

Solarzelle

Physik hinter modernen Technologien

Fachwissenschaftliche Vertiefung mit pädagogischem Fokus

ETH Zürich – Frühlingssemester 2011

Manfred Sigrist, HIT K 23.8 (Tel.: 3-2584,
Email: sigrist@itp.phys.ethz.ch)

Literatur:

- *Physics of Solar Cells - From Principle to New Concepts*, Peter Würfel, Physics Textbook, Wiley-VCH (2005)
- *The Physics of Solar Cells*, Jenny Nelson, Imperial College Press (2003).
- *The Physics of Semiconductor Solar Cells*, Brigitte Surer and Alain Hauser, Proseminar on "The Physics behind modern Technologies", WS 2005/06, on webpage: <http://n.ethz.ch/student/alhauser/download/solarcellsText.pdf> .
- *Semiconductor Devices*, S.M. Sze, John-Wiley & Sons (2002).
- *The semiconductor diode as a rectifier, a light source and a solar cell: A simple explanation*, F. Hermann and P. Würfel, Am. J. Phys. 74, 591 (2006).
- *Solid State Physics*, N.W. Ashcroft and N.D. Mermin, HRW International Edition (1981).
- *Quantenmechanik*, N. Straumann, Springer (2002).

Inhaltsverzeichnis

Einleitung und Geschichte	3
1 Halbleiter - Eine elementare Betrachtungsweise	6
1.1 Metalle und Halbleiter	6
1.2 Ladungsträger in einem Halbleiter	8
1.3 Drift- und Diffusionsströme	10
1.4 Dotierung der Halbleiter	12
2 Vom pn-Kontakt zur Solarzelle	14
2.1 pn-Kontakt im Gleichgewichtszustand	14
2.1.1 Gleichgewichtszustand	14
2.1.2 Raumladungsschicht am pn-Kontakt	15
2.1.3 Sperrschicht	17
2.1.4 Diffusionslänge für Minoritätsladungsträger	18
2.2 Strom-Spannungs-Charakteristik des pn-Kontaktes	18
2.3 Funktionen des pn-Kontaktes	20
2.3.1 Gleichrichter	20
2.3.2 Lichtquelle - LED	20
2.4 Solar-Zelle	21
2.5 Wirkungsgrad der Solarzelle	24
2.6 Experimentelle Bestimmung von E_g eines pn-Kontaktes	26
3 Licht - Klassisches Verständnis	27
3.1 Geometrische Optik	27
3.1.1 Strahlenoptik und Herons Prinzip	27
3.1.2 Lichtbrechung	28
3.1.3 Fermat'sches Prinzip	28
3.1.4 Doppelbrechung	31
3.2 Die Natur des Lichtes	31
3.2.1 Teilchen versus Welle	32
3.3 Licht als elektromagnetische Welle	33
3.3.1 Faraday-Effekt	33
3.3.2 Maxwell-Theorie	34
3.4 Licht und Materie	36
3.4.1 Transmission und Reflexion in transparenten Medien	36
3.4.2 Magneto-Optik und Faraday-Effekt	37
3.4.3 Antireflexion - Optische Beschichtung	38
3.5 Licht und Metalle	40
3.5.1 Dynamische Dielektrizitätsfunktion	40
3.5.2 Optische Eigenschaften des Metalls	41

4	Licht und Quantenmechanik	44
4.1	Planck'sche Strahlungsgesetz	44
4.2	Einsteins Lichtquanten-Hypothese	47
4.2.1	Photonengas in Hohlraumstrahlung	47
4.2.2	Photoelektrischer Effekt	48
4.3	Strahlungsabsorption und -emission	49
4.4	Übergangsraten der modernen Quantenmechanik	51
4.4.1	Goldene Regel	52
4.4.2	Harmonischer Oszillator	53
4.4.3	Übergangsrate für den harmonischen Oszillator	54
4.5	Quantisierung des elektromagnetischen Strahlungsfeldes	55
4.5.1	Elektromagnetische Strahlung als harmonischer Oszillator	55
4.5.2	Photonen sind Bosonen	56
4.5.3	Kopplung an Materie	56
4.6	Das Vakuum und der Casimir-Effekt	56
4.7	Photoemissionsspektroskopie	58
5	Elektronische Eigenschaften der Metalle und Halbleiter	61
5.1	Fermionen und Bosonen	61
5.2	Elektronenbänder	61
5.2.1	Von Atomorbitalen zur Bandstruktur	62
5.2.2	Semiklassische Beschreibung der Bandelektronen	65
5.3	Metalle - Isolatoren - Halbleiter	66
5.3.1	Metalle - Partiiell gefüllte Bänder	66
5.3.2	Isolatoren und Halbleiter	68
5.3.3	Bild der Löcher	69
5.3.4	Halbleiter bei endlicher Temperatur	69
5.4	Photoinduzierte Übergänge im Halbleiter	71
5.4.1	Direkter Übergang	72
5.4.2	Indirekter Übergang	74
5.5	Rekombination	75
5.6	Dotierung von Halbleitern	76
5.6.1	Wannier-Gleichung	76
5.6.2	Chemisches Potential	77
5.7	pn-Kontakt	78
5.8	Metall-Halbleiter-Kontakte	79
5.8.1	Raumladung	79
5.8.2	Strom-Spannungscharakteristik	80
6	Physik der globalen Erwärmung und des Treibhauseffekts	82
6.1	Energiebilanz - Kirchhoff'sches Gesetz	82
6.2	Einfachstes Modell der Erdenergie-Bilanz	84
6.3	Treibhausgase der Atmosphäre	85
6.3.1	Treibhaus-Prinzip	86
6.3.2	Treibhausgase und Absorption in der Atmosphäre	87
6.3.3	Reflektivität - Albedo	89
6.4	Einfaches Zweischichten-Modell	89
6.5	Globale Erwärmung	93

Einleitung

Als Solarzelle betrachten wir jede Vorrichtung, welche die Energie von Licht direkt in elektrische Energie umwandelt. Diese Definition schliesst Wärmekraftmaschinen aus, die mit Sonnenenergie via Sonnen-Kollektoren gespeist werden und einen elektrischen Generator antreiben. Es gibt natürlich noch viele andere Formen der Energiegewinnung aus Sonnenenergie. Als Beispiele kann man hier die ganzen biologischen Formen der Energiegewinnung anführen.

Historische Aspekte

Die erste photovoltaische Vorrichtung wurde von Alexandre-Edmond Becquerel (Vater von Henri Becquerel) entdeckt. Bei der Untersuchung von elektrolytischen Zellen mit Platin-Elektroden stellte er fest, dass ein Potentialunterschied zwischen einem belichteten und einem verdunkelten Teil entstand, der zu einem Strom zwischen den beiden Elektroden führte. Mit dem damaligen Verständnis der Physik konnte man jedoch diesen Effekt noch nicht erklären.

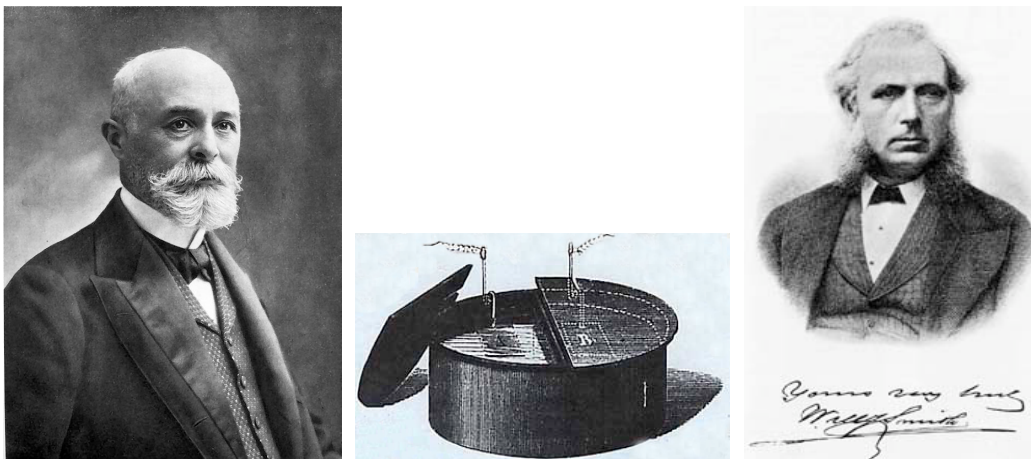


Abb. 0.1: Pioniere der Solartechnologie: Alexandre Becquerel (1820-1891); Becquerels elektrolytische Zelle; Willoughby Smith (1828-1891)

1873 stellte Willoughby Smith fest, dass man durch Lichtbestrahlung die elektrische Leitfähigkeit von Selen, das mit Platin kontaktiert wurde, verändern und bis um das 1000-fache erhöhen konnte. Man spricht hier von einem Photowiderstand oder einer Photodiode. Dieses Element konvertiert die Energie des Lichtes aber nicht in elektrische Energie, sondern es verändert nur dessen Transporteigenschaften.

Nur drei Jahre später gelang es William Grylls Adams und seinem Studenten Richard Evans Day mit Selen-Platin-Kontakten einen elektrischen Strom zu erzeugen, wenn sie belichtet wurden. Damit war die Umwandlung von Licht in elektrische Energie möglich. Die Effizienz war jedoch mit einem Wirkungsgrad von ungefähr 1% sehr gering, so dass an eine technische Ausbeutung des Effektes noch nicht zu denken war. 1883 produzierte Charles Fritts bereits Selen-Zellen mit einem Wirkungsgrad von bis zu 2%. Eine solche Photozelle bestand aus einem Selenkristall mit einer Metallschicht auf der Oberfläche. Die Grenzfläche funktionierte als sogenannte Sperrschicht.

Auf Selen basierende photoelektrische Zellen wurden bald für Belichtungsmessung und elektronische Bildübertragung verwendet. Auch die Tonfilm-Technologie profitierte, da nun optische Tonspeicherung (Tonstreifen an der Seite des Films) möglich wurde.

Gleichzeitig fing in der zweiten Hälfte des 19. Jahrhunderts die Grundlagenwissenschaft an, sich intensiver mit dem Phänomen der Photovoltaik und vor allem mit dem photoelektrischen Effekt zu beschäftigen. Interessanterweise finden gerade hier einige der entscheidenden Schritte hin zur modernen Physik statt. Diese schlossen eine Reihe von experimentellen Pionierarbeiten von Heinrich Hertz, Philipp Lenard, Wilhelm Hallwachs und vielen anderen ein. Auf der theo-

retischen Seite stehen die revolutionären Arbeiten von Max Planck und Albert Einstein, die ein neues Zeitalter der Naturwissenschaften einläuteten. Doch dazu später.



Abb. 0.2: Pioniere der Solartechnologie: Charles Pearson, Daryll Chapin und Calvin Fuller stellen die erste Solarzelle mit der pn-Struktur in den Bell Laboratories her.

Im 20. Jahrhundert wurden mit der Entdeckung der Halbleitereigenschaften von Silizium und der Fähigkeit Si-Einkristalle hoher Reinheit herzustellen neue Wege eröffnet. Ein Schritt hin zu den modernen Photozellen wurde um 1940 von Russel Ohl gemacht, der als Erster entdeckte, dass seine Siliziumproben veränderte elektrische Eigenschaften zeigten, sobald sie Licht ausgesetzt wurden. Den entscheidenden Durchbruch schafften jedoch 1954 in den Bell Laboratories Daryl Chapin, Calvin Fuller und Gerald Pearson mit ihren Solarzellen, die auf dem Konzept der Halbleiter-pn-Kontakte basierten. Letztere bilden die elementaren Bausteine der modernen Halbleiter-Elektronik und waren nur wenige Jahre früher entwickelt worden. Die heutzutage verwendeten Solarzellen basieren im Wesentlichen auf der damals entwickelten Struktur. Es gibt jedoch neuere Entwicklungen von Solarzellen, die auf anderen Prinzipien aufbauen, z.B. die sogenannten Graetzel-Zellen (dye sensitived solar cells) oder photoelektrochemische Zellen.

Prinzip der Solarzelle

Das Prinzip der Solarzellen und der Photodioden beruht darauf, dass die Bestrahlung eines Materials durch Licht Ladungsträger freisetzen kann. Diese Ladungsträger bestehen aus negativ geladenen Elektronen und positiv geladenen Löchern (Ladungserhaltung). Durch den Einfluss eines elektrischen Feldes können diese Ladungen getrennt werden, so dass ein Strom dieser Ladungen fließen kann. Falls eine geeignete Konfiguration eines elektrischen Feldes intern im Material vorhanden ist, kann die Ladungsseparation zu einer Spannung führen und als elektrische Energiequelle verwendet werden.

Die frühen Selen-Photoelemente basieren auf dem sogenannten Sperrschicht-Element. Selen als Halbleiter (oder Halbmetall) wird mit einem Metall (in moderneren Elementen Kadmium-Legierung) in Kontakt gebracht. Über der Kontaktschicht bildet sich aufgrund der Ladungsverteilung (Raumladung) eine elektrische Dipolschicht, die eine Separierung der photo-erzeugten Ladungsträger ermöglicht.¹ Solche einfachen Photoelemente können als Solarzellen verwendet

¹*Dipolschicht:* Eine Dipolschicht induziert eine "Unstetigkeit" des elektrostatischen Potentials Φ . Betrachte eine Dipolschicht, die aus zwei (unendlich) nahe beieinander liegenden, entgegengesetzt geladenen Flächen erzeugt wird. Der Abstand der Flächen sei d und die Ladungsdichte $\pm\sigma$, so dass die Dipoldichte pro Fläche durch $D = \sigma d$ gegeben ist. Als ladungsneutrale, unendliche Fläche erzeugt die Dipolschicht im Äusseren kein elektrisches Feld, sondern nur zwischen den beiden Ladungsschichten. Das elektrische Feld ist verantwortlich für den Unterschied

werden, ergeben aber einen sehr kleinen Wirkungsgrad ($\eta \sim 1\%$).² Effizientere moderne Solarzellen bestehen aus Halbleiter-Heterostrukturen, sogenannten pn-Kontakten, die durch ihre Raumladung eine Sperrschicht bilden. Hier liegen die Wirkungsgrade über 20 % (kristallines Si, III-V Halbleiter) und können durch viele Tricks noch erhöht werden.

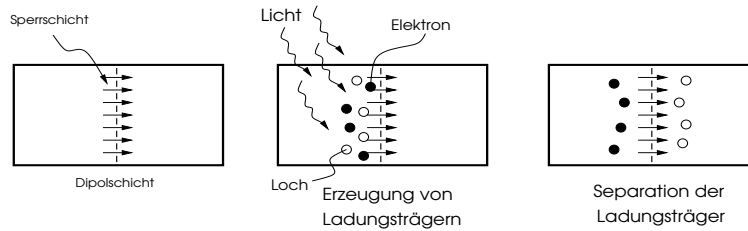


Abb. 0.4: Prinzip eines Photoelementes: Sperrschicht erzeugt eine Dipolschicht; Licht erzeugt bewegliche Ladungsträger (Elektronen und Löcher); Dipolschicht separiert Elektronen und Löcher (Elektron nach links, Loch nach rechts).

Die separierten Ladungen erzeugen eine Spannung über dem linken und rechten Ende des Elementes und können in einem geschlossenen Kreis zu einem Strom führen. Bezeichnet U_L die Spannung im offenen Stromkreis (kein Strom) und I_L den Strom im kurzgeschlossenen Stromkreis (keine Spannung), so gilt für den Wirkungsgrad $\eta < U_L I_L / P_L$. Die maximale Leistung hängt von der sogenannten Strom-Spannungskennlinie des pn-Kontaktes ab, wie wir später sehen werden.

Themen

Ausgehend von der Solarzelle können viele verschiedene physikalische Aspekte angeschnitten werden. Dazu gehört die Halbleiterphysik innerhalb des grösseren Bereiches der Festkörperphysik. Auch fundamentalere Konzepte, wie die Natur des Lichtes und die Wechselwirkung des Lichtes mit Materie müssen zur Sprache kommen, was uns geradewegs in Richtung Quantenmechanik weist. Ferner liegen sehr moderne Themen wie das Problem der globalen Erwärmung im Bereich der Solarenergie.

des elektrostatischen Potentials auf den beiden Seiten.

$$\Delta\Phi = \Phi_2 - \Phi_1 = - \int_0^d ds E = 4\pi \int_0^d \sigma = 4\pi\sigma d = 4\pi D \quad (1)$$

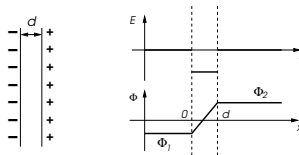


Abb. 0.3: Dipolschicht induziert Unstetigkeit im elektrostatischen Potential.

²Wirkungsgrad η einer Solarzelle: Verhältnis zwischen maximal extrahierbarer elektrischer Leistung P_E und der Leistung P_L des einfallenden Lichtes:

$$\eta = P_E / P_L . \quad (2)$$

Kapitel 1

Halbleiter - Eine elementare Betrachtungsweise

Der Grossteil der modernen Photovoltaik und praktisch alle kommerziellen Solarzellen basieren auf der Halbleiter-Technologie. Halbleiter spielen aber nicht nur im Bereich der Photovoltaik eine wichtige Rolle, sondern sie sind auch die Grundlage einer der wichtigsten Schlüsseltechnologien, die unser tägliches Leben immens beeinflusst. Daher wollen wir hier ein einfaches Bild einiger elektronischer Eigenschaften des Halbleiters und des einfachsten Bauelementes, des pn-Kontaktes, entwickeln, ohne zunächst auf die komplexeren Probleme der Bandstruktur und Phänomene der Quantenmechanik einzugehen. Dies wird Sache der späteren Kapiteln sein, in denen wir die in diesem Kapitel gemachten Aussagen in einen grösseren Rahmen einbinden werden.

1.1 Metalle und Halbleiter

Metalle und Halbleiter unterscheiden sich durch ihre elektrische Leitfähigkeit. Metalle sind im Allgemeinen gute Leiter, während Halbleiter sich wie Isolatoren verhalten. Charakteristisch für die Unterscheidung ist das Verhalten der Leitfähigkeit, wenn die Temperatur sich ändert. In Metallen steigt die elektrische Leitfähigkeit mit sinkender Temperatur an, während sie in Halbleitern absinkt und schliesslich beim absoluten Nullpunkt verschwindet.

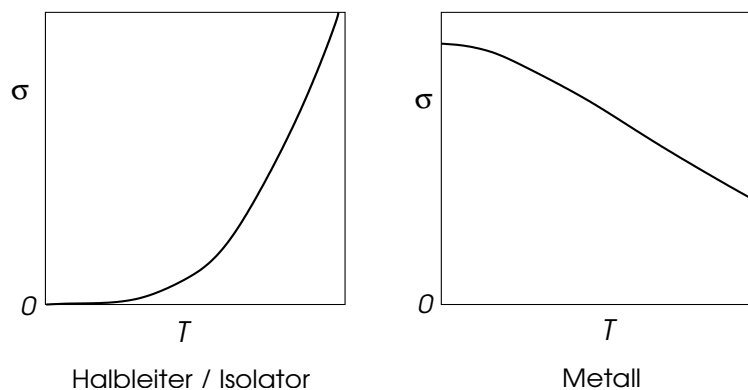


Abb. 1.1: Schematische Temperaturabhängigkeit der Leitfähigkeit eines Halbleiters (links) und eines Metalls (rechts).

Ein wesentlicher Unterschied zwischen Metallen und Halbleitern liegt in der Beweglichkeit der Elektronen. Metalle bestehen aus einem Gitter von positiv geladenen Ionen (Kern + Rumpfelektronen) und frei beweglichen, negativ geladenen Elektronen (Leitungselektronen), die zusammen ein neutrales Material bilden. Die Elektronen zwischen den Ionen führen aufgrund ihrer Bewegung und Wechselwirkung mit den Ionen zu einer Kohäsion, die das Metall zusammenhält. Die

strukturelle Stabilität der Halbleiter hingegen wird durch starke kovalente Bindungen sichergestellt. Die Valenzelektronen sind in diesen Bindungen im Wesentlichen gefangen und müssen zuerst befreit werden, um sich bewegen zu können. Um diese Eigenschaften im Detail zu verstehen, benötigen wir das Konzept der Quantenmechanik und die Bändertheorie. Wichtige elementare Halbleiter sind Silizium (Si) und Germanium (Ge).

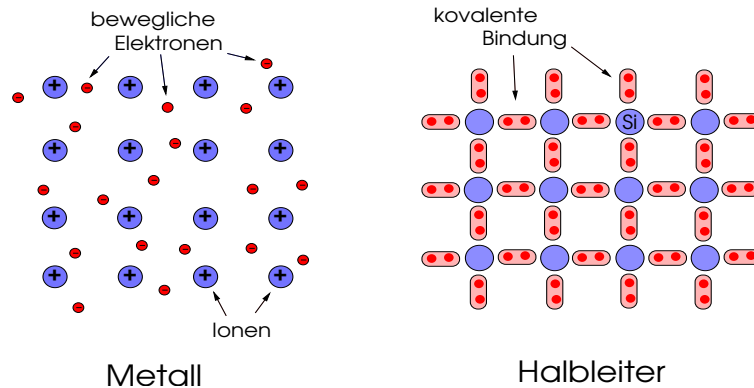


Abb. 1.2: Elektronen im Metall (links) und im Halbleiter (rechts).

Das Verhalten der Leitfähigkeit σ kann mit Hilfe einer relativ einfachen Betrachtung beschrieben werden. Die Leitfähigkeit ist durch die allgemeine Beziehung zwischen elektrischer Stromdichte \vec{j} und angelegtem elektrischem Feld \vec{E} definiert,

$$\vec{j} = \sigma \vec{E} . \quad (1.1)$$

Im Folgenden betrachten wir die Elektronen als klassische geladene Teilchen, welche die einfache Newton'sche Bewegungsgleichung befolgen und deren Bewegung durch viele mögliche Prozesse (Streuung an Verunreinigungen, Gitterverzerrungen, ...) gehemmt wird. Gemäss Newton gilt dann

$$\frac{d\vec{p}}{dt} = -e\vec{E} - \frac{\vec{p}}{\tau} , \quad (1.2)$$

wobei $-e$ die Ladung des Elektrons ist ($e = 1.6022 \times 10^{-19}$ As). Die Streuprozesse werden durch eine mittlere Zeit τ zwischen den Stößen der Elektronen beschrieben und stellen ein Dämpfung der Bewegung dar. Im stationären Zustand ändert sich der Impuls der Elektronen nicht mehr, d.h. es gilt $d\vec{p}/dt = 0$ und somit

$$\vec{p} = -e\tau \vec{E} . \quad (1.3)$$

Die Stromdichte wird durch die mittlere Dichte der beweglichen Elektronen, n , und die Geschwindigkeit $\vec{v} = \vec{p}/m$ ausgedrückt (m : Masse des Elektrons),

$$\vec{j} = -ne\vec{v} = -\frac{ne}{m}\vec{p} = \frac{ne^2\tau}{m}\vec{E} \quad \Rightarrow \quad \sigma(T) = \frac{n(T)e^2\tau(T)}{m} . \quad (1.4)$$

Die Temperaturabhängigkeit der Leitfähigkeit wird durch die Dichte $n(T)$ und der Streuzeit $\tau(T)$ bestimmt. Im Prinzip könnte sich aufgrund von Vielteilcheneffekten auch die Masse der Elektronen als Funktion der Temperatur ändern.

Für das Metall ist n im Wesentlichen temperaturunabhängig und τ steigt mit sinkender Temperatur an, d.h., die Streuprozesse werden bei tiefen Temperaturen seltener.¹ Im Halbleiter wird die Temperaturabhängigkeit durch die Dichte der beweglichen Ladungsträger dominiert. Durch

¹Einer der wichtigsten Streumechanismen ist die Elektron-Phonon-Streuung, d.h., Wechselwirkungen mit Schwingungen des Ionengitters. Letztere sind thermisch angeregte Fluktuationen und nehmen mit sinkender Temperatur ab. Beim Nullpunkt sind es nur noch Streuungen an Verunreinigungen, die einen Restwiderstand ergeben.

thermische Anregung können Ladungsträger im Halbleiter aus ihren kovalenten Bindungen herausgelöst werden. Die Ladungsträgerdichte (im Halbleiter) in Abhängigkeit der Temperatur verhält sich gemäss

$$n(T) = n_0 \left(\frac{T}{T_0} \right)^{3/2} e^{-E_g/2k_B T} \sim 10^{20} \text{ cm}^{-3} \left(\frac{T}{300\text{K}} \right)^{3/2} e^{-E_g/2k_B T}, \quad (1.5)$$

wobei E_g eine charakteristische Energie ist, die wir später als Energiebandlücke identifizieren werden ($k_B = 1.3807 \times 10^{-23} \text{ JK}^{-1}$ ist die Boltzmann-Konstante). Für verschiedene Halbleiter liegt die Energielücke im Bereich von $\sim 0.5 - 1.8 \text{ eV}$, wodurch sich in einem typischen Halbleiter bei Zimmertemperatur ($\sim 300 \text{ K}$) eine Konzentration von $10^4 - 10^{15} \text{ cm}^{-3}$ ergibt. Beachte, dass Isolatoren viel grössere Energielücken besitzen und folglich sehr viel kleinere Ladungsträgerkonzentrationen aufweisen (Diamant $E_g \approx 5.5 \text{ eV} \Rightarrow n \sim 10^{-73} \text{ cm}^{-3}$ bei Zimmertemperatur). Im Vergleich liegt die Ladungskonzentration bei Metallen bei $n \sim 10^{22} - 10^{24} \text{ cm}^{-3}$, also wesentlich höher.

1.2 Ladungsträger in einem Halbleiter

Wir werden nun einige einfache Betrachtungen zu den Ladungsträgern im Halbleiter anstellen, die uns später helfen werden, eine elementare Beschreibung der Funktion eines sogenannten pn-Kontaktes zu geben. Wie vorhin erwähnt, "entstehen" die mobilen Ladungsträger durch thermische Anregungen. Elektronen, die in den kovalenten Bindungen des Halbleiters gefangen sind, können durch Gitterschwingungen (Phononen) und elektromagnetische Strahlung (Photonen) aus ihrem gebundenen Zustand herausgelöst werden. Wir werden den Effekt der Phononen hier der Einfachheit halber beiseite lassen. Daher gehen wir davon aus, dass nur ein Photon genügend hoher Energie ein Elektron aus der Bindung befreien kann. In diesem Prozess wird jedoch eine zweite Ladung erzeugt, nämlich eine Vakanz in der betreffenden kovalenten Bindung. Auch diese kann sich frei bewegen; sie kann durch ein Elektron einer benachbarten Bindung besetzt werden, wodurch eine Vakanz in der benachbarten Bindung entsteht. Wir nennen diese zweite Art der Ladung "Loch". Sie hat die positive Ladung $+e$.

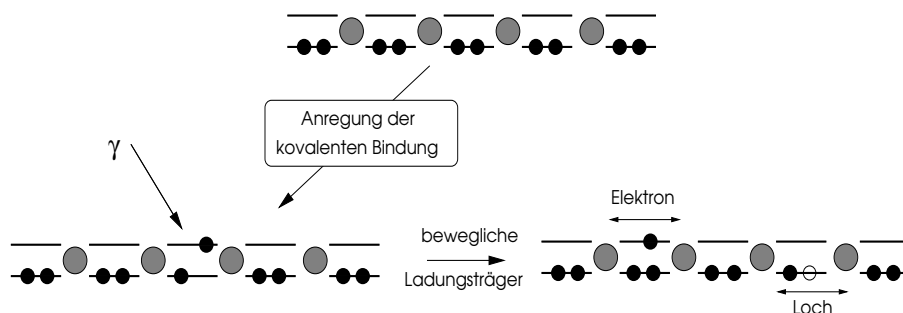


Abb. 1.3: Je zwei Elektronen sitzen im Grundzustand gerader Parität der kovalenten Bindung. Wegen des Pauli-Prinzips können diese Elektronen sich nicht zwischen verschiedenen Bindungen bewegen. Durch Anregung (Absorption eines Photons γ) wird ein Elektron in den angeregten Zustand ungerader Parität angehoben. Da dieser Zustand in anderen Bindung frei ist, kann sich das Elektron frei bewegen. Auch die Vakanz ist mobil. Damit haben wir zwei Typen freier Ladungsträger.

Die Bewegungseigenschaften der befreiten Elektronen und der erzeugten Vakanz sind im Allgemeinen unterschiedlich, so dass wir sie als zwei verschiedene Teilchensorten auffassen wollen. Wegen ihrer geringen Dichte können wir sie als ideales Gas betrachten. Beide, Elektronen und

Löcher, tragen zur Leitfähigkeit bei, indem sie sich unter einer angelegten Spannung in entgegengesetzter Richtung bewegen. Die Stromdichte kann geschrieben werden als

$$\vec{j} = -en_e\vec{v}_e + en_h\vec{v}_h = \vec{j}_e + \vec{j}_h, \quad (1.6)$$

wobei n_e und n_h die Konzentration (Dichte) der Elektronen (e) und Löcher (h : hole) bezeichnen und v_e und v_h die jeweiligen mittleren Geschwindigkeiten, die entgegengesetzt gerichtet sind.

Wir betrachten nun die Elektron-Loch-Erzeugung durch ein Photon als einen Prozess, um dessen Details wir uns zunächst nicht kümmern wollen. Vielmehr sehen wir darin eine Art chemische Reaktion,

$$\gamma \rightarrow e + h, \quad (1.7)$$

wobei γ das Photon genügender Energie bezeichnet.² Elektron und Loch können unter Abgabe eines Photons auch wieder verschwinden; sie "rekombinieren", indem sie wieder die ursprüngliche kovalente Bindung eingehen,

$$e + h \rightarrow \gamma. \quad (1.8)$$

Im Gleichgewicht sind beide Reaktionen gleich häufig,

$$\gamma \rightleftharpoons e + h. \quad (1.9)$$

Die Photonen, die wir hier betrachten, stammen aus der thermischen Strahlung des Halbleiters bei gegebener Temperatur, wobei natürlich nur diejenigen Photonen relevant sind, die genügend Energie besitzen. Aufgrund dieses Gleichgewichts enthalten Halbleiter bei endlicher Temperatur immer eine gewisse Konzentration beweglicher Ladungsträger. Für Silizium findet man bei Raumtemperatur $n_e = n_h \approx 10^{10} \text{cm}^{-3}$.

Beachte, dass bei den beiden Prozessen die elektrische Ladung erhalten bleibt, obwohl die Gesamtteilchenzahl sich verändert.

$$\text{Teilchenzahl: } n_\gamma + n_e + n_h \neq \text{konstant}, \quad (1.10)$$

$$\text{Ladung: } q_\gamma + q_e + q_h = 0,$$

mit n_γ als Photonendichte und $q_\gamma = 0$ als Photonenladung. Wir definieren die Ladungsdichte $\rho = q_en_e + q_hn_h = en_h - en_e$. Diese Ladungsdichte kann raum- und zeitabhängig sein, sie erfüllt aber stets die Kontinuitätsgleichung,

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0. \quad (1.11)$$

Im uniformen stationären Fall gilt $\rho = 0$. Für die Teilchenzahldichten hingegen ist die Kontinuitätsgleichung nicht erfüllt; die Änderung der Dichten ist komplexer:

$$\begin{aligned} \frac{\partial n_e}{\partial t} &= \Gamma_1 n_\gamma - \Gamma_2 n_h n_e - \vec{\nabla} \cdot (n_e \vec{v}_e), \\ \frac{\partial n_h}{\partial t} &= \Gamma_1 n_\gamma - \Gamma_2 n_h n_e - \vec{\nabla} \cdot (n_h \vec{v}_h), \end{aligned} \quad (1.12)$$

$$\frac{\partial n_\gamma}{\partial t} = -\Gamma_1 n_\gamma + \Gamma_2 n_h n_e + K_\gamma(T).$$

Dieses System von Gleichungen trägt dem Umstand Rechnung, dass die Elektronen, Löcher und Photonen erzeugt und vernichtet werden können. Die Symbole Γ_1 und Γ_2 bezeichnen die Reaktionsraten für die Prozesse $\gamma \rightarrow e + h$ und $e + h \rightarrow \gamma$ (beide sind positiv). Diese Ausdrücke entsprechen einer Ratengleichung, wobei die Wahrscheinlichkeit eines Prozesses von der Dichte der involvierten Teilchen abhängt. Die Dichten ändern sich jedoch auch durch die Bewegung

²Denselben Prozess kennen wir auch aus der Hochenergiephysik. In der Dirac'schen Elektronentheorie kann ein Photon ein Elektron-Positron-Paar erzeugen. Allerdings liegt die benötigte Energie bei $0.5 \text{MeV} = mc^2$.

der Teilchen, beschrieben durch den letzten Term in den ersten beiden Gleichungen. Damit haben wir einerseits die “Reaktion” und andererseits die “Diffusion” als Ursache der Änderung der Dichten (Reaktionsdiffusionsgleichung). Wir betrachten die Photonen mit der Dichte n_γ als Beitrag der thermischen Strahlung (durch K_γ ausgedrückt) und lassen den Photonenstrom weg, da die Photonen in ihrer Flugrichtung gleichverteilt sind. Beachte, dass das Gleichungssystem (1.12) die Kontinuitätsgleichung (1.11) beinhaltet (multipliziere die ersten beiden Gleichungen mit der Ladung e und subtrahiere diese anschliessend voneinander).

1.3 Drift- und Diffusionsströme

Ladungsströme im Halbleiter haben zwei Ursachen. Einerseits können die Ladungen durch ein elektrisches Feld (Kraft) bewegt werden. Andererseits treten sie durch Diffusion bei inhomogener Teilchendichte auf. Der erstere wird “Drift-Strom” genannt und ergibt sich aus der Beziehung (1.4),

$$\left. \begin{aligned} \vec{j}_{drift,e} &= \frac{n_e e^2 \tau_e}{m_e^*} \vec{E} = \sigma_e \vec{E} \\ \vec{j}_{drift,h} &= \frac{n_h e^2 \tau_h}{m_h^*} \vec{E} = \sigma_h \vec{E} \end{aligned} \right\} \Rightarrow \vec{j}_{drift} = \vec{j}_{drift,e} + \vec{j}_{drift,h} = -(\sigma_e + \sigma_h) \vec{\nabla} \Phi, \quad (1.13)$$

wobei $\tau_{e,h}$ und $m_{e,h}^*$ die Streuzeiten und effektiven Massen der Elektronen und Löcher bezeichnen. Dies ergibt unterschiedliche Leitfähigkeiten $\sigma_{e,h}$ für die beiden Teilchentypen. Wir führen hier das elektrostatische Potential Φ ein mit $\vec{E} = -\vec{\nabla} \Phi$.

Der “Diffusionsstrom” ergibt sich aus dem ersten Fick’schen Gesetz und ist die Konsequenz einer Zufallsbewegung,

$$\vec{j}_{diff,e} = e D_e \vec{\nabla} n_e \quad \text{und} \quad \vec{j}_{diff,h} = -e D_h \vec{\nabla} n_h. \quad (1.14)$$

Wir benutzen die Einstein-Relation³

$$D_e = \frac{\tau_e k_B T}{m_e^*} \quad \text{und} \quad D_h = \frac{\tau_h k_B T}{m_h^*}, \quad (1.18)$$

und erhalten

$$\begin{aligned} \vec{j}_{diff,e} &= \frac{e \tau_e n_e}{m_e^*} k_B T \frac{\vec{\nabla} n_e}{n_e} = \frac{\sigma_e}{e} k_B T \frac{\vec{\nabla} n_e}{n_e} = \frac{\sigma_e}{e} \vec{\nabla} \mu_e, \\ \vec{j}_{diff,h} &= -\frac{e \tau_h n_h}{m_h^*} k_B T \frac{\vec{\nabla} n_h}{n_h} = -\frac{\sigma_h}{e} k_B T \frac{\vec{\nabla} n_h}{n_h} = -\frac{\sigma_h}{e} \vec{\nabla} \mu_h, \end{aligned} \quad (1.19)$$

³Herleitung der Einstein-Beziehung aus der Gleichgewichtsbedingung: Wir betrachten die Situation, dass ein elektrisches Feld E für die Ladungsträger zu einem Strom $j_{drift} = n e^2 E \tau / m$ führt. Falls wir gleichzeitig erlauben, dass die Teilchendichte n räumlich abhängig ist, ergibt sich die Kontinuitätsgleichung

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial x} \frac{n e E \tau}{m} - D \frac{\partial^2 n}{\partial x^2}. \quad (1.15)$$

Im Gleichgewichtszustand wird sich die Teilchendichte so einstellen, dass $\partial n / \partial t = 0$ und somit

$$\frac{e \tau n}{m} \vec{E} - D \frac{\partial n}{\partial x} = 0, \quad (1.16)$$

d.h., der Diffusions- und der Driftstrom heben sich auf. Das Potential Φ , welches das elektrische Feld ergibt, $\vec{E} = -\vec{\nabla} \Phi$, bestimmt via der Maxwell-Boltzmann-Verteilung die Teilchendichte,

$$n(\vec{r}) = n_0 e^{-e \Phi(\vec{r}) / k_B T}, \quad (1.17)$$

mit einer Konstanten n_0 . Durch Einsetzen in (1.16) erhalten wir via Koeffizientenvergleich die Einstein-Relation (1.18).

wobei wir hier das ‘‘chemische Potential’’ $\mu_{e,h}$ der Elektronen und Locher definieren, um einen Ausdruck analog zum Driftstrom zu erhalten,

$$\vec{\nabla}\mu_e = k_B T \frac{\vec{\nabla}n_e}{n_e} = k_B T \vec{\nabla} \ln n_e \quad \Rightarrow \quad \mu_e = \mu_{e0} + k_B T \ln \left(\frac{n_e}{n_0} \right), \quad (1.20)$$

mit den Integrationskonstanten $\mu_{e,0}$ und n_0 , die wir spater bestimmen werden. Eine analoge Definition kann auch fur die Locher angewendet werden.

Mit der Einfuhrung des chemischen Potentials konnen wir nun die Strome zusammenfassen,

$$\begin{aligned} \vec{j}_e &= \frac{\sigma_e}{e} \vec{\nabla} \{ \mu_e - e\Phi \} = \frac{\sigma_e}{e} \vec{\nabla} \eta_e, \\ \vec{j}_h &= -\frac{\sigma_h}{e} \vec{\nabla} \{ \mu_h + e\Phi \} = -\frac{\sigma_h}{e} \vec{\nabla} \eta_h. \end{aligned} \quad (1.21)$$

Dies definiert die sogenannten elektrochemischen Potentiale η_e und η_h der Elektronen und Locher.

Im Gleichgewichtszustand finden die beiden Prozesse $\gamma \rightleftharpoons e+h$ gleich haufig statt, so dass sich die Dichten nicht andern. Die Dichte der Photonen wird dadurch festgelegt, dass die Photonen aus der thermischen Strahlung des Halbleiters stammen (welche im Wesentlichen ein unendliches Reservoir darstellt). Somit hangt n_γ nur von der Temperatur ab. Im Falle einer homogenen Verteilung ohne Teilchenstrome ergibt sich aus den Gleichungen (1.12),

$$n_e n_h = \frac{\Gamma_1 n_\gamma}{\Gamma_2} = n_i^2(T). \quad (1.22)$$

Dies entspricht dem Massewirkungsgesetz der Chemie.⁴ Die Konstante $n_i^2(T)$ hangt vom jeweiligen Halbleiter und der Temperatur ab. Im Falle von Silizium finden wir bei $T = 300\text{K}$

$$n_h n_e = n_i^2 \approx 10^{20} \text{cm}^{-6} \propto \left(\frac{T}{T_0} \right)^3 e^{-E_g/k_B T}, \quad (1.25)$$

so dass $n_e = n_h \approx 10^{10} \text{cm}^{-3}$, was deutlich unterhalb der Elektronkonzentration in gewohnlichen Metallen liegt.

Wir konnen diese Situation aber auch im Sinne des thermodynamischen Gleichgewichtes beschreiben, indem wir die freie Energie betrachten: $dF = dU - d(ST)$ mit U als innere Energie ($dU = TdS - pdV + \mu dN + \Phi dQ$) und S als Entropie. Es gilt

$$\begin{aligned} dF_e(T, V, N_e, Q) &= dU_e - d(S_e T) = -S_e dT - p_e dV + \mu_e dN_e + \Phi dQ, \\ dF_h(T, V, N_h, Q) &= dU_h - d(S_h T) = -S_h dT - p_h dV + \mu_h dN_h + \Phi dQ, \end{aligned} \quad (1.26)$$

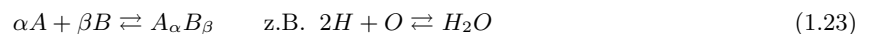
wobei dQ die anderung der Ladung bezeichnet, d.h., $dQ = -e dN_e$ fur Elektronen und $dQ = +e dN_h$ fur Locher. Fassen wir die beiden freien Energien zusammen, so erhalten wir

$$dF = dF_e + dF_h = \dots + (\mu_e - e\Phi) dN_e + (\mu_h + e\Phi) dN_h = \dots + \eta_e dN_e + \eta_h dN_h. \quad (1.27)$$

Photonen, die aus der thermischen Strahlung stammen, besitzen ein chemisches Potential $\mu_\gamma = 0$. Das thermische Gleichgewicht fur die Reaktion $\gamma \rightleftharpoons e + h$ ($dN_e = dN_h = -dN_\gamma$) ist erreicht, wenn

$$dF_{tot} = dF_e + dF_h + dF_\gamma = \dots + (\eta_e + \eta_h - \mu_\gamma) dN = 0, \quad (1.28)$$

⁴Das allgemeine Massewirkungsgesetz fur den chemischen Prozess



lautet

$$\frac{n_A^\alpha n_B^\beta}{n_{A_\alpha B_\beta}} = \text{konstant} \quad (1.24)$$

wobei die Konstante von den usseren Bedingungen, wie Temperatur und Druck, abhangt.

d.h.

$$0 = \mu_\gamma = \eta_e + \eta_h = \mu_e + \mu_h . \quad (1.29)$$

Dies ist auch verträglich mit der Bedingung $\vec{j} = 0$ wie wir unten sehen werden.

1.4 Dotierung der Halbleiter

Wegen der extrem geringen Ladungsträgerkonzentration sind die elektrischen Eigenschaften von reinen Halbleitern relativ langweilig. Interessant wird es erst, wenn man den Halbleitern spezielle Fremdatome als Verunreinigungen hinzufügt. Man spricht dabei von “dotieren”. Diese Fremdatome bringen zusätzliche Ladungsträger mit. Sie werden “Donatoren” genannt, wenn sie zusätzliche Elektronen zur Verfügung stellen und “Akzeptoren”, wenn sie Elektronen aufnehmen, d.h., zusätzliche Löcher an den Halbleiter abgeben. Für Si eignen sich Arsen-Atome (As) als Donatoren, da sie ein zusätzliches Valenzelektron besitzen, das keinen Platz in den kovalenten Verbindungen besitzt und daher nur schwach an den positiv geladenen As-Rumpf gebunden ist. In analoger Weise spielt Bor (B) die Rolle des Akzeptors mit einem Valenzelektron weniger als Si. Hier finden wir ein schwach gebundenes Loch pro B-Atom. Bei Zimmertemperatur sind alle Donatoren und Akzeptoren “ionisiert”, d.h., wir haben pro Fremdatom einen zusätzlichen frei beweglichen Ladungsträger, $n_e = n_D$ bzw. $n_h = n_A$, wobei n_D und n_A die Konzentration der Donatoren und Akzeptoren bezeichnet.

Wenn wir Donatoren einführen, sprechen wir von einem n-dotierten und für Akzeptoren von einem p-dotierten Halbleiter ($n = \text{negativ}$, $p = \text{positiv}$). Unabhängig von der Dotierung müssen die Gleichungen (1.12) immer noch gelten. Bei den üblichen kleinen Dotierungskonzentrationen von $n_{A,D} \sim 10^{17} \text{cm}^{-3}$ verändern sich die Raten Γ_1 und Γ_2 kaum, so dass die Gleichgewichtsbeziehung (1.25) gültig bleibt. Es ergibt sich daraus, dass sich das Verhältnis zwischen n_e und n_h drastisch ändert,

$$\begin{aligned} \text{n-dotiert: } n_e &\sim 10^{17} \text{cm}^{-3}, \quad n_h \sim 10^3 \text{cm}^{-3}, \\ \text{p-dotiert: } n_e &\sim 10^3 \text{cm}^{-3}, \quad n_h \sim 10^{17} \text{cm}^{-3}. \end{aligned} \quad (1.30)$$

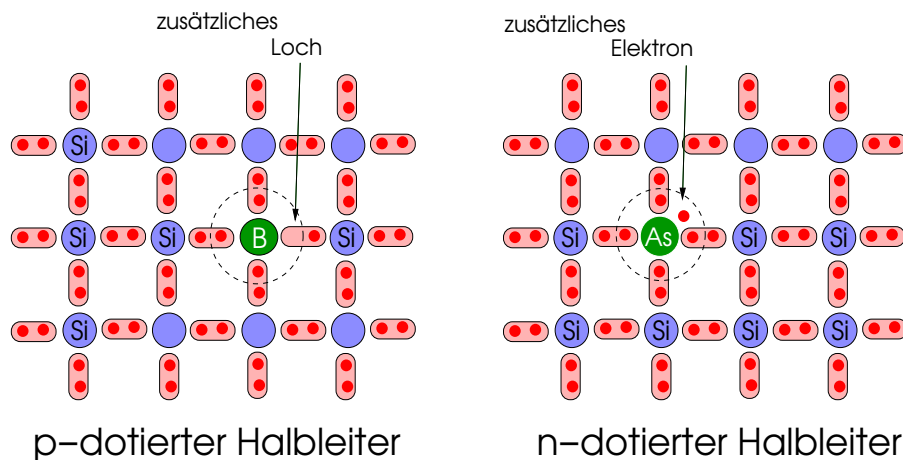


Abb. 1.4: Dotierte Halbleiter: Bor führt zu einem zusätzlichen, leichtgebundenen Loch (p-dotiert); Arsen führt zu einem zusätzlichen, leichtgebundenen Elektron (n-dotiert).

Damit wird der elektrische Stromtransport in n-dotierten Halbleitern durch die Elektronen dominiert, während Löcher kaum einen Beitrag liefern können. Das Umgekehrte gilt für p-dotierte Halbleiter. Beachte, dass auch die dotierten Halbleiter ladungsneutral sind. Legt man eine Spannung an den n-dotierten (p-dotierten) Halbleiter, dann bewegen sich die Elektronen (Löcher) vom positiven (negativen) zum negativen (positiven) Pol hin. Für kleine Felder wird dieser

Ladungstransport vom Drift-Strom dominiert und ist ohmsch, solange die Ladungskonzentration uniform bleibt. Wir werden bald sehen, dass Änderungen in der Ladungskonzentration zu nicht-ohmschem Verhalten führen.

Kapitel 2

Vom pn-Kontakt zur Solarzelle

Eines der elementarsten und wichtigsten Bauelemente der Halbleiter-Technologie ist der sogenannte pn-Kontakt. Dieses Bauelement findet auf drei verschiedene Arten Verwendung; als Stromgleichrichter (Diode), als Lichtquelle (LED - light emitting diode) und als Solarzelle.

2.1 pn-Kontakt im Gleichgewichtszustand

Der pn-Kontakt besteht aus aneinander grenzenden n-dotierten und p-dotierten Halbleitern. Wir können beide Regionen über einen ohmschen Kontakt mit einer Spannungsquelle verbinden, wobei die p-Seite mit dem ohmschen Kontakt Löcher austauscht (Löcher durch Elektronen füllt) und die n-Seite Elektronen. Falls keine Spannung anliegt, ergibt sich ein Gleichgewicht, bei dem es zu kleinen Grenzverletzungen der beiden Ladungsträgersorten kommt, aber kein Nettostrom fließt.

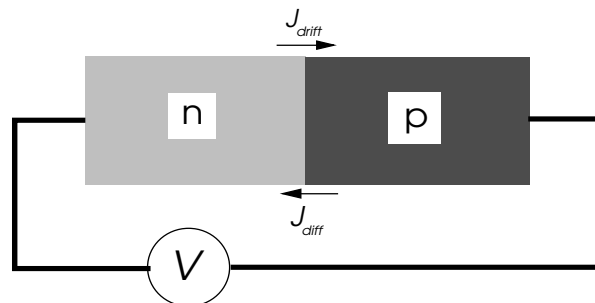


Abb. 2.1: pn-Kontakt: Im Gleichgewicht ($V = 0$) heben sich Diffusions- und Driftstrom gerade auf.

