **36**, 958.
3. Tetrode, H. M. (1912) *Annalen der Physik*, **38**, 434.
4. Würfel, P. (1982) *Journal of Physics C*, **15**, 3967.
5. (a) Shockley, W. and Read, W.T. (1952) *Physical Review*, **87**, 835; (b) Hall, R.N. (1951) *Physical Review*, **83**, 228.
6. Würfel, P., Finkbeiner, S., Daub, E. (1995) *Appl. Phys. A*, **60**, 67.
7. Landau, L.D. and Lifschitz, E.M. (1980) *Course of Theoretical Physics*, Vol. V, Butterworth-Heinemann.
8. Daub, E., Würfel, P. (1996) *Journal of Applied Physics*, **80**, 5325.
9. Kittel, C. (1958) *Elementary Statistical Physics*, John Wiley & Sons, Ltd.
10. deVos, A. (1983) *Proceedings of the 5th E.C. Photovoltaic Solar Energy Conference*, Athens, p. 186.
11. Shockley, W. and Queisser, H.J. (1961) *Journal of Applied Physics*, **32**, 510.
12. Würfel, U., Cuevas, A., and Würfel, P. (2015) *IEEE Journal of Photovoltaics*, **5**, 461.
13. O'Reagan, B. and Grätzel, M. (1991) *Nature*, **353**, 737.
14. Jaeger, K. and Hezel, R. (1985) *Proceedings of the 18th IEEE PV Specialists Conference*, Las Vegas, p. 388.
15. Tang, C. (1986) *Applied Physics Letters*, **48**, 183.
16. Walzer, K., Maennig, B., Pfeiffer, M., and Leo, K. (2007) *Chemical Reviews*, **107**, 1233.
17. Shaheen, S.E., Brabec, C.J., Padinger, F., Fromherz, T., Hummelen, J.C., and Sariciftci, N.S. (2001) *Applied Physics Letters*, **78**, 841.
18. Organic Photovoltaics, ed. by C. Brabec, O. Scherf and V. Dyakonov (2008), Wiley-VCH.
19. Yablonovitch, E. (1982) *Journal of the Optical Society of America*, **72**, 899.
20. Zhao, J., Wang, A., Green, M., and Ferrazza, F. (1998) *Applied Physics Letters*, **73**, 1991.
21. Brabec, C.J., Sariciftci, N.S., and Hummelen, J.C. (2001) *Advanced Functional Materials*, **11**, 15.
22. Green, M.A. (2003) *Third Generation Photovoltaics*, Springer-Verlag.
23. Trupke, T. and Würfel, P. (2004) *Journal of Applied Physics*, **96**, 2347.
24. Sinton, R.A., Verlinden, P.J., Crane, D.E., Swanson, R.M. (1988) *Proceedings of the 8th E.C. Photovoltaic Solar Energy Conference*, Florence, p. 1472.
25. Dimroth, F., et al. (2016) *IEEE Journal of Photovoltaics*, **6**, 343.
26. Green, M.A. (2000) *Proceedings of the 16th European Photovoltaic Solar Energy Conference*, Glasgow, p. 51.
27. Werner, J.H., Brendel, R., and Queisser, H.J. (1994) *First World Conference on Photovoltaic Energy Conversion*, Hawaii, p. 1742.

