Peter Würfel

Physics of Solar Cells

From Principles to New Concepts



WILEY-VCH Verlag GmbH & Co. KGaA

Author

Prof., Dr. rer. nat., emerit. Peter Würfel Universität Karlsruhe Institut für Angewandte Physik peter.wuerfel@physik.uni-karlsruhe.de

Cover Picture

Jörg Trebs, Berlin

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 Cataloging-in-Publication Data: A catalogue record for this book is available from the British Library

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.ddb.de>.

© 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – nor transmitted or translated into 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.

Printed in the Federal Republic of Germany Printed on acid-free paper

Printing betz-druck GmbH, Darmstadt

Bookbinding Litges & Dopf Buchbinderei GmbH,

Heppenheim

ISBN 3-527-40428-7

Contents

	List	of Sym	bols	ix
	Pref	ace		хi
1	Prol	olems of	f the Energy Economy	1
	1.1	Energ	gy economy	1
	1.2		nate of the maximum reserves of fossil energy	3
	1.3		greenhouse effect	5
		1.3.1	Combustion	5
		1.3.2	The temperature of the earth	6
2	Pho	tons		9
	2.1	Black	k-body radiation	9
		2.1.1	Photon density n_{γ} in a cavity (Planck's law of radiation)	9
		2.1.2	Energy current through an area dA into the solid angle d Ω	14
		2.1.3	Radiation from a spherical surface into the solid angle $d\Omega$	16
		2.1.4	Radiation from a surface element into a hemisphere	
			(Stefan-Boltzmann radiation law)	17
	2.2	Kirch	nhoff's law of radiation for non-black bodies	19
		2.2.1	Absorption by semiconductors	21
	2.3	The s	solar spectrum	21
		2.3.1	Air Mass	23
	2.4	Conc	centration of the solar radiation	24
		2.4.1	The Abbé sine condition	26
		2.4.2	Geometrical optics	26
		2.4.3	Concentration of radiation using the sine condition	28
	2.5	Maxi	imum efficiency of solar energy conversion	29
3	Sem	icondu		37
	3.1	Elect	crons in semiconductors	38
		3.1.1	Distribution function for electrons	38
		3.1.2	Density of states $D_{e}(\varepsilon_{e})$ for electrons	39
		3.1.3	Density of electrons	43
	3.2		S	45
	3.3	Dopi	ng	47

vi Contents

	3.4	Quasi-Fermi distributions
		3.4.1 Fermi energy and electrochemical potential
		3.4.2 Work function
	3.5	Generation of electrons and holes
		3.5.1 Absorption of photons
		3.5.2 Generation of electron-hole pairs
	3.6	Recombination of electrons and holes
		3.6.1 Radiative recombination, emission of photons
		3.6.2 Non-radiative recombination
		3.6.3 Lifetimes
	3.7	Light emission by semiconductors
		3.7.1 Transition rates and absorption coefficient
4	Con	version of Thermal Radiation into Chemical Energy 85
_	4.1	Maximum efficiency for the production of chemical energy
5		version of Chemical Energy into Electrical Energy 93
	5.1	Transport of electrons and holes
		5.1.1 Field current
		5.1.2 Diffusion current
		5.1.3 Total charge current
	5.2	Separation of electrons and holes
	5.3	Diffusion length of minority carriers
	5.4	Dielectric relaxation
	5.5	Ambipolar diffusion
	5.6	Dember effect
	5.7	Mathematical description
6	Basi	ic Structure of Solar Cells 109
	6.1	A chemical solar cell
	6.2	Basic mechanisms in solar cells
	6.3	Dye solar cell
	6.4	The pn-junction
		6.4.1 Electrochemical equilibrium of electrons in a pn-junction in the dark. 116
		6.4.2 Potential distribution across a pn-junction
		6.4.3 Current-voltage characteristic of the pn-junction
	6.5	pn-junction with impurity recombination,
	_	two-diode model
	6.6	Hetero-junctions
	6.7	Semiconductor–metal contact
		6.7.1 Schottky contact
		6.7.2 MIS contact
	6.8	The role of the electric field in solar cells

Contents vii

7	Lim	itations on Energy Conversion in Solar Cells	137			
	7.1	Maximum efficiency of solar cells	137			
	7.2	Efficiency of solar cells as a function of their energy gap	140			
	7.3	The optimal silicon solar cell	141			
		7.3.1 Light trapping	142			
	7.4	Thin-film solar cells	147			
		7.4.1 Minimal thickness of a solar cell	147			
	7.5	Equivalent circuit	149			
	7.6	Temperature dependence of the open-circuit voltage	149			
	7.7	Intensity dependence of the efficiency	15 l			
	7.8	Efficiencies of the individual energy conversion processes	151			
8	Concepts for Improving the Efficiency of Solar Cells					
	8.1	Tandem cells	155			
		8.1.1 The electrical interconnection of tandem cells	158			
	8.2	Concentrator cells	160			
	8.3					
	8.4					
		8.4.1 Hot electrons from impact ionization	165			
		8.4.2 Energy conversion with hot electrons and holes	165			
	8.5	Two-step excitation in three-level systems	168			
		8.5.1 Impurity photovoltaic effect	168			
		8.5.2 Up- and down-conversion of photons	172			
9	Prospects for the Future					
	App	pendix	181			
	Inde	ex	183			

List of symbols

$h,\hbar=h/(2\pi)$	Planck's constant	eVs
$\hbar\omega$	photon energy	eV
$a(\hbar\omega)$	absorptivity	
$r(\hbar\omega)$	reflectivity	
$t(\hbar\omega)$	transmission	
$\varepsilon(\hbar\omega) = a(\hbar\omega)$	emissivity	
$lpha(\hbar\omega)$	absorption coefficient	1/cm
k	Boltzmann's constant	eV/K
σ	Stefan-Boltzmann constant	$W/(m^2K^4)$
T	temperature	K
n_j	concentration of particle type j	1/cm ³
e	electron	
h	hole	
γ	photon	
Γ	phonon	
$n_{\mathrm{e}}, n_{\mathrm{h}}$	concentration of electrons, holes	1/cm ³
$n_{\rm i}$	intrinsic concentration of electrons and holes	1/cm ³
$N_{ m C},N_{ m V}$	effective density of states	1/cm ³
	in conduction band, valence band	
ϵ_e,ϵ_h	energy of an electron, hole	eV
$\epsilon_{ m C}$	energy of an electron	eV
	at the conduction band minimum	
$\epsilon_{ m V}$	energy of an electron	eV
	at the valence band maximum	
μ_j	chemical potential of particle type j	eV
${f \eta}_j$	electro-chemical potential of particle type j	eV
$\chi_{ m e}$	electron affinity	eV
φ	electrical potential	V
e	elementary charge	As
V	$voltage = \left[\eta_e(x_1) - \eta_e(x_2)\right]/e$	V
$\epsilon_{ ext{FC}}$	Fermi energy for electron distribution	eV
	in conduction band	
$\epsilon_{ ext{FV}}$	Fermi energy for electron distribution	eV
	in valence band	
$m_{ m e}^*, m_{ m h}^*$	effective mass of electrons, holes	g

Physics of Solar Cells: From Principles to New Concepts. Peter Würfel Copyright © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 3-527-40428-7

X List of Syn

$b_{ m e},b_{ m h}$	mobility of electrons, holes	$cm^2/(Vs)$
D_{e}, D_{h}	diffusion coefficient of electrons, holes	cm ² /s
$ au_e, au_h$	recombination life time of electrons, holes	S
$R_{ m e},R_{ m h}$	recombination rate of electrons, holes	$1/(cm^3s)$
$G_{ m e},G_{ m h}$	generation rate of electrons, holes	1/(cm ³ s)
σ_e,σ_h	cross-section for the capture	cm ²
	of an electron, hole by an impurity	
j_j	current density of particles of type j	$1/cm^2$
jo	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 non-physicists particularly with the background for a thorough understanding.

Emphasis is placed on a thermodynamic approach which 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

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

xii Preface

This book is the result of a series of lectures dealing with the physics of solar cells. I am grateful to the many students who called my attention to errors and suggested improvements. The material presented here, which differs from the usual treatment of solar cells relying on an electric field for a driving force, is the result of several years of collaboration with my teacher, W. Ruppel.

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 which 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

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 U235. Satisfying our energy needs from reserves, introduces two problems. A source of energy can continue only until it is depleted. Well before this time, that is, at the latest right now, 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 comprised not only of fossil energy carriers, but also of the oxygen in the air which is burned together with them. In addition, we will 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, 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 which 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 non-science press like kW per hour for power, since most people mistake kW 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 equ.).

```
The following relations apply:
                        3.6 \times 10^{6} \text{ J}
1 kWh
                                                                 1 kWh
                                                       =
                        29 \times 10^{9} \text{ J}
1 t coal equ.
                                                            8200 kWh
1 kg oil
                                1.4 kg coal equ.
                                                             12.0 kWh
1 \text{ m}^3 \text{ gas}
                                1.1 kg coal equ.
                                                              9.0 kWh
1 barrel oil
                               195 kg coal equ.
                                                            1670 kWh
```

Physics of Solar Cells: From Principles to New Concepts. Peter Würfel Copyright © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 3-527-40428-7

Type	Consumption	Per capita consumption	
	$(10^6 \text{ t coal equ./a})$	(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 1.1: Primary energy consumption in Germany in 2002.

Table 1.2: World primary energy consumption 2002.

Type	Consumption	Per capita consumption	
	(10 ⁹ t coal equ./a)	(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	

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 t coal equ. $/a = 8200 \, \text{kWh} / a = 0.94 \, \text{kW}$.

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

$$570 \, TWh/a = 65 \, GW \Rightarrow 0.79 \, kW/person$$
.

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

Table 1.2 shows the 2002 consumption of primary energy in the world, 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 presently 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 present prices. The actual reserves may be up to 10 times as large, of about 10×10^{12} t coal equ.

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

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

which radiates towards the earth.

Table 1.3: The world's remaining energy reserves.

Reserves in
109 t coal equ.
210
170
660
1040

In densely populated regions such as Germany, however, the balance is not so favourable if we restrict ourselves to the natural processes of photosynthesis for the conversion of solar energy to other useful forms of energy. The mean annual energy current which the sun radiates onto Germany, with an area of $0.36 \times 10^6 \,\mathrm{km^2}$, is about $3.6 \times 10^{14} \,\mathrm{kWh/a} = 4.3 \times 10^{14} \,\mathrm{kWh/a}$ 10¹⁰ t coal equ./a. Photosynthesis, when averaged over all plants, has an efficiency of about 1% and produces around 400×10^6 t coal equ./a from the energy of the sun. This is insufficient to cover the requirements of primary energy of 494×10^6 t coal equ./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 the oxygen by photosynthesis which is consumed in the combustion of gas, oil and coal. And this does not even take into consideration that the bio-mass 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 oxygen is in short supply 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 estimate¹ 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 today there are still elementary metals on the surface of the earth, even though 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 the size of the carbon reserves stored in the products of photosynthesis.

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

$$n \times (H_2O + CO_2) \Rightarrow n \times CH_2O + n \times O_2.$$

¹G. Falk, W. Ruppel, *Energie und Entropie*, Springer-Verlag, Berlin 1976.

A typical product of photosynthesis is glucose: $C_6H_{12}O_6 \equiv 6 \times CH_2O$. For this compound and also for most other carbohydrates, the ratio of free oxygen produced by photosynthesis to carbon stored in the carbohydrates is:

$$1 \mod O_2 \Rightarrow 1 \mod C$$
 or

$$32 g O_2 \Rightarrow 12 g C$$

The mass of the stored carbon $m_{\rm C}$ can therefore be found from the mass of free oxygen $m_{\rm O_2}$:

$$m_{\rm C} = \frac{12}{32} m_{\rm O_2}.$$

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/cm² on the surface of the earth resulting from the air surrounding our planet, we can calculate the mass of air from the relationship $m_{Air} \times g = p_E \times$ area:

$$m_{\rm Air}/{\rm area} = p_{\rm E}/{\rm g} = \frac{10\,{\rm N/cm^2}}{10\,{\rm m/s^2}} = 1\,{\rm kg/cm^2}.$$

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

$$m_{\rm Air} = 1 \, {\rm kg/cm^2} \times 4\pi R_{\rm earth}^2 = 5 \times 10^{15}$$
 tons of air.

Since air consists of 80% N_2 and 20% O_2 (making no distinction between volume percent and weight percent), the mass of oxygen is: $m_{O_2} = 10^{15} \,\mathrm{t} \,\mathrm{O}_2$. The maximum amount of carbon produced in photosynthesis and now present in deposits on the earth is therefore:

$$m_{\rm C} = \frac{12}{32} m_{\rm O_2} = 400 \times 10^{12}$$
 tons of carbon.

Up to now 10.4×10^{12} t coal equ. have 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 rumours 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 warmed up or is 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 400×10^{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, e.g., since the birth of Christ, Figure 1.1 gives a frightening picture. Up to the beginning of the twentieth century, the consumption of reserves was practically negligible. From this time on, it then rises exponentially up to a maximum value which will be reached in one or two decades. Consumption will then fall off again, as

5

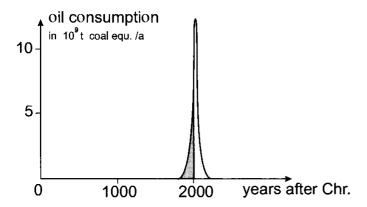


Figure 1.1: Annual consumption of oil. The area under the curve gives the estimated total oil reserves.

the reserves are gradually used up. The maximum consumption is expected when the reserves have fallen off to one half of their original levels. The reserves which have accumulated over millions of years will then literally go up in smoke over a time of only about one hundred 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 have changed over to supplying energy from regenerative sources, they will still suffer from the heritage which 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 for our climate. At the present time the atmosphere contains a fraction of 0.03 % of CO_2 . This corresponds to $2.3 \times 10^{12} \, \text{t}$ of CO_2 .

1.3.1 Combustion

Pure carbon is consumed according to the reaction $C + O_2 \Rightarrow CO_2$. 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 $m_{CO_2} = 44/12 m_C$. The combustion of 1 t of carbon thus results in $3.7 \text{ t of } CO_2$.

For different compounds of carbon other relationships apply:

- Carbohydrates:

$$30\,g\,CH_2O + 32\,g\,O_2 \Rightarrow 18\,g\,H_2O + 44\,g\,CO_2$$

resulting in $m_{\rm CO_2} = 1.47 \, m_{\rm CH_2O}$. This is the chemical reaction for the combustion of food in the human body.

- Methane (main component of natural gas):

$$16g CH_4 + 64g O_2 \Rightarrow 36g H_2O + 44g CO_2$$

resulting in $m_{\rm CO_2} = 2.75 \, m_{\rm CH_4}$.

The present global consumption of the $10^{10}\,t$ coal equ./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. At the present time the growth of one percent 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 one percent per year, the CO_2 concentration in the atmosphere will have doubled already after about 100 years. This increase is less the result of a per capita increase in energy consumption than a result of the increasing global population. The increasing CO_2 concentration in the atmosphere will have consequences for 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 will 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

$$j_{\rm E,sun}=1.3\,\rm kW/m^2\;.$$

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:

$$I_{\rm E,abs} = \pi R_{\rm e}^2 j_{\rm E,sun}$$
 with $R_{\rm e} = 6370 \, {\rm km}$ (radius of the earth).

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

$$j_{\rm E,\,earth} = \sigma T_{\rm e}^4$$
 where $\sigma = 5.67 \times 10^{-8} \, {\rm W}/({\rm m}^2 {\rm K}^4)$,

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

$$I_{\rm E,\,emit}=4\pi R_{\rm e}^2\,\sigma\,T_{\rm e}^4.$$

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

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% $(1 \, \text{kW/m}^2)$ 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 by the greenhouse and emits some of it back into the greenhouse.

We can understand the greenhouse effect of the atmosphere using a simple model. Due to a temperature of 6000 K of the sun, the solar radiation spectrum (expressed as energy current

7

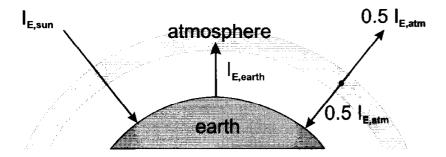


Figure 1.2: Balance of the absorbed and emitted energy currents on the surface of the earth.

per wavelength) has a maximum at about $0.5 \,\mu$ m, at which the atmosphere is transparent. Due to the lower temperature of the earth, the emission from the earth has its maximum at about $10\,\mu$ m (in the infrared). All tri-atomic 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 a maximum when the radiation from the surface of the earth is completely absorbed by the atmosphere, a situation which 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:

$$I_{\mathrm{E,earth}} = I_{\mathrm{E,sun}} + \frac{1}{2}I_{\mathrm{E,atm}}$$
.

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:

$$\frac{1}{2}I_{E, atm} = I_{E, sun}$$
 and $I_{E, earth} = 2I_{E, sun}$.

It then follows that

$$I_{\text{E,earth}} = 4\pi R_{\text{e}}^2 \, \sigma T_{\text{e,greenhouse}}^4 = 2\pi R_{\text{e}}^2 \times 1.3 \,\text{kW/m}^2$$
.

This yields a temperature of

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

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

However, the increased absorption of infrared radiation in the atmosphere due to human influences represents only half of the CO₂ production. The other half results from methane, fluorinated hydrocarbons and nitrogen oxides.

In order 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 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

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

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 comprised almost entirely of CO_2 . Treating the atmosphere of venus as a single fire screen yields a temperature of $116^{\circ}C$ for venus, a factor of $\sqrt[4]{2}$ greater than for the earth. The actual temperature of venus is, however, $475^{\circ}C$.

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 fear temperatures on the earth like 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 already occur 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 feed-back 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.

2 Photons

Photons are particles of light. Photons to which the human eye reacts, that is which we see, have energies $\hbar\omega$ between 1.5 eV and 3 eV. They always move with the velocity of light, in a vacuum with a velocity of $c_0 = 3 \times 10^8 \, \text{m/s}$ 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 behaviour 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 which 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 quantum-like transport of energy, as in this book.

2.1 Black-body radiation

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

$$a(\hbar \mathbf{\omega}) = 1. \tag{2.1}$$

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 of course 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 which is sufficiently thick and non-reflecting.

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

$$dn_{\gamma}(\hbar\omega) = D_{\gamma}(\hbar\omega)f_{\gamma}(\hbar\omega)d\hbar\omega. \qquad (2.2)$$

Physics of Solar Cells: From Principles to New Concepts. Peter Würfel Copyright © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 3-527-40428-7

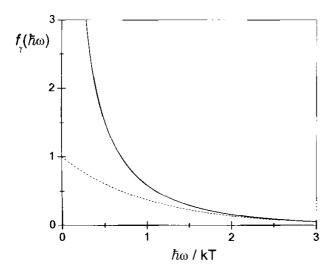


Figure 2.1: Bose–Einstein distribution function for $\mu_{\gamma} = 0$. The dashed line represents the Boltzmann distribution $\exp(-\hbar\omega/kT)$, a good approximation for $\hbar\omega > 2kT$.

Here $D_{\gamma}(\hbar\omega)$ is the density (per volume and per energy interval $d\hbar\omega$) of the states which the photons can occupy, and $f_{\gamma}(\hbar\omega)$ is the probability of occupation or the distribution function which 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:

$$f_{\gamma}(\hbar\omega) = \frac{1}{\exp\left(\frac{\hbar\omega - \mu_{\gamma}}{kT}\right) - 1} . \tag{2.3}$$

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

$$\Delta x \Delta p_x \ge h$$
. (2.4)

Particles which 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

$$\Delta x = L_x$$
 and consequently $\Delta p_x = h/L_x$

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

$$p_x = 0, \pm h/L_x, \pm 2h/L_x, \dots$$

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 $\Delta p^3 = h^3/V$ 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 $\hbar\omega$

$$\hbar \omega = c |\vec{p}|$$

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

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

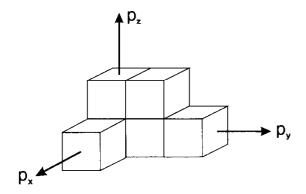


Figure 2.2: Volumes of the states in momentum space.

From (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 $d\hbar\omega$.

$$D_{\gamma}(\hbar\omega) = \frac{1}{V} \times \frac{\mathrm{d}N_{\gamma}(\hbar\omega)}{\mathrm{d}\hbar\omega} = \frac{(\hbar\omega)^2}{\pi^2\hbar^3c^3} \ . \tag{2.6}$$

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. In order 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 which can not 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 anti-reflection 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 (2.6) accounts for all photon states, regardless of their direction of motion, which is the direction of their momentum \vec{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 which are able to leave the cavity through the hole. At any location within the cavity, these are the photons moving towards 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 here 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 $d\Omega$ 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

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

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 $d\Omega$.

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

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 $\epsilon_{\gamma}=\hbar\omega$ we find the energy of the photons per volume and per photon energy interval $d\hbar\omega$ in the solid angle interval $d\Omega$

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

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

$$\hbar\omega_{\text{max}} = 2.82kT \ . \tag{2.9}$$

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

$$\hbar\omega = hv = \frac{hc}{\lambda}$$
 from which follows $d\hbar\omega = -\frac{hc}{\lambda^2}d\lambda$, (2.10)

(2.8) is transformed into

$$\frac{\mathrm{d}e_{\gamma}(\lambda)}{\mathrm{d}\lambda} = \frac{2hc\,\mathrm{d}\Omega}{\lambda^5} \,\,\frac{1}{\exp(hc/\lambda kT) - 1} \,. \tag{2.11}$$

We have left out the minus sign contained in (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

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

To convert photon energies $\hbar\omega$ to wavelengths λ we make use of Eq. (2.10):

$$\hbar\omega\lambda = hc = 1.240 \,\text{eV} \,\mu\text{m} \,. \tag{2.12}$$

From Eq. (2.8) the total energy of radiation per volume in the cavity, or the energy density, is

$$e_{\gamma} = \int_{0}^{\infty} \frac{(\hbar\omega)^{3} d\hbar\omega}{4\pi^{3}\hbar^{3}(c_{0}/n)^{3} \left[\exp(\hbar\omega/kT) - 1\right]} \int_{0}^{4\pi} d\Omega , \qquad (2.13)$$

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

$$e_{\gamma} = \frac{(kT)^4}{4\pi^3 \hbar^3 (c_0/n)^3} \underbrace{\int\limits_0^{\infty} \frac{x^3 dx}{e^x - 1}}_{\pi^4/15} 4\pi = \frac{\pi^2 k^4}{15 \, \hbar^3 (c_0/n)^3} T^4.$$
 (2.14)

The same result is of course obtained by integrating Eq. (2.11) over the wavelength. With reference to Eq. (2.7), the density of all photons in the cavity becomes

$$n_{\gamma} = \frac{(kT)^3}{4\pi^3 \hbar^3 (c_0/n)^3} \underbrace{\int_{0}^{\infty} \frac{x^2 dx}{e^x - 1}}_{2.40411} 4\pi = \frac{2.40411 \, k^3}{\pi^2 \hbar^3 (c_0/n)^3} T^3.$$
 (2.15)

The average photon energy of black-body radiation is thus given by

$$\langle \hbar \omega \rangle = e_{\gamma} / n_{\gamma} = 2.701 kT \ . \tag{2.16}$$

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 Ω 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 towards 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

$$dI_{e} = \frac{e_{\gamma}}{4\pi} \frac{dL}{dt} \frac{dA}{L^{2}} L^{2} d\Omega = e_{\Omega} c dA d\Omega , \qquad (2.17)$$

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 non-zero energy current and, therefore, cannot exist.

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

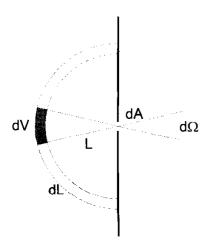


Figure 2.3: An energy current I_E originating from the volume element dV is flowing through the hole with area dA of the cavity, into the solid angle $d\Omega$.

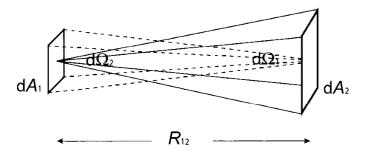


Figure 2.4: Solid angle $d\Omega_2$ under which the receiver area dA_2 is seen from the emitter dA_1 , and solid angle $d\Omega_1$ under which dA_1 is seen from dA_2 .

- 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

$$dI_{E1} = e_{\Omega 1} c \, dA_1 \, dA_2 / R_{12}^2 = dI_{E2} = e_{\Omega 2} c \, dA_2 \, dA_1 / R_{12}^2 . \tag{2.18}$$

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 which 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

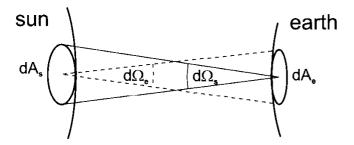


Figure 2.5: Solid angle $d\Omega_s$ under which the surface element dA_s on the sun is viewed from the earth, and solid angle $d\Omega_e$ under which the surface element dA_e on the earth is viewed from the sun.

the sun and the earth over the distance R_{se} , is given by

$$dI_{E, \text{sun-earth}} = e_{\Omega, \text{sun}} c d\Omega_{e} dA_{s}$$

$$= e_{\Omega, \text{sun}} c \frac{dA_{e} dA_{s}}{R_{\text{se}}^{2}}$$

$$= e_{\Omega, \text{sun}} c dA_{e} d\Omega_{s}.$$
(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

$$j_{E,\Omega,sun} = \frac{\mathrm{d}I_{E,sun-earth}}{\mathrm{d}A_s\,\mathrm{d}\Omega_e} = e_{\Omega,sun}\,c. \tag{2.20}$$

In accordance with the third line of Eq. (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

$$j_{E,\Omega,\,\mathrm{earth}} = \frac{\mathrm{d}I_{E,\,\mathrm{sun-earth}}}{\mathrm{d}A_{\mathrm{e}}\;\mathrm{d}\Omega_{\mathrm{s}}} = e_{\Omega,\,\mathrm{sun}}\;c = j_{E,\,\Omega,\,\mathrm{sun}}\;.$$

This is an important result. Even though the energy current per surface area, the energy current density, decreases with distance from the sun, near the sun it is of course much larger 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.

In order to find the energy current emitted from the entire sun towards 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 which 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 which *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 towards 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.

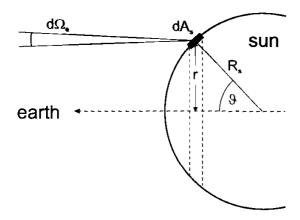


Figure 2.6: Energy current from the surface element dA_s on the surface of the sun into the solid angle $d\Omega_e$ under which the earth is viewed from the sun.

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

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

Since only the hemisphere directed towards us radiates energy to the earth, the total energy current which the sun radiates to the earth is

$$I_{E,e} = j_{E,\Omega} d\Omega_e \int_{0}^{\pi/2} 2\pi R_s^2 \sin\vartheta \cos\vartheta d\vartheta. \qquad (2.21)$$

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

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

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

$$I_E = j_{E,\Omega} \pi \, \mathrm{d}A \ . \tag{2.23}$$

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

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

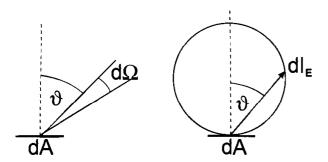


Figure 2.7: Dependence of the emitted energy current dI_E on the angle ϑ relative to the normal to the surface. The length of the arrow on the right is a measure of the magnitude of dI_E .

Making use of Eq. (2.14) for the energy density e_{γ} of black-body radiation then gives the Stefan-Boltzmann radiation law

$$j_E = \frac{\pi^2 k^4}{60\hbar^3 c^2} T^4 = \sigma T^4 \ . \tag{2.25}$$

The value of the Stefan-Boltzmann constant $\sigma = 5.67 \times 10^{-8} \, \text{W}/(\text{m}^2\text{K}^4)$. 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.

Eq. (2.24) also allows us to write the energy density per solid angle in the form

$$e_{\Omega} = \frac{\sigma}{\pi c} T^4. \tag{2.26}$$

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

$$j_{E,\Omega}=rac{\sigma T^4}{\pi}$$

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

$$j_E = \sigma T^4 \frac{\mathrm{d}\Omega}{\pi} \ . \tag{2.27}$$

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 behaviour precludes the recognition of any further details of a body by the emitted radiation, except for its contour.

This so-called Lambert behaviour cannot be taken for granted. It applies strictly only for the surfaces of bodies which 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/m²) Energy current density per solid angle $j_{E,\Omega}$ (measured in W/(m²sr)).

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 non-black bodies

In the cavity of Figure 2.8 two plates are positioned opposite each other at a distance 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 $(a_2 = 1)$. Plate 1 on the left represents a real material. It reflects (according to its reflectivity r_1), transmits (according to its transmission t_1) and absorbs (according to its absorptivity a_1) the respective part of the radiation incident from the black plate $d_{jE,2}$ 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 emitted by the black plate 2 in the energy interval $d\hbar\omega$ at the temperature T_2 with

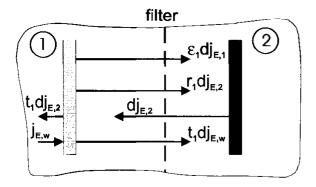


Figure 2.8: Exchange of radiation between two plates in a cavity with a filter in between, transmitting only photons with energy $\hbar\omega$.

the energy current density $dj_{E,2}(\hbar\omega, T_2)$ follows from Planck's law of radiation in (2.8) and is

$$dj_E(\hbar\omega) = \frac{d\Omega}{4\pi^3 \hbar^3 (c_0/n)^2} \frac{(\hbar\omega)^3}{\exp(\hbar\omega/kT) - 1} d\hbar\omega.$$
 (2.28)

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

$$r_1(\hbar\omega) + t_1(\hbar\omega) + a_1(\hbar\omega) = 1. \tag{2.29}$$

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

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 incident onto the black-body, plate 2, is comprised of the energy reflected by plate 1, the energy which 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 absorptivity of plate 2 is $a_2(\hbar\omega) = 1$ all the incident energy is absorbed and must be balanced by the emitted energy current $d_{jE,2}$ which is transmitted by the filter:

$$dj_{E,2}(\hbar\omega, T_2) = r_1(\hbar\omega) dj_{E,2}(\hbar\omega, T_2) + \varepsilon_1(\hbar\omega) dj_{E,1}(\hbar\omega, T_1) + t_1(\hbar\omega) dj_{E,w}(\hbar\omega, T_w)$$
(2.30)

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 Eq. (2.30) we obtain

$$r_1(\hbar\omega) + t_1(\hbar\omega) + \varepsilon_1(\hbar\omega) = 1$$
 (2.31)

Comparing with Eq. (2.29) we find

$$\varepsilon_1(\hbar\omega) = 1 - r_1(\hbar\omega) - t_1(\hbar\omega) = a_1(\hbar\omega) . \tag{2.32}$$

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

$$dj_E(\hbar\omega) = \frac{a(\hbar\omega) d\Omega}{4\pi^3 \hbar^3 (c_0/n)^2} \frac{(\hbar\omega)^3}{\exp(\hbar\omega/kT) - 1} d\hbar\omega . \qquad (2.33)$$

It is transported by the emitted photon current density

$$dj_{\gamma}(\hbar\omega) = \frac{a(\hbar\omega) d\Omega}{4\pi^{3} \,\hbar^{3} \,(c_{0}/n)^{2}} \,\frac{(\hbar\omega)^{2}}{\exp(\hbar\omega/kT) - 1} \,d\hbar\omega \quad . \tag{2.34}$$

The absorptivity $a(\hbar\omega)$ is characteristic of a body and is a function of the body's geometry. It describes the fraction of the incident light which 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 reflection, the transmission of a plate of thickness d is

$$t(\hbar\omega) = [1 - r(\hbar\omega)] \exp[-\alpha(\hbar\omega)d]. \tag{2.35}$$

From a = 1 - r - t we obtain the absorptivity of a plate of thickness d

$$a(\hbar\omega) = [1 - r(\hbar\omega)]\{1 - \exp[-\alpha(\hbar\omega)d]\}. \tag{2.36}$$

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 absorptivity 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

$$j_{E,\text{abs}} = \frac{\mathrm{d}\Omega}{4\pi^3 \hbar^3 c^2} \int_{\hbar\omega_{\mathrm{G}}}^{\infty} \frac{(\hbar\omega)^3}{\exp(\hbar\omega/kT) - 1} \,\mathrm{d}\hbar\omega \,. \tag{2.37}$$

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

$$\frac{1}{\exp(x) - 1} = \frac{\exp(-x)}{1 - \exp(-x)} = \exp(-x) \sum_{n=0}^{\infty} \exp(-nx) = \sum_{n=1}^{\infty} \exp(-nx) ,$$

which is valid for x > 0, Eq. (2.37) becomes

$$j_{E,\text{abs}} = \frac{\mathrm{d}\Omega (kT)^4}{4\pi^3 \hbar^3 c^2} \sum_{n=1}^{\infty} \int_{x_0}^{\infty} x^3 \exp(-nx) \, \mathrm{d}x.$$

Partial integration then gives

$$j_{E,\text{abs}} = \frac{\mathrm{d}\Omega (kT)^4}{4\pi^3 \hbar^3 c^2} \sum_{n=1}^{\infty} \exp(-nx_{\rm G}) \left(\frac{x_{\rm G}^3}{n} + \frac{3x_{\rm G}^2}{n^2} + \frac{6x_{\rm G}}{n^3} + \frac{6}{n^4} \right) . \tag{2.38}$$

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

$$j_{\gamma, abs} = \frac{d\Omega (kT)^3}{4\pi^3 \hbar^3 c^2} \sum_{n=1}^{\infty} \exp(-nx_G) \left(\frac{x_G^2}{n} + \frac{2x_G}{n^2} + \frac{2}{n^3} \right) . \tag{2.39}$$

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 which is incident on the earth before entering the atmosphere. It agrees well with a spectrum calculated from Eq. (2.33) if we assume that the sun is a black body and

22 Photons

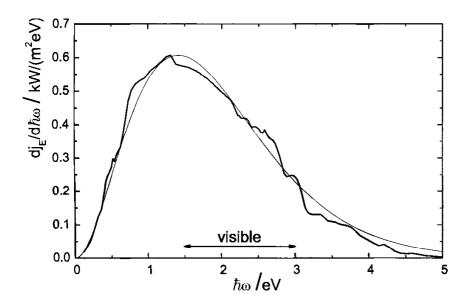


Figure 2.9: Energy current density per photon energy from the sun as a function of the photon energy just outside the earth's atmosphere (heavy line) compared with a black body at a temperature of 5800 K (thin line).