2.1.1 Gleichgewichtszustand

Im Gleichgewichtszustand, d.h., ohne angelegte Spannung und ohne Zufuhr von zusätzlichen Photonen (Beleuchtung), fließt kein elektrischer Strom durch den Kontakt, $\vec{j} = 0$. Darüber hinaus gilt die allgemeine Gleichgewichtsbedingung (1.29),

$$\eta_e + \eta_h = \mu_\gamma = 0 \quad \Rightarrow \quad \eta_e = -\eta_h. \quad (2.1)$$

Zusammen mit dem Ausdruck (1.21) für den Strom folgt sofort

$$\vec{j} = \vec{j}_e + \vec{j}_h = \frac{\sigma_e + \sigma_h}{e} \vec{\nabla} \eta_e = 0 \quad \Rightarrow \quad \vec{\nabla} \eta_e = 0. \quad (2.2)$$

Das bedeutet, dass η_e und folglich η_h keine räumliche Abhängigkeit haben.

2.1.2 Raumladungsschicht am pn-Kontakt

Genügend weit weg von der Grenzschicht ergibt sich für die Konzentration der Elektronen auf der n-Seite $n_e^n \approx n_D$ und gemäss (1.25) auf der p-Seite $n_e^p = n_i^2/n_A$ ($n_h^p \approx n_A$) gilt. Unter Verwendung der Definition des elektrochemischen Potentials (1.20) und der Gleichgewichtsbedingung (2.2) finden wir für die Elektronen in der p- und n-Region die Beziehung

$$\eta_e^p = \mu_{e0}^p + k_B T \ln \left(\frac{n_e^p}{n_0} \right) - e\Phi^p = \mu_{e0}^n + k_B T \ln \left(\frac{n_e^n}{n_0} \right) - e\Phi^n = \eta_e^n, \quad (2.3)$$

mit dem Referenzwert des chemischen Potentials $\mu_{e0}^{p,n} = \mu_e(n_e^{p,n} = n_0)$ (welcher auf beiden Seiten identisch ist, da er nur von der Dichte und nicht von der Art der Ladungsträger abhängt). Dies führt auf folgende elektrostatische Potentialdifferenz zwischen der n- und p-Seite,

$$\Delta\Phi = \Phi^n - \Phi^p = \frac{k_B T}{e} \ln \left(\frac{n_e^n}{n_e^p} \right) = \frac{k_B T}{e} \ln \left(\frac{n_D n_A}{n_i^2} \right). \quad (2.4)$$

Der daraus resultierende Driftstrom wird gerade durch den Diffusionsstrom aufgehoben, der durch die Differenz im chemischen Potential auf beiden Seiten angetrieben wird. Die Gleichung (2.4) ergibt folgende Beziehungen,

$$n_i^2(T) = n_A n_D e^{-e\Delta\Phi/k_B T} \quad \text{und} \quad n_e^p = n_D e^{-e\Delta\Phi/k_B T}. \quad (2.5)$$

Die Konzentration der Elektronen auf der p-Seite ist also durch die thermische Anregung bestimmt und folglich klein. In analoger Betrachtungsweise können die entsprechenden Resultate für die Löcher hergeleitet werden.

Die Differenz des elektrostatischen Potentials kommt, wie wir früher gesehen haben, dadurch zustande, dass sich an der Grenzschicht eine Ladungsdipolschicht ausbildet. Betrachten wir nun den Verlauf des elektrostatischen Potentials Φ durch die Grenzfläche. Dieser wird grundsätzlich durch die Poisson-Gleichung,

$$\vec{\nabla} \cdot \vec{D} = 4\pi\rho = \epsilon\vec{\nabla} \cdot \vec{E} = -\epsilon\vec{\nabla}^2\Phi, \quad (2.6)$$

bestimmt. Die Lösung der Poisson-Gleichung verlangt nun, dass wir die Ladungsdichte als Funktion des Potentials ausdrücken, $\rho[\Phi(x)]$. Dies kann man im Prinzip tun, indem man eine lokale Beziehung zwischen Potential und Ladungsdichte annimmt, analog zur Thomas-Fermi-Näherung in der Quantenmechanik. Da die resultierenden Gleichungen nicht leicht lösbar sind, folgen wir hier einer relativ groben, aber sehr einsichtsreichen Approximation von Schottky: wir nähern die Ladungsverteilung als stückweise konstante Funktion an (siehe Abb. 2.2). Hierfür wählen wir die x -Achse normal zur Grenzfläche (diese befinde sich bei $x = 0$). Unter der Annahme eines in der yz -Ebene homogenen Systems schreiben wir die Ladungsdichte als

$$\rho(x) = \begin{cases} en_D & -l_n < x < 0, \\ -en_A & 0 \leq x < l_p, \\ 0 & \text{sonst.} \end{cases} \quad (2.7)$$

Das Potential erhalten wir nun via Integration der Poisson-Gleichung (2.6) unter Annahme der Randbedingungen

$$\begin{aligned} x = -l_n : \quad & \Phi(-l_n) = \Phi^n, & \left. \frac{d\Phi}{dx} \right|_{x=-l_n} &= 0, \\ x = l_p : \quad & \Phi(l_p) = \Phi^p, & \left. \frac{d\Phi}{dx} \right|_{x=l_p} &= 0, \\ x = 0 : \quad & \Phi(0_-) = \Phi(0_+), & \left. \frac{d\Phi}{dx} \right|_{x=0_-} &= \left. \frac{d\Phi}{dx} \right|_{x=0_+}, \end{aligned} \quad (2.8)$$

wobei die letzte Bedingung die kontinuierliche Form des Potentials an der Grenzschicht sicherstellt (da die Poisson-Gleichung eine Differentialgleichung zweiter Ordnung ist und keine Oberflächenladungen vorhanden sind, müssen auch die ersten Ableitungen überall stetig sein). Beachte, dass für $x < -l_n$ und $x > l_p$ das Potential konstant ist. Damit erhalten wir das Potential

$$\Phi(x) = \begin{cases} -\frac{2\pi en_D}{\epsilon}(x + l_n)^2 + \Phi^n & x \leq 0, \\ \frac{2\pi en_A}{\epsilon}(x - l_p)^2 + \Phi^p & x \geq 0, \end{cases} \quad (2.9)$$

welches offensichtlich die ersten zwei Randbedingungen in (2.8) erfüllt. Die Bedingung bei $x = 0$ ergibt¹,

$$\Phi^n - \Phi^p = \frac{2\pi e}{\epsilon}(n_D l_n^2 + n_A l_p^2), \quad (2.12)$$

$$n_D l_n - n_A l_p = 0. \quad (2.13)$$

Diese zwei Gleichungen ergeben

$$l_n = \sqrt{\frac{\epsilon(\Phi^n - \Phi^p)n_A}{2\pi en_D(n_D + n_A)}}, \quad l_p = \sqrt{\frac{\epsilon(\Phi^n - \Phi^p)n_D}{2\pi en_A(n_D + n_A)}}. \quad (2.14)$$

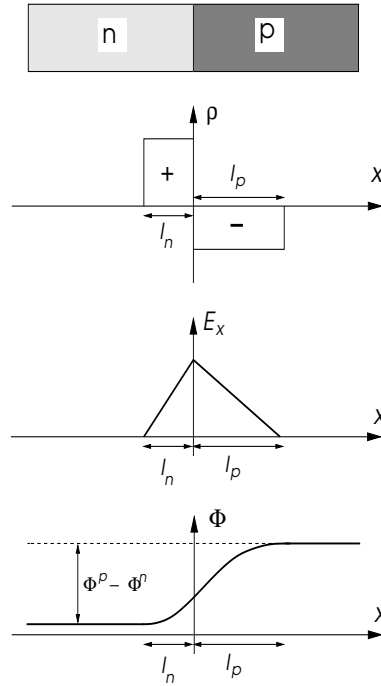


Abb. 2.2: Verhalten von Raumladungsdichte, elektrischem Feld und Potential am pn-Kontakt, beschrieben in Schottky-Näherung.

¹Beachte, dass Gleichung (2.12) dem Resultat des Ladungsdipols (1) entspricht, wobei die Dipoldichte gegeben ist durch

$$D = \int dx \rho(x)x = \frac{e}{2\epsilon}(n_D l_n^2 + n_A l_p^2). \quad (2.10)$$

Gleichung (2.13) enthält die Aussage, dass die Gesamtraumladung verschwindet,

$$\int dx \rho(x) = e(l_n n_D - l_p n_A) = 0. \quad (2.11)$$

Als Beispiel betrachten wir einen asymmetrischen pn-Kontakt (d.h. $n_D \neq n_A$) mit $n_A = 10^{17} \text{cm}^{-3}$ und $n_D = 10^{19} \text{cm}^{-3}$. Für Si mit der Permittivität $\epsilon = 12$ und der Potentialdifferenz $\Phi^n - \Phi_p \approx 0.9 \text{V}$ finden wir $l_n \sim 0.01 \mu\text{m}$ und $l_p \sim 1 \mu\text{m}$. Die Raumladung ist also sehr nahe an der Grenzschicht konzentriert; sie bildet eine Ladungsdipolschicht. Interessanterweise ist dieses Resultat in Schottky-Näherung gar nicht so weit von der realen Ladungsverteilung entfernt.

2.1.3 Sperrschicht

Die Positionsabhängigkeit der Elektronenkonzentration $n_e(x)$ kann näherungsweise mit Hilfe des (ortsunabhängigen) elektrochemischen Potentials $\eta_e(x) = \eta_e$ erhalten werden,

$$\begin{aligned} \eta_e &= \mu_{e0} + k_B T \ln \left(\frac{n_e(x)}{n_0} \right) - e\Phi(x) = \text{konstant} & (2.15) \\ \Rightarrow n_e(x) &= n_0 e^{(\eta_e - \mu_{e0} + e\Phi(x))/k_B T} = n_e^n e^{-e(\Phi^n - \Phi(x))/k_B T} \approx n_D e^{-e(\Phi^n - \Phi(x))/k_B T}, \end{aligned}$$

und analog für die Löcher,

$$n_h(x) = n_h^p e^{e(\Phi^p - \Phi(x))/k_B T} \approx n_A e^{e(\Phi^p - \Phi(x))/k_B T}. \quad (2.16)$$

Unter Verwendung von Gleichung (2.5) sehen wir, dass

$$n_e(x < -l_n) = n_D, \quad n_e(x > l_p) = \frac{n_i^2}{n_A}, \quad (2.17)$$

sowie

$$n_h(x < -l_n) = \frac{n_i^2}{n_D}, \quad n_h(x > l_p) = n_A. \quad (2.18)$$

In der Region $-l_n < x < l_p$ nimmt die Ladungskonzentration stark ab. Zum Beispiel gilt am Ort $x' = (l_p - l_n)/2$

$$n_e(x') + n_h(x') = (n_D + n_A) e^{-e\Delta\Phi/2k_B T} = n_i(T) \frac{n_D + n_A}{\sqrt{n_D n_A}} \ll n_D, n_A, \quad (2.19)$$

da $n_i^2 \ll n_D n_A$ bei $T = 300 \text{K}$. Man nennt daher diesen Bereich im Englischen “depletion layer” und im Deutschen “Sperrschicht”. Der letztere Begriff hebt einen zusätzlichen Aspekt hervor, nämlich dass der Ladungstransport in der Sperrschicht durch die niedrige Ladungsträgerkonzentration stark unterdrückt ist. Folglich wird bei einer angelegten äusseren Spannung der Spannungsabfall sich auf diesen Bereich beschränken.

Für den pn-Kontakt im Gleichgewicht gibt es daher zwei Bereiche: die Sperrschicht ($-l_n < x < l_p$) und den homogenen Bereich ($x < -l_n$ und $x > l_p$). Im homogenen Bereich wird die Ladungsträgerkonzentrationen fixiert. Beim Anlegen einer Spannung U über das ganze Element hinweg ändert sich die Konzentration der Minoritätsladungsträger am Rande der Sperrschicht. Die Spannung ändert das Potential $\Phi(x)$, so dass $\Phi(x \ll -l_n) - \Phi(x \gg l_p) = \Delta\Phi - U$, wobei $\Delta\Phi$ die Potentialdifferenz im Gleichgewichtszustand (d.h., bei $U = 0$) bezeichnet. Unter Verwendung der Näherungen in (2.15) und (2.16) finden wir

$$\begin{aligned} n_e(l_p) &\approx n_D e^{-e(\Delta\Phi - U)/k_B T} = \frac{n_i^2}{n_A} e^{eU/k_B T}, \\ n_h(-l_n) &\approx n_A e^{-e(\Delta\Phi - U)/k_B T} = \frac{n_i^2}{n_D} e^{eU/k_B T}. \end{aligned} \quad (2.20)$$

Die Gleichgewichtswerte der Minoritätsladungsträger im homogenen Bereiches werden durch Diffusion und Rekombination erreicht. Daher werden wir im nächsten Abschnitt den Aspekt der Diffusion betrachten.

2.1.4 Diffusionslänge für Minoritätsladungsträger

Die Diffusionslänge ist eine weitere wichtige Längenskala im pn-Kontakt, vor allem im Hinblick auf die Solarzellen. Am pn-Kontakt können aufgrund der Differenz im chemischen Potential Ladungsträger von einer Seite auf die andere diffundieren. Dabei handelt es sich um einen Diffusionsstrom, der die Ladungsträger von der Region hoher Konzentration (als Majoritätsladungsträger) in die Region niedriger Konzentration (als Minoritätsladungsträger) führt, z.B. Elektronen von der n- in die p-Region. Bei diesem Prozess kommt es neben der Diffusion auch zur Rekombination, so dass der Prozess durch Diffusion-Reaktion-Gleichungen vom Typ (1.12) beschrieben wird. Wir betrachten hier nur die Ladungsträger in ihrer Minoritätsregion. Als Beispiel nehmen wir die Elektronen in der p-Region. Die relevante Gleichung für den Gleichgewichtszustand ($\partial n_e / \partial t = 0$) lautet

$$\begin{aligned} 0 &= \Gamma_1 n_\gamma - \Gamma_2 n_h n_e + D_e \vec{\nabla}^2 n_e \\ &= \Gamma_2 (n_i^2 - n_A n_e) + D_e \vec{\nabla}^2 n_e \\ &= \Gamma_2 n_A \tilde{n}_e + D_e \vec{\nabla}^2 \tilde{n}_e, \end{aligned} \quad (2.21)$$

wobei wir benutzt haben, dass $\Gamma_1 n_\gamma$ durch die Gleichgewichtsbedingung ausgedrückt werden kann ($\Gamma_1 n_\gamma = \Gamma_2 n_i^2(T)$) und $n_h \approx n_A$. Ferner definieren wir hier $\tilde{n}_e = n_e - n_i^2/n_A$ als die Abweichung der Elektronenkonzentration vom Gleichgewichtswert in der p-Region.

Für die Grenzschichtgeometrie von (2.7) können wir uns wieder auf eine Dimension beschränken und erhalten

$$\tilde{n}_e(x) = \tilde{n}_e(l_p) e^{-(x-l_p)/L_e}, \quad x > 0, \quad (2.22)$$

wobei wir $L_e = \sqrt{D_e / \Gamma_2 n_A}$ als Diffusionslänge definiert haben. Beachte, dass $\tilde{\tau}_e = (\Gamma_2 n_A)^{-1}$ der mittleren Lebensdauer der Elektronen entspricht, d.h., deren mittleren Rekombinationszeit. Die Lebensdauer hängt also von der Konzentration der Akzeptoren in der p-Region ab. Für reines Si gilt $D_e \approx 35 \text{ cm}^2 \text{ s}^{-1}$. Bei einer realistischen Lebensdauer von $\tilde{\tau}_e \sim 10^{-6} \text{ s}$ finden wir $L_e \approx 60 \mu\text{m}$, was sehr viel grösser ist als $l_{p,n}$.

Wegen $L_e \gg l_p$ taucht natürlich die Frage auf, wie es sich mit der Raumladung verhält. Müsste sich diese nicht bis L_e ausdehnen? Dies ist nicht der Fall, da die zusätzliche Ladung der diffundierenden Elektronen sehr schnell kompensiert (abgeschirmt) wird. Die entsprechende Zeitskala kann abgeschätzt werden, indem wir von der Kontinuitätsgleichung, der Maxwell-Gleichung und der Gleichung für den Driftstrom ausgehen,

$$\left. \begin{aligned} \frac{\partial \rho}{\partial t} &= -\vec{\nabla} \cdot \vec{j} \\ \epsilon \vec{\nabla} \cdot \vec{E} &= 4\pi \rho \\ \vec{\nabla} \cdot \vec{j} &= \sigma_h \vec{\nabla} \cdot \vec{E} + \vec{E} \cdot \vec{\nabla} \sigma_h \end{aligned} \right\} \Rightarrow \frac{\partial \rho}{\partial t} = -\frac{4\pi \sigma_h}{\epsilon} \rho, \quad (2.23)$$

wobei wir verwendet haben, dass die Leitfähigkeit der Löcher σ_h nicht von deren Position abhängt. Folglich passt sich die Ladungsdichte auf einer Zeitskala $\tau = \epsilon / 4\pi \sigma_h$ einer Ladungsdichteänderung an. Für $n_A \sim 10^{17} \text{ cm}^{-3}$ findet man in Si $\tau \sim 10^{-14} \text{ s}$, was sehr viel schneller ist als die Rekombinationszeit.

2.2 Strom-Spannungs-Charakteristik des pn-Kontaktes

Für die Bestimmung der Strom-Spannungs-Charakteristik verwenden wir die Erkenntnis, dass der Spannungsabfall (fast ausschliesslich) in der Sperrschicht stattfindet, d.h., im Bereich des Raumladungsdipols. Eine angelegte Spannung U wird daher das elektrochemische Potential ausserhalb der Grenzschicht im Wesentlichen konstant lassen.

Wie wir oben gesehen haben, heben sich im Gleichgewicht die Beiträge des Drift- und des Diffusionsstromes gerade auf. Wir können uns fragen, welcher der beiden Ströme durch eine angelegte Spannung verändert wird und so einen endlichen Strom erzeugt. Der Driftstrom, auch Erzeugungsstrom genannt, beschreibt den durch das Dipolfeld angetriebene Transport der Minoritätsladungsträger. Jedes Elektron, das auf der p-Seite erzeugt wird und es bis zur Sperrschicht schafft, wird auf die n-Seite "hinübergesogen". Analoges gilt für die Löcher (n \rightarrow p). Dieser Prozess ist nur schwach von der angelegten Spannung abhängig, da sich die Zahl der erzeugten Elektronen (Löcher) auf der p-Seite (n-Seite) als Funktion von U nur schwach ändert.

Beim Diffusionsstrom hingegen müssen die Ladungsträger von ihrer Majoritätsseite durch thermische Fluktuationen sozusagen "bergauf" auf die Minoritätsseite getrieben werden. Die Spannung beeinflusst die Höhe der zu überwindenden Barriere. Der Anteil der Ladungsträger, die dies schaffen, wird durch den Boltzmann-Faktor, $\exp[-e(\Delta\Phi - U)/k_B T]$, beschrieben. Diese starke (exponentielle) Abhängigkeit widerspiegelt sich auch in der starken U -Abhängigkeit der Minoritätsladungskonzentration am Rande der Sperrschicht in (2.20) und dominiert die Strom-Spannungs-Charakteristik.

Betrachten wir nun den Diffusionsstrom der Elektronen bei $x = l_p$,

$$j_e(x = l_p) = eD_e \left. \frac{dn_e}{dx} \right|_{x=l_p} . \quad (2.24)$$

Aus der Diffusionsgleichung wissen wir, dass

$$n_e(x) = n_e^p + \tilde{n}_e(l_p)e^{-(x-l_p)/L_e} , \quad (2.25)$$

wobei $L_e \gg l_p$. Wir können also davon ausgehen, dass die Rekombination innerhalb der Sperrschicht vernachlässigbar ist. Setzen wir diesen Ausdruck für $n_e(x)$ in Gleichung (2.24) ein, so erhalten wir für den Beitrag der Elektronen

$$j_e = j_e(x = l_p) = -\frac{eD_e}{L_e} \tilde{n}_e(l_p) = -\frac{eD_e}{L_e} \left[n_e(l_p) - \frac{n_i^2}{n_A} \right] . \quad (2.26)$$

Analog finden wir für den Beitrag der Löcher (bei $x = -l_n$)

$$j_h = -\frac{eD_h}{L_h} \left[n_h(-l_n) - \frac{n_i^2}{n_D} \right] . \quad (2.27)$$

Mit der Spannungsabhängigkeit von $n_e(l_p)$ und $n_h(-l_n)$ gemäss (2.20) folgt für den Gesamtstrom,

$$J(U) = \left[\frac{eD_h}{L_h n_D} + \frac{eD_e}{L_e n_A} \right] n_i^2 (e^{eU/k_B T} - 1) . \quad (2.28)$$

Diese Beziehung entspricht der Strom-Spannungs-Charakteristik des pn-Kontaktes. Offensichtlich ist sie nicht-ohmsch und auch nicht symmetrisch unter Umkehr der Spannung.

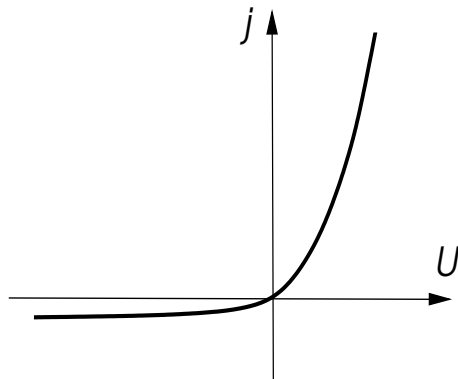


Abb.2.3: Strom-Spannungs-Charakteristik des pn-Kontaktes.

2.3 Funktionen des pn-Kontaktes

Der pn-Kontakt eignet sich für drei Funktionen. Er wird als Gleichrichter-Diode eingesetzt, da er im Wesentlichen nur für eine Stromrichtung durchlässig ist. Dank die Verknüpfung zwischen Elektron-Loch-Erzeugung und Photonen lässt sich der pn-Kontak als Lichtquelle (LED=light emitting diode) und als Solarzelle verwenden.

2.3.1 Gleichrichter

Die Funktion des pn-Kontaktes als Gleichrichter wird unmittelbar aus der Strom-Spannungs-Charakteristik klar. Verbinden wir den pn-Kontakt mit einer Spannungsquelle so, dass an der n-Seite der Minus-Pol und auf p-Seite der Plus-Pol liegt, dann ist U positiv und es fließt ein “grosser” Strom (Durchlassrichtung). Wird die Spannungsquelle umgepolt (U negativ), dann fließt ein “kleiner” Strom in die Gegenrichtung (Sperrichtung). Der Sättigungsstrom in Sperrichtung entspricht gerade

$$J_0 = \left(\frac{eD_h}{L_h n_D} + \frac{eD_e}{L_e n_A} \right) n_i^2, \quad (2.29)$$

und sollte für einen guten Gleichrichter möglichst klein sein².

Der Unterschied in der Stromstärke (Durchlassrichtung vs. Sperrichtung) kann am Verhalten der Sperrschicht verstanden werden. Diese spielt die Rolle einer Barriere für den Ladungstransport. Die Ausdehnung der Sperrschicht $l = l_p + l_n$ wird durch die Gleichungen (2.14) bestimmt,

$$l \propto \sqrt{\Phi^n - \Phi^p - U}. \quad (2.31)$$

Für positive U wird die Sperrschicht dünner, so dass der Diffusionsstrom leichter fließen kann. Für negative U wird sie dagegen dicker. Gleichzeitig wird die Ladungsträgerkonzentration stärker unterdrückt, was den Ladungsaustausch erschwert.

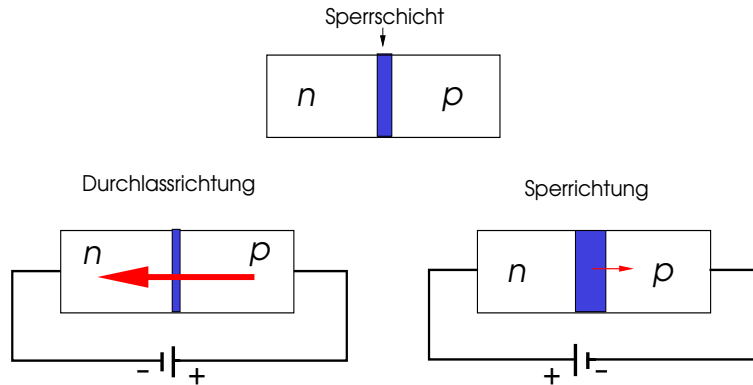


Abb. 2.4: Gleichrichter: Verhalten der Sperrschicht.

2.3.2 Lichtquelle - LED

Die Funktion des pn-Kontaktes als Lichtquelle beruht darauf, dass bei Strom in Durchlassrichtung ein Überschuss an Minoritätsladungsträgern besteht. Diese Ladungsträger rekombinieren unter Emission von Photonen einer bestimmten Wellenlänge. Betrachten wir die Photon-Produktionsrate in (1.12) und konzentrieren uns auf die p-Seite. Mit (2.20) und (2.25) finden

²Eine Abschätzung für die Grössenordnung von J_0 eines typischen pn-Kontaktes in Si ergibt mit den typischen Parametern $n_A = 5 \times 10^{16} \text{ cm}^{-3}$, $n_D = 10^{16} \text{ cm}^{-3}$, $n_i^2 = 10^{20} \text{ cm}^{-6}$, $D_e = 20 \text{ cm}^2 \text{ s}^{-1}$, $D_h = 10 \text{ cm}^2 \text{ s}^{-2}$, $\tau_h = \tau_e = 5 \times 10^{-7} \text{ s}$ die Diffusionslängen $L_e = \sqrt{D_e \tau_e} = 3.2 \times 10^{-3} \text{ cm} = 32 \mu\text{m}$ und $L_h = 2.2 \times 10^{-3} \text{ cm} = 22 \mu\text{m}$. Dies führt zu einem Sättigungsstrom (Sperrichtung)

$$J_0 = 9.3 \times 10^{-12} \text{ A cm}^{-2}. \quad (2.30)$$

wir

$$\frac{\partial n_\gamma}{\partial t} = -\Gamma_1 n_\gamma + K_\gamma + \Gamma_2 n_h n_e \approx -\Gamma_1 n_\gamma + K_\gamma + \Gamma_2 n_A n_e^p + \Gamma_2 n_A \tilde{n}_e(l_p) e^{-(x-l_p)/L_e} . \quad (2.32)$$

Betrachten wir die “überschüssig” erzeugten Photonen, dann kommen wir zur Abschätzung

$$\frac{\partial n_\gamma}{\partial t} \approx \Gamma_2 n_i^2 (e^{eU/k_B T} - 1) e^{-(x-l_p)/L_e} , \quad (2.33)$$

für $U > 0$. Diese Photonen werden vor allem im Bereich ausserhalb der Sperrschicht und innerhalb der Diffusionslänge L_e erzeugt. Die Anzahl der Photonen, die aus dem Halbleiter entkommen kann, hängt von der Struktur und den Brechungsindizes ab. Normalerweise ist das Verhältnis zwischen erzeugten und emittierten Photonen relativ klein. Dennoch hat ein LED (light emitting diode) eine viel höhere Umwandlungsrate von elektrischer Energie in sichtbares Licht als eine Glühbirne, welche in einem sehr breiten Spektrum (Schwarzkörperstrahler \rightarrow Planck’sches Spektrum) elektromagnetische Strahlung erzeugt.

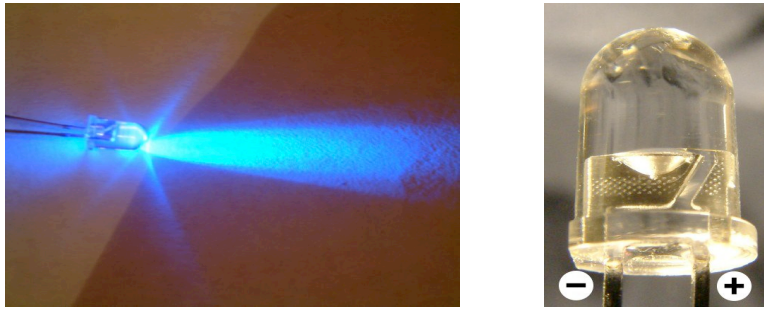


Abb. 2.5: LED: Light Emitting Diode.

Für LEDs eignen sich vor allem Halbleiter mit einer direkten Energielücke (siehe später). Die Frequenz des emittierten Lichtes wird durch die Grösse der Energielücke bestimmt: $E_g = \hbar\omega$. Vorgreifend wollen wir anmerken, dass Silizium im Gegensatz zu GaAs und dazu verwandten Halbleitern eine indirekte Lücke besitzt. Letztere werden oft für LEDs verwendet.

Halbleiter	GaAs	GaAs _{0.6} P _{0.4}	GaAs _{0.4} P _{0.6}	GaP	GaN
Wellenlänge (nm)	940	660	620	550	340
Farbe	infrarot	rot	gelb	grün	ultraviolett

2.4 Solar-Zelle

Der Prozess der LEDs lässt sich natürlich auch umkehren. In diesem Sinne wollen wir betrachten, was geschieht, wenn wir einen dotierten Halbleiter mit Licht bestrahlen. Als konkretes Beispiel wählen wir einen p-dotierten Halbleiter mit der Akzeptor-Konzentration n_A . In der bestrahlten Region werden Elektron-Loch-Paare erzeugt, zusätzlich zu denjenigen, die aufgrund des thermischen Gleichgewichtes bestehen. Die Löcher, die sowieso in grosser Konzentration vorhanden sind, haben näherungsweise eine homogene Verteilung, die Elektronen hingegen werden sich aufgrund der Diffusion verteilen. Wir wollen diesen Prozess nun unter der Annahme betrachten, dass das Licht nur an einer bestimmten Stelle auftrifft. Dies sei bei $x = 0$ in einer eindimensionalen Konfiguration (Annahme: Homogenität in die anderen Richtungen). Für die Photonen nehmen wir folgende räumliche Abhängigkeit an, $n_\gamma = n_{\gamma 0} + \tilde{n}_\gamma \delta(x)$, mit der Konzentration der thermischen Photonen $n_{\gamma 0}$. Daraus folgt für die Elektron- und Loch-Konzentrationen: $n_e(x) = n_{e0} + \tilde{n}_e(x)$ und $n_h(x) = n_{h0} + \tilde{n}_h(x)$, wobei $n_{\gamma 0}, n_{e0}$ und $n_{h0} \approx n_A$ den Gleichgewichtskonzentrationen entsprechen. Die Reaktion-Diffusion-Gleichung (1.12) für die Elektronen lautet dann

$$\begin{aligned}\frac{\partial n_e}{\partial t} &= \Gamma_1(n_{\gamma 0} + \tilde{n}_\gamma \delta(x)) - \Gamma_2(n_A + \tilde{n}_h)(n_{e0} + \tilde{n}_e) + D_e \frac{d^2}{dx^2} \tilde{n}_e \\ &\approx \Gamma_1 \tilde{n}_\gamma \delta(x) - \Gamma_2 n_A \tilde{n}_e + D_e \frac{d^2}{dx^2} \tilde{n}_e .\end{aligned}\quad (2.34)$$

Im stationären Fall $\partial n_e / \partial t = 0$ erhalten wir

$$\tilde{n}_e(x) \propto e^{-|x|/L_e} , \quad (2.35)$$

wobei $L_e = \sqrt{D_e / \Gamma_2 n_A}$ die Diffusionslänge der Elektronen im p-dotierten Halbleiter ist (siehe Abschnitt 2.1.4).

Aus dieser Betrachtung schliessen wir, dass die durch die einfallenden Photonen erzeugten Elektronen vom Entstehungsort wegdiffundieren und innerhalb der Diffusionslänge wieder rekombinieren, da sehr viele Löcher zur Verfügung stehen. Wenn das Licht nun auf einen pn-Kontakt auftrifft und Elektron-Loch-Paare auf der p-Seite in einem Bereich erzeugt, der innerhalb einer Diffusionslänge von der Sperrschicht liegt, so werden alle Elektronen, die bis zur Sperrschicht diffundieren, durch das elektrostatische Potential auf die n-Seite hinübergezogen. Damit können diese Elektronen aus dem p-Bereich entkommen und erzeugen einen Strom. Die erzeugten Löcher werden von der Sperrschicht abgestossen (da für sie das Potential ansteigt) und sie bleiben im p-Bereich.

Um das Verhalten der Elektronen zu modellieren, betrachten wir wiederum die Gleichung (2.34) mit der Sperrschicht bei $x = l > 0$. Wenn alle Überschusselektronen bei $x = l$ auf die n-Seite “verschwinden”, muss $\tilde{n}_e(x = l) = 0$ gelten. Mit dem Ansatz,

$$\tilde{n}_e(x) = \begin{cases} n_1 e^{x/L_e} & x < 0 , \\ n_2 e^{-x/L_e} + n_3 e^{x/L_e} & 0 < x < l , \end{cases} \quad (2.36)$$

und den Randbedingungen,

$$\begin{aligned}x = 0 : & \quad \begin{cases} n_1 = n_2 + n_3 , \\ n_1 + n_2 - n_3 = \Gamma_1 \tilde{n}_\gamma L_e / D_e , \end{cases} \\ x = l : & \quad n_2 e^{-l/L_e} + n_3 e^{l/L_e} = 0 , \end{aligned} \quad (2.37)$$

erhalten wir

$$n_1 = \frac{\Gamma_1 \tilde{n}_\gamma L_e}{2D_e} \left(1 - e^{-2l/L_e}\right) , \quad n_2 = \frac{\Gamma_1 \tilde{n}_\gamma L_e}{2D_e} , \quad n_3 = -\frac{\Gamma_1 \tilde{n}_\gamma L_e}{2D_e} e^{-2l/L_e} . \quad (2.38)$$

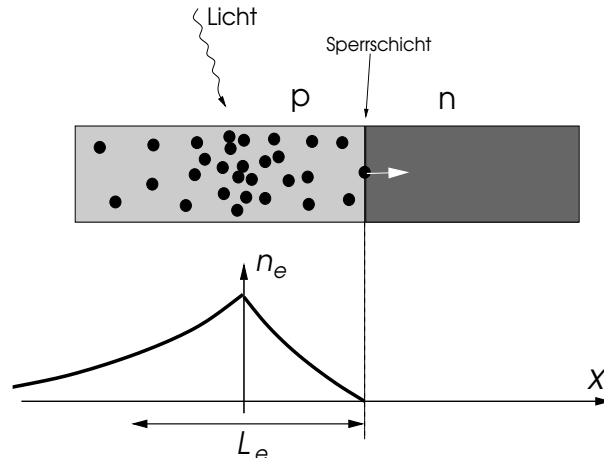


Abb. 2.6: Diffusion der erzeugten Überschusselektronen nahe der Sperrschicht. Elektronen, welche die Sperrschicht erreichen, verschwinden in den n-Bereich.

Mit dieser Lösung erhalten wir einen Nettostrom, der durch die Sperrschicht Richtung n-Bereich fließt,

$$j_e(l) = eD_e \left. \frac{d\tilde{n}_e}{dx} \right|_{x=l} = -e \frac{\Gamma_1 \tilde{n}_\gamma}{2} e^{-l/L_e} . \quad (2.39)$$

Die Stärke des Stromes hängt exponentiell von der Distanz zwischen Sperrschicht und Erzeugungsort der Elektron-Loch-Paare ab. Innerhalb dieses Modells erhalten wir bei homogener Bestrahlung den Gesamtstrom

$$J_e = \int_0^\infty dl j_e(l) = -e \frac{\Gamma_1 \tilde{n}_\gamma L_e}{2} . \quad (2.40)$$

Der Zusammenhang zwischen J_e und \tilde{n}_γ ist mit der Quanteneffizienz verknüpft. Diese gibt an, welcher Anteil der absorbierten Photonen zum elektrischen Strom beitragen. Damit ist klar, dass für diesen Strom diejenige Konzentration der Elektron-Loch-Paare verantwortlich ist, welche innerhalb einer Diffusionslänge von der Sperrschicht durch die Bestrahlung erzeugt wird. Eine analoge Betrachtung ist natürlich auch aus der Perspektive der Löcher für die n-Seite möglich. Der durch das Licht induzierte Strom J_L liefert nun einen zusätzlichen Beitrag zur Strom-Spannungs-Charakteristik,

$$J'(U) = J(U) - J_L = J_0 \left[e^{eU/k_B T} - 1 \right] - J_L . \quad (2.41)$$

Es ist wichtig zu erkennen, dass J_L in Sperrichtung des pn-Kontaktes fließt und daher ein negatives Vorzeichen besitzt. Ohne Beleuchtung ist $J_L = 0$; wir nennen daher $J(U)$ auch *Dunkelstrom*. Damit weist dieses Element bei Nullstrom (offener Stromkreis) eine Spannung U_L und bei Null-Spannung (kurzgeschlossener Stromkreis) einen Strom $-J_L$ auf,

$$eU_L = k_B T \ln \left(\frac{J_L}{J_0} + 1 \right) . \quad (2.42)$$

Die Leistung entspricht

$$P_E(U) = -U \cdot J'(U) , \quad (2.43)$$

und kann durch geeignete Wahl von U über der Sperrschicht optimiert werden (Abb. 2.7), so dass $P_{E,\max} = -J_m U_m < -J_L U_L$. Wir bezeichnen das Verhältnis

$$FF = \frac{J_m U_m}{J_L U_L} \quad (2.44)$$

als "Füllungsfaktor" (fill factor).

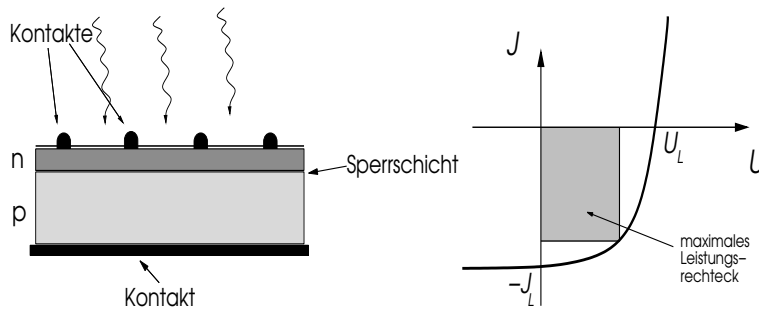


Abb.2.7: Schematische Struktur der Solarzelle und Strom-Spannungs-Charakteristik. Die grösste Leistung entspricht dem grössten eingeschriebenen Rechteck. Der Füllungsfaktor entspricht dem Bruchteil, den die schattierte Fläche von der ganzen durch die Strom-Spannungskurve abgesteckten Fläche überdeckt.

2.5 Wirkungsgrad der Solarzelle

Um den Wirkungsgrad ($\eta = P_E/P_L$) zu erhöhen wird die Struktur der Solarzelle so optimiert, dass möglichst viel Licht in die Nähe der Sperrschicht gelangt und dort in Elektron-Loch-Paar umgewandelt wird. Zudem wird die Oberfläche durch eine Antireflexionsschicht abgedeckt, die der Zelle einen dunkelbläulichen Schimmer verleiht. Es gibt aber verschiedene Verlust-Mechanismen, welche die Optimierung des Wirkungsgrades begrenzen:

- Nur derjenige Teil des Lichtspektrums kann verwendet werden, dessen Photonen genügend Energie besitzen um Elektron-Loch-Paare zu erzeugen. Es gibt also eine untere Schwellenfrequenz, $\hbar\omega \geq E_g$.
- Der “Hochenergie”-Anteil des Spektrums kann auch nicht genutzt werden, da die überschüssige Energie schnell “thermalisiert” wird, d.h., als Wärme verpufft.
- Die Anzahl der Kontakte muss optimiert werden; mehr Kontakte führen zu mehr Schattenwurf auf die Oberfläche, ergeben aber auch eine bessere Ausbeute der elektrischen Energie.
- Verunreinigungen im Halbleitermaterial kann zu schnellerer Rekombination führen und damit den Diffusionsbereich verringern.

Betrachtet man die einfachste Form einer Solarzelle (pn-Kontakt mit einem Typ von Halbleiter (single-gap device)) und das Strahlungsspektrum der Sonne, so führen die ersten beiden Faktoren auf den theoretisch maximalen Wirkungsgrad von ungefähr 30%.

Diese Abschätzung beruht auf einigen optimistischen Annahmen. Im Folgenden wollen wir letztere diskutieren und zeigen, dass für den Wirkungsgrad einer Solarzelle tatsächlich ein Maximum in Abhängigkeit von Materialparametern existiert. Zunächst betrachten wir den Photonenstrom, der von der Sonne ausgehend die Erde trifft. Der solare Energiestrom, auch bekannt als “Solarkonstante”, beträgt $J_\gamma = 1.3 \times 10^{-4} \text{ kW cm}^{-2} = w_\gamma \langle \hbar\omega \rangle$. Hier bezeichnet w_γ die Photonstromdichte und $\langle \hbar\omega \rangle$ die mittlere Photonenergie. Die Photonstromdichte ist abhängig von der Photondichte n_γ via $w_\gamma = cn_\gamma$ und wird daher durch die Planck’sche Verteilung bestimmt,

$$\frac{dn_\gamma}{d\hbar\omega} = \frac{2}{(2\pi\hbar c)^3} \frac{(\hbar\omega)^2}{e^{\hbar\omega/k_B T} - 1}. \quad (2.45)$$

Für die Energiedichte gilt

$$\frac{dE_\gamma}{d\hbar\omega} = \hbar\omega \frac{dn_\gamma}{d\hbar\omega}. \quad (2.46)$$

Wir definieren nun die Dichte der Photonon mit Frequenzen oberhalb der Schranke E_g/\hbar als

$$n_\gamma(E_g) = \int_{E_g}^{\infty} \frac{dn_\gamma}{d\hbar\omega} d\hbar\omega. \quad (2.47)$$

Mit der gesamten Energiedichte

$$E_\gamma = \int_0^{\infty} \frac{dE_\gamma}{d\hbar\omega} d\hbar\omega = \frac{\pi^2 (k_B T)^4}{60 (\pi \hbar c)^3}, \quad (2.48)$$

ergibt sich eine mittlere Energie

$$\frac{E_\gamma}{n_\gamma(0)} = \langle \hbar\omega \rangle = 2.7 k_B T = 1.39 \text{ eV}, \quad (2.49)$$

wobei $T = 6000 \text{ K}$ als Temperatur der Sonne angenommen wurde ($k_B T = 0.52 \text{ eV}$). Dies entspricht einer Photonstromdichte von $w_\gamma = J_\gamma / \langle \hbar\omega \rangle = 5.8 \times 10^{17} \text{ cm}^{-2} \text{ s}^{-1}$.

Zur Erzeugung eines Teilchen-Loch-Paares benötigen die Photonen eine Energie $\hbar\omega$ grösser als die materialspezifische Energie E_g (Bandenergielücke). Im Idealfall erzeugt jedes einfallende

Photon mit Frequenz $\omega \geq E_g/\hbar$ ein Elektron-Loch-Paar (Quanten-Effizienz = 1, d.h., 100% der Photonen tragen zum Strom bei) und es fließt ein Strom

$$J_L = -ew_\gamma(E_g) = ecn_\gamma(E_g) . \quad (2.50)$$

In Wirklichkeit liegt die Quanteneffizienz natürlich unterhalb von 1. Vergleichen wir J_L mit J_0 , dann sehen wir $|J_L| \gg J_0$ wobei

$$J_0(E_g) = Ce^{-E_g/k_B T} \quad (2.51)$$

und der Vorfaktor C als nur schwach von E_g abhängig angenommen wird ($E_g \gg k_B T$). Mit (2.42) finden wir

$$U_L \approx \frac{k_B T}{e} \ln \left(\frac{ew_\gamma}{Ce^{-E_g/k_B T}} \right) = \frac{E_g}{e} + \frac{k_B T}{e} \ln \left(\frac{ecn_\gamma}{C} \right) . \quad (2.52)$$

Um die (open circuit) Spannung U_L zu optimieren, ist es von Vorteil den Sättigungsstrom in der Sperrichtung möglichst klein zu machen (analog zur Optimierung der Gleichrichter-Diode). Die Arbeit, die pro absorbiertem Photon geleistet wird, kann nun abgeschätzt werden durch

$$W = \frac{-J_m U_m}{w_\gamma(E_g)} = \frac{-J_L U_L}{cn_\gamma} FF = eU_L FF , \quad (2.53)$$

mit FF als Füllungsfaktor (fill factor), d.h., $FFU_L J_L = J_m U_m$ mit $FF < 1$. Der Wirkungsgrad ergibt sich aus dem Verhältnis

$$\eta(E_g) = \frac{n_\gamma(E_g)W(E_g)}{n_\gamma(0)\langle \hbar\omega \rangle} , \quad (2.54)$$

wobei im Nenner die gesamte von der Sonne einfallende Leistung pro Fläche steht. Einsetzen von Gleichungen (2.52) und (2.53) in den Ausdruck (2.54) ergibt

$$\eta(E_g) = \frac{n_\gamma(E_g)}{n_\gamma(0)} FF \left\{ \frac{E_g}{\langle \hbar\omega \rangle} + \frac{k_B T}{\langle \hbar\omega \rangle} \ln \left[\frac{en_\gamma(E_g)}{C} \right] \right\} . \quad (2.55)$$

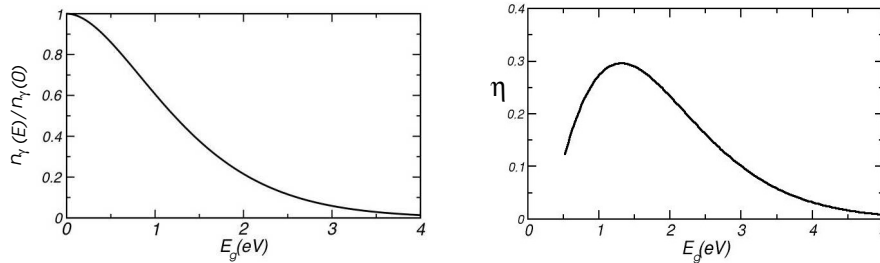


Abb. 2.8: Photondichte n_γ der Sonnenstrahlung als Funktion der Energieschranke E_g (links). Wirkungsgrad η einer fiktiven Solarzelle als Funktion von E_g (rechts). Werte aus der Literatur: $J_0(E_g = 1.14 \text{ eV}) = 6 \times 10^{-16} \text{ A cm}^{-2}$, $FF = 0.9$ und $T = 300 \text{ K}$.

Wie in Abb. 2.8 ersichtlich tritt ein Effizienz-Maximum für $E_g \approx 1 \sim 2 \text{ eV}$ auf. Für Silizium beträgt $E_g = 1.14 \text{ eV}$ und für GaAs $E_g = 1.43 \text{ eV}$; diese Materialien liegen vom Wirkungsgrad her in einem guten Regime.³ Beachte, dass wir bei dieser groben Abschätzung einige wichtige Aspekte ausser Acht gelassen haben.

Es ist interessant zu sehen, wie sich eine Erhöhung der Lichtintensität auf den Wirkungsgrad auswirkt. Nehmen wir an, dass sich die Intensität um den Faktor R ändert. Dann folgt

$$\eta_R(E_g) - \eta_{R=1}(E_g) = \frac{n_\gamma(E_g)}{n_\gamma(0)} FF \frac{k_B T}{\langle \hbar\omega \rangle} \ln R . \quad (2.56)$$

³Diese Betrachtung folgt dem Artikel von C.H. Henri, *Limiting efficiencies of ideal single and multiple gap terrestrial solar cells*, J. Appl. Phys. 51, 4494 (1980).

Die Zunahme ist also nur logarithmisch. Bei einer Erhöhung der Lichtintensität um den Faktor $R = 1000$ steigt der Wirkungsgrad bloss um $\sim 7\%$ (für eine Energielücke der Grösse $E_g = 1 \text{ eV}$). Daher kann man durch Lichtbündelung und andere intensitätserhöhende Massnahmen keine gigantische Steigerung des Wirkungsgrades erwarten.

2.6 Experimentelle Bestimmung von E_g eines pn-Kontaktes

Die Energie E_g ist eine wichtige Grösse, die die Eigenschaften eines Halbleiters bestimmt. Insbesondere geht sie in den Sättigungsstrom ein und stellt die Schwellenenergie für die Elektron-Loch-Paarzeugung dar. Wir betrachten hier eine einfache Beziehung, aus der in einem Experiment diese Grösse bestimmt werden kann.