28. Würfel, P. (1996) *Solar Energy Materials and Solar Cells*, **46**, 43.
29. Ross, R.T. and Nozik, A.J. (1982) *Journal of Applied Physics*, **53**, 3813.
30. Luque, A. and Marti, A. (1997) *Physical Review Letters*, **78**, 5014.
31. Trupke, T., Green, M.A., and Würfel, P. (2002) *Journal of Applied Physics*, **92**, 4117.
32. Trupke, T., Green, M.A., and Würfel, P. (2002) *Journal of Applied Physics*, **92**, 1668.
33. Holman, Z.C., Descoeuadres, A., de Wolf, S., and Ballif, C. (2013) *IEEE J. Photovoltaics*, **3**, 1243.
34. Sze, S.M. (1981) *Physics of Semiconductor Devices*, 2nd edn, John Wiley & Sons, Inc., New York, p. 803.
35. Basore, P.A. (1993) *Conference Proceedings of the 23rd IEEE Photovoltaic Specialists Conference*, p. 147.
36. Sinton, R.A., Cuevas, A., and Stuckings, M. (1996) *Proceedings of the 25th IEEE Photovoltaic Specialists Conference*, p. 457.
37. Feldmann, F., Bivour, M., Reichel, C., Steinkemper, H., Hermle, M., and Glunz, S.W. (2014) *Sol. Energy Mater. Sol. Cells*, **131**, 46.
38. Richter, A., Glunz, S.W., Werner, F., Schmidt, J., and Cuevas, A. (2012) *Phys. Rev. B*, **86**, 165202.
39. Pysch, D., Mette, A., and Glunz, S. (2007) *Sol. Energy Mater. Sol. Cells*, **91**, 1698.
40. Daub, E. and Würfel, P. (1995) *Phys. Rev. Lett.*, **74**, 1020.
41. (a) Bowden, S. and Sinton, R. (2007) *J. Appl. Phys.*, **102**, 124501; (b) Hovel, H.J. (1975) *Semiconductor and Semimetals*, vol. II, Academic Press, New York, p. 19.
42. Würfel, P., Trupke, T., Puzzer, T., Schäffer, E., Warta, W., and Glunz, S.W. (2007) *J. Appl. Phys.*, **101**, 123110.
43. Kampwerth, H., Trupke, T., Weber, J.W., and Augarten, Y. (2008) *Appl. Phys. Lett.*, **93**, 202102.
44. Trupke, T. et al. (2007) *Proceedings of the 22nd European Solar Energy Conference*, Milan, p. 22.
45. Breitenstein, O., Warta, W., and Langenkamp, M. (2003) *Lock-in Thermography, Basics and Application to Functional Diagnostics of Electronic Components*, Advanced Microelectronics, vol. **10**, Springer, Heidelberg, Berlin.
46. Breitenstein, O., Bauer, J., Bothe, K., Hinken, D., Müller, J., Kwapił, W., Schubert, M.C., and Warta, W. (2011) *IEEE J. Photovoltaics*, **1**, 159.
47. Carstensen, J., Popkirov, G., Bahr, J., and Föll, H. (2003) *Sol. Energy Mater. Sol. Cells*, **76**, 599.
48. Nguyen, M., Schütt, A., Carstensen, J., and Föll, H. (2012) *Proceedings of the 2012 MRS Fall Meeting, Symposium E*, p. 1493.
49. Sinton, R.A. and Cuevas, A. (2000) *Proceedings of the 16th European Photovoltaic Solar Energy Conference*, p. 1152.
50. Wolf, M. and Rauschenbach, H. (1963) *Adv. Energy Convers.*, **3**, 455.
51. Granek, F., Hermle, M., Huljic', D.M., Schultz-Wittmann, O., and Glunz, S.W. (2009) *Prog. Photovoltaics Res. Appl.*, **17**, 47.
52. Schiefer, S., Glunz, S.W., Zimmermann, B., and Würfel, U. (2014) *IEEE J. Photovoltaics*, **4**, 271.
53. Schiefer, S., Zimmermann, B., and Würfel, U. (2014) *J. Appl. Phys.*, **115**, 044506.
54. Würfel, U., Neher, D., Spies, A., and Albrecht, S. (2015) *Nat. Commun.*, **6**, 6951.
55. Köhler, A. and Bässler, H. (2015) *Electronic Processes in Organic Semiconductors: An Introduction*, Wiley-VCH Verlag GmbH. ISBN: 978-3-527-33292-2
56. Kniepert, J., Schubert, M., Blakesley, J.C., and Neher, D. (2011) *J. Phys. Chem. Lett.*, **2**, 700.
57. Juska, G., Arlauskas, K., Viliunas, M., Genevicius, K., Österbacka, R., and Stubb, H. (2000) *Phys. Rev. B*, **62**, R16235.
58. Lorrmann, J., Badada, B.H., Inganäs, O., Dyakonov, V., and Deibel, C. (2010) *J. Appl. Phys.*, **108**, 113705.