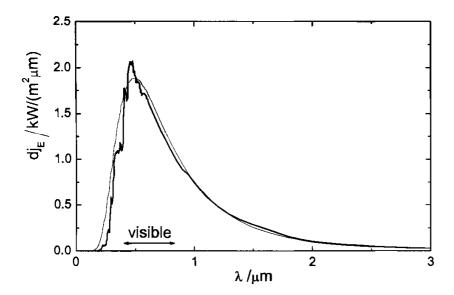


Figure 2.10: Energy current density per wavelength from the sun as a function of the wavelength outside the earth's atmosphere (heavy line) compared with a black body at a temperature of 5800 K (thin line).

has a temperature T_S of 5800 K. At this temperature of the surface of the sun, $kT_S = 0.5 \,\text{eV}$. The solar spectrum has a maximum at $\hbar\omega_{\text{max}} = 2.82kT_S = 1.41 \,\text{eV}$, i. e., in the infrared part of the spectrum, as seen in Figure 2.9. This is just outside of the visible range extending from 1.5 eV to 3 eV. The average energy of the solar photons is $2.7kT_S = 1.35 \,\text{eV}$.

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 $\lambda_{max} = 0.5 \mu m$, corresponding to a photon energy of

 $\hbar\omega = 2.48$ eV and distinctly different from the value $\hbar\omega_{max}$ found 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

$$j_{E,AM0} = 1353 \,\mathrm{W/m^2}$$
 (2.40)

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 $T_{\rm S} = 5800\,\rm K$. While 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 $\Omega_{\rm S} = 6.8 \times 10^{-5}$.

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 vapour (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

$$l=l_0/\cos\alpha$$
.

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, 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

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

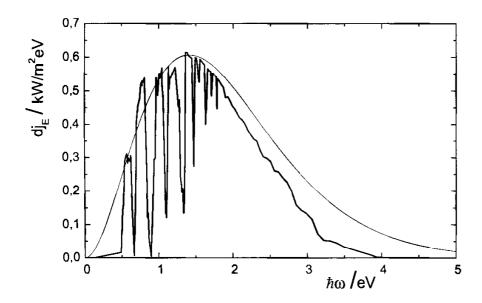


Figure 2.11: The AM1.5 spectrum (heavy line) compared with a black body at a temperature of 5800 K (thin line).

This maximum energy current density is only negligibly larger 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 $1000 \, \text{kWh/}(\text{m}^2\text{a})$. We therefore speak of $1000 \, \text{sun-hours}$ (with $1 \, \text{kW/m}^2$) per year.

Averaged over the year, the mean energy current density in Germany is $115 \,\mathrm{W/m^2}$, and roughly a factor of ten 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 $2500 \,\mathrm{kWh/(m^2a)}$. This corresponds to an average energy current density of $285 \,\mathrm{W/m^2}$. The average value over the entire earth is $230 \,\mathrm{W/m^2}$.

For the AMO 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 $I_e = 1353 \times \pi R_{\rm earth}^2 \, {\rm W/m^2}$ incident upon the earth by the total surface area of the earth $(4\pi R_{\rm earth}^2)$ to obtain the value $\langle j_e \rangle = 1/4 \times 1353 \, {\rm W/m^2} = 338 \, {\rm W/m^2}$ 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

$$\Omega_{S} = 2\pi \int_{0}^{2\pi} \sin \vartheta \, d\vartheta = 2\pi \left(1 - \cos \frac{\alpha_{S}}{2} \right)$$

$$= 6.8 \times 10^{-5} . \tag{2.41}$$

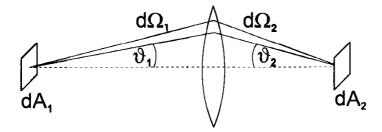


Figure 2.12: An imaging system projecting two area elements onto each other by redirecting their thermally emitted radiation must not disturb the thermal equilibrium between them.

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 $1 \,\mathrm{kW/m^2}$. 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 towards 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 not to violate 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 towards 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 towards 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 therefore exactly the same at the location of the lens in the direction towards dA_2 as at the location of dA_2 from the direction of the lens. Due 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 non-absorbing and non-emitting 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 Eq. (2.3) remains the same.

2.4.1 The Abbé sine condition

The energy current emitted in Fig. 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

$$dI_{E,1} = j_{E,\Omega,1} dA_1 \cos \theta_1 \sin \theta_1 d\theta_1 d\phi_1. \tag{2.42}$$

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 $(\vartheta_2 \varphi_2)$ into the solid angle element $d\Omega_2 = \sin \vartheta_2 d\vartheta_2 d\varphi_2$

$$dI_{E,2} = j_{E,\Omega,2} dA_2 \cos \vartheta_2 \sin \vartheta_2 d\vartheta_2 d\varphi_2 = dI_{E,1}. \tag{2.43}$$

In order to generalize the condition in (2.43) 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 Eq. (2.33) for the energy current densities per solid angle,

$$\frac{j_{E,\Omega,1}}{j_{E,\Omega,2}} = \frac{c_2^2}{c_1^2} = \frac{n_1^2}{n_2^2} \ . \tag{2.44}$$

The entire energy current absorbed by the area element dA_2 or emitted by it towards the lens is obtained by integrating over the entire lens or over the angle ϑ_2 from 0 to v, where v 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

$$I_{E,2} = j_{E,\Omega,2} \, dA_2 \, \pi \sin^2 v \,. \tag{2.45}$$

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

$$I_{E,1} = j_{E,\Omega,1} \, \mathrm{d}A_1 \, \pi \, \sin^2 u \,\,, \tag{2.46}$$

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 Eq. (2.44) we obtain the Abbé sine condition for optical imagery

$$n_1^2 dA_1 \sin^2 u = n_2^2 dA_2 \sin^2 v . (2.47)$$

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

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

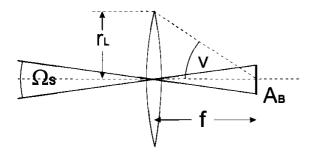


Figure 2.13: Image of the sun with an area A_B in the focal plane of a lens with radius r_L and focal length f.

where $j_{E,sun}$ 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

$$j_{E,\text{image}} = I_E/A_B = j_{E,\text{sun}} \pi r_L^2/A_B$$
.

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

$$C = \frac{\pi}{\Omega_{\rm S}} \tan^2 v \ . \tag{2.48}$$

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 $v > 45^{\circ}$, that is $r_L > f$. For this lens $C > \pi/\Omega_{\rm S}$.

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 Eq. (2.27) and Eq. (2.48) the energy current density in the image of the sun would be

$$j_{E, \mathrm{image}} = C \, j_{E, \mathrm{sun}} > rac{\pi}{\Omega_{\mathrm{S}}} rac{\Omega_{\mathrm{S}}}{\pi} \sigma T_{\mathrm{S}}^4$$

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 .

28 2 Photons

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

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

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

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

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 $j_{E,\Omega,sun}$. 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 Eq. (2.45) the energy current density in the image of the sun is

$$j_{E,\text{image}} = j_{E,\Omega,\text{sun}} \pi \sin^2 v$$
.

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

$$C = \frac{j_{E,\text{image}}}{j_{E,\text{sun}}} = \frac{\pi}{\Omega_{S}} \sin^{2} v. \tag{2.49}$$

The maximum concentration is found for $v = 90^{\circ}$

$$C_{\text{max}} = \frac{\pi}{\Omega_{\text{S}}} = 46200.$$
 (2.50)

The contradiction between Eq. (2.48), based on geometrical optics of an undistorted image, and Eq. (2.49), 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 towards 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

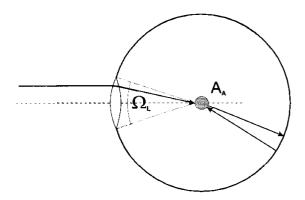


Figure 2.14: A lens projects an image of the sun onto an absorber in a perfectly reflecting cavity.

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 which 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 towards 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.

2.5 Maximum efficiency of solar energy conversion

In order to utilize 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. Due to the extraction of energy, the temperature of the absorber T_A will be lower than the temperature of the sun T_S . It must be lower in order not to emit as much energy as is absorbed.

The useful net absorbed energy current $I_{E,\text{util}}$ which 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 towards 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:

$$I_{E,\mathrm{util}} = I_{E,\mathrm{abs}} - I_{E,\mathrm{emit}} = \sigma \left(\frac{\Omega_{\mathrm{abs}}}{\pi} T_{\mathrm{S}}^4 - \frac{\Omega_{\mathrm{emit}}}{\pi} T_{\mathrm{A}}^4 \right) A_{\mathrm{A}} \ .$$

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

$$\eta_{\text{abs}} = \frac{I_{E,\text{util}}}{I_{E,\text{abs}}} = 1 - \frac{I_{E,\text{emit}}}{I_{E,\text{abs}}} = 1 - \frac{\Omega_{\text{emit}}}{\Omega_{\text{abs}}} \frac{T_{\text{A}}^4}{T_{\text{S}}^4}.$$

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.

30 2 Photons

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_0I_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 $I_{E,\text{el}}$,

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

The efficiency for this process is

$$\eta_{\rm C} = \frac{I_{E, \, \rm el}}{I_{E, \, \rm util}} = \frac{T_{\rm A}I_{\rm S} - T_{\rm 0}I_{\rm S}}{T_{\rm A}I_{\rm S}} = 1 - \frac{T_{\rm 0}}{T_{\rm A}} \,. \tag{2.51}$$

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.

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

$$\eta_{bC} = \frac{I_{E,el}}{I_{E,abs}} = \frac{I_{E,el}}{I_{E,util}} \frac{I_{E,util}}{I_{E,abs}} = \eta_{abs} \eta_C$$

$$\eta_{bC} = \left(1 - \frac{\Omega_{emit}}{\Omega_{abs}} \frac{T_A^4}{T_S^4}\right) \left(1 - \frac{T_0}{T_A}\right).$$
(2.52)

Figure 2.15 displays η_{bC} as a function of T_A for $T_S = 5800 \,\mathrm{K}$ and $T_0 = 300 \,\mathrm{K}$ 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 $T_A = 2478 \,\mathrm{K}$. 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 $T_S = 5800 \,\mathrm{K}$ 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 not be unjustified, since it would actually prolong the sun's lifetime). Under these conditions maximal concentration must be chosen and also a temperature of the absorber which is only slightly (by dT) smaller than T_S . The net absorbed energy current then is

$$I_{E,\mathrm{abs}} = \left[rac{\Omega_{\mathrm{abs}}}{\pi} \, \sigma T_{\mathrm{S}}^4 - rac{\Omega_{\mathrm{emit}}}{\pi} \, \sigma \left(T_{\mathrm{S}} - \mathrm{d}T
ight)^4
ight] A_{\mathrm{A}} pprox 4 rac{\Omega_{\mathrm{abs}}}{\pi} \, \sigma T_{\mathrm{S}}^3 \, \mathrm{d}T A_{\mathrm{A}} \; .$$

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

$$I_{S,\mathrm{abs}} = \frac{4}{3} \frac{\Omega_{\mathrm{abs}}}{\pi} \, \sigma \left[T_{\mathrm{S}}^3 - \left(T_{\mathrm{S}} - \mathrm{d}T \right)^3 \right] A_{\mathrm{A}} \approx 4 \, \frac{\Omega_{\mathrm{abs}}}{\pi} \, \sigma T_{\mathrm{S}}^2 \, \mathrm{d}T A_{\mathrm{A}} \; .$$

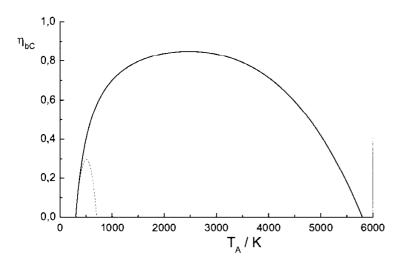


Figure 2.15: Efficiency η_{bC} for the conversion of solar heat energy by a black absorber and a Carnot engine as a function of the absorber temperature T_A , for maximum concentration ($\Omega_{emit} = \Omega_{abs}$, heavy line) and for non-concentrated radiation ($\Omega_{emit} = \pi$, $\Omega_{abs} = 6.8 \times 10^{-5}$, broken line).

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 no surprise 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 which 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 which are in a narrow energy interval $d\hbar\omega$ around a photon energy of, e. g., $\hbar\omega=1.5\,\mathrm{eV}$ 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 Eq. (2.34), the net absorbed photon current, the difference between absorbed and emitted photon currents is

$$dj_{\gamma, \text{net}}(\hbar\omega) = dj_{\gamma, \text{abs}} - dj_{\gamma, \text{emit}}$$

$$= m \left[\frac{(\hbar\omega)^2}{\exp(\hbar\omega/kT_S) - 1} - \frac{(\hbar\omega)^2}{\exp(\hbar\omega/kT_A) - 1} \right] d\hbar\omega, \qquad (2.53)$$

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 $\hbar\omega(1-T_0/T_{\rm A})$ per photon, resulting in an energy current delivered by the Carnot engine of

$$\mathrm{d}j_{E,\mathrm{Carnot}}(\hbar\omega) = \mathrm{d}j_{\gamma,\mathrm{net}}(\hbar\omega) \,\hbar\omega \,\left(1 - \frac{T_0}{T_\mathrm{A}}\right) \,.$$

32 Photons

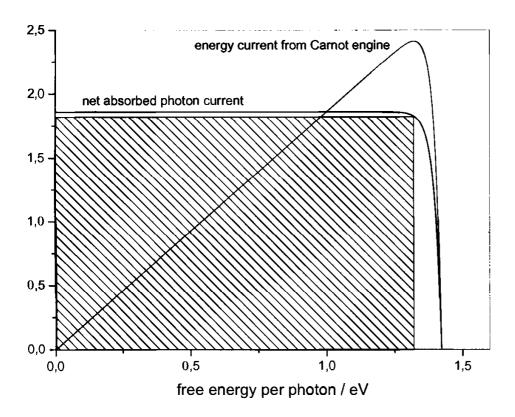


Figure 2.16: Net absorbed photon current of 1.5 eV photons as a function of the free energy per photon produced by a Carnot engine. The hatched rectangle, the largest rectangle for any point on the photon current curve, belongs to the point of maximum power, which also follows from the maximum of the energy current curve.

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 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 $dj_{E,Carnot}(\hbar\omega)$ shown by the thin line.

The efficiency of energy conversion is found by dividing the energy current $d_{jE,Carnot}(\hbar\omega)$ by the incident monochromatic energy current $\hbar\omega d_{j\gamma,abs}(\hbar\omega)$. 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%, 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 towards 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,

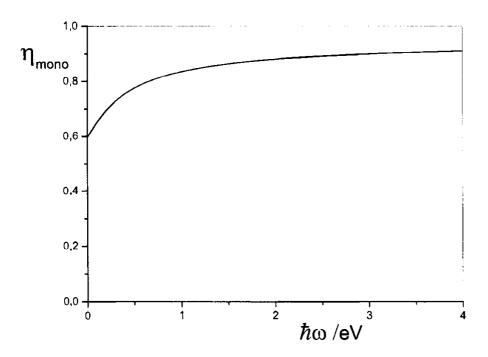


Figure 2.17: Efficiency of monochromatically absorbing Carnot engines as a function of the photon energy for fully concentrated black-body radiation of 5800 K.

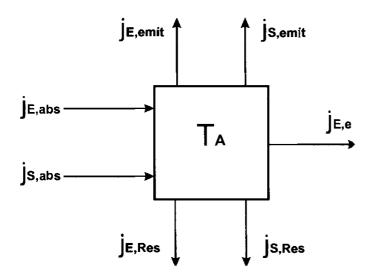


Figure 2.18: Balance of absorbed and emitted energy and entropy currents for a reversible heat engine working with radiation.

we might ask ourselves how high the efficiency would be if it were possible for a black body at a temperature $T_A < T_S$ 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.

34 2 Photons

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

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

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

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

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 $j_{S, res}$ 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

$$j_{S,abs} = j_{S,emit} + j_{S,res}$$
.

This leads to the relationships

$$j_{S,\text{res}} = \frac{4}{3}\sigma \left(T_{S}^{3} - T_{A}^{3}\right)$$

and

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

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

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

The Landsberg efficiency is therefore

$$\eta_{L} = \frac{j_{E,el}}{j_{E,abs}} = 1 - \frac{\sigma T_{A}^{4} + 4/3 \sigma T_{0} (T_{S}^{3} - T_{A}^{3})}{\sigma T_{S}^{4}}$$

$$= 1 - \frac{T_{A}^{4}}{T_{S}^{4}} - \frac{4}{3} \frac{T_{0}}{T_{S}} \left(1 - \frac{T_{A}^{3}}{T_{S}^{3}} \right).$$
(2.54)

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.

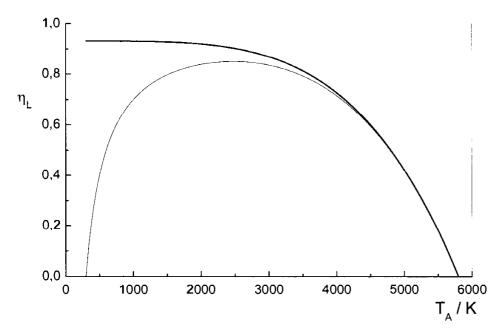


Figure 2.19: Landsberg efficiency η_L (heavy line) as a function of the absorber temperature T_A compared with the efficiency η_{bC} (thin line) of a solar heat engine with a single black absorber.

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 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 towards 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.

The thermal energy which 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 eV 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 aluminium pot with some scouring powder and a cloth. The tiny aluminium particles scrubbed off the pot make the cloth totally black. Due to the continuous energy range for electrons in a metal, an electron being excited by having absorbed 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.

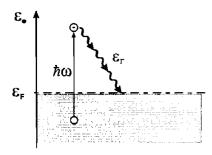


Figure 3.1: Excitation of an electron in the conduction band of a metal by the absorption of a photon with energy $\hbar\omega$ and the subsequent loss of the excitation energy by the generation of single phonons with energy ϵ_{Γ} .

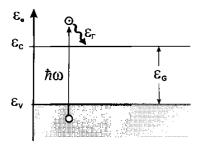


Figure 3.2: Excitation of an electron from the valence band to the conduction band of a semiconductor by the absorption of a photon with energy $\hbar\omega$.

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. In order 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 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 to lose their energy as quickly as in a metal by step-wise 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 step-wise 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 dn_e with energy ε_e distributed over the energy interval $d\varepsilon_e$ is comprised of the density of states $D_e(\varepsilon_e)$ and a distribution function $f_e(\varepsilon_e)$, which defines the occupation of the electron states:

$$dn_{e}(\varepsilon_{e}) = D_{e}(\varepsilon_{e}) f_{e}(\varepsilon_{e}) d\varepsilon_{e} . \tag{3.1}$$

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 non-integral 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 .

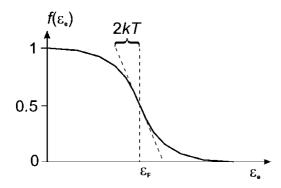


Figure 3.3: The Fermi distribution function $f(\varepsilon_e)$ defines the probability that an electron occupies a state with energy ε_e .

- 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 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 cm³ of the valence band and exciting it to any of 10^{23} states per cm³ 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 which satisfies all of these conditions is the Fermi distribution

$$f_{\rm c}(\varepsilon_{\rm c}) = \frac{1}{\exp\left(\frac{\varepsilon_{\rm e} - \varepsilon_{\rm F}}{kT}\right) + 1} \ . \tag{3.2}$$

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 $D_{\rm e}(\varepsilon_{\rm e})$ for electrons

In isolated atoms the electrons have discrete energy values, separated by large gaps on the energy scale. Decreasing the inter-atomic 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 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.

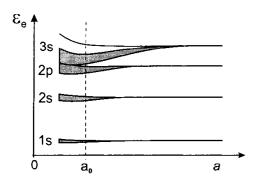


Figure 3.4: Energy of electron states in sodium as a function of the interatomic distance a between the sodium atoms. a_0 is the inter-atomic distance in solid sodium. States in the shaded energy ranges are occupied by electrons.

The stronger the interaction of the electrons with neighbouring atoms, the wider the bands become. The interaction is of course strongest for the outer shell valence electrons. Since these electrons have the 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 behaviour 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 non-conductors, because (at least at $T=0\,\mathrm{K}$) 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 $T=0\,\mathrm{K}$) is unoccupied.

In order 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 $|p'| \le |p|$ 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:

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

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

$$N_{\rm e}(|p|) = \frac{8\pi |p|^3 V}{3h^3} \ . \tag{3.4}$$

In order 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 momentum. For this reason, the states must have the same density per energy at positive and negative momentum. 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

$$\varepsilon_{\rm e} = \varepsilon_{\rm C} + \alpha p^2 + \dots$$

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

$$\varepsilon_{\rm e,kin} = \varepsilon_{\rm e} - \varepsilon_{\rm C} = \frac{p^2}{2m_{\rm e}^*} \ . \tag{3.5}$$

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

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

$$N_{\rm e}(\varepsilon_{\rm e}) = \frac{8\pi V (2m_{\rm e}^*)^{3/2}}{3h^3} (\varepsilon_{\rm c} - \varepsilon_{\rm C})^{3/2} . \tag{3.6}$$

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 Eq. (3.6).

$$D_{\rm e}(\varepsilon_{\rm e}) = \frac{1}{V} \frac{\mathrm{d}N_{\rm e}}{\mathrm{d}\varepsilon_{\rm e}} = 4\pi \left(\frac{2m_{\rm e}^*}{h^2}\right)^{3/2} (\varepsilon_{\rm e} - \varepsilon_{\rm C})^{1/2} . \tag{3.7}$$

Figure 3.5 shows the density of states of the conduction and valence bands for germanium. Important for the following discussion is only the fact that the densities of states at the upper

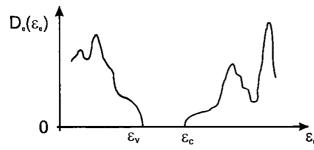


Figure 3.5: Density of states for electrons in the conduction and valence bands of the semiconductor germanium.

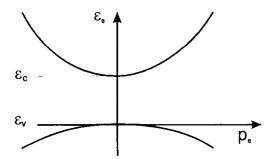


Figure 3.6: Energy ε_e of the electrons for a direct semiconductor as a function of their momentum p_e . The minimum of the conduction band and the maximum of the valence band occur at the same momentum $p_e = 0$.

and lower boundaries of a band show square-root behaviour as a function of the energy ε_e in agreement with Eq. (3.7), provided that we choose m_e^* properly. For the upper boundary of a band, at ε_V for the valence band, for example, it should not come as a surprise that we must choose $m_e^* < 0$. Treating electrons as quasi-free particles although they are in the field of the atoms ($\approx 10^{10}$ V/m) will lead to more peculiar properties. For electrons in the conduction band close to ε_C we actually expect the relationship between energy and momentum shown in Figure 3.6,

$$\varepsilon_{\rm e} - \varepsilon_{\rm C} = \frac{p_{\rm e}^2}{2m_{\rm e}^*} \,, \tag{3.8}$$

and obtain the effective mass of the electrons by matching

$$\frac{1}{m_e^*} = \frac{\mathrm{d}^2 \varepsilon_\mathrm{e}}{\mathrm{d} p_e^2}$$

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

For the $\varepsilon_e(p_e)$ relationship in Figure 3.6 the excitation of an electron from the valence band to the conduction band with the smallest possible energy, $\varepsilon_C - \varepsilon_V = \varepsilon_G$, 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 behaviour is unexpected from the idea of free particles. Here, the electrons have their lowest energy $\varepsilon_{\rm C}$ in the conduction band at a momentum different from zero, leading to the relation

$$\varepsilon_{\rm e} - \varepsilon_{\rm C} = \frac{(p_{\rm e} - p_{\rm e,0})^2}{2m_{\rm e}^*} \ . \tag{3.9}$$

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 $\varepsilon_e(p_e)$ relationship, must be symmetric with respect to the energy axis at

43

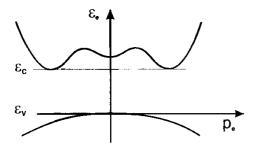


Figure 3.7: Energy of the electrons for an indirect semiconductor as a function of their momentum. The minimum of the conduction band and the maximum of the valence band occur at different values of the momentum p_e of the electrons.

 $p_e = 0$. 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 Ge (germanium) and Si (silicon).

3.1.3 Density of electrons

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

$$dn_{e}(\varepsilon_{e}) = D_{e}(\varepsilon_{e}) f_{e}(\varepsilon_{e}) d\varepsilon_{e}. \tag{3.10}$$

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:

$$n_{\rm e} = \int_{\epsilon_{\rm C}}^{\infty} D_{\rm e}(\epsilon_{\rm e}) f_{\rm e}(\epsilon_{\rm e}) d\epsilon_{\rm e} . \tag{3.11}$$

For the integration, we can make use of the density of states from Eq. (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 $\to \infty$), 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 Eq. (3.2), giving

$$n_{\rm e} = N_{\rm C} \, \exp\left(\frac{\varepsilon_{\rm C} - \varepsilon_{\rm F}}{kT}\right) \,.$$
 (3.12)

$$N_{\rm C} = 2\left(\frac{2\pi m_{\rm e}^* kT}{h^2}\right)^{3/2} \ . \tag{3.13}$$

 $N_{\rm C}$ is called the effective density of states of the conduction band. For $m_{\rm e}^*=m_{\rm e}$ it has the value

$$N_{\rm C} = 2 \times 10^{19} \,\rm cm^{-3}$$
 (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

$$n_{\rm h} = \int_{-\infty}^{\varepsilon_{\rm V}} D_{\rm e}(\varepsilon_{\rm e}) \left[1 - f_{\rm e}(\varepsilon_{\rm e})\right] d\varepsilon_{\rm e} = N_{\rm V} \exp\left(-\frac{\varepsilon_{\rm F} - \varepsilon_{\rm V}}{kT}\right) , \qquad (3.15)$$

where $N_{\rm V} = 2 \left(2\pi m_{\rm e}^* kT/h^2 \right)^{3/2}$ is the effective density of states of the valence band.

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

$$n_{\rm e}n_{\rm h} = N_{\rm C} \exp\left(-\frac{\varepsilon_{\rm C} - \varepsilon_{\rm F}}{kT}\right) N_{\rm V} \exp\left(-\frac{\varepsilon_{\rm F} - \varepsilon_{\rm V}}{kT}\right)$$
$$= N_{\rm C}N_{\rm V} \exp\left(-\frac{\varepsilon_{\rm C} - \varepsilon_{\rm V}}{kT}\right) = N_{\rm C}N_{\rm V} \exp\left(-\frac{\varepsilon_{\rm G}}{kT}\right). \tag{3.16}$$

The product of the electron density and the hole density does not depend on the position of the Fermi energy and therefore neither does it depend on the individual densities of 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 .

$$n_{\rm e}n_{\rm h} = n_{\rm i}^2 = N_{\rm C}N_{\rm V}\exp\left(-\frac{\varepsilon_{\rm G}}{kT}\right). \tag{3.17}$$

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 band. It rather follows from the condition $n_e = n_h$ in Eq. (3.12) and Eq. (3.15):

$$\varepsilon_{\rm F} = \frac{1}{2} (\varepsilon_{\rm V} + \varepsilon_{\rm C}) + \frac{1}{2} kT \ln \frac{N_{\rm V}}{N_{\rm C}}$$
(3.18)

or expressed in terms of the effective masses contained in $N_{\rm C}$ and $N_{\rm V}$

$$\varepsilon_{\rm F} = \frac{1}{2} (\varepsilon_{\rm V} + \varepsilon_{\rm C}) + \frac{3}{4} kT \ln \frac{m_{\rm h}^*}{m_{\rm e}^*} . \tag{3.19}$$

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

$$\langle \varepsilon_{\rm e} \rangle = \frac{1}{n_{\rm e}} \int_{\varepsilon_{\rm C}}^{\infty} \varepsilon_{\rm e} D_{\rm e}(\varepsilon_{\rm e}) f_{\rm e}(\varepsilon_{\rm e}) d\varepsilon_{\rm e} = \varepsilon_{\rm C} + \frac{3}{2} kT . \tag{3.20}$$

For the electrons, $\varepsilon_{\rm C}$ is a potential energy. The mean kinetic energy of the electrons (in the conduction band) is $\langle \varepsilon_{\rm e} - \varepsilon_{\rm C} \rangle = 3/2kT$, showing that the electrons in a non-degenerate semi-conductor ($\varepsilon_{\rm F} < \varepsilon_{\rm C} - 3kT$), where $n_{\rm e} \ll N_{\rm C}$, form an ideal gas.

3.2 Holes 45

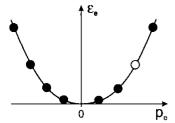


Figure 3.8: A band with a missing electron of $p_e > 0$ has a total momentum $\Sigma p_e < 0$ resulting in an electrical current $j_Q > 0$.

3.2 Holes

For each electron state with momentum \vec{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

$$j_{Q} = -\frac{e}{\text{Vol}} \sum_{\substack{\text{occupied} \\ \text{states}}} v_{e,i} = -\frac{e}{\text{Vol}} \sum_{\substack{\text{all states} \\ = 0 \text{ (filled band)}}} -\left(-\frac{e}{\text{Vol}}\tilde{v}_{e}\right) = \frac{e}{\text{Vol}}\tilde{v}_{e} . \tag{3.21}$$

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:

$$j_{Q} = -\frac{e}{\text{Vol}} \underbrace{\sum_{\substack{\text{all states} \\ \text{=0 (filled band)}}} + \underbrace{\left(\frac{q_{\text{h}}}{\text{Vol}}\tilde{v}_{\text{h}}\right)}_{\text{hole}}.$$
(3.22)

so that

$$j_Q = \frac{e}{\text{Vol}} \, \tilde{v}_e = \frac{q_h}{\text{Vol}} \, \tilde{v}_h \,. \tag{3.23}$$

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

$$q_{\rm h} = +{\rm e}$$
.

From Eq. (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

$$a_{\mathrm{e}} = -\frac{\mathrm{e}E}{m_{\mathrm{e}}^*} = \frac{\mathrm{e}E}{m_{\mathrm{h}}^*} = a_{\mathrm{h}} .$$

The effective mass of the hole is therefore

$$m_{\rm h}^*=-m_{\rm e}^*$$
.

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

$$ar{p} = \sum_{\mathrm{filled\ band}} ar{p_{\mathrm{i}}} \ - ar{p}_{\mathrm{e}} = \sum_{\mathrm{filled\ band}} ar{p_{\mathrm{i}}} \ + ar{p}_{\mathrm{h}} \ ,$$

$$\tilde{p}_{\rm h} = -\tilde{p}_{\rm e}$$
.

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

$$\epsilon_h = -\epsilon_e$$
 .

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, since 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

$$\langle \varepsilon_{\mathbf{h}} \rangle = -\varepsilon_{\mathbf{V}} + \frac{3}{2}kT \ . \tag{3.24}$$

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.

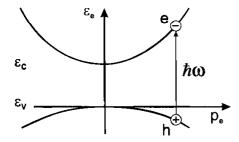


Figure 3.9: Creation of an electron-hole pair by the absorption of a photon.