Die Messung benötigt eine Stromquelle, die einen konstanten Strom I liefern kann. In einem in Durchlassrichtung angeschlossenen pn-Kontakt muss sich daher die Spannung entsprechend anpassen,

$$I = I_0 \left(e^{eU/k_B T} - 1 \right) \approx C(T) e^{-E_g/k_B T} e^{eU/k_B T}, \quad (2.57)$$

wobei $eU \gg k_B T$ angenommen wurde. Zudem gilt näherungsweise $C(T) \propto T^m$ ($1.5 \leq m \leq 4$). Mit $I = \text{konstant}$ liefert die Ableitung nach der Temperatur die Gleichung

$$0 = \frac{T}{C} \frac{dC}{dT} k_B T + E_g - eU + T \frac{deU}{dT}. \quad (2.58)$$

$(dC/dT)(T/C)$ ist von der Grössenordnung 1, so dass der erste Term für Temperaturen $T \leq 300 \text{ K}$ viel kleiner ist als E_g und daher vernachlässigt werden kann. Aus (2.57) ergibt sich

$$eU = k_B T \left(\ln \frac{I}{C} + E_g \right) \Rightarrow \frac{deU}{dT} = k_B \left(\ln \frac{I}{C} + E_g \right) - k_B T \frac{d \ln C}{dT} = K \approx \text{konstant}. \quad (2.59)$$

Daraus folgern wir, dass für die Situation eines konstanten Stromes zwischen U und T zu guter Näherung folgende lineare Beziehung besteht,

$$eU = KT + E_g, \quad (2.60)$$

so dass sich E_g als Achsenabschnitt bei $T = 0$ ergibt. Durch Extrapolation kann man diesen aus Messungen im Bereich unterhalb der Raumtemperatur erhalten.

Kapitel 3

Licht - Klassisches Verständnis

In der Diskussion der Solarzelle spielt natürlich das Licht als Energieträger eine wichtige Rolle. In der Antike wurde Licht eher im Zusammenhang mit dem Sehen analysiert. Erst im Laufe der Geschichte wurde es als eigenständige Strahlung und als Energieträger verstanden. In diesem Kapitel wollen wir die wichtigsten Konzepte rund ums Licht zusammenfassen.

3.1 Geometrische Optik

Gemäss der Platonische Schule beruhte der Prozess des Sehens darauf, dass das Auge “Seh-Strahlen” aussendet, die ein Objekt “erfassen”. Licht hatte daher keine Bedeutung ohne das Sehen. Es gab jedoch auch in der Antike andere Konzepte des Lichtes. Unabhängig davon war es aber möglich Strahlen- oder geometrische Optik zu betreiben.

3.1.1 Strahlenoptik und Herons Prinzip

Ein erstes systematisches Werk ist *Euklids Optik* im 3. Jh. v. Chr. Darin wird die Idee des Lichtstrahls eingeführt, welcher sich in einem homogenen Medium geradlinig ausbreitet. *Heron von Alexandrien* postulierte diese Eigenschaft als ein strengeres und “moderneres” Prinzip: Licht verbindet zwei Punkte auf dem kürzesten Weg. Innerhalb eines homogenen Mediums ergibt dies natürlich eine Gerade zwischen den Punkten. Darüber hinaus erhalten wir aber auch eine Aussage über das Reflexionsverhalten des Lichtes an einer spiegelnden Fläche.

Die Reflexion eines Lichtstrahl von einer glatten, ebenen Spiegelfläche hat folgende Charakteristik: der reflektierte Strahl liegt in der Einfallsebene, welche durch den einfallenden Strahl und die Oberflächennormalen bestimmt ist, und schliesst den selben Winkel mit der Normalen ein wie der einfallende Strahl. Es lässt sich leicht beweisen, dass dies die kürzeste Verbindung zwischen zwei Punkten darstellt, welche eine Reflexion an der Oberfläche beinhaltet. Der Beweis folgt der Sequenz von Darstellungen in Abb. 3.1 a)-d). Wir betrachten einen Strahl, der seinen Anfangspunkt bei A und seinen Endpunkt bei E hat und beim Punkt O die Oberfläche trifft. Die Ebene der Strahlen und Normalen liegt in der Blattebene, und O ist so gewählt, dass die beiden Winkel identisch gleich θ sind (Abb.3.1a)).

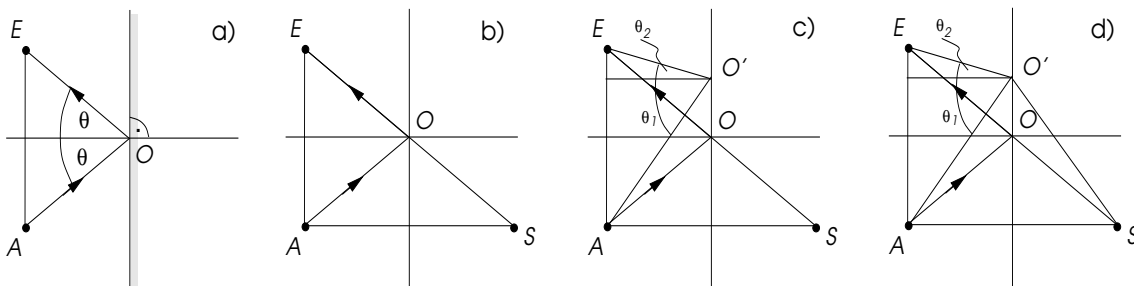


Abb. 3.1: Beweis von Herons Prinzip.

Wir verlängern die Gerade OE bis S (Spiegelpunkt von A) und stellen fest, dass die Längen der Geraden OE , AO und OS gleich sind (Abb.3.1b): $\overline{OE} = \overline{AO} = \overline{OS}$. Dies ist kompatibel mit der Charakteristik der Reflexion. Nun nehmen wir an, dass der Strahl zwischen A und E einer anderen Trajektorie folgt, die im Punkt O' auf die Oberfläche trifft (Abb.3.1c). Wenn wir nun auch die Gerade $O'S$ verbinden (beachte, dass $\overline{AO'} = \overline{O'S}$), sehen wir in Abb.3.1d) sofort, dass die Verbindung $\overline{AO'E}$ länger ist als \overline{AOE} . Diese Argumentation lässt sich auch auf jeden Punkt O' anwenden, der nicht in der Blattebene liegt. Damit wird klar, dass die Verbindung \overline{AOE} die kürzeste aller möglichen Verbindungen ist und Herons Prinzip tatsächlich die Reflexion eines Strahls richtig beschreibt.

Der wichtige Beitrag Herons ist allerdings nicht die Beschreibung der Reflexion, da das Gesetz der gleichen Winkel schon vor ihm bekannt war. Vielmehr liegt die grosse Bedeutung in der Formulierung eines Prinzips, das verschiedene Gesetze der Lichtausbreitung in einen Zusammenhang bringt. Herons Prinzip ist der Vorläufer des allgemeineren Prinzips von *Fermat*. Während sich Herons Beschreibung auf den Strahlengang in einem homogenen Medium inklusive Reflexion beschränkt (man nennt diesen Bereich der Optik auch *Katoptrics*), erlaubt Fermats Prinzip zusätzlich die Betrachtung der Lichtbrechung an Grenzschichten zwischen zwei unterschiedlichen Medien.

3.1.2 Lichtbrechung

Das Phänomen der Brechung war schon in der Antike bekannt und wurde speziell von Claudius Ptolemäus im 2. Jh. experimentell untersucht. Dabei realisierte er, dass es einen Zusammenhang zwischen Einfallswinkel und Ausfallswinkel des Lichtstrahls beim Übertritt in ein anderes Medium gibt. Er war aber nicht in der Lage das exakte Brechungsgesetz zu bestimmen. Dies gelang erst Snellius ungefähr 15 Jahrhunderte später. Ebenfalls mit Hilfe von Experimenten stellte er folgende Gesetzmässigkeit auf,

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1}, \quad (3.1)$$

wobei $\theta_{1(2)}$ der Winkel zwischen Grenzflächennormalen und Einfallswinkel (Ausfallswinkel) bezeichnet. Die Materialeigenschaften des jeweiligen Medien werden durch die Brechungsindizes n_1 und n_2 charakterisiert (Abb.3.2).

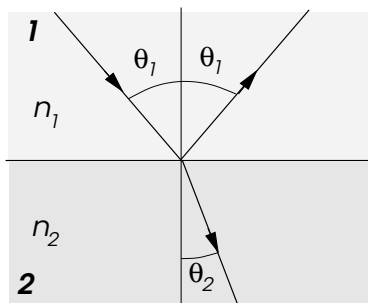


Abb. 3.2: Brechung des Lichts zwischen zwei Medien und Reflexion an der Grenzfläche.

An der Grenzfläche wird auch ein Teil des Lichtes reflektiert. Das Verhältnis zwischen reflektiertem und transmittiertem Anteil kann jedoch innerhalb der geometrischen Optik nicht bestimmt werden. Wir werden in einem späteren Abschnitt darauf zurückkommen.

3.1.3 Fermat'sches Prinzip

Fermats Prinzip erweitert das Konzept von Heron und erlaubt die Beschreibung von Brechung in einem allgemeineren Kontext. Es basiert auf der "optischen Weglänge", die er folgendermassen definiert,

$$L(A, E) = \int_A^E n ds, \quad (3.2)$$

wobei n der lokale Brechungsindex ist, und A und E bezeichnen Anfangs- und Endpunkt des Lichtstrahls. Das Fermat'sche Prinzip besagt, dass der Lichtstrahl derjenigen Trajektorie folgt, welche die optische Weglänge $L(A, E)$ minimiert. Offensichtlich ist das Prinzip von Heron darin enthalten; für ein homogenes Medium gilt $n = \text{konstant}$. Alternativ lässt sich das Prinzip auch formulieren als Minimierung der Zeit $t_{A,E}$, die der Lichtstrahl benötigt, um von A nach E zu gelangen. Dabei wird der Zusammenhang zwischen dem Brechungsindex und der Lichtgeschwindigkeit in einem Medium,

$$c' = \frac{c}{n}, \quad (3.3)$$

ausgenutzt. Es gilt dann

$$L(A, E) = ct_{A,E} = c \int_A^E \frac{ds}{c'}. \quad (3.4)$$

Hier bezeichnet c' die lokale Lichtgeschwindigkeit und $c = 299792458 \text{ ms}^{-1}$ die Vakuumslichtgeschwindigkeit ($n = 1$ im Vakuum). Dies setzt natürlich voraus, dass das Licht eine endliche Geschwindigkeit hat und sich im dichteren Medium langsamer ausbreitet ($n > 1$). Diese Tatsache war für Fermat aber irrelevant, denn es war zu seiner Zeit unmöglich, solche Aussagen experimentell zu überprüfen. Der Zusammenhang (3.3) zwischen n und c' konnte erst viel später experimentell verifiziert werden.

Es ist nicht schwierig, aufgrund dieses Prinzips die Snellius Formel der Brechung an der Grenzfläche zweier homogener Medien herzuleiten. Gemäss Abb.3.3 gilt

$$L(A, E; x) = n_1 \sqrt{x^2 + h^2} + n_2 \sqrt{(x_0 - x)^2 + h'^2}. \quad (3.5)$$

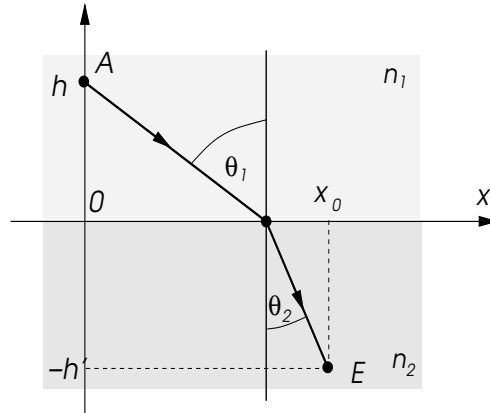


Abb. 3.3: Brechung mit der kürzesten optischen Weglänge.

Das Fermat'sche Prinzip fordert nun, dass x denjenigen Wert annimmt, bei welchem L minimal wird, d.h.,

$$0 \stackrel{!}{=} \frac{dL}{dx} = \frac{n_1 x}{\sqrt{x^2 + h^2}} - \frac{n_2 (x_0 - x)}{\sqrt{(x_0 - x)^2 + h'^2}} \Rightarrow n_1 \sin \theta_1 = n_2 \sin \theta_2. \quad (3.6)$$

Es ist aber auch möglich viel allgemeinere Situationen zu beschreiben, in denen der Brechungsindex kontinuierlich von der Position abhängt. Die variationelle Minimierung von (3.2) ergibt

$$\delta L = \int_A^E [\delta n ds + n \delta(ds)] \stackrel{!}{=} 0. \quad (3.7)$$

Wir führen $\delta \vec{r}$ als Abweichung von der Strahltrajektorie ein und erhalten

$$\left. \begin{array}{l} \delta n = \delta \vec{r} \cdot \vec{\nabla} n \\ \delta ds = \hat{s} d(\delta \vec{r}) \end{array} \right\} \Rightarrow 0 \stackrel{!}{=} \int_A^E [(\delta \vec{r} \cdot \vec{\nabla} n) ds + n \hat{s} \cdot d(\delta \vec{r})] = \int_A^E \left\{ \left[\vec{\nabla} n - \frac{d(n \hat{s})}{ds} \right] \cdot \delta \vec{r} \right\} ds, \quad (3.8)$$

wobei \hat{s} den Einheitsvektor tangential zur Strahltrajektorie bezeichnet. Die letzte Gleichung wird durch partielle Integration erhalten. Damit lautet die Variationsgleichung

$$\begin{aligned}\vec{\nabla}n &= \frac{d(n\hat{s})}{ds} = n \frac{d\hat{s}}{ds} + \hat{s} \frac{dn}{ds} = n \frac{d\hat{s}}{ds} + \hat{s} (\hat{s} \cdot \vec{\nabla}n) \\ \Rightarrow \frac{d\hat{s}}{ds} &= \frac{1}{n} \left\{ \vec{\nabla}n - \hat{s} (\hat{s} \cdot \vec{\nabla}n) \right\}.\end{aligned}\tag{3.9}$$

Schreiben wir die Ableitung $d\hat{s}/ds$ in der Form

$$\frac{d\hat{s}}{ds} = \frac{\hat{N}}{R},\tag{3.10}$$

mit dem Einheitsnormalenvektor $\hat{N} \perp \hat{s}$, $\hat{N}^2 = 1$ und dem Krümmungsradius R (siehe Abb.3.4), so folgt aus (3.9) durch Projektion auf \hat{N} die Gleichung

$$\frac{1}{R} = \hat{N} \frac{\vec{\nabla}n}{n}.\tag{3.11}$$

Der Lichtstrahl ist also immer in Richtung des ansteigenden Brechungsindexes gekrümmt.

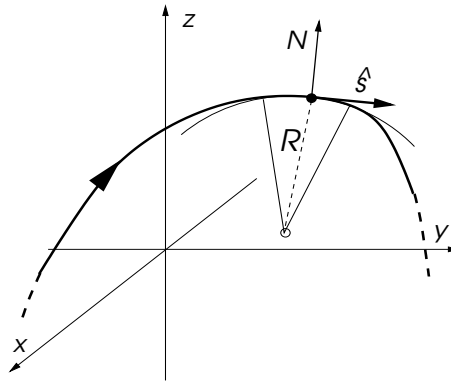


Abb. 3.4: Parameter der Lichttrajektorie: Einheitsvektoren \hat{s} und \hat{N} tangential und normal auf die Trajektorie und (momentaner) Krümmungsradius R .

Das Fermat'sche Prinzip erscheint auf den ersten Blick etwas mysteriös, wenn man sich fragt, wie das Licht sozusagen im Voraus weiss, welche Trajektorie es am schnellsten von A nach E bringt.¹

¹Beachte, dass dieser Aspekt in ähnlicher Form in der klassischen Mechanik (Variationsprinzip von Maupertuis) auftaucht. Unter allen Teilchen-Bahnen gleicher Energie E , die vom Punkt \vec{q}_A zur Zeit t_A starten und zur Zeit t_E bei \vec{q}_E enden, wird diejenige Bahn durchlaufen, deren "verkürzte Wirkung" (siehe Landau-Lifschitz Band I, Kapitel VII) gegeben ist durch

$$S' = \int_{\vec{q}_A}^{\vec{q}_E} \vec{p} \cdot d\vec{q},\tag{3.12}$$

mit dem verallgemeinertem Impuls \vec{p} und der verallgemeinerten Koordinate \vec{q} . Für eine Lagrange-Funktion der Form $L = m\dot{\vec{q}}^2/2 - V(\vec{q})$ gilt

$$S' = \int_{\vec{q}_A}^{\vec{q}_E} \sqrt{2m(E - V(\vec{q}))} ds.\tag{3.13}$$

Damit folgt die Teilchentrajektorie dem gleichen Prinzip wie der Lichtstrahl und $\sqrt{2m(E - V(\vec{q}))}$ übernimmt die Rolle des Brechungsindex. Analog gehorcht die Bahn eines Teilchens, das die Grenze zwischen zwei Bereichen unterschiedlichen Potentials (V_1 und V_2) überschreitet, dem Gesetz von Snellius (Abb.3.2,3.3),

$$\sqrt{E - V_1} \sin \theta_1 = \sqrt{E - V_2} \sin \theta_2.\tag{3.14}$$

Somit gelten die gleichen Gesetze wie für die Brechung und die Reflexion eines Lichtstrahls in der geometrischen Optik. Beachte jedoch, dass bei Fermat der Brechungsindex umgekehrt proportional zur Lichtgeschwindigkeit ist, während bei Maupertuis die Teilchengeschwindigkeit proportional zu $\sqrt{E - V}$ ist. Analog zu Gleichung (3.11) können wir schliessen, dass die Bahn eines Teilchens gegebener Energie immer in Richtung fallenden Potentials hin gekrümmt ist.

Die Antwort findet sich in der Wellennatur des Lichtes und den damit verbundenen Interferenz-Eigenschaften. Die Trajektorie wird durch die maximal konstruktive Interferenz festgelegt. Es zeigt sich zudem, dass die geometrische Optik nur Gültigkeit hat, wenn die Wellenlänge des Lichtes viel kleiner ist als die übrigen charakteristischen Längenskalen, wie z.B. die Ausdehnung möglicher Hindernisse oder die Skala verbunden mit der Änderung des Brechungsindex. Im Rahmen der geometrischen Optik bleibt aber das Fermat'sche Prinzip bis heute die wesentliche Arbeitsgrundlage.

3.1.4 Doppelbrechung

Im Jahre 1669 stellte der Däne Erasmus Bartholin (1625-1698) fest, dass die Lichtbrechung in Kalkspat oder Calcit komplizierter ist als in den meisten anderen Materialien. Ein Lichtstrahl, der dieses Material durchläuft wird in zwei Strahlen aufgespaltet.



Abb. 3.5: Doppelbrechung: (links) Lichtstrahl spaltet in zwei verschieden gebrochene Strahlen auf; (rechts) Verdopplung des Bildes durch Calcit CaCO_3 .

Dieses Phänomen legt nahe, dass Licht noch einen internen Freiheitsgrad besitzt, der nicht offensichtlich aus der geometrischen Optik folgt. Später erst hat man festgestellt, dass diese interne Struktur der Polarisation des Lichtes entspricht.

3.2 Die Natur des Lichtes

Wie schon erwähnt wurde in der Antike die Idee des Lichtes meist im Zusammenhang mit dem Sehen verstanden. Damit war "Licht" etwas, das vom Auge des Betrachters ausging. Es gab aber auch schon damals andere Vorstellungen. So vertrat Aristoteles die Idee, dass die Sicht durch dasjenige Licht entsteht, welches auf das Auge trifft und nicht umgekehrt. Ferner behauptete der Atomist Lukrez im 1. Jh. v. Chr., dass Licht und Wärme der Sonne von kleinen Teilchen übertragen werden, welche von der Sonne her zur Erde kommen. Einen entscheidenden Beitrag zum Verständnis des Lichtes lieferte der arabische Wissenschaftler Abu Ali al-Hasan ibn al-Hasan ibn al-Haytham (965-1039) oder kurz, *Alhazen*. Alhazens Konzept des Lichtes ist im Wesentlichen kompatibel mit unseren modernen Vorstellungen:

- Licht entspringt einer Lichtquelle, wird an Gegenständen reflektiert und trifft aufs Auge, wo der Gegenstand optisch wahrgenommen wird.
- Licht breitet sich geradlinig aus. Die Lichtstrahlen beeinflussen sich gegenseitig nicht.
- Licht breitet sich mit endlicher Geschwindigkeit aus.
- Die Geschwindigkeit des Lichtes ist kleiner in dichteren Medien.

Obwohl er viele dieser Thesen nicht beweisen konnte, waren die meisten seiner Erklärungen im Kerngedanken korrekt. Er hatte schon eine gute Vorstellung von der optischen Funktionsweise des Auges und experimentierte mit der Lochkamera. Seine "mikroskopischen" Vorstellungen

waren jedoch nicht richtig. Nichtsdestotrotz übte er einen wichtigen Einfluss auf die Entwicklung der Optik im Mittelalter aus. Ein entscheidender Fortschritt für das Verständnis der Optik war die quantitative Beschreibung der Lichtbrechung, welche unabhängig von Willebrod Snellius und René Decartes anfangs 17. Jh. gefunden wurde.

3.2.1 Teilchen versus Welle

Im 17. Jh. begann man sich erneut mit der Natur des Lichtes zu beschäftigen. Das Konzept des Sehstrahls, der vom Auge ausgeht, war kein Thema mehr. Vielmehr bildeten sich folgende zwei konkurrierende Konzepte heraus: Licht besteht aus einem Teilchenstrom (1); Licht ist eine Welle in einem unbekanntem Medium (2).

(1) Das Teilchenbild wurde unter anderen von Newton propagiert. Er liess sich davon leiten, dass die Teilchentrajektorie in seiner Theorie der Mechanik dem Verhalten des Lichtes in der geometrischen Optik entsprach. Die Brechung konnte wie oben gezeigt dadurch erklärt werden, dass die Teilchen sich in dichteren Medien schneller bewegten. Die hierfür notwendige Beschleunigung wurde der von Newton eingeführten Schwerkraft des Mediums zugeschrieben. Man konnte sogar das Phänomen der Lichtbeugung an scharfen Kanten mit Hilfe dieser Schwerkraft zumindest qualitativ nachvollziehen. Zudem konnte man im Teilchenbild die Farbeigenschaft mit der Existenz verschiedener Teilchenarten begründen. Die Untersuchungen der Spektralaufspaltung des Lichtes mit Prismen zählen zu den wichtigsten Beiträgen Newtons zur Physik des Lichtes.

Das Teilchenbild ist aber nicht in der Lage Interferenzeffekte zu erklären. Unter anderen hat auch Newton selbst über solche Interferenzeffekte aus eigenen Beobachtungen berichtet. Trotzdem hat er das Wellenbild verworfen, da es seiner Ansicht nach der geometrischen Optik widersprach.

(2) Einer der frühen Verfechter der Wellennatur des Lichtes war René Decartes. Ihm unterliefen allerdings einige Fehlinterpretationen. So nahm er in Analogie zum Schall in festen Körpern an, dass Licht sich in dichteren Medien schneller fortbewegte. Weitere Wellentheorien des Lichtes stammen von Francesco Grimaldi, Robert Hook und Christian Huygens. Grimaldi liess sich von seinen Beobachtungen zur Beugung leiten, während Hook vor allem mit selbstbeobachteten Interferenzeffekten argumentierte. Schliesslich entwickelte Huygens das Bild der Welle, die sich von jedem Punkt aus kugelförmig ausbreitet und sich mit anderen Wellen linear überlagern lässt. Er war auch der Erste, der eine quantitative moderne Wellentheorie entwickelte. Zudem nahm er korrekterweise an, dass Lichtwellen in dichteren Medien langsamer vorankommen. Die damalige Wellentheorie hatte keine Erklärung für die Farben und Probleme mit dem Beugungsphänomen, was der geometrischen Optik zu widersprechen schien. Das Problem lag in der Tatsache, dass die Wellenlängen des Lichtes sehr klein sind und man sich damals bei Wellen eher am vertrauten Beispiel des (langwelligen) Schalls orientierte.

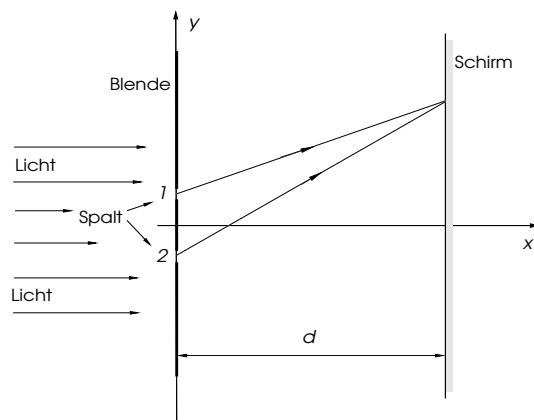


Abb. 3.6: Youngs Experiment: Das Doppelspalt-Experiment liefert ein Interferenzmuster, aus dem sich die Wellenlänge des Lichtes ablesen lässt.

Die Wellentheorie hatte neben der Teilchen-Theorie (Korpuskular-Theorie) Newtons keine Chance. Newtons Autorität, die auf dem Erfolg seiner Mechaniklehre beruhte, war zu gross. Damit war die Physik des Lichtes für über ein Jahrhundert blockiert. Erst im Jahre 1802 leitete das Doppelspalt-Experiment von Thomas Young eine Wende hin zum Wellenbild ein. Er stellte fest, dass sich ein Interferenzmuster ausbildete, wenn man Licht durch zwei dünne, sehr nahe beieinander liegende Spalten einer Blende laufen liess. Dies entsprach dem Verhalten, welches man für Wellen erwartete. Hinzu kam eine weitere entscheidende Tatsache, nämlich dass das Interferenzmuster von der Farbe des Lichtes abhing und man eine Beziehung zwischen Farbe und Wellenlänge des Lichtes herstellen konnte. Dies kann einfach nachvollzogen werden; wir nehmen an, dass eine ebene Welle der Wellenlänge λ (Wellenzahl $k = 2\pi/\lambda$) auf die Doppelspaltblende trifft und von den beiden Öffnungen je eine "Kreiswelle" ausgeht. Sind $\vec{a}_1 = (0, +y_0)$ und $\vec{a}_2 = (0, -y_0)$ die Positionen der Öffnungen und bezeichnet d die Distanz zwischen Blende und Detektorschirm, so erhalten wir im Grenzfall $d \gg \lambda, y_0$ (Schirm weit weg von Blende) für die Überlagerung der beiden Kreiswellen

$$f(\vec{r}) \propto \left| \frac{e^{ik|\vec{r}-\vec{a}_1|}}{\sqrt{|\vec{r}-\vec{a}_1|}} + \frac{e^{ik|\vec{r}-\vec{a}_2|}}{\sqrt{|\vec{r}-\vec{a}_2|}} \right|^2 \approx \frac{2}{d} \left[1 + \cos\left(\frac{y_0}{d}ky\right) \right], \quad (3.15)$$

wobei wir annehmen, dass die Distanzen (y) auf dem Schirm viel kleiner sind als d . Die Intensitätsmodulation auf dem Schirm hat dann die Wellenlänge $\lambda_S = \lambda d/y_0$. Mit diesem Experiment stellte Young fest, dass die Wellenlänge des Lichtes zwischen 400 nm (violett) und 700 nm (rot) liegt. Dies sind sehr kleine Wellenlängen, womit erklärt war, warum Beugungseffekte bei Licht (im Gegensatz zum Schall) bis anhin nicht beobachtet wurden, und warum die Strahlenoptik für "Alltagssituationen" weitgehend gültig war. Unabhängig von Young kam auch Augustin Fresnel mit ähnlichen Experimenten auf das gleiche Ergebnis.

Zu Beginn des 19. Jh. wurde das Rätsel der Doppelbrechung (siehe Abb.3.5) gelöst. Etienne-Louis Malus führte den Begriff "Polarisation" ein, der die beiden getrennten Strahlen unterscheidet, und erkannte, dass die beiden Typen "senkrecht" aufeinander stehen. Als Fresnel berichtete, dass es keine Interferenz zwischen Lichtstrahlen orthogonaler Polarisation gab, schlug Young schliesslich vor, Licht als Transversalwelle zu betrachten.

Ein weiterer, wichtiger Schritt war die Messung der Lichtgeschwindigkeit. Die erste seriöse Messung stammt aus dem 17. Jh., als Olaf Romer 1676 aus der Beobachtung der Jupitermonde einen erstaunlich guten Wert für die Lichtgeschwindigkeit erhielt: $c = 225000 \text{ km s}^{-1}$. In der Folge wurde die Messung der Lichtgeschwindigkeit mit verschiedensten Methoden immer genauer. Konzeptionell sehr wichtig war die Messung der Lichtgeschwindigkeit in dichteren Medien. 1850 stellte J.B.L. Foucault fest, dass Licht sich im Wasser langsamer bewegt als in Luft. Damit war endgültig klar, dass die Korpuskulartheorie nicht mehr haltbar war; denn mit dieser Theorie lässt sich das Snellius-Gesetz nur reproduzieren, falls die Lichtgeschwindigkeit mit der Dichte der Medien ansteigt.

3.3 Licht als elektromagnetische Welle

3.3.1 Faraday-Effekt

Trotz den tiefen Einsichten von Young, Fresnel und einigen anderen war es noch nicht klar, was Licht war und durch welches Medium die Welle getragen wurde. Ein neuer Aspekt tauchte mit der Entdeckung des Faraday-Effektes auf. 1845 untersuchte Micheal Faraday, wie Licht sich in transparenten Medien unter Einfluss von Magnetfeldern verhielt. Dabei beobachtete er, dass die Polarisationsebene des Lichtes sich in schwerem Glas dreht, wenn ein Magnetfeld in Richtung des Lichtstrahles angelegt wird (Abb.3.7). Der Drehwinkel hängt linear vom Magnetfeld B und von der Distanz d ab, die das Licht durch das Medium zurücklegt,

$$\beta = VBd, \quad (3.16)$$

wobei V die Verdet-Konstante, ein materialabhängiger Parameter, ist. Mit der Beobachtung dieses magneto-optischen Effektes wurde erstmals ein Zusammenhang zwischen Elektromagnetismus (Elektrizität und Magnetismus wurden schon seit mehr als zwei Jahrzehnten als ein gemeinsames Phänomen verstanden) und Licht hergestellt. Wichtig in diesem Kontext ist auch der Feldbegriff, welcher von Faraday eingeführt wurde. Elektromagnetismus wird nicht mehr nur über wirkende Kräfte verstanden, sondern dem elektrischen und magnetischen Feld kommt eine eigenständige Bedeutung zu.

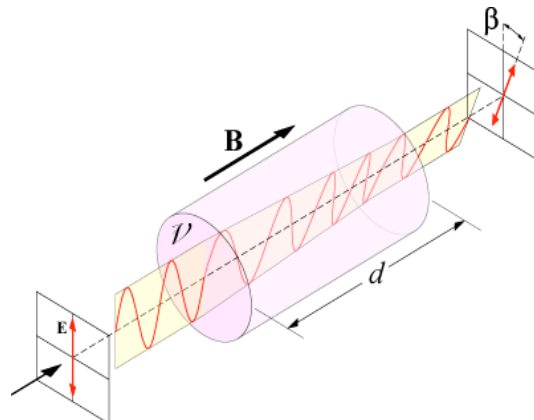


Abb. 3.7: Faraday-Effekt: Die Polarisationssebene des Lichtes rotiert um den Winkel β , wenn an das transparente Medium ein Magnetfeld in der Propagationsrichtung des Lichtes angelegt wird.

3.3.2 Maxwell-Theorie

Das tiefere theoretische Verständnis verdanken wir James Clark Maxwell, dem es gelungen ist ein System von Gleichungen aufzustellen, die den Elektromagnetismus komplett beschreiben und eine geschlossene Theorie des Lichtes beinhalten. Dabei liess er sich zunächst von den bekannten Tatsachen des Elektromagnetismus leiten,

$$\begin{aligned}
 \vec{\nabla} \cdot \vec{D} &= 4\pi\rho, & \text{Coulomb-Gesetz,} \\
 \vec{\nabla} \cdot \vec{B} &= 0, & \text{quellenfreies Magnetfeld,} \\
 \vec{\nabla} \times \vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} &= 0, & \text{Faraday Induktionsgesetz,} \\
 \vec{\nabla} \times \vec{H} &= \frac{4\pi}{c} \vec{j}, & \text{Ampère-Gesetz,}
 \end{aligned}
 \tag{3.17}$$

wobei \vec{D} die dielektrische Verschiebung, \vec{B} die magnetische Induktion, \vec{E} das elektrische Feld und \vec{H} das Magnetfeld bezeichnen. Im Vakuum gilt $\vec{E} = \vec{D}$ und $\vec{B} = \vec{H}$.

Maxwell stellte fest, dass die vier Gleichungen in sich nicht konsistent waren. Nehmen wir die Divergenz der letzte Gleichung (Ampère-Gesetz), so finden wir

$$\underbrace{\vec{\nabla} \cdot (\vec{\nabla} \times \vec{H})}_{=0} = \frac{4\pi}{c} \vec{\nabla} \cdot \vec{j}.
 \tag{3.18}$$

Dies ist aber inkompatibel mit der Kontinuitätsgleichung, $\partial\rho/\partial t + \vec{\nabla} \cdot \vec{j} = 0$. Die folgende Erweiterung der Ampère-Gleichung,

$$\vec{\nabla} \times \vec{H} - \frac{1}{c} \frac{\partial \vec{D}}{\partial t} = \frac{4\pi}{c} \vec{j},
 \tag{3.19}$$

führt auf die korrekte Form, wenn wir die Divergenz anwenden. Damit sind die Maxwell-Gleichungen komplett. Die Formulierung in Gauss'schen Einheiten enthält c als die Lichtgeschwindigkeit.²

Für die weitere Diskussion des Lichtes ist diese Erweiterung der Gleichungen entscheidend. Faradays Induktionsgesetz besagt, dass ein elektrisches Feld durch ein zeitlich variierendes Magnetfeld erzeugt werden kann. Die Erweiterung des Ampère-Gesetzes führte nun dazu, dass die reziproke Aussage auch gültig ist: ein sich zeitlich änderndes elektrisches Feld induziert ein Magnetfeld. Dies birgt die Möglichkeit eines sich "selbst erhaltenden", propagierenden elektromagnetischen Feldes in sich.

Es ist interessant zu sehen, dass die "transversalen" Maxwell-Gleichungen die Lichtgeschwindigkeit als Parameter enthalten. Betrachten wir vorderhand die Situation ohne Ladungen und Ströme, $\rho = 0$ und $\vec{j} = 0$ (Vakuum). Dann können wir die beiden Gleichungen zusammenfassen zu (wir verwenden, dass im Vakuum $\vec{B} = \vec{H}$ und $\vec{E} = \vec{D}$)

$$\left. \begin{aligned} \vec{\nabla} \times \vec{E} &= -\frac{1}{c} \frac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \times \vec{B} &= +\frac{1}{c} \frac{\partial \vec{E}}{\partial t} \end{aligned} \right\} \Rightarrow \left\{ \begin{aligned} \vec{\nabla}^2 \vec{E} - \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} &= 0 \\ \vec{\nabla}^2 \vec{B} - \frac{1}{c^2} \frac{\partial^2 \vec{B}}{\partial t^2} &= 0 \end{aligned} \right. . \quad (3.22)$$

Diese Gleichungen ergeben Wellengleichungen, deren elementare Lösungen ebene Wellen sind,

$$\vec{E}(\vec{r}, t) = \vec{E}_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)} \quad \text{und} \quad \vec{B}(\vec{r}, t) = \vec{B}_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)} . \quad (3.23)$$

Durch Einsetzen dieser Lösungen in die Maxwell-Gleichungen ergeben folgende Beziehung zwischen \vec{E} und \vec{B} ,

$$\vec{k} \times \vec{E} = \frac{\omega}{c} \vec{B} \quad \text{und} \quad \vec{k} \times \vec{B} = -\frac{\omega}{c} \vec{E} , \quad (3.24)$$

wobei gilt $\omega = c|\vec{k}|$. Zudem ist $\vec{E} \perp \vec{B}$ und beide stehen senkrecht auf \vec{k} . Youngs Hypothese war also korrekt: Licht ist eine transversale Welle.

Aus Maxwells Theorie folgt, dass Licht eine elektromagnetische Welle ist und dass elektromagnetische Wellen vielmehr als nur das sichtbare Licht einschliessen. Die Existenz von infraroter und ultravioletter Strahlung war schon bekannt: der Astronom William Herschel hatte um 1800 bei der Untersuchung des Sonnenspektrums durch Zufall die Infrarot-Strahlung entdeckt (indem er die Wärmeeigenschaften des Spektrums mit einem Thermometer analysierte). Johann Wilhelm Ritter beobachtete um 1801, dass es einen Bereich oberhalb des Violetts im Farbspektrum gibt, der besonders effektiv eine Reaktion in gewissen Silber-Salzen (z.B. AgCl) hervorruft (Schwärzung). Man nannte diese Strahlung daher zuerst "chemische Strahlung".

Im Jahre 1886 konnte Heinrich Hertz mittels eines oszillierenden elektrischen Dipols zum ersten Mal kontrolliert elektromagnetische Wellen erzeugen und diese wieder detektieren. Dabei konnte er auch die Struktur der Wellen nachweisen.

²Maxwell-Gleichungen in SI-Einheiten:

$$\begin{aligned} \vec{\nabla} \cdot \vec{D} &= \rho , & \vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0 , \\ \vec{\nabla} \cdot \vec{B} &= 0 , & \vec{\nabla} \times \vec{H} - \frac{\partial \vec{D}}{\partial t} &= \vec{j} , \end{aligned} \quad (3.20)$$

mit $\vec{D} = \epsilon_0 \vec{E}$, $\epsilon_0 = 8.8542 \times 10^{-12} \text{ As V}^{-1} \text{ m}^{-1}$ als elektrische Feldkonstante und mit $\vec{B} = \mu_0 \vec{H}$, $\mu_0 = 4\pi \times 10^{-7} \text{ Vs A}^{-1} \text{ m}^{-1}$ als magnetische Feldkonstante. Es gilt die Beziehung,

$$\epsilon_0 \mu_0 = c^{-2} . \quad (3.21)$$

In Gauss'schen Einheit, die wir normalerweise benutzen, sind die Einheit von \vec{E} und \vec{D} , sowie von \vec{B} und \vec{H} identisch.

3.4 Licht und Materie

Um Elektromagnetismus in Materie zu beschreiben, muss der Effekt der Polarisation der Materie berücksichtigt werden. Dies führt zu einer Renormierung des elektrischen Feldes und des Magnetfeldes,

$$\vec{D} = \epsilon \vec{E} \quad \text{und} \quad \vec{B} = \mu \vec{H}. \quad (3.25)$$

Diese sogenannte lineare Antwort der Materie auf das Anlegen eines äusseren Feldes wird durch die dielektrische Konstante ϵ und die magnetische Permeabilität μ beschrieben (im Vakuum $\epsilon = \mu = 1$).

3.4.1 Transmission und Reflexion in transparenten Medien

Wir betrachten nun Materialien, die keine elektrischen Ströme tragen können. Setzen wir (3.25) in die Wellengleichungen ein, ändert sich lediglich der Vorfaktor

$$\frac{1}{c^2} \rightarrow \frac{\epsilon\mu}{c^2}, \quad (3.26)$$

d.h., die Lichtgeschwindigkeit wird renormiert. Nach unseren früheren Betrachtungen identifizieren wir den Brechungsindex n mit $n^2 = \epsilon\mu \geq 1$. Dies ergibt die Lichtbrechung. Zusätzlich lässt sich nun auch die Reflexion beschreiben.

Der Einfachheit halber betrachten wir Licht, das senkrecht auf die Grenzfläche zwischen zwei Medien trifft. Die Frequenz der Welle ist ω und bleibt unverändert in den beiden Medien 1 und 2.

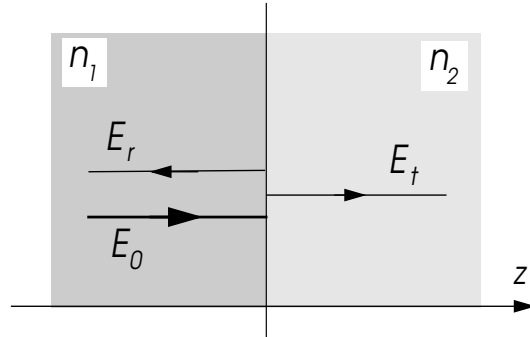


Abb. 3.8: Einfallende Welle E_0 spaltet in reflektierten Anteil E_r und transmittierten Anteil E_t auf.

Sowohl \vec{E} und \vec{H} , als auch ihre Ableitungen sind an der Grenzfläche stetig. Betrachten wir das elektrische Feld, so fällt vom Medium 1 her eine ebene Welle (E_0) ein, die sich in einen reflektierten Anteil (E_r) und einen ins Medium 2 transmittierten Anteil (E_t) aufspaltet,

$$\vec{E}(z, t) = \begin{cases} \vec{E}_0 e^{i(k_1 z - \omega t)} + \vec{E}_r e^{-i(k_1 z + \omega t)} & z < 0 \\ \vec{E}_t e^{i(k_2 z - \omega t)} & z > 0 \end{cases}, \quad (3.27)$$

mit $k_j = \omega n_j / c$ ($j = 1, 2$). Die Stetigkeitsbedingungen an $\vec{E}(z, t)$ und an die Ableitung nach z bei $z = 0$ liefert die Gleichungen

$$\left. \begin{aligned} \vec{E}_0 + \vec{E}_r &= \vec{E}_t \\ n_1(\vec{E}_0 - \vec{E}_r) &= n_2 \vec{E}_t \end{aligned} \right\} \Rightarrow \begin{cases} \vec{E}_r = \frac{n_1 - n_2}{n_1 + n_2} \vec{E}_0 \\ \vec{E}_t = \frac{2n_1}{n_1 + n_2} \vec{E}_0 \end{cases}. \quad (3.28)$$

Definieren wir die Reflektivität R und die Transmission T folgendermassen,

$$R = \left| \frac{\vec{E}_r}{\vec{E}_0} \right|^2 = \left(\frac{n_1 - n_2}{n_1 + n_2} \right)^2 \quad \text{und} \quad T = \left| \frac{\vec{E}_t}{\vec{E}_0} \right|^2 \frac{n_2}{n_1} = \frac{2n_1 n_2}{(n_1 + n_2)^2}, \quad (3.29)$$

so gilt $R + T = 1$.³ Für den senkrechten Einfall der Welle spielt die Polarisation des Lichtes keine Rolle. Dies ist nicht mehr so, falls das Licht in einem schiefen Winkel auf die Grenzfläche trifft.

3.4.2 Magneto-Optik und Faraday-Effekt

Wir betrachten nun den Faraday-Effekt in der Maxwell-Theorie. Ein isotropes, homogenes Medium in einem Magnetfeld verändert seine dielektrische Konstante, die im allgemeinen Fall als Tensor aufgefasst werden muss, ϵ_{ij} . Für ein transparentes Medium (keine Absorption) muss dieser Tensor hermitesch sein: $\epsilon_{ij} = \epsilon_{ji}^*$. Im Magnetfeld kann man diesen Tensor durch einen Gyrationvektor \vec{g} ausdrücken,

$$\vec{D} = \epsilon \vec{E} + i \vec{E} \times \vec{g} \quad \Rightarrow \quad \epsilon_{ij} = \epsilon \delta_{ij} + i \epsilon_{ijk} g_k, \quad (3.32)$$

mit ϵ_{ijk} als total antisymmetrischer Tensor.⁴ Der Gyrationvektor hängt näherungsweise linear vom Magnetfeld ab,

$$\vec{g} = f_{mo} \vec{H}. \quad (3.36)$$

Wir untersuchen nun die Situation (wie bei Faradays Experiment) mit einem Magnetfeld entlang der Propagationsrichtung des Lichtes, $\vec{H} \parallel z$ und gemäss (3.36) auch $\vec{g} = (0, 0, g)$. Die Wellengleichung für das elektrische Feld im Medium lautet dann komponentenweise,

$$\begin{aligned} \frac{\partial^2 E_x}{\partial z^2} &= \frac{\omega^2}{c^2} \mu \{ \epsilon E_x + i g E_y \}, \\ \frac{\partial^2 E_y}{\partial z^2} &= \frac{\omega^2}{c^2} \mu \{ -i g E_x + \epsilon E_y \}. \end{aligned} \quad (3.37)$$

³Die Definition der Transmission und Reflexion benutzt den Energiestrom, der über den Poynting-Vektor definiert wird,

$$\vec{S} = \frac{c}{4\pi} \vec{E} \times \vec{H} = \frac{c\sqrt{\epsilon\mu}}{4\pi} \vec{E}^2 \frac{\vec{k}}{k} \quad \text{und} \quad \vec{S}_0 + \vec{S}_r = \vec{S}_t, \quad (3.30)$$

wobei der Energiestrom erhalten ist, so dass

$$n_1 \{ |\vec{E}_0|^2 - |\vec{E}_r|^2 \} = n_2 |\vec{E}_t|^2 \quad \Rightarrow \quad 1 - R = T. \quad (3.31)$$

Dies ergibt (3.29).

⁴Die Form des Dielektrizitätstensors kann leicht verstanden werden, indem wir die Polarizierbarkeit eines transparenten (isolierenden) Mediums durch Ladungen beschreiben, die in einem harmonischen Potential gefangen sind (lokalisierte Rumpfelektronen im Atom oder Valenzelektronen in kovalenten Bindungen). Ein angelegtes, oszillierendes elektrisches Feld führt zu einer Auslenkung $\vec{s}(t)$ gemäss der Bewegungsgleichung,

$$m \ddot{\vec{s}} - m \omega_0^2 \vec{s} = e \vec{E} e^{i\omega t} - \frac{e}{c} \dot{\vec{s}} \times \vec{B}, \quad (3.33)$$

wobei wir den letzten Term, die Lorentzkraft, als kleine Störung annehmen. Somit erhalten wir für $\vec{s}(t)$ bis zu erster Ordnung in \vec{H} ,

$$\vec{s}(t) = -\frac{e \vec{E}}{m(\omega_0^2 - \omega^2)} - \frac{e^2 \mu}{c} \frac{i\omega}{m(\omega_0^2 - \omega^2)} (\vec{E} \times \vec{H}). \quad (3.34)$$

Daraus ergibt sich die Polarization $\vec{P} = -n_{\text{ion}} e \vec{s}$, und mit $\vec{D} = \vec{E} + 4\pi \vec{P}$ für gegebene Frequenz ω folgt,

$$\vec{D} = \epsilon \vec{E} + i(\vec{E} \times \vec{g}) \quad \Rightarrow \quad \begin{cases} \epsilon(\omega) = \frac{e^2 n_{\text{ion}}}{m(\omega_0^2 - \omega^2)} \\ \vec{g}(\omega) = \frac{e^3 n_{\text{ion}} \mu}{c} \frac{\omega}{(m(\omega_0^2 - \omega^2))^2} \vec{H} \end{cases}. \quad (3.35)$$

Dabei nehmen wir an, dass die Frequenz kleiner ist als die charakteristische Frequenz ω_0 der Elektronen ($\hbar\omega_0$: Anregungsenergie).

Durch Einführung der zirkulären Polarisation, $E_{\pm} = E_x \pm iE_y$, lässt sich diese Gleichung diagonalisieren,

$$\frac{\partial^2 E_{\pm}}{\partial z^2} = \frac{\omega^2 \mu}{c^2} (\epsilon \pm g) E_{\pm} = \frac{\omega^2}{c^2} n_{\pm}^2 E_{\pm} . \quad (3.38)$$

Folglich besitzen links- und rechtspolarisiertes Licht unterschiedliche Brechungsindizes und propagieren demnach mit unterschiedlichen Geschwindigkeiten im Medium.