59. Hanfland, R., Fischer, M.A., Brüting, W., Würfel, U., and MacKenzie, R.C.I. (2013) *Appl. Phys. Lett.*, **103**, 063904.
60. Barsoukov, E. and MacDonald, J.R. (eds) (2005) *Impedance Spectroscopy: Theory, Experiment, and Applications*, 2nd edn, Wiley-VCH Verlag GmbH. ISBN: 978-0-471-64749-

Index

a

Abbé sine condition 28, 29
 absorptance 23
 – negative 92
 – of semiconductors 24
 absorption coefficient 70, 88
 – measurement 92
 – negative 91
 absorption, transient 238
 acceptors 53
 air mass 26
 AM0 spectrum 26, 27, 191, 206
 AM1 spectrum 26
 AM1.5 spectrum 26, 27, 264
 ambipolar diffusion 116
 amplification of light 91
 antireflection coating 14, 177
 Auger recombination
 – *see* recombination
 – – Auger 75

b

battery 61, 154, 158, 201
 black body 11, 20
 – radiation 11
 bleaching 238, 239
 Boltzmann constant 263
 Bose–Einstein distribution 12

c

carbon dioxide 4
 – atmospheric concentration 6
 Carnot efficiency 99
 Carnot engine 33–35, 39, 41
 cavity 11, 41
 characterisation 213, 220
 chemical energy 95
 chemical equilibrium 60

chemical potential

- electrons 106
- photons 90

collision time 107
 concentration of radiation 28, 31
 concentrator 32
 concentrator cells 194
 conduction band 44, 51
 – germanium 46
 conductivity 107
 coupled equilibria 61
 current–voltage characteristic
 – for impurity recombination 143
 – pn-junction 140, 142, 143

d

Dember effect 117
 density of states
 – combined 66
 – effective 48
 – for electrons 45
 – for photons 14, 15, 88
 detailed balance 71, 72, 78
 dielectric relaxation 115
 diffusion current 107
 diffusion equation 223
 diffusion length 113, 138, 225
 diffusion voltage 134
 direct transitions 65
 distribution function
 – for electrons 43
 – for photons 12
 donors 52
 doping
 – *see* semiconductors, doping 52
 down-conversion of photons 208
 drift velocity 107
 dye solar cell 131

e

- Earth
 - energy emission 7
 - temperature 6, 7, 9
- effective mass 45, 47
- efficiency
 - as function of energy gap 174
 - intensity dependence 186
 - maximum 32, 34, 98
 - monochromatic 102
 - overall 102
- Einstein relation 108
- electric field in a solar cell 154
- electrochemical equilibrium
 - in pn-junction 133
- electrochemical potential 59, 61
 - of electrons 63
 - of holes 63
- electroluminescence 223
- electron affinity 149
- electron capture 77
 - cross-section 77, 78, 81, 83
- electron emission 78
- electron mass 263
- electron–hole pairs
 - average energy 95
 - chemical energy 96, 97
 - equilibrium with photons 99
 - generation 69, 70, 85, 87, 90
 - interaction with photons 91
 - momentum 65
 - recombination 87
 - – see recombination 71
 - temperature 95
- elementary charge 263
- emission of photons 88
 - spontaneous 88
 - stimulated 88
- emissivity 23, 89
- emittance 90
- energy
 - consumption 2, 3, 5
 - current from Sun 2
 - density 15
 - dissipation 110
 - economy 1
 - maximum reserves 4
 - of radiation 16
 - reserves 2, 3, 5
 - units 2, 263
- energy current 21
 - per solid angle 17
- energy–momentum relation 65, 66

entropy 35, 37

- discharge from solar cell 156
- of electrons 62
- of holes 62
- of photons 99
- equilibrium
 - *see* thermodynamic equilibrium 60
- equivalent circuit 183
- exciton 160
 - dissociation 161
- extraction of charge carriers 240

f

- Fermi distribution 43
- Boltzmann approximation 74
- Fermi energy 62, 64
- Fick's law 108
- field current 106
- fill factor 173, 185, 186
- f*-number 31, 32
- focusing 28
- free energy 61

g

- gallium arsenide 47, 54, 148
- generation rate 70
- geometrical optics 30
- greenhouse effect 6–8