3.3 Doping 47

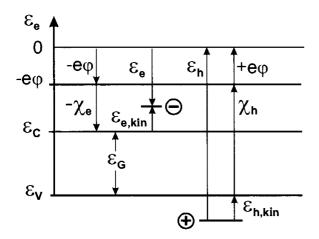


Figure 3.10: Energy scale for electrons and holes in a semiconductor. The binding energy of an electron in a state at the lower boundary of the conduction band is referred to as the electron affinity χ_e .

As a result of the excitation, the conduction band now has an extra electron with momentum \bar{p}_e and energy ε_e , while 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

$$\gamma \rightarrow e + h$$
.

This process must of course satisfy the conservation of momentum $\bar{p}_{\gamma} = \bar{p}_{e} + \bar{p}_{h}$ and the conservation of energy $\varepsilon_{\gamma} = \hbar \omega = \varepsilon_{e} + \varepsilon_{h}$.

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 $\phi=0$, having a kinetic energy $\epsilon_{e,kin}=0$. 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 upwards about the zero line. This representation is, however, complicated and 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 upwards indicate positive energies. For electron energies the arrows point downwards, since they are negative with respect to the zero line, and for hole energies the arrows point upwards. 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 which, as a rule, have one valence electron more than is necessary for chemical bonding with the neighbouring atoms. The electron not required for bonding is electrically bound to its atom by Coulomb forces as a negative charge in the field of a positive charge. To a first approximation, we expect this binding energy to be that of the

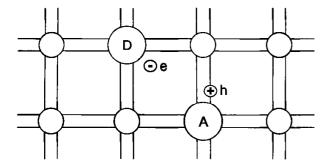


Figure 3.11: By doping lattice atoms are replaced by impurity atoms having a higher (D) or lower (A) valency.

electron in a hydrogen atom with the difference that we have to use the effective mass instead of the real mass of the electron,

$$\varepsilon_{\rm H} = \frac{m_{\rm e}^* e^4}{2(4\pi \varepsilon_0)^2 \hbar^2} = 13.5 \,{\rm eV} \frac{m_{\rm e}}{m_{\rm e}^*} \,.$$
(3.25)

Since the donor is, however, in a semiconductor and not in vacuum, the electric field binding the electron to the nucleus is somewhat weakened due to polarization of the neighbouring 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 ε_0 . For typical semiconductors such as Ge, Si and GaAs, the dielectric function ε has values > 10 for frequencies up to 10^{15} Hz, leading to a drastic reduction in the binding energy down to < 0.1 eV. For the same reason, in fact, (the chemical bond of) NaCl dissolves in H₂O.

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 which gives them their name.

Acceptors (A) are impurity atoms which, as a rule, have one valence electron less than is necessary for chemical bonding with the neighbouring atoms. An electron which 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: 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, while 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

3.3 Doping 49

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 vapour 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

$$\rho = e(n_{\rm h} + n_{\rm D}^+ - n_{\rm e} - n_{\rm A}^-) = 0. \tag{3.26}$$

Here n_e from Eq. (3.12) and n_h from Eq. (3.15) are just as much functions of the Fermi energy ε_F as

$$n_{\rm D}^+ = n_{\rm D} \left\{ 1 - \left[\exp\left(\frac{\varepsilon_{\rm D} - \varepsilon_{\rm F}}{kT}\right) + 1 \right]^{-1} \right\}$$

and

$$n_{\rm A}^- = n_{\rm A} \left[\exp \left(\frac{\varepsilon_{\rm A} - \varepsilon_{\rm F}}{kT} \right) + 1 \right]^{-1}$$
.

Eq. (3.26) therefore defines the position of the Fermi energy ε_F .

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 cm³ 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 Eq. (3.17) the hole 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

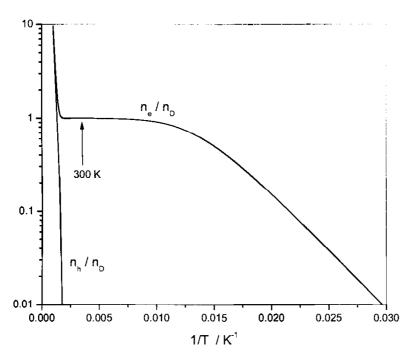


Figure 3.12: Temperature dependence of electron and hole concentrations in an n-type semiconductor with a band gap of 1 eV, containing 10^{16} donors per cm³ with an energy level $\epsilon_D = 0.05$ eV below the conduction band edge ϵ_C .

room temperature is greater than ε_A . At room temperature, donors and acceptors are almost completely ionized.

$$D \rightarrow D^+ + e$$
 $A \rightarrow A^- + h$
 $n_e \approx n_D$ $n_h \approx n_A$.

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.

The Fermi energy is obtained from Eq. (3.12) and is

$$\varepsilon_{\rm F} = \varepsilon_{\rm C} - kT \ln \frac{N_{\rm C}}{n_{\rm e}} \ ,$$

or from Eq. (3.15)

$$\varepsilon_{\rm F} = \varepsilon_{\rm V} + kT \ln \frac{N_{\rm V}}{n_{\rm h}} \ .$$

Typical doping concentrations are in the range of 10^{15} /cm³ to 10^{19} /cm³. This is relatively small compared with the lattice atom density, which is about 10^{23} /cm³. 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 which are small relative to the doping concentrations.

	$n_{\rm e}$	n_{h}	$\overline{\epsilon_{\mathrm{F}}}$
n-type	$n_{\rm e} \approx n_{\rm D}$	$n_{\rm h} = \frac{n_{\rm i}^2}{n_{\rm e}} = \frac{n_{\rm i}^2}{n_{\rm D}}$	$ \varepsilon_{\rm F} = \varepsilon_{\rm C} - kT \ln \frac{N_{\rm C}}{n_{\rm D}} $
p-type	$n_{\rm e} = \frac{n_{\rm i}^2}{n_{\rm h}} = \frac{n_{\rm i}^2}{n_{\rm A}}$		$\varepsilon_{\rm F} = \varepsilon_{\rm V} + kT \ln \frac{N_{\rm V}}{n_{\rm A}}$

Table 3.1: Electron and hole densities in n-type and p-type semiconductors.

Incorporation of both donors and acceptors in a semiconductor does not lead to increased concentrations of both electrons and holes. Eq. (3.17), $n_{\rm e}n_{\rm h}=n_i^2$ applies to doped semi-conductors as well. The electrons donated from the donors are instead largely taken up by the acceptors. (Alternatively, one could say that the holes which 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. Donor-like 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 the section on recombination via impurities on page 68.

3.4 Quasi-Fermi distributions

In a solar cell additional electrons and holes are produced by the absorption of photons from solar radiation. Figure 3.2 illustrates this process schematically. Immediately after their generation the energy distribution of the additional electrons and holes reflects the broad energy spectrum of the absorbed photons (indicating the high photon temperature) and is different from that in the dark state as shown by Figure 3.13. Due 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 $\varepsilon_{e,h,kin} = (3/2)kT_0$.

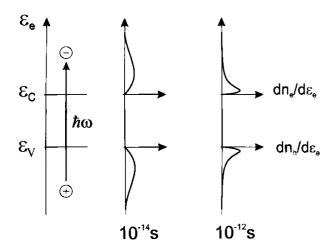


Figure 3.13: The broad energy distribution of electrons and holes, right after their generation by solar radiation, reflects the broad energy spectrum of the photons. After about 10^{-12} s electrons and holes are thermalized into a narrow room temperature distribution.

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 distribution, even for continuous irradiation. Due 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 density and the hole density 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 $\varepsilon_{kin} = (3/2)kT_0$. Due 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. Due 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

$$n_{\rm e} = N_{\rm C} \exp\left(-\frac{\varepsilon_{\rm C} - \varepsilon_{\rm FC}}{kT}\right) \tag{3.27}$$

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

$$n_{\rm h} = N_{\rm V} \exp\left(-\frac{\varepsilon_{\rm FV} - \varepsilon_{\rm V}}{kT}\right) \ . \tag{3.28}$$

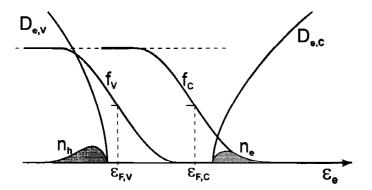


Figure 3.14: In an illuminated semiconductor the occupations of the conduction band (C) and the valence band (V) are described by different Fermi distributions $f_{\rm C}$ and $f_{\rm V}$.

It then follows that

$$n_{\rm e}n_{\rm h} = N_{\rm C}N_{\rm V} \exp\left(-\frac{\varepsilon_{\rm C} - \varepsilon_{\rm V}}{kT}\right) \exp\left(\frac{\varepsilon_{\rm FC} - \varepsilon_{\rm FV}}{kT}\right) \tag{3.29}$$

or

$$n_{\rm e}n_{\rm h} = n_{\rm i}^2 \exp\left(\frac{\varepsilon_{\rm FC} - \varepsilon_{\rm FV}}{kT}\right) \,. \tag{3.30}$$

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 on page 68.

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, which are exchanged as well.

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

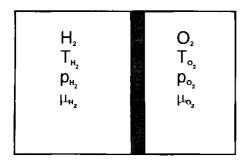


Figure 3.15: Two compartments separated by a piston contain hydrogen H_2 and oxygen O_2 at different temperatures T, pressures p and chemical potentials μ .

energy forms are in all cases products of "intensive" variables and "extensive", quantity-like, variables. In order to distinguish between these, let us imagine two identical systems which we put together to form a new system, twice as large. The intensive variables like 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 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 is hydrogen and in the other is 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 non-equilibria) 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 non-equilibrium" 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 which 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 φ dQ, 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 which is not accompanied by entropy. This part is the Free Energy $F(T, V, N_i, Q, \ldots) = E - TS$. 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

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

Due 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. $\eta_e = \mu_c - e \phi$ is the electrochemical potential of the electrons. For electrons in electrochemical equilibrium, it has the same value everywhere. $\eta_h = \mu_h + e \phi$ 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, which a solar cell delivers to a load with dN electrons and holes is then

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

The question now remains of how to relate the electrochemical potentials to the already known quantities. In Eq. (3.20) we found the mean energy of the electrons to have the value $\langle \varepsilon_e \rangle = \varepsilon_C + (3/2)kT$ 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

$$E_{\rm e} = TS_{\rm e} - p_{\rm e}V_{\rm c} + \eta_{\rm e}N_{\rm e}$$

and the mean energy per electron

$$E_{\rm e}/N_{\rm e} = \langle \varepsilon_{\rm e} \rangle = T\sigma_{\rm e} - \frac{p_{\rm e}V_{\rm e}}{N_{\rm e}} + \eta_{\rm e} \stackrel{!}{=} \varepsilon_{\rm C} + (3/2)kT. \qquad (3.33)$$

Similarly, for holes

$$E_{\rm h}/N_{\rm h} = \langle \varepsilon_{\rm h} \rangle = T \sigma_{\rm h} - \frac{p_{\rm h} V_{\rm h}}{N_{\rm h}} + \eta_{\rm h} \stackrel{!}{=} -\varepsilon_{\rm V} + (3/2)kT. \tag{3.34}$$

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 for ideal gases,

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

Substituting the effective mass m^* of the electrons and holes from Eq. (3.13) for m we find

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

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

$$pV = NkT. (3.37)$$

With Eq. (3.36) and Eq. (3.37), we find from Eq. (3.33) for electrons

$$\langle \varepsilon_{e} \rangle = \varepsilon_{C} + \frac{3}{2}kT = kT \left[\frac{5}{2} + \ln \left(\frac{N_{C}}{n_{e}} \right) \right] - kT + \eta_{e}$$

$$\varepsilon_{C} - \eta_{e} = kT \ln \left(\frac{N_{C}}{n_{e}} \right)$$

which leads to

$$n_{\rm e} = N_{\rm C} \exp\left[\frac{-(\varepsilon_{\rm C} - \eta_{\rm e})}{kT}\right] . \tag{3.38}$$

This is identical with Eq. (3.27), and we can now identify the electrochemical potential of the electrons with their Fermi energy

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

For holes a similar result applies, thus

$$-\varepsilon_{\rm V} - \eta_{\rm h} = kT \ln \left(\frac{N_{\rm V}}{n_{\rm h}} \right) \tag{3.40}$$

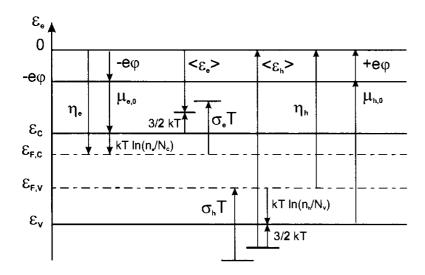


Figure 3.16: The energy forms of electrons and holes.

and

$$n_{\rm h} = N_{\rm V} \exp\left(\frac{\eta_{\rm h} + \varepsilon_{\rm V}}{kT}\right) \,. \tag{3.41}$$

Comparison with Eq. (3.28) shows that

$$\eta_{h} = -\varepsilon_{FV}$$
 (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 which 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:

$$\eta_{e}(x) + \eta_{h}(x) = \mu_{e}(x) - e\varphi(x) + \mu_{h}(x) + e\varphi(x) = \mu_{e}(x) + \mu_{h}(x).$$
 (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 which must be expended in order to excite an electron from a bound state in a semiconductor or metal thermally, i. e., by thermionic 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 $-e\phi$ 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 $\exp(-\phi/kT)$. This emission current is proportional to the concentration of electrons $n_{e,free}$ not bound to the semiconductor material, but still belonging to the same Fermi distribution, which have an energy of at least $\varepsilon_e = -e\phi$. 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 Eq. (3.12)

$$n_{\rm e,free} \propto \exp\left(-\frac{-e\phi - \varepsilon_{\rm F}}{kT}\right) = \exp\left(-\frac{-e\phi - \mu_{\rm e} + e\phi}{kT}\right) = \exp\left(\frac{\mu_{\rm e}}{kT}\right)$$

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

$$\mu_{\rm e} = \mu_{\rm e,0} + kT \ln \left(\frac{n_{\rm e}}{N_{\rm C}}\right) . \tag{3.44}$$

For semiconductors, the property which 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 Eq. (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. Due 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 which 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 phonons or even photons.

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 transitions 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 $p_{\gamma} = \hbar \omega/c$,

$$p_{\gamma} = p_{\rm e} + p_{\rm h} \approx 0$$
, and thus $p_{\rm e} = -p_{\rm h}$.

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

$$\varepsilon_{\rm e} = \varepsilon_{\rm C} + \frac{p_{\rm e}^2}{2m_{\rm e}^*}$$
 and $\varepsilon_{\rm h} = -\varepsilon_{\rm V} + \frac{p_{\rm h}^2}{2m_{\rm h}^*}$, (3.45)

leads to

$$\hbar\omega = \varepsilon_{\rm C} - \varepsilon_{\rm V} + \frac{p_{\rm e}^2}{2m_{\rm h}^*} + \frac{p_{\rm h}^2}{2m_{\rm h}^*} \ . \label{eq:epsilon}$$

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

$$\hbar\omega = \varepsilon_{\rm G} + \frac{p^2}{2} \left(\frac{1}{m_{\rm e}^*} + \frac{1}{m_{\rm h}^*} \right) = \varepsilon_{\rm G} + \frac{p^2}{2m_{\rm comb}} . \tag{3.46}$$

 $m_{\rm comb} = (m_{\rm e}^* m_{\rm h}^*)/(m_{\rm c}^* + m_{\rm h}^*)$ is the so-called combined mass.

In Eq. (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 (Eq. (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 Eq. (3.7), the energy-momentum relation in Eq. (3.46) results in the so-called combined density of states for direct optical transitions,

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

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,

$$\frac{\mathrm{d}j_{\gamma}(\hbar\omega)}{\mathrm{d}x} = -\alpha(\hbar\omega)j_{\gamma}(\hbar\omega) \tag{3.48}$$

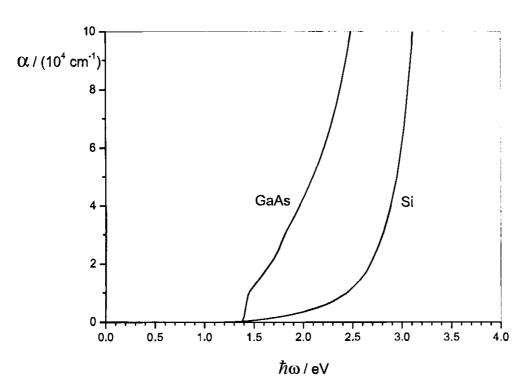


Figure 3.17: Absorption coefficient α of the "direct" semiconductor gallium arsenide and the "indirect" semiconductor silicon.

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

Integration of Eq. (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 $\hbar\omega < \epsilon_G$ are not absorbed by electron-hole generation and $\alpha(\hbar\omega < \epsilon_G) = 0$. Photons of this energy are either reflected or transmitted. For $\hbar\omega > \epsilon_G$, α shows the steep rise in accordance with the theoretical expectations $\propto (\hbar\omega - \epsilon_G)^{1/2}$ up to values of 10^4 /cm. 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 $\hbar\omega = \epsilon_G = 1.4\,\mathrm{eV}$, an exponential increase is superimposed over the square-root dependence. This is known as the Urbach 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 ϵ , ϵ is called the penetration depth of the photons. Due 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 μ m in order to absorb the absorbable part of the solar spectrum with photon energies $\hbar\omega > \epsilon_G$.

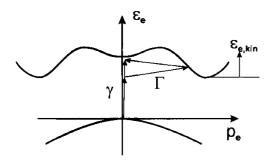


Figure 3.18: Energy ε_e of electron states in the conduction and valence bands between which transitions are possible with the absorption of a photon γ and the simultaneous absorption or emission of a photon Γ .

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 ,

$$\gamma + \Gamma \rightarrow e + h$$

$$p_{\gamma} + p_{\Gamma} = p_{e} + p_{h}$$

$$\hbar \omega + \hbar \Omega = \varepsilon_{e} + \varepsilon_{h} ,$$
(3.50)

or the phonon Γ can be emitted

$$\gamma \to e + h + \Gamma$$

$$p_{\gamma} = p_{e} + p_{h} + p_{\Gamma}$$

$$\hbar \omega = \varepsilon_{e} + \varepsilon_{h} + \hbar \Omega . \tag{3.51}$$

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 $\epsilon_{e,kin}$ which 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 $\epsilon_{e,kin}=0$ for a transition to the minimum of the conduction band and the highest possible value is $\epsilon_{e,kin}=\hbar\omega\pm\hbar\Omega-\epsilon_G$ for a transition from the maximum of the valence band to the conduction band.

$$\alpha(\hbar\omega) \propto \int_{0}^{\hbar\omega\pm\hbar\Omega-\epsilon_{G}} D_{C}(\epsilon_{e,kin}) D_{V}(\hbar\omega\pm\hbar\Omega-\epsilon_{G}-\epsilon_{e,kin}) d\epsilon_{e,kin}$$

With the energy dependence of the densities of states of Eq. (3.7) the integration yields

$$\alpha(\hbar\omega) \propto (\hbar\omega - \varepsilon_{\rm G} \pm \hbar\Omega)^2$$
 (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.

Due to the small value of α for indirect transitions, the penetration depth for photons in an indirect semiconductor is large. In order 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 100μ 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

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

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 , 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 towards the right (to larger x) is greater than that flowing in from the left (from small x), expressed by 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 due to 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

$$\operatorname{div} j_{\gamma} = \frac{\mathrm{d} j_{\gamma}}{\mathrm{d} x} = -R_{\gamma} = -\alpha j_{\gamma} .$$

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 transmission of a body of thickness d

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

the total absorbed photon current becomes

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

The absorptivity $a = (1-r)[1 - \exp(-\alpha d)]$ 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

$$G_{\rm h} = G_{\rm e} = R_{\gamma} = \alpha j_{\gamma}(x)$$
.

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 $\epsilon_{kin} \geq \epsilon_G$. 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. Due 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 $\hbar\omega > 3\epsilon_G$. This fundamental possibility requires a band structure for the semiconductor which ensures the conservation of energy and momentum. In real solar cells with band gaps $\geq 1 \text{eV}$, 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 labelled with a superscript 0. Photons of different energies contribute according to the value of the absorption coefficient $\alpha(\hbar\omega)$:

$$G_{\rm e}^{0} = G_{\rm h}^{0} = R_{\gamma}^{0} = \int_{0}^{\infty} \alpha(\hbar\omega) \, \mathrm{d}j_{\gamma}(\hbar\omega)$$

$$= \frac{\Omega}{4\pi^{3}\hbar^{3}c^{2}} \int_{0}^{\infty} \frac{\alpha(\hbar\omega)(\hbar\omega)^{2}}{\exp\left(\frac{\hbar\omega}{kT_{0}}\right) - 1} \, \mathrm{d}\hbar\omega \,. \tag{3.54}$$

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

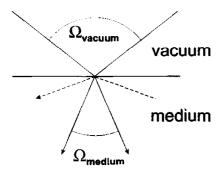


Figure 3.19: In accordance with the law of refraction at the interface between two media, only photons from a smaller solid angle in the medium with the larger refractive index, are exchanged with the other medium. Photons for which the momentum lies in the solid angle range inaccessible from outside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

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 Eq. (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 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 re-emitted inside the medium and contribute to the generation and recombination rates which then follow from Eq. (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 non-radiative 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 non-radiative 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 processes in which electrons and holes are annihilated in a reaction known as recombination. From the energy set free in this reaction, photons or phonons are produced or both are produced simultaneously. In thermal and chemical equilibrium with the radiation in the environment, in which $n_e^0 n_h^0 = n_i^2$, 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).

$$e + h \rightarrow \gamma$$
.

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

$$G_{\gamma} = R_{\rm e} = R_{\rm h} = B n_{\rm e} n_{\rm h} \ . \tag{3.55}$$

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

$$G_{\gamma}^{0} = R_{\rm e}^{0} = R_{\rm h}^{0} = R_{\gamma}^{0} = G_{\rm e}^{0} = G_{\rm h}^{0} = B n_{\rm e}^{0} n_{\rm h}^{0} . \tag{3.56}$$

It follows immediately from this equation that the product $n_{\rm e}^0 n_{\rm h}^0$ 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 $G_{\rm e}^0$ from Eq. (3.54)

$$G_{\gamma}^{0} = R_{\gamma}^{0} = G_{e}^{0} = G_{h}^{0} = \frac{\Omega}{4\pi^{3}\hbar^{3}c^{2}} \int_{0}^{\infty} \frac{\alpha(\hbar\omega)(\hbar\omega)^{2}}{\exp\left(\frac{\hbar\omega}{kT_{0}}\right) - 1} d\hbar\omega.$$
(3.57)

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

$$G_{\gamma}^{0}(\mathrm{Si}) = 3 \times 10^{5} \frac{\gamma}{\mathrm{cm}^{3} \mathrm{s}} = B n_{\mathrm{i}}^{2}.$$

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

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

In a state of non-equilibrium 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 $\langle \varepsilon_{kin} \rangle = (3/2)kT_0$. In a state of non-equilibrium 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

$$G_{\gamma} = G_{\gamma}^{0} \frac{n_{\rm e} n_{\rm h}}{n_{\rm i}^{2}} \ . \tag{3.58}$$

With reference to the quasi-Fermi energies in Eq. (3.29) we can now write

$$G_{\gamma} = G_{\gamma}^{0} \exp\left(\frac{\varepsilon_{FC} - \varepsilon_{FV}}{kT}\right) . \tag{3.59}$$

Due 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 which a solar cell must 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 Eq. (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 semi-conductor 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 $\mathrm{d} j_{\gamma}^{0}(\hbar\omega)$ 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:

$$\mathrm{d}j_{\gamma,\mathrm{em}}^0 = \mathrm{d}j_{\gamma,\mathrm{abs}}^0 = a(\hbar\omega)\,\mathrm{d}j_{\gamma}^0 \;.$$

In non-equilibrium of homogeneous electron and hole distributions, the emitted photon current density is larger by the same factor $n_{\rm e}n_{\rm h}/n_{\rm i}^2$ by which the emission rate in Eq. (3.58) is increased, so that

$$dj_{\gamma,\text{em}}(\hbar\omega) = a(\hbar\omega) \frac{n_{\text{e}}n_{\text{h}}}{n_{\text{i}}^2} dj_{\gamma}^0(\hbar\omega) . \qquad (3.60)$$

Here $a(\hbar\omega)$ is the absorptivity which, without considering multiple internal reflection, is $a(\hbar\omega) = [1 - r(\hbar\omega)]\{1 - \exp[-\alpha(\hbar\omega)d]\}$. The absorptivity is always $\leq 1 - r$ and for a thickness $d \gg 1/\alpha$, a = 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. While 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 which 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 Eq. (3.60) in terms of the quasi-Fermi energies, we obtain

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

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 (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 absorptivity 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. Eq. (3.61) is therefore only approximately true. An exact derivation is given in Section 3.7.

In place of Eq. (3.61) the exact result is 1

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

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

3.6.2 Non-radiative recombination

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

¹P. Würfel, J. Phys. C 15 (1982) 3967.

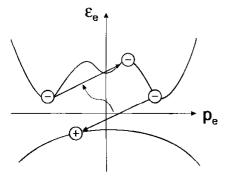


Figure 3.20: In Auger recombination, the energy set free by recombination of an electron-hole pair is absorbed by a free carrier (an electron in this figure) and is subsequently dissipated by generating phonons in collisions with the lattice.

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 which 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

$$R_{\text{Aug,e}} = C_{\text{e}} n_{\text{e}}^2 n_{\text{h}}$$

and is large for strong n-doping.

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

$$R_{\mathrm{Aug,h}} = C_{\mathrm{h}} n_{\mathrm{e}} n_{\mathrm{h}}^2.$$

This is relatively large in materials with strong p-doping.

$$R_{\text{Aug}} = n_{\text{e}} n_{\text{h}} (C_{\text{e}} n_{\text{e}} + C_{\text{h}} n_{\text{h}}) .$$

In silicon the constants have the value

$$C_{\rm e}({
m Si}) \approx C_{\rm h}({
m Si}) \approx 1 \times 10^{-30} \, {
m cm}^6 \, .$$

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 $n_{\rm e}^0 n_{\rm h}^0 = n_{\rm i}^2$ and the generation of electrons by photons is compensated by radiative recombination, non-radiative recombination processes are also compensated by the reverse processes, i. e., non-radiative 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 which offer electron states with energies approximately in the middle of the forbidden zone play the most important role. They

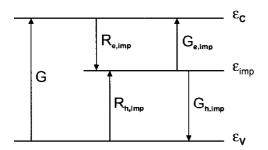


Figure 3.21: Electrons and holes produced by the absorption of photons with the generation rate G are captured by impurities with the rates $R_{e,imp}$ and $R_{h,imp}$ and they are emitted thermally with the rates $G_{e,imp}$ and $G_{h,imp}$ from the impurities back into the bands.

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 non-radiative 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 non-radiative 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.²

As an example, we will examine a single type of acceptor-like impurity, 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 distribute 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 $n_{\rm e}$ (in the conduction band) changes in time due to the rate of generation $G_{\rm e}=G_{\rm h}=G=a(\hbar\omega)j_{\gamma}/d$ by the absorption of photons averaged over the thickness d and it changes by the rate $R_{\rm e,imp}$ of capture into impurities occupied by holes and by the rate of thermal emission $G_{\rm e,imp}$ from impurities occupied by electrons. The contribution of the 300 K radiation from the environment and radiative recombination are neglected. The continuity equations for the changes in the concentrations of electrons (in the conduction band) $n_{\rm e}$, of holes (in the valence band) $n_{\rm h}$ and of occupied impurity states, i. e., of electrons trapped in impurities $n_{\rm e,imp}$ as a function of the generation and recombination rates shown in Figure 3.21 are

$$\frac{\partial n_{\rm e}}{\partial t} = G - R_{\rm e,imp} + G_{\rm e,imp} \tag{3.63}$$

$$\frac{\partial n_h}{\partial t} = G - R_{h,imp} + G_{h,imp} \tag{3.64}$$

$$\frac{\partial n_{\rm e,imp}}{\partial t} = R_{\rm e,imp} - R_{\rm h,imp} - G_{\rm e,imp} + G_{\rm h,imp}$$
(3.65)

²W. Shockley, W. T. Read, *Phys. Rev.* **87** (1952) 835. R. N. Hall, *Phys. Rev.* **83** (1951) 228.

These are three equations for the three unknown concentrations n_e , n_h and $n_{e,imp}$. 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.

$$e(n_{D^+} - n_e - n_A^- + n_h - n_{e,imp}) = 0.$$
 (3.66)

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, $(-n_{\rm e,imp})$ must be replaced by $(n_{\rm imp} - n_{\rm e,imp})$, where $n_{\rm imp}$ is the concentration of the impurities.

We now have to determine the various generation and recombination rates. The rate of recombination $R_{e,imp}$ 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

$$R_{\text{e.imp}} = \sigma_{\text{e}} v_{\text{e}} n_{\text{e}} n_{\text{h.imp}} , \qquad (3.67)$$

where $n_{h,imp}$ 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} cm².