Betrachten wir eine linear polarisierte, entlang der z -Richtung einfallende Welle mit der Polarisation $\vec{E}_0 = (E_0, 0, 0)$. Die entsprechende magnetische Komponente ergibt sich aus (3.24). Ferner sei das Medium 1 das Vakuum ($n_1 = 1$) und das Medium 2 habe in z -Richtung die Ausdehnung d . Dies ergibt zwei Grenzflächen (Vakuum - Medium 2 - Vakuum). Konzentrieren wir uns auf die transmittierte Welle, dann erhalten wir

$$E_{t\pm} = \frac{4n_{\pm}}{(1+n_{\pm})^2} e^{ik_{\pm}d} \vec{E}_{0\pm} \Rightarrow \begin{cases} E_{tx} = \frac{1}{2}(E_{r+} + E_{r-}) \approx \frac{4nE_0}{(1+n)^2} \cos(qd) e^{ikd} \\ E_{ty} = \frac{1}{2i}(E_{r+} - E_{r-}) \approx \frac{4nE_0}{(1+n)^2} \sin(qd) e^{ikd} \end{cases} , \quad (3.39)$$

mit $k = n\omega/c$ und $q = kng/2\epsilon$. Daraus wird sofort ersichtlich, dass sich die Polarisationsebene gedreht hat, und zwar um den Winkel $\beta = qd = V Hd$, so dass man für die Verdet-Konstante den Wert

$$V = \frac{kng}{2} \quad (3.40)$$

findet. Korrekturen höherer Ordnung in \vec{g} ergeben zusätzliche zirkular polarisierte Komponenten der Ordnung $O(g)$, die wir hier vernachlässigt haben.

3.4.3 Antireflexion - Optische Beschichtung

Die erste Diskussion einer reflexionshemmenden Beschichtung stammt wahrscheinlich von Rayleigh im Jahre 1886. Er betrachtete die Optimierung der Transmission für eine einzelne Schicht auf einem Medium und führte damit das Prinzip der optischen Beschichtung (optical coating) ein.

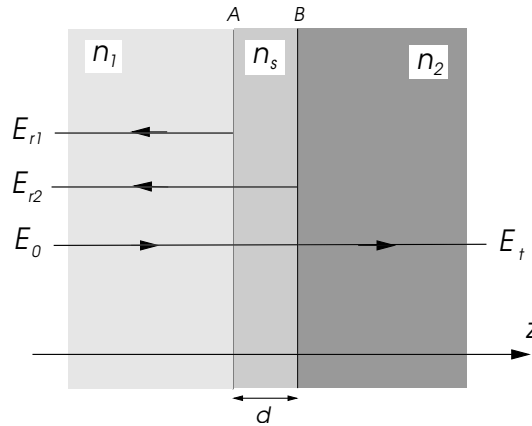


Abb. 3.9: Oberflächenschicht der Dicke d und Brechungsindex n_s führt auf zwei reflektierte Anteile je von Grenzfläche A und B, die miteinander interferieren.

Die beiden äusseren Medien besitzen die Brechungsindizes n_1 und n_2 , während die Beschichtung aus einem dünnen Film (Dicke d und Brechungsindex n_s) zwischen diesen Medien besteht. Bei gegebener Frequenz ω des Lichtes soll die Dicke des Films so gewählt werden, dass das reflektierte Licht auf dem Weg im Film ($A \rightarrow B \rightarrow A$) die Phase π oder allgemeiner $(2n+1)\pi$ (n : ganze Zahl) akkumuliert. Es muss also gelten

$$2dk_s = (2n+1)\pi \quad \text{mit} \quad k_s = \frac{\omega n_s}{c} = \frac{2\pi}{\lambda_s} . \quad (3.41)$$

Im einfachsten Fall sei $d = \pi c/2n_s\omega = \lambda_s/4$ (“ $\lambda_s/4$ -Regel”), so dass die Wellen, die an der Grenze A bzw. Grenze B reflektiert werden, destruktiv interferieren. Wir können den relativen Anteil der gesamten reflektierten Welle durch einfache Rechnung bestimmen,

$$\left| \frac{E_r}{E_0} \right| = \frac{k_1 k_2 - k_s^2}{k_1 k_2 + k_s^2} = \frac{n_1 n_2 - n_s^2}{n_1 n_2 + n_s^2}. \quad (3.42)$$

Wählen wir $n_s = \sqrt{n_1 n_2}$, so verschwindet die Reflexion vollständig. Die totale Auslöschung setzt allerdings die Bedingung $k_s d = (2n + 1)\pi/2$ voraus und gilt nicht für beliebige Frequenzen des Lichtes.

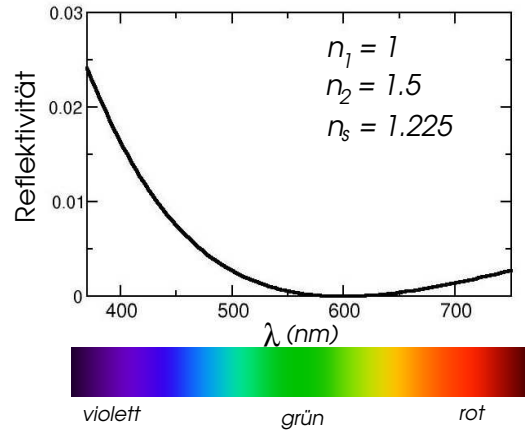


Abb. 3.10: Reflektivität einer Oberflächenschicht der Dicke $d = \lambda_s/4$, wobei λ_s der Vakuumwellenlänge $\lambda = 600$ nm (grün-gelb) entspricht, und des Brechungsindex $n_s = \sqrt{n_1 n_2}$. Die Reflektivität ohne Beschichtung beträgt $R = 1/9 = 0.111$.

Im allgemeineren Fall gilt für eine Schicht der Dicke d und $n_s = \sqrt{n_1 n_2}$,

$$\left| \frac{E_r}{E_0} \right|^2 = \frac{(n_1 - n_2)^2 \cos^2(k_s d)}{4n_s^4 \sin^2(k_s d) + (n_1 + n_2)^2 \cos^2(k_s d)}, \quad (3.43)$$

woraus sich die maximale Reflexion für $\cos(k_s d) = \pm 1$ ergibt,

$$\frac{E_r}{E_0} = \frac{n_1 - n_2}{n_1 + n_2}. \quad (3.44)$$

Beachte, dass dies demselben Faktor wie im Falle ohne Beschichtung entspricht. Damit folgt, dass die Beschichtung auf den ganzen Frequenzbereich gesehen eine bessere Transparenz aufweist, obwohl die totale Transmission nur für eine bestimmte Frequenz auftritt. Wenn wir uns auf das sichtbare Licht beschränken, können wir durch geeignete Wahl von d ein sehr gutes Antireflexionsverhalten erreichen, wie in Abb.3.10 für $n_1 = 1$, $n_2 = 1.5$, $n_s = 1.225$ und $n_s d = 150$ nm gezeigt. Beachte, dass die Reflexion für die unbeschichtete Grenzfläche $R = \{(n_1 - n_2)/(n_1 + n_2)\}^2 = 1/9 = 0.1111$ ist, während sie für die beschichtete Grenzfläche 0.025 im sichtbaren Bereich nicht übersteigt.

Natürlich ist es möglich durch bessere Formen der Beschichtung den Frequenzbereich mit optimalem Antireflexionsverhalten zu verbreitern. Solche optischen Beschichtungen werden bei optischen Instrumenten (Fernrohre, Feldstecher, etc.) verwendet. Aber auch bei Solarzellen finden wir solche Beschichtungen, um den ins Material eindringenden Anteil des Lichtes zu erhöhen. Dies erklärt den bläulichen (dunklen) Schimmer der Solarzelle, die sonst eher grauglänzend wäre; durch optische Beschichtung wird der niederfrequente Teil des sichtbaren Farbspektrums besser absorbiert und vorwiegend der höherfrequente blaue Anteil reflektiert.

3.5 Licht und Metalle

Bisher haben wir nur transparente, isolierende Medien betrachtet. Nun wenden wir uns abschliessend den Metallen zu. Polierte Metalloberflächen reflektieren das sichtbare Licht sehr gut. Untereinander unterscheiden sich Metalle jedoch in der Farbe. So kennen wir weisses Silber, gelbes Gold und rotes Kupfer. Dieser Unterschied wird von den Reflexionseigenschaften bestimmt. Da im Metall die elektrischen Ströme nicht verschwinden, sobald elektrische Felder vorhanden sind, müssen wir die Wellengleichungen erweitern. Hierfür starten wir wiederum mit den Maxwell-Gleichungen

$$\begin{aligned}\vec{\nabla} \times \vec{E} &= -\frac{1}{c} \frac{\partial \vec{B}}{\partial t}, \\ \vec{\nabla} \times \vec{B} &= \frac{1}{c} \frac{\partial \vec{E}}{\partial t} + \frac{4\pi}{c} \vec{j},\end{aligned}\tag{3.45}$$

wobei wir hier $\mu = 1$ annehmen. Im Folgenden beschränken wir uns auf die Mode $\vec{E} = \vec{E}(\omega)e^{-i\omega t}$ der (beliebig gewählten) Frequenz ω und benutzen die Ohmsche Beziehung $\vec{j}(\omega) = \sigma(\omega)\vec{E}(\omega)$. Dies führt uns auf die Wellengleichungen,

$$\begin{aligned}\vec{\nabla}^2 \vec{E} + \frac{\omega^2}{c^2} \vec{E} + \frac{4\pi i \omega \sigma(\omega)}{c^2} \vec{E} &= 0, \\ \vec{\nabla}^2 \vec{B} + \frac{\omega^2}{c^2} \vec{B} + \frac{4\pi i \omega \sigma(\omega)}{c^2} \vec{B} &= 0,\end{aligned}\tag{3.46}$$

aus welchen wir die dynamische Dielektrizitätsfunktion ablesen können,

$$\epsilon(\omega) = 1 + \frac{4\pi i \sigma(\omega)}{\omega}.\tag{3.47}$$

Beachte, dass wir in unserer Betrachtung die Beiträge der Ionenrümpfe zu $\epsilon(\omega)$ vernachlässigt haben.

3.5.1 Dynamische Dielektrizitätsfunktion

Die Leitfähigkeit $\sigma(\omega)$ für ein oszillierendes elektrisches Feld lässt sich in einem einfachen Modell basierend auf einem Gas von Elektronen bestimmen. Dabei folgen wir Drude und verwenden folgende Bewegungsgleichung,

$$\dot{\vec{p}} = -e\vec{E}e^{-i\omega t} - \frac{1}{\tau}\vec{p} \quad \Rightarrow \quad \vec{p}(t) = \frac{-e\tau\vec{E}}{1 - i\omega\tau}e^{-i\omega t} = \vec{p}(\omega)e^{-i\omega t}.\tag{3.48}$$

Mit $\vec{j}(\omega) = -ne\vec{p}(\omega)/m$ folgt

$$\sigma(\omega) = \frac{ne^2\tau}{m} \frac{1}{1 - i\omega\tau}.\tag{3.49}$$

Somit erhalten wir für die Dielektrizitätsfunktion $\epsilon = \epsilon_1 + i\epsilon_2$,

$$\epsilon(\omega) = 1 - \frac{\omega_p^2}{\omega(\omega + i/\tau)} \quad \Rightarrow \quad \begin{cases} \epsilon_1(\omega) = 1 - \frac{\omega_p^2\tau^2}{1 + \omega^2\tau^2}, \\ \epsilon_2(\omega) = \frac{\omega_p^2\tau}{\omega(1 + \omega^2\tau^2)}. \end{cases}\tag{3.50}$$

Hier bezeichnet

$$\omega_p^2 = \frac{4\pi ne^2}{m}\tag{3.51}$$

die sogenannte Plasmafrequenz; für homogene Oszillationen des Elektronengases beschreibt ω_p^{-1} die schnellste Zeitskala, in welcher Elektronen auf ein elektrisches Feld reagieren können.

3.5.2 Optische Eigenschaften des Metalls

Wir betrachten nun eine monochromatische Welle, die vom Vakuum her kommend ($n_1 = 1$) auf eine Metalloberfläche trifft. Wir behandeln diese Situation analog zu den obigen Fällen. Die Wellengleichung für das elektrische Feld innerhalb des Metalls lautet

$$\frac{d^2}{dz^2} \vec{E} = \frac{\omega^2}{c^2} n_2 \vec{E} \quad \text{mit} \quad n_2 = n' + in'' \quad \begin{cases} \epsilon_1 = n'^2 - n''^2, \\ \epsilon_2 = 2n'n'' . \end{cases} \quad (3.52)$$

Für eine aus dem Vakuum ($z < 0$) einfallende Welle ($\omega = ck$) erhalten wir

$$\begin{aligned} \vec{E}(z) &= \vec{E}_0 e^{ikz} + \vec{E}_r e^{-ikz} & z < 0, \\ \vec{E}(z) &= \vec{E}_t e^{iknz} = \vec{E}_t e^{ikn'z} e^{-kn''z} & z > 0. \end{aligned} \quad (3.53)$$

Im Metall wird die Welle offensichtlich gedämpft und zerfällt auf der Länge $\delta = 1/kn''$, der sogenannten Eindringtiefe. Durch die Stetigkeitsbedingungen an der Grenze erhalten wir die Reflektivität

$$R = \left| \frac{\vec{E}_r}{\vec{E}_0} \right|^2 = \left| \frac{1-n}{1+n} \right|^2 = \frac{(n'-1)^2 + n''^2}{(n'+1)^2 + n''^2}. \quad (3.54)$$

Untersuchen wir nun das Verhalten von R und δ als Funktion der Frequenz. Dazu unterscheiden wir drei Bereiche: (1) $\omega\tau \ll 1$, (2) $1 \ll \omega\tau \ll \omega_p\tau$ und (3) $\omega_p < \omega$.

(1) Hier finden wir mit (3.52)

$$\left. \begin{aligned} \epsilon_1 &\approx -\omega_p^2 \tau^2 \\ \epsilon_2 &\approx \frac{\omega_p^2 \tau}{\omega} \end{aligned} \right\} \Rightarrow n' \approx n'' \approx \sqrt{\frac{\epsilon_2}{2}} \approx \sqrt{\frac{\omega_p^2 \tau}{2\omega}} \gg 1. \quad (3.55)$$

Damit wird $R \approx 1$ (nahezu Totalreflexion) und

$$\delta(\omega) = \frac{c}{\omega n''(\omega)} \approx \frac{c}{\omega_p} \sqrt{\frac{2}{\omega\tau}}, \quad (3.56)$$

wobei δ die sogenannte Skineindringtiefe ist, die bei steigender Frequenz mit $\omega^{-1/2}$ abnimmt.

(2) Wir entwickeln in $(\omega\tau)^{-1} \ll 1$ und erhalten

$$\epsilon(\omega) \approx -\frac{\omega_p^2}{\omega^2} + i \frac{\omega_p^2}{\omega^3 \tau}. \quad (3.57)$$

Für den komplexen Brechungsindex ergibt sich somit

$$n'(\omega) \approx \frac{\omega_p \tau}{2\omega^2 \tau^2} \quad \text{und} \quad n''(\omega) \approx \frac{\omega_p}{\omega}. \quad (3.58)$$

Es gilt daher $n'' \gg n' \gg 1$, so dass das Metall auch in diesem Frequenzbereich totalreflektierend ist. Die Eindringtiefe wird mit $\delta(\omega) \approx c/\omega_p$ frequenzunabhängig. Dies ist die sogenannte Debye-Länge.

(3) Schliesslich gilt in diesem Bereich

$$\epsilon(\omega) \approx 1 - \frac{\omega_p^2}{\omega^2} \Rightarrow n' = \sqrt{1 - \frac{\omega_p^2}{\omega^2}}, n'' \approx 0, \quad (3.59)$$

woraus folgt, dass

$$R \approx \frac{\omega_p^2}{4\omega^2} \rightarrow 0 \quad \text{und} \quad \delta \rightarrow \infty. \quad (3.60)$$

Das bedeutet, dass in diesem Frequenzbereich das Metall transparent wird.

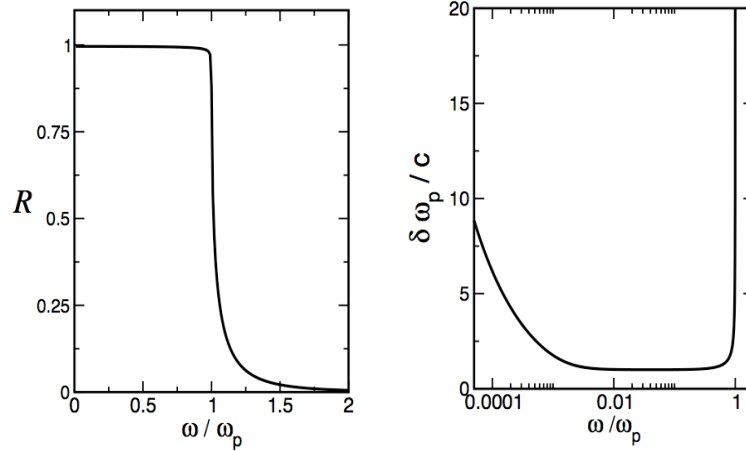


Abb. 3.11: Reflektivität und Eindringtiefe (Parameterwahl $\omega_p \bar{\tau} = 500$).

In dieser Diskussion haben wir den Beitrag vernachlässigt, der von der Polarisierbarkeit der Ionenrümpfe (Kern und Rumpfelektronen) im Metall herrührt. Da die Eigenfrequenzen ω_0 der polarisierten Ionenrümpfe sehr viel höher als die Plasmafrequenz liegen, ergibt sich ein Beitrag, der im für uns interessante Bereich im wesentlichen frequenzunabhängig ist. Unter Verwendung von (3.35) erhalten wir daher,

$$\epsilon(\omega) = 1 + \frac{4\pi i\sigma(\omega)}{\omega} + \frac{e^2 n_{ion}}{m(\omega_0^2 - \omega^2)} \approx \epsilon_\infty + \frac{4\pi i\sigma(\omega)}{\omega}. \quad (3.61)$$

Dies führt auf eine Erniedrigung der Plasmafrequenz, wenn wir den Bereich $\omega \sim \omega_p$ betrachten,

$$\epsilon(\omega) = \epsilon_\infty - \frac{\omega_p^2}{\omega^2} \Rightarrow \omega'_p = \frac{\omega_p}{\sqrt{\epsilon_\infty}}. \quad (3.62)$$

Somit wird auch die Reflektivität bei Frequenzen $\omega > \omega_p$ den Grenzwert,

$$R = \frac{(\epsilon_\infty - 1)^2}{(\epsilon_\infty + 1)^2} < 1, \quad (3.63)$$

annehmen.

Dieses relativ einfache, klassische Modell beschreibt die optischen Eigenschaften von Metallen sehr gut im niederfrequenten Bereich, wo die Reflektivität praktisch 100 % ist. Gemäss diesem Modell bleibt dies so bis zur Plasmafrequenz. In den meisten Metallen liegt die Plasmafrequenz mit $\hbar\omega_p \sim 10$ eV weit oberhalb des Bereichs des sichtbaren Lichtes (sichtbarer Bereich: $\hbar\omega \sim 1.5 - 3.5$ eV). Daher erscheinen Metalle glänzend. Manche Metalle haben aber Farben, wie etwa das rötliche Kupfer oder das gelbliche Gold. Diese Farben deuten an, dass die Reflektivität nicht für alle Frequenz bei 100% liegt. Die Abnahme der Reflektivität hat meist etwas mit elektronischen Übergängen zu tun, d.h., mit quantenmechanischen Prozessen. Sogenannte Interband-Übergänge spielen dabei eine wichtige Rolle in Edelmetallen wie Cu, Ag und Au. In Cu finden solche Absorptionen für Energien oberhalb 2.5 eV durch Quantenübergänge zwischen lokalisierten 3d-Elektronzuständen und 4s-Leitungselektronzuständen statt. Da 2.5 eV im sichtbar Frequenzbereich liegt, werden die Farben grün, blau und violett absorbiert und nur der rote Teil reflektiert. In den Figuren 3.12 sehen wir die Reflektivitäten von Ag und Cu, die stark durch Interband-Übergänge strukturiert sind. Dagegen entspricht die Reflektivität von Al in Fig. 3.12 mehr dem einfachen, klassischen Bild, obwohl auch hier Interband-Übergänge zu einer Verringerung der Reflektivität führen.

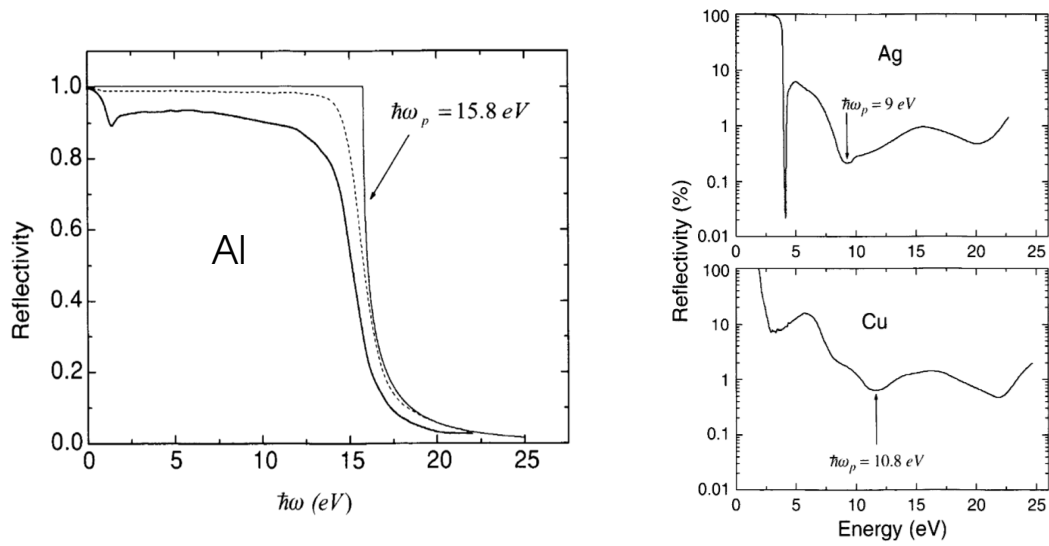


Abb. 3.12: Frequenzabhängige Reflektivität der Metalle Al, Ag und Cu. Beachte die logarithmischen Skalen bei der Reflektivität von Ag und Cu. (Quelle: *An introduction to the optical spectroscopy of inorganic solids*, J. García Solé, L.E. Bausá and D. Jaque, Wiley (2005))

Kapitel 4

Licht und Quantenmechanik

Die Maxwell'sche Theorie hat in der zweiten Hälfte des 19. Jahrhunderts ein in sich geschlossenes Verständnis des Lichtes und der gesamten elektromagnetischen Strahlung gegeben. Es gab jedoch einige offene Punkte, die vor allem mit der Wechselwirkung des Lichtes mit Materie zu tun hatten. Dazu gehörten der Photoelektrische Effekt, der von Hertz 1887 entdeckt wurde und die Schwarzkörper- oder Hohlraumstrahlung. Beide Phänomene gaben entscheidende Anstöße zur Entwicklung der Quantenmechanik.

4.1 Planck'sche Strahlungsgesetz

Schwarzkörper- oder Hohlraumstrahlung wurde bereits 1859 von G. Kirchhoff untersucht. Ein idealer schwarzer Körper absorbiert alle Strahlung, die auf ihn trifft. Eine solche Situation kann am besten durch einen Hohlraum mit einem kleinen Loch realisiert werden. Strahlung, die durch das Loch in den Hohlraum eindringt, wird nur mit sehr geringer Wahrscheinlichkeit wieder durch das Loch entkommen. Die Wände dieses Hohlraums werden auf einer gewissen Temperatur gehalten und absorbieren/emittieren elektromagnetische Strahlung auf allen Frequenzen.

Wir betrachten einen Hohlraum kubischer Form der Kantenlänge L . Die zulässigen stehenden Wellen in diesem Hohlraum besitzen die Wellenvektoren

$$\vec{k} = \frac{\pi}{L}(n_x, n_y, n_z) \quad \text{und} \quad \omega_{\vec{k}} = c|\vec{k}|, \quad (4.1)$$

wobei wir Knotenpunkte an den Wänden annehmen und $n_x, n_y, n_z \geq 0$. Die Anzahl der Moden mit Frequenzen kleiner als ω ist gegeben durch

$$N_\omega = 2 \sum_{\vec{n}} \Theta(\omega - \omega_{\vec{k}}) = \frac{1}{8} \frac{2L^3}{\pi^3} \int \Theta(\omega - \omega_{\vec{k}}) d^3k = \frac{L^3}{\pi^2 c^3} \int_0^\omega \omega'^2 d\omega' = \frac{L^3}{3\pi^2 c^3} \omega^3. \quad (4.2)$$

Durch den Faktor 2 werden beide möglichen Polarisierungen berücksichtigt. Daraus folgt die Zustandsdichte für ω

$$\rho(\omega) = \frac{1}{L^3} \frac{dN_\omega}{d\omega} = \frac{1}{\pi^2 c^3} \omega^2. \quad (4.3)$$

Die Energiedichte können wir somit schreiben als

$$u(\omega, T) = \bar{\varepsilon}(\omega, T) \rho(\omega), \quad (4.4)$$

wobei $\bar{\varepsilon}(\omega, T)$ die mittlere Energie eines Oszillators der Frequenz ω bei der Temperatur T bezeichnet. In der klassischen Statistischen Physik ist diese für alle Frequenzen gleich und bestimmt durch das Äquipartitionsgesetz,

$$\bar{\varepsilon}(\omega, T) = k_B T. \quad (4.5)$$

Diese Annahme führt zum Rayleigh-Jeans-Gesetz, welches die unphysikalische Konsequenz hat, dass die Gesamtenergie der Strahlung divergiert (Ultraviolett-Katastrophe),

$$U(T) = \int_0^\infty u(\omega, T) d\omega = \int_0^\infty \frac{k_B T}{\pi^2 c^3} \omega^2 d\omega \rightarrow \infty . \quad (4.6)$$

Aus Experimenten war um die Jahrhundertwende das Wien'sche Gesetz (1896) für grosse Frequenzen bekannt,

$$u(\omega, T) = \frac{\hbar \omega^3}{\pi^2 c^3} e^{-\hbar \omega / k_B T} \quad \Rightarrow \quad \bar{\varepsilon}(\omega, T) = \hbar \omega e^{-\hbar \omega / k_B T} , \quad (4.7)$$

wobei hier die "modernere" Schreibweise mit $\hbar = h/2\pi$ ($h = 6.661 \times 10^{-34}$ Js als Planck'sches Wirkungsquantum) und $k_B = 1.3807 \times 10^{-23}$ JK⁻¹ benutzt wurde.

Um 1900 hat Planck die Interpolation zwischen den beiden Grenzfällen (Rayleigh-Jeans und Wien) gefunden. Zunächst betrachten wir die Formulierung nach dem heutigen Verständnis. Die Annahme hinter der Planck'schen Interpolation ist das Konzept, dass Strahlung von der Materie in Quanten der Grösse $\hbar \omega$ für gegebene Frequenz emittiert/absorbiert wird. Damit erhält die mittlere Energie $\bar{\varepsilon}(\omega, T)$ die neue Form

$$\bar{\varepsilon}(\omega, T) = \frac{\sum_{n=0}^\infty \hbar \omega n e^{-\hbar \omega n / k_B T}}{\sum_{n=0}^\infty e^{-\hbar \omega n / k_B T}} = \frac{\hbar \omega}{e^{\hbar \omega / k_B T} - 1} , \quad (4.8)$$

die im niederfrequenten Bereich ins Äquipartitionsverhalten und im Hochfrequenzbereich in die Wien'sche Form übergeht.¹

Plancks frühe Überlegungen sind jedoch nicht diesem Weg gefolgt, sondern haben sich mit der Entropie beschäftigt. Aus der Thermodynamik ist folgende Beziehung zwischen der Entropie S , der inneren Energie U und der Temperatur T bekannt,

$$\frac{dS}{dU} = \frac{1}{T} . \quad (4.13)$$

Beschränken wir uns auf eine einzelne Mode der Frequenz ω , so können wir das Wien-Verhalten (4.7) verwenden, um einen Ausdruck für $1/T$ zu erhalten,

$$\frac{d\bar{s}}{d\bar{\varepsilon}} = \frac{1}{T} = -\frac{k_B}{\hbar \omega} \ln \left(\frac{\bar{\varepsilon}}{\hbar \omega} \right) \quad \Rightarrow \quad \frac{d^2 \bar{s}}{d\bar{\varepsilon}^2} = -\frac{k_B}{\hbar \omega} \frac{1}{\bar{\varepsilon}} . \quad (4.14)$$

¹Diese Form des Energiespektrums ergibt für die innere Energie der gesamten Strahlung

$$U(T) = \int_0^\infty u(\omega, T) d\omega = \frac{\pi^2 (k_B T)^4}{15 (\hbar c)^3} \propto T^4 , \quad (4.9)$$

das Stefan-Boltzmann-Gesetz. Das T^4 -Verhalten kann aus klassischen Überlegungen hergeleitet werden und ist kein Quanteneffekt. Um das zu verstehen, betrachten wir die Energiedichte, $u = (|\vec{E}|^2 + |\vec{B}|^2)/8\pi = |\vec{E}|^2/4\pi$, und die Impulsflussdichte, $|\vec{P}| = |\vec{S}| = |\vec{E} \times \vec{B}|/4\pi = |\vec{E}|^2/4\pi$, des Strahlungsfeldes im Vakuum (mit $c \equiv 1$). Für ein isotropes, homogenes (inkohärentes) Strahlungsfeld entspricht der Impulsübertrag auf eine Wand (pro Zeit- und Flächeneinheit) einem Drittel der gesamten (in alle Richtungen auftretenden) Impulsflussdichte. Da der Impulsübertrag den Druck p auf die Wand darstellt, können wir die allgemeine Beziehung ableiten,

$$p = \frac{1}{3} u \quad \text{oder} \quad pV = \frac{1}{3} U , \quad (4.10)$$

wobei wir benutzt haben, dass $|\vec{P}| = u$ und $U = uV$. Dies verwenden wir nun zusammen mit einer Beziehung aus der Thermodynamik,

$$T \left(\frac{\partial p}{\partial T} \right)_V - p = \left(\frac{\partial U}{\partial V} \right)_T = 3p = u \quad \Rightarrow \quad u = \frac{T}{3} \frac{du}{dT} - \frac{1}{3} u \quad \Rightarrow \quad \frac{du}{u} = 4 \frac{dT}{T} , \quad (4.11)$$

woraus unmittelbar folgt, dass

$$u = CT^4 . \quad (4.12)$$

Beachte, dass der Wert der Konstante C klassisch nicht festgelegt werden kann.

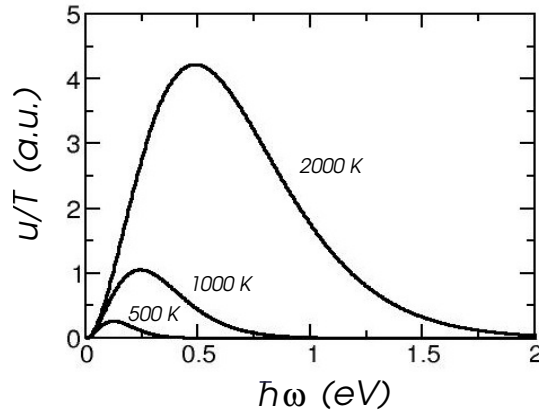


Abb.4.1: Planck'sche Strahlungsverteilungsfunktion: $u(\omega, T)/T$. Die Normulierung ergibt für kleine Frequenzen Übereinstimmung aller Kurven (Rayleigh-Jeans-Verhalten).

Diese Form gilt aber nur für kleine $\bar{\varepsilon}$ (hohe Frequenzen) bei gegebener Temperatur. Für höhere Energien ergeben sich Abweichungen, welche sich dem Gesetz von Rayleigh-Jeans annähern. Für dieses findet man

$$\frac{d\bar{s}}{d\bar{\varepsilon}} = \frac{1}{T} = \frac{k_B}{\bar{\varepsilon}} \quad \Rightarrow \quad \frac{d^2\bar{s}}{d\bar{\varepsilon}^2} = -\frac{k_B}{\bar{\varepsilon}^2}. \quad (4.15)$$

In einem "Akt der Verzweiflung" hat Planck dann zwischen (4.15) und (4.14) interpoliert,

$$\frac{d^2\bar{s}}{d\bar{\varepsilon}^2} = -\frac{k_B}{\bar{\varepsilon}(\hbar\omega + \bar{\varepsilon})}. \quad (4.16)$$

Durch Integrieren erhält man nun

$$\frac{1}{T} = \frac{d\bar{s}}{d\bar{\varepsilon}} = \frac{k_B}{\hbar\omega} \ln\left(\frac{\hbar\omega + \bar{\varepsilon}}{\bar{\varepsilon}}\right) \quad \Rightarrow \quad \bar{\varepsilon}(\omega, T) = \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1}, \quad (4.17)$$

was wieder (4.8) entspricht. Diese Form der Entropie konnte Planck später durch ein statistisches Argument begründen. Wir betrachten N identische Oszillatoren mit der mittleren Energie $\bar{\varepsilon}$ und der mittleren Entropie \bar{s} , so dass die Gesamtenergie und die Gesamtentropie gegeben sind durch

$$E_N = N\bar{\varepsilon} \quad \text{und} \quad S_N = N\bar{s}. \quad (4.18)$$

Zudem sei die Energie E_N durch P Energiequanten gegeben, d.h., $E_N = P\Delta$. Die Anzahl Möglichkeiten die P Quanten auf die N Oszillatoren zu verteilen, kann man sich auf folgende Weise überlegen: Insgesamt gibt es $P+N-1$ Objekte, wobei $N-1$ die Anzahl der Abgrenzungen zwischen den N Oszillatoren darstellt. Diese Objekte können auf $(P+N-1)!$ Arten angeordnet werden (siehe Abb.4.2). Da man aber weder die Abgrenzungen noch die Energiequanten untereinander unterscheiden kann, muss man noch durch die Permutationen der Abgrenzungen und der Quanten teilen um Doppelzählungen zu vermeiden. Somit ergibt sich die Anzahl Möglichkeiten zu

$$W = \frac{(P+N-1)!}{(N-1)!P!}. \quad (4.19)$$



Abb.4.2: Die Planck'sche Zählweise entspricht der Zählung aller (unterscheidbaren) Konfigurationen von $N-1$ Strichen und P Punkten. Diese ergibt mehr Konfigurationen als man durch eine klassische Betrachtungsweise erhalten würde.

Wenn man hingegen in der traditionellen Art zählt, kann jedes Quantum in einen beliebigen Oszillator gesteckt werden, was auf N^P Anordnungen führt (das erste Quantum hat N Möglichkeit, das zweite auch, usw.). Durch die Permutation der Quanten dividiert erhält man dann

$$W_{\text{trad}} = \frac{N^P}{P!} . \quad (4.20)$$

Betrachten wir zunächst die Planck'sche Zählweise und bestimmen die Entropie unter der Annahme, dass N und P sehr gross sind,

$$\begin{aligned} S_N &= k_B \ln W \approx k_B \{ \ln(N+P)! - \ln N! - \ln P! \} \\ &= k_B \{ (N+P) \ln(N+P) - N \ln N - P \ln P \} , \end{aligned} \quad (4.21)$$

wobei wir Stirling-Formel verwendet haben, $\ln N! \approx N(\ln N - 1)$. Mit $S_N = N\bar{s}$, $P/N = \bar{\varepsilon}/\Delta$ erhalten wir

$$\begin{aligned} \bar{s} &= k_B \left\{ \left(1 + \frac{\bar{\varepsilon}}{\Delta}\right) \ln \left(N \left(1 + \frac{\bar{\varepsilon}}{\Delta}\right) \right) - \ln N - \frac{\bar{\varepsilon}}{\Delta} \ln \left(N \frac{\bar{\varepsilon}}{\Delta} \right) \right\} \\ &= k_B \left\{ \left(1 + \frac{\bar{\varepsilon}}{\Delta}\right) \ln \left(1 + \frac{\bar{\varepsilon}}{\Delta}\right) - \frac{\bar{\varepsilon}}{\Delta} \ln \frac{\bar{\varepsilon}}{\Delta} \right\} , \end{aligned} \quad (4.22)$$

woraus folgt

$$\frac{d^2 \bar{s}}{d\bar{\varepsilon}^2} = -\frac{k_B}{\bar{\varepsilon}(\Delta + \bar{\varepsilon})} . \quad (4.23)$$

Mit dem Energiequant $\Delta = \hbar\omega$ stimmt dieser Ausdruck mit der Interpolation (4.16) überein. Wenn wir hingegen W_{trad} verwenden, erhalten wir wiederum das Resultat gemäss dem Wien'schen Gesetz,

$$\frac{d^2 \bar{s}}{d\bar{\varepsilon}^2} = -\frac{k_B}{\hbar\omega\bar{\varepsilon}} . \quad (4.24)$$

Die Planck'sche Statistik entspricht dem, was später als Bose-Einstein-Statistik der Bosonen formuliert werden wird. Damals war es aber noch nicht klar, wie dies verstanden werden sollte.

4.2 Einsteins Lichtquanten-Hypothese

Einsteins Lichtquanten-Hypothese aus dem Jahr 1905 gab Plancks Strahlungsgesetz die gültige Interpretation. Das Argument, das Einstein in seinem berühmten Artikel anbrachte, basierte auf einer Entropie-Betrachtung, die er aus dem Wien'schen Gesetz ableitete. Das Konzept, das sich daraus ergibt, kann dann auch auf den photoelektrischen Effekt angewendet werden.

4.2.1 Photonengas in Hohlraumstrahlung

Wir betrachten wieder die Hohlraumstrahlung, für die folgende Definitionen der Energiedichte und der Entropiedichte bei gegebener Frequenz und Temperatur gelten,

$$u(\omega, T) = \frac{\omega^2}{\pi^2 c^3} \bar{\varepsilon}(\omega, T) \quad \text{und} \quad \sigma(\omega, T) = \frac{\omega^2}{\pi^2 c^3} \bar{s}(\omega, T) . \quad (4.25)$$

Daraus folgt mit dem Wien'schen Gesetz,

$$\frac{d\sigma}{du} = \frac{1}{T} = -\frac{k_B}{\hbar\omega} \ln \frac{\bar{\varepsilon}}{\hbar\omega} \quad \Rightarrow \quad \sigma(\omega, T) = -\frac{k_B u}{\hbar\omega} \left(\ln \frac{\bar{\varepsilon}}{\hbar\omega} - 1 \right) . \quad (4.26)$$

Die Entropie und die Gesamtenergie für den Bereich $[\omega, \omega + \Delta\omega]$ ist gegeben durch

$$\Delta S = \sigma V \Delta\omega \quad \text{bzw.} \quad \Delta E = u V \Delta\omega , \quad (4.27)$$

so dass

$$\Delta S = -\frac{k_B \Delta E}{\hbar \omega} \left(\ln \frac{\bar{\varepsilon}}{\hbar \omega} - 1 \right) = -\frac{k_B \Delta E}{\hbar \omega} \left[\ln \frac{\pi^2 c^3 \Delta E}{\hbar \omega^3 V \Delta \omega} - 1 \right]. \quad (4.28)$$

Die Entropie zum (Referenz-)Volumen V_0 im gleichen Frequenzbereich sei ΔS_0 . Dann können wir ΔS schreiben als

$$\Delta S - \Delta S_0 = k_B \frac{\Delta E}{\hbar \omega} \ln \frac{V}{V_0} = k_B \ln \left(\frac{V}{V_0} \right)^{\Delta E / \hbar \omega}. \quad (4.29)$$

Dies ist analog zur Entropie eines idealen Gases von N Teilchen,

$$S(V, T) - S(V_0, T) = N k_B \ln \frac{V}{V_0} = k_B \ln \left(\frac{V}{V_0} \right)^N. \quad (4.30)$$

Demzufolge beschreibt der Wien'sche Grenzfall die Situation eines idealen Gases von $\Delta E / \hbar \omega$ Teilchen (eines verdünnten Gases von unabhängigen *Photonen*).

4.2.2 Photoelektrischer Effekt

Der photoelektrische Effekt wurde von Heinrich Hertz im Jahre 1887 entdeckt und später von seinem Studenten Philipp Lenard genauer untersucht. Er wird ebenfalls in Einsteins Artikel von 1905 behandelt, wenn auch am Rande, denn Lenards Experimente waren eher von qualitativer Natur. Robert Millikan verifizierte einige Jahre später Einsteins Voraussagen für den photoelektrischen Effekt quantitativ. Wichtige Voraussetzungen waren in diesem Zusammenhang die Entdeckung des Elektrons durch J.J. Thompson in Jahre 1897 und die Bestimmung der Elektronenladung durch Millikan.

Das Phänomen, dass Elektronen durch Beleuchtung mit Licht aus einem Material befreit werden können, hat einige Aspekte, die mit dem klassischen Verständnis des Lichtes nicht vereinbar sind. Zunächst wenden wir uns den experimentellen Fakten zu: Wird ein Material mit Licht bestrahlt, können Elektronen freigesetzt werden. Dabei bestimmt die Frequenz des Lichtes die Energie dieser Elektronen gemäss

$$E_e = \hbar \omega - W, \quad (4.31)$$

mit der (materialabhängigen) Austrittsarbeit W , und die Anzahl der Elektronen hängt von der Intensität des Lichtes ab. Mit Hilfe der Photonen-Hypothese lassen sich alle diese Befunde erklären: Licht wird in Quanten der Energie $\hbar \omega$ absorbiert, was zu einem Energieübertrag von $\hbar \omega$ führt. Die Lichtintensität bestimmt den Fluss der Photonen und somit auch die Anzahl der freigesetzten Elektronen.

Im folgenden wollen wir die Energie der elektromagnetischen Strahlung im Wellenbild betrachten und sie mit derjenigen des Photonbildes vergleichen. Wir verwenden von nun an das Vektorpotential \vec{A} , um die elektromagnetischen Wellen zu beschreiben. In der Coulomb-Eichung $\vec{\nabla} \cdot \vec{A} = 0$ gilt

$$\left. \begin{aligned} \vec{E} &= -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} \\ \vec{B} &= \vec{\nabla} \times \vec{A} \end{aligned} \right\} \Rightarrow \quad \square \vec{A} = \left\{ \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \right\} \vec{A} = 0. \quad (4.32)$$

Die allgemeine Lösung von dieser Gleichung kann man schreiben als eine Überlagerung von ebenen Wellen,

$$\vec{A}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}, \lambda} \left[A_{\vec{k}\lambda} \vec{e}_{\vec{k}\lambda} e^{i\vec{k} \cdot \vec{r} - i\omega t} + A_{\vec{k}\lambda}^* \vec{e}_{\vec{k}\lambda}^* e^{-i\vec{k} \cdot \vec{r} + i\omega t} \right], \quad (4.33)$$

wobei hier ein kubischer Kasten mit Volumen $V = L^3$ und periodische Randbedingungen angenommen wird und $\lambda = 1, 2$ für die beiden Polarisationsrichtungen steht, welche durch den komplexen Einheitsvektor $\vec{e}_{\vec{k}\lambda}$ beschrieben werden. Die transversale Natur der elektromagnetischen

Strahlung kommt auch aus der Eichung hervor: $0 = \vec{\nabla} \cdot \vec{A} \Rightarrow i\vec{k} \cdot \vec{e}_{\vec{k}\lambda}$. Ferner gilt $\omega = \omega_{\vec{k}} = c|\vec{k}|$. Somit folgt mit (4.32)

$$\vec{E}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}, \lambda} \frac{i\omega_{\vec{k}}}{c} \left[A_{\vec{k}\lambda} \vec{e}_{\vec{k}\lambda} e^{i\vec{k} \cdot \vec{r} - i\omega t} - A_{\vec{k}\lambda}^* \vec{e}_{\vec{k}\lambda}^* e^{-i\vec{k} \cdot \vec{r} + i\omega t} \right], \quad (4.34)$$

$$\vec{B}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}, \lambda} \left[(i\vec{k} \times \vec{e}_{\vec{k}\lambda}) A_{\vec{k}\lambda} e^{i\vec{k} \cdot \vec{r} - i\omega t} - (i\vec{k} \times \vec{e}_{\vec{k}\lambda}^*) A_{\vec{k}\lambda}^* e^{-i\vec{k} \cdot \vec{r} + i\omega t} \right]. \quad (4.35)$$

Die Energie und den Energiefluss (Betrag des Poynting-Vektors) der Strahlung im Volumen V erhalten wir aus

$$E = \int_V d^3r \frac{\vec{E}^2 + \vec{B}^2}{8\pi} = \int_V d^3r \frac{\vec{E}^2}{4\pi} = \sum_{\vec{k}, \lambda} \frac{\omega_{\vec{k}}^2}{2\pi c^2} |A_{\vec{k}\lambda}|^2, \quad (4.36)$$

$$\vec{S} = \int_V d^3r \frac{c}{4\pi} \vec{E} \times \vec{B} = \sum_{\vec{k}, \lambda} \frac{\omega_{\vec{k}} \vec{k}}{2\pi} |A_{\vec{k}\lambda}|^2. \quad (4.37)$$

Der Impuls der elektromagnetischen Strahlung² ist $\vec{P} = \vec{S}/c^2$. Wenn wir nun die Energie einer Mode (\vec{k}, λ) gegeben in (4.36) mit derjenigen von $N_{\vec{k}\lambda}$ Photonen vergleichen, erhalten wir einen Ausdruck für den Impuls $\vec{P}_{\vec{k}\lambda}$,

$$\frac{\omega_{\vec{k}}^2}{2\pi c^2} |A_{\vec{k}\lambda}|^2 = N_{\vec{k}\lambda} \hbar \omega_{\vec{k}} \quad \Longrightarrow \quad |A_{\vec{k}\lambda}|^2 = \frac{2\pi c^2 \hbar}{\omega_{\vec{k}}} N_{\vec{k}\lambda} \quad (4.39)$$

$$\Longrightarrow \quad \vec{P}_{\vec{k}\lambda} = \frac{\omega_{\vec{k}} \vec{k}}{2\pi c^2} |A_{\vec{k}\lambda}|^2 = N_{\vec{k}\lambda} \hbar \vec{k}. \quad (4.40)$$

Dies legt nahe, dass der Impuls des Photons durch $\hbar \vec{k}$ gegeben ist. Damit ist das Photon als Teilchen charakterisiert. Diese Tatsache wurde lange nicht akzeptiert und erhielt erst durch das Compton-Experiment eine solide experimentelle Basis.

4.3 Strahlungsabsorption und -emission

Wie Elektronen an die elektromagnetische Strahlung koppeln, kann in der klassischen Theorie (wie sie von Drude eingeführt wurde) nicht beschrieben werden. Hierfür liefert erst die Quantenmechanik das richtige Werkzeug.

Zunächst wollen wir Quantenübergänge in Atomen studieren. Wir halten uns dabei an die Sprache der "alten Quantenmechanik", wie sie von Niels Bohr 1913 begründet wurde und später von Arnold Sommerfeld und Paul Ehrenfest für allgemeine Einteilchensysteme mit periodischen Bahnen erweitert wurde. Diese Theorie beschreibt die Quantisierung der Energieniveaus, zwischen denen Quantenübergänge unter Absorption oder Emission eines Lichtquants möglich ist. Die Photonenenergie $\hbar\omega$ entspricht dabei gerade der Energiedifferenz der beiden Niveaus,

$$E_{n'} - E_n = \hbar\omega. \quad (4.41)$$

Die Grundlage der Theorie ist die Quantisierung der Wirkung der periodischen Bahn,

$$I = \frac{1}{2\pi} \oint \vec{p} \cdot d\vec{q} = \hbar(\nu + \alpha), \quad (4.42)$$

²Der Impuls des Strahlungsfeldes wird über die Gleichung

$$\frac{d}{dt} \vec{P} = \int_V d^3r \left\{ \rho \vec{E} + \frac{1}{c} \vec{j} \times \vec{B} \right\} \quad (4.38)$$

bestimmt, indem man die Maxwell-Gleichungen verwendet.

wobei ν eine ganze Zahl und α eine experimentell zu bestimmende Konstante ist. Beachte die Verbindung mit dem Prinzip von Maupertuis.