h

- hetero-junction 148, 153, 182
- holes 49
- homo-junction 148
- hot electrons 200

i

- impact ionization 70, 197
- impedance spectroscopy 241
- impurity photovoltaic effect 203
- impurity states 78
- index of refraction 29, 211
 - silicon 178
- indirect transitions 67

k

- Kirchhoff's law 22

l

- Lambert's law 21
- Landsberg efficiency 36, 38
- laser 91
 - laser condition 92
- LED 167

lifetime 81, 85, 218–220, 223, 230
 – for radiative recombination 86
 – upper limit 87
 light
 – polarization 13
 – trapping 177, 180
 – velocity 263
 light beam induced current (LBIC) 230
 luminescence 220

m
 maximum power point 171
 mean free path 107
 membrane
 – for electrons 128
 – for holes 128
 – semi-permeable 123
 metal contact 150
 – energy diagram 150
 MIS contact 153
 mobility 107
 momentum space 13

o
 oil consumption 5
 open-circuit voltage 142, 173, 180, 185
 optical transitions 91
 organic solar cells 160, 162

p
 Pauli Principle 43
 Peltier effect 151
 permittivity of free space 263
 phase space 13
 phonon energy 41
 photoemission 64
 photoluminescence 224
 photomultiplier 41
 photon absorption 92
 – probability 66
 photonic crystal structure 180
 photons 11
 photosynthesis 4
 photovoltage decay 237
 Planck constant 263
 Planck's radiation law 12, 15, 22
 – generalized 74, 92
 pn-junction 133, 136, 139, 141, 154, 157, 158
 – current–voltage characteristic 141
 – potential distribution 134
 Poisson's equation 134

q
 quantum efficiency (EQE) 214, 225

quasi steady state photoconductance (QSSPC)
 218
 quasi-Fermi distribution 57, 58

r
 recombination 71
 – Auger 75
 – efficiency 82
 – non-radiative 74
 – radiative 72, 87
 – surface 82
 – velocity 83, 84
 – via impurities 76
 recombination centers 56
 recombination rate 79, 87
 redox system 132
 reflection 176
 – total internal 14, 71, 177, 178
 resistance
 – parallel 184
 – reaction 140
 – series 184
 – transport 139, 140
 reverse saturation current 141, 143

s
 Schottky contact 152, 153
 semiconductors 41
 – direct 47
 – doping 52
 – ideal 99
 – III–V 85, 148, 195
 – illuminated 58
 – indirect 47, 68, 85
 – intrinsic 49
 – n-type 54
 – p-type 55
 series resistance 228
 Shockley–Queisser limit 222
 short-circuit current 141, 142, 173, 174, 180,
 189
 sine condition
 – see Abbé sine condition 29
 Snell's law of refraction 14, 70
 solar cells
 – chemical 123, 128, 140
 – dye 131
 – efficiency 175
 – energy conversion 186
 – maximum efficiency 171
 – minimal thickness 182
 – plastic 183
 – silicon 175
 – thin film 181

- solar disc 19
– solid angle 18, 263
solar spectrum 24, 25
solid angle 15, 18
space charge density 135
space charge layer 148
spectral response 213
spectrum 15
Staebler–Wronski effect 182
Stefan–Boltzmann constant 263
Stefan–Boltzmann radiation law 7, 20
Sun
– as black body 12
– radiation spectrum 7
– spectrum
– – *see* solar spectrum 25
Suns- V_{OC} 233
surface charge 151
surface photovoltage 119
surface states 82, 83
– density 83
- t**
tandem cells 189
– series connection 193
- temperature dependence of V_{oc} 185
temperature of the Earth 6
thermalization 57, 95
thermo-photonics 197
thermo-photovoltaic conversion 196
thermodynamic equilibrium 60
thermography 229
three-level system 203, 208
tunnel junction 193
two-diode model 143
two-step excitation 203
- u**
uncertainty principle 13, 45
up-conversion of photons 208
- v**
valence band 44, 51
– germanium 46
Venus atmosphere 9
vibrational quanta 41
- w**
work function 64
– measurement 64

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley's ebook EULA.