Similarly, the recombination rate for holes is given by

$$R_{\rm h,imp} = \sigma_{\rm h} \nu_{\rm h} n_{\rm h} n_{\rm e,imp} . \tag{3.68}$$

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 (3.67) and (3.68) the occupation of impurity states, and as a result the concentration of electrons $n_{e,imp}$ or holes $n_{h,imp}$ 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 $n_{e,imp}$ or holes $n_{h,imp} = n_{imp} - n_{e,imp}$ in the impurities. The generation rate of electrons is

$$G_{\rm e,imp} = \beta_{\rm e} n_{\rm e,imp} . \tag{3.69}$$

In the same way, the generation rate of holes is

$$G_{h,imp} = \beta_h n_{h,imp} = \beta_h (n_{imp} - n_{e,imp}). \tag{3.70}$$

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 Eqs. (3.63), (3.67) and (3.69), in the absence of external excitation (G = 0) and in a steady state

$$\frac{\partial n_{\rm e}}{\partial t} = -\sigma_{\rm e} \nu_{\rm e} n_{\rm e} (n_{\rm imp} - n_{\rm e,imp}) + \beta_{\rm e} n_{\rm e,imp} = 0.$$
 (3.71)

From Eq. (3.71) we can calculate the density of impurities occupied by electrons in the dark, which we already know is given by the equilibrium Fermi distribution,

$$n_{e, imp} = n_{imp} \frac{1}{\frac{\beta_e}{\sigma_e \nu_e n_e} + 1} = n_{imp} \frac{1}{\exp\left(\frac{\varepsilon_{imp} - \varepsilon_F}{kT}\right) + 1}.$$
 (3.72)

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

$$\beta_{\rm e} = \sigma_{\rm e} \nu_{\rm e} N_{\rm C} \exp\left(-\frac{\varepsilon_{\rm C} - \varepsilon_{\rm imp}}{kT}\right) , \qquad (3.73)$$

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

$$\beta_{h} = \sigma_{h} \nu_{h} N_{V} \exp\left(-\frac{\varepsilon_{imp} - \varepsilon_{V}}{kT}\right) . \tag{3.74}$$

The capture cross-sections and the coefficients of emission are likely to depend on the energy distribution of the free charge carriers. Due 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 Eqs (3.63) to (3.66), 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 Eq. (3.65) for the density of impurities occupied by electrons

$$n_{\text{e,imp}} = \frac{n_{\text{imp}} \left\{ \sigma_{e} \nu_{e} n_{e} + \sigma_{h} \nu_{h} N_{V} \exp \left[-\frac{(\varepsilon_{\text{imp}} - \varepsilon_{V})}{kT} \right] \right\}}{\sigma_{e} \nu_{e} \left\{ n_{e} + N_{C} \exp \left[-\frac{(\varepsilon_{C} - \varepsilon_{\text{imp}})}{kT} \right] \right\} + \sigma_{h} \nu_{h} \left\{ n_{h} + N_{V} \exp \left[-\frac{(\varepsilon_{\text{imp}} - \varepsilon_{V})}{kT} \right] \right\}} . (3.75)$$

This result, which is not important by itself, inserted into Eq. (3.63) gives

$$G = \frac{n_{\rm e}n_{\rm h} - n_{\rm i}^2}{\frac{n_{\rm e} + N_{\rm C} \exp\left[-(\varepsilon_{\rm C} - \varepsilon_{\rm imp})/kT\right]}{n_{\rm imp}\sigma_{\rm h}\nu_{\rm h}} + \frac{n_{\rm h} + N_{\rm V} \exp\left[-(\varepsilon_{\rm imp} - \varepsilon_{\rm V})/kT\right]}{n_{\rm imp}\sigma_{\rm e}\nu_{\rm e}}} . \tag{3.76}$$

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 $n_{\rm imp}\sigma_h v_h$ is the rate per hole of hole capture if all impurities are occupied by electrons. Its reciprocal value $\tau_{\rm h,min} = 1/(n_{\rm imp}\sigma_h v_h)$ 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, $\tau_{\rm e,min} = 1/(n_{\rm imp}\sigma_{\rm e}v_{\rm e})$ 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 splitting of the Fermi energies is the main objective of a solar cell, it is worthwhile to express Equation (3.76) in terms of the Fermi energies. We use the relations

$$n_{\rm e} = n_{\rm i} \exp\left(-\frac{\varepsilon_{\rm i} - \varepsilon_{\rm FC}}{kT}\right)$$
 (3.77)

$$n_{\rm h} = n_{\rm i} \exp\left(-\frac{\varepsilon_{\rm FV} - \varepsilon_{\rm i}}{kT}\right)$$
 (3.78)

$$N_{\rm C} \exp\left(-\frac{\varepsilon_{\rm C} - \varepsilon_{\rm imp}}{kT}\right) = n_{\rm i} \exp\left(-\frac{\varepsilon_{\rm i} - \varepsilon_{\rm imp}}{kT}\right)$$
 (3.79)

$$N_{\rm V} \exp\left(-\frac{\varepsilon_{\rm imp} - \varepsilon_{\rm V}}{kT}\right) = n_{\rm i} \exp\left(-\frac{\varepsilon_{\rm imp} - \varepsilon_{\rm i}}{kT}\right)$$
 (3.80)

where ε_i is the value of the Fermi energy in an intrinsic semiconductor in the dark, and find from (3.76)

$$G = \frac{n_{i} \left[\exp \left(\frac{\varepsilon_{FC} - \varepsilon_{FV}}{kT} \right) - 1 \right]}{\tau_{h,min} \left[\exp \left(\frac{\varepsilon_{FC} - \varepsilon_{i}}{kT} \right) + \exp \left(\frac{\varepsilon_{imp} - \varepsilon_{i}}{kT} \right) \right] + \tau_{e,min} \left[\exp \left(\frac{\varepsilon_{i} - \varepsilon_{FV}}{kT} \right) + \exp \left(\frac{\varepsilon_{i} - \varepsilon_{imp}}{kT} \right) \right]}. (3.81)$$

For a good solar cell we want the difference in the Fermi energies, i.e., the numerator of (3.81), 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 (3.81) 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 ($\varepsilon_{imp} = \varepsilon_i$) 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 ($\varepsilon_{FC} - \varepsilon_i = \varepsilon_i - \varepsilon_{FV}$). In an homogeneous material the latter condition is fulfilled only if it is intrinsic. Doping, regardless of whether n-type or p-type, will increase the splitting of 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 $\varepsilon_{FC} - \varepsilon_{FV}$ for a constant generation rate G is determined from Eq. (3.81) and shown in Figure 3.22 as a function of the impurity level ε_{imp} . As expected, $\varepsilon_{FC} - \varepsilon_{FV}$ 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. Whether the material is p-doped or n-doped, however, is irrelevant. The type and level of

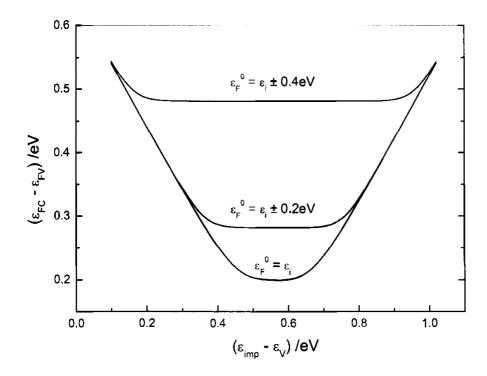


Figure 3.22: Difference in the Fermi energies as a function of the impurity energy ε_{imp} for a constant rate of generation $G = 10^{17} / (\text{cm}^3 \text{s})$ and for different doping levels. The doping concentration can be inferred from the difference between the Fermi energy in the dark ε_F^0 and the intrinsic Fermi energy ε_i from Eqs. (3.77) or (3.78). The impurities have a concentration of $n_{imp} = 10^{14} / \text{cm}^3$ with equal capture cross-sections for electrons and holes of $\sigma = 10^{-15} \, \text{cm}^2$, the velocity of electrons and holes is $v = 10^5 \, \text{m/s}$ resulting in a minimal lifetime of $\tau = 10^{-6} \, \text{s}$.

doping can be seen from the position of the Fermi energy ϵ_F^0 in the dark state. The greater the separation $|\epsilon_F^0 - \epsilon_i|$ 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 $\epsilon_F^0 - \epsilon_i > 0$ and for p-doped material $\epsilon_F^0 - \epsilon_i < 0$.

The concentrations of additional electrons and holes due to the external generation rate G can be determined analytically from Eq. (3.76) 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 Eq. (3.66) 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 Eq. (3.76) for capture cross-sections not too different for electrons and holes that

$$\Delta n = G \left\{ \tau_{\text{h,min}} \frac{n_{\text{e}}^0 + N_{\text{C}} \exp\left[-\frac{(\varepsilon_{\text{C}} - \varepsilon_{\text{imp}})}{kT}\right]}{n_{\text{e}}^0 + n_{\text{h}}^0} + \tau_{\text{e,min}} \frac{n_{\text{h}}^0 + N_{\text{V}} \exp\left[-\frac{(\varepsilon_{\text{imp}} - \varepsilon_{\text{V}})}{kT}\right]}{n_{\text{e}}^0 + n_{\text{h}}^0} \right\} . (3.82)$$

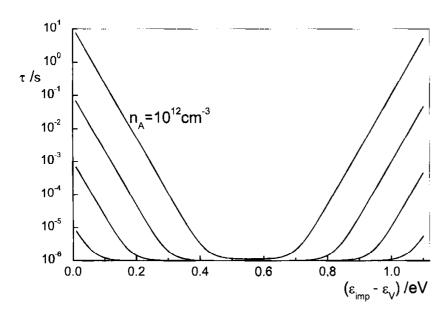


Figure 3.23: Lifetime τ of electrons and holes in p-doped silicon limited by impurity recombination as a function of the impurity energy $\varepsilon_{\rm imp}$, measured from the upper edge of the valence band. With each curve from the inside towards the outside, the acceptor density increases by a factor of 100 from $n_{\rm A} = 10^{12} / {\rm cm}^3$ to $n_{\rm A} = 10^{18} / {\rm cm}^3$. The generation rate is $G = 10^{17} / {\rm (cm}^3 \, {\rm s)}$, the impurities have a concentration of $n_{\rm imp} = 10^{14} / {\rm cm}^3$ with equal capture cross-sections for electrons and holes of $\sigma = 10^{-15} \, {\rm cm}^2$ and the velocity of electrons and holes is $v = 10^5 \, {\rm m/s}$ resulting in a minimal lifetime of $\tau_{\rm min} = 10^{-6} \, {\rm s}$.

Figure 3.23 shows the ratio $\Delta n/G$, which is also the mean lifetime τ of electrons and holes, as a function of $\varepsilon_{\rm imp}$ in p-doped silicon. $\tau_{\rm e,min} = \tau_{\rm h,min} = 10^{-6}\,\rm s$ was used for the calculation, resulting from an impurity density $n_{\rm imp} = 10^{14}\,\rm /cm^3$, a capture cross-section for electrons and holes of $\sigma = 10^{-15}\,\rm cm^2$ and a velocity of $v = 10^5\,\rm m/s$. 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 $\varepsilon_{\rm imp}$ 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 re-emission 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 splitting 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

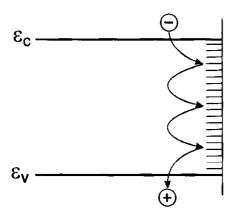


Figure 3.24: Recombination via surface states continuously distributed over energy in the energy gap of a semiconductor.

increasing splitting 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 splitting 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, then the splitting 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 Eqs. (3.63) to (3.66). But as we have seen in Figures 3.22 and 3.23, impurity states with energy in the middle of the band gap dominate the recombination properties.

Surface recombination

States with energies in the forbidden zone which are very effective mediators for non-radiative recombination are found in large concentrations on the surface of a material, where neighbouring atoms are no longer available to the lattice atoms and where impurity molecules (e. g., O_2 or H_2O) are adsorbed. These so-called surface states often have a continuous distribution over energy in the forbidden zone, as illustrated in Figure 3.24. The effect of the surface states on non-radiative 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

$$R_{\rm s.e} = \sigma_{\rm s.e} v_{\rm e} n_{\rm s.h} n_{\rm e} . \tag{3.83}$$

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 a velocity and is known as the surface recombination velocity

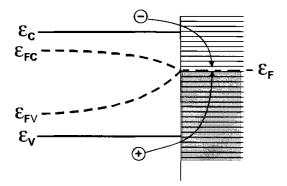


Figure 3.25: The continuous distribution of a high density of states around the Fermi energy in a metal results in a high interface recombination velocity at a semiconductor—metal contact.

 $v_{R,e}$ 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,

$$R_{\rm s,e} = v_{\rm R,e} \, n_{\rm e} \, .$$
 (3.84)

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 zone. Much less is known about the chemical nature of the surface states and their electron energies than for 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 cm/s 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 zone of the semiconductor is adjacent to the continuous distribution of states in the conduction band of the metal. To a good approximation, we can consider the surface recombination velocity at a metallic contact to be $v_{\rm R,metal} \approx \infty$. The resulting infinitely large recombination rate is compensated in the dark by a generation rate which is likewise infinitely large. Any additional generation rate at the semi-conductor/metal interface due to light absorption is negligibly small compared with the dark rate. As a result, even with additional generation, the concentrations of electrons and holes at a metallic contact do not differ from their equilibrium values in the dark state for which $n_{\rm e}^0 n_{\rm h}^0 = n_{\rm i}^2$. In the same way, the Fermi energies are not separated at a metal contact.

With silicon, good surfaces can be prepared by carefully controlled oxidation under clean room conditions. Surface state densities of $D_{\rm s} < 10^{10} / ({\rm cm}^2 {\rm \, eV})$ can be obtained at Si/SiO₂ interfaces, with recombination velocities of $v_{\rm R} \le 10 {\rm \, cm/s}$.

Such perfect surface passivation is possible only for Si/SiO_2 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 ($\varepsilon_G > 3 \, \text{eV}$) 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:

$$\frac{\partial n_{\rm e}}{\partial t} = G_{\rm c} - R_{\rm c} - {\rm div}\,j_{\rm e} \ .$$

We will restrict ourselves here to a spatially homogeneous system (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:

$$\frac{\partial (n_e^0 + \Delta n_e)}{\partial t} = G_e^0 + \Delta G_e - (R_e^0 + \Delta R_e) . \tag{3.85}$$

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

$$\frac{\partial \Delta n_{\rm e}}{\partial t} = \Delta G_{\rm e} - \Delta R_{\rm e} \ .$$

On suddenly switching off $\Delta G_{\rm e}$, we then have for radiative recombination

$$\frac{\partial \Delta n_{\rm e}}{\partial t} = -\Delta R_{\rm e} = -B n_{\rm h} \Delta n_{\rm e} \ . \tag{3.86}$$

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 re-absorption of the photons emitted by radiative recombination. The equation is therefore valid only for small semiconductor thicknesses. If we also consider the re-absorption 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 non-radiative 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 (3.86) 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 unfocussed solar radiation, this condition is fulfilled. Eq. (3.86) then yields

$$\Delta n_{\rm e}(t) = \Delta n_{\rm e}(0) \exp(-t/\tau_{\rm e,rad}) . \tag{3.87}$$

The characteristic time $\tau_{\rm e, rad} = 1/{\rm B}\,n_{\rm h}^0$ 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

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

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

$$\Delta R_{\rm e} = \Delta n_{\rm e}/\tau_{\rm e}$$
.

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

$$\Delta n_{\rm e}(t) = \Delta G_{\rm e} \tau_{\rm e} \left\{ 1 - \exp(-t/\tau_{\rm e}) \right\}. \tag{3.88}$$

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

$$\Delta n_{\rm e} = \Delta G_{\rm e} \tau_{\rm e} \ . \tag{3.89}$$

For non-radiative recombination in a steady state, once again $\Delta R_{\rm e} = \Delta G_{\rm e} = \Delta n_{\rm e}/\tau_{\rm e}$ and thus

$$\tau_{\rm e} = \frac{1}{\sigma_{\rm e} \, \nu_{\rm e} \, n_{\rm h,imp}} \,. \tag{3.90}$$

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 an n-conductor and in general for the respective minority charge carriers.

The density of the free charge carriers can be monitored by the microwave 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 $\Delta G_{\rm e} = \Delta G_{\rm h}$. 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 (3.66) then requires that $\Delta n_{\rm e} = \Delta n_{\rm h}$ and therefore $\tau_{\rm e} = \tau_{\rm h}$ 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

$$R_{\text{tot}} = \frac{\Delta n_{\text{e}}}{\tau_{\text{e,tot}}} = \sum_{i} R_{i} = \sum_{i} \frac{\Delta n_{\text{e}}}{\tau_{i}}$$

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 $\Delta G_{\rm e} = \Delta R_{\rm c} = \Delta n_{\rm c} \tau_{\rm c,tot}$. The additional concentration of electrons $\Delta n_{\rm e}$ in steady state is proportional to their lifetime.

The upper limit for the lifetime of electrons is $\tau_{e,rad}$ 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. The generalized Planck equation in (3.62) describes the emission of light by any material as a function of its absorptivity 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 $d\epsilon$ at energies ϵ_1 in the valence band and ϵ_2 in the conduction band involving photons with energy in the energy range $\hbar\omega\cdots\hbar\omega+d\hbar\omega$. The states are occupied according to two different Fermi functions $f(\epsilon_1)$ and $f(\epsilon_2)$. 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

$$dr_{\rm up}(\hbar\omega) = |M|^2 D_{12} f(\epsilon_1) \left[1 - f(\epsilon_2)\right] \frac{dj_{\gamma}(\hbar\omega)}{d\hbar\omega} d\hbar\omega , \qquad (3.91)$$

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

$$dr_{\text{stim}}(\hbar\omega) = |M|^2 D_{12} \left[1 - f(\epsilon_1)\right] f(\epsilon_2) \frac{dj_{\gamma}(\hbar\omega)}{d\hbar\omega} d\hbar\omega$$
(3.92)

and the rate of spontaneous emission, finally is

$$dr_{spont}(\hbar\omega) = |M|^2 \frac{c_0}{n} D_{\gamma}(\hbar\omega) D_{12} \left[1 - f(\epsilon_1)\right] f(\epsilon_2) d\hbar\omega , \qquad (3.93)$$

where the density of states for photons in the solid angle Ω in a medium with refractive index n is

$$D_{\gamma}(\hbar\omega) = \frac{\Omega n^3}{4\pi^3 \hbar^3 c_0^3} (\hbar\omega)^2 , \qquad (3.94)$$

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 non-absorbed photons. The net absorption rate at which photons disappear and which defines the absorption coefficient $\alpha_{12}(\hbar\omega)$ then is

$$dr_{abs}(\hbar\omega) = dr_{up}(\hbar\omega) - dr_{stim}(\hbar\omega)$$

$$= |M|^2 D_{12} [f(\varepsilon_1) - f(\varepsilon_2)] dj_{\gamma}(\hbar\omega) = \alpha_{12}(\hbar\omega) dj_{\gamma}(\hbar\omega) ,$$
(3.95)

resulting in an absorption coefficient

$$\alpha_{12}(\hbar\omega) = |M|^2 D_{12} \left[f(\varepsilon_1) - f(\varepsilon_2) \right] . \tag{3.96}$$

This relation is different from the derivation of 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. The important relation (3.96) allows us to replace the factor M and the combined density of states D_{12} in the spontaneous emission rate by the absorption coefficient $\alpha_{12}(\hbar\omega)$ and makes this treatment applicable to real materials for which only the absorption coefficient is known:

$$dr_{\text{spont}}(\hbar\omega) = \alpha_{12}(\hbar\omega) \frac{c_0}{n} D_{\gamma}(\hbar\omega) \frac{[1 - f(\epsilon_1)] f(\epsilon_2)}{f(\epsilon_1) - f(\epsilon_2)} d\hbar\omega . \tag{3.97}$$

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} :

$$f(\varepsilon_1) = \frac{1}{\exp\left(\frac{\varepsilon_1 - \varepsilon_{FV}}{kT}\right) + 1} \quad \text{and} \quad f(\varepsilon_2) = \frac{1}{\exp\left(\frac{\varepsilon_2 - \varepsilon_{FC}}{kT}\right) + 1}$$

and remembering that the energy difference $\varepsilon_2 - \varepsilon_1$ over which the transitions occur is equal to the photon energy $\hbar\omega$, (3.97) becomes

$$dr_{\text{spont}}(\hbar\omega) = \alpha(\hbar\omega) \frac{c_0}{n} D_{\gamma}(\hbar\omega) \frac{d\hbar\omega}{\exp\left(\frac{\hbar\omega - (\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}})}{kT}\right) - 1}.$$
 (3.98)

Since Eq. (3.98) 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 (3.98) 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 (3.98), if $\alpha(\hbar\omega)$ now is the absorption coefficient for all these transitions.

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

$$\operatorname{div} \operatorname{d} j_{\gamma}(x) = \operatorname{d} r_{\text{spont}}(x) - \alpha \operatorname{d} j_{\gamma}(x) . \tag{3.99}$$

We will treat only the simple case of a homogeneously excited semiconductor, in which dr_{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 $x \to \infty$. This equilibrium photon current inside the semiconductor is, from (3.99),

$$dj_{\gamma}(\infty) = dr_{\text{spont}} / \alpha . \tag{3.100}$$

Integration of Eq. (3.99) over a homogeneous semiconductor of finite thickness d yields

$$\mathrm{d}j_{\gamma}(\hbar\omega) = a(\hbar\omega) \frac{\Omega}{4\pi^{3}\hbar^{3}c_{0}^{2}} \frac{(\hbar\omega)^{2}\,\mathrm{d}\hbar\omega}{\exp\left[\frac{\hbar\omega - (\epsilon_{\mathrm{FC}} - \epsilon_{\mathrm{FV}})}{kT} - 1\right]}\,,\tag{3.101}$$

where the integration gives

$$a(\hbar\omega) = [1 - r(\hbar\omega)] [1 - \exp(-\alpha d)] \tag{3.102}$$

for the emissivity of a homogeneous semiconductor of thickness d equal to its absorptivity. Eq. (3.101) is the photon current emitted into the solid angle Ω outside the semiconductor after accounting for reflection at the surface.

Eq. (3.101) is a generalization of Kirchhoff's and Planck's laws and is valid for materials, which 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 $\varepsilon_{FC} - \varepsilon_{FV} = 0$. The difference in the Fermi energies $\mu_{eh} = \varepsilon_{FC} - \varepsilon_{FV}$ 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 which recombine, generate photons which are emitted. These photons carry the Free Energy of the electron-hole pairs, and $\mu_{\gamma} = \mu_{eh} = \varepsilon_{FC} - \varepsilon_{FV}$ 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,³ because their number n_{γ} is not conserved. According to this argument a minimum of the Free Energy F could only exist, if in $dF = \cdots + \mu_{\gamma} dn_{\gamma} + \cdots = 0$, the chemical potential $\mu_{\gamma} = 0$, since nonconservation of n_{γ} would allow $dn_{\gamma} \neq 0$. This argument is difficult to understand. Photons carry other quantities like energy, momentum and angular momentum which are strictly conserved. Furthermore, photons do not react with each other, 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

$$H_2 + \frac{1}{2}O_2 \leftrightarrow H_2O$$
.

³L. D.Landau, E. M. Lifschitz, Course of Theoretical Physics, Vol. V, Butterworth, Heinemann 1980.

Equilibrium exists, if

$$dF = \cdots + \mu_{H_2} dN_{H_2} + \mu_{O_2} dN_{O_2} + \mu_{H_2O} dN_{H_2O} + \cdots = 0.$$

With the changes in the particle numbers in the reaction $dN_{\rm H_2} = 2 \, dN_{\rm O_2} = -dN_{\rm H_2O}$ we find the equilibrium condition

$$\mu_{\rm H_2} + \frac{1}{2}\mu_{\rm O_2} = \mu_{\rm H_2O}$$
.

We conclude that non-conservation 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

$$e + h \rightleftharpoons \gamma$$
.

By the same argument as for the hydrogen-oxygen reaction, equilibrium between the electron-hole pairs and the emitted photons exists, if

$$\eta_e + \eta_h = \mu_e + \mu_h = \varepsilon_{FC} - \varepsilon_{FV} = \mu_{eh} = \mu_{\gamma}$$
.

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 which is thicker than the penetration depth of the photons, where photons are repeatedly emitted and re-absorbed before they reach the surface.

It must be emphasized that the validity of the generalized Planck equation (3.98) and (3.101) 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 $\varepsilon_{FC} - \varepsilon_{FV} > \hbar \omega$, where the denominator in (3.98) and (3.101) is negative. Under the same condition, the absorption coefficient in (3.96) is negative too, and the spontaneous emission rate in (3.97) and (3.98) 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 absorptivity a, increasing exponentially with the thickness d in Eq. (3.102). $\varepsilon_{FC} - \varepsilon_{FV} > \hbar \omega$ is also known as the condition for lasing. An equilibrium between the electron-hole pairs and the photons cannot exist if $\varepsilon_{FC} - \varepsilon_{FV} > \hbar \omega$ 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.

Eq. (3.98) describes the rate of radiative transitions. In a real material, there are also non-radiative transitions occurring in parallel and in addition. This does not invalidate Eq. (3.98), 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 Eq. (3.98) 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 Eq. (3.101) 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 absorptivity $a(\hbar\omega)$. 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 absorptivity and from it the absorption coefficient can be obtained from (3.101) 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 Eq. (3.101), if the absorptivity is known at one single photon energy allowing one to obtain the absorptivity $a(\hbar\omega)$ at all other photon energies in the measured luminescence spectrum.

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 $T_0 = 300$ 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 . 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), since 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 $T_{\rm eh} = T_S$. 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 $\mu_{\gamma} = \mu_{\rm eh} = 0$.

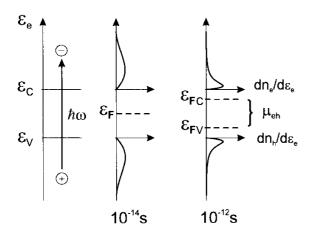


Figure 4.1: Electrons and holes generated by photons with energy $\hbar\omega$ lose energy by thermalization, which produces chemical energy per electron-hole pair $\mu_{\rm eh}$.

Physics of Solar Cells: From Principles to New Concepts. Peter Würfel Copyright © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 3-527-40428-7

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 $T_{\rm eh} = T_{\rm S}$ because of the spectrum, but $\mu_{\rm Y} = \mu_{\rm eh} < 0$.

Returning to maximum concentration, the electrons and holes are in thermal and chemical equilibrium with the sun with no splitting of the Fermi energies. In a rough estimate, the average energy of the electron-hole pairs is then $\varepsilon_G + 3kT_S$. From Eq. (3.29) it follows that, at $T_{\rm eh} = T_S$

$$\mu_{\rm eh}(T_{\rm S}) = \varepsilon_{\rm FC} - \varepsilon_{\rm FV} = \varepsilon_{\rm G} - kT_{\rm S} \ln \left(\frac{N_{\rm C}N_{\rm V}}{n_{\rm e}n_{\rm h}} \right) = 0$$

from which $kT_{\rm S} \ln (N_{\rm C}N_{\rm V}/n_{\rm e}n_{\rm h}) = \varepsilon_{\rm G}$.

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 $T_{\rm eh} = T_0$

$$\mu_{\rm eh}(T_0) = \varepsilon_{\rm FC} - \varepsilon_{\rm FV} = \varepsilon_{\rm G} - kT_0 \ln \left(\frac{N_{\rm C}N_{\rm V}}{n_{\rm e}n_{\rm h}} \right) \ .$$

Because of the constant concentrations of electrons and holes and neglecting the temperature dependence of the effective densities of states $N_{\rm C}$ and $N_{\rm V}$, this finally gives us an estimate of the chemical energy per electron-hole pair produced by the cooling process

$$\mu_{\rm eh}(T_0) = \varepsilon_{\rm FC} - \varepsilon_{\rm FV} = \varepsilon_{\rm G}(1 - T_0/T_{\rm S})$$
.

This relation looks like a Carnot efficiency, but it is different. For the efficiency we have to compare $\varepsilon_{FC} - \varepsilon_{FV}$ not with ε_G but with the energy $\varepsilon_G + 3kT_S$ 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 non-zero 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 it absorbs, but it does not fully compensate the temperature difference, since 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 non-concentrated sunlight and by 300 K background radiation under open circuit conditions, when there is only radiative recombination. Applying the generalized Planck law from Eq. (3.101), the comparison of the average energy $\langle \hbar \omega_{abs} \rangle$ of the absorbed photons which is invested for the generation of each electron-hole pair with the average energy $\langle \hbar \omega_{emit} \rangle$ 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.

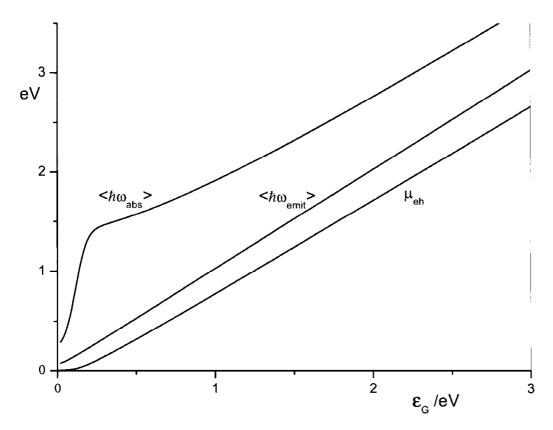


Figure 4.2: Average energy of absorbed photons $\langle \hbar \omega_{abs} \rangle$ from a non-concentrated AM0-spectrum and from the 300 K background radiation, the average energy of emitted photons $\langle \hbar \omega_{emit} \rangle$ and the chemical potential of electron-hole pairs μ_{eh} of a semiconductor 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 $\langle \hbar \omega_{abs} \rangle$ and $\langle \hbar \omega_{emit} \rangle$ is lost by thermalization of the electron-hole pairs.

The chemical energy per electron-hole pair is the sum of the electrochemical potentials of the electrons and the holes, and therefore

$$\eta_e + \eta_h = \mu_e + \mu_h = \epsilon_{FC} - \epsilon_{FV}$$
.

According to Eq. (3.29) this gives us

$$\mu_{\rm e} + \mu_{\rm h} = \varepsilon_{\rm FC} - \varepsilon_{\rm FV} = kT \ln \left(\frac{n_{\rm e} n_{\rm h}}{n_{\rm i}^2} \right)$$
 (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

$$\mu_{\rm e} + \mu_{\rm h} = kT \ln \left(1 + \frac{\Delta n_{\rm e}}{n_{\rm e}^0} \right) . \tag{4.2}$$

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 towards the sun as many photons, as much energy and as much entropy as it absorbs from the sun? As we have seen in the last 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 broad band semiconductor the electrons and holes are generated by almost monochromatic radiation, and only into those states in which they are after the thermalization, can an energy loss 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 absorptivity 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.
- 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 Eq. (2.34) and Eq. (3.101) this means that

$$dj_{\gamma,\text{emit}} = \frac{\Omega_{\text{cmit}}}{4\pi^3 h^3 c^2} \frac{\varepsilon_{\text{G}}^2 d\hbar\omega}{\exp\left(\frac{\varepsilon_{\text{G}} - \mu_{\text{eh}}}{kT_0}\right) - 1} = \frac{\Omega_{\text{abs}}}{4\pi^3 h^3 c^2} \frac{\varepsilon_{\text{G}}^2 d\hbar\omega}{\exp\left(\frac{\varepsilon_{\text{G}}}{kT_{\text{S}}}\right) - 1} = dj_{\gamma,\text{abs}}. \quad (4.3)$$

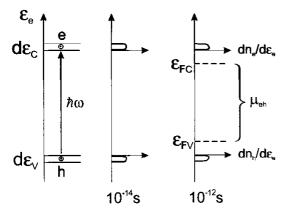


Figure 4.3: Electrons and holes generated in narrow energy ranges by monochromatic radiation have the same energy distribution after thermalization as before.

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

$$\mu_{\rm eh} = \mu_{\rm e} + \mu_{\rm h} = \varepsilon_{\rm G} \left(1 - \frac{T_0}{T_{\rm S}} \right) \,.$$
 (4.4)

The efficiency for the conversion of solar heat into chemical energy is

$$\eta = \frac{\mu_{\rm e} + \mu_{\rm h}}{\varepsilon_{\rm G}} = 1 - \frac{T_0}{T_{\rm S}} \tag{4.5}$$

This is the Carnot efficiency, a limiting value which 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 Eq. (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 Eq. (4.3), but also the energy currents which follow from the photon currents by multiplication with the photon energy which for absorption and emission have the same value $\hbar\omega = \epsilon_G$. For the absorbed and emitted radiation, the entropy per photon state is 1

$$\sigma_{\gamma} = k \left[(1 + f_{\gamma}) \ln(1 + f_{\gamma}) - f_{\gamma} \ln(f_{\gamma}) \right], \tag{4.6}$$

where f_{γ} is the Bose-Einstein distribution function

$$f_{\gamma} = \frac{1}{\exp\left(\frac{\hbar\omega - \mu_{\gamma}}{kT}\right) - 1} \tag{4.7}$$

The exponent in the distribution function has the same value for the sun and the semiconductor according to Eqs. (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 absorptivity $a(\hbar\omega = \epsilon_G) = 1$. In addition, absorption and emission involve the same number of photon states because of $\Omega_{abs} = \Omega_{emit}$. Under these conditions, the entropy absorbed with the photons from the sun is equal to the entropy emitted with the photons towards 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 Eq. (4.4), which is why we may call it a thermo-chemical equilibrium. The chemical energy produced by reversible, non-dissipative 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

¹C. Kittel, Elementary Statistical Physics, Wiley 1958.

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,

$$\frac{\partial n_{\rm e}}{\partial t} = G_{\rm e} - R_{\rm e} - \operatorname{div} dj_{\rm e} = 0 ,$$

we see that

$$\operatorname{div} \operatorname{d} j_{\mathrm{e}} = G_{\mathrm{e}} - R_{\mathrm{e}} \tag{4.8}$$

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 Eq. (4.8) over the volume of the semiconductor gives the total extracted electron current dj_e

$$\mathrm{d}j_{\mathrm{e}} = \int G_{\mathrm{e}} \mathrm{d}x - \int R_{\mathrm{e}} \, \mathrm{d}x .$$

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 re-absorption. Thus,

$$dj_e = dj_{\gamma, abs} - dj_{\gamma, emit}. (4.9)$$

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 $\mu_{\rm eh} = {\rm const}$, $j_{\rm eh}$ follows from the relations in Eq. (4.3) for the absorbed and emitted photon currents.

In an approximation to the generalized Planck law, when $\mu_{eh} < \varepsilon_G - 3kT$ allows us to neglect the "-1" in the denominator of the Planck law, we find according to Eq. (4.1)

$$dj_{e} = dj_{\gamma,abs} - dj_{\gamma}^{0} \frac{n_{e}n_{h}}{n_{i}^{2}} = dj_{\gamma,abs} - dj_{\gamma}^{0} \exp\left(\frac{\mu_{e} + \mu_{h}}{kT}\right) , \qquad (4.10)$$

where j_{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 constants and lifetimes.

For a given current of absorbed photons $dj_{\gamma,abs}$, the current dj_{eh} of extracted electrons and holes is 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

$$\mathrm{d}j_{\mu}=\mathrm{d}j_{\mathrm{eh}}\left(\mu_{\mathrm{e}}+\mu_{\mathrm{h}}\right)$$
.

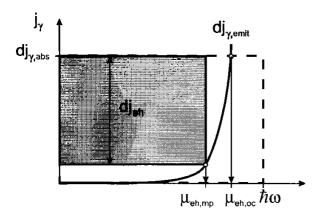


Figure 4.4: Current $dj_{\gamma,emit}$ of emitted photons as a function of the chemical energy $\mu_{eh} = \mu_e + \mu_h$ of the electron-hole pairs. The dark rectangle is the current of chemical energy extracted along with the current $dj_{eh} = dj_{\gamma,abs} - dj_{\gamma,emit}$ of extracted electron-hole pairs. It is produced from the current of absorbed energy represented by the grey rectangle.

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 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.

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.