Die Beziehung zwischen den Strahlungsübergängen und dem Planck'schen Strahlungsgesetz wurde von Einstein 1917 herausgearbeitet. Er leitete die Planck'sche Formel für die Energiedichte der Hohlraumstrahlung her unter der Annahme, dass die Boltzmann-Verteilung der Atome im Hohlraum durch die Emission und Absorption von Strahlung nicht modifiziert wird. Dies wollen wir hier nachvollziehen und betrachten Atome, die sich in den Zuständen n und n' mit den entsprechenden Energie $E_n < E_{n'}$ befinden können. Im thermodynamischen Gleichgewicht gilt für die Besetzung der Atomzustände

$$p_n = p e^{-E_n/k_B T} \quad \text{und} \quad p_{n'} = p e^{-E_{n'}/k_B T} \quad \text{mit} \quad p = N (e^{-E_n/k_B T} + e^{-E_{n'}/k_B T})^{-1}, \quad (4.43)$$

wobei N die Gesamtzahl der Atome ist. Ein Atom im Zustand n' zerfällt mit einer gewissen Wahrscheinlichkeit pro Zeiteinheit, $A_{n'}^n$, unter Abgabe eines Photons der Energie $\hbar\omega = E_{n'} - E_n$ spontan, was wir für viele Atome durch die Ratengleichung,

$$\Gamma_{n' \rightarrow n} = -A_{n'}^n p_{n'}, \quad (4.44)$$

beschreiben können. Andererseits können Übergänge durch die Hohlraumstrahlung induziert werden, und zwar mit den Raten:

$$n \rightarrow n' : \quad B_n^{n'} u(\omega, T) \quad \text{und} \quad n' \rightarrow n : \quad B_{n'}^n u(\omega, T), \quad (4.45)$$

wobei $u(\omega, T)$ die Energiedichte des Strahlungsfeldes ist. Daraus folgen die Ratengleichungen

$$\Gamma_{n \rightarrow n'} = -p_n B_n^{n'} u(\omega, T), \quad (4.46)$$

$$\Gamma_{n' \rightarrow n} = -p_{n'} \{A_{n'}^n + B_{n'}^n u(\omega, T)\}.$$

Im detaillierten Gleichgewicht, $\Gamma_{n \rightarrow n'} = \Gamma_{n' \rightarrow n}$, gilt

$$p_n B_n^{n'} u(\omega, T) = p_{n'} \{A_{n'}^n + B_{n'}^n u(\omega, T)\}. \quad (4.47)$$

Nimmt $u(\omega, T)$ die Planck'sche Form mit $\hbar\omega = E_{n'} - E_n$ an, so folgt im Limes $k_B T \rightarrow \infty$ (i.e., $p_n = p_{n'}$ und $u(\omega, T) \rightarrow \infty$)

$$B = B_n^{n'} = B_{n'}^n \quad \text{und} \quad A = A_{n'}^n = \frac{\hbar\omega^3}{\pi^2 c^3} B. \quad (4.48)$$

Damit unterscheiden wir zwischen *induzierter Absorption und Emission* mit der identischen Rate $Bu(\omega, T)$ und der *spontanen Emission* mit der Rate A . Ferner gilt die einfache Beziehung (4.48) zwischen den verschiedenen Raten.

Einsteins Betrachtungen haben zum ersten Mal den Begriff der Übergangsrates ins Spiel gebracht. Es ist interessant zu sehen, dass man eine nicht-triviale Bestimmung der Übergangsrates innerhalb der Bohr-Sommerfeld-Betrachtungen finden kann; das Bohr'sche Korrespondenzprinzip besagt nämlich, dass für hohe Quantenzahlen der Energieunterschied $E_n - E_{n-1} = \hbar\omega_{n,n-1}$ zwischen benachbarten Energieniveaus gerade der klassischen Umlauffrequenz entspricht,

$$\omega_{n,n-1} = \frac{1}{\hbar} (E_n - E_{n-1}) = \frac{E_n - E_{n-1}}{I_n - I_{n-1}} = \frac{dE(I)}{dI} = \omega_{\text{kl}}. \quad (4.49)$$

In der klassischen Betrachtungsweise kann ein gebundenes Elektron Energie mittels Dipolstrahlung verlieren. Die mittlere Leistung in einen gewissen Raumwinkel $d\Omega$ ist gegeben durch

$$\frac{d^2 E}{dt d\Omega} = \frac{1}{4\pi c^3} \left| \frac{d^2 \vec{P}}{dt^2} \right|^2 \sin^2 \theta, \quad (4.50)$$

wobei $\vec{P}(t) = e\vec{r}(t)$ das Dipolmoment ist und θ der Winkel zwischen \vec{P} und der Abstrahlrichtung \vec{k} . Andererseits gibt die spontane Emission

$$\left. \frac{d^2 E}{dt d\Omega} \right|_{n \rightarrow n'} = \hbar \omega_{n'n} A_{n'}^n \frac{3 \sin^2 \theta}{8\pi}. \quad (4.51)$$

Unter Annahme einer periodischen Zeitabhängigkeit können wir das Dipolmoment als Fourierreihe darstellen,

$$\vec{P}(t) = \sum_{m=-\infty}^{+\infty} \vec{p}_m e^{im\omega_{kl}t} \quad \text{mit} \quad \vec{p}_{-m} = \vec{p}_m^*. \quad (4.52)$$

Der Vergleich ergibt

$$A_{n'}^n = 2 \frac{2\omega_{n'n}^3}{3\hbar c^3} |\vec{p}_{n'-n}|^2, \quad (4.53)$$

wobei wir $m = n' - n$ (bzw. $m = n - n' \rightarrow$ zusätzlicher Faktor 2) und $(n' - n)\omega_{kl} = \omega_{n'n}$ benutzt haben. Falls \vec{p}_m nicht verschwindet, wird ein Übergang als erlaubt deklariert. Betrachten wir nun einen einfachen harmonischen Oszillator in einer räumlichen Dimension mit

$$z(t) = z_0 \cos(\omega_{kl}t) \quad \Rightarrow \quad p_{\pm 1} = \frac{1}{2} e z_0 \quad \text{und} \quad p_m = 0 \quad \text{für} \quad m \neq \pm 1. \quad (4.54)$$

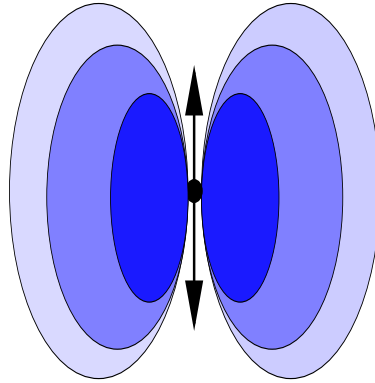


Abb.4.3: Dipolstrahlung eines Oszillators.

Mit der Korrespondenzbeziehung

$$E_n = \hbar \omega_{kl} n = \frac{m\omega_{kl}^2 z_0^2}{2} \quad \Rightarrow \quad z_0^2 = \frac{2\hbar}{m\omega_{kl}} n, \quad (4.55)$$

folgt

$$A_n^{n-1} = \frac{e^2 \omega_{kl}^3}{3\hbar c^3} z_0^2 = \frac{2e^2 \omega_{kl}^2}{3mc^3} n, \quad (4.56)$$

wodurch die Übergangsrate bestimmt ist. Dies legt fest, dass nur Übergänge möglich sind, bei welchen sich die Quantenzahl n um 1 verändert. Diese Auswahlregel gilt für die sogenannten Dipolübergänge, welche die wichtigsten Mechanismen für Übergänge darstellen. Über die Beziehungen (4.48) erhalten wir auch die Raten für induzierte Emission und Absorption.

4.4 Übergangsraten der modernen Quantenmechanik

Obwohl es interessant ist, den Gedankengängen der alten Quantenmechanik zu folgen, ist die moderne Formulierung praktischer und erlaubt allgemeinere Situationen zu behandeln. Im Folgenden werden die Kenntnisse der modernen Quantenmechanik vorausgesetzt und nur ergänzt, wo es notwendig erscheint.

Für die Analyse der Übergangswahrscheinlichkeiten zwischen verschiedenen Quantenzuständen ziehe wir die Goldene Regel heran. Das System sei beschrieben durch den Hamilton-Operator

$$\begin{aligned}\mathcal{H} &= \frac{1}{2m} \left(\widehat{\vec{p}} - \frac{e}{c} \vec{A} \right)^2 + V(\widehat{\vec{r}}) \\ &= \frac{\widehat{\vec{p}}^2}{2m} + V(\widehat{\vec{r}}) - \frac{e}{2mc} (\vec{A}(\widehat{\vec{r}}) \cdot \widehat{\vec{p}} + \widehat{\vec{p}} \cdot \vec{A}(\widehat{\vec{r}})) + \frac{e^2}{2mc^2} \vec{A}(\widehat{\vec{r}})^2 = \mathcal{H}_0 + \mathcal{H}_{\text{int}},\end{aligned}\tag{4.57}$$

wobei \mathcal{H}_0 das Quantensystem ohne äussere Störung beschreibt und \mathcal{H}_{int} die Kopplung an das Strahlungsfeld beinhaltet. Wir zerlegen die Kopplungsterme in eine *paramagnetischen* und einen *diamagnetischen* Teil,

$$\hat{V}_{\text{para}} = -\frac{e}{mc} \vec{A}(\vec{r}, t) \cdot \widehat{\vec{p}} \quad \text{und} \quad \hat{V}_{\text{dia}} = +\frac{e^2}{2mc^2} \vec{A}(\vec{r}, t)^2,\tag{4.58}$$

wobei wir benutzt haben, dass in der Coulomb-Eichung $\widehat{\vec{p}} \cdot \vec{A} = \vec{A} \cdot \widehat{\vec{p}}$ gilt. Das Spektrum des ungestörten Problems folgt aus der stationären Schrödingergleichung,

$$\mathcal{H}_0 |\phi_n\rangle = \varepsilon_n |\phi_n\rangle,\tag{4.59}$$

und kann diskret oder kontinuierlich sein. Ist dieses bekannt, so kann auch die Zustandsdichte (Anzahl Zustände im infinitesimalen Energieintervall $[E, E + dE]$),

$$\rho(E) = \sum_n \delta(E - \varepsilon_n),\tag{4.60}$$

berechnet werden.

4.4.1 Goldene Regel

Wir fassen das Strahlungsfeld als zeitabhängige (oszillierende) Störung auf, die man adiabatisch einschaltet,

$$V(t) = V_0 e^{\eta t} \cos(\omega t) \quad \text{mit} \quad \eta \rightarrow 0_+,\tag{4.61}$$

für $t < 0$. Im Folgenden werden wir den diamagnetischen Teil in \mathcal{H}_{int} vernachlässigen, da er quadratisch im störenden Strahlungsfeld ist.

Die Wahrscheinlichkeit, dass das Quantensystem durch die äussere Störung den Übergang vom Zustand $|\phi_n\rangle$ zum Zustand $|\phi_{n'}\rangle$ macht, wird durch die Goldene Regel nach Fermi und Wentzel bestimmt (die wir hier nicht herleiten),

$$\Gamma_{n \rightarrow n'} = \frac{2\pi}{\hbar} |\langle \phi_{n'} | \hat{V}_{\text{para}} | \phi_n \rangle|^2 \{ \delta(\varepsilon_{n'} - \varepsilon_n - \hbar\omega) + \delta(\varepsilon_{n'} - \varepsilon_n + \hbar\omega) \}.\tag{4.62}$$

Mit dieser Übergangsrate berücksichtigen wir sowohl den Absorptions- als auch den Emissionsprozess, abhängig davon ob $\varepsilon_{n'} > \varepsilon_n$ (Absorption) oder $\varepsilon_{n'} < \varepsilon_n$ (Emission). Dabei findet eine Energieänderung des Quantensystems um den Betrag $\hbar\omega$ statt.

Durch Summation (oder Integration bei kontinuierlichem Spektrum) über alle Endzustände n' , können wir auch die Übergangsrate des Zustandes $|\phi_n\rangle$ in einen beliebigen Zustand angeben,

$$\Gamma_n = \sum_{n'} \Gamma_{n \rightarrow n'} = \frac{2\pi}{\hbar} \sum_{n'} |\langle \phi_{n'} | \hat{V}_{\text{para}} | \phi_n \rangle|^2 \{ \delta(\varepsilon_{n'} - \varepsilon_n - \hbar\omega) + \delta(\varepsilon_{n'} - \varepsilon_n + \hbar\omega) \}.\tag{4.63}$$

Diese Rate kann durch die Zustandsdichte ausgedrückt werden, falls die Matrixelemente $|\langle \phi_{n'} | \hat{V}_{\text{para}} | \phi_n \rangle|^2 = |M|^2$ nicht von n' abhängen,

$$\Gamma_n = \frac{2\pi}{\hbar} |M|^2 \{ \rho(\varepsilon_n + \hbar\omega) + \rho(\varepsilon_n - \hbar\omega) \}.\tag{4.64}$$

4.4.2 Harmonischer Oszillator

Der harmonische Oszillator ist das ‐einfachste‐ quantenmechanische System und wird durch folgenden Hamilton-Operator beschrieben,

$$\mathcal{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{r}^2 = \sum_{\nu=x,y,z} \hbar\omega \left\{ \hat{a}_\nu^\dagger \hat{a}_\nu + \frac{1}{2} \right\} \quad \text{mit} \quad \begin{cases} \hat{r}_\nu = \frac{r_0}{\sqrt{2}} (\hat{a}_\nu^\dagger + \hat{a}_\nu) \\ \hat{p}_\nu = i \frac{m\omega r_0}{\sqrt{2}} (\hat{a}_\nu^\dagger - \hat{a}_\nu) \end{cases}, \quad (4.65)$$

und $r_0 = \sqrt{\hbar/m\omega}$. Die (Auf-/Absteige-)Operatoren $\hat{a}_\nu^{(\dagger)}$ erfllen die Heisenberg-Kommutationsrelationen,

$$[\hat{r}_\nu, \hat{p}_{\nu'}] = i\hbar\delta_{\nu\nu'} \quad \Leftrightarrow \quad [\hat{a}_\nu, \hat{a}_{\nu'}^\dagger] = \delta_{\nu\nu'} \quad \text{und} \quad [\hat{a}_\nu^\dagger, \hat{a}_{\nu'}^\dagger] = [\hat{a}_\nu, \hat{a}_{\nu'}] = 0. \quad (4.66)$$

Der Grundzustand $|\phi_0\rangle$ des eindimensionalen, harmonischen Oszillators ist definiert durch $\hat{a}_\nu|\phi_0\rangle = 0$. Fr die stationren Zustnde gilt

$$\mathcal{H}|\phi_n\rangle = \varepsilon_n|\phi_n\rangle = \hbar\omega \left(n + \frac{1}{2} \right) |\phi_n\rangle \quad \text{mit} \quad |\phi_n\rangle = \frac{1}{\sqrt{n}} \hat{a}^\dagger |\phi_{n-1}\rangle = \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |\phi_0\rangle. \quad (4.67)$$

Sie sind Eigenzustnde des Operators $\hat{a}^\dagger \hat{a}$, wie man durch mehrfaches Anwenden von (4.66) einfach zeigen kann. Die Verallgemeinerung auf drei Dimensionen ist trivial, da verschiedene Komponenten unabhngig sind.

Es gilt

$$\langle \phi_n | \hat{r} | \phi_n \rangle = \langle \phi_n | \hat{p} | \phi_n \rangle = 0, \quad (4.68)$$

was durch Symmetrieberlegungen nachvollzogen werden kann: der Eigenzustand zu den Quantenzahlen (n_x, n_y, n_z) besitzt die Paritt $(-1)^{n_x+n_y+n_z}$. Da \hat{r} und \hat{p} Tensoroperatoren vom Rang 1 sind, muss das Matrixelement zwischen zwei Zustnden gleicher Paritt verschwinden. Hingegen gilt

$$\langle \phi_{\vec{n}} | \hat{r}^2 | \phi_{\vec{n}} \rangle = r_0^2 \left(n_x + n_y + n_z + \frac{3}{2} \right). \quad (4.69)$$

Die klassische Bewegung des Teilchens wird durch einen *kohrenten Zustand* beschrieben, den Eigenzustand des Operators \hat{a} ,

$$\hat{a}|\psi_\alpha\rangle = \alpha|\psi_\alpha\rangle, \quad (4.70)$$

wobei wir uns wiederum auf eine Dimension beschrnkt haben. Die Projektion auf die stationren Zustnde $|\phi_n\rangle$ erlaubt die Darstellung

$$\langle \phi_n | \psi_\alpha \rangle = \frac{1}{\sqrt{n!}} \langle \phi_0 | \hat{a}^n | \psi_\alpha \rangle = \frac{\alpha^n}{\sqrt{n!}} \langle \phi_0 | \psi_\alpha \rangle \quad \Rightarrow \quad |\psi_\alpha\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |\phi_n\rangle. \quad (4.71)$$

Mit der zeitlichen Evolution der Zustnde ergibt sich fr die Dynamik,

$$|\psi_\alpha(t)\rangle = e^{-iHt/\hbar} |\psi_\alpha(0)\rangle = e^{-|\alpha|^2/2} e^{-i\omega t/2} \sum_n \frac{\alpha^n e^{-in\omega t}}{\sqrt{n!}} |\phi_n\rangle = e^{-i\omega t/2} |\psi_{\alpha(t)}\rangle, \quad (4.72)$$

mit $\alpha(t) = \alpha e^{-i\omega t}$. Damit erhalten wir fr den Erwartungswert

$$\begin{aligned} \langle \hat{x}(t) \rangle &= \langle \psi_{\alpha(t)} | \hat{x} | \psi_{\alpha(t)} \rangle = \frac{r_0}{\sqrt{2}} \langle \psi_{\alpha(t)} | (\hat{a} + \hat{a}^\dagger) | \psi_{\alpha(t)} \rangle = \frac{r_0}{\sqrt{2}} (\alpha(t) + \alpha(t)^*) \\ &= \sqrt{2} r_0 \cos(\omega t - \delta), \end{aligned} \quad (4.73)$$

wobei wir $\alpha = |\alpha| e^{i\delta}$ angenommen haben. Die Dynamik des kohrenten Zustandes ist mit dem Verhalten des klassischen Teilchens verknpft. Der Erwartungswert folgt den klassischen Bewegungsgleichung (Ehrenfest-Theorem),

$$\frac{d}{dt} \langle \hat{x} \rangle = \frac{1}{m} \langle \hat{p} \rangle \quad \text{und} \quad \frac{d}{dt} \langle \hat{p} \rangle = \langle F(\hat{x}) \rangle, \quad (4.74)$$

mit $F(x)$ als die Kraft. Beachte, dass im Allgemeinen $\langle F(\hat{x}) \rangle \neq F(\langle \hat{x} \rangle)$. Im Falle des harmonischen Oszillators ist jedoch die Gleichheit erfllt, denn $F(x) = -m\omega^2 x$.

4.4.3 Übergangsrate für den harmonischen Oszillator

Wir betrachten nun die Übergangsrate zwischen den Zuständen des harmonischen Oszillators, $|\phi_{\vec{n}}\rangle \rightarrow |\phi_{\vec{n}'}\rangle$. Das Strahlungsfeld sei gegeben durch die monochromatische elektromagnetische Welle (4.33) der Mode (\vec{k}, λ) , was uns auf folgendes Matrixelement führt,

$$\begin{aligned} -\frac{e}{mc} \langle \phi_{\vec{n}'} | \vec{A}(\vec{r}, t) \cdot \vec{p} | \phi_{\vec{n}} \rangle &= -\frac{e}{mc} \frac{1}{\sqrt{V}} \langle \phi_{\vec{n}'} | (A_{\vec{k}\lambda} e^{i\vec{k}\cdot\vec{r} - i\omega t} \vec{e}_{\vec{k}\lambda} + c.c.) \cdot \vec{p} | \phi_{\vec{n}} \rangle \\ &\approx -\frac{e}{mc} \frac{1}{\sqrt{V}} \sum_{\nu=x,y,z} (A_{\vec{k}\lambda} e^{-i\omega t} e_{\vec{k}\lambda}^{\nu} + c.c.) \underbrace{\langle \phi_{\vec{n}'} | \hat{p}_{\nu} | \phi_{\vec{n}} \rangle}_{= -i \frac{m\omega r_0}{\sqrt{2}} (\sqrt{n_{\nu}} \delta_{n'_{\nu}, n_{\nu}-1} - \sqrt{n_{\nu}+1} \delta_{n'_{\nu}, n_{\nu}+1})} \end{aligned} \quad (4.75)$$

Dabei haben wir die Näherung gemacht, dass $e^{i\vec{k}\cdot\vec{r}} \approx 1$. Dies rechtfertigt sich, wenn wir kr_0 für genügend kleine n betrachten,

$$kr_0 = k \sqrt{\frac{\hbar}{m\omega}} = \sqrt{\frac{k\hbar}{mc}} = \sqrt{k\lambda_C} = \sqrt{\frac{2\pi\lambda_C}{\lambda}} \ll 1, \quad (4.76)$$

wobei $\lambda_C = \hbar/mc$ die Compton-Wellenlänge des Teilchens ist. Das Matrixelement (4.75) liefert uns die Auswahlregel, dass nur Übergänge zwischen benachbarten Energie-Niveaus, d.h., $n' = n \pm 1$, erlaubt sind.

Setzen wir den Ausdruck (4.75) in die Goldene Regel (4.62) ein, so finden wir für den einzigen erlaubten Emissions-Übergang $n_z \rightarrow n_z - 1$ und $\lambda = ||$ (Dipolstrahlung ist linear polarisiert; Polarisation in Schwingrichtung) die Rate

$$\begin{aligned} \Gamma_{\vec{n} \rightarrow \vec{n}'}(\vec{k}, ||) &= \frac{2\pi}{\hbar} \frac{e^2}{m^2 c^2 V} \frac{m^2 \omega^2 r_0^2}{2} |A_{\vec{k}||}|^2 2 |e_{\vec{k}||}^z|^2 n_z \delta(\varepsilon_{n-1} - \varepsilon_n + \hbar\omega) \\ &= \frac{2\pi^2 \hbar e^2}{m} \frac{N_{\vec{k},||}}{V} 2 |e_{\vec{k}||}^z|^2 n_z \delta(\varepsilon_{n-1} - \varepsilon_n + \hbar\omega), \end{aligned} \quad (4.77)$$

mit $\vec{n}' = (n_x, n_y, n_z - 1)$. Im letzten Schritt haben wir verwendet, dass $r_0^2 = \hbar/m\omega$ und $|A_{\vec{k}\lambda}|^2 = N_{\vec{k}\lambda} 2\pi\hbar c^2/\omega$. Zusätzlich wurden die schnell-oszillierenden Anteile $\propto \exp(\pm 2i\omega t)$ vernachlässigt. Beachte, dass die Winkelabhängigkeit der Emission durch $|e_{\vec{k}||}^z|^2 = \sin^2 \theta_k$ beschrieben wird, was einer Dipolstrahlung eines in z -Richtung orientierten Dipols entspricht. Für die Abstrahlrate in den Raumwinkel $d\Omega (= d\Omega_{\vec{k}})$ finden wir

$$\begin{aligned} d\Gamma &= \frac{1}{V} \sum_{\vec{k}, n', \lambda} \Gamma_{\vec{n} \rightarrow \vec{n}'}(\vec{k}, \lambda) \xrightarrow{V \rightarrow \infty} d\Omega_{\vec{k}} \frac{1}{(2\pi)^3} \int dk k^2 \Gamma_{\vec{n} \rightarrow \vec{n}'}(\vec{k}, ||) \\ &= d\Omega_{\vec{k}} \frac{1}{(2\pi)^3 c^3} \int \omega^2 d\omega \Gamma_{\vec{n} \rightarrow \vec{n}'}(\vec{k}, ||) \\ &= d\Omega_{\vec{k}} \frac{1}{(2\pi)^3} \int \omega^2 d\omega \frac{2\pi^2 \hbar e^2}{mc^3} \frac{N_{\vec{k}||}}{V} 2 |e_{\vec{k}||}^z|^2 n_z \delta(\varepsilon_{n-1} - \varepsilon_n + \hbar\omega) \\ &= \frac{e^2 \omega_{kl}^2}{4\pi m c^3} n_z \frac{N_{\vec{k}\lambda}}{V} \sin^2 \theta d\Omega_{\vec{k}} = A_n^{n-1} \frac{N_{\vec{k}\lambda}}{V} \frac{3}{8\pi} \sin^2 \theta d\Omega_{\vec{k}}, \end{aligned} \quad (4.78)$$

wobei wir $\omega_{kl} = (\varepsilon_n - \varepsilon_{n-1})/\hbar$ definieren. Es zeigt sich, dass dieser Ausdruck mit demjenigen für A_n^{n-1} der Bohr-Sommerfeld-Theorie übereinstimmt, bis auf die Tatsache, dass wir hier stimulierte und nicht spontane Emission beschreiben. Beachte auch, dass in Gleichung (4.56) die Übergangsrate für beide Polarisationen steht, während wir hier die Polarisationen unterscheiden.

4.5 Quantisierung des elektromagnetischen Strahlungsfeldes

Während spontane Emission innerhalb der ‘alten Quantenmechanik’ auf natürliche Weise aus der Dipolstrahlung eines geladenen Oszillators folgt, liefert die Goldene Regel nur stimulierte Emission, solange das Strahlungsfeld als klassische Störung betrachtet wird, dessen Energieinhalt keine Rolle spielt. Im Folgenden werden wir daher die elektromagnetische Strahlung ebenfalls in die quantenmechanische Beschreibung miteinbeziehen.

4.5.1 Elektromagnetische Strahlung als harmonischer Oszillator

Eine Basis der Lösungen der Maxwell’sche Wellengleichung stellen die ebene Wellen

$$\vec{A}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}, \lambda} \left\{ A_{\vec{k}\lambda}(t) \vec{e}_{\vec{k}\lambda} e^{i\vec{k}\cdot\vec{r}} + A_{\vec{k}\lambda}^*(t) \vec{e}_{\vec{k}\lambda}^* e^{-i\vec{k}\cdot\vec{r}} \right\}, \quad (4.79)$$

dar, wobei die $A_{\vec{k}\lambda}(t)$ ’s die Gleichung

$$\ddot{A}_{\vec{k}\lambda}(t) + \omega_{\vec{k}}^2 A_{\vec{k}\lambda}(t) = 0, \quad (4.80)$$

mit $\omega_{\vec{k}} = c|\vec{k}|$ erfüllen. Aus (4.36) erhalten wir für die Energie den Ausdruck

$$E_{SF} = \sum_{\vec{k}, \lambda} \frac{\omega_{\vec{k}}^2}{4\pi c^2} \left\{ A_{\vec{k}\lambda}(t) A_{\vec{k}\lambda}(t)^* + A_{\vec{k}\lambda}(t)^* A_{\vec{k}\lambda}(t) \right\} = \frac{1}{2} \sum_{\vec{k}, \lambda} \left\{ P_{\vec{k}\lambda}^2 + \omega_{\vec{k}}^2 Q_{\vec{k}\lambda}^2 \right\}, \quad (4.81)$$

mit der verallgemeinerten Koordinate Q und dem kanonischen Impuls P definiert durch

$$Q_{\vec{k}\lambda} = \frac{1}{\sqrt{4\pi c^2}} (A_{\vec{k}\lambda}(t) + A_{\vec{k}\lambda}(t)^*) \quad \text{und} \quad P_{\vec{k}\lambda} = \frac{d}{dt} Q_{\vec{k}\lambda} = -\frac{i\omega_{\vec{k}}}{\sqrt{4\pi c^2}} (A_{\vec{k}\lambda}(t) - A_{\vec{k}\lambda}(t)^*). \quad (4.82)$$

Jede Mode stellt separat einen harmonischen Oszillator dar, dessen Quantisierung wir bereits gesehen haben. Wir betrachten $Q_{\vec{k}\lambda}$, $P_{\vec{k}\lambda}$ als quantenmechanische Operatoren $\hat{Q}_{\vec{k}\lambda}$, $\hat{P}_{\vec{k}\lambda}$ und führen die Auf- und Absteige-Operatoren ein:

$$\hat{a}_{\vec{k}\lambda}^{(\dagger)} = \sqrt{\frac{\omega_{\vec{k}}}{2\pi\hbar c^2}} \hat{A}_{\vec{k}\lambda}^{(\dagger)} \quad \text{mit} \quad \begin{cases} [\hat{a}_{\vec{k}\lambda}, \hat{a}_{\vec{k}'\lambda'}^\dagger] = \delta_{\vec{k}, \vec{k}'} \delta_{\lambda, \lambda'} \\ [\hat{a}_{\vec{k}\lambda}^\dagger, \hat{a}_{\vec{k}'\lambda'}^\dagger] = [\hat{a}_{\vec{k}\lambda}^\dagger, \hat{a}_{\vec{k}'\lambda'}] = 0 \end{cases}. \quad (4.83)$$

Damit erhalten wir für das Strahlungsfeld einen Hamilton-Operator, der die Form eines harmonischen Oszillators für jede Mode hat,

$$\mathcal{H}_{SF} = \sum_{\vec{k}, \lambda} \hbar\omega_{\vec{k}} \left(\hat{a}_{\vec{k}\lambda}^\dagger \hat{a}_{\vec{k}\lambda} + \frac{1}{2} \right) = \sum_{\vec{k}, \lambda} \hbar\omega_{\vec{k}} \left(\hat{N}_{\vec{k}\lambda} + \frac{1}{2} \right), \quad (4.84)$$

wobei $\hat{N}_{\vec{k}\lambda}$ der Besetzungszahl-Operator ist und die Anzahl der Photonen in der Mode (\vec{k}, λ) angibt. Damit wird das Strahlungsfeld gegebener Photonenzahlen durch der Besetzungszahl-Zustände darstellbar,

$$|N_{\vec{k}_1\lambda_1}, N_{\vec{k}_2\lambda_2}, \dots\rangle = \frac{\dots (\hat{a}_{\vec{k}_2\lambda_2}^\dagger)^{N_{\vec{k}_2\lambda_2}} (\hat{a}_{\vec{k}_1\lambda_1}^\dagger)^{N_{\vec{k}_1\lambda_1}}}{\sqrt{\dots N_{\vec{k}_2\lambda_2}! N_{\vec{k}_1\lambda_1}!}} |0\rangle, \quad (4.85)$$

mit $|0\rangle$ den Photonvakuumszustand. Die Operatoren $\hat{a}_{\vec{k}\lambda}^\dagger$ und $\hat{a}_{\vec{k}\lambda}$ erzeugen bzw. vernichten ein Photon in der Mode (\vec{k}, λ) .

Der Feldoperator des Vektorpotentials lässt sich schreiben als

$$\widehat{A}(\vec{r}) = \frac{1}{\sqrt{V}} \sum_{\vec{k}, \lambda} \sqrt{\frac{2\pi\hbar c^2}{\omega_{\vec{k}}}} \left\{ \widehat{a}_{\vec{k}\lambda} \vec{e}_{\vec{k}\lambda} e^{i\vec{k}\cdot\vec{r}} + \widehat{a}_{\vec{k}\lambda}^\dagger \vec{e}_{\vec{k}\lambda}^* e^{-i\vec{k}\cdot\vec{r}} \right\}. \quad (4.86)$$

Beachte, dass der Erwartungswert nur einem klassischen Feld entspricht, wenn wir kohärente Zustände benutzen, wie wir sie oben für den harmonischen Oszillator eingeführt haben.

4.5.2 Photonen sind Bosonen

Die Vertauschungsrelationen der Erzeugungs- und Vernichtungsoperatoren identifizieren die Photonen als Bosonen. Für eine Mode ist die Zustandssumme gegeben durch

$$Z_{\vec{k}\lambda} = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega_{\vec{k}}(n+1/2)} = \frac{1}{2 \sinh(\beta\hbar\omega_{\vec{k}}/2)} \quad (4.87)$$

mit $\beta = 1/k_B T$. Die mittlere Energie und Zahl der Photonen ist bestimmt durch

$$\langle E \rangle = \frac{1}{Z_{\vec{k}\lambda}} \frac{\partial Z_{\vec{k}\lambda}}{\partial \beta} = \hbar\omega \coth\left(\frac{\beta\hbar\omega_{\vec{k}}}{2}\right) = \hbar\omega_{\vec{k}} \left\langle n + \frac{1}{2} \right\rangle \Rightarrow \langle n \rangle = \frac{1}{e^{\beta\hbar\omega_{\vec{k}}} - 1}. \quad (4.88)$$

Letzteres entspricht der Bose-Einstein-Verteilung.

4.5.3 Kopplung an Materie

Nun verwenden wir in $\widehat{V}_{\text{para}}$ anstelle des klassischen Strahlungsfeldes die quantisierte Version, indem wir das Vektorpotential durch seinen Feldoperator (4.86) ersetzen,

$$\widehat{V}_{\text{para}} = -\frac{e}{mc} \widehat{A}(\vec{r}) \cdot \widehat{\vec{p}} = -\frac{e}{mc\sqrt{V}} \sum_{\vec{k}, \lambda} \sqrt{\frac{2\pi\hbar c^2}{\omega_{\vec{k}}}} \left\{ \widehat{a}_{\vec{k}\lambda} \widehat{\vec{p}} \cdot \vec{e}_{\vec{k}\lambda} e^{i\vec{k}\cdot\vec{r}} + \widehat{a}_{\vec{k}\lambda}^\dagger \widehat{\vec{p}} \cdot \vec{e}_{\vec{k}\lambda}^* e^{-i\vec{k}\cdot\vec{r}} \right\}. \quad (4.89)$$

Nun wollen wir wieder unser vorhergehendes Beispiel betrachten, d.h., den Übergang $|\phi_{\vec{n}'}\rangle \rightarrow |\phi_{\vec{n}}\rangle$ mit $\vec{n}' = (n_x, n_y, n_z - 1)$ unter der Erzeugung eines Photons in der Mode (\vec{k}, λ) . Für das Matrixelement erhalten wir

$$\langle \phi_{n-1}; N_{\vec{k}\lambda} | \widehat{V}_{\text{para}} | \phi_n; N_{\vec{k}\lambda} - 1 \rangle = -\hbar e \sqrt{\frac{\pi}{mV}} \sqrt{n_z} e_{\vec{k}\lambda}^z \underbrace{\langle N_{\vec{k}\lambda} | \widehat{a}_{\vec{k}\lambda} | N_{\vec{k}\lambda} - 1 \rangle}_{=\sqrt{N_{\vec{k}\lambda}}}, \quad (4.90)$$

woraus folgt

$$\Gamma_{\vec{n} \rightarrow \vec{n}'}(\vec{k}\lambda) = \frac{2\pi^2 e^2 \hbar}{m} |e_{\vec{k}\lambda}^z|^2 n_z \frac{N_{\vec{k}\lambda}}{V} \delta(\varepsilon_{n-1} - \varepsilon_n + \hbar\omega). \quad (4.91)$$

Dieses Resultat ist identisch mit dem Ausdruck (4.77). Zudem lässt sich nun nachvollziehen, dass auch ein spontaner Zerfall des Zustandes $|\phi_{\vec{n}}\rangle$ möglich ist, d.h. $N_{\vec{k}\lambda} = 1$ setzt keine äussere Störung durch ein endliches Strahlungsfeld voraus. Die Integration über alle möglichen Moden ergibt wieder das gleiche Resultat für Γ_{tot} (analog zu (4.78)).

4.6 Das Vakuum und der Casimir-Effekt

In der quantenmechanischen Situation können wir die Nullpunktsbewegung oder Vakuumfluktuationen für den spontanen Zerfall verantwortlich machen. Dies ist ein sehr interessanter Aspekt der Quantenmechanik. *Das Vakuum ist nicht "leer"*, sondern gefüllt mit Vakuumfluktuationen. Eine unmittelbare und beobachtbare Konsequenz ist der Casimir-Effekt, der auf der Existenz

einer anziehenden Kraft zwischen zwei parallelen Metallplatten im Vakuum basiert. Die Platten erzwingen Knotenpunkte für die Moden des Strahlungsfeldes in ihrem Zwischenraum, wodurch sich die Zahl der möglichen Moden im Vergleich zum uneingeschränkten Raum vermindert. Damit ist der Strahlungsdruck auf die Platten von aussen grösser als von innen.

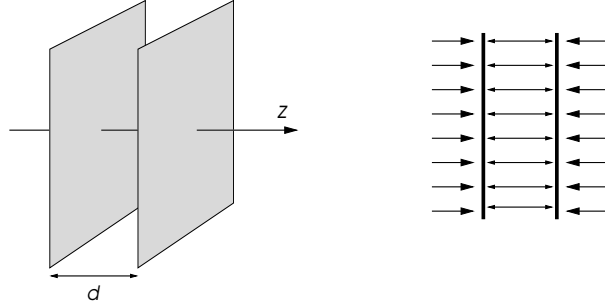


Abb.4.4: Casimir-Effekt: Zwei metallische (unendliche) Platten werden im Abstand d aufgestellt. Der Unterschied des inneren und äusseren Strahlungsdruckes ergibt eine attraktive Kraft.

Die Nullpunktsenergie pro Flächeneinheit der Platten zwischen den beiden Platten entspricht der Summe

$$E' = \frac{1}{2} \sum_{\vec{k}, \lambda} \hbar \omega_{\vec{k}}. \quad (4.92)$$

Wegen der Platten gilt für die Moden, $\omega_{\vec{k}} = c\sqrt{k_x^2 + k_y^2 + (2\pi n/d)^2}$, so dass

$$\begin{aligned} E'(d) &= \frac{\hbar c}{(2\pi)^2} \int d^2 k_{\parallel} \sum_{n=0}^{\infty'} \sqrt{k_{\parallel}^2 + (2\pi n/d)^2} e^{-\eta \sqrt{k_{\parallel}^2 + (2\pi n/d)^2}} \\ &= \frac{\hbar c}{2\pi} \sum_{n=0}^{\infty'} \int_0^{\infty} dk_{\parallel} k_{\parallel} \sqrt{k_{\parallel}^2 + (2\pi n/d)^2} e^{-\eta \sqrt{k_{\parallel}^2 + (2\pi n/d)^2}}, \end{aligned} \quad (4.93)$$

wobei wir einen Konvergenzfaktor $\eta > 0$ eingeführt haben (das Endresultat entspricht dem Limes $\eta \rightarrow 0_+$). Die Summe \sum' bedeutet, dass der Summand $n = 0$ nur halb gezählt wird. Der Term für $n = 0$ ergibt sofort

$$\frac{\hbar c}{4\pi} \int_0^{\infty} dk_{\parallel} k_{\parallel}^2 e^{-\eta k_{\parallel}} = \frac{\hbar c}{2\pi \eta^3}. \quad (4.94)$$

Für die übrigen Terme gilt

$$\int_0^{\infty} dk_{\parallel} k_{\parallel} K e^{-\eta K} = \frac{d^2}{d\eta^2} \int_0^{\infty} \frac{dk_{\parallel} k_{\parallel}}{K} e^{-\eta K} = \frac{d^2}{d\eta^2} \int_{2\pi n/d}^{\infty} dK e^{-\eta K} = \frac{d^2}{d\eta^2} \frac{e^{-\eta 2\pi n/d}}{\eta}. \quad (4.95)$$

Die Summe über n von 1 bis ∞ ist eine geometrische Reihe und führt auf

$$\begin{aligned} \frac{\hbar c}{2\pi} \frac{d^2}{d\eta^2} \left(\frac{1}{\eta(e^{2\pi\eta/d} - 1)} \right) &= \frac{(2\pi)^2 \hbar c}{d^3} \frac{d^2}{dy^2} \left(\frac{1}{y(e^y - 1)} \right) = \frac{(2\pi)^2 \hbar c}{d^3} \frac{d^2}{dy^2} \sum_{n=0}^{\infty} \frac{B_n}{n!} y^{n-2} \Big|_{y=\frac{2\pi\eta}{d}} \\ &= \frac{6\hbar c}{4\pi^2 \eta^4} d - \frac{\hbar c}{2\pi \eta^3} - \frac{\pi^2 \hbar c}{90 d^3} + O(\eta^2), \end{aligned} \quad (4.96)$$

wobei B_n die Bernoulli-Zahlen sind ($B_0 = 1, B_1 = -1/2, B_2 = 1/6, B_3 = 0, B_4 = -1/30, B_5 = 0, \dots$) und $y = 2\pi\eta/d$. Damit ergibt sich für die Energie zwischen den Platten

$$E'(d) = \frac{6\hbar c}{4\pi^2 \eta^4} d - \frac{\pi^2 \hbar c}{90 d^3} + O(\eta^2). \quad (4.97)$$

Nun bleibt noch die Energie der Strahlung im Aussenbereich zu berechnen. Dafür nehmen wir zunächst an, dass das Volumen des ganzen Systems in z -Richtung auf die Länge L ($-L/2 < z < L/2$) beschränkt sei und lassen L am Schluss gegen unendlich gehen. Man findet den Beitrag

$$E''(d) = \frac{6\hbar c}{4\pi^2\eta^4}(L-d) - 16\frac{\pi^2\hbar c}{90(L-d)^3} + O(\eta^2). \quad (4.98)$$

Für sehr grosse L ($\gg d$) lässt sich der zweite Term vernachlässigen und wir erhalten

$$E(d) = E'(d) + E''(d) = \frac{6\hbar c}{4\pi^2\eta^4}L - \frac{\pi^2\hbar c}{90d^3} + O(\eta^2), \quad (4.99)$$

wobei sich der erste Term als Energie der Nullpunktfuktuationen ohne Platten entpuppt. Der zweite Term ist die potentielle Energie einer anziehenden Kraft zwischen den Platten,

$$V_C(d) = -\frac{\pi^2\hbar c}{90d^3} \quad \Rightarrow \quad F = -\nabla_d V_C = \frac{\pi^2\hbar c}{30d^4}. \quad (4.100)$$

Diese ‘‘Kraft aus dem Nichts’’ existiert, wenn keine Photonen vorhanden sind. Solche Kräfte gibt es auch in anderen Zusammenhängen. Ein Beispiel ist die Van-der-Waals-Kraft zwischen neutralen Atomen oder zwischen Dielektrika.³

4.7 Photoemissionspektroskopie

Zum Schluss kehren wir nochmals zum Photoeffekt zurück und beschäftigen uns mit einer weitverbreiteten Anwendung in der Materialforschung. Der Photoeffekt wird heutzutage als wichtiges Instrument in der Erforschung der elektronischen Struktur vieler Materialien eingesetzt. Die durch Lichteinstrahlung freigesetzten Elektronen geben durch ihre Energieverteilung Einblick in die Eigenschaften der Elektronen im Festkörper.

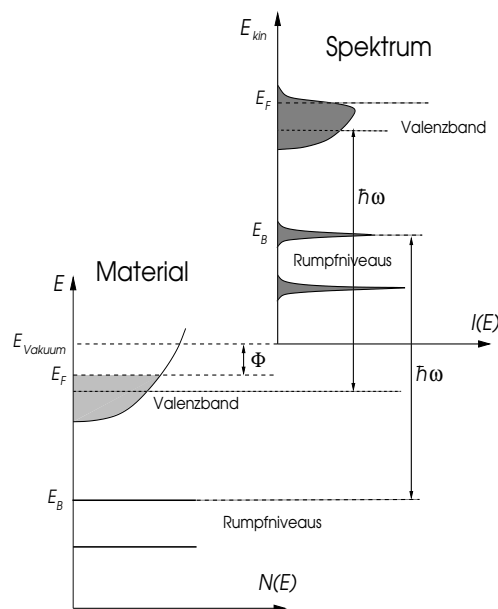


Abb.4.5: Photoemissionspektroskopie: Die Intensitätsverteilung der kinetischen Energien der emittierten Elektronen liefert ein Abbild des Spektrums im Material. Die minimale Austrittsarbeit Φ ist die Photonenergie notwendig um das Elektron höchster Energie aus einem Material herauszulösen.

³Siehe auch Christoph Bruder, *Die Kräfte aus dem Nichts*, Physik unserer Zeit 28, 149 (1997); Astrid Lambrecht, *The Casimir effect: a force from nothing*, Physics World, September 2002.

Ein von einem Material absorbiertes Photon transferiert seine Energie $\hbar\omega$ auf ein Elektron. Für genügend hohe Energien kann das Elektron aus dem Material mit der kinetischen Energie

$$E_{\text{kin}} = \hbar\omega - W = \hbar\omega - \Phi - E \quad (4.101)$$

entkommen, wobei die interessante Information in der Energie W enthalten ist. Diese Energie ist die Summe der Austrittsarbeit Φ des Materials (die minimale Energie um ein Elektron freizusetzen) und der Energie E des Elektrons im Material.

Wird ein Material mit monochromatischem Licht beschienen, können die freigesetzten Elektronen nach ihrer kinetischen Energie analysiert und statistisch gezählt werden. Das Intensitätsspektrum ergibt ein Abbild der Zustandsdichte für die besetzten Elektronenzustände.

Die Zahl der Elektronen, die mit einer gewissen Energie aus dem Material austreten, kann durch die Goldene Regel abgeschätzt werden. Dabei betrachten wir Zustände mit N Elektronen, $|\Psi_A^N\rangle$ und $|\Psi_E^N\rangle$ für den Anfangs- bzw. Endzustand, wobei wir diese Zustände in einen Teil mit $N - 1$ Elektronen und einen zweiten mit einem Elektron aufspalten,

$$|\Psi_A^N\rangle = |\phi_n\rangle \otimes |\Psi_A^{N-1}\rangle \quad \text{und} \quad |\Psi_E^N\rangle = |\phi_{\vec{k}}\rangle \otimes |\Psi_E^{N-1}\rangle. \quad (4.102)$$

Diese Näherung ist nicht trivial und geht davon aus, dass das Entreissen des Elektrons aus seiner Umgebung so plötzlich verläuft, dass das System keine Zeit hat zu relaxieren und das herausgelöste Elektron auch nicht mit dem System wechselwirkt.

Die Goldene Regel ergibt nun

$$\begin{aligned} \Gamma_{AE}(E_{\text{kin}}) &= \frac{2\pi}{\hbar} |\langle \Psi_E^N | \widehat{V}_{\text{para}} | \Psi_A^N \rangle|^2 \delta(E_E^N - E_A^N - \hbar\omega) \\ &= \frac{2\pi}{\hbar} |\langle \phi_{\vec{k}} | \widehat{V}_{\text{para}} | \psi_n \rangle|^2 |\langle \Psi_E^{N-1} | \Psi_A^{N-1} \rangle|^2 \delta(E_{\text{kin}} + E_E^{N-1} - E_A^N - \hbar\omega). \end{aligned} \quad (4.103)$$

Für die Energiebilanz nehmen wir die vereinfachende Form

$$E_{\text{kin}} = \hbar\omega + E_A^N - E_E^{N-1} = \hbar\omega - \Phi - E_n. \quad (4.104)$$

Die Intensität der austretenden Elektronen mit gegebener kinetischer Energie ist dann

$$I(E_{\text{kin}}) = \sum_{A,E} \Gamma_{AE}(E_{\text{kin}}) = \frac{2\pi}{\hbar} \sum_{A,E} |M_{AE}|^2 \delta(E_{\text{kin}} + \Phi - E_n - \hbar\omega). \quad (4.105)$$

Man kann zeigen, dass dies näherungsweise dem Ein-Elektron-Spektrum des Materials entspricht.

Detektiert man neben der Energie zusätzlich die Austrittsrichtung der Elektronen, so kann die Impulskomponente der Elektronen parallel zur Oberfläche bestimmt werden und es lässt sich die Energie-Impuls-Beziehung analysieren. Dies nennt man winkelaufgelöste Photoemissionspektroskopie.

Es ist in diesem Zusammenhang wichtig zu beachten, dass das Photon praktisch keinen Impuls auf das Elektron überträgt. Der charakteristische Impuls eines Elektrons ist von der Grössenordnung $\hbar\pi/a$ (a : Gitterkonstante). Ein solcher Impuls entspräche einer Photoenergie $\hbar\omega = \hbar c\pi/a \sim \alpha mc^2 \sim 2 - 4 \text{ keV}$ unter der Annahme, dass die Gitterkonstante $a \sim a_B \sim \text{Bohrradius}$ ist ($\alpha = 1/137$). Diese Energie liegt weit oberhalb des optischen Spektrums, welches man normalerweise für solche Experimente benutzt. Damit kann man hier den Photonimpuls vernachlässigen. Diese Technik ist speziell in den letzten 15 Jahren stark verfeinert worden. Insbesondere hat sie auch von der Verfügbarkeit von Synchrotron-Strahlungsquellen wie der SLS (Swiss Light Source) profitiert.

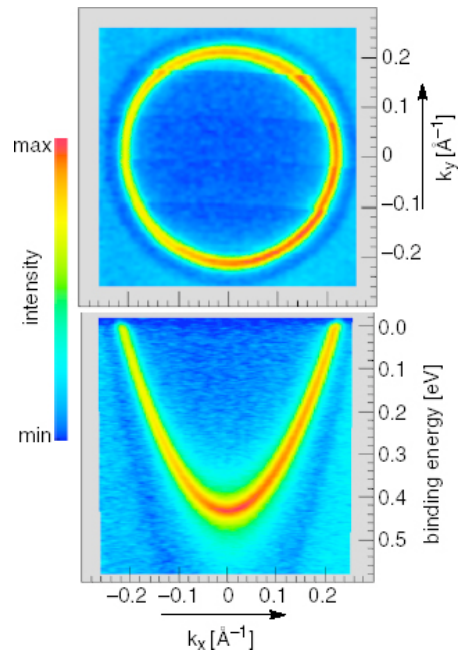


Abb.4.6: Winkelaufgelöste Photoemissionsspektroskopie: Fermifläche (Kreis) und Dispersion der Shockley-Zustände an einer Kupfer (111)-Oberfläche, wobei wir nur die Impuls senkrecht auf (111) auflösen können. Beachte die einfache Dispersion im unteren Panel, $\varepsilon_k = \hbar^2 k_x^2 / 2m$.

Kapitel 5

Elektronische Eigenschaften der Metalle und Halbleiter

In diesem Kapitel werden wir die elektronischen Eigenschaften von Metallen und von Halbleitern, die wir in den ersten beiden Kapiteln in vereinfachter Form diskutiert haben, etwas tiefer untersuchen. Insbesondere werden wir die angetroffenen Näherungen rechtfertigen. Die Beschreibung der Elektronen kann nur mit der Quantenmechanik vernünftig durchgeführt werden. Die Grundlage für das Verständnis der Elektronen ist das Bandbild.

5.1 Fermionen und Bosonen

Für die Elektronen gilt das Pauli-Prinzip, dass keine zwei Elektronen im selben Zustand sein können. Solche Teilchen sind Fermionen. Dies steht im Gegensatz zu den Bosonen. Die Gleichgewichtsverteilung der Fermionen und Bosonen hat die Form

$$n_F(E) = \frac{1}{e^{(E-\mu)/k_B T} + 1} \quad \text{bzw.} \quad n_B(E) = \frac{1}{e^{(E-\mu)/k_B T} - 1}, \quad (5.1)$$

wobei μ das chemische Potential ist. Wir können den Hamilton-Operator $\mathcal{H} = \mathcal{H}_0 - \mu \hat{N}$ annehmen (\hat{N} : Teilchenzahl-Operator). Dabei sei \mathcal{H}_0 ein Hamilton-Operator für unabhängige Teilchen. In der grosskanonischen Betrachtungsweise ist die Zahl der Fermionen oder Bosonen nicht fixiert, sondern das chemische Potential, das die Energieänderung angibt, wenn ein Teilchen hinzugefügt oder entfernt wird. Die kanonische Beschreibung hingegen legt die Teilchenzahl fest. Dann erhält das chemische Potential die Funktion eines Lagrange-Multiplikators, indem es entsprechend angepasst wird, um die Teilchenzahl-Nebenbedingung zu erfüllen,

$$N_F = \sum_n \frac{1}{e^{(\epsilon_n - \mu)/k_B T} + 1} \quad \text{und} \quad N_B = \sum_n \frac{1}{e^{(\epsilon_n - \mu)/k_B T} - 1}, \quad (5.2)$$

wobei $\{\epsilon_n\}$ das Spektrum der Einteilchenzustände in \mathcal{H}_0 ist. Hier nehmen wir an, dass die Zahl der Bosonen erhalten ist. Dies gilt für die Photonen der Hohlraumstrahlung nicht. Ein Vergleich der Verteilungsfunktion für jene Photonen,

$$n(\hbar\omega) = \frac{1}{e^{\hbar\omega/k_B T} - 1} \quad (5.3)$$

zeigt, dass wir das chemische Potential als null annehmen können.

5.2 Elektronenbänder

Die Antwort auf die Frage, warum sich Elektronen in einem Metall praktisch frei bewegen können, obwohl sie sich in einem Gitter von positiv geladenen Ionen befinden, wird durch das

Bloch'sche Theorem gegeben. Das Bloch'sche Theorem sagt aus, dass es in einem periodischen Potential Zustände gibt, deren Wellenfunktion unendlich ausgedehnt ist. Solche Zustände sind das Analogon zu den ebenen Wellen der freien Teilchen im homogenen Raum. Ihre Wellenfunktion hat die allgemeine Form

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\vec{r}} u_{\vec{k}}(\vec{r}) , \quad (5.4)$$

wobei \vec{k} den Pseudoimpuls bezeichnet. $u_{\vec{k}}(\vec{r})$ hat die Periodizität des Potentials, welche durch die Basisvektoren $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ des Gitters gegeben ist. Die Ionen-Gitterpunkte werden durch die Vektoren $\vec{R}_{\vec{n}} = \sum_{i=1,2,3} n_i \vec{a}_i$ verbunden, wobei $n_i = 0, \pm 1, \pm 2, \dots$ ($\vec{n} = (n_1, n_2, n_3)$). Damit gilt für das Potential und die Wellenfunktion,

$$V(\vec{r} + \vec{R}_{\vec{n}}) = V(\vec{r}) , \quad \psi_{\vec{k}}(\vec{r} + \vec{R}_{\vec{n}}) = e^{i\vec{k}\cdot\vec{R}_{\vec{n}}} \psi_{\vec{k}}(\vec{r}) \quad \text{und} \quad u_{\vec{k}}(\vec{r} + \vec{R}_{\vec{n}}) = u_{\vec{k}}(\vec{r}) \quad (5.5)$$

Damit beschreibt $\psi_{\vec{k}}(\vec{r})$ einen Zustand der Eigenzustand des Translationsoperators $\hat{T}_{\vec{R}_{\vec{n}}}$ ist. Da dies eine Symmetrie des Hamilton-Operators ist, d.h. $[\mathcal{H}, \hat{T}_{\vec{R}_{\vec{n}}}] = 0$, kann ein solcher Zustand auch gleichzeitig ein Eigenzustand der Energie sein,

$$\mathcal{H}\psi_{\vec{k}}(\vec{r}) = \varepsilon_{\vec{k}}\psi_{\vec{k}}(\vec{r}) , \quad (5.6)$$

mit $\varepsilon_{\vec{k}}$ als Bandenergie.

5.2.1 Von Atomorbitalen zur Bandstruktur

Betrachten wir ein einfaches Metall wie Natrium, dessen atomare Elektronkonfiguration durch $1s^2 2s^2 2p^6 3s^1$ gegeben ist. Das Atom besitzt also je 2 Elektronen im $1s$ - und $2s$ -Orbital, 6 Elektronen im $2p$ - und eines im $3s$ -Orbital, dem Pauli-Prinzip folgend aufgefüllt nach steigender Energie. Nun bringen wir diese Atome in einem Gitter zusammen, welches wir der Einfachheit halber als einfache (eindimensionale) Kette annehmen, obwohl metallisches Natrium eine kubisch raumzentrierte Struktur hat. Wenn die Atome sich auf eine Distanz annähern, die vergleichbar mit ihrem Atomradius ist, so dass sich ihre atomaren Wellenfunktionen überlappen, dann können Elektronen zwischen den Atomen transferiert werden. Die atomaren Energieniveaus verbreitern sich zu Bändern.

Wir wollen dafür einmal ein einfaches Modell aufstellen. Wir nummerieren die sechs Zustände als

$$n = (1s, 2s, 2p_x, 2p_y, 2p_z, 3s) , \quad (5.7)$$

und die Gitterpunkte als l . Wir können nun den Hamilton-Operator in Dirac-Notation schreiben als

$$\mathcal{H} = \sum_n \mathcal{H}_n + \sum_{n \neq n'} \mathcal{H}_{nn'} \quad (5.8)$$

mit

$$\mathcal{H}_n = \sum_l \epsilon_n |n, l\rangle \langle n, l| - t_n(a) \sum_l \{ |n, l+1\rangle \langle n, l| + |n, l\rangle \langle n, l+1| \} , \quad (5.9)$$

$$\mathcal{H}_{nn'} = \sum_l \{ t_{nn'}(a) |n, l+1\rangle \langle n', l| + t_{n'n}(a) |n, l\rangle \langle n', l+1| \} .$$

Die Zustände $|n, l\rangle$ bezeichnen ein Elektron im Orbital n auf dem Atom l . Wir ignorieren hier den Spin als Index, der nur für die Besetzung eine Rolle spielen wird. Die Orbitalenergien werden als ϵ_n und die Hybridisierungen mit den Nachbar-Atomen als $t_n(a)$ und $t_{nn'}(a)$ für Intra- bzw. Inter-Orbital-Hybridisierung bezeichnet. Diese Matrixelemente sind abhängig von der Gitterkonstanten a mit einer exponentiellen Abhängigkeit, welche wir der Einfachheit halber parametrisieren als

$$t_n(a) = \tilde{t}_n e^{-a/a_n} \quad \text{und} \quad t_{nn'}(a) = \tilde{t}_{nn'} e^{-a/\bar{a}} \quad (5.10)$$

mit a_n und \bar{a} als charakteristische Längenskalen, von der Größenordnung des Atomradius. Die Vorfaktoren werden durch die Form der Orbitale bestimmt.

Wir analysieren hier kurz einige qualitative Aspekte, die mit Symmetrie-Eigenschaften der Orbitale zu tun haben. Die s -Orbitale ($1s$, $2s$ und $3s$) sind rotationssymmetrisch und hybridisieren miteinander auf benachbarten Gitterpunkten (Abb.5.1). Wir fixieren die Phase so, dass die entsprechenden Matrixelemente negativ sind ($t_1, t_2, t_6 < 1$), d.h. positive Wellenfunktion trifft auf positive Wellenfunktion (Abb.5.1a). Für die p -Orbitale gibt es zwei Möglichkeiten: p_z liegt entlang der Verbindungslinie und $p_{x,y}$ stehen senkrecht darauf. Letztere Hybridisierung nur innerhalb der eigenen Spezies, während p_z auch mit s -Orbitalen hybridisieren kann. Die p_z - p_z -Hybridisierung hat ein relatives Minuszeichen, so dass das Matrixelement positiv wird (Abb.5.1d). In der Hybridisierung zwischen s - und p_z -Orbital ergeben die Konfigurationen s - p_z - und p_z - s unterschiedliches Vorzeichen des Matrixelementes (Abb.5.1b). Ferner besitzen die beiden Konfigurationen p_x - p_x - und p_y - p_y -Hybridisierung mit einem negativen Vorzeichen .

(n, n')	$1s$	$2s$	$2p_x$	$2p_y$	$2p_z$	$3s$
$1s$	-	-	0	0	-	-
$2s$	-	-	0	0	-	-
$2p_x$	0	0	-	0	0	0
$2p_y$	0	0	0	-	0	0
$2p_z$	+	+	0	0	+	+
$3s$	-	-	0	0	-	-

Tab.5.1: Vorzeichen der verschiedenen Matrixelemente $t_n (= t_{nn})$ und $t_{nn'}$.

Zunächst führen wir neue Zustände durch folgende unitäre Transformation ein,

$$|n, l\rangle = \sum_k |n, k\rangle \langle n, k | n, l\rangle = \sum_k e^{ikl} |n, k\rangle \quad (5.11)$$

Wodurch wir den Hamilton-Operator neu schreiben können als

$$\mathcal{H} = \sum_k \sum_{n, n'} \left\{ (\epsilon_n + t_n(a)\gamma_k)\delta_{nn'} + t_{nn'}(a)e^{ika} + t_{n'n}(a)e^{-ika} \right\} |n, k\rangle \langle n', k| \quad (5.12)$$

mit $\gamma_k = 2 \cos ka$. Beachte, dass $\gamma_{k+2\pi/a} = \gamma_k$ und $\exp(\pm i(k + 2\pi/a)) = \exp(\pm ika)$, so dass wir uns auf den Bereich $-\pi/a < k \leq \pi/a$ beschränken können, den wir erste Brillouin-Zone nennen.

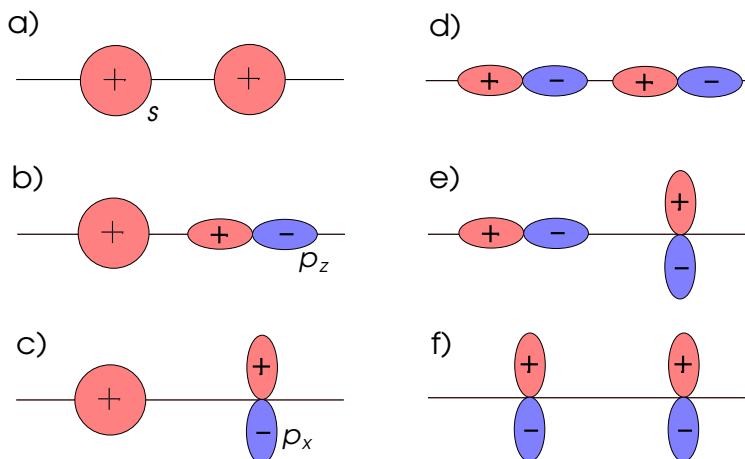


Abb.5.1: Orbital-Konfigurationen auf benachbarten Atomen: Konfigurationen a), b) und d) sind rotationssymmetrisch um die z -Achse und ergeben sogenannte σ -Hybridisierung. Konfiguration f) hat 180° -Symmetrie und entspricht π -Hybridisierung zwischen den beiden Orbitalen. Konfigurationen c) und e) haben aus Symmetriegründen keine Hybridisierung, das die Interferenz

vollständig destruktiv ist. Beachte, dass im Zusammenhang mit den p -Orbitalen die Hybridisierung vom Vorzeichen der Wellenfunktion abhängt.

Im Grenzfall $a \gg a_n, \tilde{a}$ finden wir das diskrete, k -unabhängige Spektrum der Na-Atome, wobei die Orbitale sukzessive durch Elektronen mit beiden Spins besetzt sind, das oberste ($3s$ -) Orbital wird durch nur ein Elektron besetzt. Die Einteilchen-Zustände sind hoch entartet. Durch annähern der Atome wird die Entartung aufgehoben und die Zustände gewinnen eine Dispersion.¹ Es entstehen Energiebänder, die noch in einem gewissen Mass den Charakter der ursprünglichen Orbitale mittragen, aber auch durch Hybridisierung untereinander gemischt sind. Interessant ist die Aufspaltung der Entartung für die $2p$ -Orbital, wo $2p_z$ eine andere Dispersion als $2p_{x,y}$ aufweist.

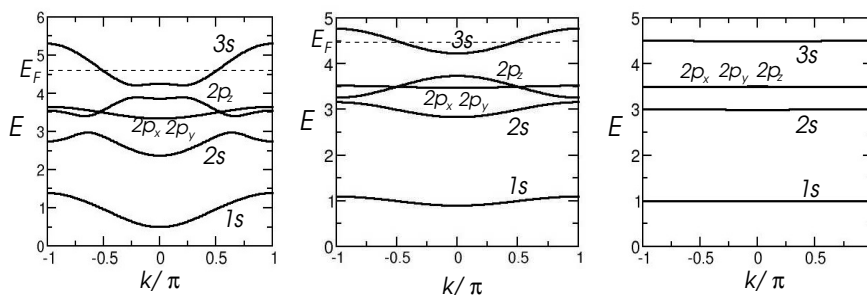


Abb.5.3: Bandstruktur für "Natrium-Kette": Wachsende Gitterkonstante a vom linken zum rechten Bild.

Die Tatsache, dass Natrium eine ungerade Anzahl von Elektronen besitzt, führt dazu, dass oberste betrachtete Band mit vorwiegend $3s$ -Charakter ein halbgefülltes Band hat. Es sind daher Anregungen beliebig kleiner Energie möglich. Insbesondere kann man mit infinitesimalem Energieaufwand dem Elektronensystem einen endlichen Impuls geben. Daher kann ein solches System elektrischen Strom führen, wenn ein elektrisches Feld angelegt wird, d.h. es ist metallisch.

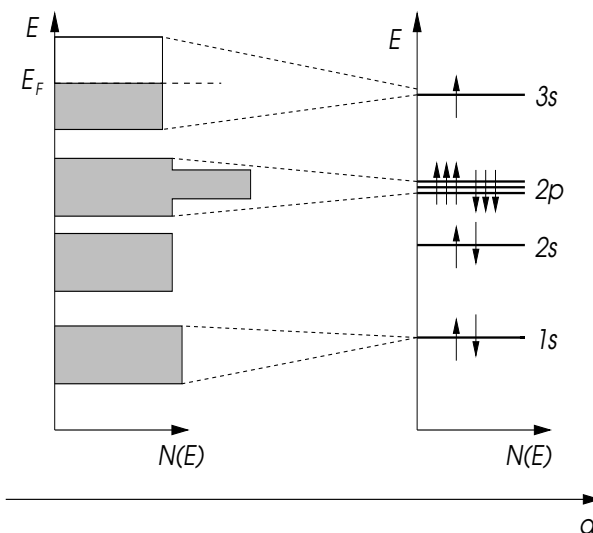


Abb.5.4: Schematische Graphik der Bandfüllung: Untere Bänder sind vollständig gefüllt. Die Elektronen in diesen Bändern bilden im Wesentlichen die Ionen-Rümpfe. Das oberste Band ist halbgefüllt. Die Elektronen dieses Bandes nehmen die Rolle der Leitungselektronen ein.

¹In dieser Betrachtung haben wir die Coulomb-Wechselwirkung ausser Acht gelassen. Natürlich ist es mit einem Energieaufwand verbunden, die Zahl der Elektronen von 11 verschieden zu machen. Daher würde der Übergang zu den Bändern nicht so verlaufen, wie hier gezeigt, sondern die Matrixelemente müssten einen Mindestwert annehmen bis Bänder auftreten. Vorher nennen wir das System einen Mott-Isolator, d.h. isolieren aufgrund der abstossenden Wechselwirkung zwischen Elektronen.

Bei Temperatur $T = 0\text{ K}$ entspricht die höchste Energie der besetzten Zustände, die Fermi-Energie E_F , dem chemischen Potential μ .

5.2.2 Semiklassische Beschreibung der Bandelektronen

In der Quantenmechanik haben wir das Ehrenfest Theorem gesehen, gemäss dem die Erwartungswerte des Orts- und Impulsoperators Gleichungen erfüllen, die der Newton'schen Mechanik ähnlich sind. Eine analoge Formulierung gilt auch für die Elektronen im periodischen Potential, wobei wir davon ausgehen, dass ein Elektron als Teilchen durch ein Wellenpaket beschrieben wird, dessen Wellenfunktion die Form hat

$$\psi(\vec{r}, t) = \sum_{\vec{k}'} g(\vec{k}') e^{i\vec{k}' \cdot \vec{r} - i\epsilon_{\vec{k}'} t}, \quad (5.13)$$

wobei $g(\vec{k}')$ um \vec{k} mit einer Breite Δk zentriert ist. Um innerhalb des Bandbildes eine sinnvolle Näherung zu haben, sollte Δk viel kleiner als die Ausdehnung der Brillouin-Zone sein: $\Delta k \ll 2\pi/a$ was mit der Heisenberg'schen Unschärfe-Relation $\Delta k \Delta x > 1$ dazu führt, dass das Wellenpaket über viele Einheitszellen des Gitters ausgedehnt ist: $\Delta x \gg a/2\pi$.

Ohne einen Beweis führen wir die Regeln der semiklassischen Formulierung der Elektronenbewegung ein, wenn ein elektrisches und ein magnetisches Feld angelegt wird:

- Der Bandindex eines Elektrons ist eine Erhaltungsgrösse. Es gibt keine Übergänge zwischen verschiedenen Bändern.
- Die Bewegungsgleichungen lauten:

$$\dot{\vec{r}} = \vec{v}_n(\vec{k}) = \frac{1}{\hbar} \frac{\partial \epsilon_n(\vec{k})}{\partial \vec{k}} \quad (5.14)$$

$$\hbar \dot{\vec{k}} = -e\vec{E}(\vec{r}, t) - \frac{e}{c} \vec{v}_n(\vec{k}) \times \vec{H}(\vec{r}, t)$$

- Alle Elektronenzustände haben ihren Wellenvektor innerhalb der ersten Brillouin-Zone, da \vec{k} und $\vec{k} + \vec{K}$ demselben Zustand entsprechen (\vec{K} : reziproker Gittervektor).
- Im thermischen Gleichgewicht ist die Dichte der Elektronen im n -ten Band im Volumenelement d^3k bei \vec{k} gegeben durch

$$n_F(\epsilon_n(\vec{k})) \frac{d^3k}{4\pi^3} = \frac{1}{e^{(\epsilon_n(\vec{k}) - \mu)/k_B T} + 1} \frac{d^3k}{4\pi^3}, \quad (5.15)$$

wobei $d^3k/4\pi^2 = 2d^3k/(2\pi)^3$ entspricht mit der "2" für die beiden möglichen Spinzustände des Elektrons.

Beachte, dass $\hbar \vec{k}$ nicht den Impuls des Elektrons darstellt, sondern den sogenannten Gitterimpuls der Bloch'schen Band-Theorie und daher mit dem Eigenwert des Zustandes bei der Gitter-Translationsoperation zusammenhängt. Daher entspricht die rechte Seite der zweiten Gleichung in (5.14) nicht der Kraft, die auf das Elektron wirkt, denn wir haben hier das periodische Potential des Ionengitters nicht miteingeschlossen. Dessen Effekt kommt in der Form der Bandenergie $\epsilon_n(\vec{k})$ ins Spiel (Das Wellenpaket ist viel grösser als die Gitterkonstante.).

Ein Plausibilitätsargument führt auf die zweite Gleichung in (5.14), indem wir nämlich die Energieerhaltung heranziehen. Die zeitliche Ableitung von

$$E = \epsilon_n(\vec{k}(t)) - e\phi(\vec{r}(t)) \quad (5.16)$$

muss verschwinden,

$$0 = \frac{dE}{dt} = \frac{\partial \varepsilon_n(\vec{k})}{\partial \vec{k}} \cdot \dot{\vec{k}} - e \vec{\nabla} \phi \cdot \dot{\vec{r}} = \vec{v}_n(\vec{k}) \cdot \left\{ \hbar \dot{\vec{k}} - e \vec{\nabla} \phi \right\} . \quad (5.17)$$

Daraus folgt (5.14), wobei der Anteil der Lorentzkraft dadurch erlaubt ist, dass er senkrecht auf der Geschwindigkeit \vec{v}_n steht.

5.3 Metalle - Isolatoren - Halbleiter

5.3.1 Metalle - Partiiell gefüllte Bänder

Die Natrium-Kette hat ein partiell gefülltes Band, da pro Einheitszelle eine ungerade Zahl von Elektronen vorhanden sind. Dies ergibt sich in diesem Fall ganz einfach durch Abzählen der Elektronen auf einem isolierten Na-Atom, das 11 Elektronen ergibt. Damit ist das 3s-artige Band halbgefüllt, d.h. wir könnten noch einmal soviel Elektronen in diesem Band unterbringen, wie schon vorhanden sind. Dies ist nicht nur für das einfache Modell der Kette der Fall, sondern für das richtige dreidimensionale Na-Gitter. Die analoge Betrachtung gilt auch für die anderen Alkali-Metalle: Li, K, Rb, Die Fermienergie E_F liegt im Bereich des 3s-Bandes und $\varepsilon_n(\vec{k}) = E_F$ definiert eine Fläche im reziproken \vec{k} -Raum, die sogenannte Fermifläche.

Die semiklassischen Bewegungsgleichungen beschreiben die Dynamik der Elektronen, wenn wir ein \vec{E} -Feld anlegen (wir lassen das Magnetfeld hier beiseite). Dies entspricht der Bewegung der Elektronen zwischen Streueignissen. Das Anlegen eines elektrischen Feldes führt zu einer Beschleunigung der Elektronen, die wegen der Streuung in einen stationären Zustand übergehen. Durch die Boltzmann-Transportgleichung kann man die Elektronenverteilung $n(\vec{k}, t)$ beschreiben, wobei wir $n(\vec{k}, t)$ als räumlich homogen annehmen. Mit $n(\vec{k}, t)$ ist die Zahl der Elektronen im Volumenelement d^3k des reziproken Raumes,

$$dN = n(\vec{k}, t) \frac{d^3k}{4\pi^3} . \quad (5.18)$$

Wir nehmen an, dass $n(\vec{k}, t)$ nur leicht von der Gleichgewichtsverteilung $n_F(\varepsilon(\vec{k}))$ abweicht. Der Einfachheit halber nehmen wir ein einzelnes Band der Energie $\varepsilon(\vec{k})$ an.

$$n(\vec{k}, t) = n_F(\varepsilon(\vec{k})) + g(\vec{k}, t) . \quad (5.19)$$

Die Abweichung vom Gleichgewicht hat verschiedene Ursachen. Falls wir ein elektrisches Feld anlegen, ergibt sich eine Beschleunigung. Die Elektronen können gestreut werden, d.h. sie ändern ihre Bewegungsrichtung. Im stationären Fall muss die Gesamtheit dieser Änderung sich aufheben:²

$$\begin{aligned} 0 &= \left. \frac{dn(\vec{k}, t)}{dt} \right|_{\text{feld}} - \left. \frac{dn(\vec{k}, t)}{dt} \right|_{\text{streu}} = \dot{\vec{k}} \cdot \frac{\partial n(\vec{k}, t)}{\partial \vec{k}} - \left. \frac{dn(\vec{k}, t)}{dt} \right|_{\text{streu}} \approx \dot{\vec{k}} \cdot \frac{\partial n_F(\varepsilon(\vec{k}))}{\partial \vec{k}} - \left. \frac{dg(\vec{k}, t)}{dt} \right|_{\text{streu}} \\ &\approx \dot{\vec{k}} \cdot \frac{\partial \varepsilon(\vec{k})}{\partial \vec{k}} \frac{\partial n_F(\varepsilon(\vec{k}))}{\partial \varepsilon(\vec{k})} + \frac{g(\vec{k}, t)}{\tau} = -e \vec{E} \cdot \vec{v}(\vec{k}) \frac{\partial n_F(\varepsilon(\vec{k}))}{\partial \varepsilon(\vec{k})} + \frac{g(\vec{k}, t)}{\tau} . \end{aligned} \quad (5.21)$$

² Boltzmann-Transport-Gleichung: Die allgemeine Transportgleichung für die Verteilungsfunktion $n(\vec{r}, \vec{k}, t)$ lautet

$$\frac{Dn}{Dt} = \frac{\partial n}{\partial t} + \dot{\vec{r}} \cdot \frac{\partial n}{\partial \vec{r}} + \dot{\vec{k}} \cdot \frac{\partial n}{\partial \vec{k}} = \left. \frac{\partial n}{\partial t} \right|_{\text{streu}} \quad (5.20)$$

wobei Dn/Dt die sogenannte substantielle Ableitung ist, die die Änderung der Verteilungsfunktion misst, wenn man sich mit dem Volumenelement um (\vec{r}, \vec{k}) im Phasenraum mitbewegt. Der Streutermin auf der rechten Seite kann in verschiedener Weise behandelt werden. Die einfachste Art ist die Relaxationszeit-Näherung, wie wir sie in diesem Abschnitt anwenden. Für unser Problem können wir die räumliche Abhängigkeit vernachlässigen und es gibt im stationären Fall auch keine explizite t -Abhängigkeit von n .

Dabei haben wir die semiklassischen Gleichungen verwendet und zudem die Relaxationszeitnäherung angenommen: $\partial g/\partial t|_{coll} = -g/\tau$ mit τ als Relaxationszeit, die charakteristische Zeit über welche das System aufgrund von Streuprozessen in seine Gleichgewichtslage relaxiert.

Die Lösung dieser Gleichung lautet:

$$g(\vec{k}) = e\tau \vec{E} \cdot \vec{v}(\vec{k}) \frac{\partial n_F(\varepsilon(\vec{k}))}{\partial \varepsilon(\vec{k})} = \frac{-e\tau \vec{E} \cdot \vec{v}(\vec{k})}{4k_B T \cosh^2[(\varepsilon(\vec{k}) - \mu)/2k_B T]} . \quad (5.22)$$

Diese Abweichung beschränkt sich auf einen dünnen Bereich der Dicke $k_B T$ um die Fermi-Energie (chemische Potential).

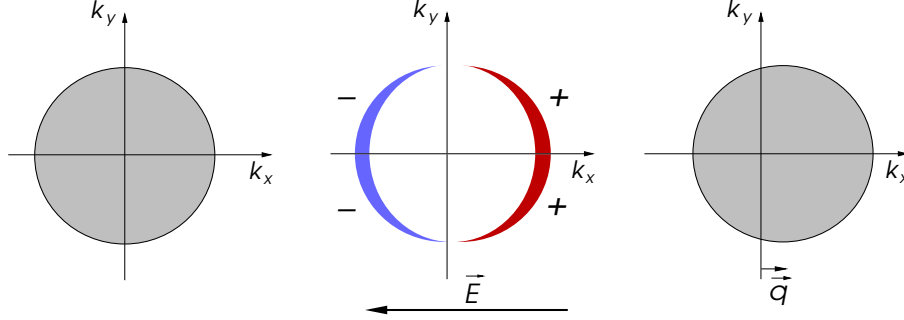


Abb.5.5: Schematische Gitterimpulsverteilung: Links: Fermisee im Gleichgewicht; Mitte: Abweichung vom Gleichgewicht $g(\vec{k})$ mit positivem (negativem) Anteil antiparallel (parallel) zu \vec{E} (beachte, dass $\vec{E} \cdot \vec{v} \propto \cos \theta = k_x/|\vec{k}|$); Rechts: resultierender verschobener Fermisee.

Dies bedeutet, dass nur der Rand des Fermisees durch das elektrische Feld beeinflusst wird. Wie Abb.5.5 jedoch zeigt, resultiert die Abweichung im Wesentlichen in einen verschobenen Fermisee. Daher kann man das Resultat auch dahingehend interpretieren, dass alle Elektronen im Mittel einen zusätzlichen Gitterimpuls \vec{q} erhalten haben. Dies entspricht dem (klassischen) Drude-Bild, dass wir früher betrachtet hatten. Wir können nämlich den Strom ausrechnen,

$$\begin{aligned} \vec{j} &= -e \int \frac{d^3 k}{4\pi^3} \vec{v}(\vec{k}) n(\vec{k}) = -e \int \frac{d^3 k}{4\pi^3} \vec{v}(\vec{k}) \left\{ n_F(\varepsilon(\vec{k})) + g(\vec{k}) \right\} \\ &= -e^2 \tau \int \frac{d^3 k}{4\pi^3} \vec{v}(\vec{k}) \vec{E} \cdot \vec{v}(\vec{k}) \frac{\partial n_F(\varepsilon(\vec{k}))}{\partial \varepsilon(\vec{k})} \end{aligned} \quad (5.23)$$

Nehmen wir nun an, dass $\vec{j} \parallel \vec{E}$ mit $\vec{j} = \sigma \vec{E}$ dann gilt

$$\sigma = \frac{e^2 \tau}{4\pi^3} \int d^3 k \left(\frac{\vec{E} \cdot \vec{v}}{|\vec{E}|} \right)^2 \left(-\frac{\partial n_F(\varepsilon(\vec{k}))}{\partial \varepsilon(\vec{k})} \right) \quad (5.24)$$

Die Winkelmittelung ergibt dann $(\vec{E} \cdot \vec{v})^2 = |\vec{E}|^2 |\vec{v}|^2 / 3$ und

$$\left(-\frac{\partial n_F(\varepsilon(\vec{k}))}{\partial \varepsilon(\vec{k})} \right) \approx \delta(\varepsilon(\vec{k}) - E_F) . \quad (5.25)$$

Damit folgt

$$\sigma = \frac{e^2 \tau}{12\pi^3} \int_{FF} \frac{\bar{v}^2 dS}{|\partial \varepsilon(\vec{k})/\partial \vec{k}|} = \frac{e^2 \tau}{12\pi^3 \hbar} \int_{FF} \bar{v} dS \approx \frac{e^2 \tau}{12\pi^2 \hbar} v_F 4\pi \bar{k}_F^2 = \frac{e^2 n_e \tau}{m^*} \quad (5.26)$$

wobei \int_{FF} das Integral über die Fermifläche bedeutet. Ferner haben wir verwendet, dass $\bar{k}_F^3 = 3\pi^2 n_e$ mit n_e als Elektronendichte. Damit erreichen wir wieder die Drude-Formel, welche hier die effektive Masse m^* definiert, so dass $\hbar \bar{k}_F = m^* \bar{v}_F$.

Partiell gefüllte Bänder haben eine Fermifläche und resultieren in einer endlichen Leitfähigkeit. Dies zeichnet metallisches Verhalten aus. Die Relaxationszeit, die wir in unserer Rechnung als Konstante eingeführt haben, ist im Allgemeinen energieabhängig, was zu einer Temperaturabhängigkeit führt. Wichtigste Beiträge zu τ kommen von der Imperfektion des Gitters. Dies kommt einerseits von Fehlstellen und Verunreinigung mit "falschen" Atomen. Dieser Beitrag ist im Wesentlichen temperaturunabhängig und beschränkt die Leitfähigkeit bei tiefen Temperaturen. Thermische Fluktuationen führen zur Bewegung des Ionengitters, wodurch sich eine starke Temperaturabhängigkeit ergibt, da diese Fluktuationen mit steigender Temperatur anwachsen. Daraus ergibt sich ein Verhalten, das wir für den spezifischen Widerstand $\rho = 1/\sigma$ hinschreiben wollen:

$$\rho(T) = \rho_0 + \begin{cases} AT^5 & T \ll T_D \\ A'T & T \gg T_D \end{cases} \quad (5.27)$$

wobei T_D die charakteristische Temperatur (Debye-Temperatur) der Gitterbewegung ist, im Normalfall einige 100 K.

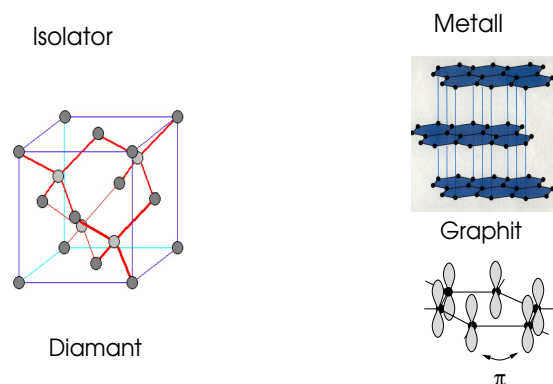


Abb.5.6: Strukturen des Kohlenstoff: Diamant und Graphit: Die mobilen Elektronen des Graphit haben 2p-Charakter und entsprechen den p-Orbitalen, die aus der Graphit-Ebene hinauszeigen. Ihre Hybridisierung hat π -Charakter.

5.3.2 Isolatoren und Halbleiter

Falls alle Bänder entweder ganz gefüllt oder leer sind, haben wir einen Isolator oder einen Halbleiter. Die Unterscheidung trifft man über die Energielücke E_g , die Differenz der Energie des untersten unbesetzten und obersten besetzten Zustands. Für grosse Energien E_g im Bereich von mehreren eV spricht man von einem Isolator. Für $E_g \lesssim 1.5eV$ ist es ein Halbleiter. Eine notwendige aber nicht hinreichende Bedingung für einen Isolator/Halbleiter ist die Eigenschaft, dass es im Mittel eine gerade Anzahl Elektronen pro Gittereinheitszelle hat. Ein Beispiel dafür finden wir bei Kohlenstoff mit der Elektronenkonfiguration C $(1s)^2(2s)^2(2p)^2$. Kohlenstoff kommt als Diamant mit einer dreidimensionalen Kristallstruktur vor, die zwei C-Atome pro Einheitszelle besitzt. Dies ist ein Isolator mit einer Energielücke $E_g \approx 5.5eV$. Andererseits gibt es auch den Graphit mit einer zweidimensionalen Honigwabengitter-Struktur. Auch hier liegen zwei C-Atome in der Einheitszelle. Aber Graphit ist metallisch, wenn auch mit einer ganz speziellen Bandstruktur. Bei Diamant sind alle vier äusseren Elektronen $(2s)^2(2p)^2$ in kovalenten Bindungen eingeschlossen. Bei Graphit sind es nur drei der vier. Ein 2p-Elektron kann sich über π -Hybridisierung entlang der Graphit-Ebene bewegen. Es kommt also sehr stark auf die Bandstruktur an. Falls sich verschiedene Bänder überlappen, können auch Systeme mit einer geraden Anzahl Elektronen pro Einheitszelle metallisch sein. Beispiele dafür sind auch die Erdalkali-Metalle Be, Mg, Ca, Sr,

Isolatoren und Halbleiter können bei $T = 0K$ keinen elektrischen Strom führen. Wir beschränken unsere Betrachtung auf das oberste gefüllte Band. Die Stromdichte ergibt sich analog zu oben,

$$\vec{j} = -e \int_{BZ} \frac{d^3k}{4\pi^3} n(\vec{k}) \vec{v}(\vec{k}) = -e \int_{BZ} \frac{d^3k}{4\pi^3} n(\vec{k}) \frac{1}{\hbar} \frac{\partial \varepsilon_n(\vec{k})}{\partial \vec{k}} = 0. \quad (5.28)$$

Da alle Zustände besetzt sind und somit $n(\vec{k}) = 1$, geht das Integral über die ganze Brillouin-Zone (BZ). Das Integral verschwindet, da $\varepsilon_n(\vec{k})$ im reziproken Raum eine periodische Funktion ist, d.h. $\varepsilon_n(\vec{k} + \vec{K}) = \varepsilon_n(\vec{k})$ mit \vec{K} als reziprokem Gittervektor.

5.3.3 Bild der Löcher

Ein interessanter Aspekt der Bandtheorie ist das Bild der Löcher. Wir berechnen wieder die Stromdichte für ein partiell gefülltes Band innerhalb der semiklassischen Näherung:

$$\begin{aligned} \vec{j} &= -e \int_{BZ} \frac{d^3k}{4\pi^3} n(\vec{k}) \vec{v}_n(\vec{k}) = -e \left\{ \int_{BZ} \frac{d^3k}{4\pi^3} \vec{v}(\vec{k}) - \int_{BZ} \frac{d^3k}{4\pi^3} [1 - n(\vec{k})] \vec{v}(\vec{k}) \right\} \\ &= +e \int_{BZ} \frac{d^3k}{4\pi^3} [1 - n(\vec{k})] \vec{v}(\vec{k}). \end{aligned} \quad (5.29)$$

Dies bedeutet, dass wir die Stromdichte über die Elektronen in den besetzten Zuständen mit Ladung $-e$ bestimmen können, oder aber über die "Löcher" (positive Ladung), die fehlenden Elektronen, in den unbesetzten Zuständen. Beide Beschreibungen sind in der Bandtheorie äquivalent. Es ist jedoch üblicherweise einfacher mit Löchern zu arbeiten, wenn ein Band fast gefüllt ist. Umgekehrt eignet sich das Elektronenbild sicher besser für ein schwach gefülltes Band.

5.3.4 Halbleiter bei endlicher Temperatur

Wir betrachten nun einen Halbleiter, in dem alle Bänder bis zum Band n gefüllt sind, und das $n + 1$ -te Band und alle darüberliegenden leer sind. Das n -te Band nennt man Valenzband und das $n + 1$ -te Band ist das Leitungsband. Wir werden von nun die Indizes v (valence) bzw. c (conduction) für diese beiden Bänder benutzen. Bei endlicher Temperatur werden einige Elektronen des Valenzbandes sich thermisch angeregt im Leitungsband aufhalten. Es gilt wegen der Elektronenzahl-Erhaltung

$$\begin{aligned} \int_{BZ} \frac{d^3k}{4\pi^3} [1 - n_F(\varepsilon_v(\vec{k}))] &= \int_{BZ} \frac{d^3k}{4\pi^3} n_F(\varepsilon_c(\vec{k})) \\ \Rightarrow 0 &= \int_{BZ} \frac{d^3k}{4\pi^3} \left\{ \frac{1}{e^{-(\varepsilon_v(\vec{k}) - \mu)/k_B T} + 1} - \frac{1}{e^{(\varepsilon_c(\vec{k}) - \mu)/k_B T} + 1} \right\} \end{aligned} \quad (5.30)$$

Mit dieser Gleichung können wir das chemische Potential festlegen (μ als Lagrange-Multiplikator). Zur Vereinfachung nehmen wir an, dass nur die Zustände an den Bandrändern wichtig sind, da die Temperatur klein sein soll verglichen mit der Energielücke E_g ($\gg k_B T$). Wir machen die Parametrisierung der Bandränder durch die effektive Massen:

$$\varepsilon_v(\vec{k}) = -\frac{\hbar^2 k^2}{2m_v} \quad \text{und} \quad \varepsilon_c(\vec{k}) = E_g + \frac{\hbar^2 k^2}{2m_c}. \quad (5.31)$$

Wir legen den Energienullpunkt an den oberen Rand des Valenzbandes. Mit der Annahme, dass $\mu \gg k_B T$ und $E_g - \mu \gg k_B T$ erhalten wir folgende Gleichung

$$\begin{aligned}
0 &= \int_{BZ} \frac{d^3 k}{4\pi^3} \left\{ \frac{1}{e^{-(-\hbar^2 k^2/2m_v - \mu)/k_B T} + 1} - \frac{1}{e^{(E_g - \mu + \hbar^2 k^2/2m_c)/k_B T} + 1} \right\} \\
&\approx \int_{BZ} \frac{d^3 k}{4\pi^3} \left\{ e^{-\mu/k_B T} e^{-\hbar^2 k^2/2m_v k_B T} - e^{-(E_g - \mu)/k_B T} e^{-\hbar^2 k^2/2m_c k_B T} \right\} \\
&\approx \frac{(2k_B T)^{3/2}}{4\pi^3 \hbar^3} \int d^3 y e^{-y^2} \left\{ m_v^{3/2} e^{-\mu/k_B T} - m_c^{3/2} e^{(\mu - E_g)/k_B T} \right\}.
\end{aligned} \tag{5.32}$$

Daraus ergibt sich unmittelbar für das chemische Potential

$$\mu = \frac{E_g}{2} + \frac{3}{4} k_B T \ln \left(\frac{m_v}{m_c} \right) \tag{5.33}$$

Beachte, dass $\mu = \mu_e = -\mu_h$ für Elektronen und Löcher gemäss unserer Definition in Kapitel 1. Bei $T = 0K$ liegt das chemische Potential genau in der Mitte der Energielücke.

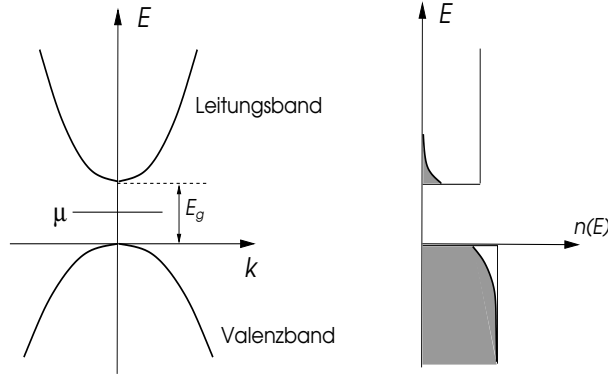


Abb. 5.7: Verteilung der Elektronen auf das Valenz- und Leitungsband bei endlicher Temperatur. Falls $k_B T \ll E_g$ finden wir eine kleine Konzentration von Elektronen im Leitungsband und die entsprechend gleiche Konzentration von Löchern (Vakanzen) im Valenzband.

Betrachten wir nun die Leitfähigkeit bei endlicher Temperatur, zunächst den Beitrag des Valenzbandes.

$$\begin{aligned}
\sigma_v &= e^2 \tau_v \int_{BZ} \frac{d^3 k}{4\pi^3} \frac{|\vec{v}|^2}{3} \left(-\frac{\partial n_F(\varepsilon_v(\vec{k}))}{\partial \varepsilon_v(\vec{k})} \right) \approx \frac{e^2 \tau_v}{12\pi^3} \int_{BZ} d^3 k |\vec{v}|^2 \frac{1}{k_B T} e^{(\mu - \varepsilon_v(\vec{k}))/k_B T} \\
&\approx \frac{e^2 \tau_v}{3\pi^2} \frac{\hbar^2}{k_B T m_v^2} e^{-\mu/k_B T} \int_0^\infty dy y^4 e^{-y^2} = \frac{e^2 \tau_v n_h}{m_v},
\end{aligned} \tag{5.34}$$

wobei τ_v die Relaxationszeit der Valenzelektronen ist und n_h die Dichte der Löcher im Valenzband bezeichnet,

$$n_h = \int_{BZ} \frac{d^3 k}{4\pi^3} [1 - n_F(\varepsilon_v(\vec{k}))] \approx \int_{BZ} \frac{d^3 k}{4\pi^3} e^{(\mu - \varepsilon_v(\vec{k}))/k_B T} \approx \frac{e^{-\mu/k_B T}}{\pi \sqrt{2\pi}} \left(\frac{m_v k_B T}{\hbar^2} \right)^{3/2}. \tag{5.35}$$

Analog finden wir für die Leitungselektronen,

$$\sigma_c = \frac{e^2 \tau_c n_e}{m_c} \quad \text{mit} \quad n_e = n_h. \tag{5.36}$$

Wenn wir (1.20) benutzen erhalten wir für n_h und n_e ,

$$n_h = n_e = \frac{1}{\pi\sqrt{2\pi}} \left(\frac{k_B T}{\hbar^2} \right)^{3/2} (m_v m_c)^{3/4} e^{-E_g/2k_B T}. \quad (5.37)$$

womit die Leitfähigkeit

$$\sigma = \sigma_c + \sigma_v = e^2 n_h(T) \left(\frac{\tau_c}{m_c} + \frac{\tau_v}{m_v} \right) \propto T^{3/2} e^{-E_g/2k_B T} \quad (5.38)$$

exponentiell unterdrückt wird, wenn T gegen Null geht. Dies ist das Verhalten eines Isolators/Halbleiters. Ferner ergibt sich für das Massewirkungsgesetz (1.25) das wir im ersten Kapitel gesehen haben

$$n_e n_h = n_i(T)^2 = \frac{1}{2\pi^3} \left(\frac{k_B T}{\hbar^2} \right)^3 (m_v m_c)^{3/2} e^{-E_g/k_B T}. \quad (5.39)$$

Ferner betrachten wir die mittlere Energie der Elektronen im Leitungs- und Löcher im Valenzband,

$$\bar{E}_v = \int_{BZ} \frac{d^3 k}{4\pi^3} \varepsilon_v(\vec{k}) [1 - n_F(\varepsilon_v(\vec{k}))] = -n_h(T) \frac{3}{2} k_B T, \quad (5.40)$$

$$\bar{E}_c = \int_{BZ} \frac{d^3 k}{4\pi^3} \varepsilon_c(\vec{k}) n_F(\varepsilon_c(\vec{k})) = n_e(T) \left[E_g + \frac{3}{2} k_B T \right].$$

Dies entspricht dem Äquipartitionsgesetz für unabhängige klassische Teilchen. Daher ist das Bild der Ladungsträger als ideales Gas, das wir im Kapitel 1 aufgestellt hatten, eine gute Näherung.

5.4 Photoinduzierte Übergänge im Halbleiter

Wie wir in Kapitel 1 beschrieben hatten, kann der Zustand des Halbleiters bei endlicher Temperatur als ein Gleichgewicht zwischen Elektron-Loch-Paaren und den Photonen der thermischen Strahlung des Material betrachtet werden. Uns interessiert nun das Absorptions-/Emissionsverhalten eines Halbleiters im Strahlungsfeld. Wir unterscheiden zwei Typen von Halbleitern: diejenigen mit einer direkten und diejenigen mit einer indirekten Energielücke.

$$E_g = \min_{\vec{k}, \vec{q}} [\varepsilon_c(\vec{k} + \vec{q}) - \varepsilon_v(\vec{k})] \quad \text{mit} \quad \begin{cases} \vec{q}_{min} = 0 & \text{direkt} \\ \vec{q}_{min} \neq 0 & \text{indirekt} \end{cases} \quad (5.41)$$

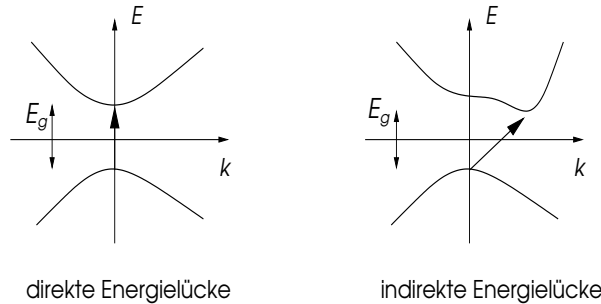


Abb.5.8: Direkte Energielücke: Übergang zwischen Leitungs- und Valenzband mit minimaler Energie ohne Impulstransfer möglich. Indirekte Energielücke: Übergang zwischen Leitungs- und Valenzband mit minimaler Energie nur mit endlichem Impulstransfer möglich.