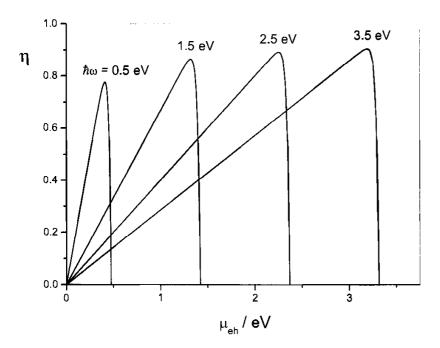


Figure 4.5: Efficiency η with which chemical energy is extracted as a function of the chemical potential μ_{ch} of the electron-hole pairs of monochromatic absorbers for different photon energies.

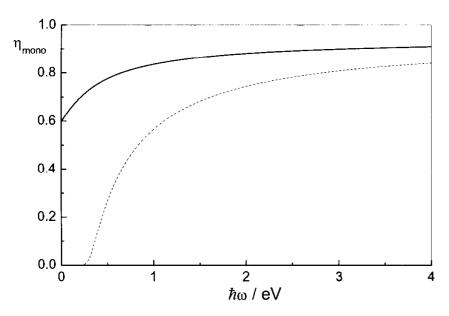


Figure 4.6: Monochromatic efficiency $\eta_{mono}(\varepsilon_G)$ for obtaining chemical energy $\mu_e + \mu_h$ as a function of the photon energy $\hbar\omega$ of fully concentrated (full line) and non-concentrated (broken line) monochromatic solar radiation.

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 Eq. (3.101) 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 non-concentrated 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 $\eta_{mono}(\epsilon_G)$ over the spectrum after weighting each interval $d\hbar\omega$ with its share of the absorbed energy current.

$$\eta = rac{\int_0^\infty \eta_{
m mono}(\epsilon_{
m G}) \hbar \omega \, {
m d} j_{\gamma, {
m abs}}}{\int_0^\infty \hbar \omega \, {
m d} j_{\gamma, {
m abs}}}$$

The total efficiency is 86% for fully concentrated radiation from a 5800 K sun and 67% for non-concentrated radiation from a solid angle of 6.8×10^{-5} .²

Although we were able to calculate with what efficiency chemical energy can be obtained from an illuminated semiconductor by an extraction of some of the generated electron—hole pairs, we still have no idea, how a device should look, which allows this extraction of electron—hole pairs and their energy. This is the subject of the next chapter.

²A. deVos, *Proceedings 5. E.C. Photovoltaic Solar Energy Conf.*, Athens 1983, p. 186.

5 Conversion of Chemical Energy into Electrical Energy

When electrons and holes are removed pair-wise 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 will find out what the forces are by which electrons and holes are driven 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 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 of particles are coupled during the exchange of electrons or holes, and consequently the forces acting on them are coupled, too, we will first treat the effects of these forces separately. For this, we will 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 not forget that the forces acting on the charge and on the quantity both act on the same particles and must therefore be added to a resultant force, which is then responsible for driving the particles.

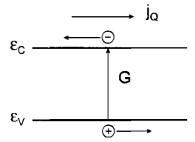


Figure 5.1: Electrons in the conduction band and holes in the valence band have to move in different directions to produce an electrical charge current j_Q .

Physics of Solar Cells: From Principles to New Concepts. Peter Würfel Copyright ©2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 3-527-40428-7

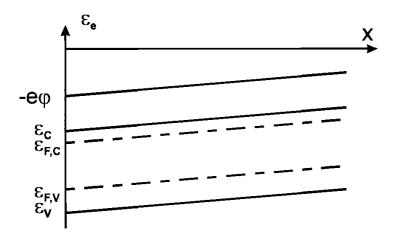


Figure 5.2: Electron energies in an electric field for uniform electron and hole concentrations.

5.1.1 Field current

The electric field $E = -\operatorname{grad} \varphi$ 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

$$\mu_{\rm e} = \mu_{\rm e,0} + kT \ln \left(\frac{n_{\rm e}}{N_{\rm C}} \right) ,$$

is independent of position and

$$\operatorname{grad}\mu_{e}=0$$
,

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 $-\varepsilon_{C}$ 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

$$j_{Q,f,i} = z_i e n_i \langle v_i \rangle . ag{5.1}$$

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 v_{th} .

Due 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 $a_i = z_i eE/m_i^*$. Because of the exponential distribution of the time intervals between collisions – a few particles undergo

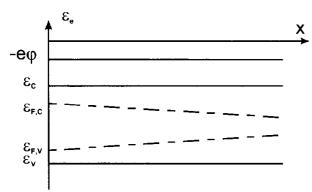


Figure 5.3: Electron energies for a position-dependent electron concentration in the absence of an electric field.

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

$$\langle v_i \rangle = z_i \frac{e}{m_i^*} \tau_{c,i} E . \tag{5.2}$$

 $b_i = e\tau_{c,i}/m_i^*$ is called the mobility of the particles of type *i*. The density of the field current is then

$$j_{O,f,i} = z_i^2 e n_i b_i E = \sigma_i E$$
.

Here $\sigma_i = z_i^2 e n_i b_i$ is the conductivity of the particles of type *i*. Making use of $E = -\operatorname{grad} \varphi$ we can also express the charge current in terms of the gradient of the electrical energy per particle $z_i e \varphi$

$$j_{Q,f,i} = -\frac{\sigma_i}{z_i e} \operatorname{grad}(z_i e \varphi) . \tag{5.3}$$

For electrons with $z_i = -1$ this is

$$j_{Q,f,e} = \frac{\sigma_e}{e} \operatorname{grad}(-e\varphi)$$
 (5.4)

and for holes with $z_i = +1$

$$j_{Q,f,h} = -\frac{\sigma_h}{e} \operatorname{grad}(e\varphi)$$
 (5.5)

5.1.2 Diffusion current

When the electrical potential is the same at every location, that is $\operatorname{grad} \varphi = 0$, a pure diffusion current flows if the concentration is non-uniform. 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 the concentration-dependent and therefore non-uniform component of the chemical potential of the electrons

$$\mu_{\rm e} - \mu_{\rm e,0} = kT \ln \frac{n_{\rm e}}{N_{\rm C}} ,$$

where $N_{\rm C}$ is the effective density of states for electrons.

In the usual representation by Fick's Law, the charge current due to a non-uniform concentration of particles of type i is given by

$$j_{O,d,i} = z_i e(-D_i \operatorname{grad} n_i) \tag{5.6}$$

with the diffusion coefficient D_i . 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

$$j_{Q,d,i} = -\frac{z_i e n_i b_i}{e} \operatorname{grad} \mu_i = -\frac{\sigma_i}{z_i e} \operatorname{grad} \mu_i .$$
 (5.7)

For electrons with $z_i = -1$, the diffusion current is given by

$$j_{Q,d,e} = \frac{\sigma_e}{e} \operatorname{grad} \mu_e \tag{5.8}$$

and for holes with $z_i = +1$, the diffusion current is

$$j_{Q,d,h} = -\frac{\sigma_h}{e} \operatorname{grad} \mu_h . ag{5.9}$$

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. To get the currents caused by the forces, they have to be multiplied by the conductivities. 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:

$$j_{Q,i} = -\frac{\sigma_i}{z_i e} \left\{ \operatorname{grad} \mu_i + \operatorname{grad}(z_i e \varphi) \right\}$$
 (5.10)

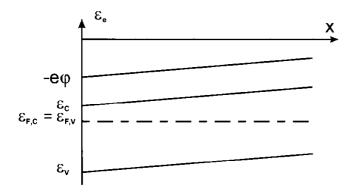


Figure 5.4: Constant Fermi energy in a semiconductor in which both an electric field and a concentration gradient exist.

or

$$j_{Q,i} = -\frac{\sigma_i}{z_i e} \operatorname{grad}(\mu_i + z_i e \varphi) = -\frac{\sigma_i}{z_i e} \operatorname{grad} \eta_i . \tag{5.11}$$

The gradient of the electrochemical potential, $\operatorname{grad} \eta_i = \operatorname{grad}(\mu_i + z_i \operatorname{e}\varphi)$, 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 which "feel" only the electric field and others which only contribute to the diffusion current.

Since the only charged particles moving in a semiconductor are electrons and holes, the total charge current is

$$j_Q = \frac{\sigma_e}{e} \operatorname{grad} \eta_e - \frac{\sigma_h}{e} \operatorname{grad} \eta_h , \qquad (5.12)$$

or, making use of the quasi Fermi energies, for which $\varepsilon_{FC} = \eta_e$ and $\varepsilon_{FV} = -\eta_h$,

$$j_Q = \frac{\sigma_e}{e} \operatorname{grad} \varepsilon_{FC} + \frac{\sigma_h}{e} \operatorname{grad} \varepsilon_{FV} .$$
 (5.13)

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 (5.13), the charge current j_Q is zero in spite of non-zero 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 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 Eq. (5.3) that

$$\frac{\mathrm{d}\rho_E}{\mathrm{d}t} = \frac{j_{Q,f,i}}{z_i \mathrm{e}} \left(-\operatorname{grad}(z_i \mathrm{e}\varphi) \right) = j_{Q,f,i}^2 / \sigma_i ,$$

and for the diffusion current it follows from Eq. (5.7) that

$$\frac{\mathrm{d}\rho_E}{\mathrm{d}t} = \frac{j_{Q,d,i}}{z_i \mathrm{e}} \left(-\operatorname{grad}\mu_i \right) = j_{Q,d,i}^2 / \sigma_i .$$

As expected, in both cases the rate of energy dissipation $d\rho_E/dt > 0$, even if the individual currents 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 grad $\varepsilon_F = 0$, 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 Eq. (5.13) 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 which 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 will 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 terminal and the right terminal

$$eV = \epsilon_{\text{F,left}} - \epsilon_{\text{F,right}} \neq 0$$
 .

We will 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.

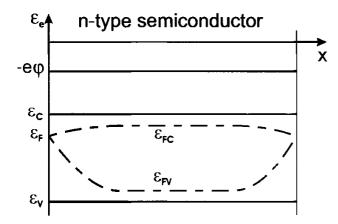


Figure 5.5: Distribution of the Fermi energies of a homogeneously illuminated n-type semiconductor with strong surface recombination on the left and on the right.

Figure 5.5 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 towards both surfaces, where they recombine. Since the absence of an external circuit implies that no charge current can flow, the particle currents of electrons and holes flowing to the same surface must be equal. Due to the larger conductivity of electrons for the n-conductor under discussion, it follows from Eq. (5.13) that the gradient of the Fermi energy for the conduction band ε_{FC} is smaller then 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.

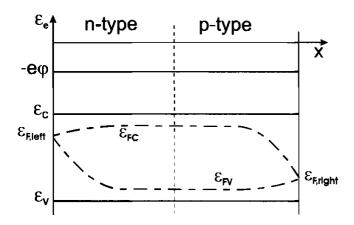


Figure 5.6: Distribution of the Fermi energies in a homogeneously illuminated pn structure.

The electrons flow towards the left when their electrochemical potential, the Fermi energy, $\eta_e = \epsilon_{FC}$, decreases towards the left. The holes flow towards the right when their electrochemical potential, $\eta_h = -\epsilon_{FV}$, decreases towards the right (when ϵ_{FV} increases towards the right). For the electron current towards the left and the hole current towards 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 which 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 due 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 $(\varepsilon_{FC} - \varepsilon_{FV})$ 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 will increase the gradients of the Fermi energies for the majority carriers and decrease 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 will 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 $Si \le 42 \,\mathrm{mA/cm^2}$), for the usual doping levels of $(10^{16}-10^{17})\,\mathrm{cm^{-3}}$ 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 in the terminals on the left and right. This produces the voltage V at the terminals of the solar cell:

$$V = \frac{1}{e} \int_{\text{left}}^{\text{right}} \text{grad}\,\varepsilon_{\text{FC}} \,dx \,. \tag{5.14}$$

Since $\varepsilon_{FC} = \varepsilon_{FV}$ in both terminals, one could also integrate over grad ε_{FV} .

That the voltage is given by the difference between the Fermi energies on the left and right with $eV = \Delta \epsilon_F$, results from the fact that $\operatorname{grad} \epsilon_F$ 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., $\operatorname{grad} \epsilon_F = 0$ within the voltmeter or $\Delta \epsilon_F = 0$ 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 $\Delta \phi$ 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.

5.3 Diffusion length of minority carriers

Our goal is that as many of the charge carriers, produced by the absorption of photons, as possible flow towards 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_c into a p-conductor as minority charge carriers, as in Figure 5.7. The electrons will move in the x-direction, in the y- and z-directions the system shall be homogeneous. The charge injected with the electrons will quickly be 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

$$\frac{\partial n_{\rm e}}{\partial t} = G_{\rm e} - R_{\rm e} - \text{div } j_{\rm e} = 0 \ . \tag{5.15}$$

For the particle diffusion current,

$$j_{\rm e} = -D_{\rm e} \frac{{
m d} n_{\rm e}}{{
m d} x}$$
 and ${
m div} j_{\rm e} = -D_{\rm e} \frac{{
m d}^2 n_{\rm e}}{{
m d} x^2}$.

In the p-conductor

$$G_{
m e} = G_{
m e}^0 = rac{n_{
m e}^0}{ au_{
m e}} \qquad {
m and} \qquad R_{
m e} = rac{n_{
m e}(x)}{ au_{
m e}} = rac{n_{
m e}^0}{ au_{
m e}} + rac{\Delta n_{
m e}(x)}{ au_{
m e}} \; .$$

It then follows from Eq. (5.15) that

$$-\frac{\Delta n_{\rm e}(x)}{\tau_{\rm e}} + D_{\rm e} \frac{{\rm d}^2 \Delta n_{\rm e}(x)}{{\rm d}x^2} = 0.$$
 (5.16)

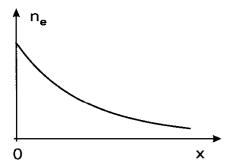


Figure 5.7: Distribution of electrons injected as minority charge carriers into a p-conductor.

The solution takes the form

$$\Delta n_{\rm e}(x) = \Delta n_{\rm e}(0) \exp\left(-x/L_{\rm e}\right). \tag{5.17}$$

The characteristic length $L_{\rm e}$ is the diffusion length (here for the electrons). Substituting this result into Eq. (5.16) we obtain

$$L_{\rm e} = \sqrt{D_{\rm e}\tau_{\rm e}} \ . \tag{5.18}$$

For electrons in pure silicon $D_e = 35 \,\mathrm{cm^2/s}$. For a lifetime of, e. g., $\tau_c = 10^{-6} \,\mathrm{s}$, $L_e = 60 \,\mu\mathrm{m}$. In very 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 will reach the n-conductor which are generated in the p-conductor at a distance not larger than their diffusion length.

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 will be explained in the next paragraph, other electrons will 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 paragraph, 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 will 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 will 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

$$\frac{\partial \rho_Q}{\partial t} = -\operatorname{div} j_Q . \tag{5.19}$$

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

$$\operatorname{div} D = \varepsilon \varepsilon_0 \operatorname{div} E = \rho_Q . \tag{5.20}$$

The electric field produces a charge current $j_Q = \sigma E$, from which we find

$$\operatorname{div} j_Q = \sigma \operatorname{div} E + E \operatorname{grad} \sigma. \tag{5.21}$$

Because of the assumed homogeneity, grad $\sigma = 0$. Substituting Eq. (5.21) into Eq. (5.19) and replacing div E from Eq. (5.20),

$$\frac{\partial \rho_Q}{\partial t} = -\sigma \operatorname{div} E = -\frac{\sigma}{\varepsilon \varepsilon_0} \rho_Q \tag{5.22}$$

with the solution

$$\rho_Q(t) = \rho_Q(0) \exp\left(\frac{-t}{\varepsilon \varepsilon_0/\sigma}\right) . \tag{5.23}$$

 $\varepsilon\varepsilon_0/\sigma$ 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 n-silicon, with a doping of $10^{17}/\text{cm}^3$, an electron mobility of $b_e = 1000\,\text{cm}^2/(\text{Vs})$ and $\epsilon = 12$, the dielectric relaxation time is $\epsilon \epsilon_0/\sigma = 6 \times 10^{-14}\,\text{s}$, 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. Due to their concentration gradients they diffuse from the surface into the interior. We will assume that the electrons have the larger mobility, larger diffusion coefficient and therefore move faster. Electrons and holes will 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, non-homogeneous 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

$$j_{e} = -D_{e} \operatorname{grad} n_{e} - \frac{\sigma_{e}}{e} E$$

$$j_{h} = -D_{h} \operatorname{grad} n_{h} + \frac{\sigma_{h}}{e} E.$$
(5.24)

Since the charge current vanishes,

$$j_O = eD_e \operatorname{grad} n_e - eD_h \operatorname{grad} n_h + (\sigma_e + \sigma_h)E = 0, \qquad (5.25)$$

we find that the electric field associated with ambipolar diffusion is given by

$$E = \frac{e}{\sigma_{\rm e} + \sigma_{\rm h}} \left(D_{\rm h} \operatorname{grad} n_{\rm h} - D_{\rm e} \operatorname{grad} n_{\rm e} \right) . \tag{5.26}$$

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 $\operatorname{grad} n_{\rm e} \approx \operatorname{grad} n_{\rm h} = \operatorname{grad} n$. The resulting approximation for the field in Eq. (5.26) substituted into Eq. (5.24) gives the particle currents in the field E

$$j_{e} = j_{h} = -\frac{D_{e}\sigma_{h} + D_{h}\sigma_{e}}{\sigma_{e} + \sigma_{h}} \operatorname{grad} n.$$
(5.27)

Since this has the form of the diffusion current equation, we can define

$$D_{\rm amb} = \frac{D_{\rm e}\sigma_{\rm h} + D_{\rm h}\sigma_{\rm e}}{\sigma_{\rm e} + \sigma_{\rm h}} \tag{5.28}$$

as the ambipolar diffusion coefficient.

5.6 Dember effect

To conclude this chapter, we will present an example to show how important it is to distinguish between electrical potential differences and electrochemical potential differences.

We will 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 Eq. (5.26) by the mobilities, apply the approximation that electrons and holes have the same distribution (grad $n_e \approx \text{grad} n_h$), multiply both numerator and denominator in Eq. (5.26) by $e(b_e + b_h)$ and find

$$E = \frac{kT}{e} \frac{b_{h} - b_{e}}{b_{e} + b_{h}} \frac{\operatorname{grad}(\sigma_{e} + \sigma_{h})}{\sigma_{e} + \sigma_{h}} = \frac{kT}{e} \frac{b_{h} - b_{e}}{b_{e} + b_{h}} \operatorname{grad} \ln(\sigma_{e} + \sigma_{h}).$$
 (5.29)

Using $E = -\operatorname{grad}\varphi$, 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,

$$\Delta \varphi_D = \varphi(0) - \varphi(\infty) = \frac{kT}{e} \frac{b_e - b_h}{b_e + b_h} \ln \left(\frac{\sigma_e(0) + \sigma_h(0)}{\sigma_e(\infty) + \sigma_h(\infty)} \right) . \tag{5.30}$$

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 then becomes positively charged and assumes a potential, which is more positive by $\Delta \phi_D$ than in the interior.

In order to correctly understand the meaning of the Dember effect, we will 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

$$\Delta \varphi_D = \frac{kT}{e} \ln \frac{n_e(0)}{n_e(\infty)} \ . \tag{5.31}$$

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

5.6 Dember effect

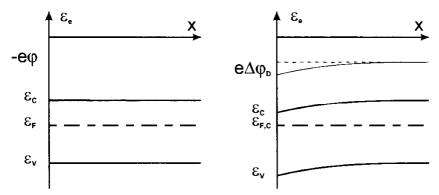


Figure 5.8: Distribution of potentials in the dark state (left) and on illumination of the free left surface of a semiconductor in which only electrons are mobile.

measurement. From Eq. (5.12) 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 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 Eq. (5.31) 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 Eq. (5.31) 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 want to be reserved 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

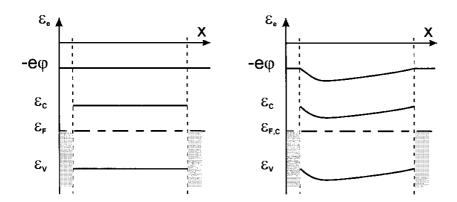


Figure 5.9: Distribution of potentials in a semiconductor with metal contacts in which only electrons are mobile, in the dark state (left) and with the left surface exposed to light (right).

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.

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 will 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. changes to the charge distribution can produce measurable voltages, which are, however, non-stationary. 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 photo-voltage. 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 non-stationary 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:

$$j_O = -e j_e + e j_h ;$$
 (5.32)

- For the particle currents:

$$j_{\rm e} = -\frac{\sigma_{\rm e}}{{\rm e}^2} \operatorname{grad} \eta_{\rm e} , \qquad (5.33)$$

$$j_{\rm h} = -\frac{\sigma_{\rm h}}{e^2} \operatorname{grad} \eta_{\rm h} ; \qquad (5.34)$$

- For the carrier concentrations:

$$\frac{\partial n_{\rm e}}{\partial t} = G_{\rm e} - R_{\rm e} - \operatorname{div} j_{\rm e} ,$$

$$\frac{\partial n_{\rm h}}{\partial t} = G_{\rm h} - R_{\rm h} - \operatorname{div} j_{\rm h} ;$$
(5.35)

$$\frac{\partial n_{\rm h}}{\partial t} = G_{\rm h} - R_{\rm h} - \operatorname{div} j_{\rm h} ; \qquad (5.36)$$

- For the electrical potential:

$$\nabla^2 \varphi = -\frac{\rho_Q}{\varepsilon \varepsilon_0} \ . \tag{5.37}$$

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.

6 Basic Structure of Solar Cells

In order to properly understand the requirements for a solar cell, which is to generate current and voltage, we will first turn to a similar problem which 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 hydrogen and oxygen pressures 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, 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 at practically the equilibrium values.

Similar to a solar cell, we want to extract the gases separately and with them their chemical energy. How can we remove the gases separately from the mixture in the reaction cell? This requires membranes which 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 permeable for oxygen on the right, blocks the hydrogen. If the partial pressure, and therefore the chemical potential, of one gas type is greater in the reaction cell

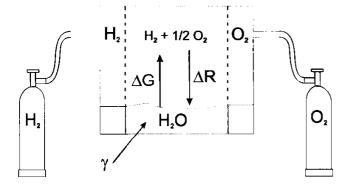


Figure 6.1: Hypothetical chemical solar cell in which water is decomposed into hydrogen and oxygen by the absorption of photons. Hydrogen and oxygen can be separately removed through membranes which selectively pass hydrogen on the left and oxygen on the right.

than outside, this gas will flow through its membrane 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. 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. In order to obtain the maximum chemical energy current, we must therefore accept a certain level of recombination. In the following, we will 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

$$dF = \mu_{\text{H}_2} dN_{\text{H}_2} + \mu_{\text{O}_2} dN_{\text{O}_2} + \mu_{\text{H}_2\text{O}} dN_{\text{H}_2\text{O}} = 0.$$
 (6.1)

From the chemical reaction

$$H_2 + \frac{1}{2}O_2 \rightleftharpoons H_2O$$

changes in the number of the molecules are related by

$$dN_{H_2} = 2 dN_{O_2} = -dN_{H_2O}$$

and Eq. (6.1) becomes

$$\mu_{\rm H_2} + \frac{1}{2}\mu_{\rm O_2} = \mu_{\rm H_2O} , \qquad (6.2)$$

which, as we have seen before, could have been read directly from the reaction equation. Eq. (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 does not change when the amount of water changes, since the concentration of water molecules in the water stays constant.

$$\mu_{\text{gas}} = \mu_0 + kT \ln \left(\frac{p_{\text{gas}}}{p_0} \right) , \qquad (6.3)$$

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 (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 Eq. (6.2).

$$kT \ln \left(\frac{p_{\text{H}_2}^0 \sqrt{p_{\text{O}_2}^0}}{p_0^{3/2}} \right) = \mu_{0,\text{H}_2\text{O}} - \mu_{0,\text{H}_2} - \frac{1}{2}\mu_{0,\text{O}_2} = -2.46\,\text{eV} \ . \tag{6.4}$$

The value on the right hand side of Eq. (6.4) is found in books on physical chemistry. As for the equilibrium concentrations of electrons or holes by doping, one of the partial pressures can be chosen at will. When the partial pressure of oxygen is $p_{O_2}^0 = 0.2$ bar as in air,

6.1 A chemical solar cell 111

the equilibrium partial pressure of hydrogen is only $p_{\rm H_2}^0 = 1.8 \times 10^{-41}\,\rm bar$. 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 re-established. Since 1 H₂ molecule reacts with $\frac{1}{2}$ O₂ molecule to give 1 molecule of H₂O, the recombination rate of hydrogen is

$$R_{\rm H_2} \sim p_{\rm H_2} \sqrt{p_{\rm O_2}}$$
.

This recombination rate for non-equilibrium partial pressures can be expressed by the equilibrium recombination rate

$$R_{\rm H_2} = R_{\rm H_2}^0 \frac{p_{\rm H_2} \sqrt{p_{\rm O_2}}}{p_{\rm H_2}^0 \sqrt{p_{\rm O_2}^0}} \ . \tag{6.5}$$

When, by illumination, the generation rate is increased from the equilibrium generation rate $G_{\rm H_2}^0$ by $\Delta G_{\rm H_2}$, hydrogen may flow out from the reaction cell through its membrane. The hydrogen current, arbitrarily counted positive if it is flowing into the cell, is

$$j_{H_{2}} = 2j_{O_{2}} = R_{H_{2}} - G_{H_{2}}$$

$$j_{H_{2}} = 2j_{O_{2}} = R_{H_{2}}^{0} \frac{p_{H_{2}}\sqrt{p_{O_{2}}}}{p_{O_{2}}^{0}} - G_{H_{2}}^{0} - \Delta G_{H_{2}}$$

$$j_{H_{2}} = 2j_{O_{2}} = R_{H_{2}}^{0} \left(\frac{p_{H_{2}}\sqrt{p_{O_{2}}}}{p_{H_{2}}^{0}\sqrt{p_{O_{2}}^{0}}} - 1\right) - \Delta G_{H_{2}}.$$

$$(6.6)$$

The interpretation is simple. Hydrogen and oxygen will 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.

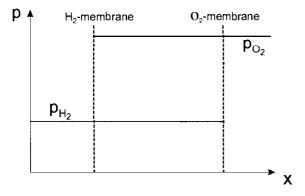


Figure 6.2: The conductivities for hydrogen in the hydrogen membrane on the left and for oxygen in the oxygen membrane on the right, are assumed to be large. The gradients of the partial pressures required to drive the currents are therefore small and do not show up in the spatial distribution of the pressures.

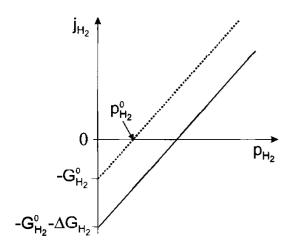


Figure 6.3: Current of hydrogen, positive if flowing from the hydrogen bottle in Figure 6.1 into the cell as a function of the hydrogen partial pressure, without illumination (broken line) and with additional generation $\Delta G_{\rm H_2}$ by illumination (solid line). The oxygen pressure is kept constant.

The hydrogen current flowing between the hydrogen bottle and the cell is shown in Figure 6.3 as a function of the hydrogen pressure in the hydrogen bottle, while the oxygen pressure is kept constant. 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.

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 Eqs. (6.2) and (6.3), the recombination rate can be written as

$$R_{\rm H_2} = R_{\rm H_2}^0 \exp\left(\frac{\mu_{\rm H_2} + 1/2\mu_{\rm O_2} - \mu_{\rm H_2O}}{kT}\right) , \qquad (6.7)$$

and the dependence of the hydrogen currents on the chemical potentials is

$$j_{\rm H_2} = 2 j_{\rm O_2} = R_{\rm H_2}^0 \left[\exp\left(\frac{\mu_{\rm H_2} + 1/2\mu_{\rm O_2} - \mu_{\rm H_2O}}{kT}\right) - 1 \right] - \Delta G_{\rm H_2} ,$$
 (6.8)

which is shown in Figure 6.4.

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

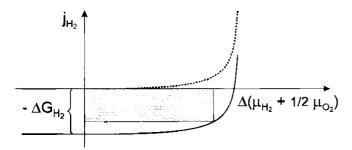


Figure 6.4: Current of hydrogen, positive if flowing from the hydrogen bottle in Figure 6.1 into the cell as a function of the deviation $\Delta(\mu_{\rm H_2} + \frac{1}{2}\mu_{\rm O_2})$ of the chemical potentials of hydrogen and oxygen from their equilibrium values, without illumination (broken line) and with additional generation $\Delta G_{\rm H_2}$ by illumination (solid line). A smaller and more realistic equilibrium generation rate $G_{\rm H_2}^0$ than in Figure 6.3 is assumed. The shaded rectangle is the largest current of chemical energy delivered by the cell.

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 semipermeable 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 which is permeable for electrons and which blocks the holes would be a material, which 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. In order 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

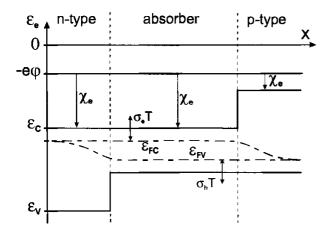


Figure 6.5: An n-type electron membrane on the left allows electrons generated in the absorber by illumination, to flow to the left while blocking the holes. A p-type hole membrane on the right allows holes to flow to the right blocking the electrons. Electrons are driven by an invisibly small gradient of $-\varepsilon_{FC}$, holes are driven by an invisibly small gradient of ε_{FV} .

larger band gap of the hole membrane on the right, combined with a smaller electron affinity, causes an energy barrier in the conduction band for the electrons. Due 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 gradients of the Fermi energies required to drive the electrons to the left and the holes to the right are negligibly small. 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. The same holds for the electrons in the hole membrane, accordingly.

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 outwards 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 will discuss different arrangements which fulfil 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 hetero-structure 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 6.6. 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 in order 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 TiO₂. As Figure 6.6 illustrates, the electrons in the dye reach the conduction band of the TiO₂ without difficulty. However, due to the large band gap of more than 3 eV of TiO₂, the holes present in the dye encounter a high barrier for the transition to the valence band of the TiO₂.

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 TiO_2 layer is therefore composed of particles, only a few nm in size, in a porous structure. All TiO_2 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 iodine ions of an iodine redox system (I^-/I_3^-) provide for charge transport. The energy of an electron in I^- differs only very slightly from the energy of an electron in the ground state of the dye, so that the flow of holes

¹B. O'Reagan, M. Grätzel, *Nature* **353** (1991) 737.

6.4 The pn-junction 115

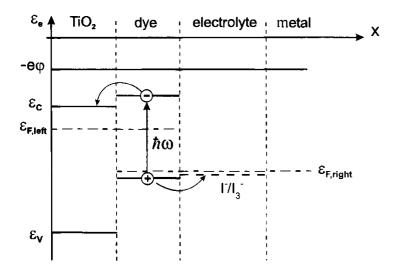


Figure 6.6: Dye solar cell in which the electron-hole pairs are produced in a ruthenium bipyridil dye. The electrons flow outwards towards the left through the n-conductor TiO_2 and the holes towards the right through the iodine ions with which the acetonitrile electrolyte is doped.

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 fulfil the requirements for the selective transport of electrons towards the left in the TiO₂ and of holes towards 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 which 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 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 remaining 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 these 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 300μ 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 1μ 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 will therefore be 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.
$$\eta_e + \eta_h = \mu_Y = 0$$
.

From Eq. (5.12) 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 $grad \eta_e = -grad \eta_h$ and therefore

$$j_Q = \frac{\sigma_e + \sigma_h}{e} \operatorname{grad} \eta_e = 0$$
.

Since $\sigma_e + \sigma_h \neq 0$, grad $\eta_e = 0$.

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 pnjunction. This is the implication of electrochemical equilibrium between electrons in the negion and electrons in the p-region. Indicating values in the p-region far away from the pnjunction by a superscript p and in the n-region far away from the pn-junction by a superscript p, we have

$$\eta_{e}^{p} = \mu_{e,0}^{p} + kT \ln \frac{n_{e}^{p}}{N_{C}} - e\phi^{p} = \eta_{e}^{n} = \mu_{e,0}^{n} + kT \ln \frac{n_{e}^{n}}{N_{C}} - e\phi^{n} .$$
 (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 (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

$$\varphi^{n} - \varphi^{p} = \frac{kT}{e} \ln \frac{n_{\rm D} n_{\rm A}}{n_{\rm i}^{2}} . \tag{6.10}$$

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 grad $\eta_e = 0$ and grad $\eta_h = 0$. The driving force then no longer exists, and 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. Still, they are mobile in random Brownian motion.

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-conductor and the p-conductor follows from the relationship with the distribution of the charge density ρ_Q .