Die Absorption elektromagnetischer Strahlung in einem Halbleiter entspricht einem Übergang eines Elektrons vom Valenz- zum Leitungsband. Der mögliche Impulsübertrag folgt vom Photon-Impuls bei gegebener Energie. Vergleichen wir diesen mit einem reziproken Gitterimpuls, wobei

wir von einer Photonenergie im Bereich von $\hbar\omega \sim 1eV$ ausgehen, was einer Wellenlänge von $\lambda \sim 1\mu m$ entspricht, so finden wir,

$$ka = 2\pi \frac{a}{\lambda} \sim 10^{-3} \quad \Rightarrow \quad k \ll K = \frac{2\pi}{a} . \quad (5.42)$$

Die Wellenlänge für diesen Frequenzbereich ist viel grösser als die Gitterkonstante a , die von der Grössenordnung $\lesssim 1nm$ ist. Daher dürfen wir annehmen, dass der Impulsübertrag null ist. Somit sind optische Übergänge in Halbleitern mit direkter Energielücke optimal. Um einen Übergang der Energie $\hbar\omega \approx E_g$ in einem Halbleiter mit indirekter Energielücke zu beobachten, muss das Photon zusätzlich mit einer Gitterschwingung, einem Phonon, kombiniert werden.³ Wie Photonen stellen Phononen Energiequanten dar, die von einem elektronischen System absorbiert oder emittiert werden können. Auch sie haben eine Energie und einen Impuls. Da die Dispersion der Phononen wegen der viel kleineren Geschwindigkeit viel flacher ist als bei den Photonen, können Phononen bei kleinem Energieübertrag einen grossen Impuls an ein Elektron weitergeben.

5.4.1 Direkter Übergang

Der Kopplungsterm zwischen Elektronen und elektromagnetischer Strahlung wurde schon in Abschnitt 4.4 und 4.5.3 diskutiert,

$$\hat{V}_{para} = -\frac{e}{mc} \vec{A}(\vec{r}) \cdot \hat{\vec{p}} = -\frac{e}{mc\sqrt{V}} \sum_{\vec{k}, \lambda} \sqrt{\frac{2\pi\hbar c^2}{\omega_{\vec{k}}}} \left\{ \hat{a}_{\vec{k}\lambda} \hat{\vec{p}} \cdot \vec{e}_{\vec{k}\lambda} e^{i\vec{k}\cdot\vec{r}} + \hat{a}_{\vec{k}\lambda}^\dagger \hat{\vec{p}} \cdot \vec{e}_{\vec{k}\lambda}^* e^{-i\vec{k}\cdot\vec{r}} \right\} . \quad (5.47)$$

Im Festkörper ersetzen wir \vec{p}/m durch die Geschwindigkeit. Wir benutzen nun wieder Fermis Goldene Regel, um die Übergangswahrscheinlichkeit zu bestimmen. Der Prozess soll ein Elektron

³Gitterschwingungen können durch Deformationsschwingung eines elastischen Mediums beschrieben werden, das eine Vereinfachung des diskreten Atomgitters darstellt. Die Deformation eines elastischen Mediums führt dazu, dass ein infinitesimales Raumelement d^3r , das sich im undefinierten Zustand am Punkt \vec{r} befand, nun bei $\vec{r}'(\vec{r})$ ist. Damit definieren wir das Auslenkungsfeld $\vec{u}(\vec{r}) = \vec{r}'(\vec{r}) - \vec{r}$ an jedem Punkt im Medium. Das Auslenkungsfeld $\vec{u} = \vec{u}(\vec{r}, t)$ hängt im Allgemeinen von Ort und Zeit ab mit einer kinetischen und einer potentiellen (elastischen) Energie,

$$E_{kin} = \frac{\rho}{2} \int d^3r \left(\frac{\partial \vec{u}}{\partial t} \right)^2 \quad \text{und} \quad E_{el} = \frac{\lambda}{2} \int d^3r \left(\vec{\nabla} \cdot \vec{u} \right)^2 \quad (5.43)$$

wobei wir hier absichtlich eine vereinfachte Form angenommen haben und Scherdeformationen vernachlässigen, da diese in niedrigster Ordnung nicht an die Elektronen koppeln. Daraus lässt sich eine Gleichung für longitudinal polarisierte Wellen herleiten,

$$\frac{1}{c_s^2} \frac{\partial^2}{\partial t^2} \vec{u} - \vec{\nabla}(\vec{\nabla} \cdot \vec{u}) = 0 \quad (5.44)$$

mit der Schallgeschwindigkeit $c_s^2 = \lambda/\rho$, wobei ρ die Massendichte des elastischen Mediums und λ die elastische Konstante (Lamé-Koeffizient) bezeichnet. Die Normalmoden der Wellengleichung sind ebene Wellen (Dichtewellen $\vec{\nabla} \times \vec{u} = 0$)

$$\vec{u}(\vec{r}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{k}} \frac{\vec{k}}{k} \left\{ q_{\vec{k}}(t) e^{i\vec{k}\cdot\vec{r}} + q_{\vec{k}}^*(t) e^{-i\vec{k}\cdot\vec{r}} \right\} \quad \Rightarrow \quad \frac{d^2}{dt^2} q_{\vec{k}} + \omega_{\vec{k}}^2 q_{\vec{k}} = 0 \quad (5.45)$$

wobei $\omega_{\vec{k}} = c_s |\vec{k}| = c_s k$. Analog zu den elektromagnetischen Wellen können wir die Normalmoden (harmonische Oszillatoren) der elastischen Wellen auch quantisieren und erhalten neue Teilchen, die durch Erzeugungs-/Vernichtungsoperatoren in zweiter Quantisierung beschrieben werden:

$$\mathcal{H} = \sum_{\vec{k}} \hbar\omega_{\vec{k}} \left\{ \hat{b}_{\vec{k}}^\dagger \hat{b}_{\vec{k}} + \frac{1}{2} \right\} \quad \text{mit} \quad [\hat{b}_{\vec{k}}, \hat{b}_{\vec{k}'}^\dagger] = \delta_{\vec{k}, \vec{k}'}, \quad [\hat{b}_{\vec{k}}, \hat{b}_{\vec{k}'}] = [\hat{b}_{\vec{k}}^\dagger, \hat{b}_{\vec{k}'}^\dagger] = 0 \quad (5.46)$$

was wie bei den Photonen, der Statistik von Bosonen entspricht mit einer Energie-Impulsbeziehung $\hbar\omega_{\vec{k}} = c_s |\vec{k}|$ wie bei Photonen, die für Wellenlängen viel grösser als der reziproke Gittervektor linear ist. Für $\lambda \sim a$ ergibt sich eine Korrektur. Die Schallgeschwindigkeit ist viel kleiner als die Lichtgeschwindigkeit $c_s \sim 10^{-5}c$ so dass die Phononen bei gleicher Wellenlänge eine viel kleinere Energie besitzen.

mit Impuls $\hbar\vec{k}$ aus dem Valenzband in einen Zustand mit Impuls $\hbar\vec{k}'$ im Leitungsband bringen unter Absorption eines Photons der Energie $\hbar\omega$.

$$\Gamma[(v, \vec{k}), (c, \vec{k}'), \omega] = \frac{2\pi}{\hbar} \left| \langle \phi_f(\vec{k}, \vec{k}'), N_{\vec{q}, \lambda} - 1 | \widehat{V}_{para} | \phi_i, N_{\vec{q}, \lambda} \rangle \right|^2 \delta(\varepsilon_c(\vec{k}') - \varepsilon_v(\vec{k}) - \hbar\omega), \quad (5.48)$$

mit $|\phi_f(\vec{k}, \vec{k}')\rangle$ als elektronischer Endzustand, bei dem aus dem Ausgangszustand $|\phi_i\rangle$ ein Elektron den Bandübergang vollzogen hat. Das Matrixelement hat die Form

$$\left| \langle \phi_f(\vec{k}, \vec{k}'), N_{\vec{q}, \lambda} - 1 | \widehat{V}_{para} | \phi_i, N_{\vec{q}, \lambda} \rangle \right|^2 = \frac{2\pi\hbar e^2}{m^2\omega_{\vec{q}}} \frac{N_{\vec{q}, \lambda}}{V} \left| \vec{e}_{\vec{q}\lambda} \cdot \vec{M}_{cv}(\vec{k}, \vec{k} + \vec{q}) \right|^2 [1 - n_F(\varepsilon_c(\vec{k}'))] n_F(\varepsilon_v(\vec{k})), \quad (5.49)$$

wobei \vec{q} für die Impulserhaltung irrelevant ist und wir annehmen $\vec{k}' = \vec{k}$. Die Faktoren der Fermi-verteilungsfunktionen messen die Wahrscheinlichkeiten, ein Elektron im Valenzband entfernen und auf einen freien Platz im Leitungsband bei gegebenen Energien setzen zu können. Diese Faktoren sind im Wesentlichen 1 und werden von nun an vernachlässigt. Falls der Übergang dipolerlaubt ist, folgt

$$\lim_{\vec{k} \rightarrow 0} \langle |\vec{e}_{\vec{q}\lambda} \cdot \vec{M}_{cv}(\vec{k}, \vec{k})|^2 \rangle_{\vec{q}} = |\mathcal{M}|^2 \neq 0, \quad (5.50)$$

wobei wir eine Mittelung über die Propagationsrichtung und Polarisation der Photonen ausgeführt haben. Wir benutzen nun die Formulierung mit den effektiven Massen, so dass die Delta-Funktion in (5.48) erfüllt wird durch

$$\hbar\omega = \varepsilon_c(\vec{k}) - \varepsilon_v(\vec{k}) = E_g + \frac{\hbar^2 k^2}{2m_c} + \frac{\hbar^2 k^2}{2m_v} = E_g + \frac{\hbar^2 k^2}{2m_r} \quad \text{mit} \quad \frac{1}{m_r} = \frac{1}{m_c} + \frac{1}{m_v}. \quad (5.51)$$

Nehmen wir nun an, dass das Matrixelement nur schwach von \vec{k} abhängt und für gegebene Frequenz die Photonen homogen und isotrop vorhanden sind, dann finden wir für die Absorptionsrate,

$$\begin{aligned} \Gamma_{abs}(\omega) &= \frac{(2\pi)^2 e^2}{m^2\omega} \rho(\omega) N_\gamma(\omega) |\mathcal{M}|^2 \frac{1}{V} \sum_{\vec{k}} \delta\left(\hbar\omega - E_g - \frac{\hbar^2 k^2}{2m_r}\right) \\ &= \frac{e^2}{m^2\omega} \left(\frac{2m_r}{\hbar^2}\right)^{3/2} \rho(\omega) N_\gamma(\omega) |\mathcal{M}|^2 (\hbar\omega - E_g)^{1/2}. \end{aligned} \quad (5.52)$$

Der Faktor $\rho(\omega)$ resultiert von der Integration über alle Richtungen von \vec{q} (Abstrahlrichtung) und den beiden Polarisationen und ist gegeben durch (4.3). $N_\gamma(\omega)$ ist die Zahl der Photonen mit Energie $\hbar\omega$. Falls der Übergang bei $\vec{k} = 0$ dipol-verboten ist, können wir folgende Näherung machen,

$$\lim_{\vec{k} \rightarrow 0} \langle |\vec{e}_{\vec{q}\lambda} \cdot \vec{M}_{cv}(\vec{k}, \vec{k})|^2 \rangle_{\vec{q}} = \langle |\vec{e}_{\vec{q}\lambda} \cdot \vec{k} M_{cv}(\vec{k} = 0)|^2 \rangle_{\vec{q}} |\mathcal{M}'|^2 k^2, \quad (5.53)$$

so dass

$$\Gamma'_{abs}(\omega) = \frac{e^2}{m^2\omega} \left(\frac{2m_r}{\hbar^2}\right)^{5/2} \rho(\omega) N_\gamma(\omega) |\mathcal{M}'|^2 (\hbar\omega - E_g)^{3/2}. \quad (5.54)$$

In beiden Fällen stellt die Energielücke E_g die Schwellenenergie für die Photonen dar. Für dipol-verbotene Übergänge ist die Übergangsrate viel kleiner für Frequenzen, die gerade oberhalb E_g liegen, als für dipol-erlaubte. Beachte, dass die Rate proportional zur Dichte der Photonen bei gegebener Frequenz ist, $\Gamma_{vc}(\omega) \propto N_\gamma(\omega)$ wie wir es in unserer früheren Beschreibung durch die Reaktionsdiffusionsgleichung in Kapitel 1 angesetzt hatten. Dabei sollte man folgende Definitionen machen

$$n_\gamma = \int_{E_g}^{\infty} d\omega \frac{N_\gamma(\omega)}{V} \quad \text{und} \quad \Gamma_1 n_\gamma = \frac{1}{V} \int_{E_g}^{\infty} d\omega \Gamma_{vc}(\omega), \quad (5.55)$$

wobei Γ_1 in (1.12) eingeführt wurde.

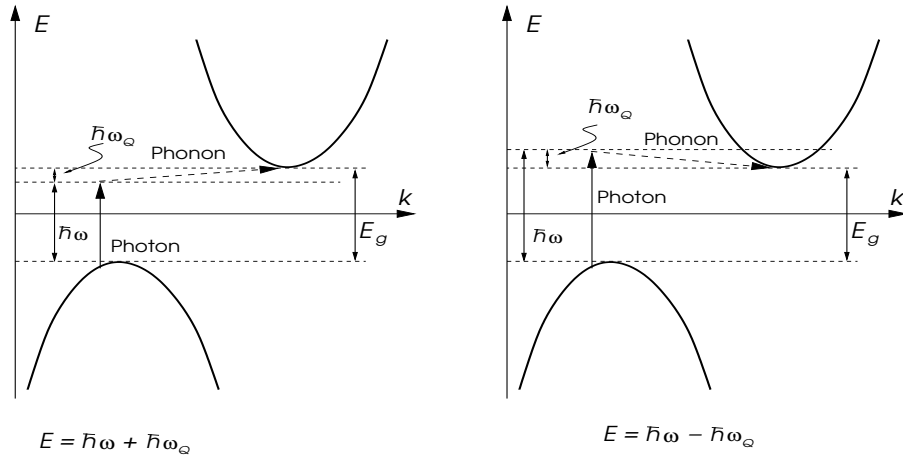


Abb.5.9: Indirekter Übergang unter Beteiligung eines Phonons: Das Photon überträgt den Grossteil der Energie aber praktisch keinen Impuls, während das Phonon den notwendigen Impuls liefert aber für die Energiebilanz weniger wichtig ist. Es gibt zwei Fälle: Phonon überträgt Energie auf Elektron $E = \hbar\omega + \hbar\omega_Q$ (linkes Panel); Phonon absorbiert vom Elektron Energie $E = \hbar\omega - \hbar\omega_Q$.

5.4.2 Indirekter Übergang

Im Falle einer indirekten Energielücke benötigen wir neben dem Energieübertrag durch die Photonen auch noch den Impulsübertrag der Phononen. Der benötigte Impulsübertrag sei \vec{Q} . Wir gehen analog zum obigen Fall vor und schreiben für die Übergangsrate

$$\Gamma_{abs}(\omega) \approx \frac{2\pi}{\hbar} \rho(\omega) N_\gamma(\omega) \sum_{\vec{k}, \vec{k}'} \left\{ |C_+|^2 \delta(\hbar\omega - \varepsilon_c(\vec{k}) + \varepsilon_v(\vec{k}') - \hbar\omega_{\vec{Q}}) + |C_-|^2 \delta(\hbar\omega - \varepsilon_c(\vec{k}) + \varepsilon_v(\vec{k}') + \hbar\omega_{\vec{Q}}) \right\}, \quad (5.56)$$

wobei gilt $\vec{Q} \approx \vec{k} - \vec{k}'$ und $\hbar\omega_{\vec{Q}} = c_s |\vec{Q}| \ll \hbar\omega$. Daher gilt

$$\Gamma_{abs}(\omega) \approx \frac{1}{16\hbar(2\pi)^2} \rho(\omega) N_\gamma(\omega) \left(\frac{2m_c}{\hbar^2} \right)^{3/2} \left(\frac{2m_v}{\hbar^2} \right)^{3/2} \times \left\{ |C_+|^2 (\hbar\omega + \hbar\omega_{\vec{Q}} - E_g)^2 + |C_-|^2 (\hbar\omega - \hbar\omega_{\vec{Q}} - E_g)^2 \right\}, \quad (5.57)$$

wobei wir implizit voraussetzen, dass der Beitrag nur endlich ist, falls der Ausdruck in der Klammer $(\dots)^2$ positiv ist. Offensichtlich ist dieser Übergang für Photonen in der Nähe der Schwellenenergie stark unterdrückt. Der erste und zweite Term entsprechen einer Absorption bzw. Emission eines Phonons. Beachte, dass die Koeffizienten $|C_\pm|^2$ von der Phonon-Konzentration $N_p(\omega)$ abhängen.⁴ Der indirekte Übergang benötigt das Zusammenspiel von Photon und Phonon, was die Angelegenheit erheblich erschwert. Daher sind die entsprechenden Matrixelemente oft mehrere Größenordnungen kleiner als bei direkten Übergängen.

⁴Analog zur Photon-Absorption/Emission gilt:

$$|C_+|^2 \propto N_p(\omega_{\vec{Q}}) = \frac{1}{e^{\hbar\omega_{\vec{Q}}/k_B T} - 1} \quad (5.58)$$

$$|C_-|^2 \propto N_p(\omega_{\vec{Q}}) + 1 = \frac{1}{1 - e^{-\hbar\omega_{\vec{Q}}/k_B T}}$$

wobei bei der Emission neben der stimulierten auch die spontane Emission dazukommt.

5.5 Rekombination

Die erzeugten Elektron-Loch-Paare können auch wieder zerfallen. Dafür gibt es verschiedene Möglichkeiten. Wir betrachten hier die strahlenden Rekombinationen, d.h. diejenigen die zur Emission eines Photons führen. Die Übergangsraten können aus den oben berechneten Ergebnissen bestimmt werden. Betrachten wir wieder den direkten Übergang zuerst,

$$\begin{aligned}\Gamma_{cv}[(v\vec{k}), (c\vec{k}'), \omega] &\approx \frac{(2\pi)^2 e^2}{m^2 \omega} \frac{N_{\vec{q}\lambda} + 1}{V} |\mathcal{M}|^2 [1 - n_F(\varepsilon_v(\vec{k}))] n_F(\varepsilon_c(\vec{k})) \delta\left(\hbar\omega - E_g - \frac{\hbar^2 k^2}{2m_r}\right) \\ &\approx \frac{(2\pi)^2 e^2}{m^2 \omega} \frac{N_{\vec{q}\lambda} + 1}{V} |\mathcal{M}|^2 e^{-E_g/k_B T} e^{-\hbar^2 k^2/2m_r k_B T} \delta\left(\hbar\omega - E_g - \frac{\hbar^2 k^2}{2m_r}\right)\end{aligned}\quad (5.59)$$

Der Faktor $N_{\vec{q}\lambda} + 1$ enthält die stimulierte und die spontane Emission. Offensichtlich sind diese Übergänge von der Verfügbarkeit von Elektron-Loch-Paaren abhängig, wie wir dies in Kapitel 1 gesehen haben. Betrachten wir die gesamte Rekombinationswahrscheinlichkeit für gegebene Photonfrequenz ω ,

$$\Gamma_{em}(\omega) = \sum_{\vec{k}} \Gamma_{cv}[(v\vec{k}), (c\vec{k}), \omega] = \frac{e^2}{m^2 \omega} \rho(\omega) (N_\gamma(\omega) + 1) |\mathcal{M}|^2 \left(\frac{2m_r}{\hbar^2}\right)^{3/2} \sqrt{\hbar\omega - E_g} e^{-\hbar\omega/k_B T} \quad (5.60)$$

für die dipolerlaubten Übergänge bei $\vec{k} = 0$. Für dipolverbotene Übergänge erhalten wir analog

$$\Gamma_{em}(\omega) = \frac{e^2}{m^2 \omega} \rho(\omega) (N_\gamma(\omega) + 1) |\mathcal{M}'|^2 \left(\frac{2m_r}{\hbar^2}\right)^{5/2} (\hbar\omega - E_g)^{3/2} e^{-\hbar\omega/k_B T}. \quad (5.61)$$

In beiden Fällen werden Photonen nur oberhalb der Energielücke emittiert und Emissionsrate wird im Bereich hoher Frequenzen exponentiell unterdrückt. Der dipolerlaubte Übergang zeigt eine deutlich stärkere Emission aufgrund des steilen Anstieges bei $\hbar\omega \approx E_g$. Beachte, dass der Beitrag der stimulierten Emission verhältnismässig klein ist, denn $\hbar\omega \gg k_B T$ bei Raumtemperatur, also $N_\gamma(\omega) \ll 1$, da $N_\gamma = (e^{\hbar\omega/k_B T} - 1)^{-1}$.

Nun wenden wir uns den Halbleitern mit indirekter Energielücke zu. Es ist relativ einfach aus der obigen Diskussion herzuleiten, dass in diesem Fall

$$\begin{aligned}\Gamma_{em}(\omega) &\approx \frac{1}{16\hbar(2\pi)^2} \rho(\omega) (N_\gamma(\omega) + 1) \left(\frac{2m_c}{\hbar^2}\right)^{3/2} \left(\frac{2m_v}{\hbar^2}\right)^{3/2} \\ &\quad \times \left\{ |C_+|^2 (\hbar\omega + \hbar\omega_{\vec{Q}} - E_g)^2 + |C_-|^2 (\hbar\omega - \hbar\omega_{\vec{Q}} - E_g)^2 \right\} e^{-\hbar\omega/k_B T}.\end{aligned}\quad (5.62)$$

Dabei ist wieder zu berücksichtigen, dass die beiden Terme nur beitragen, wenn das Argument in $(\dots)^2$ positiv ist. Zudem ist auch die Besetzung der Phononen wichtig. Da die charakteristische Energie vergleichbar mit der Raumtemperatur ist, steht immer eine grosse Zahl von Phononen zur Verfügung.

Wie schon vorher erwähnt sind die Matrixelemente bei indirekten Übergänge viel kleiner als im Fall direkter Übergänge. Somit eignen sich Halbleiter mit direkter Energielücke viel besser für Lumineszenz-Dioden (LED), insbesondere bei dipolerlaubten Übergängen. Die Emission hat ein klares Maximum leicht oberhalb der Energielücke. GaAs und seine Verwandten sind Halbleiter mit einer direkten Energielücke, während Si und Ge eine indirekte Energielücke besitzen.

Es gibt noch andere Rekombinationsmöglichkeiten. Falls die Konzentration der Elektronen im Leitungsband genügend gross ist, so dass die Wechselwirkung zwischen den Elektronen eine Rolle spielt, dann kann das Elektron-Loch-Rekombination auch ohne Abgabe eines Photons geschehen, indem die Energie an ein Elektron abgegeben wird. Dieser Prozess wird Auger-Rekombination genannt und ist intrinsisch.

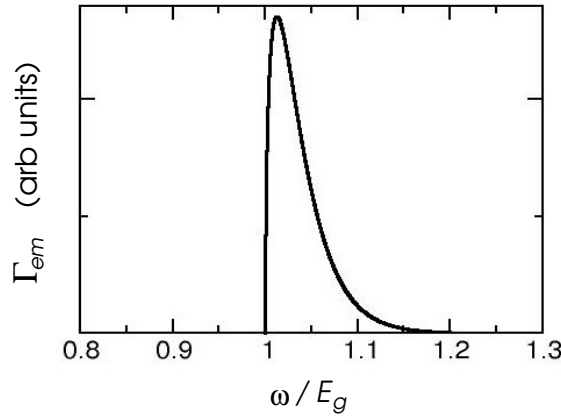


Abb.5.10: Frequenzabhängigkeit der strahlenden Rekombinationsrate für einen Halbleiter mit direkter Energielücke.

Es existieren auch extrinsische Rekombinationskanäle, zum Beispiel über Verunreinigungen oder Oberflächen. Verunreinigungen können zu lokalisierten Zuständen in der Energielücke führen und Elektron-Loch-Paar können dort eingefangen werden. Diese Übergangsraten sind schneller als die oben diskutierten strahlenden Übergänge. Die Konzentration der Verunreinigungen spielt aber auch eine wichtige Rolle.

5.6 Dotierung von Halbleitern

Wir haben im ersten Kapitel bereits über die Dotierung der Halbleiter gesprochen. Wenn wir in einem Si-Halbleiter Si-Atome durch Atome aus der Gruppe III im Periodensystem ersetzen, z.B. Al, oder durch Atome aus der Gruppe V, z.B. P, ersetzen, dann führt dies zu Störstellen im Kristallgitter, die zusätzliche Ladungsträger einführen. Betrachten wir die Elektronenkonfigurationen,

$$\text{Si: } (1s)^2(2s)^2(2p)^6(3s)^2(3p)^2 \quad \left\{ \begin{array}{l} \text{Al: } (1s)^2(2s)^2(2p)^6(3s)^2(3p) = [\text{Si}] + 1 \text{ Loch} \\ \text{P: } (1s)^2(2s)^2(2p)^6(3s)^2(3p)^3 = [\text{Si}] + 1 \text{ Elektron} \end{array} \right. \quad (5.63)$$

Al agiert als Akzeptor und P als Donator.

5.6.1 Wannier-Gleichung

Schauen wir uns die P-Störstelle genauer an. Das zusätzliche Elektron hat die effektive kinetische Energie eines Leitungsband-Elektrons,

$$\varepsilon_c(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m_c} \quad (5.64)$$

Der Ionenrumpf besitzt jedoch die positive Ionenladung $+e$, so dass das überschüssige Elektron gebunden werden kann. Die Wannier-Gleichung liefert eine einfache Beschreibung für die Wellenfunktion $\psi(\vec{r})$ dieses Elektrons,

$$\left\{ -\frac{\hbar^2 \vec{\nabla}^2}{2m_c} - \frac{e^2}{\epsilon |\vec{r}|} \right\} \psi(\vec{r}) = E\psi(\vec{r}), \quad (5.65)$$

wobei ϵ die Dielektrizitätskonstante des Halbleiters ist, die die Abschirmeffekte des Halbleiters beschreibt (ein Halbleiter ist ein Dielektrikum). Diese Gleichung hat die gleiche Form wie die

Schrödinger-Gleichung für das Wasserstoff-Atom und folglich auch das gleiche Spektrum, jedoch mit renormierten Parametern. Wir finden die Bindungsenergien,

$$E_n = -\frac{m_c e^4}{2\hbar^2 \epsilon^2 n^2} = -\frac{m_c}{m_c \epsilon^2 n^2} Ry . \quad (5.66)$$

Der effektive Bohrradius ist gegeben durch

$$a_B^* = \frac{\hbar^2 \epsilon}{m_c e^2} = \frac{\epsilon m}{m_c} a_B . \quad (5.67)$$

Für Si haben wir die Parameter $m_c \approx 0.2m$ und $\epsilon \approx 12$ ($1Ry = 13.6eV$ und $a_B = 0.53\text{\AA}$), so dass

$$E_1 \approx -20meV \quad \text{und} \quad a_B^* \approx 30\text{\AA} . \quad (5.68)$$

Damit besitzen wir einen schwach gebundenen elektronischen Zustand, gemessen vom unteren Rand des Leitungsbandes, der in der Energielücke liegt. Eine ähnliche Betrachtung kann für Al als Störstelle durchgeführt werden, wobei wir Energiezustände für ein gebundenes Loch unmittelbar oberhalb des oberen Randes des Valenzbandes sehen.

Diese Bindungsenergie entspricht einer Temperatur von 230 K. Damit können wir diese effektiven "Wasserstoff-Atome" bei Raumtemperatur als grossenteils ionisiert betrachten, so dass die P-Atome ihre Elektronen an das Leitungsband ($n_e \sim n_d$) und die Al-Atome ihre Löcher an das Valenzband ($n_h \sim n_A$) abgeben.

5.6.2 Chemisches Potential

Wir betrachten eine Konzentration von n_D Donatoren. Die Konzentration der Elektronen im Leitungsband entspricht näherungsweise (5.35),

$$n_e \approx \frac{e^{-(E_g - \mu)/k_B T}}{\pi \sqrt{2\pi}} \left(\frac{m_c k_B T}{\hbar^2} \right)^{3/2} = N_c e^{-(E_g - \mu)/k_B T} \quad (5.69)$$

mit $E_g > \mu$ und die Konzentration der Elektronen auf den Donatoren ist gegeben durch

$$n_d = \frac{n_D}{e^{(E_D - \mu)/k_B T} + 1} \quad \Rightarrow \quad n_D \approx n_d + n_e . \quad (5.70)$$

Daraus folgt

$$e^{\mu/k_B T} = \frac{n_e}{N_c} e^{E_g/k_B T} \quad (5.71)$$

woraus mit etwas Rechnung folgt, dass

$$\begin{aligned} \mu &= E_D + k_B T \ln \left(\frac{n_D}{n_e} - 1 \right) > E_D \\ \mu &= E_g + k_B T \ln \left(\frac{n_e}{N_c} \right) < E_g \end{aligned} \quad (5.72)$$

Damit liegt das chemische Potential zwischen E_D und E_g . Bei Temperatur $T = 0K$ liegt das chemische Potential zwischen der Grundzustandsenergie und der Energie des ersten angeregten Zustandes der Störstelle. Die analoge Überlegung gilt für Akzeptoren, so dass

$$\mu = -k_B T \ln \left(\frac{n_h}{N_v} \right) > 0 \quad \text{mit} \quad N_v = \frac{1}{\pi \sqrt{2\pi}} \left(\frac{m_v k_B T}{\hbar^2} \right)^{3/2} , \quad (5.73)$$

und es gilt $\mu < E_A$.

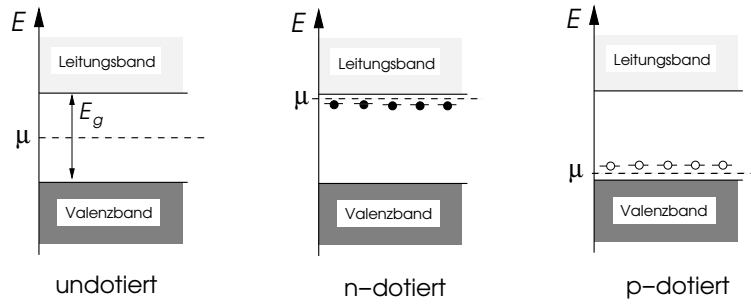


Abb.5.11: Position des chemisches Potentials.

5.7 pn-Kontakt

Wenn wir nun einen Kontakt zwischen einem n- und einem p-dotierten Halbleiter herstellen, verlangt das Gleichgewicht, dass sich das chemische Potential auf beiden Seiten angleicht. Dies geschieht durch Einführung einer Differenz im elektrostatischen Potentials für die Elektronen,

$$\eta^p = \mu^p - e\Phi^p = \eta^n = \mu^n - e\Phi^n, \quad (5.74)$$

wobei wir hier wieder das elektrochemische Potential eingeführt haben, das wir schon in Kapitel 1 und 2 angetroffen haben. Wir gehen nun davon aus, dass auf der p-Seite $n_e \sim n_D$ und auf der n-Seite $n_h \sim n_A$, dann erhalten wir aus obigen Gleichungen für das chemische Potential,

$$e(\Phi^p - \Phi^n) = \mu^p - \mu^n = E_g + k_B T \ln \left(\frac{n_D n_A}{N_v N_c} \right) = k_B T \ln \left(\frac{n_D n_A}{n_i(T)^2} \right), \quad (5.75)$$

wie wir es in Kapitel 2 gefunden hatten. Die Potentialdifferenz rührt vom Raumladungsdipol her, der sich am Kontakt bildet, wie aus Abb.5.11/5.12 ersichtlich ist. Dazu verbiegen sich die Bänder in der Übergangsregion. Da die entsprechenden Störstellen in der Raumladungszone ihre Elektronen oder Löcher verloren haben, muss die Ladungsdichte der Akzeptor/Donator-Dichte entsprechenden, wie wir es in (2.7) angenommen hatten: $\rho \approx -en_D$ oder en_A im jeweiligen Bereich.

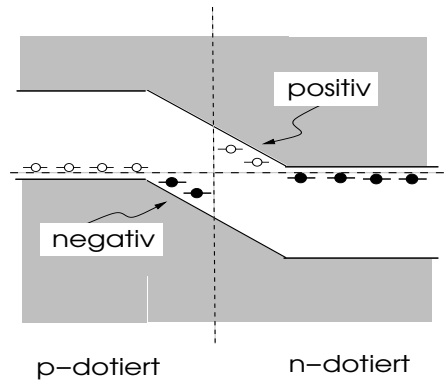


Abb.5.12: Schematisches Verhalten der Energielücke am pn-Kontakt. Aufbau eines Ladungsdipols im Bereich, in dem sich die Störstellenniveaus durch das chemische Potential verschieben.

Damit sind die wesentlichen Elemente des pn-Kontaktes durch die mikroskopische Betrachtungsweise beschrieben. Die Eigenschaften des pn-Kontaktes folgt nun aus der Diskussion in Kapitel 2.

5.8 Metall-Halbleiter-Kontakte

Betrachten wir nun noch den Kontakt zwischen einem Metall und einem Halbleiter, den wir als n-dotiert annehmen. Wiederum geht es um das Angleichen der elektrochemischen Potentiale. Zudem ist es wichtig hier die Austrittsarbeit der Elektronen miteinzubeziehen, die in beiden Systemen sehr unterschiedlich sein kann.

Die Austrittsarbeit $-e\Phi_M$ des Metalls entspricht der minimalen Energie, die benötigt wird um ein Elektron an der Fermienergie zu entnehmen und ins Vakuum zu befördern. Für den n-dotierten Halbleiter ist die Austrittsarbeit $-e\Phi_H$ die minimale Energie ein Elektron von einer Störstelle ins Vakuum zu bringen. Diese Energie ist etwas grösser als die Energie $-e\Phi_c$, die Differenz zwischen dem untersten Vakuumniveau und unteren Bandrand des Leitungsbandes. Wenn wir die elektrochemischen Potentiale der beiden Seiten angleichen, dann findet in der Kontaktregion wieder eine Verbiegung der Bandstruktur statt, wie in Abb.5.13 gezeigt.

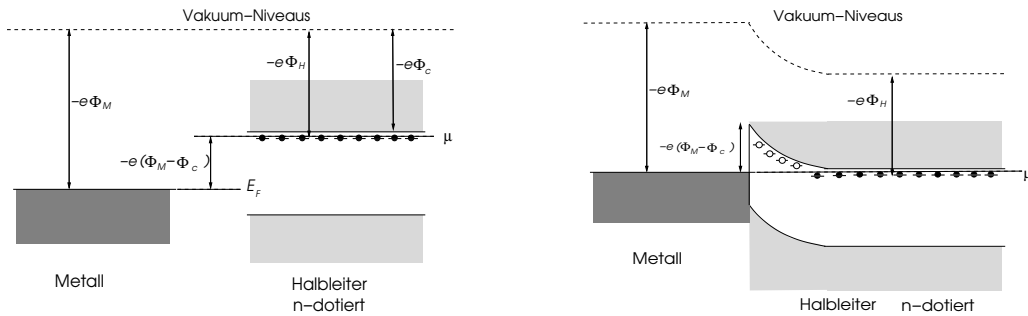


Abb.5.13: Schematisches Struktur des Metall-Halbleiter-Kontaktes: Links: Energie-Niveau-Schema des Metalls und des n-dotierten Halbleiters; Rechts: Im Kontakt gleicht sich das elektrochemische Potential an, indem sich ein Ladungsdipol ausbildet. An der Grenze findet man eine Potentialbarriere für die Elektronen, die Schottky-Barriere: $-e(\Phi_M - \Phi_c)$.

5.8.1 Raumladung

Durch die Verbiegung der Bänder werden die Störstellen in der Nähe des Kontaktes entleert, da sie über dem chemischen Potential zu liegen kommen. Damit finden wir dort eine positive Ladungsdichte $\rho = en_D$ auf einer Länge l_n . Betrachten wir das entsprechende elektrostatische Potential auf der Halbleiterseite,

$$\frac{d^2\Phi}{dx^2} = -\frac{4\pi en_D}{\epsilon} \quad \Rightarrow \quad \Phi(x) = \begin{cases} -\frac{2\pi en_d}{\epsilon}(x - l_n)^2 + \Phi_H, & 0 \leq x < l_n \\ \Phi_H & l_n < x \end{cases} \quad (5.76)$$

Auf der Seite des Metalls kommt der Abschirmungseffekt der beweglichen Elektronen ins Spiel. Die entsprechende Gleichung für $x < 0$ lautet,

$$\frac{d^2\Phi}{dx^2} = \frac{\Phi - \Phi_M}{l_{TF}^2} \quad \Rightarrow \quad \Phi(x) = \Phi_0 e^{x/l_{TF}} + \Phi_M, \quad (5.77)$$

wobei l_{TF} die Thomas-Fermi-Abschirmlänge bezeichnet.⁵ Das Potential und seine Ableitung müssen stetig sein bei $x = 0$, so dass folgt,

$$\left. \begin{aligned} \Phi_0 + \Phi_M &= -2\pi en_D l_n^2 + \Phi_H \\ \frac{\Phi_0}{l_{TF}} &= 4\pi en_D l_n \end{aligned} \right\} \Rightarrow \left\{ \begin{aligned} l_n &\approx \sqrt{\frac{\Phi_M - \Phi_H}{4\pi en_D}} \\ \Phi_0 &\approx l_{TF} \sqrt{4\pi en_D (\Phi_M - \Phi_H)} \end{aligned} \right. . \quad (5.82)$$

Beachte, dass $l_{TF} \ll l_n$ und $\Phi_0 \ll \Phi_M - \Phi_H$. Das bedeutet, dass die Ladungsschicht im Metall wesentlich dünner ist als im Halbleiter.

5.8.2 Strom-Spannungscharakteristik

Im Gegensatz zu den Halbleiter pn-Kontakten ist der Ladungstransport im Metall-Halbleiter-Kontakt von den Majoritätsladungsträgern dominiert, d.h. in unserem Beispiel durch die Elektronen. Dabei ist es wichtig, dass die Schottky-Barriere viel höher als die Temperatur ist: $E_S = -e(\Phi_M - \Phi_c) \gg k_B T$. Betrachten wir den Strom der Elektronen vom Halbleiter zum Metall. Elektronen können thermisch ins Metall emittiert werden, falls sie die Barriere überwinden können (wir vernachlässigen hier quantenmechanisches Tunneln und betrachten die Elektronen im Halbleiter wieder als Gas klassischer Teilchen). Dies geschieht durch thermische Anregung. Dafür kommen Elektronen im Leitungsband mit der Eigenschaft

$$\frac{\hbar^2 k_x^2}{2m_c} > E_S - E_c + \mu - eU \quad (5.83)$$

in Frage, wobei U die angelegte Spannung bezeichnet und die Energie der Leitungsbandelektronen durch $\hbar^2 k^2/2m_c + E_c$ definiert ist. Für die Stromdichte, die der Diffusionsstromdichte

⁵Betrachten wir eine externe Ladungsdichte $\rho_{ex}(\vec{r})$. Durch die Coulomb-Wechselwirkung wird diese Ladung die Verteilung der Elektronen im Metall verändern und eine induzierte Ladung erzeugen, die die Abweichung von der uniformen Elektronendichte $n_e(E_F)$ entspricht (beachte neutralisierende positiv geladenes Ionengitter). Diese induzierte Ladung können wir in der sogenannten Thomas-Fermi Näherung schreiben als

$$\rho_{ind}(\vec{r}) = -e \{n_e(E_F + e\Phi(\vec{r})) - n_e(E_F)\} \quad \text{mit} \quad n(E_F) = \frac{k_F^3}{3\pi^2} = \frac{1}{3\pi^2 \hbar^2} (2mE_F)^{3/2} \quad (5.78)$$

wobei $E_F = \hbar^2 k_F^2/2m$. Bei dieser Näherung wird angenommen, dass das Potential $\Phi(\vec{r})$ sich im Vergleich zu k_F^{-1} im Raum langsam verändert, so dass lokal das Elektronengas durch eine gefüllte Fermikugel beschrieben werden kann. Die Poissongleichung erhält nun die Form

$$\begin{aligned} \nabla^2 \Phi(\vec{r}) &= -4\pi [\rho_{ind}(\vec{r}) + \rho_{ex}(\vec{r})] \approx 4\pi e^2 \left. \frac{\partial n(E)}{\partial E} \right|_{E=E_F} \Phi(\vec{r}) - 4\pi \rho_{ex}(\vec{r}) \\ &= \frac{1}{l_{TF}^2} \Phi(\vec{r}) - 4\pi \rho_{ex}(\vec{r}) \end{aligned} \quad (5.79)$$

mit der Thomas-Fermi-Abschirmlänge l_{TF} definiert durch,

$$l_{TF}^{-2} = 4\pi e^2 \left. \frac{\partial n(E)}{\partial E} \right|_{E=E_F} = \frac{6\pi e^2 n_e}{E_F} . \quad (5.80)$$

mit $n_e = n_e(E_F)$ als die uniforme Elektronendichte. In einem einfachen Metall ist diese Länge einige Angström. Für eine externe Punktladung q am Ursprung erhalten wir als Lösung für das Potential

$$\Phi(\vec{r}) = q \frac{e^{-r/l_{TF}}}{r} . \quad (5.81)$$

Das Potential hat nicht mehr Coulomb-Form, sondern ist auf der Länge l_{TF} abgeschirmt und hat die Form eines Yukawa-Potentials.

entspricht, gilt daher

$$\begin{aligned}
j_{H \rightarrow M} &= 2 \int \frac{d^3k}{(2\pi)^3} \Theta \left(\frac{\hbar^2 k_x^2}{2m_c} - (E_S - E_c + \mu - eU) \right) e \frac{\hbar k_x}{m_c} \exp \left\{ -\frac{1}{k_B T} \left(\frac{\hbar^2 k^2}{2m_c} + E_c - \mu \right) \right\} \\
&= \frac{2em_c k_B T}{4\pi^2 \hbar^3} \int_{E_S - E_c + \mu - eU} d \left(\frac{\hbar^2 k_x^2}{2m_c} \right) \exp \left\{ -\frac{1}{k_B T} \left(\frac{\hbar^2 k_x^2}{2m_c} + E_c - \mu \right) \right\} \\
&= \frac{2m_c e (k_B T)^2}{4\pi^2 \hbar^3} e^{-(E_S - eU)/k_B T}
\end{aligned} \tag{5.84}$$

Der Strom der vom Metall in den Halbleiter fließt hat die Eigenschaft des Driftstromes, indem nämlich ein Elektron das die Barriere überquert in den n-dotierten Halbleiter abgesogen wird. Für verschwindende Spannung U muss dieser Strom gerade den Strom $j_{H \rightarrow M}$ aufheben als Gleichgewichtsbedingung: $j_{H \rightarrow M}(U = 0) = j_{M \rightarrow H}$. Damit ergibt sich die Strom-Spannungscharakteristik

$$j = j_0 \left\{ e^{eU/k_B T} - 1 \right\} \quad \text{mit} \quad j_0(T) = A (k_B T)^2 e^{-E_S/k_B T} . \tag{5.85}$$

Auch ein Metall-Halbleiter-Kontakt eignet sich als Gleichrichter-Diode. Man nennt sie auch Schottky-Diode. Im Prinzip gibt es auch noch den Strom der Minoritätsladungsträger, im gegebenen Beispiel sind es die Löcher. Diese Verhalten sich wie bei einem pn-Kontakt. Der Beitrag zur Strom-Spannung-Charakteristik ist jedoch sehr klein verglichen mit dem Majoritätsladungsanteil. Daher ist der Metall-Halbleiter-Kontakt eine sogenannte unipolares Element. Die Schottky-Diode reagiert daher auch auf kürzeren Zeitskalen als die pn-Diode.

Auch die frühen photoelektrischen Zellen, die aus metalbeschichtetem Selen bestanden, können auf die beschriebenen Eigenschaften des Metall-Halbleiter-Kontaktes zurückgeführt werden. Löcher, die durch Photonabsorption auf der Halbleiter-Seite erzeugt werden und durch Diffusion die Raumladungsschicht erreichen, werden durch den Ladungsdipol sofort auf die Metallseite hinübergesogen. Damit entsteht ein Strom analog zur Situation im pn-Kontakt.

Kapitel 6

Physik der globalen Erwärmung und des Treibhauseffekts

Abschliessend möchten wir uns noch einem anderen Aspekt der Sonnenenergie zuwenden, der in letzter Zeit sehr stark ins öffentliche Interesse gerückt ist. Sonnenenergie spielt nicht nur im Zusammenhang mit alternativen Energiequellen eine wichtige Rolle, sondern ist die Basis für Leben überhaupt auf unserem Planeten. Sie ist verantwortlich für das Klima, das auf der Erde vorherrscht und die Existenz der Biosphäre. Die wichtigste Energiequelle für die Erde ist die Strahlung der Sonne und die wichtigste Energiesenke ist die elektromagnetische Strahlung der Erde ins All. Die Eigenwärme, die aus dem Inneren der Erde an die Oberfläche tritt, spielt in dieser Energiebilanz eine sekundäre Rolle. In diesem Kapitel wollen wir betrachten, welche Faktoren die Energiebilanz der Erde mitbestimmen. In diesem Zusammenhang werden wir auch auf den Treibhaus-Effekt zu sprechen kommen. Das Klima eines Planeten ist jedoch eine äusserst komplexe Angelegenheit und kann hier nur in sehr vereinfachter Form behandelt werden.

6.1 Energiebilanz - Kirchhoff'sches Gesetz

Das Kirchhoff'sche Gesetz sagt aus, dass ein System im Gleichgewichtszustand die gleiche Energiemenge emittieren wie absorbieren muss. Ansonsten würde sich das System entweder abkühlen oder erwärmen und wäre nicht im thermischen Gleichgewicht, denn die Temperatur T ist eine Gleichgewichtsvariable.

Analysieren wir nun ein Beispiel, bei dem der Gleichgewichtszustand durch Strahlungsausgleich aufrechterhalten wird. Wir betrachten einen kugelförmigen Hohlraum, dessen Oberfläche wie ein schwarzer Strahler wirkt. Darin schwebt frei und konzentrisch eine Kugel, die ein "grauer Körper" ist, d.h. sie absorbiert nur einen Teil der Strahlung während der Rest reflektiert wird. Die Oberfläche der inneren Kugel (1) habe den Radius R_1 und die Innenfläche des Hohlraumes (2) hat den Radius R_2 . Der Hohlraum sei nach aussen vollständig isoliert, so dass wir ein abgeschlossenes System betrachten und ein Energiefluss nur zwischen den beiden Flächen stattfindet.

Gemäss dem Stefan-Boltzmann-Gesetz (4.9) strahlt ein schwarzer Körper der Temperatur T Energie im ganzen Spektrum mit der Leistung pro Fläche

$$S(T) = \sigma T^4 \quad \text{mit} \quad \sigma = 5.6705 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4}. \quad (6.1)$$

Zunächst behaupten wir, dass unabhängig von den Absorptionseigenschaften (Reflexionseigenschaften) des grauen Körpers, im Gleichgewichtszustand beide Kugeln dieselbe Temperatur annehmen. Falls dies nämlich nicht so wäre, könnten wir einen materiellen Wärmekontakt (Draht, ...) zwischen den beiden Kugeln anbringen, so dass Wärmeenergie von der heisseren in die kältere Kugel fliesst (Fick'sches Gesetz). Um das Gleichgewicht aufrecht zu erhalten, müsste dann Energie durch Strahlung von der kälteren zur wärmeren Kugel übertragen werden. Dies widerspricht jedoch dem zweiten Hauptsatz.

Der Energietransport durch Strahlung von Kugel (1) nach (2) ist bestimmt durch

$$\left. \frac{dE}{dt} \right|_{1 \rightarrow 2} = A_1 a_1 S_1 \chi_{1 \rightarrow 2} - A_2 (1 - r_1) S_2 \chi_{2 \rightarrow 1} \quad (6.2)$$

wobei die Flächen $A_i = 4\pi R_i^2$ eingehen. Ferner ist a_1 der Bruchteil, der auf Kugeloberfläche (1) absorbiert wird, und r_1 der Bruchteil, der reflektiert wird, wobei gelten muss, dass $1 = a_1 + r_1$ mit $0 \leq a_1, r_1 \leq 1$. Die beiden Faktoren $\chi_{1 \rightarrow 2}$ und $\chi_{2 \rightarrow 1}$ sind geometrische Konstanten, die den Anteil der Energie angeben, der von (1) ausgehend (2) erreicht, bzw. umgekehrt. Für unseren Fall gilt, dass $\chi_{1 \rightarrow 2} = 1$. Das Kirchhoff'sche Gesetz sagt aus, dass die Kugel (1) nur so viel Energie emittieren kann, wie sie absorbiert. Daher ist der Vorfaktor des ersten Terms auf der rechten Seite a_1 . Der zweite Term beschreibt den Teil der Strahlung von (2) nach (1), der die Oberfläche von (1) erreicht und nicht reflektiert wird, was sich durch den Faktor $1 - r_1$ ausdrücken lässt. Dabei kommt die Vereinfachung ins Spiel, dass die gesamte von Kugel (1) reflektierte Strahlung (unabhängig von der Art der Streuung) die Kugel (2) erreicht und dort wieder vollständig absorbiert wird. Ansonsten müssten wir noch das Problem der Mehrfachreflexion mitberücksichtigen. Aus (6.2) folgt, dass im Gleichgewicht (verschwindender Energietransport) unabhängig von a_1 gilt,

$$A_1 S_1 = A_2 \chi_{2 \rightarrow 1} S_2. \quad (6.3)$$

Für identische Temperaturen, $T_1 = T_2$, gilt gemäss (6.1) $S_1 = S_2$ und somit folgt $\chi_{2 \rightarrow 1} = A_1/A_2 = R_1^2/R_2^2$.¹ Für das abgeschlossene System, das wir hier analysiert haben, findet kein Energieaustausch mit der Aussenwelt statt. Dies ist anders, wenn wir uns die Situation der Erde ansehen, die einem offenen thermodynamischen System entspricht.