From the Maxwell equation

$$\operatorname{div} D = \rho_O$$

with $D = \varepsilon \varepsilon_0 E$ and with $E = -\operatorname{grad} \varphi$ we find Poisson's Equation

$$\operatorname{div} E = -\operatorname{div}\operatorname{grad}\varphi = -\nabla^2\varphi = \frac{\rho_Q}{\varepsilon\varepsilon_0} \ . \tag{6.11}$$

The interface between the n- and p-regions is assumed to extend much farther in the yand z-directions than the width of the space charge layer to be derived. This allows a onedimensional treatment, giving

$$\frac{\mathrm{d}^2 \varphi}{\mathrm{d}x^2} = -\frac{\rho_Q}{\varepsilon \varepsilon_0} \ . \tag{6.12}$$

In the n-region the density of the space charge is

$$\rho_{Q}^{n}(x) = e\left(n_{D}^{+} - n_{e}(x)\right) = e n_{D}^{+} \left(1 - \exp\left\{\frac{e\left[\phi(x) - \phi^{n}\right]}{kT}\right\}\right). \tag{6.13}$$

This relationship follows from Figure 6.7, which schematically depicts the potential distribution and the charge distribution.

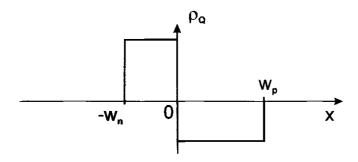
Unfortunately, however, for this space charge density, Poisson's Equation can only be solved numerically. According to Schottky, we can approximate the charge distribution by a spatially constant positive space charge density in a region up to an, as yet, undefined depth w_n on the n-side. A constant 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 6.7.

$$\rho_Q^n = e n_D^+ \approx e n_D$$
 for $-w_n < x \le 0$

$$\rho_Q^p = -e n_A^- \approx -e n_A$$
 for $0 \le x < w_p$.

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

$$w_{\rm p} = \frac{n_{\rm D}}{n_{\rm A}} w_{\rm n} \tag{6.14}$$



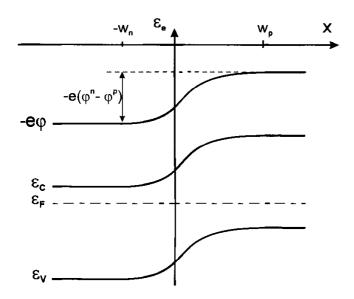


Figure 6.7: Distribution of the space charge density ρ_Q and of the electrical energy $-\epsilon \varphi$ per electron in a pn-junction.

and the entire thickness of the space charge layer is

$$w = w_{\rm n} + w_{\rm p} = \left(1 + \frac{n_{\rm D}}{n_{\rm A}}\right) w_{\rm n}.$$
 (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 $-w_n < x \le 0$ of the n-region yields

$$\varphi_{n}(x) = -\frac{en_{D}}{2\varepsilon\varepsilon_{0}}(x + w_{n})^{2} + \varphi^{n}, \qquad (6.16)$$

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 $0 \le x < w_p$ of the p-region yields

$$\varphi_{\mathbf{p}}(x) = \frac{\mathbf{e}n_{\mathbf{A}}}{2\varepsilon\varepsilon_{\mathbf{0}}} \left(x - w_{\mathbf{p}} \right)^2 + \varphi^{\mathbf{p}} . \tag{6.17}$$

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

$$\varphi^n - \varphi^p = \frac{e}{2\varepsilon\varepsilon_0} \left(n_D w_n^2 + n_A w_p^2 \right) . \tag{6.18}$$

With reference to Eqs (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 Eq. (6.10):

$$w = \sqrt{\frac{2\varepsilon\varepsilon_0}{e} \frac{n_A + n_D}{n_A n_D} (\varphi^n - \varphi^p)}. \tag{6.19}$$

For the asymmetrical pn-junction of a Si solar cell, with $n_D = 10^{19} / \text{cm}^3$ and $n_A = 10^{16} / \text{cm}^3$, Eq. (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 Eq. (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 favour Auger recombination and should be avoided in a solar cell. Are there minimum doping levels required for a pn-junction solar cell? The problem is that the conversion of chemical energy of the electron-hole pairs into electrical energy is incomplete, when the difference between the Fermi energies in the contacts $\epsilon_{F,left} - \epsilon_{F,right}$ is smaller than the difference between the Fermi energies $\epsilon_{FC} - \epsilon_{FV}$ in the absorber, as shown in Figure 5.6. Incomplete conversion has to do with the generation of entropy. A comparison of Figure 5.6 with Figure 6.5, where the conversion is complete, indicates that the entropy per electron σ_e , which increases with the distance of the Fermi energy ϵ_{FC} from the conduction band, increases from the interior towards the left surface in Figure 5.6. Entropy generation is obvious, since the entropy rises at the expense of the free energy of the electrons $\eta_e = \epsilon_{FC}$. The same happens to the entropy per hole σ_h increasing with the distance of the Fermi energy ϵ_{FV} from the valence band and rising towards the right surface. This entropy generation is avoided, if care is taken that the concentration of the electrons in a pn-junction is at least the same everywhere on the path of the carriers, towards their membrane.

From the considerations above, it follows that the majority carrier concentrations in the dark in the regions functioning as membranes must be at least as large as the additional concentrations generated by illumination. Since the dark concentrations of the majority carriers are equal to the doping concentrations, this condition is expressed as

$$n_{\rm D}n_{\rm A} \ge n_{\rm e}^p n_{\rm h}^n$$

equivalent to the condition

$$kT \ln \frac{n_{\rm D} n_{\rm A}}{n_{\rm i}^2} = e\left(\mathbf{\phi}^n - \mathbf{\phi}^p\right) \geq \varepsilon_{\rm FC} - \varepsilon_{\rm FV} = kT \ln \frac{n_{\rm c}^p n_{\rm h}^n}{n_{\rm i}^2}. \tag{6.20}$$

This requires suitable doping 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 $e(\varphi^n - \varphi^p) = \varepsilon_{FC} - \varepsilon_{FV}$, 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 $e(\varphi^n - \varphi^p) < \varepsilon_{FC} - \varepsilon_{FV}$. 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:

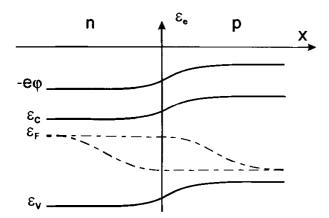


Figure 6.8: Potential distribution in an illuminated solar cell in which the diffusion voltage is $e(\varphi^n - \varphi^p) > \varepsilon_{FC} - \varepsilon_{FV}$.

so chemical energy is wasted. Figure 6.8 gives the potential distribution for a more strongly doped cell, for which $e(\varphi^n - \varphi^p) > \epsilon_{FC} - \epsilon_{FV}$. 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.

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 towards the pn-junction, and the reverse direction, in which electrons and holes flow away from the pn-junction.

In the forward direction, shown in the upper part of Figure 6.9, 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.

In the reverse direction, shown in the lower part of Figure 6.9, 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. But only those electrons generated in the p-region which 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 $L_{\rm e}$ to the right or a hole diffusion length $L_{\rm h}$ 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

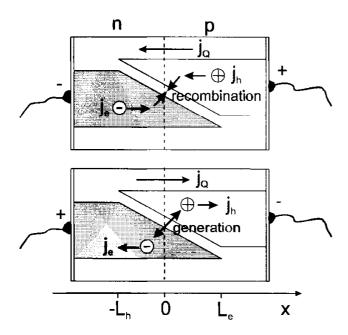


Figure 6.9: Electron and hole currents in a pn-junction. For a negative polarity of the n-region with respect to the p-region, i.e., in the forward direction (upper part), electrons and holes flow towards the pn-junction, where they recombine. In the reverse direction (lower part), for a positive polarity of the n-region, electrons and holes flow away from the pn-junction, where they are produced.

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 towards the pn-junction),

$$j_Q = -e \int_{-L_h}^{L_c} \operatorname{div} j_h \, \mathrm{d}x \,,$$
 (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

$$rac{\partial n_{
m h}}{\partial t} = G_{
m h} - R_{
m h} - {
m div}\, j_{
m h} = 0 \; , \qquad \qquad {
m div}\, j_{
m h} = G_{
m h} - R_{
m h} \; .$$

Again, we divide the generation rate into the components G_h^0 in the dark and ΔG_h from the illumination

$$G_{\rm h}=G_{\rm h}^0+\Delta G_{\rm h}\;.$$

The recombination rate for radiative recombination is accordingly

$$R_{\rm h} = R_{\rm h}^0 \frac{n_{\rm e} n_{\rm h}}{n_{\rm i}^2} = R_{\rm h}^0 \exp\left(\frac{\eta_{\rm e} + \eta_{\rm h}}{kT}\right) ,$$
 (6.22)

where, because of $\eta_e + \eta_h = 0$ in equilibrium with the 300 K photons and phonons

$$G_{\rm h}^0 = R_{\rm h}^0$$
.

We can now write the charge current of Eq. (6.21) as

$$j_Q = -e \int_{-L_h}^{L_e} \left\{ G_h^0 \left[1 - \exp\left(\frac{\eta_e + \eta_h}{kT}\right) \right] + \Delta G_h \right\} dx . \tag{6.23}$$

In principle, the sum of the electrochemical potentials $\eta_e + \eta_h$ is a function of the position and depends on the resistances which 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

$$e + h \rightleftharpoons \gamma, n\Gamma$$
.

In this case

$$\operatorname{div} j = G - R.$$

If the chemical reaction is limiting the current, as it did 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 towards 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 is a maximum of $42 \,\mathrm{mA/cm^2}$ for silicon in non-focussed solar radiation. For a doping concentration of $n_{\rm A} = 10^{16} \,\mathrm{/cm^3}$ and a mobility of $b_{\rm h} = 470 \,\mathrm{cm^2/(Vs)}$ the conductivity is $\sigma_{\rm h} = 0.75 \,\mathrm{/(\Omega\,cm)}$.

The voltage drop is then $1/e \operatorname{grad} \eta_h = j_Q/\sigma_h = 56 \, \mathrm{mV/cm}$, for a thickness of $300 \, \mu \mathrm{m}$ thus less than $2 \, \mathrm{mV}$. This is negligible compared with $(\eta_e + \eta_h)/e$, which is of the order of $1 \, \mathrm{V}$. We conclude that the transport resistance is negligible. The current through a pn-junction is instead limited by the reaction resistance.

We therefore have

grad
$$\eta_h \approx 0$$
 for $x > -L_h$

and

grad
$$\eta_e \approx 0$$
 for $x < L_e$

so that

$$\eta_e + \eta_h \neq f(x)$$
 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-region and the p-region.

From the considerations above, it follows that the exponential function in the integrand in Eq. (6.23) is constant over the range of the integration limits. This greatly simplifies the integration of Eq. (6.23), and we obtain the current-voltage characteristic of the pn-junction as

$$j_Q = e G_h^0 (L_e + L_h) \left[exp \left(\frac{eV}{kT} \right) - 1 \right] - e \int_{-L_h}^{L_e} \Delta G_h \, dx \,.$$
 (6.24)

An external short-circuit (V = 0) defines the short-circuit current j_{sc}

$$j_Q = -e \int_{-L_h}^{L_e} \Delta G_h \, dx = -e \int_{-L_h}^{L_e} \Delta G_e \, dx = j_{sc} .$$
 (6.25)

In the dark ($\Delta G_{\rm e,h}=0$) and for large negative voltages ($\exp({\rm e}V/kT)\ll 1$) we find the reverse saturation current $j_{\rm S}$, which is independent of the voltage,

$$j_O = -eG_{e,h}^0(L_e + L_h) = -j_S$$
 (6.26)

The short-circuit current j_{sc} and the reverse saturation current j_{sc} are the essential elements of the current-voltage characteristic of a pn-junction

$$j_Q = j_S \left[\exp\left(\frac{eV}{kT}\right) - 1 \right] + j_{sc} . \tag{6.27}$$

With Eq. (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 Eq. (3.54) and the lifetime from Section 3.6.3.

In real pn-junctions, we must also consider the electron-hole generation by non-radiative transitions, as the reverse process of non-radiative recombination. The rate of generation cannot then be given in general form. However, it can be expressed in terms of the diffusion lengths, since the lifetimes which these contain represent a parameter which 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

$$G_{\mathrm{e,h}}^{0} = R_{\mathrm{e,h}}^{0} = \frac{n_{\mathrm{e}}^{p}}{\tau_{\mathrm{e}}} = \frac{n_{\mathrm{h}}^{n}}{\tau_{\mathrm{h}}}.$$

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

$$j_{\rm S} = \mathrm{e}n_{\rm i}^2 \left(\frac{D_{\rm e}}{n_{\rm A}L_{\rm e}} + \frac{D_{\rm h}}{n_{\rm D}L_{\rm h}} \right) . \tag{6.28}$$

This expression for the reverse saturation current is valid even for real pn-junctions in which recombination is predominantly non-radiative, when empirically determined values are used for the diffusion lengths of the electrons and holes.

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 $L_{\rm h}$ away from the surface. In fact, the n-layer at the surface is chosen to be very thin, 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 $\mathrm{d}j_{\gamma}(\hbar\omega,x=0)$. This yields

$$\Delta G_{\mathrm{e,h}}(x) = \int_0^\infty \alpha(\hbar\omega)(1 - r(\hbar\omega)) \, \exp(-\alpha(\hbar\omega)x) \, \mathrm{d}j_{\gamma}(\hbar\omega,0) \; .$$

The contribution to the short-circuit current in the interval $d\hbar\omega$ is

$$dj_{sc}(\hbar\omega) = -e \left[1 - r(\hbar\omega)\right] dj_{\gamma}(\hbar\omega, 0) \alpha(\hbar\omega) \int_{0}^{L_{e}} e^{-\alpha x} dx$$

$$= -e \left[1 - r(\hbar\omega)\right] \left\{1 - \exp\left[-\alpha(\hbar\omega)L_{e}\right]\right\} dj_{\gamma}(\hbar\omega, 0) .$$
(6.29)

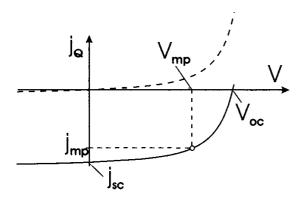


Figure 6.10: Charge current of the pn-junction in the dark (dashed line) and with illumination (solid line) as a function of the voltage. The sign of the voltage corresponds to the polarity of the p-region. The shaded rectangle represents the maximum power delivered by the illuminated pn-junction.

 $[1-r(\hbar\omega)]\{1-\exp[-\alpha(\hbar\omega)L_e]\}$ is the familiar absorptivity $a(\hbar\omega,L_e)$ 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 absorptivity for the actual thickness must be used.² 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. This gives

$$j_{
m sc} = -{
m e} \int_0^\infty a(\hbar\omega, L_{
m e})\,{
m d}j_{m{\gamma}}(\hbar\omega, 0) \;.$$

Figure 6.10 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 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 $V_{\rm oc}$ is important. From

$$j_Q = j_S \left[\exp\left(\frac{eV_{oc}}{kT}\right) - 1 \right] + j_{sc} = 0$$
(6.30)

we obtain

$$V_{\rm oc} = \frac{kT}{e} \ln \left(1 - \frac{j_{\rm sc}}{j_{\rm S}} \right) . \tag{6.31}$$

It is very important for the 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 between the Fermi energies for Auger recombination or for recombination via impurities, which leads to a different dependence of the current on the voltage.

²This result requires that a perfect hole membrane at the rear side of the cell prevents the electrons from reaching the back contact where they would otherwise recombine, without contributing to the short-circuit current.

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 which 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 know that impurities with an electron energy in the middle of the forbidden zone contribute particularly strongly to recombination. In order to evaluate the influence of impurity recombination on the current-voltage relationship, we will therefore restrict ourselves to this type of impurity, with $\varepsilon_{imp} = \varepsilon_i$ in the middle of the forbidden zone. Furthermore, it is assumed that electrons and holes have equal thermal velocities ν and are captured with the same cross-sections, σ . With this simplification, Eq. (3.76) leads to the relationship

$$R_{\text{imp}} = n_{\text{imp}} \sigma v n_{\text{i}} \frac{\exp\left[\frac{\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}}}{kT}\right] - 1}{\exp\left[\frac{\varepsilon_{\text{FC}} - \varepsilon_{\text{i}}}{kT}\right] + \exp\left[\frac{\varepsilon_{\text{i}} - \varepsilon_{\text{FV}}}{kT}\right] + 2}.$$
(6.32)

For a given voltage V and thus a given difference $\varepsilon_{FC} - \varepsilon_{FV} = eV$ between the Fermi energies, Figure 6.11 shows the recombination rate R_{imp} as a function of the mean value of the Fermi energies relative to the position ε_i of the impurity level in the middle of the forbidden zone. It can be seen that the recombination rate has a pronounced maximum, when the impurity level lies in the middle between the Fermi energies.

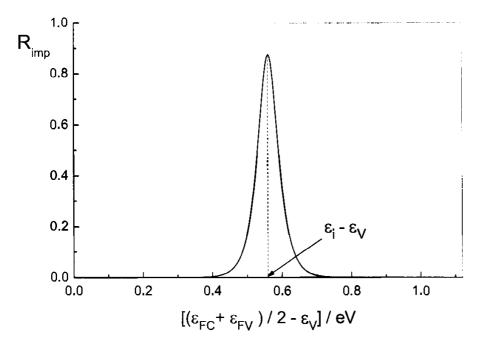


Figure 6.11: Impurity recombination rate R_{imp} as a function of the mean value of the Fermi energies $(\varepsilon_{\text{FC}} + \varepsilon_{\text{FV}})/2$ relative to the edge of the valence band ε_{V} in a semiconductor with an energy gap of $\varepsilon_{\text{G}} = 1.12\,\text{eV}$ calculated for an applied voltage of $V = 0.4\,\text{V}$.

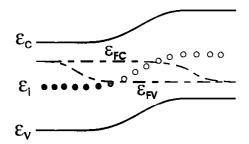


Figure 6.12: Potential distribution in a pn-junction with impurities at ε_i in the middle of the forbidden zone.

This behaviour is understandable from Figure 6.12, 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. There $\varepsilon_{FC} - \varepsilon_i = \varepsilon_i - \varepsilon_{FV} = eV/2$. This means that the recombination rate in Eq. (6.32) is

$$R_{\text{imp}} = n_{\text{imp}} \sigma v n_{\text{i}} \frac{\exp\left(\frac{eV}{kT}\right) - 1}{2\left[\exp\left(\frac{eV}{2kT}\right) + 1\right]}.$$
(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

$$j_{Q,\text{imp}} = \frac{ew\sigma v n_{\text{imp}} n_{\text{i}}}{2} \left\{ \exp\left(\frac{eV}{2kT}\right) - 1 \right\} = j_{S2} \left\{ \exp\left(\frac{eV}{2kT}\right) - 1 \right\}. \tag{6.34}$$

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

$$j_Q = j_{S1} \exp\left\{\left(\frac{eV}{kT}\right) - 1\right\} + j_{S2} \exp\left\{\left(\frac{eV}{2kT}\right) - 1\right\} + j_{sc}.$$
 (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.

6.6 Hetero-junctions 127

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-side, 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 between 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 homo-junction, 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 material, outside of 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 10¹⁶/cm³ acceptors, typical for crystalline silicon solar cells. For this lifetime the diffusion length is $L_e \approx 1$ cm. In order to preserve the function of the hole membrane on the p-side which 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 $\hbar\omega \geq \varepsilon_G$, a waste of precious material. We will see in Section 7.3, 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.

6.6 Hetero-junctions

In Figure 5.6 we saw 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 it is justified to neglect minority carrier currents in the wrong direction, namely an electron current towards the p-side and a hole current towards 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.

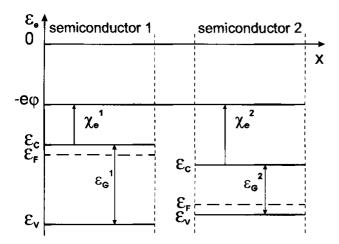


Figure 6.13: Two different semiconductors prior to making contact.

In addition to preventing interface recombination at metal contacts, hetero-junctions are important, when it is not possible to prepare a pn-junction from a single material, a so-called homo-junction. There are in fact many materials which 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 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 $D = \varepsilon \varepsilon_0 E$ and the electrical potential φ are continuous over the interface. Due to 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 6.13 we see two different semiconductors, at the left semiconductor 1, which is n-doped, and at the right semiconductor 2, which is p-doped. In order to determine the potential distribution in the dark, we will begin in Figure 6.14 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 6.13.

The distribution of the electrical potential is determined by the charge distribution and is independent of the electron affinity or the energy gap. Due to the continuity of the electrical potential across the interface, in the case without charge in interface states, the resulting

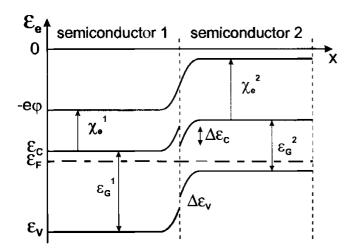


Figure 6.14: The two different semiconductors of Figure 6.13 in contact.

potential distribution is, apart from the different slope arising from the ratio of the dielectric constants, the same as for the normal pn-junction in Section 6.4.

According to Eq. (6.9), the electrical potential difference between the neutral regions of the n-conductor and the p-conductor is

$$e(\varphi^{1} - \varphi^{2}) = \mu_{e,0}^{1} + kT \ln\left(\frac{n_{e}^{1}}{N_{C}^{1}}\right) - \mu_{e,0}^{2} - kT \ln\left(\frac{n_{e}^{2}}{N_{C}^{2}}\right) , \qquad (6.36)$$

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.

$$\begin{array}{lcl} \Delta\epsilon_C & = & \epsilon_C^1 - \epsilon_C^2 = \chi_e^2 - \chi_e^1 \\ \Delta\epsilon_V & = & \epsilon_V^1 - \epsilon_V^2 = \chi_e^2 - \chi_e^1 + \epsilon_G^2 - \epsilon_G^1 \end{array}.$$

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 to control the charge transport.

6.7 Semiconductor-metal contact

Metal contacts are to 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 6.15 shows a semiconductor and a metal before making contact and in contact.

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,

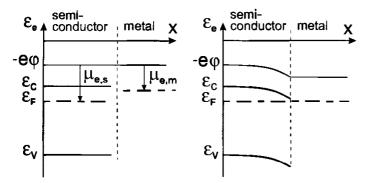


Figure 6.15: Schematic energy diagram of a semiconductor and metal before making contact (left) and in contact (right).

an electrical potential difference arises, as for the case of two semiconductors, corresponding to the difference in the work functions. Due 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 6.15, 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 favourable 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 travelling 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, however, 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 due to the presence of surface states on the silicon surface, which are charged due to the 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.

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	5.65
Electron affinity /eV	4.01	4.07					

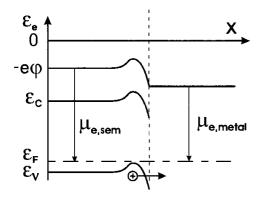


Figure 6.16: Holes from the valence band can tunnel through a thin potential barrier of a strongly p-doped depletion layer into the metal.

Contacts with good carrier exchange can also be prepared based on 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. Aluminium 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 aluminium 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 6.16. 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 tunnelling. Generating a strongly doped p⁺-layer in front of the metal contact also improves the membrane character of the p-region.

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 which 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 6.17 shows the potential distribution at a Schottky contact on a p-conductor in the dark. It must, however, be noted that Schottky contacts are 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 hetero-junctions 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 Cu₂S is formed.

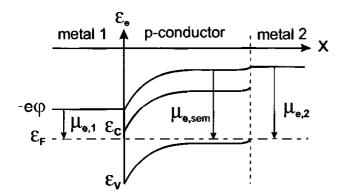


Figure 6.17: A metal with a small work function (left) forms a Schottky contact on a p-conductor with a larger work function. On the right side, a metal with a large work function forms an ohmic contact.

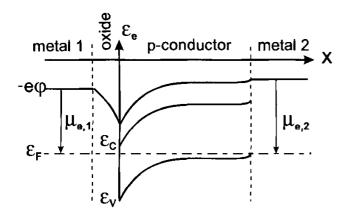


Figure 6.18: In an MIS (Metal Insulator Silicon) structure a very thin oxide layer between the metal and the silicon prevents surface recombination. The depletion of holes in the p-conductor is the result of a positive charge trapped in the oxide near the silicon.

6.7.2 MIS contact

In order 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 MIS (Metal Insulator Silicon) structures³ this disadvantage is compensated by exploiting a 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 6.18 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.

³K. Jaeger, R. Hezel, Proc 18. IEEE PV Spec. Conf., Las Vegas (1985), p. 388.

6.8 The role of the electric field in solar cells

The reader may find it confusing that the electric field which exists in the dark and, although somewhat reduced, also in the light in a 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. When 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 fulfil the condition of conservation of entropy. The dye solar cell in Section 6.3 is a good example. The intimate mixture of electron membrane (TiO₂), 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 will be given at the end of this section.

We frequently read that it is just the electric field of a pn-junction which 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 which 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 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 which 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, 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 6.19 shows water which 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

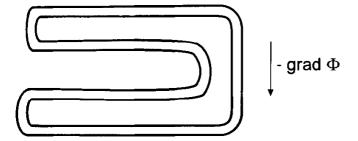


Figure 6.19: Water in a closed pipe does not flow, although the gradient of the gravitational potential Φ drives it downwards.

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 which 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.

Figure 6.20 shows a pn-junction with contacts on the n- and p-side. The contacts are made from the same metal as the wire (not shown) which 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, due 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, $\eta_e = \epsilon_{FC}$ and $\eta_h = -\epsilon_{FV}$, 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.

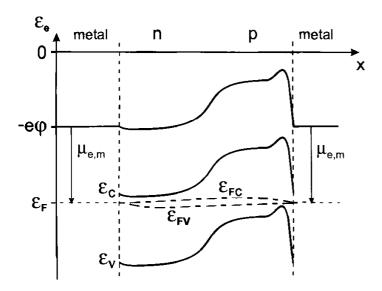


Figure 6.20: The distribution of the electrical potential φ in a pn-junction with metal contacts shows that a charge cannot gain energy from moving around a closed circuit.

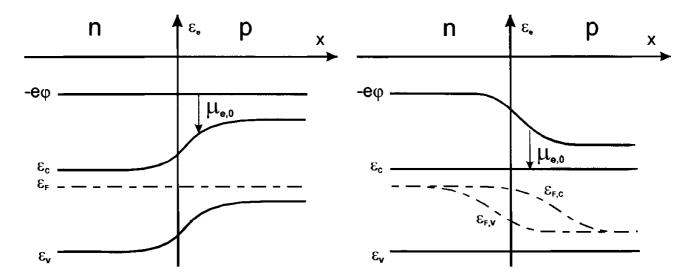


Figure 6.21: Potential distribution in a pn-junction, where the electron affinity $\chi_{\mathbf{e}}(x) = -\mu_{\mathbf{e},0}(x)$ compensates the concentration dependent part $kT \ln [n_{\mathbf{e}}(x)/N_{\mathbf{C}}]$ of the chemical potential. Left: in the dark and in electrochemical equilibrium. Right: illuminated and at open-circuit.

In order 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 hetero-junctions. According to Eq. (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 , also increasing from the n-to the p-region. In the dark, in electrochemical equilibrium, this structure has the potential distribution shown on the left in Figure 6.21.

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 6.21. 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 6.21 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 between the Fermi energies and that it is therefore the same for the hypothetical pn-junction in Figure 6.21 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 negative ions within the battery flow from the plus pole to the minus pole. They are able to do so, because a gradient of their chemical potential overcompensates the gradient of the electrical potential giving them a gradient of their electrochemical potential in the right direction, their only driving force.

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 which function as semi-permeable 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, this is exactly the number of electron—hole pairs 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 4.1. It defines the "maximum power point" for the charge current density $j_{\rm mp}$ and the voltage $V_{\rm mp}$. For a given current-voltage characteristic, it is therefore important to have an algorithm in order to find the maximum power point.

Independently of the form of the characteristic, the functional relationship between j_Q and V, the condition for maximum power yields

$$d(j_Q V) = dj_Q V + j_Q dV = 0 ,$$

and thus

$$\left(\frac{\mathrm{d}j_{Q}}{\mathrm{d}V}\right)_{\mathrm{mp}} = -\left(\frac{j_{Q}}{V}\right)_{\mathrm{mp}}.\tag{7.1}$$

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.

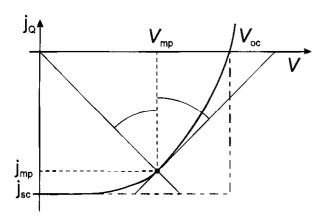


Figure 7.1: Geometrical construction of the maximum power point.

For a certain form of the characteristic, namely for radiative recombination, in Eq. (6.27)

$$\frac{\mathrm{d}j_Q}{\mathrm{d}V} = j_\mathrm{S} \frac{\mathrm{e}}{kT} \exp\left(\frac{\mathrm{e}V_\mathrm{mp}}{kT}\right) = -\frac{j_\mathrm{mp}}{V_\mathrm{mp}} \,. \tag{7.2}$$

With

$$j_{\mathrm{mp}} = j_{\mathrm{S}} \left\{ \exp \left(\frac{\mathrm{e}V_{\mathrm{mp}}}{kT} \right) - 1 \right\} + j_{\mathrm{sc}}$$

and

$$\frac{j_{\rm se}}{j_{\rm S}} = 1 - \exp\left(\frac{{\rm e}V_{\rm oc}}{kT}\right)$$

we find from Eq. (7.2)

$$V_{\rm mp} = \frac{kT}{e} \left\{ \exp\left(\frac{e(V_{\rm oc} - V_{\rm mp})}{kT}\right) - 1 \right\} . \tag{7.3}$$

Solutions of this equation, and thus the maximum power point, can be found by numerical techniques. With the fill factor

$$FF = \frac{j_{\rm mp}V_{\rm mp}}{j_{\rm sc}V_{\rm oc}} \tag{7.4}$$

we define a measure for how well the maximum power rectangle fits under the characteristic. For only radiative recombination $j_{\rm sc}$ $V_{\rm oc}$ 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 Eq. (7.3). This gives us

$$V_{\mathrm{mp}} = V_{\mathrm{oc}} - \frac{kT}{\mathrm{e}} \ln \left(1 + \frac{\mathrm{e}V_{\mathrm{mp}}}{kT} \right) \approx V_{\mathrm{oc}} - \frac{kT}{\mathrm{e}} \ln \left(1 + \frac{\mathrm{e}V_{\mathrm{oc}}}{kT} \right) .$$

Since the logarithm depends only weakly on its argument, we have substituted $V_{\rm oc}$ for $V_{\rm mp}$ in the logarithm. With this result for $V_{\rm mp}$ we derive the current $j_{\rm mp}$ at the maximum power point from the characteristic and find an approximate fill factor

$$FF = \frac{eV_{\rm oc}/kT - \ln(1 + eV_{\rm oc}/kT)}{1 + eV_{\rm oc}/kT} . \tag{7.5}$$

Values of the fill factor 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 absorptivity. With anti-reflection coatings, we can theoretically reduce the reflection to r=0. For a cell with large thickness and at the same time a large diffusion length, the absorptivity over the diffusion length is given by $a(\hbar\omega \ge \epsilon_G) \approx 1$. The short-circuit current produced by the absorbed photon current is then

$$j_{\rm sc} = -e \int_0^\infty a(\hbar\omega) \, dj_{\gamma,\,\rm sun}(\hbar\omega) = -e \int_{\varepsilon_{\rm G}}^\infty dj_{\gamma,\,\rm sun}(\hbar\omega) \,. \tag{7.6}$$

Maximum open-circuit voltage

The open-circuit voltage V_{oc} defines the separation $\epsilon_{FC} - \epsilon_{FV}$ of the Fermi energies at which recombination is in equilibrium with electron-hole generation throughout the entire cell. Due 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 everywhere equal to the averaged generation rate. When the thickness of the solar cell is reduced and surface recombination is prevented, the recombination rate (per volume) and with it $\epsilon_{FC} - \epsilon_{FV}$ 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, not only in order 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 absorptivity 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

$$j_{\gamma,\text{emit}} = \int_0^\infty a(\hbar\omega) \, \mathrm{d}j_{\gamma}^0(\hbar\omega) \, \exp\left(\frac{\varepsilon_{\text{FC}} - \varepsilon_{\text{FV}}}{kT}\right) \,.$$
 (7.7)

For the maximum short-circuit current, $a(\hbar\omega \ge \epsilon_G) = 1$, and replacing $\epsilon_{FC} - \epsilon_{FV}$ by eV, the total rate of radiative recombination is

$$j_{\gamma,\text{emit}} = \exp\left(\frac{eV}{kT}\right) \int_{\epsilon_{G}}^{\infty} dj_{\gamma}^{0}(\hbar\omega) ,$$
 (7.8)

independent of the thickness.

Defining the charge current from the p-region to the n-region (towards the left in Figure 6.5) as positive, the charge current delivered by the solar cell is

$$j_Q = e j_{\gamma, \text{emit}}(V) - e j_{\gamma, \text{abs}}$$
(7.9)

or

$$j_{Q} = e \left[\exp \left(\frac{eV}{kT} \right) - 1 \right] \int_{\varepsilon_{G}}^{\infty} dj_{\gamma}^{0}(\hbar\omega) - \int_{\varepsilon_{G}}^{\infty} dj_{\gamma, \text{sun}}(\hbar\omega) . \tag{7.10}$$

Since the spectra of the 300 K background radiation $dj_{\gamma}^{0}(\hbar\omega)$ and of the sun $dj_{\gamma,sun}(\hbar\omega)$ (outside the atmosphere AM0 and on the surface of the earth AM1.5) are known, the maximum energy current $(j_{Q}V)_{max}=j_{mp}V_{mp}$ can be determined from Eq. (7.10) and, from this, the efficiency

$$\eta = \frac{j_{\rm mp} V_{\rm mp}}{\int_0^\infty \hbar \omega \, \mathrm{d}j_{\gamma, \, \rm sun}(\hbar \omega)} \,. \tag{7.11}$$

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 $\varepsilon_G = 0$ and decreases with increasing ε_G . The open-circuit voltage V_{oc} is, however, zero for $\varepsilon_G = 0$ and increases with increasing energy gap. The efficiency η is therefore zero at $\varepsilon_G = 0$ and at $\varepsilon_G \to \infty$. Somewhere in between is its maximum. From Eq. (7.10) and Eq. (7.11) we can calculate the efficiency η as a function of the energy gap ε_G when only radiative recombination takes place for the case of thick cells, in which $a(\hbar\omega < \varepsilon_G) = 0$ and $a(\hbar\omega \ge \varepsilon_G) = 1$.

Figure 7.2 gives the result for the AM0 spectrum outside the atmosphere, and Figure 7.3 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 eV and 1.5 eV are suitable for

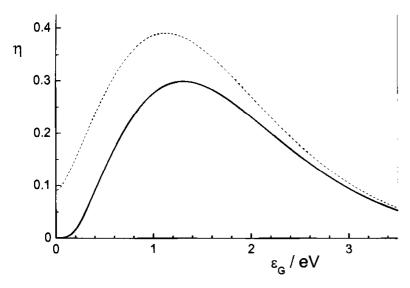


Figure 7.2: Efficiency of solar cells with radiative recombination only as a function of their energy gap for the AM0 spectrum, non-concentrated (solid line) and for full concentration (dashed line).

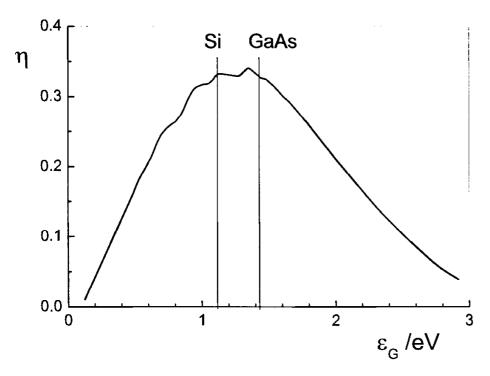


Figure 7.3: Efficiency of solar cells with radiative recombination only as a function of the energy gap for the AM1.5 spectrum.

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 $\hbar\omega < 1 \,\mathrm{eV}$, 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 which fully protects it and prevents any further corrosion. The interface between Si and SiO₂ when grown under clean-room conditions has a very low density of surface states, with a very low surface recombination velocity. With $\epsilon_G = 1.12\,\mathrm{eV}$, silicon has a favourable energy gap for the conversion of solar energy. Besides all these advantages, silicon with its indirect optical transitions has the serious disadvantage of weak absorption. Consequently, silicon must be much thicker than a semiconductor with direct transitions. Moreover, due 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. As a result of the poor absorption, not only is more silicon needed, but silicon 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 non-transparent contact poses a problem. Metal contacts are therefore arranged in narrow strips, in comb-like structures leaving most of the surface uncovered. The

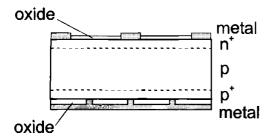


Figure 7.4: Cross-section of a silicon pn solar cell.

charge must then flow in the membrane layer towards 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 surface is strongly n-doped, with the designation n^+ , to function as an electron membrane. In order 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 (grad η_e) and the smaller recombination probability is instead due to the reduced concentration of electrons in the p^+ -layer.

Due 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 which reflects the photons not yet absorbed and thus enhances absorption in the cell. The oxide 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 7.4 illustrates a cross-section through this structure. The current-voltage characteristic in Figure 7.5 was calculated for the AM1.5 spectrum, assuming that the front side is non-reflecting and only the unavoidable radiative recombination and Auger recombination corresponding to the necessary doping levels are considered, but not surface recombination. This cell has an efficiency of 25%.

7.3.1 Light trapping

The absorptivity of a body increases as the reflectivity is reduced and the path of the photons within the body becomes longer. This triviality enables us to consider still another way to improve the absorptivity than by an anti-reflection coating and a large thickness. The reflectivity 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 7.6 makes this possible.

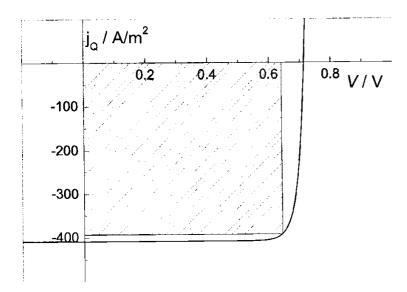


Figure 7.5: Charge current j_Q as a function of the voltage V for a Si solar cell with a thickness of $400\,\mu\text{m}$ illuminated by the AM1.5 spectrum. The maximum power given by the rectangle corresponds to an efficiency of 25%.

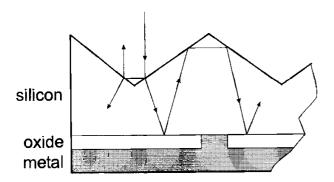


Figure 7.6: Surface texture for reducing reflection and increasing the length of the light path.

For light reflected twice, the total reflectivity 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 due to 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.

At the critical angle for total internal reflection

$$\sin \alpha_{\rm T} = \frac{1}{n_{\rm Si}} \ .$$

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 which 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 reflection when they happen to 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 which 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 absorptivity is defined as

$$a(\hbar\omega) = \frac{I_{E,\,\mathrm{abs}}(\hbar\omega)}{I_{E,\,\mathrm{inc}}(\hbar\omega)} = \frac{I_{\gamma,\,\mathrm{abs}}(\hbar\omega)}{I_{\gamma,\,\mathrm{inc}}(\hbar\omega)} \; .$$

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.¹

$$(1-r)I_{\gamma,inc} = I_{\gamma,em} + I_{\gamma,abs}$$
.

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

$$I_{\gamma,\,\mathrm{abs}} = 4\pi\alpha V \, j_{\gamma,\Omega} \; .$$

The photon current leaving through the surface is

$$I_{\gamma,\text{em}} = A \left(1 - r\right) \frac{\pi}{n^2} j_{\gamma,\Omega} .$$

The absorptivity for perfect light trapping is therefore

$$a_{\text{trap}} - \frac{I_{\gamma,\text{abs}}}{I_{\gamma,\text{inc}}} = (1 - r) \frac{I_{\gamma,\text{abs}}}{I_{\gamma,\text{em}} - I_{\gamma,\text{abs}}}$$

$$a_{\text{trap}} = (1 - r) \frac{4\pi\alpha L}{(1 - r)\pi/n^2 + 4\pi\alpha L}$$

$$a_{\text{trap}} = \frac{1 - r}{(1 - r)/(4n^2\alpha L) + 1}.$$
(7.12)

For small values of the absorption coefficient α we find the absorptivity $a_{\rm trap}=4n^2\alpha L$. For a reflectivity r which is not too large, it is surprising that the absorptivity does not depend on the reflectivity. For silicon, $a_{\rm trap}$ is a factor of $4n_{\rm Si}^2\approx 50$ greater than the absorptivity 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 $4n_{\rm Si}^2\approx 50$.

¹E. Yablonovitch, J. Opt. Soc. Am. 72 (1982) 899.

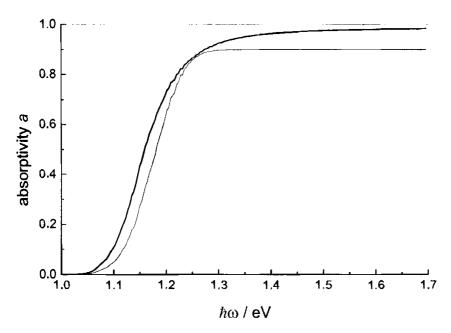


Figure 7.7: Absorptivity a as a function of the photon energy $\hbar\omega$ for 20μ m thick silicon with light trapping (heavy line) and for 400μ m thick silicon without light trapping (thin line), assuming a reflectivity of r = 0.1 in both cases.

Although the above derivation is based on a homogeneous distribution of photons, which is a good approximation only when $\alpha \ll 1/L$, Eq. (7.12) is also valid to a high degree of accuracy for larger values of α since the saturation value $a_{\text{trap}} = 1 - r$ is already reached for $\alpha < 1/L$ and cannot increase further, even for larger values of α . Figure 7.7 shows that, by light trapping, high absorptivity values can be obtained even for thin silicon layers in spite of silicon's small absorption coefficient.

The enhancement of the absorptivity 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 behaviour 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 2.1.1, 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 towards 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 towards the sun. The density of the photons in the semiconductor which are not absorbed or only weakly absorbed, then increases to a value which 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 which only exchanges radiation with the sun would of course have to track the position of the sun. If we make this effort, the absorptivity for absorbed photons with $\hbar\omega > \epsilon_G$ would increase much more steeply than shown in Figure 7.7, and silicon solar cells could be made still thinner. For radiative recombination only, 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.

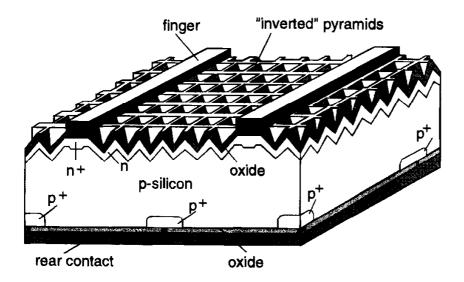


Figure 7.8: Structure of the best silicon solar cell manufactured to date with an efficiency of 24.4%, developed by M.Green's group.

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 7.7 we see that a $20\,\mu\rm m$ thick silicon film with light trapping has a higher absorptivity than a $400\,\mu\rm m$ thick wafer without light trapping. A higher absorptivity 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 absorptivity. Decreasing the thickness, while maintaining a high absorptivity makes the solar cell more ideal.

A high-efficiency, thin-film silicon solar cell would be a great achievement because of silicon's favourable 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, poly-crystalline silicon films have been grown on a variety of substrates including glass, a very promising development.

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.² Its efficiency is 24.4%. Figure 7.8 illustrates its structure.

²J. Zhao, A. Wang, M. Green, F. Ferrazza, Appl. Phys. Lett. **73** (1998) 1991.

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 which can be absorbed, a thickness of only a few μ m 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 reductions for the production of solar cells.

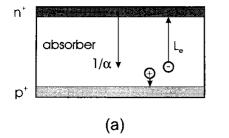
Due to the close proximity to the surface with its high surface recombination, a well developed membrane is required at least on the front surface. Since electron—hole pairs should not be generated in this layer, it must have a large energy gap. It is called a window layer, through which photons pass 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 favourable 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 hetero-junctions. 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 Eq. (2.3), 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 zone. 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 zone 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 in the efficiency of solar cells made of a-Si:H.

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, 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 of good luminescent quantum yields, indicating dominant radiative



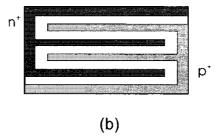


Figure 7.9: (a) In the plane arrangement of an absorber between electron and hole membranes, the diffusion lengths $L_{e,h}$ must be larger than the thickness of the absorber and the thickness must be larger than the penetration depth $1/\alpha$ of the photons. (b) Many absorbing layers in a meander-like structure combine good absorption with a small distance between the membranes.

recombination. Their drawback are the very small mobilities and diffusion lengths of electrons and holes. In the common plane arrangement of an absorber between two membranes, shown on the left of Figure 7.9, 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 which is twice 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 7.9 (b) fulfils 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, TiO₂ for the electrons and the I^-/I^{3-} -redox system for the holes, has a minimal value, the thickness of the mono-molecular dye layer. Nevertheless, many dye layers provide for sufficient absorption.

The organic or plastic solar cell is another example.³ In this solar cell a p-type polymer serves both as absorber and as 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 7.9 (b), 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 indium-tin-oxide electrode on the TiO₂ and the redox system is very poor, but is excellent with the small platinum islands on the counter electrode.

With the small distance between the membranes and the resulting large interface area, interface recombination is enhanced and may be a problem.

³C. J. Brabec, N. S. Sariciftci, J. C. Hummelen, Adv. Funct. Mater. 11 (2001) 15.

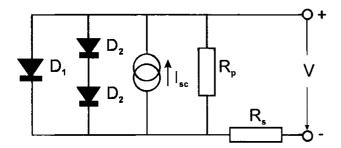


Figure 7.10: Equivalent circuit for a solar cell consisting of (from left) diode D_1 with direct recombination, diodes D_2 with impurity recombination, current source I_{sc} , parallel resistance R_p and series resistance R_s .

The plastic solar cell is again an example of the case where the mixing on a nanometer scale of electron and hole membranes prevents the build-up of space charge and electric fields. Charge carriers are, nevertheless, separated and a diode current—voltage characteristic is observed, demonstrating that the electric field and a potential barrier, both absent in plastic solar cells, are not the causes for the properties of diodes.

7.5 Equivalent circuit

In the current-voltage characteristic for the solar cell in Eq. (6.35) we can regard the current I_Q as the sum of the current through the pn-junction in the dark and the current I_{sc} from a current source, connected in parallel for the currents to add.

Figure 7.10 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 which 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

$$I_{Q} = I_{S1} \left[\exp \left(\frac{e(V - I_{Q}R_{S})}{kT} \right) - 1 \right] + I_{S2} \left[\exp \left(\frac{e(V - I_{Q}R_{S})}{2kT} \right) - 1 \right] + I_{sc} + \frac{V - I_{Q}R_{S}}{R_{p}} . (7.13)$$

Figure 7.11 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 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 and the solar cell must therefore have a higher temperature than the environment. For solar irradiation of $1 \, \text{kW/m}^2$ the temperature difference to the environment may be some $10 \, \text{K}$.

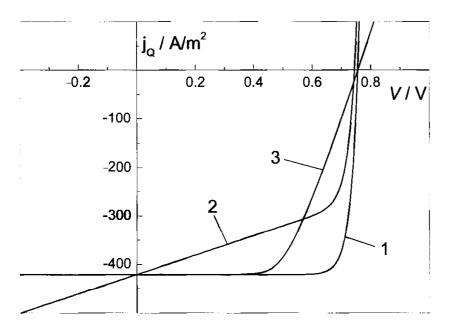


Figure 7.11: Current-voltage characteristic of a solar cell with: (2) $R_S = 0\Omega$, $R_p = 50\Omega$; and (3) $R_S = 5\Omega$, $R_p = \infty$, compared with (1) $R_S = 0\Omega$ and $R_p = \infty$.

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

$$V_{\rm oc} = \frac{1}{e} (\eta_e + \eta_h) = \frac{kT}{e} \ln \left(\frac{n_e n_h}{n_i^2} \right) , \qquad (7.14)$$

we find for the temperature dependence

$$\frac{\mathrm{d}V_{\mathrm{oc}}}{\mathrm{d}T} = \frac{k}{\mathrm{e}} \ln \left(\frac{n_{\mathrm{e}} n_{\mathrm{h}}}{n_{\mathrm{i}}^2} \right) + \frac{kT}{\mathrm{e}} \left[\frac{1}{n_{\mathrm{e}}} \frac{\mathrm{d}n_{\mathrm{e}}}{\mathrm{d}T} + \frac{1}{n_{\mathrm{h}}} \frac{\mathrm{d}n_{\mathrm{h}}}{\mathrm{d}T} - \frac{1}{n_{\mathrm{i}}^2} \frac{\mathrm{d}\left(n_{\mathrm{i}}^2\right)}{\mathrm{d}T} \right] . \tag{7.15}$$

Here

$$n_{\rm i}^2 = N_{\rm C} N_{\rm V} \exp\left(-\frac{\varepsilon_{\rm G}}{kT}\right)$$

and

$$\frac{\mathrm{d}\left(n_{\mathrm{i}}^{2}\right)}{\mathrm{d}T} = \frac{\varepsilon_{\mathrm{G}}}{kT^{2}} n_{\mathrm{i}}^{2} .$$

It follows that

$$\frac{\mathrm{d}V_{\mathrm{oc}}}{\mathrm{d}T} = \frac{V_{\mathrm{oc}} - \varepsilon_{\mathrm{G}}/\mathrm{e}}{T} + \frac{kT}{\mathrm{e}} \left(\frac{1}{n_{\mathrm{e}}} \frac{\mathrm{d}n_{\mathrm{e}}}{\mathrm{d}T} + \frac{1}{n_{\mathrm{h}}} \frac{\mathrm{d}n_{\mathrm{h}}}{\mathrm{d}T} \right) . \tag{7.16}$$

No general statements can be made about the expressions in parentheses, except that both are probably < 0, and that the first expression is negligible in an n-conductor and the second expression in a p-conductor. The temperature dependence is caused essentially by $(V_{\rm oc} - \varepsilon_{\rm G}/{\rm e})/T$. The temperature dependence is more pronounced in bad cells, where $V_{\rm oc}$ is small.

For a silicon cell with $V_{\rm oc} = 0.6\,\rm V$ and $\varepsilon_{\rm G} = 1.12\,\rm eV$ at $T = 300\,\rm K$, ${\rm d}V_{\rm oc}/{\rm d}T = -1.7\,\rm mV/K$. 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 all the better.

The short-circuit current j_{sc} is simply given by the absorbed photon current and increases, as expressed by Eq. (7.6), proportional to the intensity.

$$j_{\rm sc} = -e j_{\gamma, \rm abs} \tag{7.17}$$

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 Eq. (7.8)

$$V_{\rm oc} = \frac{kT}{e} \ln \left(1 + \frac{ej_{\gamma,abs}}{j_{\rm S}} \right) , \qquad (7.18)$$

where j_S is the reverse saturation current. For $j_{\gamma,abs} \gg j_S$, true for any observable light intensity, the open-circuit voltage V_{oc} increases logarithmically with the intensity.

The fill factor, finally, as given by Eq. (7.5) can be approximated for $V_{\rm oc} \gg kT$ roughly by

$$FF = 1 - \frac{kT}{V_{\text{oc}}} \tag{7.19}$$

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

$$\eta = \frac{FF j_{\rm sc} V_{\rm oc}}{\int_0^\infty \hbar \omega \, \mathrm{d}j_{\gamma, \, \rm sun}(\hbar \omega)} \,, \tag{7.20}$$

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 the 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 AMO spectrum, energy conversion with a solar cell is still well away from the theoretical limit of $\eta_{max}=0.85$ for the solar heat engine of Section 2.1.1. It is very instructive once again 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 7.12 shows the individual processes schematically.

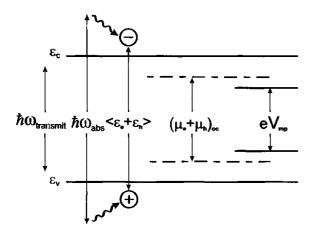


Figure 7.12: Individual processes in a solar cell.

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 $\langle \hbar\omega_{abs} \rangle$ of the absorbed photons. Let us assume that each absorbed photon generates just one electronhole pair which, at short-circuit, contributes to the current. Then

$$j_{E,abs} = j_{\gamma,abs} \langle \hbar \omega_{abs} \rangle = -\frac{j_{sc}}{e} \langle \hbar \omega_{abs} \rangle$$
 (7.21)

The absorption efficiency is

$$\eta_{abs} = \frac{j_{E,abs}}{j_{E,inc}} \,. \tag{7.22}$$

The second process is the thermalization of the electron-hole pairs, each produced initially with the mean energy $\langle \hbar \omega_{abs} \rangle$ which by thermalization is reduced to $\langle \epsilon_e + \epsilon_h \rangle = \epsilon_G + 3kT$. The efficiency of this process is

$$\eta_{\text{thermalization}} = \frac{\langle \varepsilon_e + \varepsilon_h \rangle}{\langle \hbar \omega_{\text{abs}} \rangle} \ . \tag{7.23}$$

The third factor defining the maximum chemical energy $(\mu_e + \mu_h)_{oc} = eV_{oc}$ which 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

$$\eta_{\text{thermodynamic}} = \frac{eV_{\text{oc}}}{\langle \varepsilon_e + \varepsilon_h \rangle} . \tag{7.24}$$

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 $-j_{\rm sc}V_{\rm oc}$ the solar cell delivers at the maximum power point as an electrical energy current $j_{\rm mp}V_{\rm mp}$.

$$FF = \frac{j_{\rm mp}V_{\rm mp}}{j_{\rm sc}V_{\rm oc}} \ . \tag{7.25}$$

The product of all these efficiencies gives the overall efficiency

$$\eta = \underbrace{\frac{j_{E,abs}}{j_{E,in}}}_{\eta_{abs}} \quad \underbrace{\frac{\langle \varepsilon_e + \varepsilon_h \rangle}{\langle \hbar \omega_{abs} \rangle}}_{\eta_{thermodynamic}} \quad \underbrace{\frac{eV_{oc}}{\langle \varepsilon_e + \varepsilon_h \rangle}}_{\eta_{thermodynamic}} \quad \underbrace{\frac{j_{mp}V_{mp}}{j_{sc}V_{oc}}}_{FF} = \frac{-j_{mp}V_{mp}}{j_{E,in}} .$$
(7.26)

For silicon, and in particular, for the $20\,\mu\text{m}$ thick cell with light trapping, whose absorptivity is shown in Figure 7.7, exposure to the AM1.5 spectrum gives the following values

$$\langle \hbar \omega_{abs} \rangle = 1.80 \,\mathrm{eV}$$

 $\langle \epsilon_e + \epsilon_h \rangle = \epsilon_G + 3kT = 1.2 \,\mathrm{eV}$
 $j_{sc} = 413 \,\mathrm{A/m^2}$ $j_{mp} = 401 \,\mathrm{A/m^2}$
 $V_{oc} = 0.770 \,\mathrm{V}$ $V_{mp} = 0.702 \,\mathrm{V}$.

The efficiencies are therefore

$$\begin{split} &\eta_{abs} = 0.74 \\ &\eta_{thermalization} = 0.67 \\ &\eta_{thermodynamic} = 0.64 \end{split}$$

$$FF = 0.89$$
.

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 electronhole pairs into chemical energy are particularly small and thus in need of improvement.

8 Concepts for Improving the Efficiency of Solar Cells

As was shown in the previous chapter, even avoiding all non-radiative 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 non-absorption of low-energy photons. In order to improve the efficiency, we must focus primarily on reducing these losses. We will 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 recently presented by M.Green. \(^1

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 $\epsilon_G < \hbar \omega < \epsilon_G + d\epsilon$ and processing the other photons by solar cells with a different band gap. Cells operated in this way are known as tandem cells.

For a black-body solar spectrum, a solar cell with the energy gap ε_G and the idealized absorptivity $a(\hbar\omega < \varepsilon_G) = 0$, $a(\hbar\omega \ge \varepsilon_G) = 1$, has the short-circuit current

$$j_{\rm sc} = -e \frac{\Omega_{\rm S}}{4\pi^3 \hbar^3 c^2} \int_{\epsilon_{\rm G}}^{\infty} \frac{(\hbar \omega)^2}{\exp\left(\frac{\hbar \omega}{kT_{\rm S}}\right) - 1} \, \mathrm{d}\hbar \omega \ . \tag{8.1}$$

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

$$j_{E,eh} = -j_{sc} \varepsilon_G / e = j_{\gamma,abs} \varepsilon_G$$
 (8.2)

Using $\langle \varepsilon_e + \varepsilon_h \rangle = \varepsilon_G$ instead of $\varepsilon_G + 3kT$ we here divide the overall efficiency η somewhat differently (and not entirely correct) into the thermalization and the thermodynamic efficiencies.

The shaded rectangle in Figure 8.1 indicates the energy current $j_{E,eh}$ transferred to the electron-hole pairs after thermalization, for the energy gap at which $j_{E,eh}$ has its maximum value.

¹M. A. Green, Third Generation Photovoltaics, Springer Verlag (2003).

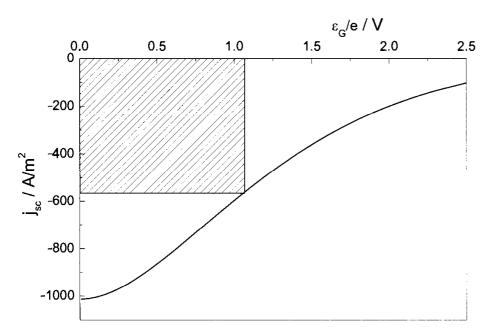


Figure 8.1: Short-circuit current of a solar cell as a function of the energy gap of its absorber ε_G , for a black-body spectrum with $T_S = 5800 \,\mathrm{K}$.

The area under the curve $j_{sc}(\varepsilon_G/e)$ is the entire incident energy current $j_{E,sun}$ 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

$$dj_{sc} = -e \frac{\Omega_S}{4\pi^3 \hbar^3 c^2} \frac{\varepsilon_G^2}{\exp\left(\frac{\varepsilon_G}{kT_S}\right) - 1} d\varepsilon_G$$
(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

$$j_E = -\int_0^\infty \frac{\varepsilon_G}{e} \, dj_{sc} = \frac{\Omega_S}{4\pi^3 \hbar^3 c^2} \int_0^\infty \frac{\varepsilon_G^3}{\exp\left(\frac{\varepsilon_G}{kT_S}\right) - 1} \, d\varepsilon_G . \tag{8.4}$$

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), Eq. (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 $j_{sc}(\epsilon_G/e)$ curve is the energy current, which is not utilized, because it is not absorbed by the solar cell. A solar cell with $\epsilon_G = 1.1 \, \text{eV}$ 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 efficiency by Shockley and Queisser² 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.

²W. Shockley, H. J. Queisser, J. Appl. Phys. **32** (1961) 510.

8.1 Tandem cells

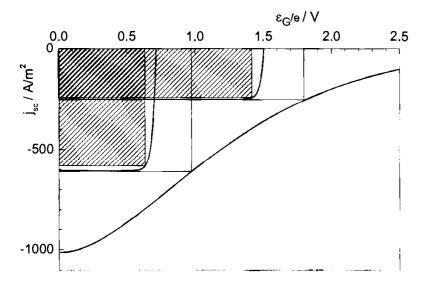


Figure 8.2: Current-voltage characteristics of two solar cells with energy gaps $\varepsilon_{G1} = 1.8 \,\text{eV}$ and $\varepsilon_{G2} = 0.98 \,\text{eV}$.

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 which 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 *AMO* spectrum the optimal combination consists of $\varepsilon_{G2} = 1.0 \, \text{eV}$ and $\varepsilon_{G1} = 1.9 \, \text{eV}$, yielding an overall efficiency of $\eta = 0.44$.

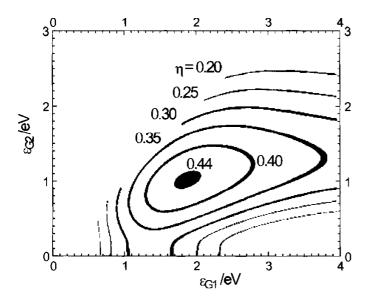


Figure 8.3: Efficiency for two solar cells in tandem operation, with energy gaps ε_{G1} and ε_{G2} for the AMO spectrum when their energy currents are added.

8.1 Tandem cells

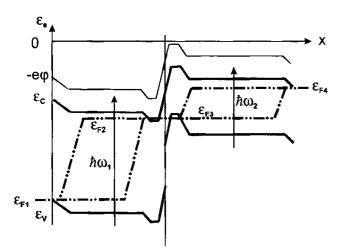


Figure 8.4: Electrical series connection of two solar cells by a tunnel junction provides efficient recombination of electrons and holes without requiring a difference between their Fermi energies.

solar cells. Figure 8.4 shows two cells which 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 ($\varepsilon_{F2} = \varepsilon_{F3}$) 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 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

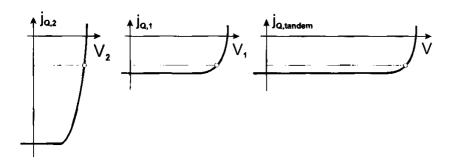


Figure 8.5: Current-voltage characteristic for the series connection of two solar cells with different short-circuit currents and open-circuit voltages. For each value of the current j_Q the voltages V_1 and V_2 are added to give the total voltage V.

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.

In order 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.³

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 delivered by a solar cell with smaller area than for non-concentrated radiation. Another advantage is that concentrated radiation can be processed with greater efficiency, as discussed in Chapter 7.

In areas with much more 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. In order 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

³T. Trupke, P. Würfel, J. Appl. Phys. **96** (2004) 2347.

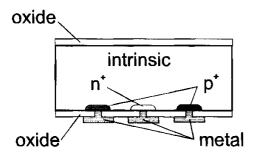


Figure 8.6: The p- and n-type membranes for holes and electrons and the metal contacts of the point-contact cell for concentrated radiation, are both placed on the back side of the cell.

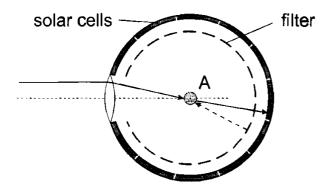


Figure 8.7: In the thermo-photovoltaic converter, the intermediate absorber is surrounded in an evacuated cavity by solar cells illuminated by its thermal radiation.

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.⁴ For a tandem a record efficiency of 38.9% for a monolythic configuration of three cells consisting of GaInP, GaInAs and Ge operating at a concentration of the solar radiation by a factor of 490 was reported.⁵

8.3 Thermo-photovoltaic energy conversion

The solar-thermal conversion method of Section 2.1.1 can be modified to be applicable to solar cells. Figure 8.7 illustrates the principle.

A focussing optical system is used to concentrate the solar radiation onto an intermediate absorber which, 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 $\varepsilon_G \leq \hbar \omega \leq \varepsilon_G + d\varepsilon$ without loss and reflects all other photons, which cannot be used optimally, back to the intermediate absorber. These photons, together with the photons emitted by the solar cell and transmitted by the filter, help to maintain the temperature T_A of the intermediate absorber. If the recombination is entirely radiative and the photons emitted by the cells are re-absorbed by the absorber and are therefore not lost, the cells can be operated close to their open-circuit voltage and thus

⁴R. M. Swanson, *Proc. 8. E.C. Photovoltaic Solar Energy Conference*, Florence 1988.

⁵M. Yamaguchi et al., Proc. 19. EU Photovoltaic Solar Energy Conference, Paris 2004.

have the efficiency derived in Eq. (4.5) for the conversion of the absorbed photon energy, i. e.,

$$\eta_{\text{cell}} = 1 - \frac{T_0}{T_A} \ . \tag{8.6}$$

Since this is the Carnot efficiency, which was used for the efficiency of the heat engine of the solar-thermal conversion process described in Section 2.1.1, we find, for the thermo-photovoltaic conversion process, the same efficiency as for the solar-thermal conversion, thus

$$\eta = \left[1 - \left(\frac{T_{A}}{T_{S}}\right)^{4}\right] \left(1 - \frac{T_{0}}{T_{A}}\right) \tag{8.7}$$

with a maximum value of $\eta = 0.85$ at an absorber temperature of $T_A = 2478 \, \mathrm{K}$.

The implementation of this concept in practice is difficult for two reasons. At the optimal temperature of the intermediate absorber of $T_A = 2478\,\mathrm{K}$ 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 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 $\hbar\omega \geq \epsilon_G$. 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 favourable for thermo-photovoltaic 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. Emission of only photons with $\hbar\omega \geq \epsilon_G$ 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, which works at about $1000\,^{\circ}\text{C}$ 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 way 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.

⁶M. A. Green, Proceedings 16th E.C. PV Solar Energy Conf. (2000) p. 51.

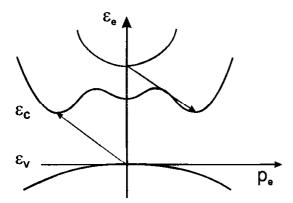


Figure 8.8: Transition of an electron from a higher band to the minimum of the conduction band by impact ionization in an indirect semiconductor, resulting in the additional generation of an electron and a hole at the band edges.

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, 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.⁷

In order 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.