¹Der Konstanten $\chi_{1 \rightarrow 2}$ und $\chi_{2 \rightarrow 1}$ lassen sich durch einfache geometrische Betrachtungen ableiten. Von einem Punkt A auf der Kugelinnenfläche (2) wird in alle Richtung abgestrahlt. Es gibt einen Bereich des Abstrahlungswinkels α , für den der Strahl wieder auf der Kugeloberfläche (2) auf einem Punkt B landet. Wir bestimmen α_0 als kleinsten Winkel, bei dem der Strahl von A ausgehend die Kugel (1) nicht trifft: $\sin \alpha_0 = R_1/R_2$. Aus Abb. 6.1 ist ersichtlich, dass $\theta = \pi - 2\alpha_0$, der Winkel vom Zentrum zu den Punkten A und B .

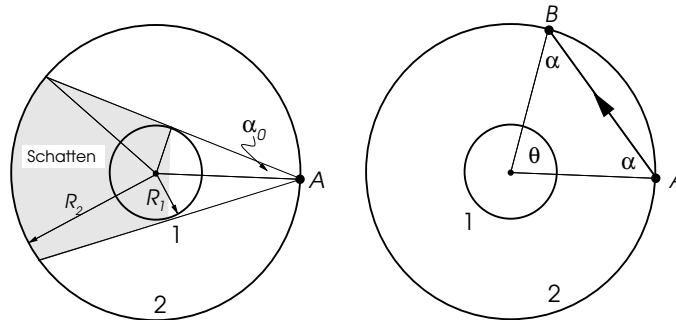


Abb.6.1: Geometrie des Strahlungsübertrags von Kugel (2) nach (1).

Der Schatten, den Kugel (1) für die Strahlen von einem Punkt A auf Kugel (2) wirft, verglichen mit der ganzen Innenfläche der Kugel (2) entspricht $\chi_{2 \rightarrow 1}$. Es gilt nun

$$I_{2 \rightarrow 1}(\alpha_0) = \int_{\text{Schatten}} d\phi d\theta \sin \theta = 2\pi \int_{\theta_0}^{\pi} d\theta \sin \theta = -2\pi \cos \theta \Big|_{\theta_0=\pi-2\alpha_0}^{\pi} = 4\pi \sin^2 \alpha_0 = 4\pi \frac{R_1^2}{R_2^2}. \quad (6.4)$$

Für die ganze Intensität erhalten wir $I_{2 \rightarrow 1}(\alpha_0 = \pi/2) = 4\pi$, so dass

$$\chi_{2 \rightarrow 1} = I_{2 \rightarrow 1}(\alpha_0)/I_{2 \rightarrow 1}(\pi/2) = R_1^2/R_2^2. \quad (6.5)$$

Andererseits ist offensichtlich, dass jeder Strahl von jedem Punkt auf Kugel (1) die Kugel (2) erreicht: $\chi_{1 \rightarrow 2} = 1$. Aus (6.3) folgt $S_1 = S_2$, so dass in umgekehrter Richtung auf $T_1 = T_2$ geschlossen werden kann.

6.2 Einfachstes Modell der Erdenergie-Bilanz

Die Sonne liefert der Erde eine enorme Energiemenge, die wir als Solarkonstante bezeichnen.

$$S_0 = 1370 \text{ Wm}^{-2} . \quad (6.6)$$

Für die Gesamtleistung, die die Erde trifft, erhalten wir folglich

$$P_0 = S_0 \pi R_E^2 = 1.66 \times 10^{17} \text{ W} , \quad (6.7)$$

bei einem Radius von $R_E = 6370 \text{ km}$ unter Berücksichtigung, dass die Erde nur von einer Seite von der Sonne beleuchtet wird. Im Vergleich dazu ist der mittlere Energieverbrauch der gesamten Weltbevölkerung mit $\sim 1.5 \times 10^{13} \text{ W}$ um zirka 4 Grössenordnungen kleiner.

Um ein thermisches Gleichgewicht zu erreichen und nicht stetig zu erhitzen, muss die Erde die Energie, die die Sonne liefert, wieder abbauen. Die einzige Möglichkeit dies zu tun ist die Abstrahlung ins Weltall, abgesehen von der Umsetzung der Sonnenenergie in andere Energieformen, wie z.B. in Biomasse, fossile Brennstoffe, Solar-Elektrizität, ... und auch in latente Wärme bei Aggregatzustandsänderungen (Eis \leftrightarrow Wasser \leftrightarrow Dampf). Anders als das vorhin betrachtete Beispiel stellt die Konstellation Sonne und Erde kein geschlossenes System dar. Ein wichtiger Aspekt ist die Tatsache, dass die Sonne durch Nuklearprozesse Energie produziert. Aber nur ein kleiner Teil dieser Energie erreicht die Erde. Das meiste geht ins Weltall. Auch ist der Energiefluss von der Erde zur Sonne irrelevant, was nur schon aus geometrischen Gründen ersichtlich wird: $\chi_{S \rightarrow E}, \chi_{E \rightarrow S} \ll 1$.

Analysieren wir die Energiebilanz dieses Systems zunächst am einfachsten Modell, in welchem die Erde als ein schwarzer Körper betrachtet wird. Dann gilt für die Energieströme,

$$J_{S \rightarrow E} = A_S \chi_{S \rightarrow E} S_S = \int_{\text{Halbkugel}} d\Omega R_E^2 S_0 \cos \theta = \pi R_E^2 S_0 , \quad (6.8)$$

wobei A_S die Oberfläche der Sonne ist und $S_S = \sigma T_S^4$. Die Abstrahlung der Erde ins Weltall ist,

$$J_{E \rightarrow \infty} = A_E \chi_{E \rightarrow \infty} S_E = 4\pi R_E^2 \sigma T_E^4 , \quad (6.9)$$

wobei A_E die Oberfläche der Erde ist und $\chi_{E \rightarrow \infty}$ der Anteil der emittierten Strahlung, die ins Weltall entkommt, d.h. $\chi_{E \rightarrow \infty} = 1$ bis auf ganz kleine Korrekturen.

Die Gleichgewichtsbedingung $J_{S \rightarrow E} = J_{E \rightarrow \infty}$ ergibt

$$S_E = \frac{1}{4} S_0 \quad \Rightarrow \quad T_E = \left(\frac{S_0}{4\sigma} \right)^{1/4} \approx 279 \text{ K} . \quad (6.10)$$

Dies liegt verblüffend nahe beim tatsächlichen Mittelwert, den wir bei zirka 288 K (15°C) ansetzen.²

Diese Übereinstimmung ist jedoch purer Zufall, denn das Modell ist viel zu einfach. Zuerst einmal ist die Erde kein schwarzer Körper, sondern eher ein grauer Körper mit einer gemittelten

²In einer etwas verfeinerten Betrachtung kann man auch die Abhängigkeit der Erdoberflächentemperatur von den Breitengraden berücksichtigen. Der Energiefluss auf die Erde hängt lokal vom Auftretswinkel der Strahlung ab und ist proportional zur Projektion auf den lokalen Normalvektor der Erdoberfläche. Daher erhalten wir für den Breitenwinkel θ (vom Nordpol aus gemessen),

$$J_{S \rightarrow E}(\theta) d\theta = \sin \theta d\theta \int_{-\pi/2}^{+\pi/2} d\phi S_0 R_E^2 n_x = \sin \theta d\theta \int_{-\pi/2}^{+\pi/2} d\phi S_0 R_E^2 \cos \phi \sin \theta = 2R_E^2 S_0 \sin^2 \theta d\theta , \quad (6.11)$$

wobei die Sonne genau über dem Äquator steht und in x -Richtung strahlt. Von jedem Punkt auf der Oberfläche die Erde in alle Richtungen (Halbkugel über dem Punkt) ab. Zudem rechnen für gegebenes θ die Abstrahlung auf dem ganzen Breitenkreis,

$$J_{E \rightarrow \infty}(\theta) d\theta = \sin \theta d\theta \int_0^{2\pi} d\phi S_E = 2\pi S_E R_E^2 \sin \theta d\theta . \quad (6.12)$$

Reflektivität von zirka 30% ($r_E = 0.3$). Nehmen wir dies in Rechnung, dann folgt, dass die Erde nur $1 - r_E$ des Solarenergiestromes absorbiert,

$$S_E = \frac{1}{4}(1 - r_E)S_0 \quad \Rightarrow \quad T_E = \left(\frac{S_0(1 - r_E)}{4\sigma} \right)^{1/4} \approx 255 \text{ K}, \quad (6.14)$$

was als -18°C schon recht kalt ist. Diese Reflektivität stammt von der Oberflächenreflexion, insbesondere von Schnee und Eis, aber auch durch die Reflexion an Wolken zustande.

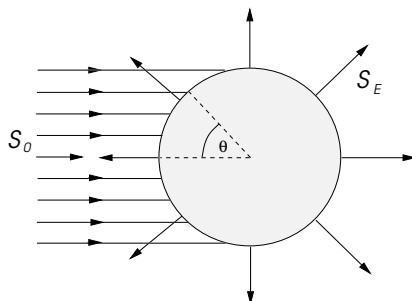


Abb.6.3: Absorption und Emission der Erde.

Damit wird klar, dass wir ein verbessertes Modell brauchen. Was wir bis anhin vernachlässigt haben ist die Atmosphäre der Erde, die eine wichtige Rolle spielt. Die Atmosphäre absorbiert, und folglich emittiert, reflektiert und streut Strahlung, und sie kann auch durch Konvektion zum Wärmetransport beitragen.

6.3 Treibhausgase der Atmosphäre

Wenn wir das Spektrum der einfallenden und emittierten Strahlung betrachten, fällt auf, dass dabei sehr unterschiedliche Frequenzbereiche eine Rolle spielen. Das Strahlungsspektrum der Sonne mit $T_S = 5800 \text{ K}$ ist maximal im sichtbaren Bereich (bezeichnen wir später als "UV"),

Damit gilt im Gleichgewicht, $J_{S \rightarrow E}(\theta) = J_{E \rightarrow \infty}(\theta)$,

$$T(\theta) = \left(\frac{S_0 \sin \theta}{\pi \sigma} \right)^{1/4}, \quad (6.13)$$

mit der maximalen Temperatur am Äquator $T(\pi/2) = 296 \text{ K}$ und einer mittleren Temperatur $\overline{T(\theta)} = 0.931 \times T(\pi/2) = 275 \text{ K}$. Diese mittlere Temperatur unterscheidet sich stark vom vereinfachten Wert, den wir oben gefunden haben, 279 K.

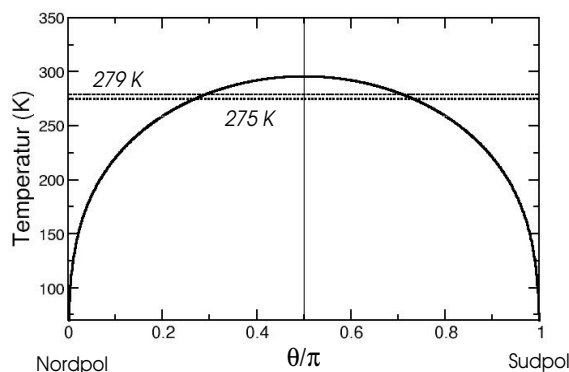


Abb.6.2: Temperatur-Verteilung aufgelöst nach Breitengrade.

Das Modell berücksichtigt natürlich keine horizontalen Wärmeausgleich.

während die Strahlung der Erde mit $T_E \approx 280$ K tief im Infraroten (IR) liegt. Das Maximum in der Spektralverteilung findet man sofort mittels des Wien'schen Verschiebungsgesetzes,

$$\lambda_m(\mu\text{m}) = \frac{2897}{T(\text{K})} \quad \Rightarrow \quad \begin{cases} 0.5 \mu\text{m} & T = 5800 \text{ K} \\ 10 \mu\text{m} & T = 280 \text{ K} \end{cases} \quad (6.15)$$

Dies hat weitreichende Folgen, denn im Spektralbereich der Sonne (UV und sichtbar) ist die Atmosphäre der Erde grösstenteils transparent. Hingegen findet in der Atmosphäre eine beträchtliche Absorption im infraroten Bereich statt, in dem die Erde abstrahlt. Dies ist die Grundlage des Treibhauseffektes.

6.3.1 Treibhaus-Prinzip

Nehmen wir uns zunächst einmal das Treibhaus eines Gärtners an. Ein einfaches Modell ist ein ebener Boden (schwarzer Körper) mit einer parallelen Glasdecke auf einer gewissen Höhe, durch die das Sonnenlicht senkrecht einstrahlt. Die Fläche A sei gross genug, dass wir Randeffekte vernachlässigen können. Das Glas sei vollständig transparent für die Sonnenstrahlung, aber nur teilweise transparent für die Infrarotstrahlung (aber nicht reflektierend), die der Treibhausboden abstrahlt. Betrachten wir als den Energiefluss:

$$\begin{aligned} \text{Sonne} \rightarrow \text{Boden:} \quad & J_{S \rightarrow B} = AS_0 \\ \text{Boden} \rightarrow \text{Glas:} \quad & J_{B \rightarrow G} = Aa_G\sigma T_B^4 \\ \text{Glas} \rightarrow \text{Boden} \quad & J_{G \rightarrow B} = Aa_G\sigma T_G^4 \\ \text{Glas} \rightarrow \infty : \quad & J_{G \rightarrow \infty} = Aa_G\sigma T_G^4 \\ \text{Boden} \rightarrow \infty \quad & J_{B \rightarrow \infty} = A(1 - a_G)\sigma T_B^4 \end{aligned} \quad (6.16)$$

Dabei ist a_G der Absorptionsanteil des Glases und $1 - a_G$ entspricht dem transmittierten Teil. T_G und T_B sind die Temperatur des Glases bzw. des Bodens. Beachte, dass das Glas auf beide Seiten (unten und oben) gleich abstrahlt, d.h. es hat eine doppelt so grosse Fläche, auf der es abstrahlen kann, als der Treibhausboden. Damit gilt für die Bilanz:

$$\text{Boden:} \quad J_{S \rightarrow B} + J_{G \rightarrow B} = J_{B \rightarrow G} + J_{B \rightarrow \infty} \quad \Rightarrow \quad S_0 + a_G\sigma T_G^4 = \sigma T_B^4 \quad (6.17)$$

$$\text{Glas:} \quad J_{B \rightarrow G} = J_{G \rightarrow B} + J_{G \rightarrow \infty} \quad \Rightarrow \quad T_B^4 = 2T_G^4$$

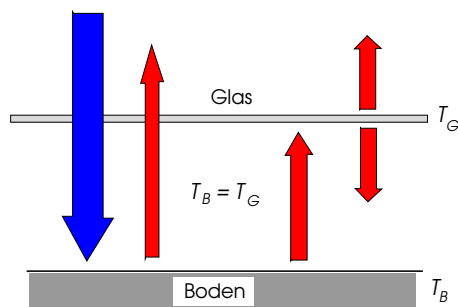


Abb.6.4: Treibhaus-Prinzip: Glasüberdachung absorbiert im IR (vom Boden) aber lässt UV (von Sonne) durch. Die Hälfte der absorbierten Strahlung wird wieder zum Boden zurückreflektiert.

Die zweite Gleichung ergibt, dass $T_B = 2^{1/4} T_G$, so dass wir sofort aus der ersten Gleichung finden, dass

$$T_B(a_G) = \left(\frac{2S_0}{\sigma(2 - a_G)} \right)^{1/4} = \frac{T_B(0)}{(1 - a_G/2)^{1/4}}, \quad (6.18)$$

wobei $T_B(0)$ die stationäre Temperatur des Bodens ohne Treibhaus ist, d.h. $a_G = 0$, da alle Abstrahlung vom Boden direkt ins Unendliche wegfließt. Wenn das Glas im infraroten Teil des Spektrums vollkommen undurchlässig wirkt, dann erhalten wir eine Erhöhung um einen Faktor $2^{1/4} = 1.1892$. Dies interpoliert sich kontinuierlich im Zwischenbereich. Die Atmosphäre ist um den Faktor $2^{-1/4} = 0.84$ kälter als die Erdoberfläche.

Es ist einfach das Modell zu erweitern durch Reflektivitäten in den beiden Spektralbereichen für das Glas und den Boden. Fazit ist jedoch immer, dass die selektive Durchlässigkeit eine zusätzliche Erwärmung, einen Treibhauseffekt, bewirkt.

6.3.2 Treibhausgase und Absorption in der Atmosphäre

Die Atmosphäre mit ihrer Zusammensetzung erfüllt gerade die Bedingung, die wir für einen Treibhauseffekt benötigen. Sie ist im UV und sichtbaren Bereich der elektromagnetischen Strahlung praktisch transparent. Im Infraroten hingegen gibt es sogenannte Treibhausgase, die stark absorbieren. Dazu gehören CO_2 , Methan CH_4 und Stickstoff-Oxid N_2O . Wasserdampf ist das wichtigste "Treibhausgas". Es sind aber vor allem die ersten drei, auf die der Mensch einen Einfluss hat.

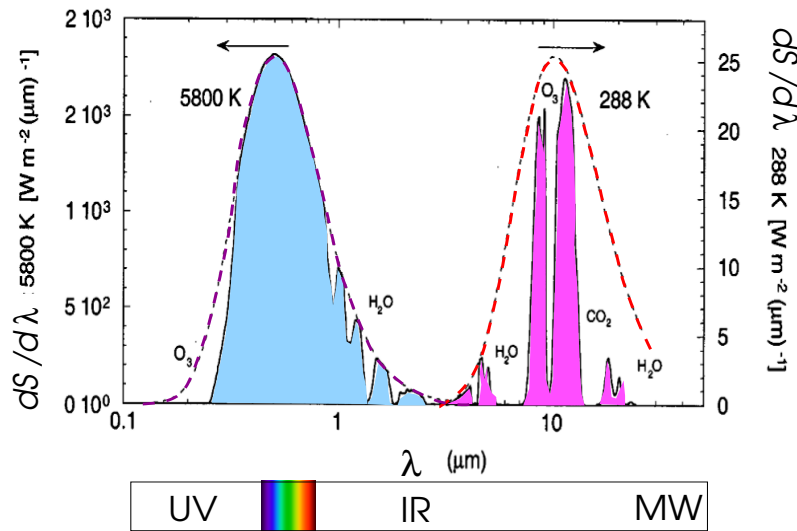


Abb.6.5 Transmissionsspektrum für die Solareinstrahlung (links) und die Erdatstrahlung (rechts). (J.R. Barker und Marc H. Ross, *Am. J. Phys.* 67, 1216 (1999)).

Abb. 6.5 zeigt, dass das Spektrum der Sonne durch Ozon O_3 im UV-Bereich eine starke Absorption aufweist. Daneben gibt es einige Absorptionslinien des Wasserdampfs bei grösseren Wellenlängen. Im IR-Bereich der Erdstrahlung sind sowohl Wasser, Ozon als auch CO_2 sehr wichtig für die Absorption.

Als bekanntestes Beispiel diskutieren wir hier CO_2 . Das CO_2 -Molekül ist linear bestehend aus drei Atomen, C im Zentrum und auf beiden Seiten kovalent gekoppelt ein O-Atom. Dies gibt ein reichhaltiges Molekül-Spektrum. Das elektronische Spektrum liegt bei sehr hohen Energien wie die atomaren Spektren und ist wenig relevant im Spektralbereich, der für uns von Interesse ist. Die Rotationen haben sehr niedrige Energie für

$$E(J) = hcB J(J + 1) \quad J = 0, 1, 2, 3, \dots \quad \text{mit} \quad hcB \approx 0.6K \approx 0.4\text{cm}^{-1}. \quad (6.19)$$

Als nächstes betrachten wir die Vibrationsmoden. Es gibt vier Normalschwingungen des Moleküls, wobei zwei entartet sind. Es gibt die symmetrische (1) und antisymmetrische (3) Streckschwingung und zwei Knickschwingungen (2). Die beiden Streckschwingungen erhalten die Achse des Moleküls, während die Knickschwingung eine transversale Auslenkung erzeugt, wobei zwei Auslenkungsrichtungen (x und y) möglich sind.

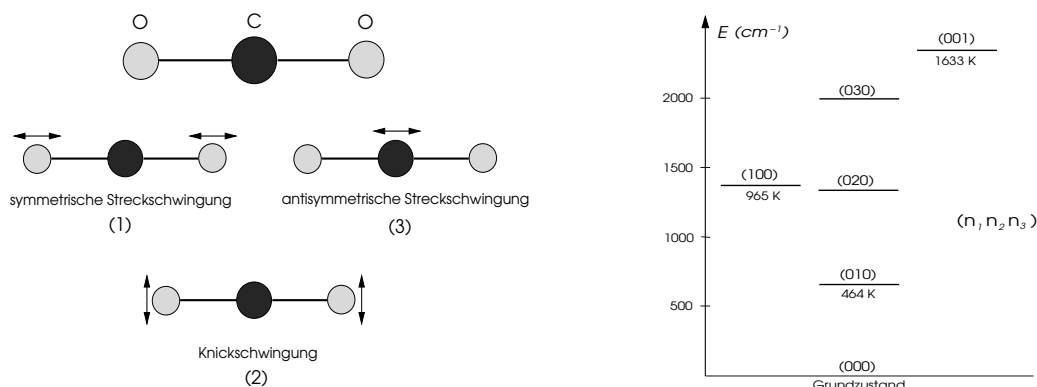


Abb.6.6: Vibrationsmoden des CO_2 -Moleküls: (links) 3 verschiedene Moden, wobei die Knickschwingung zwei möglich Polarisierungen hat; (rechts) Spektrum der drei Moden bezeichnet durch die Quantenzahlen (n_1, n_2, n_3) .

Diese Schwingungen besitzen das Spektrum von harmonischen Oszillatoren:

$$E = \hbar\omega_1 \left(n_1 + \frac{1}{2} \right) + \hbar\omega_2 (n_{2x} + n_{2y} + 1) + \hbar\omega_3 \left(n_3 + \frac{1}{2} \right) \quad (6.20)$$

mit $\hbar\omega_1 = 965\text{K} = 1388\text{cm}^{-1}$, $\hbar\omega_2 = 464\text{K} = 667\text{cm}^{-1}$ und $\hbar\omega_3 = 1633\text{K} = 2349\text{cm}^{-1}$. Dabei sind nur Mode (2) und (3) dipolaktiv. Ähnliche Spektren lassen sich für andere Moleküle der Treibhausgase finden, wie im Abb. 6.7 dargestellt. Für CO_2 finden wir wichtige Absorptionslinien bei $\lambda \approx 15\mu\text{m}$ und $\lambda \approx 4.25\mu\text{m}$, die von der Knick- bzw. der antisymmetrischen Streckschwingung herrühren.

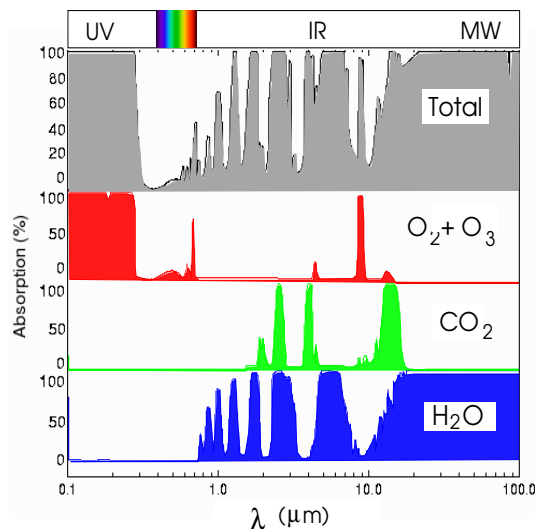


Abb.6.7. Spektrum der Absorption für die Treibhausgase Ozon, CO_2 und Wasserdampf (Quelle: www.brneurosci.org/co2.htm).

Vergleichen wir mit Abb.6.5 dann sehen wir, die dominante Absorptionen der Erdstrahlung durch CO_2 bei $\lambda = 15\mu\text{m}$, Ozon bei $\lambda = 10\mu\text{m}$ und Wasser bei $\lambda \approx 6\mu\text{m}$ erzeugt wird. Die

Verbreiterung der Linien kommen durch verschiedene Mechanismen zustande. Die Linie bei der Resonanzfrequenz ω_0 kann durch eine Lorentzkurve beschrieben werden,

$$a(\omega - \omega_0) \propto \frac{\Gamma}{\pi(\omega - \omega_0)^2 + \Gamma^2} \quad (6.21)$$

wobei Γ mit der Lebensdauer, mit Kollisionen mit anderen Molekülen oder mit der Dopplerverschiebung der bewegten Moleküle zusammenhängt.

6.3.3 Reflektivität - Albedo

Die Erde ist nicht einfach ein schwarzer Körper. Ein Blick auf ein Bild der Erde aus dem Weltall zeigt, dass wenigstens im sichtbaren Bereich des Strahlungsspektrums eine beträchtliche Reflexion, auch Albedo genannt, vorherrscht. Die Reflexion von der Erdoberfläche kommt von Schnee und Eis, Sand oder überbauten Flächen. In der Atmosphäre sind es vor allem die Wolken. Die Reflexion ist daher durchaus keine unveränderbare Grösse. Ferner hängen die Reflexions- und Absorptionseigenschaften auch vom Frequenzbereich ab.

Oberfläche	$r^{(UV)}$	$r^{(IR)}$
Schnee	0.8-0.9	0.005
Wasser	0.03-0.01	0.04
Sand	0.25-0.35	0.1
Rasen	0.25	0.016
Wald	0.05-0.18	

Tab. 6.1: Reflektivität verschiedener Formen der Erdoberfläche für den UV- und IR-Bereich.

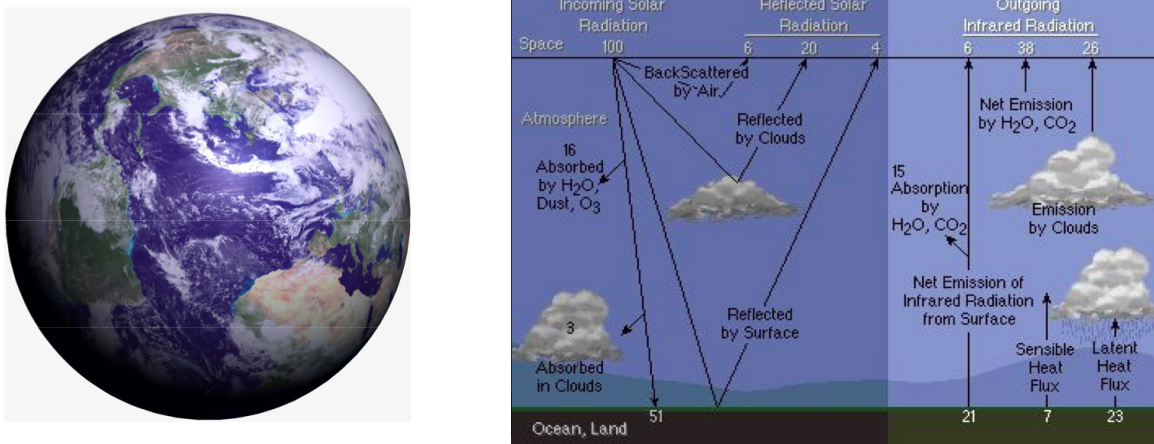


Abb.6.8: Energie-Umsatz im System Erdoberfläche-Atmosphäre (Quelle: Gidon Eshel "Global Warming Physics and Climate Modeling, Predictability and Uncertainty", University of Chicago.)

6.4 Einfaches Zweischichten-Modell

Wir stellen hier ein relativ einfaches Modell für das System Erde-Atmosphäre auf. Die Grössen, die für uns von Interesse sind, sind die beiden Temperaturen, T_E auf der Erdoberfläche und T_A in der Atmosphäre. Die Vereinfachung liegt hier darin, dass wir die komplexe Schichtung der Atmosphäre vernachlässigen und sie einfach als eine Schicht auffassen.³

³Dabei folgen wir R.S. Know, Am. J. Phys. 67, 1227 (1999).

Beginnen wir damit die Systemparameter zu definieren:

UV-Bereich: (Sonnenstrahlung)	Reflektivität der Erdoberfläche r_E Reflektivität der Atmosphäre r_A Absorptivität der Atmosphäre a_A
IR-Bereich (Erd-/Atmosphären-Strahlung)	Absorptivität der Atmosphäre g_A Absorptivität der Atmosphäre $g_E = 1$

Der Energietransport geschieht einerseits durch Strahlung, wobei S_0 die Solarkonstante, $S_E = \sigma T_E^4$ die Erdoberstrahlung und $S_A = g_A \sigma T_A$ die Atmosphären-Abstrahlung bezeichnet. Andererseits, gibt es einen Energieaustausch zwischen Erde und Atmosphäre, der nicht-strahlend ist, S_K aufgrund von Konvektion und latenter Wärme (Eisschmelze, Wasser \leftrightarrow Dampf). Der letztere Beitrag ist nicht klein und liegt bei einem mittleren Wert $S_K \sim 100 \text{ W m}^{-2}$.

Betrachten wir nun die Energieflüsse. Im Gleichgewicht finden wir für die Erdoberfläche,

$$\frac{(1 - r_E)(1 - a_A)(1 - r_A) S_0}{1 - r_E r_A} \frac{1}{4} + g_A S_A = S_E + S_K \quad (6.22)$$

wobei wir $g_E = 1$ gesetzt haben. Für die Atmosphäre gilt entsprechend,

$$\frac{a_A(1 - a_A)r_E(1 - r_A)^2 S_0}{1 - r_E r_A} \frac{1}{4} + a_A(1 - r_A) \frac{S_0}{4} + g_A S_E + S_K = 2g_A S_A. \quad (6.23)$$

In beiden Fällen erhalten wir einen komplizierten Vorfaktor für den solaren Beitrag im UV-Bereich, der von der mehrfachen Reflexion zwischen Erdoberfläche und Atmosphäre stammt.⁴ Dies sind zwei Gleichungen für die beiden Unbekannten S_A und S_E . Wir können sie als lineares Gleichungssystem schreiben,

$$\begin{pmatrix} 2g_A & -g_A \\ -g_A & 1 \end{pmatrix} \begin{pmatrix} S_A \\ S_E \end{pmatrix} = \begin{pmatrix} A \frac{S_0}{4} + S_K \\ B \frac{S_0}{4} - S_K \end{pmatrix} = \begin{pmatrix} A + h \\ B - h \end{pmatrix}, \quad (6.28)$$

⁴Die Einstrahlung der Sonne auf die Erdoberfläche wird durch den Effekt der Mehrfach-Reflexion verstärkt. Wir beschränken uns auf den absorbierten UV-Anteil der Strahlung für die Erdoberfläche,

$$\begin{aligned} & \frac{S_0}{4} (1 - r_A)(1 - r_E)(1 - a_A) + r_E r_A (1 + r_E r_A + (r_E r_A)^2 + \dots) (1 - r_A)(1 - r_E)(1 - a_A) \frac{S_0}{4} \\ & = \frac{S_0}{4} (1 - r_A)(1 - r_E)(1 - a_A) + r_E r_A S_0^* = S_0^*, \end{aligned} \quad (6.24)$$

wobei wir die effektive Solarkonstante S_0^* einführen. Der erste Term auf der linken Seite ist der Energiefluss der die Erdoberfläche direkt von der Sonne her erreicht und absorbiert wird. Der zweite Term beschreibt, den Teil der durch Mehrfachreflexion dazukommt. Daraus ergibt sich für die effektive Solarkonstante,

$$S_0^* = \frac{(1 - r_E)(1 - a_A)(1 - r_A) S_0}{1 - r_E r_A}. \quad (6.25)$$

Analog finden wir für die Atmosphäre den Teil der UV-Strahlung, der nicht direkt von der Sonne her absorbiert, sondern wenigstens einmal an der Erdoberfläche reflektiert und dann absorbiert wird,

$$\begin{aligned} & \frac{S_0}{4} a_A (1 - r_A)^2 (1 - a_A) r_E + r_A r_E (1 + r_E r_A + (r_E r_A)^2 + \dots) \frac{S_0}{4} a_A (1 - r_A)^2 (1 - a_A) r_E \\ & = \frac{S_0}{4} a_A (1 - r_A)(1 - a_A) r_E + r_A r_E \tilde{S}_0^* = \tilde{S}_0^*, \end{aligned} \quad (6.26)$$

so dass

$$\tilde{S}_0^* = \frac{a_A r_E (1 - r_A)^2 (1 - a_A) S_0}{1 - r_E r_A} \frac{1}{4}. \quad (6.27)$$

Der zweite Term der linken Seite in (6.23) entspricht der direkten Absorption der Sonnenstrahlung.

wobei $h = 4S_K/S_0$ und

$$A = \frac{a_A(1 - a_A)r_E(1 - r_A)^2}{1 - r_E r_A} + a_A(1 - r_A) \quad \text{und} \quad B = \frac{(1 - r_E)(1 - a_A)(1 - r_A)}{1 - r_E r_A}. \quad (6.29)$$

Dieses Gleichungssystem ist einfach zu lösen,

$$\begin{pmatrix} S_A \\ S_E \end{pmatrix} = \sigma \begin{pmatrix} T_A^4 \\ T_E^4 \end{pmatrix} = \frac{S_0}{4(2 - g_A)} \begin{pmatrix} A + h + g_A(B - h) \\ A + h + 2(B - h) \end{pmatrix} \quad (6.30)$$

Daraus können wir die früheren Resultate einfach reproduzieren. Nehmen wir an, dass die Atmosphäre in keinem Strahlungsbereich absorbiert, $a_A = g_A = 0$, dann folgt $A = 0$ und

$$T_E^4 = 2T_A^4 = T_0^4 \left\{ \frac{(1 - r_E)(1 - r_A)}{1 - r_E r_A} - h \right\} < T_0^4, \quad (6.31)$$

wobei $0 < h < 1$ und $0 < (1 - r_E)(1 - r_A)/(1 - r_E r_A) \leq 1$. Beachte, dass der Ausdruck in der Klammer $\{\dots\}$ aus physikalischen Gründen nie negativ wird, da im stationären Zustand nie mehr Energie konvektiv abgeführt wird als strahlend der Erdoberfläche zugeführt wird. T_0 ist die Temperatur an der Erdoberfläche ohne Albedo und Treibhauseffekt ($T_0 = 279K$ wie in (6.10)). Das bedeutet, dass Reflexion allein nicht zu einer Temperaturerhöhung führt, selbst wenn wir die UV-Strahlung teilweise durch Mehrfach-Reflexion zwischen Erdoberfläche und Atmosphäre "einfangen" können. Auch der Energiefluss über Konvektion usw. wirkt kühlend.

Gehen wir nun das Problem an, die geeigneten Parameter zu finden. Die Standardversion des Zweischichten-Modells vernachlässigt S_K , obwohl dieser Beitrag offensichtlich nicht klein ist. Wir wollen jetzt einmal versuchen innerhalb unseres Modells vernünftige Parameter zu suchen. Mit Hilfe von Abb. 6.7 betrachten wir nun näherungsweise die Energiebilanz:

Anteile	Wm^{-2}	%	Modell-Parameter
Solarkonstante/4	342	100	$S_0/4$
UV reflektiert durch A	89	26	$r_A S_0/4$
UV reflektiert durch E (netto)	14	4	$(1 - a_A)^2 r_E (1 - r_A) S_0/4(1 - r_E r_A)$
UV absorbiert durch A	65	19	$A S_0/4$
UV absorbiert durch E	174	51	$B S_0/4$
IR emittiert durch E	390	114	S_E
IR absorbiert durch A	368	108	$g_A S_E$
IR emittiert durch A (oben)	218	64	$g_A S_A$
IR emittiert durch A (unten)	318	93	$g_A S_A$
IR emittiert ins All	240	70	$g_A S_A + (1 - g_A) S_E$
Konvektion,	100	29	$h S_0/4$

Tab. 6.2: Energiebilanz der verschiedenen Energiefluss-Komponenten in Zahlen und der Vergleich mit den Ausdrücken in unserem Modell.

Es ist offensichtlich, dass unser Parameterraum nicht ausreicht, um die Energiebilanz genau wie in Tab. 6.2 zu reproduzieren. Dies tritt besonders deutlich bei IR-Emission der Atmosphäre hervor, bei der sehr unterschiedliche Abstrahlungen nach oben und unten existieren. Dies liegt an der komplexen Schichtung der Atmosphäre, die in unserem einfachen Modell vollständig vernachlässigt wird. Um dennoch eine vernünftige Beschreibung mit Hilfe unseres Modells zu finden, müssen wir einige Kompromisse eingehen.

Wir führen die Wahl der Parameter so durch, dass eine mittlere Oberflächentemperatur von $T_E = 288K$ resultiert. In der folgenden Tabelle geben wir die möglichen Parameter-Sets für das Standard-Modell ($h = 0$) und das erweiterte Zweischichten-Modell.

Grösse	Standard-Modell	erweitertes Modell	maximale Erwärmung
a_A	0.08	0.09	0
g_A	0.89	0.965	1
h	0	0.117	0
r_A	0.255	0.270	0
r_E	0.160	0.10	0
T_E	288K	288K	331K
T_A	246K	252K	279K

Tab. 6.3: Tabelle der Modell-Parameter für das Standard- und das erweiterte Modell, sowie das Modell mit dem maximalen Treibhauseffekt.

Innerhalb des Zweischicht-Modells ist es nur bedingt möglich die detaillierte Energiebilanz zu erhalten. Dies gilt auch für den konvektiven Anteil h , der beinahe dreimal zu klein ausfällt, wenn wir versuchen T_E auf einen "vernünftigen" Wert zu fixieren. Es ist jedoch interessant zu sehen, was geschieht, wenn wir die verschiedenen Parameter variieren. Betrachten wir nur kleine Änderungen, dann können wir die Ableitung von T_E nach den entsprechenden Parametern berechnen.

$\frac{1}{T_E} \frac{\partial T_E}{\partial x}$	Standard-Modell	erweitertes Modell
$\frac{1}{T_E} \frac{\partial T_E}{\partial g_A}$	0.23	0.24
$\frac{1}{T_E} \frac{\partial T_E}{\partial a_A}$	-0.095	-0.12
$\frac{1}{T_E} \frac{\partial T_E}{\partial r_A}$	-0.3	-0.32
$\frac{1}{T_E} \frac{\partial T_E}{\partial r_E}$	-0.21	-0.21
$\frac{1}{T_E} \frac{\partial T_E}{\partial h}$	-0.2	-0.21

Die Erdtemperatur wird erniedrigt, falls die UV-Absorption und -Reflektivität der Atmosphäre und die UV-Reflektivität der Erdoberfläche zunimmt. Andererseits ergibt eine erhöhte IR-Absorption der Atmosphäre einen stärkeren Treibhauseffekt, so dass T_E ansteigt. Allgemein gilt

$$\Delta T_E = \left\{ \frac{\partial T_E}{\partial g_A} \Delta g_A + \frac{\partial T_E}{\partial a_A} \Delta a_A + \frac{\partial T_E}{\partial r_A} \Delta r_A + \frac{\partial T_E}{\partial r_E} \Delta r_E + \frac{\partial T_E}{\partial h} \Delta h \right\}. \quad (6.32)$$

Beachte, dass beide Modelle ungefähr die gleichen Koeffizienten liefern.

6.5 Globale Erwärmung

Der Effekt von Spurengasen auf das Erdklima wurde schon sehr früh entdeckt. Erste Arbeiten dazu stammen von Jean-Baptiste Fourier im Jahre 1824. In den Sechzigerjahren des 19. Jahrhunderts beschäftigte sich John Tyndall mit dem Treibhauseffekt, der vorallem von Wasserdampf herrührt. 1896 schlug Svante Arrhenius CO_2 als wichtiges Treibhausgas vor und schätzte ab, dass eine Verdopplung der CO_2 -Konzentration in der Atmosphäre eine Temperatur-Erhöhung von 4 - 6°C nach sich ziehen würde. Die Rolle von CO_2 als Treibhausgas konnte jedoch erst sehr viel später nachgewiesen werden. Neuere Studien legen die Erwärmung bei einer CO_2 -Verdoppelung in der Atmosphäre zwischen 1.5 und 4.5K.

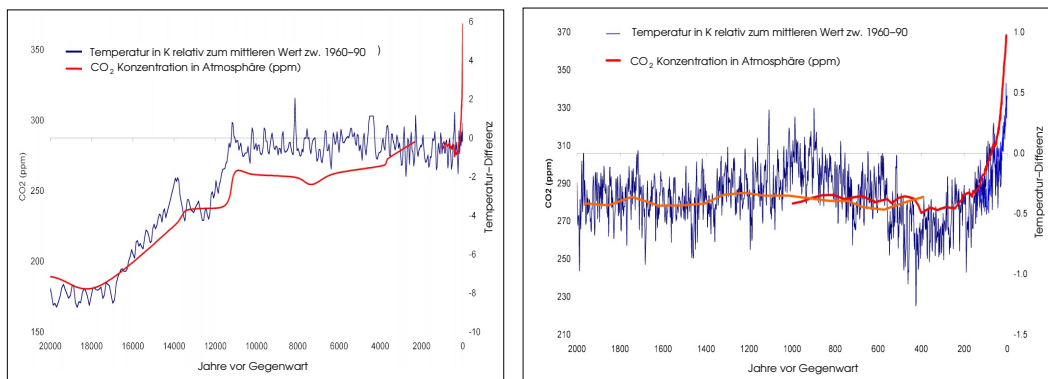


Abb.6.9: Korrelation zwischen CO_2 und Temperatur für die letzten 20000 Jahre (lnks) und der letzten 2000 Jahre (rechts). Beachte den markanten Knick bei ca. 10000 Jahren, der das Ende der letzten Eiszeit markiert. (Quelle: www.brighton73.freeseerve.co.uk/globalwarmingfaq.htm).

In der Vergangenheit ist die Korrelation der Temperatur und der CO_2 -Konzentration leicht phasenverschoben im Sinne, dass die Änderung der CO_2 -Konzentration eher eine Folge als die Ursache der Temperaturänderung zu sein scheint. Die neue Situation, die wir gegenwärtig beobachten, ist die Tatsache, dass in den letzten zwei Jahrhunderten ein markanter Anstieg der CO_2 -Konzentration stattgefunden hat (siehe auch Abb. 6.9), wie diverse Untersuchungen nachweisen. Aufgrund der Isotopen-Zusammensetzung konnte ein klarer Zusammenhang zur Verbrennung fossiler Brennstoffen hergestellt werden. Nie in der belegbaren Geschichte des CO_2 -Haushalts der Atmosphäre hat die Konzentration einen derart hohen Wert erreicht. Mit der Konzentrationsänderung geht eine nachweisbare allmähliche Erhöhung der mittleren jährlichen Temperatur einher, die Anlass zu Besorgnis gibt, obwohl die Konsequenzen nur schwer abschätzbar sind. Das wichtigste Treibhausgas bleibt nämlich Wasserdampf, dessen Konzentration von den klimatischen Verhältnissen stark beeinflusst wird. Dies erschwert die Abschätzung der Effekte, die eine Erhöhung der Konzentrationen von CO_2 und anderen Treibhausgasen, wie Methan und N_2O mit sich bringt. Dabei spielen Rückkopplungseffekte, die nur schwer vorhersagbar sind, eine besondere Rolle. Aufgrund der obigen Modellbetrachtungen wollen wir hier einige Zusammenhänge beleuchten.

- Eine Erhöhung der CO_2 -, CH_4 - und N_2O -Konzentration erhöht die Infrarot-Absorption g_A ($\Delta g_A > 0$). Dies resultiert in einer Temperaturerhöhung.
- Ein offensichtlicher Rückkopplungseffekt ergibt die Tatsache, dass eine erhöhte Temperatur den Albedo der Erdoberfläche erniedrigen kann, indem nämlich grosse Schnee- und Eisflächen abschmelzen. Dies ist beobachtbar an Gletschern und an den Eisflächen an beiden Polen. Die freiwerdende Fläche ist stärker absorbierend und erniedrigt folglich r_E ($\Delta r_E < 0$). Gemäss unserem Modell führt dies zu einer Erwärmung, die den Fortgang solchen Abschmelzens noch beschleunigen kann. Dies ist eine verstärkende Rückkopplung.

- Die Erwärmung führt zu einer Erhöhung des Wasserdampfs in der Atmosphäre. Dies hat einerseits die Folge, dass der Treibhauseffekt verstärkt wird durch Anstieg von g_A . Andererseits wird die UV-Reflektivität der Atmosphäre durch Wolkenbildung erhöht, d.h. r_A steigt. Der erste Aspekt ($\Delta g_A > 0$) lässt die Temperatur ansteigen, während der zweite ($\Delta r_A > 0$) zu einer Reduktion der Temperatur führt. In welche Richtung der Trend kippt kann nur durch sehr komplexe Modellierung vorhergesagt werden, und stellt eine der grössten Unsicherheiten bei der Klimaprognose dar. Diese Rückkopplung ist unklar.
- Die "Verdunklung" der Atmosphäre durch Vulkanäusbrüche erhöht deren Absorption und möglicherweise der Reflexion im UV-Bereich ($\Delta a_A > 0$ und $\Delta r_A > 0$). Beide Effekte haben eine Erniedrigung der Temperatur zur Folge.
- Das Abschmelzen von Eis und Schnee kann zwischenzeitlich die Erwärmung etwas dämpfen, da der nicht-strahlende Anteil der Wärmeabfuhr von der Erdoberfläche zunimmt ($\Delta h > 0$).

Rückkopplungseffekte machen das Klimasystem äusserst kompliziert und könnten durchaus zu Überraschungen führen. Zum Beispiel ist ungewiss, in welchem Umfang das im Wasser gelöste CO_2 bei Erwärmung der Meere und Seen in die Atmosphäre übergeht. Enorme Mengen von gefrorenem Methan lagert auf dem Meeresboden und könnte sich auch in grösserem Umfang irreversibel in die Atmosphäre begeben. Die Rückkopplungseffekte machen das Klimamodell auch stark nichtlinear, so dass eine Vielzahl von verschiedenen stationären Zuständen existieren können. Vom Studium der Eiszeiten weiss man, dass Klimaänderungen sich oft in sehr kurzer Zeit vollziehen und unabsehbare Folgen haben können. Die Rolle der Atmosphäre für den Energiehaushalt der Erde ist aber wichtig dafür, dass eine Biosphäre auf der Erdoberfläche überhaupt existieren kann.

Die Gewalt des Treibhauseffektes kann jedoch sehr eindrücklich am Beispiel der Venus illustriert werden. Die Solarkonstante der Venus liegt bei 2580 W m^{-2} wobei aufgrund eines besonders starken Albedo der Atmosphäre zirka 80 % der Strahlung reflektiert wird, so dass ein wesentlich kleinerer Energiefluss die Planetenoberfläche erreicht als bei der Erde. Dies würde eine äusserst geringe Oberflächentemperatur führen. In Wahrheit liegt die Temperatur aber bei $T_V \approx 740 \text{ K}$ ($\sim 465^\circ \text{C}$). Diese hohe Temperatur hat ihren Ursprung im Treibhauseffekt der Atmosphäre, die zu 96 % aus CO_2 besteht.

Die hier diskutierten einfachen Prinzipien des globalen Wärmehaushaltes sind offensichtlich stark vereinfacht. Trotzdem können wir davon ausgehen, dass die qualitativen Aussagen stimmen. Wie sich die globale Erwärmung und der Treibhauseffekt auf die Entwicklung des Klimas auf der Erde auswirken, muss mit komplizierteren und ausgefeilteren Modellen untersucht werden.