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: while

⁷J. H. Werner, R. Brendel, H. J. Queisser, First World Conf. on Photovoltaic Energy Conversion, Hawaii, 1994.

⁸P. Würfel, Solar Energy Materials and Solar Cells, **46** (1996) 43.

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

$$dF = \ldots + \eta_e dN_e + \eta_h dN_h + \ldots = 0.$$

With impact ionization and Auger recombination, electrons and holes are always created or annihilated in pairs, that is $dN_e = dN_h = dN$ and

$$dF = \ldots + (\eta_e + \eta_h) dN + \ldots = 0.$$

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 $T_{\rm eh} = T_0$ and $\eta_{\rm e} + \eta_{\rm h} = 0$, 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 Eq. (2.32) 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 one contact and p-type for the transport of holes to the other 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 $\frac{3}{2}kT_A$ in the absorber to $\frac{3}{2}kT_0$ 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 can only occupy states over a narrow range $\Delta \varepsilon_e$ at the energy ε_e , as shown in Figure 8.9.

For $\Delta \varepsilon_{\rm e} < kT_0$, 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 $\varepsilon_{\rm h}$ in the hole membrane. The isentropic cooling therefore produces the chemical and electrochemical energy per electron-hole pair,

$$\mu_{\rm e} + \mu_{\rm h} = \eta_{\rm e} + \eta_{\rm h} = (\varepsilon_{\rm e} + \varepsilon_{\rm h})(1 - T_0/T_{\rm A})$$
 (8.8)

The arrangement of Figure 8.9 is a working solar cell. The voltage is

$$V = (\eta_e + \eta_h)/e , \qquad (8.9)$$

and the current is

$$j_Q = e(j_{E, absorbed} - j_{E, emitted}) / (\varepsilon_e + \varepsilon_h)$$
 (8.10)

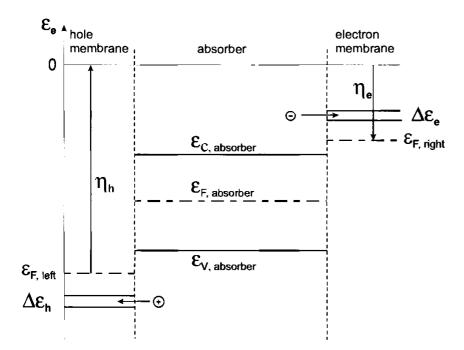


Figure 8.9: Energies of electrons and holes in the absorber, in which impact ionization and Auger recombination are in equilibrium at $T_A > T_0$, and in membranes through which electrons and holes flow outwards and where they are in temperature equilibrium with the environment at $T = T_0$.

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 current the energy in the electron-hole system decreases and with it the electron-hole temperature T_A in the absorber. Due to the lower temperature, the emitted energy current is reduced and also 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 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 Eqs. (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 Eq. (2.24), absorbs the energy current σT_S^4 and emits the energy

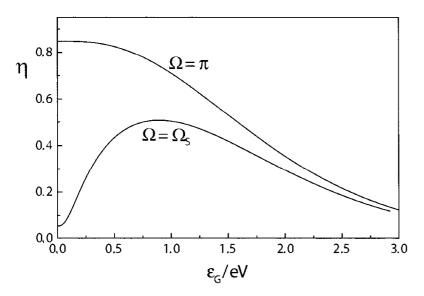


Figure 8.10: Efficiency for a hot carrier cell with impact ionization for non-concentrated incident solar radiation with $\Omega = \Omega_S$ and for maximum concentration with $\Omega = \pi$.

current σT_A^4 , at the temperature T_A . With Eqs. (8.9) and (8.10), the efficiency with which electrical energy is delivered is

$$\eta = \frac{j_{Q}V}{\sigma T_{S}^{4}} = \frac{\sigma(T_{S}^{4} - T_{A}^{4})}{\sigma T_{S}^{4}} \left(1 - \frac{T_{0}}{T_{A}}\right) = \left(1 - \frac{T_{A}^{4}}{T_{S}^{4}}\right) \left(1 - \frac{T_{0}}{T_{A}}\right)$$
(8.11)

and is therefore identical with the efficiency of the ideal solar heat engine in Eq. (2.52) in Section 2.1.1 and of the thermo-photovoltaic conversion process discussed in the previous section. The efficiency has its maximum value of $\eta_{max} = 0.85$ at a temperature of the electronhole system of $T_A = 2478 \, \text{K}$, if we assume a temperature of $T_S = 5800 \, \text{K}$ 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 AMO spectrum, however, a non-zero energy gap is preferable, otherwise more photons would be emitted than absorbed at small photon energies. For $\varepsilon_G > 0$ the balance becomes more favourable.

An earlier proposal for a hot-carrier solar cell by Ross and Nozik⁹ which did not account for impact ionization and Auger recombination, finds even higher efficiencies for narrow-gap semiconductors under less then 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 which keep the carrier concentrations constant and impact ionization and Auger recombination which don't, is physically impossible in narrow gap semiconductors. The problem becomes obvious for the case where electron-hole pairs are withdrawn with an

⁹R. T. Ross, A. J. Nozik, J. Appl. Phys. **53** (1982) 3813.

energy that is smaller than the average energy of the absorbed photons. The more electron—hole pairs are withdrawn, the more energy per pair is piling 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 non-radiative transitions between the bands and an impurity level. Impurities with energies for electrons in the middle of the 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 non-radiative recombination.¹⁰

The model, as shown in Figure 8.11, has states in the valence band with $\varepsilon_e \leq \varepsilon_V$, at the impurity level ε_{imp} and in the conduction band with $\varepsilon_e \geq \varepsilon_C$. In order to optimally utilize 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 ε_{imp} in the lower half of the energy gap (not in the middle), photons having energies $\varepsilon_{imp} - \varepsilon_V \leq \hbar \omega < \varepsilon_C - \varepsilon_{imp}$ are exclusively absorbed in transitions from the impurity. Photons having $\varepsilon_C - \varepsilon_{imp} \leq \hbar \omega < \varepsilon_C - \varepsilon_V$ are exclusively absorbed in transitions from the impurity to the conduction band, and photons having $\hbar \omega \geq \varepsilon_C - \varepsilon_V$ 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} cm² is a typical value. Although optical cross-sections vary with energy, we assume them to be constant over the energy range of absorbable photons.

¹⁰A. Luque, A. Marti, Phys. Rev. Lett. 78 (1997) 5014.

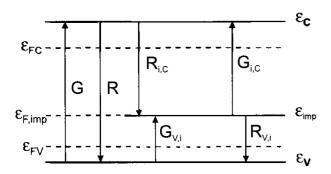


Figure 8.11: In addition to radiative band-band transitions with the rates G and R, radiative transitions between the bands and the impurity are taken into account. Non-radiative transitions are excluded.

The electrons and holes are assumed to have a high mobility 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

$$\frac{\partial n_{\rm e}}{\partial t} = G_{\rm bb} + G_{\rm i,C} - R_{\rm bb} - R_{\rm i,C} - \text{div } j_{\rm e} = 0$$
 (8.12)

$$\frac{\partial n_{h}}{\partial t} = G_{bb} + G_{V,i} - R_{bb} - R_{V,i} - \text{div } j_{h} = 0$$
(8.13)

$$\frac{\partial n_{\text{e,imp}}}{\partial t} = R_{\text{i.C}} - G_{\text{i,C}} + G_{\text{V,i}} - R_{\text{V,i}} = 0$$
(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. As a result 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 unfavourable 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 donor-like and half acceptor-like, and

$$\rho_O = e(n_h - n_e + n_{imp}/2 - n_{e,imp}) = 0.$$
(8.15)

Alternatively, if the impurities are donor-like, an occupation probability of $\frac{1}{2}$ 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

$$G_{bb} = \frac{1}{d} \int_{\epsilon_{C} - \epsilon_{V}}^{\infty} a_{bb} \, dj_{\gamma}(\hbar\omega)$$

$$G_{i,C} = \frac{1}{d} \int_{\epsilon_{C} - \epsilon_{imp}}^{\epsilon_{C} - \epsilon_{V}} a_{i,C} \, dj_{\gamma}(\hbar\omega)$$

$$G_{V,i} = \frac{1}{d} \int_{\epsilon_{imp} - \epsilon_{V}}^{\epsilon_{C} - \epsilon_{imp}} a_{V,i} \, dj_{\gamma}(\hbar\omega)$$
(8.16)

The absorptivity for the band-band transitions is assumed to be $a_{bb} = 1$, whereas the absorptivities for the impurity transitions depend on the concentration and occupation of the impurities,

$$a_{i,C} = 1 - \exp(-\alpha_{i,C} d)$$
 $a_{V,i} = 1 - \exp(-\alpha_{V,i} d)$

where the absorption coefficients follow from Eq. (3.96)

$$\alpha_{i,C} = \sigma_{i,C} n_{imp} (f_i - f_C)$$
 and $\alpha_{V,i} = \sigma_{V,i} n_{imp} (f_V - f_i)$

These equations determine the generation rates 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 Eq. (3.101),

$$\frac{\mathrm{d}R_{i,j}}{\mathrm{d}\hbar\omega} = \frac{1}{d} a_{i,j} \frac{1}{4\pi^2 \hbar^3 c_0^2} \frac{(\hbar\omega)^2}{\exp\left[\frac{\hbar\omega - (\varepsilon_{\mathrm{F},j} - \varepsilon_{\mathrm{F},i})}{kT} - 1\right]}$$
(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

$$j_Q = -e \operatorname{div} j_e \ d = -e \operatorname{div} j_h \ d$$
 (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 $V = (\varepsilon_{FC} - \varepsilon_{FV})/e$. 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.

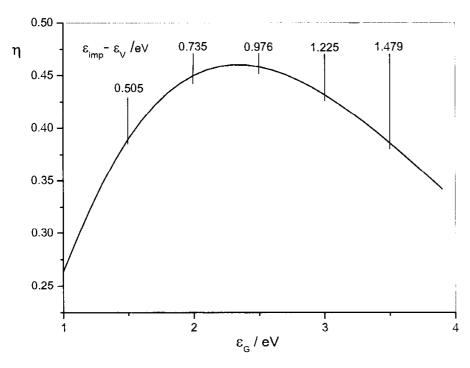


Figure 8.12: Efficiency as a function of the energy gap $\varepsilon_C - \varepsilon_V$ for radiative band-band transitions and radiative transitions between the bands and an impurity level at ε_{imp} . Non-radiative transitions are excluded. The numbers at the curve give the optimal position of the impurity level with regard to the valence band for selected band gaps.

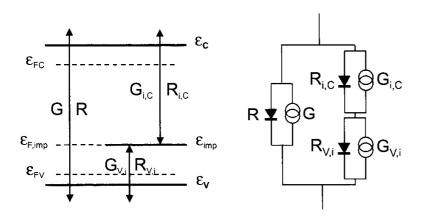


Figure 8.13: Equivalent circuit for a solar cell with an impurity level between valence and conduction bands as shown in Figure 8.11.

The efficiency reaches a maximum value of $\eta=0.46$ for an energy gap $\epsilon_C-\epsilon_V=2.4 \, eV$ and an impurity level at $\epsilon_{imp}-\epsilon_V=0.93 \, eV$.

This result reminds us of the improvement of the efficiency by tandem cells, since 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 absorptivities 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 non-radiative 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 non-radiative recombination is most probable, including non-radiative 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 which 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.

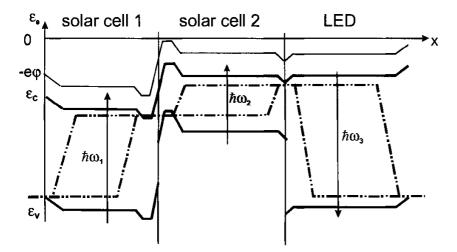


Figure 8.14: Two solar cells with small band gaps drive a LED with a large band gap, to emit photons useful for a large band gap solar cell, thereby up-converting two small energy photons into one higher energy photon.

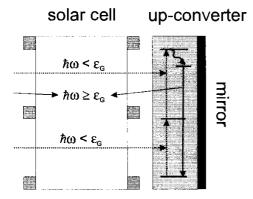


Figure 8.15: An up-converter behind a solar cell absorbs small energy photons, transmitted by the solar cell, in a 2-step excitation process. Higher energy photons with $\hbar\omega \geq \epsilon_G$ emitted by the up-converter generate additional electron—hole pairs in the solar cell.

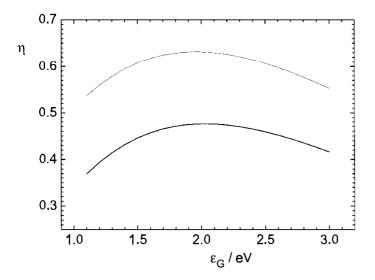


Figure 8.16: Efficiency of a solar cell as a function of its band gap ε_G operating with directly absorbed and with up-converted photons from a 6000 K black body spectrum, from a solid angle $\Omega_S = 6.8 \times 10^{-5}$ (thick line) and with maximum concentration from $\Omega = \pi$ (thin line).

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 towards the solar cell. A closer inspection of the impractical 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 along the line 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 re-emission of two small-energy photons.¹¹

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 non-concentrated radiation. The possible efficiencies of a solar cell with

¹¹T. Trupke, M. A. Green, P. Würfel, J. Appl. Phys. **92** (2002) 4117.

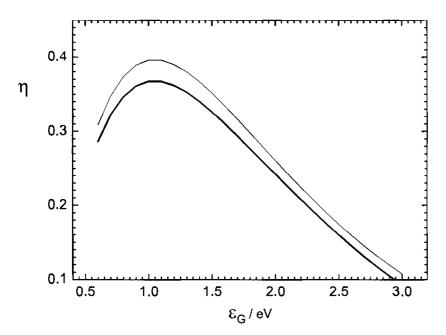


Figure 8.17: Efficiency of a solar cell as a function of its band gap ε_G operating with directly absorbed and with down-converted photons from a non-concentrated 6000 K black-body spectrum for a down-converter placed on the rear side of the solar cell (thin line). The efficiency for operation only with photons from the down-converter when it is placed on the front side is smaller (thick line).

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 has a high quantum efficiency. Second, the up-converter is a separate device which can be applied to existing well developed bifacial solar cells. Third, since the up-converter is separated from the solar cell, it would very little interfere with the recombination processes in the solar cell. Applying an up-converter to a solar cell would not do any harm, it 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. 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 $\hbar\omega > 2\,\epsilon_G$ by splitting these photons into two photons with $\hbar\omega > \epsilon_G$. 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.

¹²T. Trupke, M. A. Green, P. Würfel, J. Appl. Phys. **92** (2002) 1668.

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 solar photons, which the solar cell could absorb, are absorbed by the down-converter. Photons with $\epsilon_G \leq \hbar\omega < 2\epsilon_G$ 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 $\hbar\omega > \epsilon_G$ 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 towards 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 which has the same large index of refraction as the solar cell, all the photons emitted towards the solar cell enter the solar cell without reflection, whereas most of the photons emitted towards 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 favourable with an up-converter. For both systems, the calculations consider only radiative transitions, which may, however, be closer to reality in materials which do not require good transport properties for electrons or holes.

9 Prospects for the Future

In the first chapter we saw that our present energy economy cannot continue in its current form, because we are in the process of changing the environment in which we live. Mankind and all other forms of life have developed over millions of years by adapting to the conditions around us in a continuous process of evolution. We have good reason to describe changes to these conditions, which are too rapid for life to adapt to, as natural disasters. Life on earth is an extremely complex system, and even today we do not fully understand the interrelations within this system. If we do not wish to risk life as a whole or parts thereof by significant departures from the present equilibrium, we should tolerate only minor changes in the prevailing conditions. Every system reacts linearly to minor changes, even our complex ecosystem. Small changes to the natural processes result in small changes of the properties of the environment. The restriction to minor changes means that we may make use only of processes existing in the natural state of equilibrium. The burning of wood, coal, oil and gas occurs as a natural process as well, resulting in the natural production of CO₂, CO and SO₂. Burning of wood, coal, oil and gas could thus be tolerated, if their burning by man only caused minor changes to natural processes. However, we are presently proceeding rapidly to violate this condition. The amounts produced by mankind are no longer minor changes. The condition of causing only minor changes is, in any case, violated when processes are used or substances produced which do not exist in nature and it is almost impossible to predict how the environment will react. This is particularly true for many waste products from nuclear energy use. An example which shows what can happen if formerly non-existing substances are introduced into the environment is that of chlorinated fluoro-hydrocarbons (CFC), which do not occur at all in nature. They were regarded as entirely harmless, since they are non-toxic and chemically inert. It came as a great surprise to learn that they in fact destroy the ozone layer and also are, to a large extent, responsible for the greenhouse effect. Many more examples exist, where the violation of the condition of minor changes has led to unpleasant surprises.

No such surprises are to be expected when electricity is generated from solar energy by solar cells. In making use of the processes taking place in a solar cell, we are only linking ourselves into processes which would in any case occur without us. In our absence, solar radiation would be absorbed by the earth and, in part, be reflected back into space. In the course of this process, the earth heats up to just the temperature at which it can re-emit the energy current absorbed from the sun. It is very important that we do not alter this process significantly. Viewed from the standpoint of thermodynamics, the solar heat, very valuable at the sun's temperature of roughly 6000 K, is cooled down to the earth's temperature, where it is practically worthless and is then emitted into space. What is changed, if the solar radiation is processed by solar cells? Part of the absorbed heat (in fact, most of it in real systems with

Physics of Solar Cells: From Principles to New Concepts. Peter Würfel Copyright ©2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 3-527-40428-7

efficiencies of 20% or less) is cooled at the site of the solar cells down to the temperature of the environment in the same way as it would be without the solar cells. The electrical power generated by the solar cells is then re-routed through consumers before friction and other dissipative processes finally degrade it to heat at the temperature of the environment, from where it is emitted into space. The energy balance between the absorbed and the emitted energy currents remains unchanged. With the use of solar cells, we simply allow the natural process of cooling the solar heat to take place in ways of greater benefit to us.

The preceding chapters have not only shown that solar cells are well suited for obtaining electrical power from solar energy, they have also shown that there is no better way to do this than with the intelligent utilization of solar cells, e.g., in tandem cell arrangements, because the possible efficiencies of these systems coincide with general efficiency limits predicted by thermodynamics. This has two consequences. On the one hand, there is no need to continue searching for other methods of harnessing solar energy which are fundamentally more efficient. On the other hand, since we cannot hope for future discoveries and technologies with significantly greater efficiencies, there is no reason to wait any longer to begin the serious development of a solar-energy economy.

Our present energy economy consumes oxygen and produces CO₂. Because of the fast and extensive spreading of gases in the atmosphere, this is a global, and not simply a local, problem. For heavily populated and industrialized areas such as Germany, where so much oxygen is burnt that none would be left to breathe, this spreading of the gases is very fortunate. But it also makes it very difficult for politicians to decide on radical changes in the present energy economy, because continuing with the present energy consumption has little consequences locally, or, the other way round, the required very considerable local efforts are not rewarded on a local basis if they are not implemented globally.

A global energy supply by solar energy on the present level must be easily possible, otherwise we would already be suffering from substantial global warming. If our power requirements were not small compared with the solar energy current reaching the earth, their coverage from resources would result in an increase in the temperature of the earth, even without the greenhouse effect in order to allow for the emission of this additional energy into space.

A quick estimate indicates that a solar-based global energy economy could, in principle, be implemented relatively easily. The greater part of the globally consumed $10 \times 10^{13} \, \text{kWh/a}$, is used to produce low-temperature heat, for heating buildings and for cooking. Even in countries with less than average sunshine like Sweden, most of these requirements can be met using well-insulated solar warm-water collectors. The rest, roughly $5 \times 10^{13} \, \text{kWh/a}$, could be generated by solar cells. Most of this would be used for the production of hydrogen, since this is an easily transported and easily stored form of chemical energy. In sunny areas, with incident solar radiation of more than $2000 \, \text{kWh/(a m}^2)$, a total efficiency of no more than $10 \, \text{percent}$ over an area of $500 \, \text{km}$ times $500 \, \text{km}$ equal to $2.5 \times 10^{11} \, \text{m}^2$, would be sufficient. Much larger areas are available in the sunny deserts. Nevertheless, covering such an enormous area with solar cells is presently unimaginable. The problems with a future solar-based energy economy would be alleviated, if we could reduce our energy requirements, or at least maintain them at their present level.

This vision of a solar-energy future, at the present time hardly possible for political reasons alone, must not let us lose sight of what solar energy can presently contribute to our energy requirements. Fortunately, there is enormous potential in the industrialized countries

themselves, because otherwise the technologies required for use in the deserts would hardly be developed. As an example of an industrialized country, we will assess the situation for solar generation of electrical power in Germany.

In Germany, there are about 80 million people living in an area of 357000 km². This gives a population density of 226 persons per km², with 4425 m² available to each person. In Germany, the sun supplies about 1000 kWh/(a m²), i.e., 115 W/m² averaged over the year. Over the area of 4425 m² per person, the sun supplies around 500 kW. Compared with this, the current power requirement of 5.7 kW/person, 0.76 kW per person of this as electrical power, appear as almost negligibly small. Not all of this area will be used with high efficiency. The Germans allow themselves the "luxury" of using 180000km², that is half of the total area of Germany, to satisfy an energy requirement of merely 0.1 kW per person. This, however, is our most important energy requirement, our food, produced on agricultural farmland and meadowland.

To satisfy the electrical power requirements with solar cells, assuming an efficiency of 20%, which will become possible in the near future, we would need an area of 33 m² per person. This is very nearly as much as the average area of 35 m² living space available per person in Germany. Industrial buildings additionally account for at least the same area. Assuming three-story buildings on average gives a floor space of 23 m² per person for buildings in Germany. The roof surface areas will be somewhat greater. Only those oriented towards the north are unsuitable for solar cells. Furthermore, especially with multi-storied buildings, the wall areas oriented towards the south are very well suited.

From this estimate, we can see that the areas at and on top of existing buildings are already sufficient for nearly covering our present energy requirements by the use of solar cells, even in a country which does not receive the most sunshine. There is no reason whatever to speak of replacing forests with solar cells. Since there is more sunshine in summer than in winter and more energy is needed in winter than in summer, it is necessary to store energy in the summer for use in winter. This problem remains to be solved, and will certainly entail storage losses. This, of course, makes additional roof surface necessary. At present, though, a large amount of electrical power is wasted in producing low-temperature heat for heating buildings and for hot water. We could, in fact, manage with considerably less electrical power without any loss of comfort. Even for the relatively poor amount of sunshine in Germany, the enormous potential of solar energy for generating power without harm to the environment fully justifies the most intensive efforts to develop a solar-energy economy.

In view of the high population density and high rate of power consumption in industrialized countries like Germany, it seems more probable that not all of the power requirements will be met by utilizing only the solar energy captured in these countries. Even in an age of solar energy it is, in fact, more likely that industrialized countries will import energy from sunnier and less heavily populated countries, in very much the same way as today.

Appendix

Fundamental constants

Boltzmann constant $k = 1.3807 \times 10^{-23} \text{ W s/K} = 8.617 \times 10^{-5} \text{ eV/K}$

Planck constant $h = 6.626 \times 10^{-34} \,\text{W s}^2 = 4.136 \times 10^{-15} \,\text{eV s}$

 $\hbar = 1.0546 \times 10^{-34} \,\mathrm{W \ s^2} = 6.582 \times 10^{-16} \,\mathrm{eV \ s}$

velocity of light $c_{vac} = 2.998 \times 10^8 \, \text{m/s}$

elementary charge $e = 1.602 \times 10^{-19} \, A \, s$

permittivity of free space $\epsilon_0 = 8.85 \times 10^{-12} \, \text{A s/(V m)}$

Stefan–Boltzmann constant $\sigma = 5.67 \times 10^{-8} \, \text{W}/(\text{m}^2 \, \text{K}^4)$

solid angle of solar disc $\Omega_s = 6.8 \times 10^{-5}$

 $\hbar\omega\lambda = hc_{\rm vac} = 1.240\,{\rm eV}\,\mu{\rm m}$

 $\frac{1}{4\pi^3\hbar^3c_{\text{vac}}^2} = 5.04 \times 10^7 \, \frac{\text{W}}{(\text{eV})^4 \, \text{m}^2}$

Units of energy

 $1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$

1 J = 1 Ws = 1 Nm

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

Material constants at 300 K

	Ge	Si	GaAs
$\epsilon_{\rm G}/{\rm eV}$	0.66	1.12	1.42
χ/eV	4.13	4.01	4.07
ε	16	11.9	13.1
$N_{\rm C}/{\rm cm}^{-3}$	1×10^{19}	3×10^{19}	5×10^{17}
$N_{\rm V}/{\rm cm}^{-3}$	6×10^{18}	1×10^{19}	7×10^{18}
$n_{\rm i}$ / cm ⁻³	2.3×10^{13}	1×10^{10}	2.1×10^{6}
$m_{ m c}^*/m_{ m e}$	0.88	1.08	0.067
$m_{ m h}^*/m_{ m e}$	0.29	0.55	0.47
$b_{\rm e}/[{\rm cm}^2/({\rm V}{\rm s})]$	3800	1450	8500
$b_h/[cm^2/(V s)]$	1800	480	400

Physics of Solar Cells: From Principles to New Concepts. Peter Würfel Copyright ©2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

ISBN: 3-527-40428-7

Standard global AM1.5 spectrum

with $1000\,\mathrm{W/m^2}$ total radiation

	d i /d)		d i= /d)		
$\frac{\lambda}{\mu m}$	$\frac{\mathrm{d}j_E/\mathrm{d}\lambda}{\mathrm{W}/(\mathrm{m}^2\mu\mathrm{m})}$	$\frac{\lambda}{\mu m}$	$\frac{\mathrm{d}j_E/\mathrm{d}\lambda}{W/(m^2 \mathrm{cm})}$	$\frac{\lambda}{\mu m}$	$\frac{\mathrm{d}j_E/\mathrm{d}\lambda}{W/(m^2 cm)}$
			$W/(m^2 \mu m)$		$W/(m^2 \mu m)$
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
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

Index

Abbé sine condition, 26, 28	coupled equilibria, 55
absorption coefficient, 60, 79, 80	current_voltage characteristic
measurement, 83	for impurity recombination, 125
negative, 82	pn-junction, 120, 122, 124
absorptivity, 20, 21	1 3
negative, 82	Dember effect, 104
of semiconductors, 21	density of states
acceptors, 48	combined, 59
air mass, 23	effective, 43
AM0 spectrum, 23, 24, 157, 170	for electrons, 39, 41
AM1 spectrum, 23	for photons, 10, 12, 80
AM1.5 spectrum, 23, 24, 182	detailed balance, 64, 65, 71
ambipolar diffusion, 103	dielectric relaxation, 102
anti-reflection coating, 12, 142	diffusion current, 95
, , , , , , , , , , , , , , , , , , ,	diffusion length, 101, 120
battery, 55, 133, 136, 166	diffusion voltage, 116
black body, 9, 18	direct transitions, 59
radiation, 9	distribution function
Boltzmann constant, 181	for electrons, 38
Bose–Einstein distribution, 10	for photons, 10
	donors, 47
carbon dioxide, 3	doping, see semiconductors, doping
atmospheric concentration, 5	down-conversion of photons, 172
Carnot efficiency, 89	drift velocity, 95
Carnot engine, 29–31, 35, 37	·
cavity, 9, 37	earth
chemical energy, 85	energy emission, 6
chemical equilibrium, 54	temperature, 6, 8
chemical potential	effective mass, 41, 42
electrons, 94	efficiency
photons, 81	as function of energy gap, 140
collision time, 94	intensity dependence, 151
concentration of radiation, 24, 28	maximum, 29, 30, 88
concentrator, 29	monochromatic, 92
concentrator cells, 160	overall, 92
conduction band, 40, 46	Einstein relation, 96
germanium, 41	electric field in a solar cell, 133
conductivity, 95	electrochemical equilibrium

Physics of Solar Cells: From Principles to New Concepts. Peter Würfel Copyright © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 3-527-40428-7

in pn-junction, 116	f-num ber , 27, 28
electrochemical potential, 53, 55	focussing, 24
of electrons, 56	free energy, 55
of holes, 57	
electron affinity, 129	gallium arsenide, 42, 48, 128
electron capture, 70	generation rate, 63
cross-section, 70, 71, 73, 75	geometrical optics, 26
electron emission, 71	global warming, 178
electron-hole pairs	greenhouse effect, 5–7, 177, 178
average energy, 86	
chemical energy, 86, 87	hetero-junction, 127, 131, 147
equilibrium with photons, 89	holes, 45
generation, 58, 62–64, 81	homo-junction, 128
interaction with photons, 82	hot electrons, 165
momentum, 59	
recombination, see recombination	impact ionization, 63, 163
recombination, 79	impurity photovoltaic effect, 168
temperature, 85	impurity states, 70
elementary charge, 181	index of refraction, 25, 175
emission of photons, 79	silicon, 143
spontaneous, 79	indirect transitions, 61
stimulated, 79	
emissivity, 20, 81	Kirchhoff's law, 19
•	
energy	Lambert's law, 18
consumption, 2, 5	Landsberg efficiency, 33, 34
current from sun, 2	laser, 82
density, 13	laser condition, 82
dissipation, 98	lifetime, 74, 77
economy, 1, 177	for radiative recombination, 77
maximum reserves, 3	upper limit, 79
of radiation, 13	light
reserves, 2–4	polarization, 11
units, 1, 181	trapping, 142, 146
energy current, 19	velocity, 181
per solid angle, 15	
energy–momentum relation, 59, 60	maximum power point, 137
entropy, 30, 34	mean free path, 94
of electrons, 56	mem bran e
of holes, 56	for electrons, 113
of photons, 89	for holes, 113
equilibrium, see thermodynamic equilibrium	semi-permeable, 109
equivalent circuit, 149	metal contact, 129
	energy diagram, 130
Fermi distribution, 39	MIS contact, 132
Boltzmann approximation, 67	mobility, 95
Fermi energy, 56, 58	momentum space, 11
Fick's law, 96	
field current, 94	oil consumption, 5
fill factor, 138, 149, 151	open-circuit voltage, 124, 139, 146, 15

Index 185

optical transitions, 79	ideal, 89
ozone layer, 177	III-V, 76, 128, 160
	illuminated, 53
Pauli Principle, 38	indirect, 43, 61, 62
Peltier effect, 130	intrinsic, 44
permittivity of free space, 181	n-type, 49
phase space, 10	p-type, 50
phonon energy, 37	short-circuit current, 123, 124, 139, 140, 146
photoemission, 58	155
photomultiplier, 37	sine condition, see Abbé sine condition
photon absorption, 59	Snell's law of refraction, 12, 64
probability, 59	solar cells
photonic crystal structure, 145	chemical, 109, 113, 122
photons, 9	dye, 114
photosynthesis, 3	efficiency, 140
Planck constant, 181	energy conversion, 137
Planck's radiation law, 9, 13, 20	maximum efficiency, 137
generalized, 67, 79	minimal thickness, 147
pn-junction, 115, 116, 119, 120, 123, 133–135	plastic, 148
current-voltage characteristic, 120	silicon, 141
potential distribution, 117	thin film, 147
Poisson's equation, 117	solar disc, 16
• ,	solid angle, 15, 181
quasi-Fermi distribution, 51, 52	solar spectrum, 21, 22
•	solid angle, 12, 15
recombination, 65	space charge density, 117
Auger, 67	space charge layer, 128
efficiency, 74	spectrum, 13
non-radiative, 67	Staebler–Wronski effect, 147
radiative, 65, 79	Stefan–Boltzmann constant, 181
surface, 75	Stefan–Boltzmann radiation law, 6, 17, 18
velocity, 76	sun
via impurities, 68	as black body, 9, 21
recombination centers, 51	radiation spectrum, 6
recombination rate, 71, 78	spectrum, see solar spectrum
redox system, 114	surface charge, 130
reflection, 142	surface photo-voltage, 106
total internal, 12, 64, 143, 144	surface states, 75, 76
resistance	density, 75
parallel, 149	•
reaction, 122	tandem cells, 155
series, 149	series connection, 159
transport, 122	temperature dependence of V_{oc} , 149
reverse saturation current, 123, 124	temperature of the earth, 6
	thermalization, 52, 85
Schottky contact, 131, 132	thermo-photonics, 162
semiconductors, 37	thermo-photovoltaic conversion, 161
direct, 42	thermodynamic equilibrium, 54, 55
doping, 47	three-level system, 168, 173
A	· · · · · · · · · · · · · · · · · · ·

tunnel junction, 159 two-diode model, 125 two-step excitation, 168

uncertainty principle, 10, 11, 40 up-conversion of photons, 172

valence band, 40, 46

germanium, 41 venus atmosphere, 8 vibrational quanta, 37

warm-water collectors, 178 work function, 57 measurement, 58